

22 - 24 NOVIEMBRE 2023 | CARTAGENA



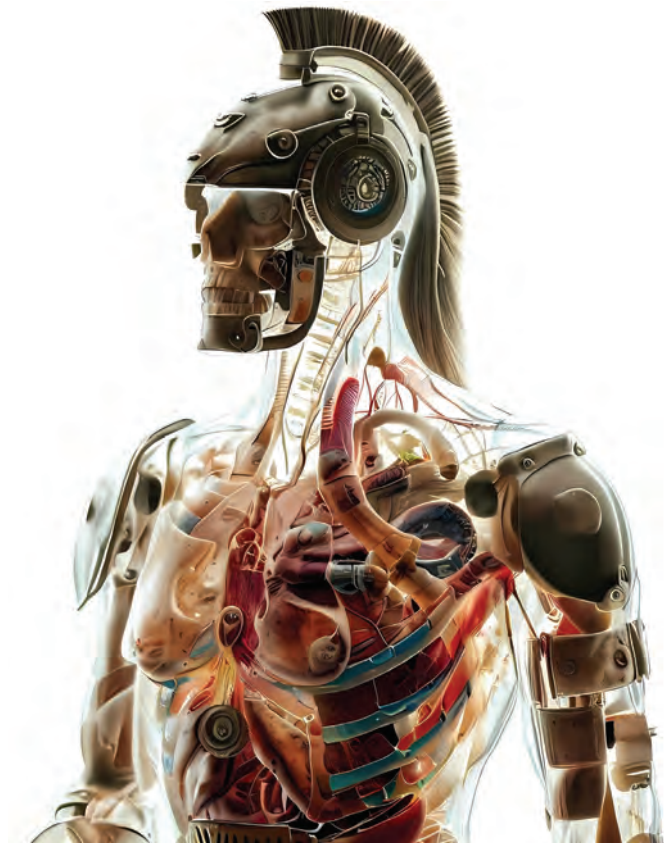
CASEIB 2023 XLI

Congreso Anual de la Sociedad Española de

Ingeniería Biomédica

Contribuyendo a la salud basada en valor

LIBRO DE ACTAS



Universidad
Politécnica
de Cartagena

edicionesUPCT



Universidad
Politécnica
de Cartagena

25
Aniversario UPCT
1998-2023

Sociedad Española de
INGENIERÍA BIOMÉDICA



Tecnología
Sanitaria



Co-funded by the
European Union

Región de Murcia



f SéNeCa(+)

Agencia de Ciencia y Tecnología
Región de Murcia

22 - 24 NOVIEMBRE 2023 | CARTAGENA



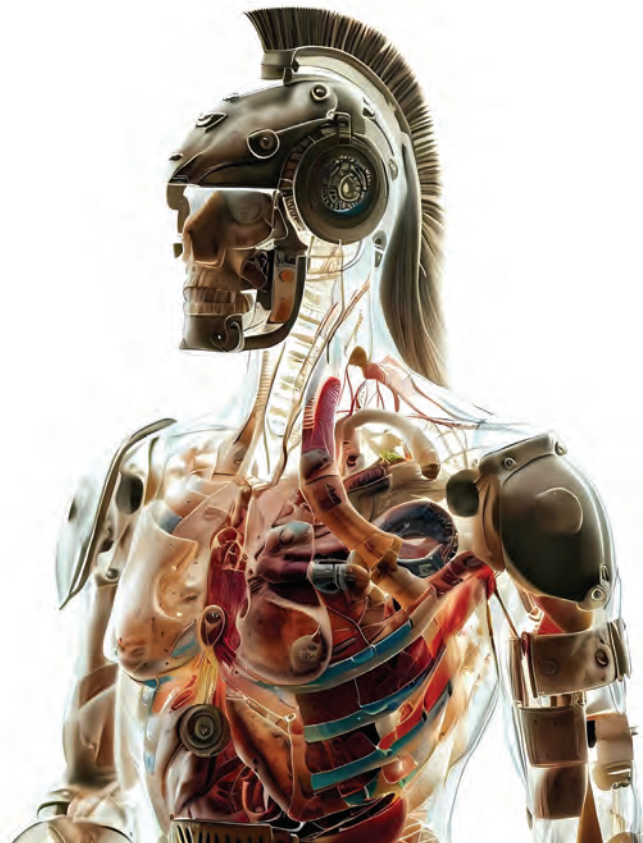
CASEIB 2023 XLI

Congreso Anual de la Sociedad Española de

Ingeniería Biomédica

Contribuyendo a la salud basada en valor

LIBRO DE ACTAS



AITADIS



CONGRESO (EXPEDIENTE 22193/0C/23) FINANCIADO POR LA CONSEJERÍA DE MEDIO AMBIENTE, MAR MENOR, UNIVERSIDADES E INVESTIGACIÓN, A TRAVÉS DE LA FUNDACIÓN SÉNECA-AGENCIA DE CIENCIA Y TECNOLOGÍA DE LA REGIÓN DE MURCIA CON CARGO AL PROGRAMA REGIONAL DE MOVILIDAD, COLABORACIÓN E INTERCAMBIO DE CONOCIMIENTO "JIMÉNEZ DE LA ESPADA"



CASEIB 2023 XLI

Congreso Anual de la Sociedad Española de

Ingeniería Biomédica

Contribuyendo a la salud basada en valor

22 - 24 NOVIEMBRE 2023 | CARTAGENA

LIBRO DE ACTAS

Los trabajos incluidos en esta publicación han sido evaluados por el Comité Científico que en ella se relaciona y según el procedimiento que se recoge en www.caseib.es/2023/

Presidente del Comité Organizador:

Joaquín Roca González

Secretaria del Comité Organizador:

Dolores Ojados González

Editores:

Joaquín Roca González

Dolores Ojados González

Juan Suardíaz Muro

Maquetación:

Lucía Martínez Abellán

Marta García Fuentes

Diseño de portada:

Lucía Martínez Abellán

Cita bibliográfica:

Roca González, J., Ojados González, D. y Suardíaz Muro, J. Eds. CASEIB 2023. Libro de Actas del XLI Congreso Anual de la Sociedad Española de Ingeniería Biomédica. Cartagena: Universidad Politécnica de Cartagena, 2023. 742 p. ISBN: 978-84-17853-76-1



Libro de Actas del XLI Congreso Anual de la Sociedad Española de Ingeniería Biomédica

Se distribuye bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional.

ISBN: 978-84-17853-76-1

URI: <http://hdl.handle.net/10317/12893>

ORGANIZADORES



Universidad
Politécnica
de Cartagena



COLABORADORES



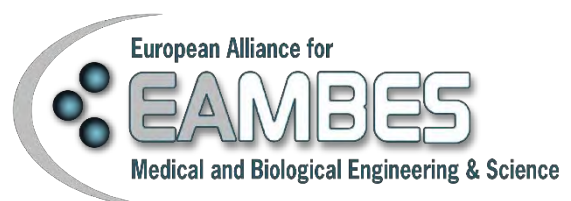
PATROCINADORES



Congreso (Expediente 22193/OC/23) financiado por la Consejería de Medio Ambiente, Mar Menor, Universidades e Investigación, a través de la Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia con cargo al Programa Regional de Movilidad, Colaboración e Intercambio de Conocimiento "Jiménez de la Espada"



AVALADO POR



ÍNDICE

BIENVENIDA	vii
COMITÉ DE HONOR	ix
COMITÉ ORGANIZADOR	x
COMITÉ CIENTÍFICO	xi
PROGRAMA MIÉRCOLES 22 DE NOVIEMBRE	xii
PROGRAMA JUEVES 23 DE NOVIEMBRE	xiii
PROGRAMA VIERNES 24 DE NOVIEMBRE	xiv
CONFERENCIA PLENARIA INAUGURAL	xv
CONFERENCIA PLENARIA DE CLAUSURA	xvii
MESA REDONDA I	xix
MESA REDONDA II	xxi
PREMIOS	xxiii
ACTOS	xxvi
ÍNDICE DE SESIONES CIENTÍFICAS	xxvii
Señales Biomédicas (I)	1
Imágenes Médicas (I)	26
Biomecánica / Ingeniería de Rehabilitación	51
Telemedicina y e-Salud	72
Bioinstrumentación	93
Modelado y Simulación	110
Nanomedicina	135
Realidad Virtual	148
Premios SEIB-FENIN	173
Póster I	226
Señales Biomédicas (II)	326
Imágenes Médicas (II)	351
Neurotecnologías (I)	376
Impresión 3D	401
Premios José María Ferrero Corral	426
Póster II	447
Señales Biomédicas (III)	544
Biomateriales	561
Neurotecnologías (II)	584
Póster III	605
ÍNDICE DE AUTORES	690

BIENVENIDA

Estimados congresistas,

En nombre del Comité Organizador, es un honor daros la bienvenida al **XLI Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2023)**, que se celebra en Cartagena del 22 al 24 de noviembre de 2023. Esta edición ha sido organizada por *Universidad Politécnica de Cartagena (UPCT)* con la colaboración de la *Sociedad Española de Ingeniería Biomédica (SEIB)*. Para la UPCT, es un honor poder recibirlos como ya hicimos en 2000 y 2007 en el año en el que celebramos el 25 aniversario de la creación de nuestra Universidad. La organización de este evento ha sido apadrinada por la *International Federation for Medical and Biological Engineering (IFMBE)* y la *European Alliance for Medical and Biological Engineering and Sciences (EAMBES)*.

CASEIB 2023 se configura como el punto de encuentro más importante para todos los agentes implicados en la investigación interdisciplinar en ingeniería biomédica, contando con el apoyo de empresas, instituciones y sociedades médicas. Además, y con la intención de garantizar el crecimiento y el futuro de la Ingeniería Biomédica, en esta edición se continuarán celebrando los concursos destinados a alumnos de Grado, Máster y Doctorado, así como la reunión de los Coordinadores de los Grados en Ingeniería Biomédica de las universidades españolas que imparten dicha titulación.

En esta nueva edición de CASEIB hemos programado una serie de actividades científicas que esperamos sean de vuestro agrado. Con el objetivo de que el evento sea un lugar de enriquecimiento profesional y personal, **CASEIB 2023** presenta los siguientes actos de interés:

- Conferencia Inaugural del Congreso a cargo de *Prof. Dr. Jurijs Dehtjars, Director del RTU FMETA Institute of Biomedical Engineering and Nanotechnologies (BINI) de la Riga Technical University (Letonia)*, con el título “*Nanoscaled electrical functionalization of materials for bio needs*”.
- Conferencia de Clausura del Congreso a cargo del *Dr. Jose Manuel Baena*, ingeniero emprendedor en el campo de la bioimpresión con el título “*Biomedical applications of 3D Bioprinting*”.
- Mesa Redonda titulada “*Aceleración de empresas en Ingeniería Biomédica en el contexto de la Salud Basada en Valor*”, para debatir sobre la importancia del impulso a empresas emergentes de base tecnológica en el entorno de la Ingeniería Biomédica y su contribución a la mejora de la salud.
- Mesa Redonda de Sociedades Médicas, titulada “*El ingeniero biomédico y el profesional sanitario: Aliados en el uso de la IA en el cuidado de la salud*”, que abordará desde un enfoque clínico la llegada de nuevos métodos basados en la inteligencia artificial a la práctica sanitaria y el rol de los ingenieros biomédicos.
- Asamblea General de la Sociedad Española de Ingeniería Biomédica, para debatir las futuras actividades y objetivos de la Sociedad.
- Reunión de Coordinadores de Ingeniería Biomédica, con el fin de establecer políticas comunes para el impulso de esta disciplina desde las distintas universidades españolas.
- *Premio José María Ferrero Corral*, en la que se presentan los cinco mejores trabajos realizados por estudiantes de Máster y Doctorado.
- *Competición SEIB- FENIN de estudiantes de Grado en Ingeniería Biomédica*, en la que se exponen los trece mejores trabajos de estudiantes de este Grado en diferentes Universidades españolas.
- *Premio SEIB-EIT Health*, donde se presentarán los trabajos más emprendedores e innovadores de recientes graduados en Ingeniería Biomédica, en las modalidades “*Patient-Focused Vision*”, “*Healthcare-Delivery Model*” y “*Audience Choice*”.

CASEIB 2023 acoge también una amplia muestra de trabajos científicos en diversas líneas de investigación de la Ingeniería Biomédica. Así, se contempla la presentación de 204 trabajos firmados por un total de 877 autores, incluyendo los 5 trabajos finalistas del *Premio “José María Ferrero Corral”* y los 13 trabajos de estudiantes del *Premio SEIB-FENIN* para estudiantes del Grado en Ingeniería Biomédica. Los 186 trabajos restantes se dividen en los 150 artículos de conferencia y 36 resúmenes extendidos aceptados una vez revisados por el Comité Científico y repartidos entre las diferentes áreas temáticas del congreso. De ellos, 90 se presentan como comunicaciones orales y 96 en formato de póster.

- Comunicaciones orales: Se han planificado 15 sesiones orales organizadas en sesiones paralelas cada uno. En concreto, tres sesiones se dedican a Señales Biomédicas, dos a Imágenes Biomédicas, dos a Neurotecnologías, una a Biomecánica/Ing. Rehabilitación, una a Telemedicina/e-salud, una a Bioinstrumentación, una a Modelado y Simulación, una a Nanomedicina, una a Realidad Virtual, una a Impresión 3D y una a Biomateriales.
- Comunicaciones tipo póster: Se han organizado tres sesiones de pósteres en las que se presentan un total de 96 comunicaciones en diversas áreas de trabajo en Ingeniería Biomédica.

Si bien la mayor parte de trabajos han sido presentados por investigadores españoles, algunos incluyen como coautores a ciudadanos de otros 23 países: Bélgica (12), Francia (9), Colombia (8), Estados Unidos (8), Países Bajos (8), Reino Unido (8), Argentina (6), Irlanda (6), Italia (6), Portugal (6), República Checa (5), Suecia (5), México (4), Chile (3), Finlandia (2), Japón (2), Rumania (2), Suiza (2), Canadá (1), Dinamarca (1), Honduras (1), Líbano (1) y Mongolia (1). De igual forma, entre los asistentes al congreso se ha cuenta con investigadores de Bélgica, Francia, Polonia, Rumanía, Letonia, Bolivia, Colombia, México y Estados Unidos.

Desde el Comité Organizador de **CASEIB 2023** nos gustaría mostrar nuestro agradecimiento a la *Fundación Séneca Agencia de Ciencia y Tecnología de la Región de Murcia* y a las instituciones que han colaborado en la realización del congreso, en especial a la *Federación de Empresas de Tecnologías Sanitarias (FENIN)* y el nodo español del *European Institute of Innovation & Technology in Health (EIT Health Spain)*; así como a la *Sociedad Española de Neumología (SEPAR)*, la *Sociedad Española de Cardiología (SEC)*, la *Sociedad Española de Radiología Médica (SERAM)* y la *Sociedad Española de Directivos de la Salud (SEDISA)*. Este agradecimiento se extiende también al *Centro Europeo de Empresas e Innovación de Cartagena*, el *Instituto de Fomento de la Región de Murcia* y el *Servicio Murciano de Salud* así como a las empresas invitadas a las mesas redondas (*Flongy*, *Kronohealth*, *Inbentus* y *Cella Medical Solutions*), ejemplos del potencial de creación de riqueza del sector.

También queremos reconocer la ardua labor del Comité Científico en la revisión de los trabajos y en la moderación de las sesiones científicas. Asimismo, agradecemos la generosidad de los Conferenciantes invitados y Ponentes de las Mesas Redondas por compartir su experiencia con nosotros. Gracias a la SEIB por confiar en la Universidad Politécnica de Cartagena para celebrar por tercera vez el CASEIB. Finalmente nos gustaría agradecer a todos los miembros del Comité Organizador, en especial a los estudiantes del Grado en Ingeniería Biomédica por su esfuerzo, dedicación e implicación en hacer posible que CASEIB 2023 sea una realidad.

Agradecidos por vuestra confianza, esperamos que disfrutéis de un programa elaborado desde el cariño y nuestro firme compromiso con la Sociedad Española de Ingeniería Biomédica y todos aquellos que hacéis posible que CASEIB siga creciendo en calidad y en cantidad de trabajos y participantes año tras año.

Un abrazo de todo el Comité Organizador,

Joaquín Roca González
Presidente del Comité Organizador del CASEIB 2023

COMITÉ DE HONOR

Excmo. Sr. Fernando López Miras

Presidente de la Comunidad Autónoma de la Región de Murcia

Sra. D^a. Beatriz Miguel Hernández

Rectora Magnífica de la Universidad Politécnica de Cartagena

Excma. Sra. D^a. Noelia María Arroyo Hernández

Alcaldesa del Excmo. Ayuntamiento de Cartagena

Excmo. Sr. D. Juan José Pedreño Planes

Consejero de Salud de la Comunidad Autónoma de la Región de Murcia

Sra. D^a. María Fuensanta Martínez Lozano

Directora de la Fundación para la Formación e Investigación Sanitarias de la Región de Murcia (FFIS)

Sr. D. Pablo Ramirez Romero

Director del Instituto Murciano de Investigación Biosanitaria (IMIB)

Sr. José Francisco Sedes Romero

Director Gerente del Área II de Salud - Servicio Murciano de Salud

Sra. D^a. Catalina Egea Gilabert

Vicerrectora de Investigación, Transferencia y Divulgación de la Universidad Politécnica de Cartagena

COMITÉ ORGANIZADOR

Presidente

Joaquín Roca González

Secretaria

Dolores Ojados González

Vocales UPCT

Joaquín Arias Pardilla
M^a Dolores Avilés González
Manuel Burgos Olmos
Francisco Cavas Martínez
José Fernando Cerdán Cartagena
Javier Cerrillo Moreno
Julián F. Conesa Pastor
Oscar de Francisco Ortiz
Marcos Egea Gutiérrez-Cortines
José Manuel Ferrández Vicente
Patricio Franco Chumillas
Francisco Javier Garrigós Guerrero
José Hernández Grau
José Jódar Ferrández
Jorge Larrey Ruiz
Fulgencio Marín García
Jesús Martínez Frutos
Juan Morales Sánchez
Stella Moreno Grau
Francisco José Ortiz Zaragoza
M^a Dolores Parras Burgos
Nieves Pavón Pulido
José Luis Sancho Gómez
Juan Francisco Sánchez Pérez
Juan Suardiáz Muro
José Alfonso Vera-Repullo
Rafael Verdú Monedero
Rosendo Zamora Pedreño

Vocales Universidad de Murcia

M^a Trinidad Herrero Ezquerro
Virginia Reverte Ribó

Colaboradores SEDIC/UPCT

Israel Baños Rivas
Héctor Flores Aparicio
Gustavo Adolfo Salcedo Eugenio
Marta García Fuentes
Lucía Martínez Abellán
Kamilia Essamhi Saidi
Pilar Sánchez Alcaraz

Estudiantes Grado en Ingeniería Biomédica

Candela Alarcón García
Marta Argudo López
Marta Gea Ponce
Paula Hernández Alegría
Alba López De Lamo
Victoria López De Lamo
Marta Pilar Marín López
Alejandro Martínez Guillermo
Carmen Ortega Escolar
Natalia Sánchez Bastida
Gabriel Sarria Gómez
Joaquín Solano Martínez

Asesores Clínicos

Jose Manuel Allegue Gallego
Pablo Conesa Zamora
Antonio Esquinas Morales
María Galindo Martínez
José Francisco Garrido Peñalver
Claudia Gómez Cárdenas
Ana María Hurtado López
Roberto Jiménez Sánchez
Ginés Luengo Gil
María Teresa Roldán Chicano
Juan Salinas Ramos
Trinitario Sánchez Montón
Gorka Sánchez Nanclares
Carmen Santiago García
Fernando Santonja Medina
José Francisco Sedes Romero
Alfredo Serna Berna
Inmaculada Selles Navarro

Asesores Internacionales

Filip Górski, *Poznan University of Technology (Polonia)*
Jurijs Dehtjars, *Riga Technical University (Letonia)*
Aleksejs Kataševs, *Riga Technical University (Letonia)*
Răzvan Păcurar, *Technical University of Cluj-Napoca (Rumanía)*
Ancuta Păcurar, *Technical University of Cluj-Napoca (Rumanía)*
Miroslav Trajanovic, *University of NIS, Serbia)*
Nikola Vitkovic, *University of NIS, Serbia)*

COMITÉ CIENTÍFICO

Presidente

Enrique J. Gómez Aguilera, *Universidad Politécnica de Madrid.*

Secretario

Joaquín Roca González, *Universidad Politécnica de Cartagena.*

Vocales

Begoña Acha Piñero, *Universidad de Sevilla*

Raúl Alcaraz Martínez, *Universidad de Castilla-La Mancha*

Erik Alonso, *Universidad del País Vasco*

Juan Carlos Antony, *GE Healthcare*

Elisabete Aramendi Ecenarro, *Universidad del País Vasco*

José Luis Arce Diego, *Universidad de Cantabria*

M^a Gloria Bueno García, *Universidad de Castilla-La Mancha*

Ana de los Reyes Guzmán, *Servicio de Salud de Castilla-La Mancha*

Manuel Desco Menéndez, *Universidad Carlos III de Madrid*

Félix Fanjul Vélez, *Universidad de Cantabria*

Javier García Casado, *Universidad Politécnica de Valencia*

María García Gadañón, *Universidad de Valladolid*

Beatriz F Giraldo Giraldo, *Universidad Politécnica de Cataluña*

Carlos Gómez Peña, *Universidad de Valladolid*

María Elena Hernando Pérez, *Universidad Politécnica de Madrid*

Roberto Hornero Sánchez, *Universidad de Valladolid*

Raimon Jané Campos, *Universidad Politécnica de Cataluña*

Pablo Laguna Lasaosa, *Universidad de Zaragoza*

Dolores Ojados González, *Universidad Politécnica de Cartagena*

Jesús Poza Crespo, *Universidad de Valladolid*

Gema Prats Boluda, *Universidad Politécnica de Valencia*

Javier Reina Tosina, *Universidad de Sevilla*

José Joaquín Rieta Ibáñez, *Universidad Politécnica de Valencia*

Isabel Román Martínez, *Universidad de Sevilla*

Francisco Javier Saiz Rodríguez, *Universidad Politécnica de Valencia*

Patricia Sánchez González, *Universidad Politécnica de Madrid*

Daniel Sánchez Morillo, *Universidad de Cádiz*

Andrés Santos Lleó, *Universidad Politécnica de Madrid*

Carmen Serrano Gotarredona, *Universidad de Sevilla*

Ruth Solana Gracia, *Hospital Universitario Infanta Leonor*

Juan Suardiaz Muro, *Universidad Politécnica de Cartagena*

Abel Torres Cebrián, *Universidad Politécnica de Cataluña*

PROGRAMA MIÉRCOLES 22 DE NOVIEMBRE

8:30 - 9:30	REGISTRO (Hall del Antiguo Hospital de Marina)				
9:30 - 11:00	Señales Biomédicas I (Salón de Grados)	Imágenes Médicas I (Paraninfo)	Biomecánica - Ing. Rehab (Aula Sebastián Ferignan)	Telemedicina y e-Salud (Aula PB3)	
11:00 - 12:00	Sesión Póster I - Café - Networking (Claustro)				
12:00 - 13:00	Conferencia Plenaria Inaugural Prof. Dr. Jurijs Dehtjars (RTU IBEN) "Nanoscaled electrical functionalization of materials for bio needs" (Paraninfo)				Exhibición Empresas (Claustro)
13:00 - 13:30	Acto Inaugural (Paraninfo)				
13:30 - 15:00	Comida (Cafetería ETSII)				
15:00 - 16:30	Bioinstrumentación (Salón de Grados)	Modelado y Simulación (Paraninfo)	Nanomedicina (Aula Sebastián Ferignan)	Realidad Virtual (Aula PB5)	
16:30 - 19:00	Premio SEIB FENIN Estudiantes Grado Ingeniería Biomédica (Paraninfo)				

Acto Inaugural

Acto inaugural de CASEIB2022 con la presencia de autoridades.

Intervienen:

Dra. Dña. Beatriz Miguel Hernández, Rectora Magnífica Universidad Politécnica de Cartagena.

Excmo. Sr. D. Juan María Vázquez Rojas, Consejero Medio Ambiente, Universidades, Investigación y Mar Menor.

Dr. D. Enrique Gómez Aguilera, Presidente de la Sociedad Española de Ingeniería Biomédica.

Dr. D. Joaquín Roca González, Presidente del Comité Organizador del CASEIB 2023

Asisten:

Dr. D. Antonio Caballero Pérez, Director general de Universidades de la Región de Murcia

Dr. D. José Sedes Romero Director Gerente Área II – Servicio Murciano de Salud

Sra. D^a. Beatriz Sánchez del Álamo – Concejala de Turismo del Excmo. Ayuntamiento de Cartagena

Premios SEIB-FENIN

Presentaciones de los trece finalistas del Premio SEIB-FENIN de Estudiantes de los Grados en Ingeniería Biomédica.

Conferencia Plenaria Inaugural

"Nanoscaled electrical functionalization of materials for bio needs", Prof. Dr. Jurijs Dehtjars *Riga Technical University (Letonia)*

PROGRAMA JUEVES 23 DE NOVIEMBRE

8:30 - 10:00	Señales Biomédicas II (Salón de Grados)	Imágenes Médicas II (Paraninfo)	Neurotecnologías I (Aula Sebastián Ferignan)	Impresión 3D (Aula PB3)	
10:00 - 11:00	Sesión Póster II - Café - Networking (Claustro)				
11:00 - 12:00	Mesa Redonda I: Aceleración de empresas en Ingeniería Biomédica (Paraninfo)				Exhibición Empresas (Claustro)
12:00 - 13:30	Premio José María Ferrero Corral (Paraninfo)				
13:30 - 15:30	Comida (Cafetería ETSII)				
15:30 - 17:00	Premios SEIB EIT HEALTH (Paraninfo)	Reunión Coordinadores Ingeniería Biomédica (Salón de Grados)			
17:00 - 18:30	Asamblea SEIB (Paraninfo)				

Mesa Redonda I

“Aceleración de empresas en Ingeniería Biomédica en el contexto de la Salud Basada en Valor” Moderada por el Dr. Juan Carlos Antony y el Dr. Joaquín Roca González.

Premio José María Ferrero Corral

Presentaciones de los cinco finalistas del concurso de comunicaciones de estudiantes de Máster y Doctorado.

Premios SEIB-EIT Health

Presentaciones de los trabajos más emprendedores e innovadores de recientes graduados en Ingeniería Biomédica, en las modalidades *"Patient-Focused Vision"*, *"Healthcare-Delivery Model"* y *"Audience Choice"*.

Reunión de Coordinadores de Ingeniería Biomédica

Reunión en formato híbrido de los coordinadores académicos de la titulación del Grado en Ingeniería Biomédica de las distintas universidades españolas.

Asamblea de la SEIB

Asamblea Anual de la Sociedad Española de Ingeniería Biomédica.

PROGRAMA VIERNES 24 DE NOVIEMBRE

8:30 - 10:00	Señales Biomédicas III (Salón de Grados)	Biomateriales (Paraninfo)	Neurotecnologías II (Aula Sebastián Ferignán)	Exhibición Empresas (Claustro)
10:00 - 11:00	Sesión Póster III - Café - Networking (Claustro)			
11:00 - 12:15	Mesa Redonda II: El ingeniero biomédico y el profesional sanitario: Aliados en el uso de la IA en el cuidado de la salud (Paraninfo)			
12:15 - 13:00	Conferencia de Clausura Dr. José Manuel Baena "Biomedical applications of 3D Bioprinting" (Paraninfo)			
13:00 - 14:00	Acto de Clausura y Entrega de Premios (Paraninfo)			

Mesa Redonda II

“El ingeniero biomédico y el profesional sanitario: Aliados en el uso de la IA en el cuidado de la salud” Moderada por el Dr. Enrique J. Gómez Aguilera y la Dra. Ruth Solana,

Conferencia Plenaria de Clausura

“Biomedical applications of 3D Bioprinting” por el Dr. Jose Manuel Baena

Acto de Clausura y Entrega de Premios

Acto de clausura de CASEIB 2023 y entrega del Premio José María Ferrero Corral, las distinciones de la Competición SEIB-FENIN y los premios SEIB a la mejor Tesis Doctoral en 2022 y premio SEIB al mejor artículo científico publicado en 2022 por jóvenes investigadores.

CONFERENCIA PLENARIA INAUGURAL

Miércoles 22 noviembre 12:00 - 13:00, Paraninfo



Nanoscaled electrical functionalization of materials for bio needs

Prof. Dr. Jurijs Dehtjars

Institute of Biomedical Engineering and nanotechnologies
Riga Technical University Riga, Latvia
jurijs.dehtjars@rtu.lv

1. Motivation

A microorganism or biomolecule communicates with the surface of a substrate at the nanoscale. There are several approaches to functionalize the surface of the substrate aimed to control the attachment of microorganisms or biomolecules. Following the data by SCOPUS (2023), surface coating and morphology engineering are mostly exploited. However, in the last decade, the amount of publications on electrical functionalization of the surface increased significantly. The present report reviews physical fundamentals [1] and technological approaches for surface electrical functionalization as well as its influence on attachment and properties of microorganisms and biomolecular structures.

2. Functionalization, microorganisms and biomolecular structures

Electrical functionalisation (EF) is applicable for materials which do not have electrical conductivity. EF impacts the wettability of the substrate surface [2]. In particular, the polar component of the contact angle measurements has the strongest connection with the electrical potential of the functionalized surface. To measure the charge at the nanolayer of the surface the near-threshold photoelectron spectroscopy is demonstrated for use [3]. The easiest way to engineer the electrical charge of the surface is based on the application of radiation [4]. In this case the electrical charge is excited to the specific local states. Another way focuses reconstruction of a crystalline lattice of the substrate. The example is given for the hydroxyapatite ($\text{Ca}_5(\text{PO}_4)_3\text{OH}$) crystal which is widely used for bone implants. Rearrangement of the crystal ions may be done because of the proton shift against the oxygen atom [5,6]. Doping of the hydroxyapatite (HAP) with magnesium, manganite, silver or strontium also demonstrates a capacity to engineer the surface electrical charge [7]. Reconstruction of HAP nanoparticle lattice is possible because of the engineering of its size. This results surface density of the positively charged protons: the higher curvature of the particle surface more its positive charge increases [8]. Following such the result the surface roughness induces the specific distribution of the charge of the surface [9]. Sharp peaks of the surface roughness are similar to the small nanoparticles. However, the surface having peaks with a low curvature is like the large nanoparticle. Therefore, the sharp peaks are more positively charged against the peaks having low curvature [9]. Different microorganisms (osteoblastic cells, *Bacillus subtilis*, *Saccharomyces cerevisiae* (yeast)), virus-like particles (VLP) demonstrate a higher capacity to be attached to the substrate (HAP, PMAA, SiO_2 nanoparticles, Si) if the latter is charged negatively, the proliferation of the osteoblastic cells, uptake of heavy metal ions by yeast cells, and induction of the specific antibodies by VLP attached to SiO_2 nanoparticles increasing [10-14]. The coupling of the charged substrate with yeast cells depends on their dimensions because the latter corresponds to the dipole moment of the cell [15].

3. Conclusion

Electrical functionalization of the substrates may become a tool to control cell and biomolecule attachment as well as their use.

References

- [1] K. Baltacis, V. Bystrov, A. Bystrova, Yu. Dekhtyar, T. Freivalds, J. Raines, K. Rozenberga, H. Sorokins, M. Zeidaks. Physical Fundamentals of Biomaterials Surface Electrical Functionalization. *Materials* 2020, V13(20), 4575; <https://doi.org/10.3390/ma13204575>, 28 p
- [2] Dekhtyar Y., Lancere L. Polyaka N., Sudnikovich A., Tyulkin F., Valters V. PMMA wettability caused by ultraviolet radiation// *Advanced Materials Research* 222 – 2011 – p.322-325.
- [3] Yu. Dekhtyar. Weak electron emission current for characterization of nanomaterials, gas and radiation sensing towards medical applications. *Proc. of Estonian Academy of Sciences*, 2014, V63 (3), pp 258-266.
- [4] A. Rapoport, D. Borovikova, A. Kokina, A. Patmalnieks, N. Polyak, I. Pavlovska, G. Mezinskis, Yu. Dekhtyar. Immobilisation of yeast cells on the surface of hydroxyapatite ceramics. *Process Biochemistry*, 2011, 46, 665–670.
- [5] V. Bystrov, N. Bystrova, Yu. Dekhtyar, S. Filippov, A. Karlov, A. Katashev, C. Meissner, E. Paramonova, A. Patmalnieks, N. Polyaka, A Saprionova. Size Depended Electrical Properties of Hydroxyapatite Nanoparticles. *IFMBE proceedings*. V 14, 2006, pp.3149-3150.
- [6] V. Bystrov, Yu. Dekhtyar, A. Karlov, A. Katashev, C. Meissner, A. Patmalnieks, E. Paramonova, N. Polyaka. Size related electron work function of hydroxyapatite nanoparticles. *Scientific reports. Journal of the Univeristy of Applied Sciences Mittweida*. 2006. Nr. 10. Pp 19-21.
- [7] V S Bystrov, J Coutinho, A V Bystrova, Yu. Dekhtyar, R C Pullar, A Poronin, E Palcevskis, A Dindune, B Alkan, C Durucan, E V Paramonova. Computational study of Hydroxyapatite structures, properties and defects. *J. Phys. D: Appl. Phys.* 48 (2015) 195302 (20pp) doi:10.1088/0022-3727/48/19/195302,
- [8] V S Bystrov, Yu. Dekhtyar, E Paramonova, A. Katashev, A. Karlov, N Polyaka, A V Bystrova, A Patmalnieks, A L Kholkin. Computational and experimental studies of size and shape related physical properties of hydroxyapatite nanoparticles. *J. Phys.: Condens. Matter* 2011, V23, 065302 (10pp)
- [9] Yu. P. Sharkeev, K. S. Popova, K. A. Prosolov, E. Freimanis, Yu. Dekhtyar, I. A. Khlusov. Electrical Potential and Topography of the Surface of a Calcium-Phosphate Coating Deposited with RF-Magnetron Discharge., *Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques*, 2020, Vol. 14, No. 1, pp. 200–206.
- [10] Yu. Dekhtyar, Polyaka N., Sammons R. Electrically charged hydroxyapatite enhances immobilization and proliferation of osteoblasts. *14th Nordic -Baltic conference on Biomedical engineering and medical physics*. Riga, Latvia, 2008. 134.
- [11] Yu. Dekhtyar, M. Romanova, A. Kachanovska, D. Skrastina, R. Renhofa, P. Pumpens, A. Patmalnieks. Silica nanoparticles for improving efficiency of virus-like particle-based hepatitis B vaccine. *Int conference on nanotechnologies and biomedical engineering*, Chisinau, Moldova, 7-8 July, 2011, pp344-346.
- [12] Yu. Dekhtyar, A. Kachanovska, G. Mezinskis, A. Patmalnieks, P. Pumpens, R. Renhofa Self assembled system of semiconductor and virus like nanoparticles., grāmatā pp 347-350. In book: *Functionalized Nanoscale Materials, Devices, and Systems for chem.-bio Sensors, Photonics, and Energy Generation and Storage*", Springer, 2008.
- [13] D. Skrastina, I. Petrovskis, I. Lieknina, J. Bogans, R. Renhofa, V. Ose, A. Dishlers, P. Pumpens, Yu. Dekhtyar. Silica Nanoparticles as the Adjuvant for the Immunisation of Mice Using Hepatitis B Core Virus-Like Particles. *PLOS ONE* (2014) <https://doi.org/10.1371/journal.pone.0114006>, pp 1-17.
- [14] Dekhtyar Yu., Lancere L., Polyaka N., Sudnikovich A., Tyulkin F., Valters V. Ultraviolet radiation influences wettability and biological properties of PMMA surface// *Biomedical Engineering conference, Kaunas, Lithuania – 2009* - p.69-71.
- [15] Bondareva, E., Dekhtyar, Y., Gorosko, V., Sorokins, H., Rapoport, A. Correlation of Statistical Distributions of the Dimension of Yeast Cells Attached to the Substrate and Its Surface Electrical Potential. *Materials* 2022, 15 (6), <https://doi.org/10.3390/ma15010006>; 10.

CONFERENCIA PLENARIA DE CLAUSURA

Viernes 24 noviembre 12:15 - 13:00, Paraninfo



Biomedical applications of 3D Bioprinting

Dr. Jose Manuel Baena

Biofabrication and cell culture lab, REGEMAT 3D
Avenida del conocimiento 41, 18016 Granada, Spain;
ceo@regemat3d.com

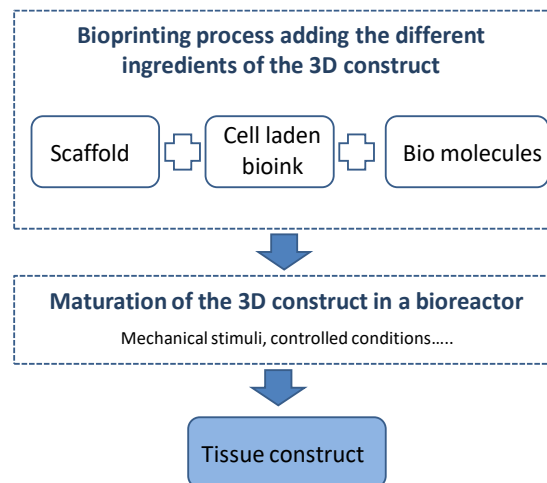
Recent developments in additive manufacturing (AM) or 3D printing technologies have made it possible to manufacture meshed structures in a wide range of biomaterials and to generate three-dimensional cell cultures in the laboratory with high cell viability, a wide range of technologies known as 3D bioprinting.

However, this is far from being a functional tissue, with its histological structure and behavior. There is still a lot of work to be done to stimulate these structures in such a way that the extracellular matrix (ECM) of the desired tissues is generated. Although it may seem that the clinical use of bioprinted tissues is part of the day to day activity of a hospital or that we are close to produce fully functional organs in the lab, this is unfortunately not the true. The potential of Bioprinting in creating functional tissues is huge but we may be nowadays in a technology hype as it happened before with 3D printing. Nowadays after many years and clinical cases 3D printing are recognized amongst the different technologies available to manufacture sanitary products, the technology is adding value and finding a space in the day to day activity of clinical centers.

The potential of bioprinting is immense but we in the industry may be missing one part of the procedure, the maturation, a 3D printed scaffold with cells is not a tissue, neither an organ, and should be ethical and care more about the news and expectations of people reading it. Marketing is needed in promoting emerging technologies but care has to be taken when talking about problems that affect to many patients around the world. The lack of tissue regeneration in human beings, the deficiency of allogeneic transplants, and longer average life expectancy make the creation of functional tissues in the laboratory one of the most important problems of humanity right now. However, truth be told, the results obtained are still less than desired. Many new bioprinting groups have been created around the world in the past few years, and a variety of commercial systems are now available to the researchers. More and more publications on the matter are growing steady. Nevertheless, the results obtained are still far from true clinical application. For example, the organogenesis efforts well covered by the media, including the regeneration of the skin, cornea, or the heart are still too far away from human testing. In addition, a common misconception by the industry is that we need to directly create functional tissue, but in reality, we are creating a matrix loaded with cells. The key is to make these cells to behave as they do in vivo and to promote the creation of functional tissues. This requires defining the right biofabrication strategy, but also the right maturation strategy.

Here we present two blocks of processes to be defined to design a successful strategy to create a functional living tissue. The first block relates to the ingredients to be printed and the bioprinting part. The second block relates to the maturation of the construct using physiological and anatomical like stimuli and conditions. For the creation of living tissue, both the bioprinting process and the maturation of the construct are crucial. Recreating human adult conditions in the lab or the stimuli that occur during embryogenesis will move the results of tissue engineering closer to clinical application. In our research we have proposed a unique 2-blocks approach to create functional tissues: 1) Bioprinted constructs and a device that applies the right mechanical conditions on the structure to be replaced, mimicking physiology. 2) The maturation procedure, applying the right stimuli, creating the final functional tissue. We think that the best stress distribution is the real one, and other approaches failed as they do not closely mimic the true physiologic conditions happening in nature.

For the creation of a living tissue, it is crucial the bioprinting process and the ingredients selected to achieve the objective to create a functional specific tissue (first block) the scaffold, the cells, the bio inks and the other ingredients to be printed that will promote the formation of the right tissue. But also, the maturation procedure applied to the 3D cell-laden constructs, that is even more important (second block). If we think about bioprinting as a technology to recreate all the structure in the same form as shown in living tissue, we are going to fail. We have to think of bioprinting as a way of creating cell-laden 3D constructs as a precursor of functional tissue. The maturation and tissue formation process will be as important or even more than the bioprinting step. Focusing on the strategies related to both blocks in the diagram will be important to obtain the desired functional knee cartilage tissue, for example.



Literature search and real-life experience tell us that mechanical stress distribution is crucial as stimuli to create the right tissue. Nonloaded muscle is absorbed, the non loaded bone loses its matrix. Also, the scaffold architecture as it will affect the stimuli distribution and other important parameters as the biodegradation time. The selection of the right ingredients and the bioprinting procedure is very important in the success of the creation of functional tissues, as well as the maturation procedure applied to the 3D cell-laden constructs is even more important. This approach will open a wide research area for tissue engineers to develop protocols with different stimuli to create functional tissues, either using direct or indirect (using molds as temporal containers, a fiber structure holding loads and a cell-friendly matrix, even adipose tissue containing blood vessels) bioprinting methods. Bioprinting is going to be losing importance while the maturation process will increase in publications and results in the next years. This will make possible the generation of functional vascularized, ready to be used tissues and organs.

There is still a lot of work to do but the results could improve the life of millions of people. Let's see what the future brings us.

- [1] Murphy, S., Atala, A. 3D bioprinting of tissues and organs. *Nat Biotechnol* 32, 773–785 (2014). <https://doi.org/10.1038/nbt.2958>.
- [2] Baena, J.M.; Jiménez, G.; López-Ruiz, E.; Antich, C.; Griñán-Lisón, C.; Perán, M.; Gálvez-Martín, P.; Marchal, J.A. Volume-by-volume bioprinting of chondrocytes-alginate bioinks in high temperature thermoplastic scaffolds for cartilage regeneration. *Exp. Biol. Med.* 2019, 244, 13–21. <https://doi.org/10.1177/1535370218821128>.
- [3] J. Melchor; E. López-Ruiz; J. Soto; G. Jiménez; C. Antich; M. Perán; J.M. Baena; J.A. Marchal; G. Rus (2018). "In-bioreactor ultrasonic monitoring of 3D culture human engineered cartilage." *Sensors and Actuators B: Chemical* 266: 841-852.
- [4] Oberdiek, F.; Vargas, C.I.; Rider, P.; Batinic, M.; Görke, O.; Radenković, M.; Najman, S.; Baena, J.M.; Jung, O.; Barbeck, M. Ex Vivo and In Vivo Analyses of Novel 3D-Printed Bone Substitute Scaffolds Incorporating Biphasic Calcium Phosphate Granules for Bone Regeneration. *Int. J. Mol. Sci.* 2021, 22, 3588. <https://doi.org/10.3390/ijms22073588>
- [5] Ferrís-Oñate, J., Morales-Martín, I., Blanco-Fernández M, Baena-Martínez JM, Delgado-Gordillo J, Hurtós- Casals E, Atorrasagasti-Goyalde G, Bermejo-Bosch I, Atienza-Vicente CM, Solera-Navarro MJ (2010). "FABIO project: Development of a new generation of customized medical devices. *Revista de Biomecánica* 53: 75-78.
- [6] D. Martínez-Moreno, G. Jiménez, P. Gálvez-Martín, G. Rus, J. A. Marchal, Cartilage biomechanics: A key factor for osteoarthritis regenerative medicine. *Biochim. Biophys. Acta Mol. Basis Dis.* 2019, 1865, 1067.

MESA REDONDA I

Jueves 23 noviembre 11:00 - 12:00, Parainfo

Aceleración de empresas en Ingeniería Biomédica en el contexto de la Salud Basada en Valor

La ingeniería biomédica, centrada en el diseño, desarrollo, fabricación y aplicación de tecnologías para la salud está experimentando un importante desarrollo en España, impulsado por la creciente demanda de soluciones innovadoras para la atención sanitaria. En este contexto, la aceleración de empresas de base tecnológica en el ámbito de la ingeniería biomédica se presenta como una oportunidad clave para propiciar el desarrollo de nuevas soluciones sanitarias que contribuyan a mejorar la salud y el bienestar de la población bajo el enfoque de la salud basada en valor (SBV).

La SBV busca optimizar los resultados de salud para los pacientes a un coste razonable, planteando nuevos retos para las empresas de ingeniería biomédica en el desarrollo de productos y servicios que sean eficaces, seguros y rentables. Así, estas empresas deben demostrar el valor de sus soluciones, buscando siempre el beneficio real (para los pacientes y para el sistema sanitario). Por otra parte, y debido a la creciente complejidad de la atención sanitaria, las empresas deben ser capaces de adaptarse a los nuevos modelos de atención sanitaria, ofreciendo soluciones integradas y personalizadas. Finalmente, estas empresas deben colaborar con otros actores del sistema sanitario, como los hospitales, los centros de investigación y los proveedores de atención sanitaria con el fin de ofrecer productos innovadores que respondan a las necesidades reales del sector.

En este contexto, las empresas de base tecnológica (EBT) en el ámbito de la tecnología médica (TM) tienen un papel clave que jugar; desarrollando soluciones innovadoras que ayuden a mejorar los resultados de salud de los pacientes a un coste menor. Sin embargo, la realidad nacional hace que estas empresas se enfrentan a distintos problemas que dificultan su introducción en el mercado de la tecnología médica:

- La financiación es uno de los principales retos a los que se enfrentan las EBT en España, pues los inversores tradicionales, suelen tener dificultades para comprender la complejidad, alcance y valor introducido por las soluciones de TM.
- Las soluciones de TM suelen estar sujetas a una regulación compleja, lo que puede dificultar su desarrollo y comercialización.
- La competencia en el sector de la TM es cada vez mayor, por lo que las EBT españolas tienen que competir con empresas multinacionales que disponen de mayores recursos y experiencia.

La aceleración de EBTs puede ser una herramienta eficaz para ayudar a las empresas de ingeniería biomédica a afrontar los retos de impuestos por la SBV; ayudando a las empresas emergentes a desarrollar su negocio y a escalar su actividad mediante el asesoramiento, la mentoría de expertos, y el acceso a recursos financieros y de networking. En España, existen diversas iniciativas que promueven la aceleración de empresas de base tecnológica en el ámbito de la ingeniería biomédica que les permiten:

- Aprovechar la experiencia de los expertos en aceleración para desarrollar soluciones que sean eficaces, seguras y rentables.
- Acceder a recursos para realizar ensayos clínicos y análisis de coste-efectividad.

- Colaborar con otros actores del sistema sanitario mediante programas específicos de aceleración (Hospitales, Universidades, Centros de Investigación y Servicios de Salud).

Es de esperar que los programas de aceleración EBTs en el entorno de las tecnologías médicas favorezca la introducción de soluciones innovadoras centradas en el valor y creadas desde la demanda en los sistemas de salud.

En esta mesa redonda se pretende presentar los distintos aspectos clave del proceso específico de aceleración de EBTS en el entorno de las tecnologías médicas gracias a la experiencia de los ponentes invitados y de cuatro EBTs que participaran presentando su trayectoria.

Moderadores

- **Juan Carlos Antony**, Iberia Regional Executive Manager PCS at GE HealthCare, Vicepresidente Adjunto de la Fundación Tecnología y Salud de FENIN, Vocal de Industria, emprendimiento y transferencia tecnológica en salud SEIB.
- **Joaquín Roca González**, Coordinador del Grado en Ingeniería Biomédica en la UPCT, Presidente del Comité Organizador CASEIB 2023.

Ponentes

- **Sergio Muñoz Lezcano**, Director de Innovación, Salud Digital y Tecnologías Emergentes en Fenin (Federación Española de Empresas de Tecnología Sanitaria).
- **Izabel Alfany**, Directora General de EIT Health Spain
- **Carlos Hurtado Mengual**, Director del Centro Europeo de Empresas e Innovación de Cartagena y responsable del programa “CreceStartup Health”. Experto en creación y promoción de startups y EBTs, y sistematización de la innovación en pymes
- **Gorka Sánchez Nanclares**, Responsable de Innovación del Servicio Murciano de Salud. Experto en Co-creación desde la Demanda y Compra Pública Innovadora.

Empresas invitadas

- **Manuel A. Burgos**, CEO en Flowgy (www.flowgy.com). Modelado y simulación digital para diagnóstico y planificación quirúrgica de la vía aérea.
- **María José Martínez Madrid**, CEO en Kronohealth (www.kronohealth.com). Asesoría circadiana orientada a mejorar tu descanso y resolver problemas de sueño.
- **Rafael Valverde**, CEO en Inbentus (www.inbentus.com) Ventiladores mecánicos de altas prestaciones basados en tecnología de turbina.
- **Miguel Rodríguez**, CTO en Cella Medical Solutions (www.cellams.com). Soluciones tecnológicas 3D a medida para cirugías complejas.

MESA REDONDA II

Viernes 24 noviembre 11:00 - 12:15, Paraninfo

El ingeniero biomédico y el profesional sanitario: Aliados en el uso de la IA en el cuidado de la salud

Las soluciones de inteligencia artificial (IA) están transformando la forma de brindar los cuidados de salud, generando grandes expectativas en la medicina en todo el mundo e igualmente abriendo un debate sobre el modo en que la ética y los derechos humanos deber ocupar un lugar central durante su concepción, despliegue y utilización. Las organizaciones sanitarias han acumulado vastos conjuntos de datos en forma de registros, señales e imágenes y se ha producido un acercamiento de las nuevas tecnologías al ciudadano con diversidad de dispositivos de monitorización en tiempo real.

La combinación de la ingeniería biomédica y la IA ha permitido avances significativos en la atención médica. La ingeniería biomédica, como disciplina que aplica los principios de la ingeniería en medicina, se enfoca en el diseño y desarrollo de dispositivos médicos, terapias y tecnologías adaptadas a las necesidades clínicas para mejorar la calidad de vida de los pacientes. Por otro lado, la IA permite automatizar tareas repetitivas, apoyar la detección temprana de enfermedades. Mejorar la capacidad del diagnóstico y la eficacia de los tratamientos, ayudar en la toma de decisiones clínicas, complementar la experiencia formativa de los profesionales, reducir tiempos de investigación, optimizar tiempo y costes, y aumentar la calidad de vida de los pacientes, mejorando su grado de adhesión a tratamientos.

Todas estas capacidades de la IA se basan en su capacidad de análisis de datos y la creación de modelos predictivos para mejorar la precisión del diagnóstico, el tratamiento y la prevención de enfermedades. La IA puede analizar grandes cantidades de datos médicos para identificar patrones y predecir resultados, que ayudan a los sanitarios a tomar decisiones más informadas sobre el diagnóstico y el tratamiento de patologías e identificar pacientes que tienen un mayor riesgo de desarrollar ciertas enfermedades, así como favorecer una intervención temprana. En el ámbito del análisis de la imagen y las señales, los algoritmos de IA son capaces de detectar anomalías que podrían pasar desapercibidas para los clínicos. Además, la IA facilita ayuda a los médicos para desarrollar planes de tratamiento personalizados en pacientes con enfermedades crónicas.

La IA también ha tenido un impacto significativo en el campo educativo. Los estudiantes de ciencias médicas y de ingeniería biomédica pueden utilizar herramientas basadas en IA para mejorar su aprendizaje y comprensión del material. De igual modo, la IA aporta valor en áreas como la gestión de la salud, mejorando notablemente la eficiencia y reduciendo los costos en el sector salud. Por ejemplo, permite el análisis de los flujos de los procesos asistenciales optimizando los recursos y facilitando la accesibilidad a la atención, la generación de chatbots que ayuden reducir el tiempo que los pacientes pasan esperando para hablar con un profesional sanitario y reducir el número de visitas innecesarias al hospital o clínica, etc.

Por ello, en 2035 se espera que la atención sanitaria sea hiperpersonalizada, digitalizada y alimentada por IA. Los hospitales del futuro serán centros más pequeños y capacitados para lo más complejo; con menos camas y menos profesionales; con estancias más cortas; crecientemente especializados; y, sobre todo, escalables. Los hospitales estarán diseñados en torno al consumidor, en lugar de exigir que sea el paciente el que se adapte a sus condiciones. Los pacientes podrán recibir atención sanitaria desde la comodidad de sus hogares y la tecnología esté perfectamente integrada en nuestra vida cotidiana.

El trabajo conjunto entre el profesional sanitario, el ingeniero biomédico y los pacientes será imprescindible a la hora de abordar los grandes retos de transformación del modelo sanitario que nos depara el futuro, si no ya presente.

Con esta mesa, pretendemos dar una visión introductoria sobre la IA en ingeniería biomédica y realizar un recorrido por las principales aplicaciones que nos ofrece esta tecnología en las diferentes áreas de conocimiento invitadas a debate.

Moderadores

- **Prof. Dr. Enrique J. Gómez Aguilera**, Catedrático de la UPM y Presidente de la Sociedad Española de Ingeniería Biomédica (SEIB)
- **Dra. Ruth Solana**, Facultativo del Hospital Universitario Infanta Leonor y Vocal de Sociedades Médicas SEIB

Ponentes

- **Dr. Carlos Almonacid Sánchez**, Jefe de Servicio de Neumología Hospital Universitario Puerta de Hierro Majadahonda, Vicepresidente Neumólogo de la Sociedad Española de Neumología (SEPAR), Vocal de la Junta Directiva de la Federación de Asociaciones Científico Médicas de España (FACME).
- **Dr. José María de la Torre Hernández**, Jefe del Servicio de Cardiología en el Hospital Universitario Marqués de Valdecilla, Santander. Vicepresidente de la Sociedad Española de Cardiología (SEC).
- **Dr. Luis Concepción Aramendía**, Jefe Servicio Radiodiagnóstico Hospital General Universitario Dr. Balmis de Alicante, Responsable de Asuntos Profesionales en Junta Directiva de la Sociedad Española de Radiología Médica (SERAM).
- **Dr. Francisco Soto Bonel**, Presidente de la Sociedad Española de Directivos de la Salud (SEDISA)

PREMIOS

Premio José María Ferrero Corral

Jueves 23 de noviembre, 12:00h – Paraninfo

CASEIB 2023 organiza con la colaboración de la SEIB el Premio José María Ferrero Corral para estudiantes de Máster y Doctorado en el área de la Ingeniería Biomédica. Los cinco finalistas presentarán sus trabajos durante 10 minutos, a los que le seguirán 5 minutos de debate con el tribunal

Tribunal:

- Prof. Dr. Enrique Gómez Aguilera. Universidad Politécnica de Madrid.
- Prof. Dr. Luis Javier Reina Tosina. Universidad de Sevilla.
- Prof. Dr. Roberto Hornero Sánchez. Universidad de Valladolid.

Trabajos seleccionados:

- Felipe Arias-González, Alejandra Rodríguez-Contreras, Miquel Punset, José María Manero, Óscar Barro, Mónica Fernández-Arias, Fernando Lusquiños, Javier Gil y Juan Pou, “*Very low elastic modulus Ti alloys obtained by Laser Directed Energy Deposition to avoid bone resorption in bone implants*”.
- Violeta Puche-García, Laura Martínez-Mateu, David Filgueiras-Rama, Lucía Romero y Javier Saiz, “*Comparación in-silico de la eficacia de los fármacos flecainida y vernakalant en fibrilación auricular paroxística*”.
- Amaia Sanz, Erik Alonso, Iraia Isasi, Andima Larrea, Ruth Salaberria y Elisabete Aramendi. “*Análisis wavelet de la señal de oximetría cerebral para el cálculo de la frecuencia de las compresiones del masaje de resucitación cardiopulmonar*”.
- Gema Quiñonero López, Juan Gallo, Andreia Magalhaes, Josep Samitier y Aranzazu Villasante. “*Development of a neuroblastoma-on-a-chip bioengineered platform enabling the translational study of emerging nanotherapies*”.
- Verónica López González, Alicia Pose Díez de La Lastra, Juan de León-Luis y Javier Pascau. “*Augmented reality simulator for laparoscopic surgery on a realistic phantom*”.

Premios SEIB - FENIN de Estudiantes de Grado en Ingeniería Biomédica

Miércoles 22 de noviembre, 16:30h – Paraninfo

CASEIB 2023 acoge un año más la competición de estudiantes de Grado en Ingeniería Biomédica coordinada por la SEIB y la FENIN. En esta edición del CASEIB, hay trece finalistas seleccionados de entre todos los trabajos enviados con resultados relevantes en el campo de la Ingeniería Biomédica de diferentes Universidades, ya sea como parte de su Trabajo Fin de Grado o de trabajos de asignaturas de la titulación. Durante la sesión, los estudiantes expondrán oralmente su trabajo durante 7 minutos y debatirán con el tribunal durante 3 minutos.

Tribunal:

- Prof. Dra. Elisabete Aramendi Ecenarro. Universidad del País Vasco.
- Prof. Dr. Javier Saiz Rodríguez. Universidad Politécnica de Valencia.
- Prof. Dr. Luis Javier Reina Tosina. Universidad de Sevilla.

Trabajos seleccionados:

- Ángela Romo Marín, Carla De Gregorio Bermejo, Beatriz De la Calle, Cristina Simón-Martínez y Mario Martínez-Zarzuela, “*Segmentación de patrones de marcha en pacientes con parálisis cerebral mediante aprendizaje profundo y visión artificial*”. Universidad de Valladolid.
- Marta Pareja Jiménez, Enric Casanova Batlle, Joaquim Ciurana Gay, Emma Polonio Alcalá y Teresa Puig Miquel. “*Preliminary screening of polycaprolactone-polystyrene electrospun scaffolds for the formation of tumoroids*”. Universidad de Girona.
- Ramón Domínguez, Cristina Soguero Ruiz y David Chushig Muzo. “*Modelos basados en redes neuronales artificiales para la generación y predicción de series temporales de glucosa*”. Universidad Rey Juan Carlos.
- Ivan Luque Del Toro, Mireia Masias Bruns y Gemma Piella Fenoy. “*Computer-aided detection system for pulmonary embolism with integrated cardiac assessment based on embolic burden*”. Universitat Pompeu Fabra.
- Lourdes Carreras-Vidal, África Martínez-Blanco, Sergio Noé, Jordi Alcaraz, Ramon Farré y Nuria Gavara. “*Study of fibroblasts activation kinetics and identification of fibroblasts subpopulations in physiological and pathological conditions*”. Universitat de Barcelona.
- Martín Ruiz Gutiérrez, Ainoa Tejedera Villafranca, Beatriz Giraldo, Javier Ramón Azcón y Juan Manuel Fernández Costa. “*Development of an electric stimulation system to analyze muscle damage in 3D Duchenne muscular dystrophy muscle culture*”. Universitat Politècnica de Catalunya.
- David Morales Loro, Tomás Palacios Gutiérrez y Andrés Díaz Lantada. “*Chamber design of a portable breathalyzer for disease diagnosis*”. Universidad Politécnica de Madrid.
- Ángel Rodríguez Ballabriga, Albert Fabregat-Sanjuan, Agnes Rigo Vidal, Rosa Pamies-Vilà, Vanesa Rius Costa, Susana Larrosa Capacés y Vicenç Pascual Rubio. “*Análisis de electroencefalografía de amplitud integrada en neonatos: Evaluación del rendimiento de aCUP-E frente a los electrodos de gel líquido*”. Universitat Rovira i Virgili de Tarragona.

- Celina Salamanca-Gonzalez, Ilaria Sorrentino, Angie K. Molina-Oviedo, Eduardo Arevalo-Nunez de Arenas y Iria Medrano-Fernandez. *“Analyses of H₂O₂ production in time during wound healing”*. Universidad Carlos III de Madrid.
- Neus Pons Andrés, Estela Villanueva Bádenas, Gloria Gallego Ferrer y Laia Tolosa Pardo. *“Nuevos modelos multicelulares para la predicción de la hepatotoxicidad idiosincrática”*. Universitat Politècnica de València.
- Elena Agulló Mertens y David Bru Orts. *“Estudio numérico para consolidación mediante fijación de fracturas verticales en cuello de fémur”*. Universidad de Alicante.
- Alicia Aniorte Gacía, Lola Ojados González, Gustavo Salcedo Eugenio. *“Aplicación de Realidad Aumentada enfocada a la Salud Bucodental en edades tempranas”*. Universidad Politécnica de Cartagena.
- Laiene Azkue, Jon Kerexeta Sarriegi, Nekane Larburu Rubio and Unai Ayala Fernandez. *“Optimización de Servicios de Urgencias mediante Modelos de Inteligencia Artificial: Revisión Sistemática de la Predicción de Ingresos Hospitalarios desde Urgencias”*. Mondragon Unibertsitatea.

ACTOS

Acto inaugural

Miércoles 22 de noviembre, 13:00 h – Paraninfo

La XLI edición del Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2023) contará en su inauguración con la presencia de las siguientes autoridades, en representación de la Universidad Politécnica de Cartagena, la Consejería de Medio Ambiente, Universidades, Investigación y Mar Menor de la Región de Murcia, la Sociedad Española de Ingeniería Biomédica y el Comité Organizador de CASEIB 2023:

- Dra. Dña. Beatriz Miguel Hernández, Rectora Magnífica Universidad Politécnica de Cartagena.
- Excmo. Sr. D. Juan María Vázquez Rojas, Consejero Medio Ambiente, Universidades, Investigación y Mar Menor de la Región de Murcia.
- Dr. D. Enrique Gómez Aguilera, Presidente de la Sociedad Española de Ingeniería Biomédica.
- Dr. D. Joaquín Roca González, Presidente del Comité Organizador del CASEIB 2023

Asamblea de la SEIB

Jueves 23 de noviembre, 17:00 h – Paraninfo

Durante el CASEIB 2023 se celebrará la asamblea anual de la Sociedad Española de Ingeniería Biomédica.

Acto de clausura

Viernes 24 de noviembre, 13:00 h – Paraninfo

El XLI Congreso Anual de la Sociedad Española de Ingeniería Biomédica será clausurado por los representantes de la Sociedad Española de Ingeniería Biomédica y del Comité Organizador de CASEIB 2022. Durante la ceremonia de clausura se hará entrega del Premio José María Ferrero Corral, las distinciones de la Competición SEIB-FENIN y los premios SEIB a la mejor Tesis Doctoral en 2022 y premio SEIB al mejor artículo científico publicado en 2022 por jóvenes investigadores.

- Dr. D. Enrique Gómez Aguilera, Presidente de la Sociedad Española de Ingeniería Biomédica.
- Dr. D. Joaquín Roca González, Presidente del Comité Organizador del CASEIB 2023

ÍNDICE DE SESIONES CIENTÍFICAS

Miércoles 22 de noviembre de 2023

Sesión Señales Biomédicas (I)

Presentación Oral

9:30 - 11:00

Salón de Grados

Moderan: *Raquel Bailón Luesma y José Francisco Garrido Peñalver*

- | | |
|--|----|
| Los cambios en la morfología del pico al fin de la onda T, cuantificados mediante "time-warping", están asociados con la aparición de fibrilación ventricular en un modelo porcino de isquemia aguda | 2 |
| Dispersión espacial de los tiempos de activación y repolarización asociada a diferentes modos de estimulación cardíaca | 6 |
| Detección automática de residuos de la actividad ventricular presentes en las ondas fibrilatorias extraídas de registros electrocardiográficos de fibrilación auricular | 10 |
| ECG-ENET: Red neuronal convolucional explicable para la ayuda en el diagnóstico de la apnea del sueño infantil | 14 |
| Análisis de la variabilidad de la frecuencia cardíaca en maratonianos | 18 |
| Valoración de la actividad muscular asociada al uso de herramientas laparoscópicas mediante EMG. Nueva propuesta de análisis bivariado amplitud-frecuencia. | 22 |

Sesión Imágenes Médicas (I)

Presentación Oral

9:30 - 11:00

Paraninfo

Moderan: *Carmen Serrano Gotarredona y Andrés Santos Lleó*

- | | |
|--|----|
| Automatic calculation of pelvis morphology from CT images | 27 |
| Deep Learning methodologies for direct image reconstruction and integrated attenuation correction in brain PET/MRI | 31 |
| Análisis cuantitativo del corazón en imágenes 4DCT en pacientes con taquicardia ventricular recurrente o refractaria | 35 |
| Deep Learning-Based Generation of Synthetic CT from MR Images for Craniosynostosis Planning | 39 |
| DeepCeres: red neuronal artificial para la segmentación de lóbulos del cerebelo a partir de imágenes multimodales de RM de alta resolución | 43 |
| DeepThalamus: Red neuronal para la segmentación de núcleos talámicos a partir de imágenes de RM de alta resolución. | 47 |

Sesión Biomecánica / Ingeniería de Rehabilitación

Presentación Oral

9:30 - 11:00

Aula Sebastián Ferignan

Moderan: Ana De Los Reyes Guzmán y Antonio José Del Ama Espinosa

Análisis del riesgo de rotura de aorta aneurismática con comportamiento anisotrópico	52
Clasificación de movimientos del tren superior con visión artificial y aprendizaje profundo en adulto sano y paciente pediátrico con parálisis cerebral	56
Desarrollo de nuevos efectores finales para robots asistivos en la rehabilitación de miembro superior	60
Diseño de un sistema integral de atención domiciliaria con robot móvil asistencial	64
Digitalización del test de Tinetti mediante técnicas de Visión Artificial.	68
Diseño conceptual de un sistema robótico ambulatorio para la asistencia de la marcha de personas con baja movilidad	

Sesión Telemedicina y e-Salud

Presentación Oral

9:30 - 11:00

Aula PB3

Moderan: Enrique Javier Gómez Aguilera y Javier Reina Tosina

Diseño e implementación de una aplicación basada en procesos para la gestión de consentimientos de pacientes	73
EMMA: una nueva plataforma para la monitorización y acompañamiento de personas con endometriosis	77
Arquitectura de Microservicios para Ensayos Clínicos en Pérdida de Peso con Plataformas mHealth	81
Herramienta de salud digital IMIDoc para la monitorización de pacientes con enfermedades inflamatorias inmunomediadas	85
Interoperabilidad de Bluetooth Low Energy con ISO/IEEE X73-PHD: estudio de propuestas basadas en encapsulación de datos	89
Accompain: monitorización ambulatoria del dolor oncológico	

Sesión Bioinstrumentación

Presentación Oral

15:00 - 16:30

Salón de Grados

Moderan: *José Millet Roig, Antoni Ivorra Cano y Juan Suardíaz Muro*

Piezoelectric paint-based force sensor for hand rehabilitation	94
Optimización de un sistema de seguimiento y monitorización de órganos trasplantados durante la fase de transporte	98
Development of a stand-alone device to simultaneously record cardiocography biosignals in an open format for testing methods of assessing fetal heart rate variability	102
Prototipo de un sistema integral de nariz-lengua-ojo electrónicos para la detección de enfermedades en muestras de orina	106
Evaluación de tres prendas sensorizadas con textiles conductores para el registro de ECG durante la vida cotidiana	
Design and development of a mock pulmonary circulation system to test intravascular sensors	

Sesión Modelado y Simulación

Presentación Oral

15:00 - 16:30

Paraninfo

Moderan: *Félix Fanjul Vélez y Manuel Antonio Burgos Olmos*

Rol de la dirección de fibras y heterogeneidades iónicas en la simulación de arritmias auriculares	111
Metodología para la obtención de modelos 3D de cavidades nasales con parámetros funcionales especificados	115
Modelado computacional de la ablación de tejidos biológicos por electroporación irreversible utilizando campo eléctrico pulsado de alta intensidad	119
Computer modeling of radiofrequency thermocoagulation (RF-TC) using the recording intracerebral electrodes implanted for stereo electroencephalography (SEEG) monitoring	123
Validación de un algoritmo para el cálculo automático de la distancia inter lesión en la ablación por catéter de radiofrecuencia de Fibrilación Auricular	127
Perturbación eléctrica y térmica provocada por sensores de temperatura en estudios preclínicos de ablación cardiaca por radiofrecuencia	131

Sesión Nanomedicina

Presentación Oral

15:00 - 16:30

Aula Sebastián Ferignan

Moderan: Oscar Castaño Linares y Javier Cerrillo Moreno

- Laser re-irradiation of palladium nanoparticles for antibacterial applications. 136
- Improving the efficacy of magnetic nanoparticle-mediated hyperthermia using trapezoidal pulsed electromagnetic fields as an in vitro anticancer treatment in uveal melanoma and glioblastoma multiforme cell lines 140
- Effectiveness of gold nanorods of different sizes in photothermal therapy to eliminate melanoma and glioblastoma cells. 144
- Effect of hyperthermia-based nanotherapies in three- dimensional models of neuroblastoma tumors
- Biosensor for the neuronal activity monitoring and neurofilament light detection in a brain-on-a-chip
- Quantum Control of Colour Centres in Diamond for nano-NMR

Sesión Realidad Virtual

Presentación Oral

15:00 - 16:30

Aula PB5

Moderan: Ana De Los Reyes Guzmán, Lola Ojados González y Dolores Parras Burgos

- HC3DU: Un nuevo plugin para ofrecer retroalimentación háptica en simulaciones de Realidad Virtual 149
- Diseño e implementación de un planificador quirúrgico para cirugía colorrectal junto con un visualizador basado en realidad mixta 153
- Valoración de la Experiencia del Usuario en Entornos Virtuales de Rehabilitación de Miembros Superiores 157
- M3Display: Sistema de realidad aumentada para la rehabilitación de la función motora del miembro superior 161
- OTOVIRT: Cirugía virtual para el entrenamiento por simulación y el ensayo preoperatorio en cirugía otológica 165
- Automatización para la evaluación ergonómica integrando las tecnologías de realidad virtual, captura de movimiento, electromiografía de superficie y uso de exoesqueleto 169

Sesión Premios SEIB-FENIN

Presentación Oral

16:30 - 19:00

Paraninfo

Tribunal: Elisabete Aramendi Ecenarro, Javier Sáiz y Javier Reina Tosina

Segmentación de patrones de marcha en pacientes con parálisis cerebral mediante aprendizaje profundo y visión artificial	174
Preliminary screening of polycaprolactone-polystyrene electrospun scaffolds for the formation of tumoroids	178
Modelos basados en redes neuronales artificiales para la generación y predicción de series temporales de glucosa	182
Computer-aided detection system for pulmonary embolism with integrated cardiac assessment based on embolic burden	186
Study of fibroblasts activation kinetics and identification of fibroblasts subpopulations in physiological and pathological conditions	190
Development of an electric stimulation system to analyze muscle damage in 3D Duchenne muscular dystrophy muscle culture	194
Chamber design of a portable breathalyzer for disease diagnosis	198
Análisis de electroencefalografía de amplitud integrada en neonatos: Evaluación del rendimiento de aCUP-E frente a los electrodos de gel líquido	202
Analyses of H ₂ O ₂ production in time during wound healing	206
Nuevos modelos multicelulares para la predicción de la hepatotoxicidad idiosincrática	210
Estudio numérico para consolidación mediante fijación de fracturas verticales en cuello de fémur	214
Aplicación de Realidad Aumentada enfocada a la Salud Bucodental en edades tempranas	218
Optimización de Servicios de Urgencias mediante Modelos de Inteligencia Artificial: Revisión Sistemática de la Predicción de Ingresos Hospitalarios desde Urgencias	222

Sesión Póster I

Póster

11:00 - 12:00

Claustro

Modera: Joaquín Roca González

Clasificación Supervisada de Pacientes con Síndrome de Brugada con Biomarcadores Derivados del ECG	227
Autómata celular para simulación de fibrilación auricular: acercando el tiempo de cómputo a la práctica clínica	231
Importancia del Pre-entrenamiento en el Desarrollo de una Red Neuronal Convolutiva Ligera para la Evaluación de la Calidad del ECG	235
Algoritmo de Sincronización del Electrocardiograma Para Análisis Retrospectivos Basados en Vectocardiograma	239
Clasificación EEG del estrés mental inducido por un Serious Game mediante Deep Learning	243
Diagramas de Recurrencia General para la Detección Automática de Apnea Obstruktiva del Sueño	247
Análisis integral de la vorticidad en la región anterior de las vías aéreas nasales	251
Análisis computacional del flujo espiratorio nasal: Investigación de las diferencias geográficas y de sexo en la resistencia y la asimetría	254
Aplicación de técnicas de Inteligencia Artificial a datos de Diabetes de un repositorio OMOP Common Data Model	258
Control por bloques Adaptativo por Modelo de referencia In Silico para pacientes con Diabetes Tipo 1	262
Colimación Automática para Sistemas de Radiología mediante Aprendizaje Profundo	266
Hacia una protonterapia adaptativa: Evaluando el potencial de la nnUNet para la segmentación de órganos en riesgo	270
Enhancing MRI Reconstruction Efficiency Through Multi-GPU Parallelization	274
Modelo de predicción de la respuesta al tratamiento de quimio-radioterapia en pacientes con cáncer de pulmón de células no pequeñas localmente avanzado irresecable mediante la aplicación de radiómica en imágenes de TC	278
Estimación de la proporción de hidrops endolinfático en pacientes con la enfermedad de Ménière	282
Estudio comparativo de algoritmos de segmentación de imágenes médicas para la generación de modelos para la planificación quirúrgica	286

Desarrollo de nuevos biomarcadores de imagen para la detección de cáncer de mama	290
Synchronization in Retrospective Respiratory Gating using Deep Learning	294
Propuesta de modelo basado en redes neuronales para la clasificación de regiones del esqueleto a partir de imágenes TC	298
Herramienta para la evaluación de los márgenes de resección en cirugía conservadora de cáncer de mama	302
Beyond DTI: Insights in Heart Infarcted Tissue	306
Uso de cuestionarios digitales en el aula y realización de seminarios en la unidad de cuidados intensivos como innovaciones docentes en Ingeniería Biomédica	310
Predictores de adherencia a una app de salud en el proyecto europeo PHARAON	314
Aplicación para la gestión personalizada de pacientes con diabetes	318
Diferenciación por métodos estadísticos convencionales y machine learning entre tejido pulmonar sano y patológico de medidas de impedancia eléctrica	322
Proyecto para la validación de medida de ángulo de lordosis en embarazadas mediante imagen no invasiva versus test de flechas	
A Model of Mechanical Dyssynchrony based on ECG features	
Deep Learning CT segmentation for dosimetry in postoperative endometrial carcinoma treatment	
A calibrated classifier for Alzheimer's disease and Frontotemporal dementia using MRI	
Crystal resolvability evaluation of ASIC for hexagonal SiPM	
A low-cost device for conducting respiratory oscillometry, based on open-source software and Internet of Things technology, designed for home use in developing countries	

Jueves 23 de noviembre de 2023

Sesión Señales Biomédicas (II)

Presentación Oral

8:30 - 10:00

Salón de Grados

Moderan: *José Joaquín Rieta Ibáñez, Jesús Poza Crespo y José Manuel Allegue Gallego*

New Approach for the Assessment of Sarcopenia Using Core Shape Modelling	327
Comparación de métodos manuales para el estudio del jitter neurofisiológico.	331
Análisis del proceso de llenado de la señal sEMG a medida que aumenta gradualmente la fuerza en el cuádriceps	335
Coefficiente de coactivación muscular como indicador del control motor: caso de estudio para flexo-extensión de codo en adultos	339
Monitorización de la depresión mediante el análisis de la circadianidad del ritmo cardíaco proporcionado por un dispositivo wearable	343
Aplicación de inteligencia artificial para la predicción del estrés en cirugía robótica mínimamente invasiva	347

Sesión Imágenes Médicas (II)

Presentación Oral

8:30 - 10:00

Paraninfo

Moderan: *José Luis Arce Diego y Begoña Acha Piñero*

Predicción automática del diagnóstico del sarcoma de Ewing en una serie extensa de tumores de células redondas incluidas en micromatrices de tejidos	352
Evaluation of machine learning algorithms and relevant biomarkers for the diagnosis of multiple sclerosis based on optical coherence tomography.	356
Enhancing Ulcerative Colitis Histological Image Segmentation with an Active Learning Framework	360
Discriminación Histológica mediante Espectroscopia Óptica de Onda Parcial de Muestras Tisulares de Cerebro	364
Estudio comparativo entre los modelos geométricos de córneas 3D paciente-específico generados a partir de los topógrafos Sirius y MS-39	368
Aprendizaje auto-supervisado para mejorar el diagnóstico del síndrome de Sjögren con ultrasonografía	372

Sesión Neurotecnologías (I)

Presentación Oral

8:30 - 10:00

Aula Sebastián Ferignan

Moderan: Mario Ortiz García y Roberto Hornero Sánchez

Metodología de análisis de la actividad cerebral en pacientes con ictus en fase de rehabilitación	377
Análisis de señales EEG en movimientos de flexión plantar y dorsal mediante el empleo de un exoesqueleto de bajo coste para la caracterización de la acción motora.	381
Aplicación de modelo de Deep Learning IFNet en decodificación de imaginación motora del pedaleo para señales EEG	385
Evaluación de una interfaz cerebro-ordenador basada en potenciales relacionados con eventos para la detección de estímulos visuales en una tarea de vigilancia	389
Propuesta de BCI basado en RSVP con número reducido de estímulos usando el modo de escritura T9	393
Evaluación del Impacto del Aprendizaje Auto-Supervisado en la Precisión de Interfaces Cerebro-Ordenador basadas en Imaginación Motora	397

Sesión Impresión 3D

Presentación Oral

8:30 - 10:00

Aula PB3

Moderan: Lola Ojados González y Dolores Parras Burgos

Metodología de planificación, diseño y fabricación de modelos anatómicos y guías quirúrgicas a medida mediante impresión 3D en caso de hemipelvectomía por osteosarcoma	402
3D printed implant structure for breast reconstruction and systematic drug delivery for preventing the reoccurrence of cancer	406
Simulador de bronoscopias mediante fabricación aditiva para la exploración médica personalizada	410
Obtención de modelos 3D de nuevos prototipos de hisopos nasofaríngeos mediante la tecnología de impresión 3D en materiales biocompatibles	414
Adaptación de pieza para tubos de respirador - tubuladora segura para la ventilación no invasiva	418
Asiento o lecho de posición infantil personalizado mediante técnicas de escaneado 3D y fabricación aditiva	422

Sesión Premios José María Ferrero Corral

Presentación Oral

12:00 - 13:30

Paraninfo

Tribunal: *Enrique Javier Gómez Aguilera, Javier Reina Tosina y Roberto Hornero Sánchez*

Very low elastic modulus Ti alloys obtained by Laser Directed Energy Deposition to avoid bone resorption in bone implants	427
Comparación in-silico de la eficacia de los fármacos flecainida y vernakalant en fibrilación auricular paroxística	431
Análisis wavelet de la señal de oximetría cerebral para el cálculo de la frecuencia de las compresiones del masaje de resucitación cardiopulmonar	435
Development of a neuroblastoma-on-a-chip bioengineered platform enabling the translational study of emerging nanotherapies	439
Augmented reality simulator for laparoscopic surgery on a realistic phantom	443

Sesión Póster II

Póster

10:00 - 11:00

Claustro

Moderador: *Juan Suardíaz Muro*

Análisis del Volumen de Ablación en Diagnóstico Médico mediante Espectroscopía de Plasma Inducido	448
Calibración de simulaciones de fibrilación auricular para reproducir biomarcadores clínicos	452
Cinemática del movimiento del miembro superior durante una tarea funcional realizada en entorno virtual y entorno físico real en población sana y con lesión medular cervical	456
Data augmentation study for rare diseases assessment with Deep Learning: Confocal Imaging analysis of Congenital Muscular Dystrophy	460
Evaluating Techniques for Neuron Identification in Complex Cultures: A Deep Learning Approach	464
Un enfoque explicativo para el diagnóstico de la degeneración macular asociada a la edad mediante técnicas de deep learning	468
Aplicación de MobileNet para el diagnóstico temprano del glaucoma: Un enfoque binocular	472
Ensamblaje de redes neuronales para el establecimiento de una prioridad en la atención del paciente con lesión en la piel	476

Comparación de múltiples redes neuronales convolucionales para el diagnóstico automático de la degeneración macular asociada a la edad usando retinografías	480
Métodos de ayuda al diagnóstico de glaucoma basado en diferencias de características extraídas de ambos ojos en tomografías de coherencia óptica	484
Monitorización sin contacto en tiempo real: estudio piloto mediante vehículo aéreo no tripulado	488
Análisis del Color, Textura y Temperatura para el Diagnóstico y Clasificación de Hemangiomas Infantiles	492
Concordancia entre especialistas. ¿Cómo podemos saber si tenemos una GT representativa para entrenar una red neuronal? Un estudio centrado en el Carcinoma Basocelular (BCC)	496
Influencia de los sistemas Brain-Computer Interface basados en Neurofeedback en las características de la red cerebral	500
Validación de un protocolo de adquisición de señales electroespino-gráficas en posición decúbito prono	504
Conceptual design of a functional electrical stimulation cycling platform as a rehabilitation therapy for spinal cord injury or stroke	508
An inexpensive arbitrary waveform neurostimulator for the selective activation of neurons in retinal prosthesis	512
Análisis del estrés durante procedimientos quirúrgicos mediante laparoscopia convencional y robótica	516
Diseño e implementación de un modelo basado en inteligencia artificial para estimar el consumo máximo de oxígeno en adultos de mediana edad	520
Clasificación de patologías mediante análisis acústico de la voz	524
Deformación de la presión de las vías respiratorias durante la ventilación mecánica invasiva detectada mediante técnicas de aprendizaje automático supervisado	528
Caracterización de las interacciones complejas en pacientes que pasan con éxito la prueba SBT	532
Aplicación de técnicas de Inteligencia Artificial Explicable para la identificación de factores relacionados con la calidad de sueño en adultos sanos	536
Calibración automática para el pegado de camas en TAC	540
Cardiac Ablation with Electroporation: Numerical Simulations of Contact Force Effects in Lesion Size	

Simulación de arritmias reentrantes en isquemia aguda: importancia del sistema His-Purkinje

Caracterización de canales de conducción en ventrículos infartados: Un enfoque para la predicción de arritmias

High-Resolution Sleep Position Monitoring in Adolescents with Smartphone Accelerometers

Noninvasive assessment of neuromechanical/neuroventilatory coupling and mechanical efficiency in COPD

Thoracic Bioimpedance for Breathing Pattern Estimation in COPD Patients

Preprocesamiento de vídeo para extracción de frecuencia cardíaca sin contacto utilizando fotopleletismografía remota

Rapid detection of viral particles using imaging techniques

MRI Compatibility of Injectable Microstimulators: Preliminary Validation in a 3T Scanner

Metodología para la evaluación objetiva del estado cognitivo en neurorrehabilitación con exoesqueleto de miembro inferior.

Proyecto "Aprender enseñando: Interacciones educativas entre asignaturas del Grado en Ingeniería Biomédica"

Sesión Premios SEIB-EIT HEALTH

Presentación Oral

15:30 - 17:00

Parainfo

Moderadora: Isabel Alfany

[Nanorobotics and Magnetic Hyperthermia in Cancer \(NRB & MHT\)](#)

[e-Health platform for the monitoring of elderly \(eHealthPlt\)](#)

[Design of a breathalyzer for disease diagnosis \(DBDD\)](#)

[Differentiation of lung tissue through EIS \(Lung EIS\)](#)

[Cognireal](#)

[Non-invasive Technique Study of Lordosis \(LORDOSIS\)](#)

[CleftCare3D Grafting & Bone Resorption Study \(Cleft3D\)](#)

Viernes 24 de noviembre de 2023

Sesión Señales Biomédicas (III)

Presentación Oral

8:30 - 10:00

Salón de Grados

Moderan: *Raimon Jané Campos, Pablo Laguna Lasaosa y María Galindo Martínez*

Caracterización del Acoplo Cardiopulmonar en Pacientes Pediátricos con Apnea Obstructiva del Sueño 545

Detección automática de la apnea del sueño infantil utilizando técnicas de deep learning y explainable artificial intelligence en señales de flujo aéreo 549

Interpretabilidad de Redes Convolucionales para el Análisis de Episodios de Tos en Enfermedades Respiratorias 553

Uso de la actividad electrodermal (EDA) para la clasificación de situaciones de estrés en cirugía robótica de mínima invasión 557

Análisis de la dinámica respiratoria en sujetos sanos para diferentes patrones

Efecto de la función difusa en la estimación de la entropía muestral difusa con valores de tolerancia fijos (fFuzzyEn)

Sesión Biomateriales

Presentación Oral

8:30 - 10:00

Paraninfo

Moderan: *Gustavo Guinea Tortuero y Joaquín Arias Pardilla*

Axonal guidance using biofunctionalized silk fibroin fibers manufactured using the SFS technique 562

Microfluidic Model of the Alternative Vasculature in Neuroblastoma 566

A microphysiological setup to mimic, test and analyze myocardial tissue for cardiovascular diseases 570

Lab-on-a-Chip Analysis Using Benchtop NMR Technology 574

Versatility of Activated Vapor Silanization (AVS) functionalization for biomedical applications 576

Hydrogels with tunable characteristics to simulate the neuronal environment 580

Sesión Neurotecnologías (II)

Presentación Oral

8:30 - 10:00

Aula Sebastián Ferignan

Moderan: Antonio José Del Ama Espinosa y Claudia Gómez Cárdenas

Evaluación del paradigma de adquisición de la actividad electroencefalográfica en estado de reposo	585
Modelo basado en redes neuronales convolucionales y altamente conexas para la ayuda al diagnóstico de pacientes en coma tras paro cardíaco	589
Análisis de las alteraciones en la estructura neurofisiológica en la migraña	593
Análisis múltiple de la red funcional cerebral en sujetos con hipertensiones de la sustancia blanca	597
Implementing Emotion Detection from Speech for Psychological Assessment of Elderly People: A Comparative Study of Python-based Approaches and Existing Solutions	601
Advanced Neuromuscular Stimulation with Injectable Wireless Microstimulators based on Volume Conduction	

Sesión Póster III

Póster

10:00 - 11:00

Claustro

Modera: Francisco José Ortiz Zaragoza

Estimadores del retardo entre las series de QT y RR en registros ECG de prueba de esfuerzo: evaluación en simulación	606
Aplicación multiplataforma para gestión farmacológica del servicio de Neonatología (SAFINeo)	610
Modelado, diseño y fabricación de un sistema inyector de aerosoles basado en modelos anatómicos	614
A case study: Improving face detection in Cartagena Vessel Traffic Service Operators to detect fatigue onset using thermographic images through the YOLO5Face model	618
GB3 as a Novel Therapeutic Target in Neuroblastoma's Alternative Vasculature through Stiffness Modeling	622
Potential role of Calcineurin isoform A- β 1 in neurovascular niche stem cell dynamics as a therapeutic target	626
A microphysiological model designed for the research of angiogenic sprouting in an extracellular matrix	630

Producción de membranas de Bruch artificiales de fibroína de seda mediante Dynamic Dope Destabilization Spinning para la creación un Epitelio Pigmentario de la Retina funcional	634
Analysis of Peroxiporin-Mediated H ₂ O ₂ Fluxes in HeLa Cellular Membrane: A Therapeutic Potential in Redox signaling disorders	638
Predicción de complicaciones postoperatorias a partir de la actividad neuronal de pacientes pediátricos sometidos a anestesia mediante modelos de Machine Learning	642
The importance of integral time length windows for the classification of activities of daily living based on Machine Learning techniques	646
Impact of integrated signals for doing HAR using Deep Learning models	650
Detector automático de artefactos en señales neuronales basado en técnicas de Inteligencia Artificial	654
Measuring cognitive load in a simulation game	658
Protocolo para el entrenamiento cognitivo de la memoria semántica mediante tDCS	662
Self-Supervised Audio Spectrogram Transformer para el Reconocimiento de Emociones en Señales de Voz: Un Enfoque de Entrenamiento Semi-Supervisado	666
Mejorando el Analisis de Emociones en texto mediante Llama 2	670
Aplicación de Realidad Virtual para Simulación de Recorridos de Evacuación de espacios sanitarios	674
Realidad virtual aplicada a la ingeniería del factor humano. Caso de estudio "Manipulación de paciente en enfermería"	678
Holograma Piramidal 360 aplicado a la ingeniería del factor humano. Caso de Estudio Ergonómico "manipulación de paciente en Enfermería"	682
Aplicación de la Realidad Virtual para la capacitación en materia de comunicación en crisis sanitarias	686
Probabilistic model-based tool for assessing exercise functional capacity in patients with COPD	
Fusing Clinical and Metabolomic Data for the Prediction of Ventilatory Therapies in ICU COVID Patients	
Proyecto para la validación de medida de ángulo de lordosis en embarazadas mediante imagen no invasiva versus test de flechas	
Electroencephalogram Based Alzheimer Disease Classification using Machine and Deep Learning Methods	

Exactitud y precisión absoluta mediante sistema CAD del ángulo preoperatorio HKA en radiografías coronales de miembros inferiores

Sistema de calidad para la fabricación "in-house" de hisopos nasofaríngeos impresos en 3D

Bioimpresión 3D: Tecnologías y Aplicaciones para aplicaciones Point-Of-Care Manufacturing

Cryosectioning of hydrogels as a reliable approach to increase yield and further tune mechanical properties

Perfusion Evaluation of Biocompatible Scaffolds for Lab-on-a-Chip Systems Using Magnetic Resonance Imaging

Señales Biomédicas (I)

Miércoles 22 9:30 - 11:00

Salón de Grados

Los cambios en la morfología del pico al fin de la onda T, cuantificados mediante “time-warping”, están asociados con la aparición de fibrilación ventricular en un modelo porcino de isquemia aguda

N. Gómez^{1,2}, J. Ramírez^{1,3,4}, A. Martín-Yebra^{1,3}, MM. Demidova⁵, P. Platonov⁵, JP. Martínez^{1,3}, P. Laguna^{1,3}

¹ Grupo BSICoS, I3A, IIS Aragón, Universidad de Zaragoza, Zaragoza, España, ngomez@unizar.es

² Universidad de Oriente, Santiago de Cuba, Cuba

³ Centro de Investigación Biomédica en Red – BBN (CIBER-BBN), Zaragoza, España

⁴ William Harvey Research Institute, Queen Mary University of London, London, United Kingdom

⁵ Department of Cardiology, Clinical Sciences, Lund University, Lund, Sweden

Resumen

En este trabajo se propone un índice, d_w , para cuantificar la variación morfológica de la zona del pico al fin de la onda T (T_{pe}) del ECG, basado en la medida de la deformación temporal (en inglés *time-warping*) y se evalúa su asociación con la aparición de episodios de fibrilación ventricular (FV) en condiciones isquémicas. Se analizaron los registros de ECG de 26 cerdos sometidos a una oclusión coronaria de 40 minutos. La serie d_w a lo largo del tiempo se obtuvo cuantificando las diferencias morfológicas en T_{pe} entre distintas fases de la oclusión y un T_{pe} de referencia tomado en el estado basal. En los registros basales, d_w permaneció estacionario con mediana intra-registro que oscila entre 0.11 y 4.98 ms [mediana: 1.76; IQR: 1.80], mientras que durante la oclusión arterial siguió una tendencia gradual creciente bien marcada a medida que progresaba la isquemia con valores entre 1.58 y 47.80 ms [mediana: 15.47; IQR: 18.53]. En el periodo de 20 a 25 min desde el inicio de la oclusión (y durante los 5 min previos al episodio de FV), las medias de d_w en el grupo de cerdos que tuvieron FV fueron significativamente superiores a las del grupo sin FV, con valores de mediana de 40.0 (y 34.4) ms frente a 7.8 (y 7.7) ms, con p -valores de 0.002 (y 0.001), respectivamente. En conclusión, este estudio experimental en un modelo porcino muestra que los aumentos dinámicos del índice d_w durante la progresión de la isquemia están asociados con el riesgo de aparición de FV.

1. Introducción

Las alteraciones de la repolarización ventricular se han asociado con la presencia de un sustrato arritmogénico [1], lo que ha llevado al estudio de varios índices derivados del ECG para cuantificar el aumento de la dispersión de la repolarización ventricular y predecir arritmias ventriculares. En concreto, la dispersión del área de la onda T [2], la dinámica de la repolarización periódica (PRD) [3], la aparición de un patrón de onda J [4], la variabilidad del intervalo QT [5] la distancia desde el pico hasta el final de la onda T (intervalo T_{pe}) [6], las variaciones de la morfología de la onda T respecto a una referencia normal (TMV) [7], y la restitución de la morfología de la onda T (TMR) [8] son algunos ejemplos de estos índices.

En condiciones isquémicas, la delineación de la onda T, y en particular el inicio de la onda, está muy expuesta

a errores de anotación como consecuencia de la elevación/depresión del ST. La dispersión de la repolarización ventricular reflejada en el intervalo T_{pe} ha demostrado potencial en la predicción del riesgo arritmico [9], y no requiere la determinación del inicio de la onda T, por lo que resulta ser un biomarcador apropiado en condiciones de isquemia. Además, los índices basados en intervalos de tiempo que se sustentan únicamente en medidas de anotación de ondas, no capturan todos los posibles cambios morfológicos contenidos en el ECG. Para superar estas restricciones se evaluó la capacidad del índice basado en el análisis de *time-warping* propuesto por Ramírez *et al.* [10], adaptado al intervalo entre el pico y el final de la onda T para restringir su aplicación, durante un modelo humano de isquemia inducida por corto tiempo mediante intervención coronaria percutánea (PCI, de las siglas en inglés) [11]. En ese estudio se informó la capacidad de dicho índice para captar la modificaciones de la formas de onda desde el pico hasta el final de la onda T inducidas por la isquemia. El presente estudio tiene como objetivo evaluar si este índice d_w , restringido de pico al fin de la onda T, se asocia con el riesgo de fibrilación ventricular en un modelo porcino de infarto.

2. Materiales y métodos

2.1. Materiales: La población de estudio incluyó 26 cerdos a los que se indujo un infarto del miocardio a tórax cerrado mediante PCI, ver más detalles en [12]. En cada animal se infló un globo en la arteria descendente izquierda (Left Descending Artery, LAD) para inducir la isquemia. Se utilizaron registros de ECG de 12 derivaciones, digitalizados a una frecuencia de muestreo de 1024 Hz y una resolución de amplitud de 1.18 μ V por bit, para monitorizar a los cerdos antes del inflado del globo (fase basal) y durante todo el periodo de oclusión de 40 minutos (fase de oclusión). Durante el estudio, 16 cerdos no sufrieron FV (grupo sin FV) y 10 cerdos sufrieron FV (grupo con FV) después del minuto 15 desde el inicio de la oclusión, siendo desfibrilados posteriormente (aquellos cerdos con episodios de FV antes del minuto 15 no fueron incluidos en este estudio). El

tiempo de aparición de los eventos de FV tras el inicio de la oclusión osciló entre 17:40 y 30:50 min, con una media de 21:20 min. El estudio se ajustó a la Guía para el cuidado y uso de animales de laboratorio de los *National Institutes of Health* de EE.UU. (NIH, Publicación nº 85-23, revisada en 1996) y recibió la aprobación del comité local de ética en investigación animal.

2.2. Cuantificación de los cambios en la morfología del pico al fin de la onda T: El pre-procesado del ECG incluyó los siguientes pasos: filtrado paso-bajo (frecuencia de corte de 40 Hz) para la eliminación del ruido eléctrico y muscular, seguido de un paso-alto (frecuencia de corte de 0.5 Hz) para la atenuación de la desviación de la línea base, en ambos casos filtros de Butterworth de sexto orden aplicados de forma bidireccional. La delineación de cada derivación del ECG se realizó utilizando el método descrito en [13]. Posteriormente, se aplicó una estrategia basada en reglas de selección multiderivacional [14] sobre los conjuntos de marcas de las 8 derivaciones estándar para obtener una única marca multiderivación. A continuación, se realizó una transformación espacial basada en análisis de componentes principales sobre las 8 derivaciones estándar, aprendido sobre la onda T antes delineada con el fin de enfatizar el contenido de la onda T. Por último, se delineó de nuevo la primera componente principal (PC1) de la derivación transformada, se segmentaron las ondas T y se filtraron nuevamente paso-bajo a una frecuencia de corte de 20 Hz con un filtro bidireccional Butterworth de sexto orden para eliminar los restantes componentes de alta frecuencia fuera de banda.

Los cambios en la morfología entre el pico y el final de la onda T a lo largo del tiempo se cuantificaron mediante el índice d_w propuesto por Ramírez en [10], y adaptado en [11] a la restricción de la segunda mitad de la onda T. Para cada s -ésima ventana móvil de señal de 15 segundos a lo largo del registro (con solapamiento de 10 s entre ventanas), se extrajeron las ondas T y se calculó una onda media (Mean Warped T-peak to T-end Wave, MWTPE). Inicialmente, todas las ondas T dentro de una ventana dada se transformaron en ondas de polaridad positiva. La clase dominante entre las ondas T bifásicas o monofásicas se definió para cada ventana como la clase que tenía el mayor número de ocurrencias. Para calcular la MWTPE sólo se tuvieron en cuenta las ondas T que pertenecían a la misma clase que la dominante. La MWTPE de referencia se calculó en una ventana de tiempo que comprende los primeros 60 segundos de la fase basal, de modo que d_w representa los cambios morfológicos del pico al fin de la onda T en relación con el estado basal. Para cada ventana s , el marcador $d_w(s)$ se estima como la reparametrización temporal entre dos ondas, una MWTPE, $\mathbf{f}^s(\mathbf{t}^s) = [f^s(t^s(1)), \dots, f^s(t^s(N_s))]^T$, donde $\mathbf{t}^s = [t^s(1), \dots, t^s(N_s)]^T$, y una MWTPE de referencia seleccionada, $\mathbf{f}^r(\mathbf{t}^r)$. El índice d_w es la cantidad media de *warping* necesaria para minimizar las diferencias en el dominio del tiempo entre la onda objeto de estudio $\mathbf{f}^s(\mathbf{t}^s)$, y la de referencia $\mathbf{f}^r(\mathbf{t}^r)$.

Estimación de series d_w a través de funciones *warping*: Sea $\gamma(\mathbf{t}^r)$ la función de *warping* que relaciona \mathbf{t}^r y \mathbf{t}^s tal que $[\mathbf{f}^s \circ \gamma](\mathbf{t}^r) = \mathbf{f}^s(\gamma(\mathbf{t}^r))$ es la señal reparametrizada, o el *time-warping* de $\mathbf{f}^s(\mathbf{t}^s)$ utilizando $\gamma(\mathbf{t}^r)$. Al igual que en [10], se utilizó la función de pendiente de raíz cuadrada (SRSF, por sus siglas en inglés) en lugar de las señales originales, para encontrar la función de *warping* óptima y evitar así el llamado “efecto pellizco”. La función *warping* óptima, $\gamma^*(\mathbf{t}^r)$, es la que minimiza la diferencia de amplitud entre la SRSF de $\mathbf{f}^r(\mathbf{t}^r)$ y $\mathbf{f}^s(\gamma(\mathbf{t}^r))$. El nivel de *warping* representa la cantidad de estiramiento necesario para ajustar de manera óptima la onda en estudio en relación con la de referencia. El biomarcador d_w cuantifica este nivel de *warping* requerido como el promedio del valor de la diferencia absoluta entre $\gamma(\mathbf{t}^r)$ y \mathbf{t}^r :

$$d_w = \frac{1}{N_r} \sum_{n=1}^{N_r} |\gamma^*(\mathbf{t}^r(n)) - \mathbf{t}^r(n)|. \quad (1)$$

Los cambios en el curso del tiempo inducidos por la isquemia capturados por $d_w(s)$ se calcularon para cada $\mathbf{f}^s(\mathbf{t}^s)$ MWTPE estimada en cada ventana s de 15 segundos en la PC1, lo que da como resultado una serie $d_w^{PCA}(s)$ ($s \in \{1, \dots, S\}$). Se compararon los resultados obtenidos entre el grupo con FV y el grupo sin FV mediante la prueba de Kruskal-Wallis. Se asumió significancia estadística cuando el p -valor ≤ 0.05 .

3. Resultados y discusión

No se observaron cambios significativos en la forma del pico al fin de la onda T (cuantificados por d_w^{PCA}) durante los registros basales en la población de estudio, con mediana intra-registro que oscilaba entre 0.11 y 4.98 ms [mediana: 1.76; IQR: 1.80], representativo de la pequeña variabilidad natural latido a latido del ECG. Por el contrario, se observaron fuertes cambios inducidos por la isquemia, reflejados en los cambios de forma del pico al final de la onda T (aumentos/disminuciones de anchura y amplitud) y capturados por un aumento de la magnitud de d_w^{PCA} durante la fase de inflado del globo en la mayoría de los cerdos, con mediana intra-registro que oscilaba entre 1.58 y 47.80 ms [mediana: 15.47; IQR: 18.53].

La Figura 1 muestra un ejemplo de la evolución temporal de $d_w^{PCA}(s)$ durante las fases basal y de oclusión para dos cerdos concretos, uno del grupo que sufrió FV y otro del grupo que no sufrió FV. Los cambios dinámicos en la morfología del pico al fin de la onda T, inducidos por la oclusión arterial, quedaron bien reflejados por la evolución del índice $d_w^{PCA}(s)$, con una tendencia creciente a medida que progresaba la isquemia, la cual comienza inmediatamente después de los minutos iniciales de la oclusión. La tendencia creciente es más acentuada en el cerdo con FV que en el cerdo sin FV, con valores que oscilan entre 4.92 y 44.51 y entre 1.94 y 7.52 ms, respectivamente. Por el contrario, en ambos cerdos, la magnitud $d_w^{PCA}(s)$ permanece casi estacionaria durante la fase basal, oscilando entre 0 y 8.67 y entre 0 y 3.15 ms, respectivamente.

Durante la fase de oclusión, el índice $d_w^{PCA}(s)$ presentó un comportamiento similar en todos los cerdos al comparar las tendencias temporales, caracterizado por un aumento desde el inicio de la exposición a la isquemia, con los mayores cambios en los primeros 5 minutos y permaneciendo relativamente estable a partir de entonces. El curso temporal medio (línea azul) y la desviación estándar (línea verde) del índice $d_w^{PCA}(s)$ en todos los cerdos para el grupo sin FV y el grupo con FV se muestran en la Figura 2, tanto durante las fases basal como de oclusión, alineadas con el inicio del registro o con el inicio del inflado del globo, respectivamente. Obsérvese cómo el curso temporal medio de $d_w^{PCA}(s)$ es capaz de captar las variaciones en la forma de onda del pico a fin

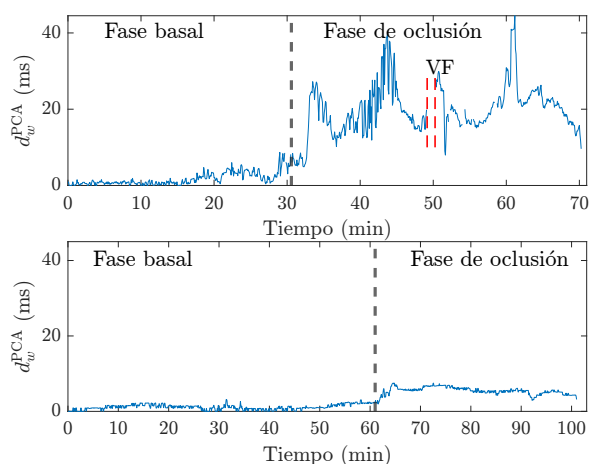


Figura 1. Evolución temporal de $d_w^{PCA}(s)$ a lo largo del tiempo durante las fases basal y de oclusión de dos cerdos concretos. El panel superior muestra la evolución para un cerdo del grupo con FV y el panel inferior para un cerdo del grupo sin FV. La línea negra discontinua indica el inicio de la oclusión arterial y la línea roja discontinua la aparición del episodio de FV.

En la Figura 3 se presenta la media del índice $d_w^{PCA}(s)$ medido en diferentes intervalos de 5 minutos (primeros y últimos 5 minutos antes del inicio de la oclusión, en los intervalos de 1-5, 5-10, 10-15, 15-20, 20-25 minutos tras el inicio de la oclusión y durante los 5 minutos previos al episodio de FV) agrupados para los dos grupos de cerdos. Estos valores medios de $d_w^{PCA}(s)$ fueron, excepto para los primeros 5 minutos basales, significativamente mayores en el grupo con FV que en el grupo sin FV, con un valor de mediana de 0.82, 0.83, 6.25, 7.56, 8.02, 7.76, 7.81 y 7.65 frente a 0.44, 2.35, 12.45, 18.78, 26.71, 23.96, 30.95 y 34.37, y un p -valor de 0.223, 0.011, 0.035, 0.001, 0.001, 0.007, 0.002 y 0.001, para cada intervalo, respectivamente. El índice d_w^{PCA} mostró cambios estadísticamente significativos ya desde los 5 minutos anteriores al inicio de la oclusión, posible consecuencia de las modificaciones ya inducidas por la inserción del catéter. A medida que avanzaba el tiempo de oclusión, se encontraron diferencias interindividuales significativas en la magnitud de los cambios de d_w^{PCA} para el grupo con FV desde los primeros 5 minutos del inflado del globo, que iban desde una variación insignificante al

inicio de la oclusión hasta una magnitud pronunciada a medida que transcurría el tiempo. El aumento muy significativo de d_w^{PCA} para el grupo con FV en relación con el grupo sin FV a medida que progresa la isquemia indica que los aumentos de la magnitud de d_w^{PCA} por encima de algún umbral se pueden asociar con el riesgo de aparición de episodios de FV.

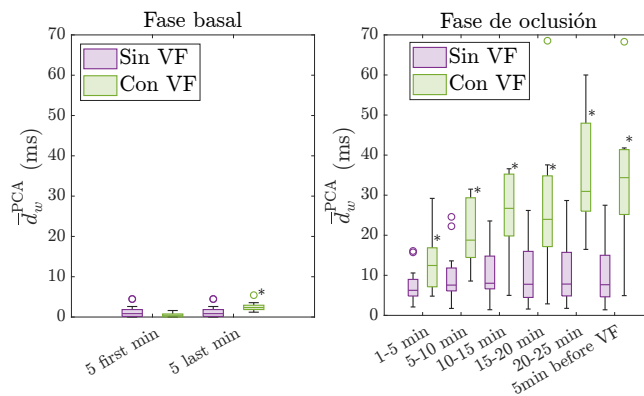


Figura 3. Comparación de las medias de d_w^{PCA} , para el subgrupo con FV (verde) y el subgrupo sin FV (morado), medidas en diferentes segmentos de 5 minutos. * indica significación estadística para la comparación entre grupos.

4. Conclusiones

En un modelo de infarto de miocardio porcino, el marcador de forma basado en *time-warping*, d_w^{PCA} , restringido al intervalo entre el pico y el final de la onda T, permite monitorizar los cambios de repolarización inducidos por la isquemia. Un mayor incremento de d_w^{PCA} durante la progresión de la isquemia se asocia con un mayor riesgo de aparición de fibrilación ventricular y sugiere estudios de evaluación en humanos.

Agradecimientos

Este trabajo ha sido financiado por los proyectos PID2021-128972OA-I00, LMP14121, TED2021-130459B-I00 y PID2022-140556OB-I00, respaldados por el Ministerio de Ciencia e Innovación (MCIN) y FEDER, y por el Gobierno de Aragón al Grupo *Biomedical Signal Interpretation and Computational Simulation* (BSICoS) T39-20R cofinanciado por FEDER 2014-2020 “*Building Europe from Aragón*”. JR agradece la financiación de la beca RYC2021-031413-I del MCIN.

Referencias

- [1] Veronique MF Meijborg, Chantal E Conrath, T Opthof, Charly NW Belterman, Jacques MT de Bakker, and R. Coronel. Electrocardiographic T wave and its relation with ventricular repolarization along major anatomical axes. *Circ Arrhythm Electrophysiol*, 7:524–531, 2014.
- [2] Tuomas V Kenttä, Moritz F Sinner, Bruce D Nearing, Rebecca Freudling, Kimmo Porthan, Jani T Tikkanen, Martina Müller-Nurasyid, Katharina Schramm, Matti

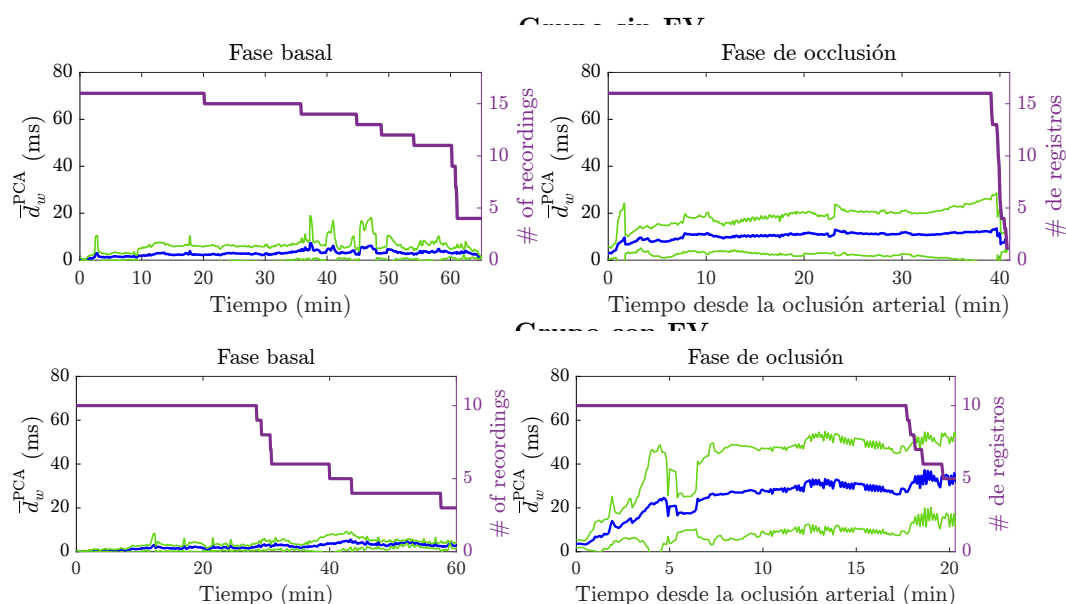


Figura 2. Promedio del curso temporal de $d_w^{PCA}(s)$ (línea azul) \pm desviación estándar (línea verde) alineado con el inicio del registro en la fase basal y con el inicio de la oclusión durante la fase de oclusión, para el grupo sin FV (panel superior) y para el grupo con FV (panel inferior). La línea morada y el eje-y derecho representan el número de registros promediados en cada momento dado.

- Viitasalo, Antti Jula, et al. Repolarization heterogeneity measured with T-wave area dispersion in standard 12-lead ECG predicts sudden cardiac death in general population. *Circulation: Arrhythmia and Electrophysiology*, 11(2):e005762, 2018.
- [3] Saúl Palacios, Iwona Cygankiewicz, Antoni Bayés de Luna, Esther Pueyo, and Juan Pablo Martínez. Periodic repolarization dynamics as predictor of risk for sudden cardiac death in chronic heart failure patients. *Scientific Reports*, 11(1):20546, 2021.
- [4] Marina M Demidova, Alba Martín-Yebra, Jesper van der Pals, Sasha Koul, David Erlinge, Pablo Laguna, Juan Pablo Martínez, and Pyotr G Platonov. Transient and rapid QRS-widening associated with a J-wave pattern predicts impending ventricular fibrillation in experimental myocardial infarction. *Heart Rhythm*, 11(7):1195–1201, 2014.
- [5] Mathias Baumert, Alberto Porta, Marc A Vos, Marek Malik, Jean-Philippe Couderc, Pablo Laguna, Gianfranco Piccirillo, Godfrey L Smith, Larisa G Tereshchenko, and Paul GA Volders. QT interval variability in body surface ECG: measurement, physiological basis, and clinical value: position statement and consensus guidance *Europace*, 18(6):925–944, 2016.
- [6] Jan E Azarov, Marina M Demidova, Sasha Koul, Jesper Van Der Pals, David Erlinge, and Pyotr G Platonov. Progressive increase of the T peak-T end interval is associated with ischaemia-induced ventricular fibrillation in a porcine myocardial infarction model. *EP Europace*, 20(5):880–886, 2018.
- [7] Julia Ramírez, Antti Kiviniemi, Stefan van Duijvenboden, Andrew Tinker, Pier D Lambiase, Juhani Junttila, Juha S Perkiömäki, Heikki V Huikuri, Michele Orini, and Patricia B Munroe. ECG T-wave morphologic variations predict ventricular arrhythmic risk in low-and moderate-risk populations. *Journal of the American Heart Association*, 11(17):e025897, 2022.
- [8] Julia Ramírez, Michele Orini, Ana Mincholé, Violeta Monasterio, Iwona Cygankiewicz, Antonio Bayes de Luna, Juan Pablo Martínez, Esther Pueyo, and Pablo Laguna. T-wave morphology restitution predicts sudden cardiac death in patients with chronic heart failure. *Journal of the American Heart Association*, 6(5):e005310, 2017.
- [9] Charles Antzelevitch, José M Di Diego, and Mariana Argenziano. Tpeak-Tend as a predictor of ventricular arrhythmogenesis. *International Journal of Cardiology*, 249:75–76, 2017.
- [10] Julia Ramírez, Michele Orini, J Derek Tucker, Esther Pueyo, and Pablo Laguna. Variability of ventricular repolarization dispersion quantified by time-warping the morphology of the T-waves. *IEEE Transactions on Biomedical Engineering*, 64(7):1619–1630, 2016.
- [11] Neurys Gómez, Julia Ramírez, Juan Pablo Martínez, and Pablo Laguna. Time-warping based end-of-T-wave shape marker reflects repolarization changes during ischemia. In *2022 Computing in Cardiology (CinC)*, volume 498, pages 1–4. IEEE, 2022.
- [12] Marina M Demidova, Jesper van der Pals, Joey FA Ubachs, Mikael Kanski, Henrik Engblom, David Erlinge, Victor M Tichonenko, and Pyotr G Platonov. S-segment dynamics during reperfusion period and the size of myocardial injury in experimental myocardial infarction. *Journal of electrocardiology*, 44(1):74–81, 2011.
- [13] Juan Pablo Martínez, Rute Almeida, Salvador Olmos, Ana Paula Rocha, and Pablo Laguna. A wavelet-based ECG delineator: evaluation on standard databases. *IEEE Transactions on biomedical engineering*, 51(4):570–581, 2004.
- [14] Pablo Laguna, Raimon Jané, and Pere Caminal. Automatic detection of wave boundaries in multilead ECG signals: Validation with the CSE database. *Computers and biomedical research*, 27(1):45–60, 1994.

Dispersión espacial de los tiempos de activación y repolarización asociada a diferentes modos de estimulación cardíaca

S. Palacios¹, R. Smisek², K. Curila³, U. Nguyen⁴, F.W. Prinzen⁴, J. Halamek², F. Plesinger², P. Jurak², J.P. Martínez^{1,5}, E. Pueyo^{1,5}

¹ BSICoS Group, I3A, IIS Aragón, Universidad de Zaragoza, Zaragoza, Spain, {spalacios,jpmart,epueyo}@unizar.es

² Institute of Scientific Instruments, The Czech Academy of Sciences, Brno, Czechia

³ Charles University and University Hospital Kralovske Vinohrad, Cardiocenter, 3rd Faculty of Medicine, Prague, Czechia

⁴ Maastricht University Medical Centre (MUMC), Maastricht, Netherlands (The)

⁵ CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Spain

Resumen

En pacientes con indicación de marcapasos permanente se aplican distintos tipos de estimulación ventricular. Los denominados fisiológicos estimulan el sistema de conducción cardíaca induciendo una activación fisiológica eficiente. Entre estos se encuentran la estimulación selectiva del haz de His (HBP selectiva, sHBP, y HBP no selectiva, nsHBP, por sus siglas en inglés) y las estimulaciones selectiva y no selectiva de la rama izquierda (sLBBP y nsLBBP). Otras regiones cardíacas que también suelen estimularse mediante el marcapasos son el septo del ventrículo izquierdo (LVSP) o del ventrículo derecho (RVSP) y el ápex del ventrículo derecho (RVAP). En este trabajo se analizaron 695 electrocardiogramas de muy alta frecuencia (UHF-ECG) obtenidos de 176 pacientes con complejo QRS estrecho y con indicación de marcapasos. Se caracterizaron los tiempos de activación (TA) y de repolarización (TR) y se agruparon en tres regiones según las derivaciones en las que se evaluaron (R1: derivaciones V1-V2; R2: V3-V4; R3: V5-V6). Globalmente en la población, las estimulaciones sHBP, nsLBBP y LVSP proporcionaron los valores de AT y RT más similares a los obtenidos durante ritmo espontáneo. Los valores absolutos de las medias para las diferencias R1-R2 y R3-R2 en TA resultaron menores a 3, 16 y 10 ms para sHBP, nsLBBP y LVSP, respectivamente, con respecto al ritmo espontáneo. Para TR estas diferencias fueron menores a 11, 34 y 24 ms para sHBP y nsLBBP y LVSP. En conclusión, las estimulaciones HBP, LBBP y LVSP inducen los tiempos de activación y repolarización ventricular más similares a los hallados en ritmo espontáneo en pacientes con conducción fisiológica (QRS estrecho).

1. Introducción

En pacientes con bradicardia sintomática con indicación de implantación de marcapasos pueden aplicárseles distintos tipos de estimulación cardíaca. La estimulación del sistema de conducción cardíaca, estimulación Hisiana (*His bundle pacing*, HBP) o estimulación de la rama izquierda (*left bundle branch pacing*, LBBP), ha surgido como una alternativa a la estimulación convencional en el ventrículo derecho (*right ventricular pacing*, RVP) [1], la cual se había descrito que podía causar disincronía ventricular, aumentando el riesgo de insuficiencia cardíaca y disminuyendo la fracción de eyección del ventrículo izquierdo.

Aunque HBP es el modo de estimulación que mejor replica la conducción fisiológica, presenta algunas limitaciones relacionadas con el posicionamiento de los electrodos, que lleva a un 10-20% de fallo en la implantación, y un umbral alto de captura de la zona de estimulación, con su consecuente consumo elevado de energía. LBBP, que se define como la captura del haz de la rama izquierda (HRI), parece una alternativa más efectiva para superar las limitaciones de la HBP [2]. Dentro de este tipo de estimulación, se distingue LBBP selectiva (cuando únicamente se llega a capturar el HRI) o no selectiva (cuando se capturan tanto el HRI como el tejido miocárdico del septo adyacente). Se ha reportado una mayor tasa de implantes exitosos con LBBP así como bajos umbrales de captura [3]. La estimulación en el haz de His junto con células del miocardio se denomina no selectiva (nsHBP), mientras que, si se estimula únicamente el haz de His, se denomina estimulación selectiva (sHBP).

Hasta la fecha no se han caracterizado de manera exhaustiva los efectos del LBBP selectiva (sLBBP) y no selectiva (nsLBBP) en comparación con sHBP, nsHBP y otros modos de estimulación que actúan en el septo interventricular, como la estimulación septal del ventrículo izquierdo (*left ventricular septal pacing*, LVSP), del ventrículo derecho (*right ventricular septal pacing*, RVSP) o la estimulación apical del ventrículo derecho (*right ventricular apex pacing*, RVAP).

Se han propuesto distintos marcadores no invasivos derivados del electrocardiograma (ECG) para cuantificar las características de la activación ventricular en respuesta a la estimulación cardíaca. Entre otros, cabe señalar el índice de disincronía eléctrica (eDYS), la duración de la depolarización local (Vdx), la media de estas duraciones entre todas las derivaciones (Vd) [4] y la dispersión de la depolarización ventricular obtenida de mapas epicárdicos [5].

El objetivo de este estudio es la evaluación de los cambios en la sincronía ventricular tanto en relación a los tiempos de activación (TA) como a los de repolarización (TR) que se producen tras la aplicación de los distintos tipos de estimulación cardíaca. Para ello se han analizados registros ECG de diferentes pacientes con QRS estrecho y con

indicación de implante de marcapasos, adquiridos antes y después del implante.

2. Población de estudio

Este estudio incluye pacientes a quienes se les indicó la implantación de un marcapasos. Se adquirieron 695 electrocardiogramas de muy alta frecuencia (5 kHz), correspondientes a 176 pacientes, en el International Clinical Research Center del St Anne's University Hospital, Brno, República Checa, el Cardiocenter of Faculty Hospital Kralovske Vinohrady y el Third Medical Faculty of Charles University, Praga, República Checa.

La distribución de los registros según el tipo de estimulación aplicada durante los mismos fue la siguiente: 50 registros durante sHBP, 160 durante nsHBP, 87 durante LVSP, 47 durante nsLBBP, 13 durante sLBBP, 102 registros de RVSP, 37 durante RVAP y 199 registrados grabados durante ritmo espontáneo.

Los registros ECG de 14 derivaciones se grabaron usando un equipo Ventricular Dyssynchrony Imaging (VDI, ISI Brno, Cardion, FNUSA, CZ, 2018). Todos los sujetos fueron sometidos a una estimulación de prueba antes de aplicar los distintos modos concretos de estimulación, mientras eran monitorizados mediante un ECG de 12 derivaciones. De esta manera se identificaron los patrones correspondientes de activación ventricular y se comprobó si llegaba a capturar parte del miocardio o de las ramas de conducción [6].

3. Métodos

3.1. Preprocesado

Las señales ECG obtenidas directamente del equipo VDI fueron preprocesadas para eliminar la línea de base y para filtrar, aplicando un filtro notch, el ruido de red de 50 Hz.

Los artefactos generados por los impulsos procedentes del tipo de estimulación fueron eliminados por un algoritmo semiautomático similar al descrito en [7]. El algoritmo aplicado consiste en los siguientes pasos:

1. Se obtuvieron las derivaciones ortogonales, XYZ, a partir de las 12 derivaciones estándar de la señal ECG mediante la matriz de transformación de Kors [8]. El vector magnitud se calculó como:

$$v(n) = \sqrt{x(n)^2 + y(n)^2 + z(n)^2}$$

donde $x(n)$, $y(n)$ y $z(n)$ representan cada una de las componentes ortogonales.

2. La pendiente del vector magnitud, dv/dt , es aproximadamente proporcional a la diferencia entre muestras consecutivas. El inicio y el final del impulso de la estimulación se identificaron mediante las siguientes condiciones:

$$\frac{dv}{dt} > \alpha_{onset}; \quad \frac{dv}{dt} < \alpha_{end}$$

donde α_{onset} y α_{end} son los umbrales elegidos para la detección del inicio y final del impulso asociados al estímulo.

3. El artefacto procedente del impulso se eliminó, reemplazando la duración del mismo por la resultante de una interpolación lineal de la señal entre el inicio y final identificados del paso anterior como parte del artefacto.

Para este estudio se seleccionaron 0.5 y -0.5 mV/ms como umbrales de inicio, α_{onset} , y final, α_{end} , del impulso.

Los complejos QRS para cada registro fueron detectados y clasificados según su morfología, diferenciando los patrones estimulados e irregulares [9]. La delineación del resto de ondas del ECG se realizó por medio de un sistema automático basado en la transformada de wavelet [10].

3.2. Latido promedio

A partir de la morfología predominante de entre todos los latidos de cada registro se definió el latido mediano como representativo de dicho registro. El algoritmo aplicado fue el siguiente:

1. Los intervalos RR, RR_i , se calcularon usando las marcas fiduciales de los complejos QRS para cada latido i .
2. Para cada registro, se seleccionó el valor de la moda observada en el histograma de intervalos RR, utilizando intervalos de 20 ms de anchura. Aquellos latidos asociados a la moda del RR se usaron para obtener el primer candidato a latido mediano.
3. Se calculó el coeficiente de correlación entre cada latido y el candidato a latido mediano. Si este coeficiente era menor a 0.85, el latido se descartó.
4. Se alinearon los latidos disponibles maximizando la correlación cruzada respecto al latido mediano.
5. Se calculó un nuevo latido mediano como mediana de los latidos.

Cualquier variable o medida relacionada con la señal original se calculó sobre este latido mediano obtenido.

3.3. Tiempos de activación y repolarización

La dispersión de TA entre las derivaciones correspondientes a diferentes regiones ventriculares se calculó como un marcador para medir la disincronía eléctrica ventricular en cada estrategia de estimulación. Los TAs para cada una de las derivaciones individuales se obtuvieron como el tiempo donde dv/dt alcanza el valor más negativo, es decir, la pendiente más negativa de la relación entre el voltaje y el tiempo, en el complejo QRS [11]. Del mismo modo, TR es un marcador que mide el tiempo de recuperación ventricular y se determinó como el tiempo asociado al máximo del valor de dv/dt , es decir, a la máxima pendiente de la onda T (figura 1) [11]. Ambos tiempos se miden desde el inicio del complejo QRS de cada latido correspondiente.

Se definieron tres regiones que comprenden el área de los ventrículos. Estas regiones se identificaron a partir de las derivaciones precordiales: derivaciones septales (V1 y V2, primera región, R1), derivaciones anteroapicales (V3 y V4, segunda región, R2) y derivaciones anterolaterales (V5 y V6, tercera región, R3). En este trabajo, la media de los

tiempos calculados para las dos derivaciones se tomó como valor para cada una de las regiones asociadas. Además, la región R2 se tomó como referencia. Así, podemos diferenciar el comportamiento de la parte derecha, asociada a R1, y el correspondiente a la parte izquierda, R3. Finalmente, se evaluaron las diferencias R1-R2 y R3-R2.

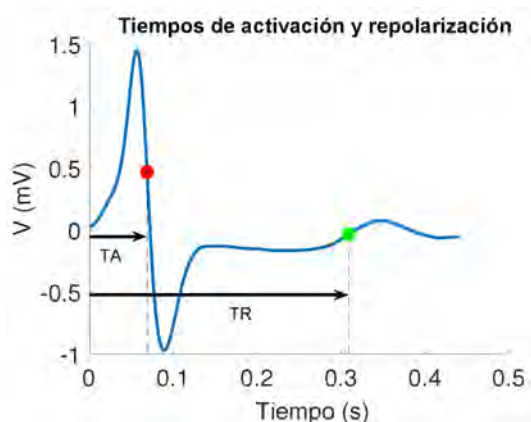


Figura 1. Un latido con las anotaciones utilizadas para calcular el tiempo de activación, TA (en rojo) y el tiempo de repolarización, TR (en verde).

3.4. Análisis estadístico

Mediante el análisis de los tiempos de activación, TA, y de repolarización, TR, se encontraron diferencias entre las regiones ventriculares según el tipo de estimulación cardíaca aplicada. Las diferencias entre las regiones 1 y 3 respecto a la región 2 se presentan en forma de media de todos los registros correspondientes a cada tipo de estimulación. Para realizar la comparación entre diferentes técnicas de estimulación y el ritmo basal, la referencia fue la señal no estimulada (ritmo espontáneo). Se utilizó la prueba U de Mann-Whitney (también llamada prueba de suma de rangos de Wilcoxon) para comparar los tiempos de activación o de repolarización entre cada una de las técnicas de estimulación. Diferencias estadísticas con p-valor menor a 0.05 se consideraron significativas. Para el análisis de los datos se utilizó MATLAB R2020a (9.8).

4. Resultados

4.1. Tiempos de activación y repolarización para cada tipo de estimulación

La figura 2 muestra el tiempo medio de activación y repolarización para cada una de las técnicas de estimulación cardíaca analizadas y el ritmo espontáneo en las regiones septal, R1, y anterolateral, R3, respecto a la región R2, anteroapical.

No se encontraron diferencias significativas entre R1-R2 ($p > 0.8$) para AT y RT entre los registros sHBP y los registros espontáneos (-8.22 y -8.30 ms para TA y 8.58 y 18.31 ms para TR, respectivamente).

Comparando los dos tipos de estimulación LBBP, sLBBP y nsLBBP, se observó que ambos modos tenían un comportamiento similar en TR, especialmente considerando las diferencias R3-R2 (-25.01 y -20.66, $p > 0.5$).

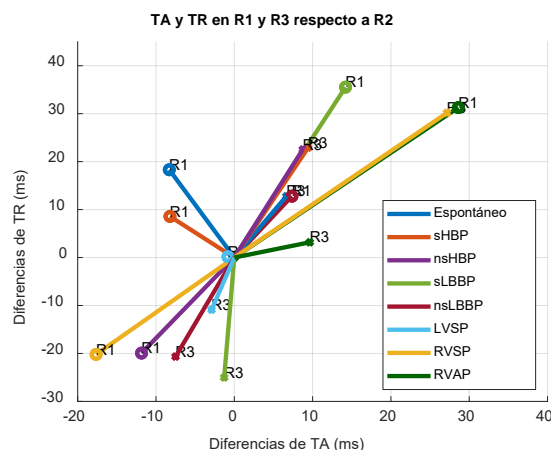


Figura 2. TA y TR medios en las regiones R1 y R3 respecto a R2 para los registros espontáneos y los procedentes nsHBP, nsHBP, sLBBP, nsLBBP, LVSP, RVSP y RVAP.

Cuando se analizó el patrón de depolarización y repolarización de los registros asociados a RVAP se encontró que los TA y TR medios eran mayores en R1 (28.62 y 31.27 ms) que en R3, mientras que para los registros RVSP se observó lo contrario (-17.65 ms para TA y -20.21 ms para TR), siendo estas diferencias estadísticamente significativas ($p < 0.01$). En el caso de LVSP, tanto en la región R1 como R2 AT y RT presentaron valores similares, según puede verse en la Figura 2, con -0.82 ms para TA medio y 0.15 ms para TR medio.

4.2. Comparación entre estimulación cardíaca y ritmo espontáneo

Modos de estimulación	R1-R2		R3-R2	
	ΔTA (ms)	ΔTR (ms)	ΔTA (ms)	ΔTR (ms)
sHBP	0.09	-9.73	2.84*	10.12*
nsHBP	-3.55*	-38.24*	2.10	9.76
sLBBP	22.50*	17.21*	-7.93*	-37.79*
nsLBBP	15.66*	-5.52	-14.14*	-33.44*
LVSP	7.48*	-18.16	-9.54*	-23.67*
RVSP	-9.35*	-38.52*	20.52*	17.38*
RVAP	36.92*	12.96*	3.00	-9.60*

Tabla 1. Diferencias entre los tiempos de activación y repolarización medios para R1-R2 y R3-R2 respecto al ritmo basal. *p-valor < 0.05

A continuación, los valores de TA y TR medios de cada tipo de estimulación se compararon con los valores medios correspondientes al ritmo espontáneo. Estas diferencias se denotan como ΔTA y ΔTR . Valores próximos a cero corresponden a una respuesta similar al ritmo espontáneo.

Los valores de ΔTA y ΔTR , evaluados en R1 y R3 (respecto a R2), se muestran en la Tabla 1. Las mayores diferencias en dispersión de TA se encontraron para los registros de RVAP (36.92 ms entre R1-R2) y de dispersión de TR en registros RVSP (-38.52 ms, entre R3-R2).

Este análisis también confirmó lo observado en la Figura 2, esto es, que las menores diferencias entre el ritmo

espontáneo y las técnicas de estimulación cardíacas aplicadas se encontraron para los registros de sHBP y nsLBBP tanto en el caso de TA como el de TR en la primera región (R1-R2). Para la tercera región (R3-R2) las estimulaciones fisiológicas, sHBP y nsHBP, y RVAP presentaron un comportamiento similar al de los registros de ritmo espontáneo en términos de dispersión de TA y TR.

5. Discusión y conclusiones

En pacientes con indicación de implante de marcapasos permanente se están aplicando diferentes estrategias de estimulación cardíaca. La disincronía eléctrica está asociada con fallo cardíaco y con un incremento de la mortalidad [12] y, por lo tanto, es importante determinar cuál de las técnicas presenta un comportamiento eléctrico más parecido a los patrones fisiológicos en pacientes con conducción normal (esto es, con QRS estrecho).

En este trabajo se mide la disincronía eléctrica calculando los valores medios de TA y TR, así como la diferencia de estos en las regiones R1 y R3 respecto a R2. HBP, LBBP y LVSP inducen valores de TA y TR más cercanos a los obtenidos en ritmo espontáneo. En promedio, los sujetos a los que se les aplicó sHBP tenían menores diferencias ΔTA y ΔTR . LBBP o LVSP tienen asociadas ventajas en términos de mayor éxito de implantación y menores umbrales de captura y, sin embargo, presentan valores solo ligeramente superiores en los marcados evaluados que los encontrados en HBP. Otras técnicas, tal como RVSP, muestran mayores diferencias respecto al ritmo basal.

Los resultados presentados, en términos de dispersión de TA, están en línea con trabajos previos. En [4,6], la disincronía ventricular se caracterizó usando marcadores de registros UHF-ECG y se encontró que HBP conseguía preservar la sincronía ventricular de manera similar al ritmo espontáneo, en concordancia a los hallazgos de este trabajo. En [13], la estimulación del ventrículo derecho se comparó con técnicas fisiológicas, incluyendo nsHBP. En línea con nuestros resultados, los autores encontraron que la estimulación del ventrículo derecho produce una mayor disincronía eléctrica que las estrategias basadas en la estimulación del haz de His. La dispersión de TR está, sin embargo, poco caracterizada en la literatura, lo que no permite establecer una comparación de nuestros resultados con los de estudios previos.

Aunque LBBP y LVSP inducen un ligero incremento en la dispersión de TA y TR, en comparación a HBP, esta no es estadísticamente significativa, por lo que podrían ser las técnicas a aplicar, ya que podrían conseguir superar las limitaciones de HBP, como son la dificultad en el acceso de la zona de estimulación o el aumento de la energía necesaria en la estimulación.

Agradecimientos

Este trabajo ha sido realizado con el apoyo de los proyectos PID2019-105674RB-I00, PID2019-104881RB-I00, TED2021-130459B-I00 y la ayuda BES-2017-080587 (Ministerio de Ciencia e Innovación), el proyecto LMP94_21 y el grupo de referencia BSICoS T39-23R (Gobierno de Aragón cofinanciado por el FEDER 2014-

2020 “Construyendo Europa desde Aragón”) y el proyecto ERC G.A. 638284 (European Research Council). Los cálculos computacionales se han realizado en la ICTS NANBIOSIS (HPC Unit at University of Zaragoza).

Referencias

- [1] Burri, H et al. EHRA clinical consensus statement on conduction system pacing implantation: endorsed by the Asia Pacific Heart Rhythm Society (APHRs), Canadian Heart Rhythm Society (CHRS), and Latin American Heart Rhythm Society (LAHRS). *Europace*. 2023; 25(4), 1208–1236.
- [2] Tun HN et al. Conduction system pacing: promoting the physiology to prevent heart failure. *Heart Failure Reviews*. 2023; 28, 379–386.
- [3] Zhou Y et al. Left ventricular septal pacing versus left bundle branch pacing in the treatment of atrioventricular block. *Annals of Noninvasive Electrocardiology*, 2022;27, e12944.
- [4] Jurak P et al. Novel ultra-high-frequency electrocardiogram tool for the description of the ventricular depolarization pattern before and during cardiac resynchronization. *J Cardiovasc Electrophysiol*. 2020;31:300–307.
- [5] Arnold AD et al. Left ventricular activation time and pattern are preserved with both selective and nonselective His bundle pacing. *Heart Rhythm O2*. 2021;2:439–445.
- [6] Curila K et al. Both selective and nonselective His bundle, but not myocardial, pacing preserve ventricular electrical synchrony assessed by ultra-high-frequency ECG. *Heart Rhythm*. 2020;17:607–14.
- [7] Haq KT et al. Detection and removal of pacing artifacts prior to automated analysis of 12-lead ECG. *Computers in Biology and Medicine*. 2021; 133:104396.
- [8] Kors JA et al. Reconstruction of the Frank vectorcardiogram from standard electrocardiographic leads: diagnostic comparison of different methods. *Eur. Heart J*. 1990, 11, 1083–1092.
- [9] Plesinger F, Jurco J, Jurak P. Robust multichannel QRS detection. *Comput Cardiol*. 2014;2014:557-560.
- [10] Martinez JP et al. A wavelet-based ECG delineator: evaluation on standard databases. *IEEE Trans. Biomed. Eng*. 2004; 51, 570–581.
- [11] Wyatt RF et al. Estimation of ventricular transmembrane action potential durations and repolarization times from unipolar electrograms. (Abstract) *Am. J. Cardiol*. 1981; 97:488
- [12] Miyajima K et al. Assessing cardiac mechanical dyssynchrony in left bundle branch area pacing and right ventricular septal pacing using myocardial perfusion scintigraphy in the acute phase of pacemaker implantation. *J. Cardiovasc. Electrophysiol*. 2022;33(8), 1826–1836
- [13] Curila K et al. Ventricular activation pattern assessment during right ventricular pacing: Ultra-high-frequency ECG study. *J Cardiovasc Electrophysiol*. 2021; 32(5):1385–1394.

Detección automática de residuos de la actividad ventricular presentes en las ondas fibrilatorias extraídas de registros electrocardiográficos de fibrilación auricular

P. Escribano¹, J. Ródenas¹, M. García¹, F. Ravelli², M. Masè², J.J. Rieta³, R. Alcaraz¹

¹ Grupo de Investigación en Ingeniería Electrónica, Biomédica y de Telecomunicación, DIEEAC, Universidad de Castilla-La Mancha, Albacete, España, {pilar.escribano, juan.rodenas, manuel.garcia, raul.alcaraz}@uclm.es

² Department of Cellular, Computational and Integrative Biology - CIBIO, University of Trento, Trento, Italia

³ BioMIT.org, Departamento de Ingeniería Electrónica, Universitat Politècnica de Valencia, Valencia, España

Resumen

El desarrollo de técnicas de extracción de las ondas fibrilatorias (ondas-f) del electrocardiograma (ECG) de superficie ha posibilitado su análisis para detectar y mejorar el tratamiento de la fibrilación auricular (FA). Sin embargo, estas técnicas tienen un rendimiento limitado por la presencia de artefactos y complejos QRST de morfología variable. Por ello, se propone una metodología para optimizar la selección de tramos de ondas-f libres de residuos de la cancelación ventricular. Primeramente, la detección de valores atípicos fuera del margen comprendido entre 2,5 veces los percentiles de 5 y 95% de la señal de ondas-f extraída del ECG preoperatorio al procedimiento de ablación por catéter (AC) de 148 pacientes de FA persistente permitió la detección de los artefactos registrados. Seguidamente, se caracterizaron los intervalos QRST cancelados para estimar la actividad ventricular remanente en las ondas-f mediante diferentes métricas. \mathcal{R}_{uVR} fue la más capacitada para tal fin (71,66% de exactitud), obtenida del producto del valor cuadrático medio de las ondas-f en el intervalo QRST por su valor máximo. No obstante, su combinación con otras métricas mediante un modelo de ensamblado de árboles de decisión mejoró la exactitud en la detección de residuos ventriculares hasta alcanzar casi un 82%. Finalmente, esta metodología demostró la relevancia de asegurar la calidad de los tramos de ondas-f extraídos del ECG de superficie, mejorando la capacidad discriminatoria de la frecuencia dominante de las ondas-f para predecir el resultado del procedimiento de AC.

1. Introducción

La fibrilación auricular (FA) es la arritmia cardíaca más común en adultos, con un prevalencia entre 2 y 4% que se espera que aumente 2,3 veces en los próximos años [1]. Sus síntomas reducen la calidad de vida de los afectados y es el factor de riesgo más común de sufrir un accidente cerebrovascular isquémico, incrementando su mortalidad [2]. Además, suele evolucionar a estados sostenidos con el paso del tiempo y está ligada a un fenómeno de remodelado del sustrato auricular que favorece su perpetuación [3]. Todo ello hace que sea recomendable detectar y tratar esta arritmia lo antes posible [1, 4].

Múltiples estudios se han enfocado en detectar eventos y predecir el resultado de diferentes tratamientos de la FA mediante el registro y análisis de la actividad fibrilatoria del corazón [5, 6, 7]. Actualmente, la predicción del tratamiento de ablación por catéter (AC) de FA persistente

es de interés clínico por su todavía alta tasa de recaída a FA (>35% en el primer año), a pesar de los avances en tecnologías de ablación y mapeo [8]. Esto optimizaría la selección de pacientes sometidos a AC, evitaría riesgos innecesarios, y disminuiría la carga y el coste del sistema sanitario para tratar la FA [9], lo que se considera uno de los principales retos de salud pública [10].

Hasta el momento, los predictores clínicos analizados han proporcionado resultados limitados o contradictorios, mientras que la caracterización de las ondas fibrilatorias (ondas-f) registradas en el electrocardiograma (ECG) de superficie ha revelado un potencial predictivo prometedor, pero todavía clínicamente insuficiente [7]. Independientemente de su aplicación, la caracterización de las ondas-f requeriría de su previa extracción del ECG, puesto que la FA es un proceso continuo que queda camuflado por la mayor amplitud de la actividad ventricular también registrada [11].

Aunque se han desarrollado diversas estrategias de extracción de las ondas-f, la arritmia de FA suele presentar variaciones en la morfología de los complejos QRST que acotan el rendimiento de estas técnicas. Además, la presencia de ruido y artefactos como la actividad muscular o el movimiento de los electrodos también acotan su capacidad [11]. Todo ello favorece la presencia de residuos de la actividad ventricular y otras distorsiones indeseadas en la señal de ondas-f extraída que introducen variabilidad y restan fiabilidad a su posterior análisis.

El objetivo principal de este trabajo es desarrollar una herramienta que permita detectar tramos de ondas-f de calidad, es decir, aquellos que mejor se aproximen a la actividad fibrilatoria pura y libre de otros efectos. Además, se presenta un caso de aplicación de esta herramienta para mejorar el rendimiento predictivo de la caracterización de las ondas-f para anticipar el resultado del procedimiento de AC en pacientes de FA persistente.

2. Materiales y métodos

2.1. Población de estudio

Este estudio involucra a 148 pacientes de FA persistente sometidos a su primera AC de radiofrecuencia en los hospitales universitarios de Toledo y Albacete. Todos ellos fueron intervenidos bajo indicaciones clínicas estándar [7]. Finalmente, tras un periodo de seguimien-

to de 9 meses, 48 pacientes recayeron a FA, mientras que los 100 restantes mantuvieron el RS.

2.2. Adquisición y preprocesado del ECG

Antes de comenzar el procedimiento de AC, se registró la señal superficial del ECG de 12 derivaciones con una frecuencia de muestreo de 977 Hz y una resolución de 16 bits. La extracción de las ondas-f se realizó en la derivación V1 por ser en la que presentan mayor amplitud respecto a la actividad ventricular [12]. Previamente, esta señal se preprocesó mediante técnicas de filtrado comunes y basadas en la transformada wavelet para atenuar el desvío de la línea base, la interferencia de la red eléctrica y las componentes de ruido de alta frecuencia [13, 14]. Finalmente, la extracción de las ondas-f se llevó a cabo mediante un método de cancelación del valor singular adaptativo de los complejos QRST [15].

2.3. Optimización de la extracción de ondas-f

El registro de artefactos durante la adquisición del ECG de superficie es muy frecuente y puede que sus efectos sobre la señal de ondas-f permanezcan tras su preprocesado y extracción. Generalmente, esta alteración se manifiesta con valores atípicos de mayor amplitud de las ondas-f. Por ello, su exclusión se basó en acotar la señal entre los percentiles de 5 y 95 % multiplicados por un factor de 2,5, empíricamente seleccionado.

Aunque esta técnica también permitió eliminar los residuos ventriculares de mayor amplitud, se realizó una estimación más detallada para asegurar la selección de señales de ondas-f de alta calidad. Para ello, se computaron las métricas propuestas para evaluar y comparar el rendimiento de diferentes técnicas de extracción de ondas-f en señales reales y simuladas en el trabajo de Alcaraz, R. *et ál.* [16].

En primer lugar, se extrajeron los intervalos de ondas-f de mayor longitud libres de artefactos para cada uno de los pacientes, obteniendo un total de 9844 intervalos QRST cancelados. Todos ellos fueron chequeados visualmente junto con el equipo médico, que clasificó 441 de ellos como intervalos QRST que presentaban residuos tras la cancelación de la actividad ventricular y los restantes se consideraron bien cancelados. Seguidamente, se calcularon los estimadores de la actividad residual ventricular en cada uno de los complejos QRST cancelados, tomando su duración de 350 ms y 70 ms antes de la posición original de los picos R como su origen.

Concretamente, se calculó el estimador signal-to-noise ratio (\mathcal{R}_{SNR}) como el logaritmo del ratio de la amplitud de la señal de ECG preprocesada y de la señal de ondas-f extraída en la posición del pico R. También se obtuvo el índice unnormalized ventricular residue (\mathcal{R}_{uVR}) como el producto del valor cuadrático medio de la señal de ondas-f en el intervalo QRST por su valor máximo y se normalizó por el valor cuadrático medio de toda la señal, obteniendo \mathcal{R}_{VR} . Por otro lado, el cómputo del parámetro modified ventricular residue \mathcal{R}_{mVR} se basó en el valor de los cuartiles de la señal de ondas-f en el

intervalo QRST y de los 30 intervalos QT que preceden y siguen al intervalo QRST analizado. En el caso de que las señales contengan menos de 60 intervalos QT, se tomaron todos los intervalos QT disponibles. Por último, se evaluaron las métricas ventricular activity (\mathcal{R}_V), atrial activity (\mathcal{R}_A), y ventricular and atrial activity (\mathcal{R}_{VA}). Estos estimadores se basan en la transformada Wavelet continua de la señal de ondas-f mediante el cómputo de la entropía Wavelet de las escalas que contienen frecuencias de la actividad auricular, ventricular y ambas, respectivamente, tal y como se recoge en [16].

2.4. Análisis estadístico

Las métricas propuestas se expresaron a través de su mediana y rango intercuartílico (RI) para el conjunto de intervalos QRST mal cancelados y los que estaban libres de residuos ventriculares. Por otro lado, se empleó la prueba no paramétrica U de Mann-Whitney para evaluar la diferenciación estadística entre las distribuciones de ambos grupos de datos, tras verificar su normalidad y homocedasticidad mediante las pruebas de Lilliefors y de Levene, respectivamente. Concretamente, se consideró un valor de significación estadística de $p < 0,05$ como estadísticamente significativo.

La evaluación del rendimiento de los estimadores de la actividad residual ventricular se llevó a cabo dividiendo la base de datos en un 65 % para entrenamiento (96 pacientes) y un 35 % para test (52 pacientes). Las señales de ondas-f disponibles tenían una duración de entre 6 s y 5 min, de manera que el grupo de entrenamiento incluyó 310 y 4969 intervalos QRST con y sin residuos ventriculares, mientras que los restantes formaron parte del grupo de test. La curva característica operativa del receptor (ROC) se empleó para estimar el umbral óptimo para discernir entre los dos tipos de intervalos QRST. Se tomó como umbral óptimo aquel que proporcionaba el mejor balance entre especificidad (E_s) y sensibilidad (E_s) [17], considerando estas métricas como la proporción de intervalos QRST con y sin residuos ventriculares correctamente clasificados, respectivamente. Finalmente, también se calculó la máxima exactitud (E_x), que es el porcentaje total de intervalos QRST correctamente clasificados, y el área bajo la curva ROC (AROC), como medida agregada del rendimiento de un estimador en todos los umbrales de clasificación posibles.

Para explorar información complementaria entre estimadores individuales y mejorar el rendimiento en la clasificación de los intervalos QRST, se realizó un análisis multivariante. Se empleó el algoritmo random under-sampling boosting (RUSBoost), adecuado para trabajar con bases de datos desbalanceadas, para ensamblar 40 árboles de decisión con un máximo de 20 nodos, que fueron los parámetros optimizados para mejorar la eficiencia del modelo [18]. Se empleó un algoritmo de selección de características secuencial progresivo que minimizó el error de clasificación del modelo. Este algoritmo tomó \mathcal{R}_{SNR} , \mathcal{R}_{uVR} , \mathcal{R}_{VR} y \mathcal{R}_{mVR} como el conjunto óptimo de métricas a combinar óptimo para construir el modelo

de clasificación, cuyo rendimiento se evaluó con el mismo criterio seguido para los estimadores individuales.

2.5. Ejemplo de evaluación del método propuesto

Con el objetivo de testear el desempeño de la herramienta de optimización propuesta, se tomaron las señales incluidas en el grupo de test para abordar el problema de predicción del resultado del procedimiento de AC en pacientes de FA persistente. En particular, se evaluó la frecuencia dominante (FD) de las ondas-f, que es un predictor ampliamente utilizado en este ámbito. Se comparó su rendimiento predictivo mediante la caracterización de tramos de señal seleccionados por la herramienta de optimización propuesta en este trabajo y a partir de la señal directamente extraída del ECG preoperatorio.

La FD de las ondas-f se obtuvo como la frecuencia asociada a la máxima amplitud de su espectro de densidad espectral de potencia (DEP) en el rango frecuencial de 3 a 25 Hz [7]. La DEP se estimó a través del periodograma de Welch con una ventana de Hamming de 4000 puntos, una superposición de 3000 puntos entre ventanas adyacentes y una resolución espectral de 0,1 Hz.

Para evaluar el impacto de la herramienta de selección propuesta, la FD se calculó sobre un intervalo de 6 s de ondas-f elegido aleatoriamente de entre los intervalos libres de artefactos y residuos de la actividad ventricular seleccionados por la herramienta. Los resultados se compararon con aquellos obtenidos al promediarla FD de 5 intervalos de 6 s aleatoriamente extraídos de las señales de ondas-f originales.

3. Resultados

La tabla 1 muestra los valores de los estimadores de los residuos de la actividad ventricular calculados para los dos tipos de intervalos QRST. Se puede apreciar que todos ellos proporcionaron diferencias estadísticamente significativas, obteniendo p -valores inferiores a 0,05, excepto en el caso de \mathcal{R}_A . Por otro lado, el rendimiento de clasificación de cada estimador se muestra en la Tabla 2, donde \mathcal{R}_{uVR} presenta la mayor capacidad discriminativa entre los dos grupos de intervalos QRST. No obstante, su combinación con \mathcal{R}_{SNR} , \mathcal{R}_{VR} y \mathcal{R}_{mVR} mediante el modelo RUSBoost mejoró el rendimiento de clasificación con valores de Se, Es, Ex y AROC del 81,64%, 81,68%, 81,64% y 88,17%, respectivamente.

Finalmente, se compararon los resultados obtenidos en el ejemplo propuesto para predecir el resultado de la AC en pacientes de FA persistente a través de la FD (Figura 1). Se puede apreciar que se logra mejorar la diferenciación estadística entre los pacientes que mantuvieron el RS y los que recayeron a FA mediante el uso de la herramienta para la optimización de la selección de tramos de ondas-f. Además, los parámetros de Se, Es, Ex, y AROC mejoran más de un 10% ya que pasan de tomar los valores 55,56, 54,84, 55,10 y 62,10 a alcanzar un 66,67, 67,74, 67,35 y 67,83, respectivamente, cuando se eliminan los efectos de los artefactos y residuos de la cancelación de la actividad ventricular.

Índice	Intervalos QRST de ondas-f		
	con residuos	sin residuos	p -valor
	Mediana (RI)	Mediana (RI)	
\mathcal{R}_{SNR}	11,53 (7,08)	15,79 (6,35)	< 0,001
\mathcal{R}_{VR}	30,01 (43,44)	9,56 (12,67)	< 0,001
\mathcal{R}_{uVR} (mV ²)	0,030 (0,052)	0,007 (0,014)	< 0,001
\mathcal{R}_{mVR}	1,18 (0,40)	1,29(0,38)	< 0,001
\mathcal{R}_A	5,39 (0,12)	5,40 (0,11)	0,242
\mathcal{R}_V	4,56 (0,49)	4,65 (0,51)	< 0,001
\mathcal{R}_{VA}	5,30 (0,12)	5,31 (0,10)	< 0,001

Tabla 1. Resultados y diferenciación estadística de los estimadores de residuos de la actividad ventricular.

Índice	Se (%)	Es (%)	Ex (%)	AROC (%)
\mathcal{R}_{SNR}	66,18	66,21	66,18	68,40
\mathcal{R}_{VR}	68,14	68,03	68,13	70,50
\mathcal{R}_{uVR}	71,66	71,66	71,66	73,77
\mathcal{R}_{mVR}	58,66	58,73	58,67	57,50
\mathcal{R}_A	50,59	50,57	50,59	51,65
\mathcal{R}_V	55,10	55,10	55,10	55,98
\mathcal{R}_{VA}	54,20	54,20	54,20	56,01

Tabla 2. Resultados de clasificación de los estimadores de residuos de la actividad ventricular.

3. Discusión y conclusiones

El análisis de las ondas-f ha permitido detectar episodios y predecir el resultado de diferentes tratamientos de la arritmia de FA [5, 6, 7]. Para ello, se han desarrollado diversas técnicas de extracción de las ondas-f del electrocardiograma (ECG) de superficie. Sin embargo, estas técnicas tienen un rendimiento limitado por la presencia de artefactos y complejos QRST de morfología variable [11]. Por ello, en este trabajo se propone una metodología para optimizar la selección de tramos de ondas-f de mejor calidad que permita dar mayor fiabilidad a su análisis y caracterización posterior.

Concretamente, entre los parámetros evaluados en este trabajo, \mathcal{R}_{uVR} fue el más capacitado para caracterizar el rendimiento de técnicas de extracción de ondas-f en señales reales en el trabajo de Alcaraz, R. *et ál.* [16]. De igual manera, este índice ha demostrado la más alta correlación con la presencia de actividad residual ventricular en los intervalos QRST de las ondas-f extraídas y liberadas de artefactos. Además, su combinación con los estimadores \mathcal{R}_{SNR} , \mathcal{R}_{VR} y \mathcal{R}_{mVR} mediante el modelo RUSBoost mejoró la exactitud en la detección de residuos ventriculares hasta alcanzar casi un 82%.

Por otro lado, el ejemplo propuesto para la predicción del resultado de la AC en pacientes de FA persistente muestra una mejor correlación de la FD con el resultado del procedimiento gracias al uso de la metodología propuesta en el estudio. La técnica de promediado de la FD para 5 tramos de 6 s ondas-f permite reducir la variabilidad de las distribuciones obtenidas para ambos grupos

de pacientes (Figura 1(a)). A pesar de ello, se aprecia una mejor separación entre ambas distribuciones cuando solo se calcula la FD en un tramo de 6 s seleccionado con la técnica propuesta en este estudio (Figura 1(b)).

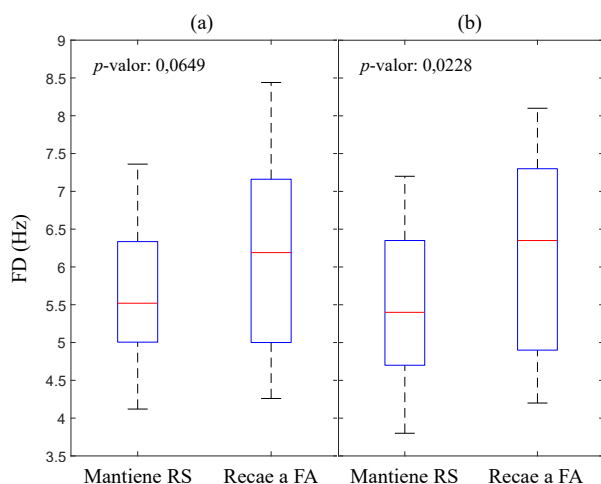


Figura 1. FD de las ondas-f promediada a partir de 5 intervalos de 6 s (a) y calculada en un intervalo de 6 s seleccionado por el método propuesto en el estudio (b).

Este estudio revela la importancia de asegurar la calidad de las ondas-f extraídas del ECG de superficie para su análisis. Se propone un método de selección de tramos de ondas-f libres de artefactos y residuos de la actividad ventricular. En particular, este método ha permitido mejorar la predicción del éxito de la AC en pacientes de FA persistente a través de la FD de las ondas-f. No obstante, estos hallazgos deberían ser corroborados en el futuro con una base de datos más amplia.

Agradecimientos

Esta investigación ha sido financiada por las subvenciones públicas PID2021-00X128525-IV0, PID2021-123804OB-I00 y TED2021-130935B-I00 del Gobierno de España, 10.13039/501100011033 junto con el Fondo Europeo de Desarrollo Regional (UE), SBPLY/17/180501/000411 y SBPLY/21/180501/000186 de la Junta de Comunidades de Castilla La Mancha, y AICO/2021/286 de la Generalitat Valenciana. Además, Pilar Escribano es beneficiaria de un contrato predoctoral 2020 PREDUCLM-15540, cofinanciado por el programa operativo del Fondo Social Europeo (FSE) 2014-2020 de Castilla-La Mancha y una ayuda de movilidad 2023-univers-11618 del Vicerrectorado de Política Científica de la UCLM.

Referencias

- [1] G. Hindricks and T. Potpara *et al.* 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). *Eur Heart J*, 42(5):373–498, 2021.
- [2] P. Warmus and N. Niedziela *et al.* Assessment of the manifestations of atrial fibrillation in patients with acute cerebral stroke - a single-center study based on 998 patients. *Neurol Res*, 42(6):471–476, 2020.
- [3] U. Schotten and D. Dobrev *et al.* Current controversies in determining the main mechanisms of atrial fibrillation. *J Intern Med*, 279(5):428–38, 2016.
- [4] S. Nattel and E. Guasch *et al.* Early management of atrial fibrillation to prevent cardiovascular complications. *Eur Heart J*, 35(22):1448–56, 2014.
- [5] A. Petrénas and L. Sörnmo *et al.* Detection of occult paroxysmal atrial fibrillation. *Med Biol Eng Comput*, 53(4):287–297, 2015.
- [6] F. Holmqvist and M. Stridh *et al.* Atrial fibrillation signal organization predicts sinus rhythm maintenance in patients undergoing cardioversion of atrial fibrillation. *Europace*, 8(8):559–65, 2006.
- [7] P. Escribano and J. Ródenas *et al.* Preoperative prediction of catheter ablation outcome in persistent atrial fibrillation patients through spectral organization analysis of the surface fibrillatory waves. *Journal of Personalized Medicine*, 12(10):1721, 2022.
- [8] B. Schmidt and J. Brugada *et al.* Ablation strategies for different types of atrial fibrillation in Europe: results of the ESC-EORP EHRA Atrial Fibrillation Ablation Long-Term registry. *Europace*, 22(4):558–566, 2020.
- [9] K. Walsh and F. Marchlinski. Catheter ablation for atrial fibrillation: current patient selection and outcomes. *Expert Review of Cardiovascular Therapy*, 16(9):679–692, 2018.
- [10] G. Lippi *et al.* Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge. *Int J Stroke*, 16(2):217–221, 2021.
- [11] L. Sörnmo and A. Petrénas *et al.* Extraction of f waves. In *Atrial Fibrillation from an Engineering Perspective*, pages 137–220. Springer International Publishing, 2018.
- [12] A. Petrénas *et al.* Lead systems and recording devices. In *Atrial Fibrillation from an Engineering Perspective*, pages 25–48. Springer International Publishing, 2018.
- [13] Leif Sörnmo and Pablo Laguna. Chapter 7 - ECG Signal Processing. In *Bioelectrical Signal Processing in Cardiac and Neurological Applications*, Biomedical Engineering, pages 453–566. Academic Press, Burlington, 2005.
- [14] M. García and M. Martínez-Iniesta *et al.* A novel wavelet-based filtering strategy to remove powerline interference from electrocardiograms with atrial fibrillation. *Physiol Meas*, 39(11):115006, 2018.
- [15] R. Alcaraz and J.J. Rieta. Adaptive singular value cancellation of ventricular activity in single-lead atrial fibrillation electrocardiograms. *Physiol Meas*, 29(12):1351–69, 2008.
- [16] R. Alcaraz, L. Sörnmo, and J. J. Rieta. Reference database and performance evaluation of methods for extraction of atrial fibrillatory waves in the ECG. *Physiol Meas*, 40(7):075011, 2019.
- [17] F. Habibzadeh *et al.* On determining the most appropriate test cut-off value: the case of tests with continuous results. *Biochem Med*, 26(3):297–307, 2016.
- [18] C. Seiffert and T.M. Khoshgoftaar *et al.* Rusboost: A hybrid approach to alleviating class imbalance. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 40(1):185–197, 2010.

ECG-ENET: Red neuronal convolucional explicable para la ayuda en el diagnóstico de la apnea del sueño infantil

C. García Vicente¹, G. C. Gutiérrez Tobal^{1,2}, J. Jiménez García^{1,2}, A. Martín Montero^{1,2}, D. Gozal³, y R. Hornero^{1,2}

¹Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, clara.garciav@uva.es

²CIBER de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Valladolid, España.

³ Office of The Dean, Joan C. Edwards School of Medicine, Marshall University, 1600 Medical Center Dr, Huntington, WV 25701, gozald@health.missouri.edu

Resumen

La apnea obstructiva del sueño (AOS) consiste en un trastorno respiratorio, que en niños se ha vinculado con el sistema cardíaco y un aumento del riesgo cardiovascular. El diagnóstico estándar es la polisomnografía (PSG), pero su coste, complejidad e incomodidad, especialmente en niños, limitan su disponibilidad y contribuyen a un infra diagnóstico de la enfermedad. Para abordar esta situación, se propone por primera vez una alternativa simplificada utilizando el electrocardiograma (ECG) nocturno y una red neuronal convolucional (CNN) que estima la severidad de la AOS pediátrica. Además, se plantea el método Gradient-weighted Class Activation Mapping (GradCAM) para interpretar los resultados de la CNN. Para ello, se han analizado 1610 registros de ECG de niños. El rendimiento de nuestro enfoque superó los mejores resultados de estudios previos (Cohen's kappa de 4 clases 0,359 vs. 0,166 y precisión de 4 clases 56,52% vs. 41,89%). Además, GradCAM identificó patrones bradicardia-taquicardia en las zonas de transición desde un evento de apnea a zonas post-apnea, destacando las regiones entre ondas T y P. Nuestros resultados muestran que la implementación de una CNN explicable mediante el ECG puede ser útil en el diagnóstico de la AOS pediátrica y ayudaría a los facultativos a mejorar la confianza en sistemas automatizados e identificar patrones cardíacos asociados con la enfermedad. Todo ello convierte nuestra propuesta en una alternativa prometedora a la PSG, con el potencial de facilitar un diagnóstico objetivo, rápido, de menor coste y preciso de la AOS.

1. Introducción

La apnea obstructiva del sueño (AOS) pediátrica es un trastorno respiratorio caracterizado por la presencia de múltiples eventos de apnea e hipopnea debido a la obstrucción de las vías aéreas superiores durante el sueño. Actualmente, su prevalencia oscila entre el 1% y el 5% [1]. La AOS está relacionada estrechamente con el sistema cardiovascular, provocando una respuesta fisiopatológica con cambios en la frecuencia cardíaca (FC) en forma de patrones característicos de bradicardia-taquicardia [2]. Además, esta condición se relaciona con la alteración de la regulación del sistema nervioso autónomo, hipertrofia ventricular e hipertensión arterial, aumentando el riesgo de padecer enfermedades cardiovasculares en niños, principalmente en casos de AOS severa [2].

El método de diagnóstico estándar es la polisomnografía (PSG) nocturna, mediante la cual se monitorizan múltiples señales biomédicas, como el electroencefalograma, el electrocardiograma (ECG), el flujo aéreo o la saturación de oxígeno en sangre [1]. Estas señales se utilizan para

calcular el índice de apnea-hipopnea (IAH), que estima el número de eventos apneicos por hora y sirve para establecer la gravedad de la enfermedad [3]. A pesar de su eficacia, la PSG es una prueba costosa, incómoda y de disponibilidad limitada, especialmente en el caso de los niños, lo que resulta en largas listas de espera y un marcado infra diagnóstico de la enfermedad [1]. Para hacer frente a estas complicaciones, diversos estudios han desarrollado estrategias simplificadas que utilizan un número reducido de señales de la PSG junto con técnicas de inteligencia artificial [4]. Sin embargo, ninguno de los enfoques ha utilizado la señal en crudo de ECG para el diagnóstico de AOS en niños. Esta señal contiene información que podría ser de utilidad en el diagnóstico de la AOS, como el acoplamiento cardiorrespiratorio y el ya mencionado patrón de bradicardia-taquicardia [2]. Estos fenómenos, junto con el aumento del riesgo de desarrollar enfermedad cardiovascular, hacen que pueda resultar especialmente interesante el análisis de la señal de ECG.

Existen estudios de AOS en adultos en los que se han propuesto métodos automáticos de *deep learning* (DL) utilizando la señal de ECG, resultando las redes neuronales convolucionales (CNNs) especialmente útiles [5]. Sin embargo, si nos centramos en la AOS pediátrica, existe una carencia demostrada en la aplicación de métodos de DL a la información cardíaca en general, y al ECG en particular [6]. En este grupo de población únicamente se han desarrollado alternativas de diagnóstico automático mediante técnicas de *feature-engineering* y utilizando información cardíaca, como la señal de fotopletimografía o la variabilidad de la frecuencia cardíaca (*heart rate variability*, HRV), pero no se ha utilizado la señal de ECG en crudo [7], [8]. Estas técnicas requieren mucho tiempo debido al preprocesado y a la extracción de características, y aunque producen resultados de interés, tienden a omitir información crucial del ECG potencialmente útil para un diagnóstico preciso. Además, otra importante limitación en toda la literatura previa relacionada es la falta de interpretabilidad de los modelos utilizados [6], obstaculizando así una mejor aceptación de los modelos automáticos en entornos clínicos reales.

Por estos motivos, en este trabajo se propone la implementación y evaluación de métodos de DL junto con técnicas de *eXplainable Artificial Intelligence* (XAI) [9]. Por una parte, los métodos de DL permiten utilizar una

gran cantidad de datos en crudo y extraer automáticamente características relevantes para un buen rendimiento del modelo [10]. Por otro lado, para aumentar la confianza en los métodos de DL es esencial entender en qué se basan sus predicciones automáticas. En este sentido, el uso de técnicas de XAI podría ayudar a extraer los patrones específicos del ECG que usan los algoritmos de DL para determinar la severidad de la AOS pediátrica. Además, también podría ayudar a establecer relaciones entre los hallazgos fisiopatológicos del ECG y la enfermedad.

Así pues, este estudio parte de una hipótesis doble. En primer lugar, que una arquitectura basada en CNN puede simplificar el diagnóstico de la AOS pediátrica utilizando toda la información relevante de la señal de ECG para estimar la gravedad de la enfermedad. En segundo lugar, que la aplicación de XAI mediante *Gradient-weighted Class Activation Mapping* (GradCAM) puede ayudar en la interpretación clínica de los resultados del modelo y a extraer nuevos conocimientos de la enfermedad asociados con el sistema cardíaco [9]. En consecuencia, el objetivo principal del estudio es evaluar un modelo interpretable basado en CNN con la señal de ECG para estimar el IAH y establecer la severidad de la AOS pediátrica.

2. Material y métodos

2.1. Sujetos y señales

En este estudio se utilizó la base de datos *Childhood Adenotonsillectomy Trial* (CHAT) disponible en el *National Sleep Research Resource* (<https://sleepdata.org/datasets/chat>) [11]. CHAT cuenta con 1610 registros de PSG nocturna realizados a pacientes pediátricos con edades entre los 5 y 9.9 años. Para este estudio, se han analizado los registros de ECG incluidos en dichas PSGs. Los sujetos se dividieron en tres subconjuntos independientes: entrenamiento (60%), validación (20%) y test (20%). La información clínica y sociodemográfica sobre los sujetos incluidos en el estudio se muestra en la Tabla 1.

Los registros de una única derivación de ECG se remuestrearon a una frecuencia de muestreo común de 100 Hz. También se realizó una corrección de la componente continua eliminando el valor medio de la señal mediante ventanas de 30 segundos. Además, se utilizó un filtro paso bajo utilizando una ventana de Hamming y frecuencia de corte de 25 Hz para reducir el ruido. Por último, dividimos las señales ECG en segmentos de 10 minutos, utilizando un solapamiento del 50%, para aumentar el número de datos, y se normalizó cada segmento individualmente restándole su valor medio y dividiéndolo por su desviación típica.

Variables	Entrenamiento	Validación	Test
Sujetos	988	323	299
Edad (años)	7,00 [2,00]	7,00 [2,00]	6,90 [2,00]
Mujeres (%)	48,28%	50,77%	53,85%
IMC (kg/m ²)	17,31 [5,92]	17,12 [6,25]	17,43 [6,04]
IAH (e/h)	2,64 [4,77]	2,45 [4,77]	2,32 [5,11]
IAH _{≥1} (e/h)	488	167	144
IAH _{≥5} (e/h)	159	44	49
IAH _{≥10} (e/h)	129	45	41

Tabla 1: Información demográfica y clínica de los sujetos de CHAT. Los datos se presentan como n (%) o mediana [rango intercuartil].

2.2. Arquitectura CNN

La Figura 1 muestra la arquitectura CNN implementada en este estudio, compuesta por tres bloques convolucionales (b_1 , b_2 y b_3) repetidos n_{sb} veces. Las capas de cada bloque se seleccionaron en base a arquitecturas CNN propuestas en trabajos previos [12]. La capa de salida se implementó con activación lineal para abordar una regresión y obtener como salida de la CNN la estimación de eventos de apnea por segmento (\hat{y}_{est}).

2.3. Estimación IAH

Posteriormente a la estimación de eventos con la CNN, se calculó el IAH por sujeto sumando los eventos estimados en todos los segmentos de un mismo sujeto y dividiendo entre el tiempo total de registro. Es importante destacar que el valor de esta tasa infraestima el valor del IAH de la PSG original, ya que para nuestra estimación utilizamos el tiempo total de registro, que es mayor que el tiempo total de sueño. Para minimizar este desajuste, re-estimamos el IAH final implementando un modelo de regresión de vectores soporte (SVR) obtenido en el grupo de entrenamiento para corregir este sesgo [12].

2.4. Interpretabilidad del modelo CNN

En este estudio se ha aplicado XAI mediante el método GradCAM con el fin de interpretar y comprender los procesos internos de la CNN en relación con la detección de eventos apneicos e identificar patrones cardíacos asociados con la AOS pediátrica [9]. El cálculo de GradCAM se realizó utilizando los gradientes obtenidos de cada capa convolucional para calcular un mapa de calor específico por capa. El mapa de calor final se obtuvo mediante el promedio de todos los mapas de calor generados en cada capa convolucional resaltando así las regiones más relevantes [13].

2.5. Análisis estadístico

La capacidad diagnóstica de la AOS pediátrica se evaluó mediante matrices de confusión, el valor de precisión de 4 clases (Acc_c) y el coeficiente Cohen's kappa de 4 clases (k) siguiendo los 4 grupos de severidad típicos: no AOS ($IAH < 1$ e/h), AOS leve ($1 \leq IAH < 5$ e/h), AOS moderada ($5 \leq IAH < 10$ e/h) y AOS severa ($IAH \geq 10$ e/h). Además, calculamos la precisión (Acc), sensibilidad (Se), especificidad (Sp), valor predictivo positivo y negativo (VPP y VPN) y razón de verosimilitud positiva y negativa (LR^+ y LR^-) para los umbrales de severidad IAH 1, 5 y 10 e/h.

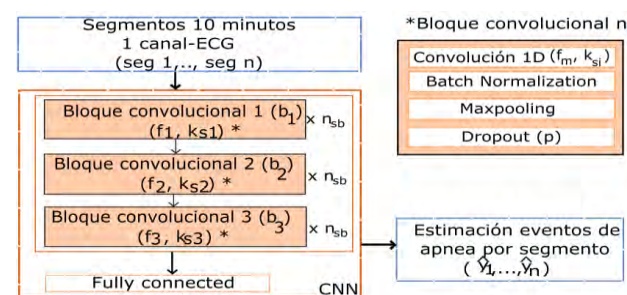


Figura 1: Arquitectura CNN propuesta. N_{sb} : repeticiones de cada bloque convolucional; f_m : filtros; k_s : tamaño del kernel; p : probabilidad.

3. Resultados

3.1. Configuraciones óptimas para CNN y SVR

La configuración óptima de hiperparámetros del algoritmo se obtuvo de manera heurística entrenando el modelo en el grupo de entrenamiento y evaluándose el rendimiento de cada configuración en el grupo de validación. Para ello, se utilizó el valor de k aplicado a la matriz de confusión resultante de comparar el grado de severidad real de la AOS con el derivado de la estimación con nuestro IAH. El modelo CNN con el que se alcanzó el mayor k en validación (0.3857) se obtuvo con los valores de $n_{sb}=5$, $f_i=16$ y $k_{s,i}=17$ para b_1 , $f_2=32$ y $k_{s,2}=9$ para b_2 y $f_3=64$ y $k_{s,3}=7$ para b_3 . Además, se obtuvo un dropout ($p=0$). Los hiperparámetros óptimos en la SVR fueron $C=1,54$, $\epsilon=0,10$ y $\gamma=0,02$.

3.2. Evaluación del rendimiento diagnóstico

La Figura 2 muestra la matriz de confusión tras la clasificación del grado de severidad de AOS pediátrica. En términos globales, se obtuvo un valor de $Acc_4=56,52\%$ y un $k=0.3586$ en el grupo de test. Además, en la Tabla 2 se puede observar como a medida que aumenta la severidad de la enfermedad, la precisión del algoritmo es mayor.

3.3. Identificación de patrones ECG

La Figura 3 muestra los mapas de calor obtenidos con el método GradCAM en segmentos de ECG de 10 minutos. Las Figuras 3 (a.1), 3 (a.2) y 3 (a.3) y 3 (b.1), 3 (b.2) y 3 (b.3) representan los mapas de calor asociados a segmentos en los que la red ha hecho una buena estimación del número de eventos de apnea. En este caso, GradCAM destaca las regiones donde la FC muestra fases de bradicardia y taquicardia en las zonas de transición de un evento de apnea a regiones post-apnea. Además, en las Figuras 3 (a.3) y 3 (b.3) se observa como GradCAM destaca las regiones que abarcan la onda T y el inicio de la siguiente onda P de diferentes latidos. Por otro lado, en las Figuras 3 (c.1) y 3 (c.2) se representa el mapa de calor asociado a un segmento de ECG en el que el modelo ha hecho una predicción poco precisa. A pesar de que GradCAM destaca las regiones entre las que se encuentran

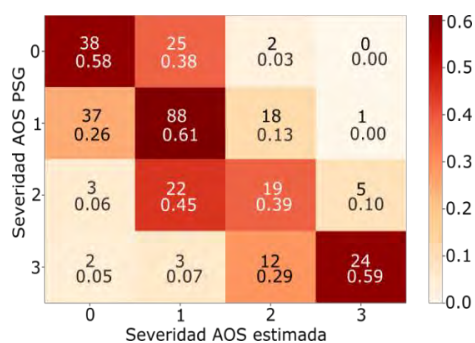


Figura 2: Matriz de confusión para las 4 severidades en el grupo de test. 1: Sin AOS, 2: AOS leve, 3: AOS moderada, 4: AOS severa.

	Se	Sp	VPP	VPN	LR ⁺	LR ⁻	Acc
1 e/h	82,1	58,5	87,7	47,5	2,0	0,3	76,9
5 e/h	66,7	90,0	74,1	86,2	6,6	0,4	82,9
10 e/h	58,5	97,7	80,0	93,7	25,2	0,4	92,3

Tabla 2: Figuras de mérito (%) de la clasificación binaria de la CNN.

contenidos todos los eventos, el modelo no encuentra patrones distintivos del ECG para discriminar entre las zonas con evento y sin él.

4. Discusión y conclusiones

En este trabajo se ha desarrollado un modelo DL interpretable basado en CNN y el ECG que ha alcanzado un alto rendimiento diagnóstico en la AOS pediátrica. Al igual que en la literatura existente, el rendimiento fue superior conforme aumenta el grado de severidad de la enfermedad [6]. Además, puesto que $LR^+>10$ es un indicador robusto para determinar la presencia de una enfermedad, el valor de esta métrica para 10 e/h ($LR^+=25,2$) indica que nuestro modelo presenta una fiabilidad elevada para detectar AOS severa [14]. Por otro lado, GradCAM ha mostrado su utilidad como herramienta de apoyo en la comprensión del comportamiento del ECG en episodios pre/post-apneicos destacando la importancia en las regiones de la onda T y el inicio de la onda P del latido posterior. Estos hallazgos preliminares abren la puerta para investigar en profundidad el comportamiento de estas ondas y evaluar la relación entre la AOS infantil y la aparición de arritmias cardíacas.

Analizando la explicación de los resultados realizada con GradCAM, el uso que hace el modelo CNN de la información de los patrones bradicardia-taquicardia parece una consecuencia natural de lo que se conoce de la respuesta fisiológica del corazón ante los eventos apneicos [2]. Del mismo modo, el uso de la información procedente de las ondas T y P coincide con las evidencias en la dispersión de la onda P y el intervalo QT en AOS pediátrica, principalmente en los casos más severos [15]. En concreto, la duración prolongada de la onda P podría indicar un retardo en la conducción auricular relacionado con los mecanismos fisiopatológicos que contribuyen al desarrollo de fibrilación auricular en AOS adulta [15]. En relación con la onda T, ésta presenta una propiedad de contracción asociada al aumento de la FC; sin embargo, anomalías en esta onda y cambios en el segmento ST podrían sugerir posibles alteraciones cardíacas [16]. Asimismo, intervalos QT muy prolongados están vinculados con una mayor probabilidad de muerte súbita y este hecho también ha sido evidenciado en AOS pediátrica severa [16], [17].

Estudios previos analizaron características derivadas de la señal HRV mediante técnicas convencionales de *machine learning* con el propósito de detectar la AOS pediátrica utilizando los mismos umbrales de corte que en este estudio [8]. Con sus metodologías alcanzaron valores de Acc entre 63,40% y 80,07% para 1 e/h, entre 81,00% y 84,95% para 5 e/h y entre 89,30% y 91,64% para 10 e/h. Nuestra propuesta logró un rendimiento diagnóstico más elevado en 5 y 10 e/h, indicando una mayor aplicabilidad clínica en el diagnóstico de los casos más graves de AOS pediátrica. Estos resultados son especialmente prometedores, ya que son precisamente estos pacientes quienes obtendrán mayor beneficio de un diagnóstico y tratamientos tempranos. Además, en términos generales, nuestra propuesta también obtuvo un valor más alto de k (0,359 vs. 0,166) y Acc_4 (56,52% vs. 41,89%).

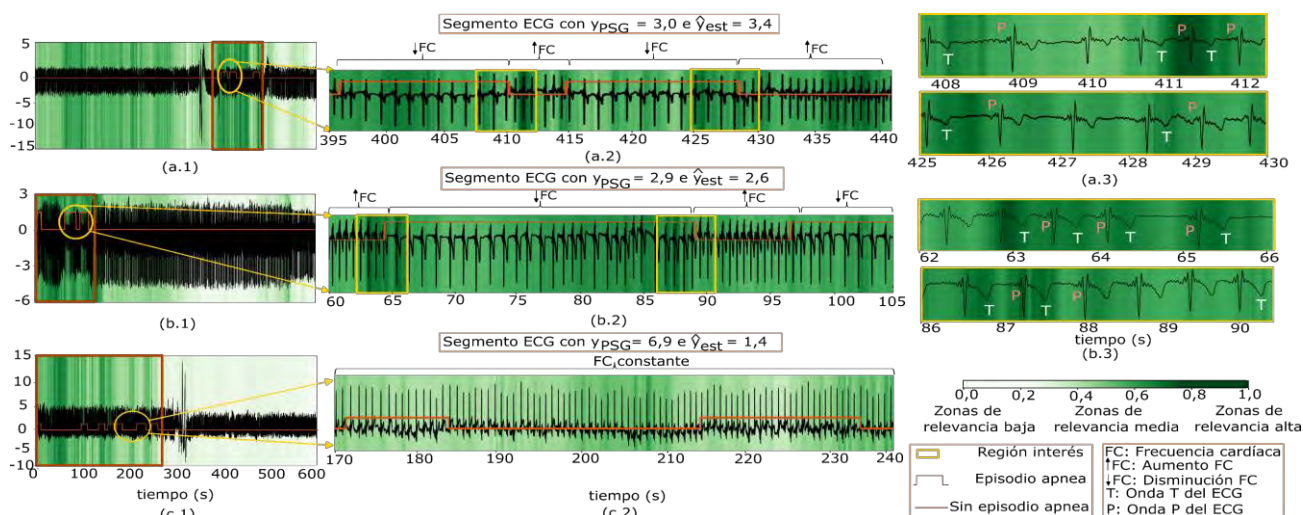


Figura 3: Visualización GradCAM de segmentos de ECG. Las Figuras 3 (a.1), 3 (b.1) y 3 (c.1) muestran la visualización GradCAM de segmentos de 10 minutos. Las Figuras 3 (a.2), 3 (b.2) y 3 (c.2) son ampliaciones de las zonas indicadas con un círculo en las Figuras 3 (a.1), 3 (b.1) y 3 (c.1). En ellas se destacan las regiones más relevantes de episodios pre-post apneicos (verde oscuro) con cuadros amarillos y aparecen indicadas entre llaves las zonas de cambio en el aumento y disminución de la FC (\uparrow FC y \downarrow FC). Las Figuras 3 (a.3) y 3 (b.3) corresponden a imágenes aumentadas de las regiones de interés indicadas con rectángulos en amarillos en las Figuras 3 (a.2) y 3 (b.2). En ellas aparecen indicadas las ondas P y T de las regiones más destacadas por GradCAM. y_{PSG} : número de eventos reales marcados en la PSG; y_{est} : número de eventos estimados por la CNN.

Entre las limitaciones del estudio cabe mencionar que se ha utilizado una sola base de datos, por lo que sería necesario evaluar la capacidad de generalización del modelo propuesto en otras bases de datos de AOS pediátrica. Además, GradCAM fue creado inicialmente para su aplicación en imágenes, por lo que la aplicabilidad en el ámbito de señales fisiológicas y, en concreto la interpretación en el ECG, es más compleja. De este modo, una línea futura de investigación podría centrarse en evaluar el modelo con otros enfoques de XAI.

Como conclusión, la implementación de un modelo CNN interpretable utilizando el ECG de un solo canal puede ser útil para la ayuda en el diagnóstico de AOS pediátrica, así como para identificar patrones específicos en el ECG relacionados con la enfermedad. Por consiguiente, este enfoque plantea una prometedora alternativa a la PSG con un gran potencial para reducir las listas de espera hospitalarias y permitir un diagnóstico y tratamiento rápido, objetivo y preciso de la AOS.

Agradecimientos

Este estudio ha sido financiado por el MCIN/AEI10.13039/501100011033, ERDF, y 'NextGenerationEU/PRTR' (PID2020-115468RB-I00 y PDC2021-120775-I00), por el CIBER-BBN (CB19/01/00012) y el proyecto TinyHeart (Early Stage 2022). El NSRR fue financiado por el NHLBI (R24 HL114473, 75N92019R002). C. García-Vicente cuenta con un contrato "Ayudas para contratos predoctorales para la Formación de Doctores" del Ministerio de Ciencia, Innovación y Universidades (PRE2021-100792).

Referencias

- [1] H. L. Tan *et al.*, "Obstructive sleep apnea in children: A critical update," *Nat. Sci. Sleep*, vol. 5, pp. 109–123, Sep. 2013.
- [2] C. Guilleminault *et al.*, "Cyclical variation of the heart rate in sleep apnoea syndrome. Mechanisms, and Usefulness of 24 h Electrocardiography as a Screening Technique," *Lancet*, vol. 323, no. 8369, pp. 126–131, Jan. 1984.
- [3] C. L. Marcus *et al.*, "Diagnosis and Management of Childhood Obstructive Sleep Apnea Syndrome," *Pediatrics*, vol. 130, no. 3, pp. e714–e755.
- [4] P. Moridian *et al.*, "Automatic diagnosis of sleep apnea from biomedical signals using artificial intelligence techniques: Methods, challenges, and future works," *WIREs Data Min. Knowl. Discov.*, vol. 12, no. 6, Nov. 2022.
- [5] C. Sun *et al.*, "A systematic review of deep learning methods for modeling electrocardiograms during sleep," *Physiol. Meas.*, vol. 43, no. 8, p. 08TR02, Aug. 2022.
- [6] G. C. Gutiérrez-Tobal *et al.*, "Reliability of machine learning to diagnose pediatric obstructive sleep apnea: Systematic review and meta-analysis," *Pediatr. Pulmonol.*, vol. 57, no. 8, pp. 1931–1943, Aug. 2022.
- [7] E. Gil *et al.*, "PTT Variability for Discrimination of Sleep Apnea Related Decreases in the Amplitude Fluctuations of PPG Signal in Children," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 5, pp. 1079–1088, May 2010.
- [8] A. Martín-Montero *et al.*, "Pediatric sleep apnea: Characterization of apneic events and sleep stages using heart rate variability," *Comput. Biol. Med.*, vol. 154, p. 106549, Mar. 2023.
- [9] R. R. Selvaraju *et al.*, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 336–359, Feb. 2020.
- [10] I. Goodfellow *et al.*, *Deep learning*. MIT Press, 2016.
- [11] C. L. Marcus *et al.*, "A Randomized Trial of Adenotonsillectomy for Childhood Sleep Apnea," *N. Engl. J. Med.*, vol. 368, no. 25, pp. 2366–2376, Jun. 2013.
- [12] J. Jiménez-García *et al.*, "A 2D convolutional neural network to detect sleep apnea in children using airflow and oximetry," *Comput. Biol. Med.*, vol. 147, no. April, 2022.
- [13] F. Vaquerizo-Villar *et al.*, "An explainable deep-learning model to stage sleep states in children and propose novel EEG-related patterns in sleep apnea," *Comput. Biol. Med.*, vol. 2, no. 2, p. 107419, Aug. 2023.
- [14] J. J. Deeks *et al.*, "Diagnostic tests 4: likelihood ratios," *BMJ*, vol. 329, no. 7458, pp. 168–169, Jul. 2004.
- [15] C. Kraikriangsri *et al.*, "P-wave dispersion as a simple tool for screening childhood obstructive sleep apnea syndrome," *Sleep Med.*, vol. 54, pp. 159–163, Feb. 2019.
- [16] L. Sörnmo *et al.*, "The Electrocardiogram—A Brief Background," in *Bioelectrical Signal Processing in Cardiac and Neurological Applications*, L. Sörnmo and P. B. T.-B. S. P. in C. and N. A. Laguna, Eds. Burlington: Elsevier, 2005, pp. 411–452.
- [17] C. L. Marcus, "Sleep-disordered Breathing in Children," *Am. J. Respir. Crit. Care Med.*, vol. 164, no. 1, pp. 16–30, Jul. 2001.

Análisis de la variabilidad de la frecuencia cardíaca en maratonianos

M. Javierre¹, M. Kloostermann², P. Van Dam², T. Eijsvogels, E. Pueyo^{1,4}, R. Bailón^{1,4}

¹ BSICoS Grupo, I3A, IIS Aragón, Universidad de Zaragoza, España, {mjavierre, rbailon, epueyo}@unizar.es.

² University Medical Center Utrecht, Utrecht, Países Bajos

³ Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, Países Bajos

⁴ CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

Resumen

Diversos estudios han descrito que un exceso en la cantidad e intensidad de ejercicio físico puede aumentar el riesgo de sufrir arritmias, algunas de las cuales pueden conducir a la muerte súbita cardíaca (MSC). La medida no invasiva de la actividad eléctrica del corazón y, en concreto, los análisis de variabilidad de la frecuencia cardíaca (HRV, por sus siglas en inglés), son unas de las principales vías de investigación de los mecanismos que llevan a la MSC, ya que la información aportada por este análisis está directamente relacionada con la modulación del sistema nervioso autónomo. En este estudio se analizaron señales electrocardiográficas adquiridas de 10 deportistas mediante “Body Surface Potential Mapping” (BSPM) de 64 derivaciones en tres instantes temporales: una semana antes, justo después y dos semanas después de correr una maratón. Se desarrollaron métodos de procesamiento de señal para seleccionar las derivaciones más robustas para la evaluación de HRV, que resultaron ser las situadas en el apéndice del corazón. Mediante el estudio de índices temporales (HRM, RMSSD, SDNN, SDDSD y PNN50) y frecuenciales de HRV (PHF, PLFN, LFHF y PHFN) se observó que: 1) tras la maratón hay una activación tardía del sistema nervioso parasimpático y una desactivación tardía del sistema nervioso simpático y 2) dos semanas después de la maratón todos los índices volvieron a sus valores iniciales. Este desequilibrio autonómico, observado justo después del ejercicio, podría contribuir a la generación de arritmias tras la realización de la maratón.

1. Introducción

La muerte súbita cardíaca (MSC) es una de las principales causas de muerte en países occidentales. En el subgrupo de la población que practica deporte de alta intensidad algunos estudios han descrito una incidencia de MSC de 1,6 muertes por cada 100.000 individuos, frente a la incidencia de 0,75 muertes por 100.000 individuos en el subgrupo de la población que no lo practica [1]. Otros estudios que han analizado registros de un gran número de participantes (> 14 millones) que han disputado carreras de larga distancia entre 1976 y 2009 han concluido que la incidencia de MSC es entre 0.6 y 1.9 por cada 100.000 atletas [2].

Bajo la hipótesis de que la práctica de ejercicio físico muy intenso, como correr una maratón, puede aumentar el riesgo de MSC, varios estudios han intentado encontrar marcadores tanto de forma invasiva como no invasiva que pudieran explicar la predisposición a sufrir MSC. En su mayoría, estos estudios utilizan resonancias magnéticas y muestras de sangre, intentando demostrar que la concentración de troponina y de miosina aumen-

ta después de la práctica de ejercicio intenso [3]. Se han encontrado marcadores que indican tanto daño cardíaco reversible como daño en la integridad del tejido del miocardio, que dos semanas después de la fecha del maratón no se ha reparado [4]. Esta integridad está directamente relacionada con la permeabilidad del miocardio, la concentración de troponina y la filtración de las reservas citosólicas a la circulación, marcadores que están estrechamente conectados con problemas cardiovasculares que pueden acabar en MSC [5].

El análisis de la variabilidad de la frecuencia cardíaca (heart rate variability, HRV) ha sido utilizado en otros trabajos como predictor de MSC, utilizando marcadores temporales y frecuenciales. En [6], se extrajeron 20 parámetros del análisis de HRV, pudiendo predecir MSC 2 minutos antes de que ocurra con una precisión del 90%. En [7], se propuso la combinación de 24 parámetros diferentes y se clasificaron sucesos de MSC 12 minutos antes de su ocurrencia con una precisión del 84%. En este trabajo se demuestra que mediante la medida no invasiva de la actividad cardíaca es posible ver diferencias en la modulación del sistema nervioso autónomo tras la realización de actividad física intensa, las cuales podrían relacionarse con daño cardíaco reversible, ya que al cabo de dos semanas todas las medidas analizadas vuelven a niveles basales. Además, a partir de registros de alta resolución espacial, se identificó la zona que otorga más robustez a la obtención de la frecuencia cardíaca, haciendo más viable este tipo de análisis en un futuro.

2. Materiales

2.1. Población de estudio

La población de estudio estaba compuesta por 12 maratonianos bien preparados físicamente, sin ninguna patología cardiovascular y sin ningún factor de riesgo. De estos 12, uno no pudo terminar la carrera debido a una complicación en la rodilla, por lo que fue excluido de la población de estudio. De los 11 hombres restantes, uno no pudo ser evaluado debido a un implante metálico en la espalda, por lo que se analizó un subgrupo de 10 corredores varones con una edad mediana (rango intercuartílico) de 51,1 [50,1-55,12] años.

2.2. Señales electrocardiográficas

Las adquisiciones realizadas consistieron en señales electrocardiográficas (ECG) de 64 electrodos, 61 electrodos procedentes del mapeo del potencial en la superficie corporal (BSPM, por sus siglas en inglés) y 3 más colocados en posiciones típicas de algunas derivaciones del

ECG estándar. Las señales fueron tomadas antes de la carrera (1 semana antes), justo al finalizar la misma (4 ± 2 horas post-maratón) y 2 semanas después de la carrera a una frecuencia de muestreo de 2048 Hz. El BSPM de 64 electrodos se adquirió mediante 12 tiras de parches colocadas horizontalmente en el torso del sujeto. La duración de las señales varió entre 5 y 10 minutos.

3. Métodos

3.1. Preprocesado de las señales

Se aplicó un filtro paso-alto de Butterworth con frecuencia de corte de 0.25 Hz para eliminar la línea de base de las señales electrocardiográficas. Después se aplicaron filtros notch de frecuencia 50, 100, 150 y 200 Hz para mitigar los efectos de la red eléctrica y, por último, se aplicó un filtro paso-bajo tipo Butterworth con frecuencia de corte de 200 Hz.

3.2. Delineación y elección de la derivación

En cada registro se identificó la derivación donde la detección de los puntos fiduciales del complejo QRS fuese más robusta, comparando la serie temporal de intervalos RR (inverso de la frecuencia cardíaca) correspondiente a la primera componente de las extraídas mediante análisis de componentes principales (principal component analysis, PCA) con la serie temporal de RR extraída de cada derivación.

En la Figura 1 se representa la primera componente extraída del análisis PCA, que es una combinación lineal de las 64 derivaciones originales, maximiza la varianza y enfatiza el complejo QRS [8]. Se delinearon tanto la señal obtenida por PCA como las 64 originales para obtener los puntos R y S con un delineador uniderivacional basado en la transformada wavelet [9].

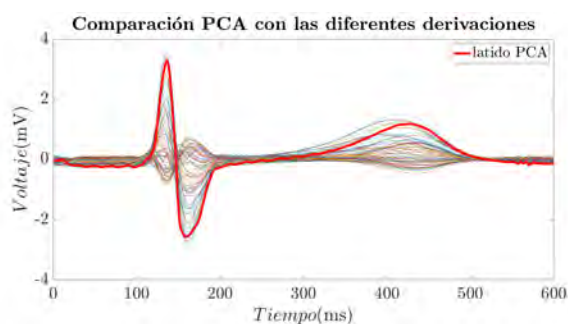


Figura 1. Comparación de la señal obtenida mediante PCA y las derivaciones originales.

Si en una derivación se detectaron los puntos R en al menos un 75 % de los latidos y los puntos S en al menos un 75 % de los latidos, se seleccionó el tipo de punto fiducial con mayor amplitud media para el posterior análisis HRV. Si no se cumplió la condición mencionada, se seleccionó el tipo de punto fiducial más frecuente. A partir de las detecciones se calcularon los intervalos RR como la diferencia de posición entre dos puntos R o S contiguos de manera sucesiva, por una parte para la señal PCA y por otra, para las 64 derivaciones. Utilizando la serie de

RR procedente de PCA, se calculó la raíz cuadrada del error cuadrático medio (root mean square error, RMSE) de esta con las otras 64 derivaciones y se escogió aquella derivación con menor RMSE. Un ejemplo de este procedimiento puede verse en la Figura 2.

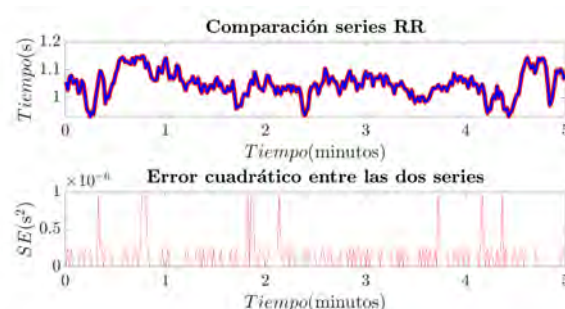


Figura 2. Comparación de la serie RR extraída de la primera componente obtenida mediante PCA (rojo) y la serie RR de la derivación asociada con menor RMSE (azul) de una sesión, en este caso proveniente de la derivación 42. El RMSE se establece en $(3,35) \cdot 10^{-4}$ s.

3.2. Variabilidad de la frecuencia cardíaca y análisis estadístico

Para el análisis HRV se utilizaron: 1) la serie RR mencionada anteriormente, que se corrigió para evitar el efecto tanto de latidos ectópicos como errores en detecciones (falsos positivos/falsos negativos) utilizando los métodos descritos en [10]. A esta serie se le denomina serie NN y se usó para extraer los índices temporales de HRV; 2) la señal moduladora $m(t)$, que se obtuvo a partir del modelo IPFM [10] y cuya densidad espectral de potencia se estimó mediante el periodograma de Welch utilizando ventanas de 100 segundos con un solapamiento del 50 %, considerando la señal previamente muestreada a 4 Hz y normalizada para evitar la influencia de la frecuencia cardíaca media [11].

Del dominio temporal se extrajeron los siguientes marcadores: frecuencia cardíaca media (HRM), desviación estándar de los intervalos NN corregidos ($SDNN$), raíz de la media cuadrática de la diferencia entre intervalos NN ($RMSSD$), desviación estándar de la diferencia entre intervalos NN adyacentes ($SDSD$), porcentaje de los intervalos NN consecutivos que difieren más de 50 ms entre sí ($PNN50$) [12]. Del dominio frecuencial se analizaron los siguientes marcadores: potencia de la componente de baja frecuencia (LF) en la banda [0.04-0.15] Hz, que mide tanto la actividad del sistema nervioso simpático como parasimpático (PLF); potencia de la componente de alta frecuencia (HF), comprendida en la banda [0.15-0.4] Hz, que mide la actividad del sistema nervioso parasimpático (PHF); potencia de la componente HF normalizada por la potencia total en las bandas HF y LF ($PLFN$); potencia de la componente HF normalizada por la potencia total en las bandas HF y LF ($PHFN$); cociente de la potencia de LF entre la potencia de HF, que detecta si uno de los sistemas nerviosos predomina sobre el otro ($LFHF$) [12].

Como no todos los marcadores siguen una distribución normal, la prueba elegida para determinar la diferencia estadística entre los tres tipos de sesiones fue la prueba de los rangos con signo de Wilcoxon, fijando el umbral de los rangos con signo de Wilcoxon, fijando el umbral de la p_{valor} para la significación estadística en 0,05.

4. Resultados y discusión

Se determinó cuáles de las 64 derivaciones registradas mediante BSPM llevaban a series de RR más próximas a la serie de RR obtenida a partir de la señal calculada aplicando PCA. En la Figura 3 se muestran las derivaciones que llevaron a menor RMSE en dicha comparación. Los resultados indican que las derivaciones que llevan a un cálculo más robusto del frecuencia cardíaca son las correspondientes a electrodos localizados próximos al apéndice del corazón, en la zona baja del pectoral izquierdo. Si comparamos la mejor derivación (derivación 42), con la peor derivación (derivación 10, situada en la parte inferior del costado derecho) la diferencia en términos de RMSE es, en mediana entre sesiones de registro, de 11,3 [1,8-34,3] ms.

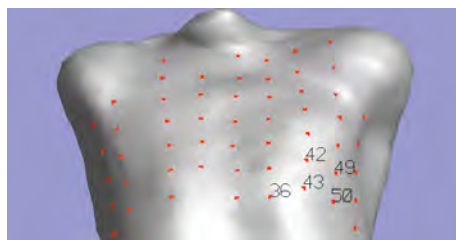


Figura 3. Cinco mejores derivaciones para realizar el análisis de la variabilidad de la frecuencia cardíaca y su posición sobre el torso.

En la Figura 5 se muestran los diagramas de caja de los diferentes marcadores frecuenciales y temporales de la variabilidad de la frecuencia cardíaca. Se indican con un asterisco las comparaciones que dieron lugar a diferencias estadísticamente significativas entre las mediciones realizadas en dos sesiones de registro diferentes.

Como puede verse en la Figura 5 y confirmaron los resultados de las pruebas estadísticas no paramétricas, los índices frecuenciales del análisis de HRV fueron capaces de detectar una bajada de potencia en la banda HF tras el ejercicio, mientras que la potencia de la banda LF, relacionada con el sistema simpático, permaneció estable. Se apreció una subida en los índices que relacionan las potencias de HF y LF tras la maratón. En la tercera sesión todos los marcadores volvieron a la normalidad. Los marcadores temporales marcan una disminución de la HRV inmediatamente después de la maratón, llegando a una estabilización en la tercera sesión registrada dos semanas más tarde. Los resultados obtenidos concuerdan con otros trabajos [13]. Takahashi *et al* concluyeron que la reducción del índice PHF tras el ejercicio aeróbico está directamente relacionada con el decrecimiento de la actividad parasimpática como respuesta a la intensidad de la maratón. Como se demuestra en este trabajo, esta

reducción se mantiene durante un periodo de tiempo (4 ± 2 horas) y este retraso en la activación del sistema nervioso parasimpático puede estar relacionada con el aumento de la frecuencia cardíaca respecto a las sesiones uno y tres, ya que, al no haberse activado el sistema nervioso parasimpático de forma rápida, el cuerpo del atleta no está consiguiendo llevar a una situación de reposo a su sistema cardiovascular, causando un desequilibrio entre el sistema nervioso simpático y parasimpático, tal y como indican los parámetros que relacionan las dos bandas de frecuencia. Resultados similares se obtuvieron en [14, 15] donde se describieron tanto un retraso en la activación del sistema nervioso parasimpático como un retraso en la bajada de actividad del sistema nervioso simpático. Una elevación en la actividad simpática o una reducción en la actividad vagal tras el ejercicio se han relacionado con el desarrollo de arritmias ventriculares malignas, mientras que la rápida recuperación simpática parece ejercer un efecto protector en relación al riesgo arritmico [16].

Las causas de estos efectos todavía no están claras, pero pueden estar relacionadas con los niveles de neuropéptidos, liberados por los nervios simpáticos durante el ejercicio, que pueden atenuar la actividad del sistema nervioso parasimpático durante la recuperación [13]. Otros estudios relacionan el aumento de la frecuencia cardíaca con el consumo elevado de oxígeno que exige el cuerpo humano durante y después de la carrera [17]. Además, este fenómeno ha demostrado estar relacionado con la intensidad del ejercicio, determinando que, cuanto mayor es la intensidad del ejercicio, más tarda el cuerpo en volver a una situación estable [18]. Los resultados del análisis de índices temporales de HRV concuerdan con los índices frecuenciales. La disminución de estos índices tras la maratón con respecto a los valores basales antes de realizar el ejercicio está relacionada con el retardo en la activación del sistema nervioso parasimpático. Estos marcadores vuelven a su normalidad tras dos semanas de haber realizado la maratón.

5. Conclusiones

Se evaluaron registros de la actividad eléctrica cardíaca sobre la superficie corporal con alta resolución espacial (BSPM) para caracterizar los efectos de la maratón y analizar si pueden estar relacionados con el daño cardíaco. Se desarrollaron métodos de procesamiento de señal y se confirmó que hay cambios en la actividad eléctrica cardíaca justo después de realizar una actividad aeróbica, volviendo dicha actividad a su situación basal tras dos semanas de la realización de la maratón. Los cambios observados indicaron un desequilibrio entre las dos ramas del sistema nervioso autónomo que podría contribuir a la generación de arritmias tras la realización de la maratón, si bien dichos cambios fueron revertidos poco tiempo después de la realización del ejercicio. Cabe destacar que la población del análisis es reducida, de tan solo 10 corredores con las mismas condiciones, lo que supone una limitación en el estudio.

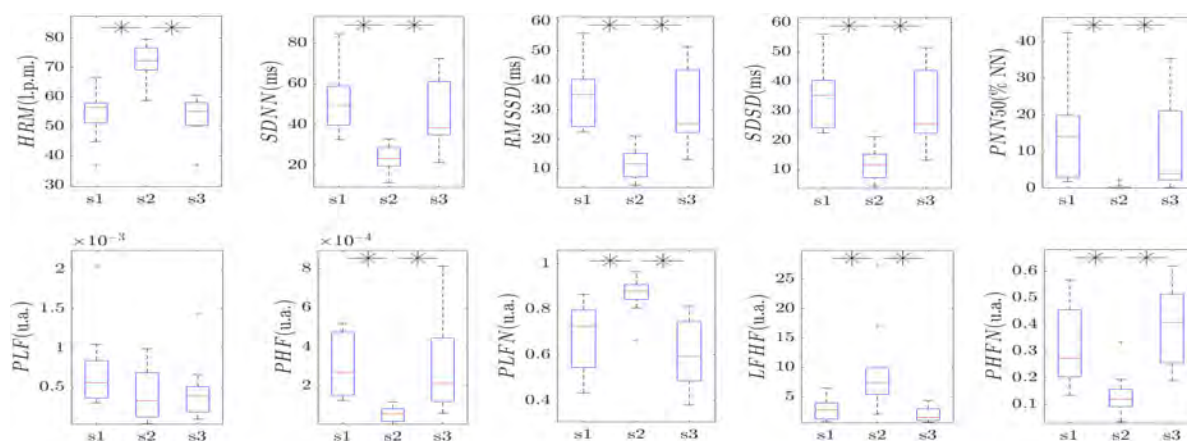


Figura 5. Diagramas de caja de los diferentes marcadores de HRV en las tres sesiones registradas. La parte superior corresponde a marcadores de análisis temporal y la parte inferior al análisis frecuencial. Las diferencias significativas entre sesiones están marcadas con un asterisco.

Agradecimientos

Este trabajo ha sido parcialmente financiado por los proyectos TED2021-131106B-I00, TED2021-130459B-I00, PID2021-126734OB-C21 y PID2022-140556OB-I00 (Ministerio de Ciencia e Innovación y Fondo Social Europeo), España, el grupoBSICoS T39_23R (Gobierno de Aragón y Fondo Social Europeo).

Referencias

- [1] A. Boraita. Muerte súbita y deporte. ¿Hay alguna manera de prevenirla en los deportistas? *Revista Española de Cardiología*, 55(4):333–336, January 2002.
- [2] O. Waite *et al.* Sudden cardiac death in marathons: a systematic review. *The Physician and Sportsmedicine*, 44(1):79–84, January 2016.
- [3] V. Aengevaeren *et al.* Marathon running transiently depletes the myocardial lipid pool. *Physiological Reports*, 8(17):e14543, September 2020.
- [4] V. Aengevaeren *et al.* Myocardial injury and compromised cardiomyocyte integrity following a marathon run. *JACC Cardiovascular Imaging*, 13(6):1445–1447, June 2020.
- [5] V. Aengevaeren *et al.* Cardiac biomarker kinetics and their association with magnetic resonance measures of cardiomyocyte integrity following a marathon run: Implications for postexercise biomarker testing. *Journal of the American Heart Association*, 10(13):e020039, July 2021.
- [6] M. Ebrahimzadeh, E. Pooyan. Early detection of sudden cardiac death by using classical linear techniques and time-frequency methods on electrocardiogram signals. *Journal of Biomedical Science and Engineering*, 04(11):699–706, 2011.
- [7] M. Shi *et al.* Early detection of sudden cardiac death by using ensemble empirical mode decomposition-based entropy and classical linear features from heart rate variability signals. *Frontiers in Physiology*, 11:118, February 2020.
- [8] F. Castells *et al.* Principal component analysis in ECG signal processing. *EURASIP Journal on Advances in Signal Processing*, 2007(1), December 2007.
- [9] J. P. Martínez *et al.* A wavelet-based ECG delineator: evaluation on standard databases. *IEEE Transactions on Biomedical Engineering*, 51(4):570–581, April 2004.
- [10] J. Mateo, P. Laguna. Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal. *IEEE Transactions on Biomedical Engineering*, 50(3):334–343, March 2003.
- [11] R. Bailón *et al.* The integral pulse frequency modulation model with time-varying threshold: Application to heart rate variability analysis during exercise stress testing. *IEEE Transactions on Biomedical Engineering*, 58(3):642–652, March 2011.
- [12] L. Sörnmo, P. Laguna. *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. Elsevier, 2005.
- [13] T. Takahashi *et al.* Difference in human cardiovascular response between upright and supine recovery from upright cycle exercise. *European Journal of Applied Physiology*, 81(3):233–239, February 2000.
- [14] W. Niewiadomski *et al.* Suppression of heart rate variability after supramaximal exertion. *Clinical Physiology and Functional Imaging*, 27(5):309–319, September 2007.
- [15] T. Peçanha *et al.* Methods of assessment of the post-exercise cardiac autonomic recovery: A methodological review. *International Journal of Cardiology*, 227:795–802, January 2017.
- [16] K. Heffernan *et al.* Cardiac autonomic modulation during recovery from acute endurance versus resistance exercise. *European Journal of Cardiovascular Prevention Rehabilitation*, 13(1):80–86, February 2006.
- [17] J. Smolander *et al.* Estimating oxygen consumption from heart rate and heart rate variability without individual calibration. *Clinical Physiology and Functional Imaging*, 31(4):266–271, July 2011.
- [18] T. Mann *et al.* Effect of exercise intensity on post-exercise oxygen consumption and heart rate recovery. *European Journal of Applied Physiology*, 114(9):1809–1820, September 2014.

Valoración de la actividad muscular asociada al uso de herramientas laparoscópicas mediante EMG. Nueva propuesta de análisis bivariado amplitud-frecuencia

I. Alandí-Rocafull¹, J.L. Martínez-de-Juan¹, A. Conejero Rodilla², H.M. Pace Bedetti², G. Prats-Boluda¹

¹ Centro de Investigación e Innovación en Bioingeniería, Universitat Politècnica de València, Valencia, Spain, {irene.alandi@gmail.com, jlmartinez@ci2b.upv.es, gprats@ci2b.upv.es}

² Instituto de Diseño y Fabricación (IDF), Universitat Politècnica de València, Valencia, Spain, {ancoro@idf.upv.es, horaciopacebedetti@gmail.com}

Resumen

El avance de la medicina dirige las técnicas de intervención hacia aquellas mínimamente invasivas (CMI). Su instrumental, en ocasiones, carece de ergonomía lo que puede provocar lesiones al personal sanitario. El objetivo de este trabajo es evaluar, a través de biomarcadores obtenidos de registros electromiográficos de superficie (sEMG), dos herramientas de laparoscopia: un prototipo y otra empleada en la práctica clínica habitual. Para ello, se dispone de registros de sEMG de 17 sujetos de los músculos bíceps, deltoides, trapecio y flexor radial del carpo, obtenidos mediante un protocolo isométrico. Se han computado parámetros temporales, espectrales y de complejidad como biomarcadores de la actividad muscular determinando la existencia de diferencias estadísticamente significativas entre herramientas y posiciones para cada uno de los biomarcadores. Se ha completado el estudio a través de un Joint Analysis of Spectrum and Amplitude (JASA) tradicional y con una nueva propuesta que considera otros parámetros. Finalmente, los biomarcadores parecen apuntar hacia una menor tendencia a la aparición de fatiga muscular con la herramienta prototipo para el bíceps, deltoides y trapecio, y con la herramienta control para el flexor radial del carpo. Los nuevos análisis bivariados propuestos permiten una caracterización más exhaustiva del estado muscular.

1. Introducción

La cirugía laparoscópica se clasifica como una técnica quirúrgica mínimamente invasiva. Sus aplicaciones más usuales son cirugías de la cavidad abdominal, siendo más prevalentes las del aparato digestivo. Éstas consisten en operar dicha cavidad sin la necesidad de realizar grandes incisiones de apertura. Para ello, se introducen de forma transparietal diversos instrumentos quirúrgicos, a través de unos elementos conocidos como trócares [1]. Esto ha permitido reducir los riesgos, costes y complicaciones asociados a este tipo de operaciones [2].

A pesar de sus ventajas, estas técnicas también poseen limitaciones entre las que destacan las debidas a la ergonomía del material quirúrgico [3]: las posturas que los cirujanos adoptan para alcanzar las estructuras anatómicas son más extremas y requieren un mayor esfuerzo muscular que aquellas propias de las cirugías abiertas. Uno de los principales factores que contribuyen a este hecho es el tipo de agarre de las herramientas. Por ello resulta tan esencial la realización de estudios de mejora para las mismas.

Como consecuencia en el ámbito clínico son comunes dolencias tales como sobrecarga muscular de los miembros superiores, dolor en el cuello, hombro, brazo y muñeca.

Los grupos musculares más afectados son [4]: los flexores y extensores del brazo y antebrazo, destacando el flexor y extensor de los dedos, el bíceps braquial, los flexores *carpi radialis* y *carpi ulnaris* y, finalmente, los efectores del hombro: deltoides y trapecio.

El Instituto de Diseño y fabricación automatizada (IDF) en colaboración con el centro de Investigación e Innovación en Bioingeniería (CI2B), ambos pertenecientes a la Universitat Politècnica de València (UPV), han desarrollado una herramienta prototipo bajo el concepto de *Postural Freedom* con el objetivo de reducir la carga muscular que el cirujano presenta tras la cirugía.

El objetivo del presente trabajo es caracterizar y evaluar el estado muscular asociado al uso de dos herramientas laparoscópicas (control y prototipo desarrollado) a partir registros de electromiografía de superficie (sEMG) y el cómputo y análisis estadístico comparativo de biomarcadores temporales, espectrales y no lineales. Además, se evalúa el estado muscular mediante un análisis combinado de amplitud frecuencia tradicional, (JASA), que realiza una representación en cuatro cuadrantes de las diferencias de valor de dos parámetros (temporal y espectral) [5]. Finalmente, se proponen análisis bivariados con otros parámetros, suponiendo una alternativa a los JASA [5] tradicionales.

2. Materiales y métodos

2.1. Registros de sEMG

Se dispone de una base de datos de registros de sEMG simultáneos de los músculos bíceps, deltoides, trapecio y flexor radial del carpo de 17 sujetos, sin experiencia previa, sometidos a un protocolo isométrico procedentes de un estudio anterior [6]. Para el registro de cada músculo se colocaron dos electrodos de Ag-AgCl siguiendo las recomendaciones del programa SENIAM [7]. Las señales sEMG se obtuvieron empleando el amplificador de bioseñales de propósito general GrassP511 AC® (10V/mV en bíceps, 5V/mV en deltoides, 20V/mV en trapecio y 20V/mV en el flexor radial del carpo y un filtro paso-banda con frecuencias de corte fijadas en 3Hz y 1 kHz). Finalmente, la señal se adquirió a una frecuencia de muestreo de 5 kHz a través de la tarjeta NIUSB-6229 (National Instruments®, Texas USA). El protocolo consistió en colocar la herramienta en una posición estática durante 55 segundos, hasta completar un total de 12 registros: 10 posiciones, cuyo orden se puede observar en

la *Fig.1*. Todos los participantes emplearon ambas herramientas en orden aleatorio para evitar el posible sesgo por aprendizaje.

2.2. Herramientas evaluadas

La primera herramienta evaluada es usualmente empleada en la práctica clínica (control): AdTec® de un solo uso de la empresa B.Braun (Alemania). La herramienta prototipo, por su parte, se basa en el concepto de *Postural Freedom* [6] el cual sostiene que los cirujanos mantienen posiciones forzadas durante los procedimientos de laparoscopia, las cuales pueden derivar en lesiones a largo plazo, debido a la falta de ergonomía del instrumental quirúrgico. Si estas herramientas permitiesen un mayor rango de movimiento, las posturas adoptadas no serían tan extremas, reduciéndose así los niveles de fatiga a los que se ven sometidos. Esto se implementa en este nuevo prototipo a través de una articulación tipo rótula, la cual permite un mayor rango de movimiento durante el procedimiento, tal y como se puede observar en la *Fig 1*.

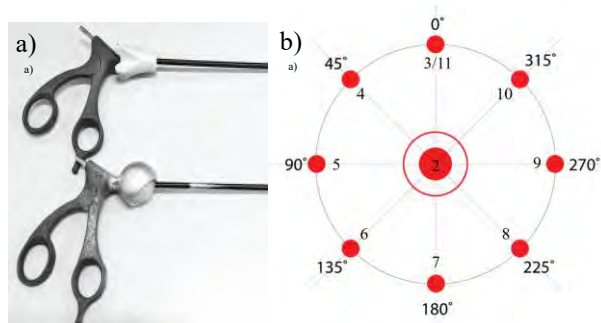


Figura 1. Herramientas control (superior) y prototipo (inferior) (a) y plantilla empleada durante el protocolo (b).

2.3. Preprocesado y biomarcadores considerados

Se desarrolla un paquete de análisis propio mediante el entorno de programación Matlab R2022b. El preprocesado elegido para los registros de sEMG consiste en cinco pasos: (1) filtrado paso-banda de fase cero, tipo Butterworth, de orden 5, (frecuencia de corte inferior de 10 Hz y superior de 500 Hz); (2) diezmo (promediado) por un factor de 4; (3) descarte de los primeros 10 y últimos 5 segundos de cada uno de los registros; (4) eliminación del valor medio y (5) obtención de biomarcadores temporales y *zscore*, para los biomarcadores espectrales y de complejidad [8].

Como biomarcadores univariados, se ha computado el parámetro temporal valor eficaz (RMS); espectrales: frecuencia mediana (MPF) y Spectral Moment Ratio (SMR, considerando frecuencias entre 5 y 200 Hz y momentos de orden -1 y 5), y de complejidad: entropía muestral (SE, $r=0.15$ y $m=2$), dimensión fractal de Higuchi (HFD, $n=13$ y $m=50$) y Lempel-Ziv binario (LZ).

La literatura ha asociado la aparición de fatiga muscular a un aumento del RMS [9], una disminución de la MPF [9], un aumento del SMR [10], y una disminución de la SE [11], y del HFD [12] y el LZ [13].

Para la determinación de las diferencias estadísticas se emplea el estudio de muestras pareadas Wilcoxon, con un intervalo de confianza de $p=0.05$. Se comparan, por un lado, las diferencias estadísticas entre herramientas para

una misma posición en la *Fig. 3*. Por otro lado, se evalúan en la *Tabla 1* las diferencias para cada herramienta por separado entre las posiciones de entrada y salida, análogas en cuanto a colocación del instrumento.

2.4. JASA y alternativa propuesta

Los análisis JASA permiten caracterizar de una forma más completa el estado muscular en estudios ergonómicos a través del estudio combinado de dos biomarcadores: temporal (RMS) y espectral (MF o MPF) [5]. La propuesta realizada en el presente trabajo sugiere la sustitución del parámetro espectral tradicionalmente empleado por el SMR, que ha demostrado ser más sensible a cambios en el espectro asociados a la posible aparición de fatiga muscular [8]. Atendiendo al comportamiento de estos biomarcadores asociado en la bibliografía con la aparición de fatiga muscular, se puede comprobar que una disminución de la MF o MPF es equivalente a un aumento del SMR [7,8]. Por tanto, la distribución de cuadrantes tradicional de los JASA cambiará para la representación RMS vs. SMR: los cuadrantes de la parte superior pasarán a la parte inferior y viceversa, se puede observar la nueva distribución de cuadrantes en la *Fig. 2*.

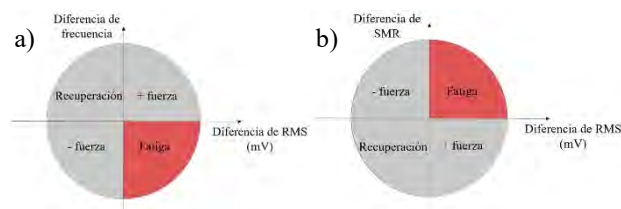


Figura 2. JASA tradicional (a) y alternativa propuesta (b).

Para el presente estudio, se comparará en la *Fig. 4* el valor de cada parámetro en la salida, por músculos, con respecto a su análogo para la posición de entrada, calculando las diferencias de éstos.

3. Resultados y discusión

En la *Fig. 3* representa la evolución de los valores de los biomarcadores a lo largo del protocolo. Se evalúan las tendencias asociadas con la fatiga muscular: disminución de RMS [8] y SMR [9] y aumento de MPF [8] y de parámetros de complejidad [10-12].

En cuanto al RMS, destacan los valores elevados del trapecio (verde) con la herramienta control (línea discontinua), respecto a los correspondientes a la herramienta prototipo, asociando mayores valores de RMS a una posible aparición de la fatiga muscular [9]. El resto de los músculos y herramientas presentan valores inferiores en magnitud, similares entre ellos, en todos los casos presentándose siempre por encima las líneas de tendencia correspondientes al control (discontinuas), siguiendo el comportamiento del músculo anterior. Para los biomarcadores espectrales (MPF y SMR) el comportamiento de los músculos y herramientas es análogo al anteriormente mencionado.

Se aprecian valores más extremos, alejados de la tendencia típicamente asociada a la posible aparición de fatiga [9,10], para el flexor radial del carpo (rosa) siendo más notable en el control (línea discontinua). En cambio, el bíceps (azul)

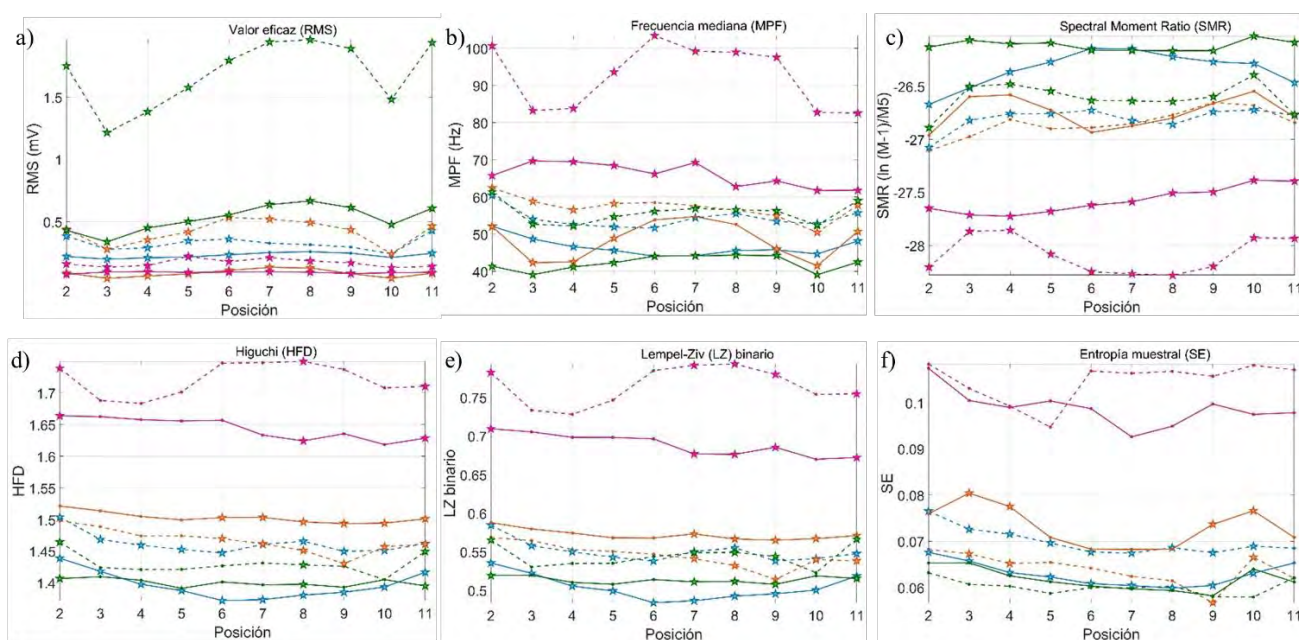


Figura 3. Representación temporal del promedio de los biomarcadores univariados para las herramientas control (línea discontinua) y prototipo (continua) en los músculos bíceps (azul), deltoides (naranja), trapecio (verde) y flexor radial del carpo (rosa); con estrellas destacando las posiciones que presentan diferencias estadísticamente ($p < 0.05$) significativas entre

y trapecio (verde) presentan el comportamiento contrario, pudiéndose sugerir una mayor fatiga de estos músculos con el prototipo (línea continua).

Finalmente, los biomarcadores de complejidad (SE, LZ y HFD) reducen sus valores ante la aparición de dicha fatiga [11-13], apuntando a ello en mayor grado en el bíceps con el prototipo (azul continua) y deltoides con el control (naranja discontinua) y presentando pocas diferencias para el trapecio (verde) y flexor del carpo (rosa). Como se puede observar, los resultados obtenidos para la valoración del estado muscular en algunos casos como el trapecio o deltoides a través de este es controvertido, por lo que se proceden a realizar los estudios de la *Tabla 1* y *Fig. 4*.

La *Tabla 1* recoge la existencia de diferencias estadísticamente significativas para cada una de la herramientas y biomarcadores entre las posiciones de entrada y salida del protocolo (misma posición estados de inicio y final del protocolo). Se aprecia una reducción del número de biomarcadores que presentan diferencias estadísticamente significativas entre estas dos posiciones con la herramienta prototipo, para los músculos bíceps, deltoides (de 5 biomarcadores control a 3 prototipo) y trapecio (de 3 biomarcadores control a ninguno prototipo) mientras que se observa un aumento para el flexor radial del carpo (de un parámetro en control a 4 con el prototipo).

El análisis JASA tradicional [5] y la nueva propuesta efectuada en la *Fig. 4* arrojan resultados acordes con lo observado para las posiciones de entrada y salida de forma aislada: los valores asociados a los músculos bíceps (azul) y deltoides (naranja) se encuentran en zona de mayor predisposición a la fatiga con ambas herramientas [5]. Lo mismo ocurre para la herramienta control en el caso del trapecio (puntos huecos verdes). Para el bíceps, los valores correspondientes a la herramienta prototipo (rellenos) se

encuentran más cercanos al origen de coordenadas lo que indicaría que las diferencias entre biomarcadores, y por tanto la variación de estado muscular, es menos acusado en este caso. Para el flexor del carpo (rosa) con la herramienta prototipo se encuentra en zona de fatiga mientras que con el control (puntos huecos) se sitúa en zonas de reducción de fuerza, de acuerdo con lo que parecían indicar los resultados de entrada y salida. Finalmente, indicar que la diferencia entre ambas representaciones JASA es notoria en la localización de los puntos correspondientes al trapecio en el prototipo (puntos rellenos verdes): en el JASA tradicional, se encuentra en la zona de aumento de fuerza [5], mientras que, en la nueva propuesta, se sitúa en zona de fatiga. Esto puede deberse a que, tal y como se ha demostrado, el SMR es un parámetro más sensible que la MPF ante la aparición de cambios en el espectro [9], por lo que su potencial sensibilidad a la detección de indicios de fatiga muscular es mayor.

4. Conclusiones

El análisis los biomarcadores univariados derivados del sEMG apuntan hacia una menor predisposición a la fatiga muscular con la herramienta prototipo para el bíceps, deltoides y trapecio, y un aumento para el flexor radial del carpo. La propuesta de análisis bivariado JASA realizada en el presente trabajo empleando el SMR permite realizar una valoración más exhaustiva del estado muscular, puesto que su mayor sensibilidad a la fatiga muscular indicada por la literatura parece caracterizar el estado del trapecio de forma más fiable y forma acorde con lo sugerido por los análisis univariados de las posiciones de entrada y salida. Estos resultados preliminares deben ser corroborados y ampliados en estudios futuros

		RMS	MPF	SMR	LZ	HFD	SE
Bíceps	Control	-	↓***	↑***	↓***	↓***	↓***
	Prototipo	-	↓***	↑*	↓*	-	-
Deltoides	Control	-	↓*	↑***	↓**	↓**	↓***
	Prototipo	-	-	-	↓*	↓**	↓*
Trapecio	Control	↑*	↓*	↑*	-	-	-
	Prototipo	-	-	-	-	-	-
Flexor radial del carpo	Control	-	-	↑*	-	-	-
	Prototipo	↑*	-	-	↓*	↓*	↓***

Tabla 1. Comparativa entre las posiciones de entrada (nº2) y salida (nº 11) por músculo y herramienta. Se representan con flechas las tendencias de los biomarcadores en la posición de salida, en comparación con la entrada, en color rojo las que siguen el comportamiento asociado a la aparición de fatiga, con * para p-valores < 0.05, ** en casos < 0.01 y *** < 0.005.

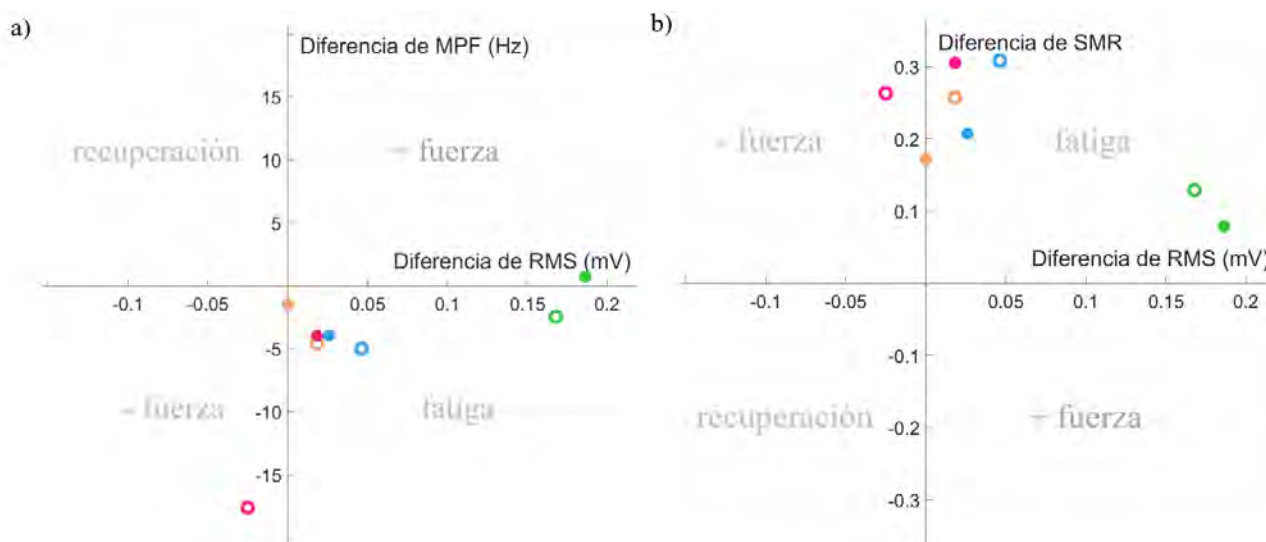


Figura 4. JASA tradicional (a) y nueva propuesta de análisis bivariado considerando el parámetro espectral SMR en el eje de ordenadas (b). Se representan los resultados con la herramienta control (puntos huecos) y para la herramienta prototipo (puntos rellenos) para los músculos bíceps (azul), deltoides (naranja), trapecio (verde) y al flexor radial del carpo (rosa).

Referencias

- [1] P. Chauvet et al. "Laparoscopy and laparoscopic surgery: general principles and instrumentation." In: EMC-Gynecology-Obstetrics 54.2 (2018), 1-17.
- [2] Lobato, R. F., de Adana Belbel, J. C. R., Morales, F. A., Septiem, J. G., Lucas, F. J. M., & Esteban, M. L. (2014). A cost-benefit study comparing open and laparoscopic ventral hernia repair. *Cirugía Española*, 92(8), 553-560.
- [3] R. Berguer, DL Forkey, and WD Smith. "Ergonomic problems associated with laparoscopic surgery." In: *Surgical Endoscopy* 13 (1999), 466-468.
- [4] DAG Reyes, B Tang, and A Cuschieri. "Minimal access surgery (MAS)-related surgeon morbidity syndromes." In: *Surgical Endoscopy and Other Interventional Techniques* 20 (2006), 1-13.
- [5] Luttmann, A., Jäger, M., & Laurig, W. (2000). Electromyographical indication of muscular fatigue in occupational field studies. *International journal of Industrial ergonomics*, 25(6), 645-660.
- [6] H. M. Pace-Bedetti, J. L. Martinez-de-Juan, A. Conejero, and G. Prats-Boluda. "A Surface Electromyogram Evaluation of the Postural Freedom Effects in Laparoscopic Surgery." In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 3143-3146. IEEE, July 2019.
- [7] H. J. Hermens and B. Freriks, "SENIAM 5. Recommendations for Sensors and Sensor Placement Procedures." Roessingh Research and Development. Enschede, the Netherlands, 1997.
- [8] Tanaka, T., Nambu, I., Maruyama, Y., & Wada, Y. (2021). Verification of Normalization Method to Improve Usability and Versatility among Users of Applications that Predict Continuous Motion Using Electromyography. In 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 178-18. Octubre 31 - Noviembre 4, 2021. Virtual Conference.
- [9] Javier Villar Valero. "Time-frequency analysis of fatigue parameters in surface electromyography (sEMG) signals: an ergonomic study in laparoscopic surgery." Doctoral thesis. Universitat Politècnica de València, 2019.
- [10] Kahl, L., Eger, M., & Hofmann, U. G. (2015). Effects of sampling rate on automated fatigue recognition in surface EMG signals. *Current Directions in Biomedical Engineering*, 1(1), 80-84.
- [11] Moslem, B., Khalil, M., Marque, C., Diab, M.O.. «Complexity analysis of the uterine electromyography». In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. IEEE. 2010, 2802-2805.
- [12] Marri, K. and Swaminathan, R. (2016). Classification of muscular nonfatigue and fatigue conditions using surface emg signals and fractal algorithms. In *Dynamic Systems and Control Conference*, volume 50695, page V001T10A002. American Society of Mechanical Engineers.
- [13] Mehran Talebinejad, Adrian DC Chan y Ali Miri. «A Lempel–Ziv complexity measure for muscle fatigue estimation». In: *Journal of Electromyography and Kinesiology* 21.2 (2011), 236-241.

Imágenes Médicas (I)

Miércoles 22 9:30 - 11:00

Paraninfo

Automatic calculation of pelvis morphology from CT images

P. Guijarro Martínez¹, D. Sánchez García², L. Cubero¹, J. De Leon-Luis^{3,4}, J. Pascau^{1,5},

¹ Departamento de Bioingeniería, Universidad Carlos III de Madrid, Madrid, Spain pguijarr@pa.uc3m.es, jpascau@ing.uc3m.es

² Departamento de Radiología, Hospital Universitario Gregorio Marañón, Madrid, Spain

³ Servicio de Obstetricia y Ginecología, Hospital Gregorio Marañón, Madrid, Spain

⁴ Departamento de Salud Pública y Materno-Infantil, Universidad Complutense de Madrid, Madrid, Spain

⁵ Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain

Abstract

Pelvimetry is the study of the pelvis morphology in women for labor planning and medical assessment. This can be achieved by manually annotating pelvic CT images for extracting several measures of interest, which can be both time-consuming and subjective. While machine learning has achieved significant success in 2D landmarking applications, results in pelvic CT images are still limited, particularly with small datasets. This paper presents a two-step approach for detecting 3D landmarks in pelvic CT images. First, a simple CNN coarsely estimates landmark locations, serving as a starting point for further refinement. Then, higher resolution 3D patches and independent neural networks are used to obtain the final position for each landmark. Our model has shown promising results, obtaining an average distance error of 6.71 mm across 7 landmarks. These values allowed the calculation of the morphological measurements, demonstrating a strong correlation with the manual values. The proposed model has shown promising results, offering efficient and accurate predictions of the anatomical landmarks in CT examinations.

1. Introduction

As described in the literature [1, 2], pelvimetry is a common procedure for pelvic analysis in women. It allows assessing the shape and dimensions of the pelvic bone, enabling clinical evaluation and labor planning. Understanding pelvic anatomy and morphology requires calculating several morphometric measurements such as the subpubic angle, the anatomical transverse diameter and the obstetric conjugate diameter. These measurements provide valuable information about the overall size of the pelvis and whether they may limit natural birth delivery. According to size and shape, the female pelvis can be classified into four distinct types: anthropoid, gynecoid, android and platypelloid [3]. The above-mentioned measures could also be related to the Symphysis–Ischial Spine Distance (SID), which has been shown to depend on the sex and height of each patient [4]. The SID is of great importance for determining the type of pelvis in women, as well as its size and shape. To replicate these results and understand the relationship between SID and labor complications, radiologists at Hospital Gregorio Marañón are manually annotating a large CT image dataset, identifying several anatomical landmarks. Automating this process would facilitate their research and provide robust and replicable measurements.

In comparison to the clinical approach [5], Computed Tomography (CT) pelvimetry has emerged as a key

improvement, providing precise and reliable measurements. Analysing high-resolution 3D images offers radiologists a better understanding of the overall structure of the pelvis. However, this process requires manually annotating each CT scan, which can be time-consuming. Furthermore, it demands the knowledge of skilled specialists, heavily relying on their criteria and expertise. Hereby, Machine Learning (ML) algorithms have gained importance in this area, providing accurate positioning of anatomical landmarks and therefore relieving the workload of the professionals. These studies have revolutionized multiple fields of image analysis, achieving great success in tasks such as facial landmark detection in 2D images [6] and anatomical landmark detection in medical datasets [7]. These advances have noticeably increased the efficiency and accuracy of landmark identification. However, when it comes to identifying pelvic landmarks in CT scans, the number of published studies is limited. This can be attributed to the high computational cost associated with 3D images and the unavailability of annotated medical datasets. However, Payer et al [6] predicted landmarks in T1-weighted 3D gradient echo hand MR scans using a U-Net where each landmark was represented by a Gaussian. The results on 85 training sets resulted in an average error of 1.18 mm. Bier et al. [8] detected multiple pelvic X-ray landmarks independently of their viewing direction by utilizing a CNN, which achieved an average error of 5.6 mm. Luo et al. [7] used decision forests followed by morphological image processing techniques on thoraco-abdominopelvic CT scans, achieving an average distance error of 13.7mm.

In this work, we have developed a model for the automatic calculation of pelvis morphology based on the prediction of 3D anatomical landmarks. The detected landmarks allowed calculating morphometric measures of interest, offering objective and quantitative data about the overall dimensions of the pelvis.

2. Materials and Methods

2.1 Subjects

The selected dataset for this study was provided by Hospital General Universitario Gregorio Marañón (HGUGM), containing 87 pelvic CT images in DICOM format from women aged from 20 to 44 years. These images were collected from women who undergone a medical examination previous to birth delivery. All images were carefully selected and annotated by three expert

radiologists, who then classified them into one of the four types of pelvis mentioned. The images were acquired on several CT devices: Ingenuity (Philips), Brilliance 64 (Philips) and Revomaxima (General Electric). All images were pseudo-anonymized to protect personal data.

Type of pelvis	Number of samples
Anthropoid	48
Gynecoid	31
Android	5
Platypelloid	3

Tabla 1. Number of samples for each type of pelvis in women

As shown in Table 1, a significant class imbalance was observed, which is certainly expected given the prevalence of specific pelvic types in women. There were a total of 7 annotated landmarks, corresponding to the subpubic angle (L0, L1, L2), the obstetric conjugate diameter (L3, L4) and the longitudinal diameter (L5, L6), as shown in Figure 1. For instance, the subpubic angle, is formed by the inferior pubic rami and the lower part of the pubic symphysis. The anatomical transverse diameter represents the largest distance of the superior aperture of the pelvis, also known as the pelvic inlet. Finally, the obstetric conjugate diameter, measures the distance from the sacral promontory to the point below the upper inner edge of the symphysis. The importance of these measures relies on understanding the overall size of the pelvis so as to determine its dimensions as they may limit natural birth delivery.

The annotations were initially provided as 2D snapshots on the 3D volume render of each CT. To obtain the corresponding 3D coordinate for each landmark, we manually annotated and saved their positions on 3D Slicer [9], a software platform used for medical image analysis and visualization. In some cases, the provided images did not show the full pelvic bone, as they were cropped at the pubic rami. For these cases, the landmarks of the subpubic angle were strategically positioned at key points due to the absence of bone in those images.

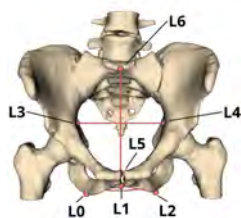


Figure 1. Landmarks location annotated in 3D Slicer.

2.2 Data pre-processing

Voxel size and field of view (FOV) differed among patients depending on the original diagnostic purpose of each scan. To standardize image size, manual cropping was performed in the z-axis using 3D Slicer. A reference point was established at the L5 vertebrae, beyond which all images were cropped for homogenization.

Then, the pelvic bone was segmented to eliminate unwanted structures and improve the overall performance of our model. First of all, a gaussian filter was applied with a standar deviation of 2 mm for reducing image noise. Then, we applied intensity-based segmentation with min and max thresholds of 150 and 1050 Hounsfield Units, corresponding to the pixel values where pelvic bone is observed [10]. The resulting mask was then processed with morphological dilation and erosion techniques and finally applied to the original image.

To address variations in image size and resolution, intensity-based registration was applied as part of the preprocessing. Resolution in the X and Y axis ranged from 0.52 to 0.97 mm, with mean 0.76 mm. For the Z axis resolution ranged from 0.5 to 5 mm, with mean 1.11 mm. After analyzing the sizes and resolutions of the dataset, four images were discarded as the overall resolution was significantly worse compared to the rest of the images, so they were removed from the dataset. The image with the largest size and highest resolution ratio was chosen as the target image for the registration process. This reference image had matrix size (582, 444, 280) pixels and voxel spacing (0.625, 0.625, 1.0) mm. All images were normalized in intensity and aligned to the target image with rigid registration with Mean Squared Error as cost function. This process aimed to align all images without introducing scaling or deformation, preserving the relative positions of the landmarks.

The large size of the resulting volumes did not allow training the models at full resolution in the available GPUs due to memory limitations, so all images were resized by a factor of (3, 3, 2) in x,y and z, respectively, increasing the voxel size and therefore reducing the overall resolution of the images. This resizing factor was also chosen to achieve similar voxel size in each dimension.

2.2 Training

Our proposed landmark prediction model involves a two-step approach based on Convolutional Neural Networks. All tests were run in a workstation equipped with AMD Ryzen 7 5800X 8-Core Processor and 64 GB RAM, along with one Quadro RTX 8000 GPU with 48 GB memory. For training our model the following libraries were installed: keras v2.8.0, numpy v1.22.3, pandas v2.0.0, SimpleITK v2.1.1.2, matplotlib v3.5.2 and TensorFlow v2.8.0.

Our proposed approach consists of two steps. First, a simple CNN, named CNN_LR, takes as input the resized images of size (194, 148, 105) pixels with (1.875, 1.875, 2) mm spacing. In this step, low-resolution images will be used for an initial prediction of the landmarks. The final step involves training a CNN for each landmark, named CNN_HR, using the same architecture on patches of size 100x100x40 pixels around each landmark obtained from the original high-resolution images. This second model focuses on the local volume surrounding each landmark and preserves the original resolution, but fits in the GPU memory due to the decreased image size. During training of the CNN_HR, the generated patches were centered around the ground truth labels of the orig-

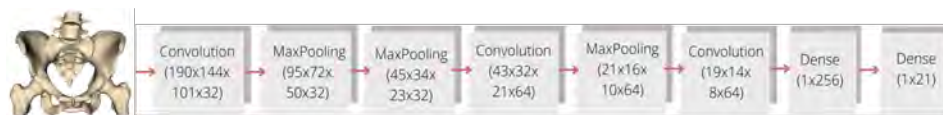


Figura 2. Model architecture.

inal image. To better simulate the inference, where the predicted landmark from the CNN_LR will have some error, we applied a random translation between -10 and 10 pixels in all axis. This translation led to variations in position within the patch, so the training landmarks were not always at the center of the patch, and the network could learn the spatial relationships and true positions of the landmarks. For the test subset, the generated patches were obtained from the predictions of the CNN_LR rather than the real annotations from the original images. To determine the corresponding position of these predictions in the original image, we applied a reverse scaling operation by multiplying the predicted positions by the scaling factors (3, 3, 2). For this subset, random translation was not necessary as the real predictions were not in the center of the generated patch.

In both phases of the pipeline, we used the same model architecture, as shown in Figure 2. This architecture consisted of conventional 3D convolutions and max pooling layers to capture relevant features in the input data. Additionally, we included several dropout layers to regularize the network and reduce overfitting. The loss function was the Mean Squared Error which penalized large errors due to the squaring of the differences between the predicted and true values. For evaluating the results obtained in each stage, the Euclidean distance between the prediction and the real coordinate was calculated for each landmark.

We divided the dataset into train, test and validation sets using a 60-20-20 relation, maintaining proportional class distribution across sets, ensuring representative sampling and enhancing model evaluation reliability. This supports robust CNN training and accurate performance assessment on unseen test data.

Furthermore, data augmentation was applied to deal with limited data and interpatient variability, increasing the training data by a factor of 10. Given the importance of preserving the structural and compositional characteristics of the original data, we implemented random scaling and rotation. First, images were randomly scaled with a probability of 0.5 and a scaling factor from 0.9 to 1.15, which was centered at the image midpoint. In addition, random rotation from -15 to 15 degrees was applied around X and Y axes. The rotation was also centered in the middle of the image. Data augmentation noticeably increased the training data by making minor changes in the orientation and spatial positioning of the pelvic structures.

For the evaluation of the models performance we used the correlation coefficient metric. This measures the strength of the linear relationship between the mor-

phometric measures of interest and their corresponding predictions. It ranges from -1 to 1, where negative values correspond to inverse correlation and values near zero indicate that there is no linear relationship between values.

3. Results and discussion

In the first stage, the mean distance error obtained between the predicted points and the ground truth labels was 9.32 mm with a standar deviation of 4.18 mm and a median of 9.17 mm. The distance errors for each landmark are illustrated in Figure 3. For the final step, the average prediction error for the corresponding landmarks was significantly smaller, achieving a mean value of 6.71 mm, with a standar deviation of 3.28 mm. The median obtained in this stage was 6.58 mm. As observed, the second CNN improves the results decreasing the error for all landmarks. L0 and L2 obtained the largest errors, possibly due to their location within the pubic rami. For cropped images, these coordinates were no longer anatomical points but virtual ones, potentially hindering the model's ability to learn their accurate positioning.

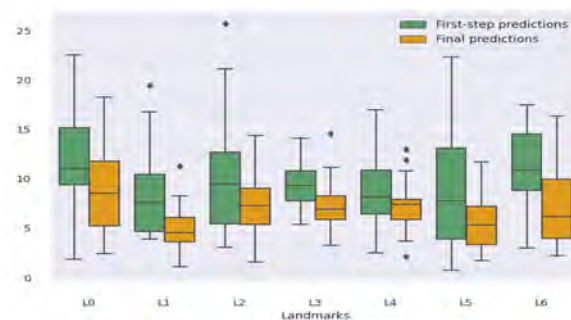


Figure 3. Landmarks location annotated in 3D Slicer.

The main goal of this study was the prediction of anatomical landmarks from pelvic CT images, but also calculate the corresponding measures, thus enabling pelvimetry studies on large datasets without manual labelling. To better analyze the similarity between the predicted and the ground truth measurements, scatter plots are provided to analyze the results achieved in each stage quantitatively. The scatter plots show that the ground truth labels and the predictions are more correlated after the refinement of the landmarks. This is especially noticeable in the subpubic angle, displayed in Figure 4, for which the coarse predictions were highly uncorrelated, with correlation coefficient 0.01, while the linear relationship significantly increased after the refinement, with a final r value of 0.57. Nevertheless, the

calculation of the subpubic angle is highly dependent on the exact positions of the predicted landmarks, as small changes in the coordinates could lead to significant differences in the angle. Due to the fact that L0 and L2 had large errors as observed in the boxplot representation, we also expected higher errors in the subpubic angle calculation.

In the case of the transverse diameter, Figure 5, the correlation coefficient improved from 0.37 to 0.85. This had the most correlated predictions with respect to the real annotations, which matches the observed results in the boxplot. Finally, the correlation coefficient obtained for the obstetric conjugate diameter was improved from 0.35 to 0.58 (Figure 6). In this case, the low correlation could be associated with the large errors obtained for L6 predictions due to the position of this landmark in the vertebrae.

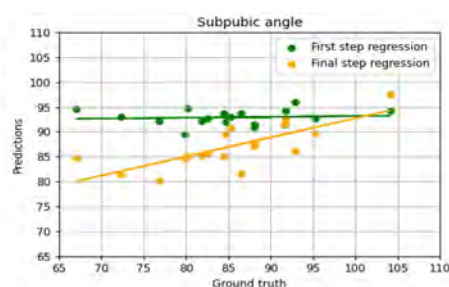


Figure 4. Scatter plot and linear regression for the subpubic angle measurements.

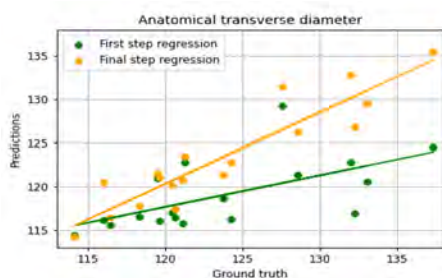


Figure 5. Scatter plot and linear regression for the transverse diameter.

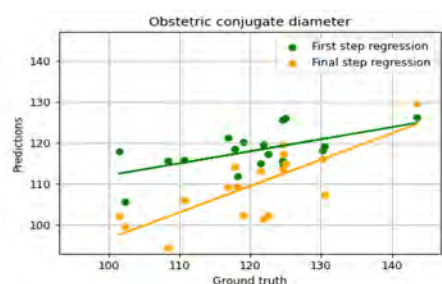


Figure 6. Scatter plot and linear regression for the obstetric conjugate diameter..

4. Conclusions

This study highlighted the potential of machine learning algorithms for improving pelvic landmark detection. The lack of available literature in this area made it difficult to develop an efficient method for predicting 3D landmarks on our limited dataset. Additionally, cropped images difficulted the manual annotation

in landmarks L0 and L2, limiting the learning possibilities of the neural network. However, our model showed promising results, providing efficient and accurate predictions of the anatomical landmarks in CT scans. We have successfully developed a two-step approach that will reduce the workload of the radiologists, serving as a valuable tool for pelvimetry. Future investigations will explore testing in larger datasets and classifying the different types of pelvis based on our predictions.

Acknowledgments

Research supported by projects PI122/00601 and AC20/00102 (Ministerio de Ciencia, Innovación y Universidades, Instituto de Salud Carlos III, Asociación Española Contra el Cáncer and European Regional Development Fund “Una manera de hacer Europa”), project PerPlanRT (under the frame of ERA PerMed), TED2021-129392B-I00 and TED2021-132200B-I00 (MCIN/AEI/10.13039/501100011033 and European Union “NextGenerationEU”/PRTR).

References

- [1] P. Frémondrière *et al.* Which foetal-pelvic variables are useful for predicting caesarean section and instrumental assistance. *Medical Principles and Practice*, 26(4):359–367, 2017.
- [2] Miriam S Lenhard *et al.* Pelvimetry revisited: analyzing cephalopelvic disproportion. *European journal of radiology*, 74(3):e107–e111, 2010.
- [3] S. Kilpatrick *et al.* Normal labor and delivery. *Obstetrics: Normal and problem pregnancies*, 5:303–321, 2007.
- [4] D. Sánchez *et al.* Effects of sex, age and height on symphysis–ischial spine distance measured on a pelvic ct. *Journal of Clinical Medicine*, 11(9):2395, 2022.
- [5] I. Salk *et al.* Pelvimetry by three-dimensional computed tomography in non-pregnant multiparous women who delivered vaginally. *Polish journal of radiology*, 81:219, 2016.
- [6] C. Payer *et al.* Regressing heatmaps for multiple landmark localization using cnns. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17–21, 2016, Proceedings, Part II 19*, pages 230–238. Springer, 2016.
- [7] A. Jimenez-Pastor *et al.* Automated vertebrae localization and identification by decision forests and image-based refinement on real-world ct data. *La radiologia medica*, 125:48–56, 2020.
- [8] B. Bier *et al.* Learning to detect anatomical landmarks of the pelvis in x-rays from arbitrary views. *International journal of computer assisted radiology and surgery*, 14:1463–1473, 2019.
- [9] A. Fedorov *et al.* 3d slicer as an image computing platform for the quantitative imaging network. *Magnetic Resonance Imaging*, 30(9):1323–1341, Nov 2012.
- [10] Tami D DenOtter *et al.* Hounsfield unit. 2019.

Deep Learning methodologies for direct image reconstruction and integrated attenuation correction in brain PET/MRI

P. Díaz Serrano¹, J. E. Ortuño Fisac^{2,3}, J. López Santiago¹, F. Panetsos Petrova^{4,5}, G. Kontaxakis³

¹ Universidad Carlos III de Madrid (UC3M), Madrid, España; pameladiaz2318@gmail.com, jlopez@tsc.uc3m.es

² CIBER de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, Spain

³ ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain, {je.ortuno, g.kontaxakis}@upm.es

⁴ Neurocomputing and Neurorobotics Research Group, Faculty of Biology and Faculty of Optics, Universidad Complutense de Madrid (UCM), Madrid, Spain; fivos@ucm.es

⁵ Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain

Abstract

This study proposes the ATDeepPET model, a novel deep learning architecture crafted specifically for advancing positron emission tomography (PET) image reconstruction in PET/MR scanners. By incorporating magnetic resonance (MR) images into its learning process, ATDeepPET addresses the persistent challenges associated with attenuation effects in PET/MR scanners, eliminating the need for simulated transmission scans. ATDeepPET's performance is assessed alongside the deep learning model DeepPET, as well as established methods such as FBP, ML-EM, and ML-EMR for comparison. The findings reveal noteworthy achievements since ATDeepPET accomplishes competitive image quality compared to FBP, ML-EM, and ML-EMR when applied to brain phantoms while also demonstrating a reduction in reconstruction times. Nevertheless, when dealing with real PET images, ATDeepPET does exhibit some performance variability, underscoring the increased complexity of real-set scenarios and the importance of employing diverse datasets to enhance its robustness. Moreover, ATDeepPET, despite inherent limitations, including heightened memory requirements and sensitivity to dataset variations, presents a promising path forward for PET image reconstruction. Its hallmark traits include exceptional execution speed, liberation from the prerequisite of prior physics knowledge, and the prospect of obviating the need for an additional CT scan for attenuation correction. These attributes hold transformative potential in terms of enhancing diagnostic precision and curtailing patient radiation exposure.

1. Introduction

Positron emission tomography (PET) is a valuable tool in medical imaging, providing insights into an organism's metabolic activity [1], particularly benefiting cancer diagnosis and treatment [2]. During PET image acquisition, data is recorded as sinograms, requiring reconstruction to produce interpretable images. However, solving the inverse problem for this imaging modality is complex, leading to reconstructed images that are highly susceptible to noise.

Analytical methods, like Filtered Backprojection (FBP), lack noise and physics modelling [3]. Conversely, iterative methods, such as Maximum Likelihood Expectation Maximization (ML-EM), offer the capability to model the Poisson statistical noise present in the acquired data and incorporate more intricate physical models [3]. Furthermore, these iterative methods allow

for the inclusion of regularization terms during the iterative process, considering prior knowledge.

Nonetheless, it is essential to note that iterative methods come with significant computational demands, particularly as the complexity and accuracy of the employed physical model increase. Also, PET imaging suffers from attenuation artefacts, which are caused due to varying tissue photon absorption coefficients, leading to a distorted tracer distribution depiction.

Current attenuation correction methods have limitations, such as the practical challenges and regulatory complexities associated with gamma sources [4]. Alternatively, using attenuation maps from X-ray computerized tomography (CT) scans exposes patients to additional ionizing radiation.

As Deep Learning (DL) methodologies continue to expand, novel approaches have emerged within the field. These methods possess the capability to grasp the underlying physics of PET data generation, addressing noise and attenuation effects directly through the training data [5]. DL methods find applications in various aspects of image reconstruction, often replacing specific components, such as regularization, within conventional methods. However, this work focuses on DL architectures that aim for end-to-end mapping, involving direct reconstruction by deep networks from raw sinogram data to a reconstructed image. This approach requires assimilating all aspects of the image reconstruction process, including physics, imaging models, and statistics, demanding a substantial amount of training data. With this objective, encoders-decoders architectures are of relevance, capturing essential data features in the encoder stage and transforming them into high-quality images in the decoder stage [5].

Traditional practice involves the use of CT images for attenuation correction in PET scans. However, there is a growing interest in replacing CT with magnetic resonance (MR) images due to the advantages offered by PET/MR systems, such as improved soft tissue contrast and reduced ionizing radiation [6]. However, since MR images do not inherently provide attenuation information, generating attenuation maps from PET/MR scanners is challenging. Various methodologies, including machine learning, have been proposed to address this issue, such as creating synthetic CT (sCT) images from MR data [7].

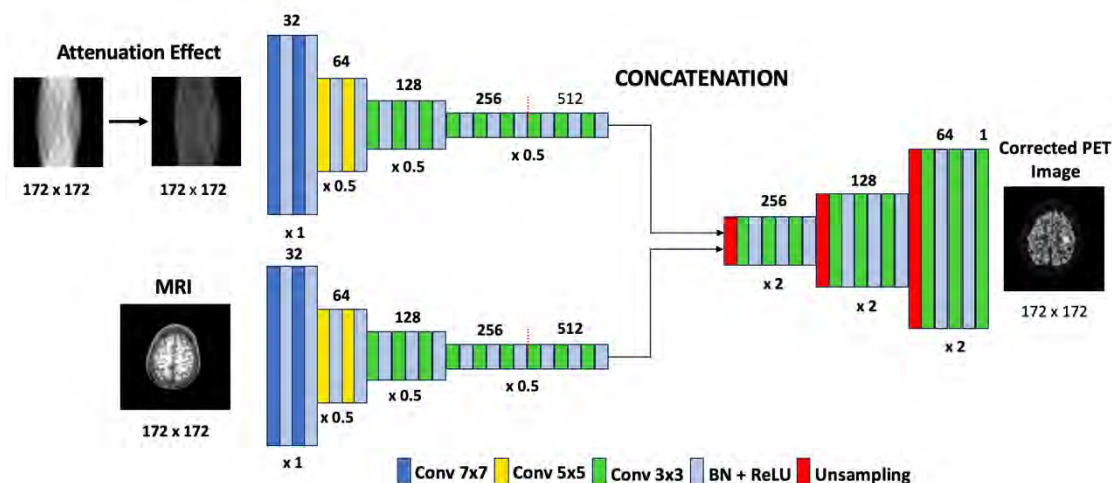


Figure 1. Network ATTDDeepPET image reconstruction method.

This study introduces an alternative approach within the DeepPET [8] architecture. Instead of creating intermediate sCT images, this proposed method directly learns from MR images to perform both attenuation correction and PET image reconstruction. This approach eliminates the need to convert MR images to sCT, simplifying the process by enabling the network to learn the complex relationship between attenuation correction and image reconstruction directly from MR inputs.

2. Methodology

2.1. Deep Learning methods

A modified version of the DeepPET method [9], was implemented. Several improvements were made to the original architecture, including reducing the total number of layers from 31 to 22, which helped lower the memory requirements. Additionally, the optimizer was switched from Adam stochastic gradient descent to RMSProp from Keras¹. The loss function remained as the mean squared error (MSE). The resulting architecture comprised an encoding phase where sinogram input is compressed into feature maps through convolution layers, and a decoding phase introducing up-sampling layers to produce the final reconstructed image. After each convolution in the encoding stage, a ReLU activation function and batch normalization are applied.

Then, a second architecture was implemented, where a second encoder was added to incorporate co-registered MR images (Fig. 1), introducing anatomical details to the reconstruction and attenuation correction process. We named this architecture ATTDDeepPET. For both networks, sinograms with attenuation effects were introduced to evaluate the ability of the networks to reconstruct and correct the attenuation effect.

Furthermore, in order to reduce the memory demands of the method, a customized batch generator was employed. This batch generator dynamically loads GPU input data from the disk as required, operating in a volatile and adaptable manner.

2.2. Traditional methods

To evaluate the efficacy of the DL models in comparison to conventional approaches, a comparative analysis involving three key methodologies was implemented: FBP, ML-EM and its regularized variant (ML-EMR) were implemented in a code using the ASTRA Toolbox library.

2.3. Datasets

To test the performance of the network, the following datasets were used, all conformed by PET, CT, and MR co-registered images. The CT sinogram was used to generate the attenuation effect in the input PET sinograms, which were obtained using the ASTRA toolbox with 2D ideal geometry and parallel projectors. Lastly, random noise following a Poisson distribution was introduced into the sinograms to replicate the statistical characteristics of PET data acquisition. Importantly, this simulation does not incorporate attenuation or scattering effects:

- Brain phantom dataset consisting of 64596 2D axial slice images extracted from 20 brain phantom 3D volumes. These volumes were originally designed for MRI simulation and were subsequently adapted for PET imaging. This dataset is publicly accessible [10]. To broaden the analysis, some of the images were induced with artificial lesions.
- 15725 real PET images, including 2D axial, sagittal, and coronal slices, derived from the 3D volumes of 37 healthy patients [11].

For all datasets, 80% of the images are destined for training, 10% for validation, and 10% for testing.

¹ <https://keras.io/>

3. Results and discussion

The two distinct architectures were trained and tested with the same datasets. The evaluation phase involved the reconstruction of the test image set through the three distinct methodologies: FBP, ML-EM, ML-EMR, DeepPET, and the ATTDDeepPET model. The resultant reconstructed images were subsequently subjected to a comparative analysis against a reference ground truth image with the structural similarity metric (SSIM), peak signal-to-noise ratio (PSNR) and normalized root mean square error (NRMSE) [12].

3.1. Brain phantom dataset

Initially, the DeepPET and ATTDDeepPET networks were trained using the training set of brain phantoms of size 172x172. The learning rate was set to 10⁻⁴, and the batch size to 32. The GPU used was PNY QUADRO RTX-A6000 48GB GDDR6.

The mean error of images in the test set reconstructed with the different methods can be seen in Table 1. The results demonstrate that ATTDDeepPET consistently outperforms other methods across all three metrics. DeepPET also exhibits strong performance, closely trailing behind ML-EMR and ML-EM, surpassing the performance of FBP, which consistently score lower across all metrics.

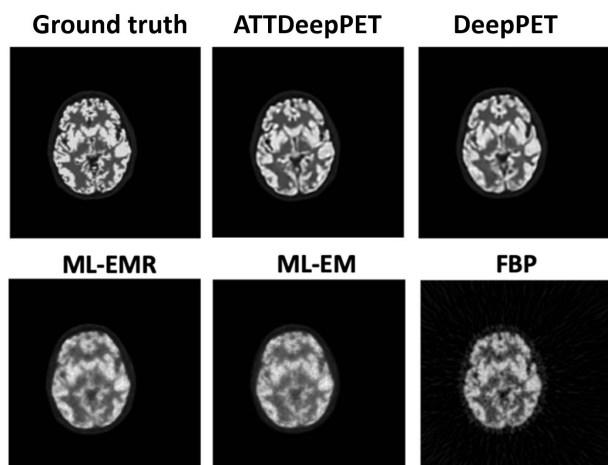


Figure 2. Phantom reconstructed axial slices employing FBP, ML-EM, ML-EMR, DeepPET and ATTDDeepPET

A Tukey’s Honestly Significant Difference (HSD) test was conducted, a statistical analysis that provided a nuanced understanding of the performance dynamics. In the context of brain phantoms, ATTDDeepPET emerges as the standout performer, showcasing significantly higher SSIM and PSNR values, with statistical significance supported by the rejection of the null hypothesis ($p < 0.05$). Furthermore, ATTDDeepPET demonstrates the lowest nRMSE, again with statistical significance compared to the other methods. Conversely, while DeepPET delivers commendable results, surpassing the performance of ML-EM and FBP methods, it’s important to note that it does not exhibit a statistically significant improvement over the ML-EMR method.

Method	SSIM	PSNR	NRMSE
FBP	0,5461	26,4315	0,2554
ML-EM	0,9545	29,0161	0,2198
ML-EM regularized	0,9634	29,5834	0,2041
DeepPET	0,9677	30,2062	0,2082
ATTDeepPET	0,9753	31,5333	0,1811

Table 1. Average error for brain phantom dataset reconstruction.

3.2. Real brain images dataset

Subsequently, DeepPET and ATTDDeepPET architectures were trained using the real PET images dataset. The learning rate was set to 10⁻⁴, and the batch size to 32. The same GPU was used. Obtained quantitative metrics SSIM, PSNR, and NRMSE were computed between each of the reconstructed images and the ground truth. The ensuing averages of these metrics are compiled in Table 2.

Both ATTDDeepPET and DeepPET showed lower performance compared to ML-EMR and ML-EM when trained using real data. However, upon examining the reconstructed images (Fig. 3), it becomes apparent that the deep learning models produce smoother images with reduced noise compared to the ML-EMR and ML-EM approaches. This suggests that the reconstructions may bear a closer resemblance to the ground truth. However, there is a noticeable reduction in the definition of cerebral gyri in these reconstructions, which results in an overall decrease in image fidelity to the ground truth.

Method	SSIM	PSNR	NRMSE
FBP	0,3966	25,1741	0,2213
ML-EM	0,8939	30,3253	0,1358
ML-EM regularized	0,9078	30,5839	0,1211
DeepPET	0,8030	26,8114	0,1707
ATTDeepPET	0,8776	27,1281	0,1648

Table 2. Average error for the brain phantom dataset containing lesion reconstruction.

To further investigate this phenomenon, a region of interest (Fig. 4) was selected. It becomes apparent that ATTDDeepPET struggles to precisely reconstruct the defining boundaries of the cerebral cortex. Measurements of SSIM, PSNR, and NRMSE were computed by comparing the encoder-decoder models with the ML-EMR, which showed the best accuracy (Table 3).

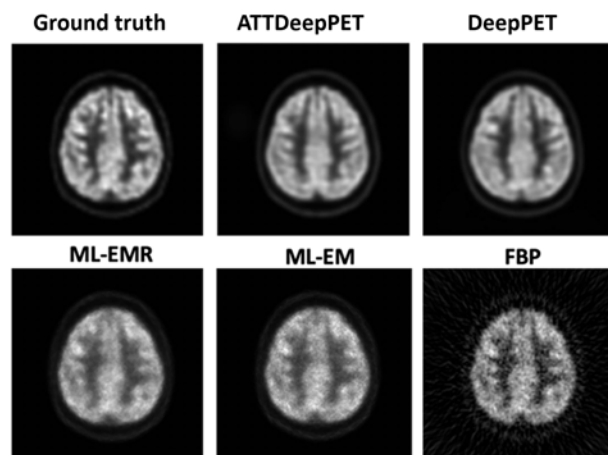


Figure 3. Real PET reconstructed axial slices employing FBP, ML-EM, ML-EMR, DeepPET and ATTDDeepPET.

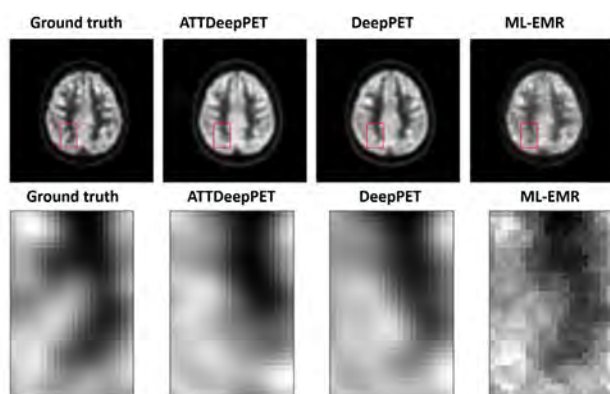


Figure 4. Cerebral gyri comparison.

Additionally, the analysis revealed the presence of notable outliers characterized by elevated nRMSE values. A closer examination of these outlier images uncovered a consistent pattern: they predominantly featured images with minimal brain content, particularly in the axial plane. These outlier images were primarily composed of slices located at the extremities of the head structure.

Method	SSIM	PSNR	NRMSE
ML-EMR	0,7539	17,4505	0,1350
DeepPET	0,5609	15,8724	0,1896
ATTDeepPET	0,5002	15,6831	0,1797

Table 3. ROI’s SSIM, PSNR and NRMSE averages.

3.3. Reconstruction times

The temporal efficiency of each reconstruction method was quantified by measuring and calculating the average execution times across all images. The encoder-decoder architectures, DeepPET and ATTDeepPET, demonstrated significant reductions in reconstruction time, highlighting their notable computational efficiency. DeepPET exhibited a remarkable processing speed. Compared to FBP, DeepPET achieved a substantial speed enhancement of approximately 1.72 times. In comparison to ML-EM and ML-EMR, DeepPET demonstrated acceleration rates of approximately 735.13 times and 1299.23 times, respectively. Similarly, ATTDeepPET presented acceleration rates of approximately 1.20 times compared to FBP, 511.47 times compared to ML-EM, and 906.73 times compared to ML-EMR.

4. Conclusions

The results of this study emphasize the effectiveness of direct image reconstruction methods utilizing DL. These methods demonstrate the capability of performing attenuation correction directly from sinogram data, eliminating the need for synthetic CT images in favor of MRI. This transition is advantageous as it reduces patient exposure to ionizing radiation and significantly reduces reconstruction times, potentially enabling real-time studies. However, while these models perform well with simulated brain data, their performance diminishes when applied to real-world image datasets, highlighting the challenges posed by authentic data. Enhancing the robustness of these models for real-world scenarios should be a priority. On the downside, the computational

complexity of CNNs requires substantial resources and time for network training. These models also rely heavily on extensive and diverse training datasets, limiting their applicability in scenarios with limited or suboptimal data. Traditional methods, grounded in established mathematical principles, offer stability and predictability across various contexts. However, the opacity of modern techniques using encoder-decoder raises interpretability concerns. Furthermore, there is a limitation in terms of model generalization. Models trained specifically with brain data may perform poorly when applied to other anatomical regions. A promising direction could involve hybrid approaches, merging deep learning with traditional methods. These hybrid methods, including unfolded techniques for iterative image reconstruction or adaptive learning of frequency filters, aim to balance the strengths of both paradigms and might serve to overcome the limitations inherent in singular methodologies.

Acknowledgments

This work has been partially funded by the project S2022/BMD-7236 MINA-CM of the Autonomous Community of Madrid and project IEL.GKA.100 DIANIA of FUNDETEL-UPM.

References

- [1] P. Suetens, “Nuclear medicine imaging,” in *Fundamentals of Medical Imaging*, 2nd ed. Cambridge University Press, 2009, pp. 105–127. doi: 10.1017/CBO9780511596803.006.
- [2] D. Delbeke, “Oncological applications of FDG PET imaging: Brain tumors, colorectal cancer lymphoma and melanoma,” *J. Nucl. Med.*, 40(4): 591–603, 1999.
- [3] D. W. Wilson and B. M. W. Tsui, “Noise properties of filtered-backprojection and ML-EM reconstructed emission tomographic images,” *IEEE Trans. Nucl. Sci.*, 40(4): 1198–1203, 1993, doi: 10.1109/23.256736.
- [4] P. Kinahan, B. Hasegawa, and T. Beyer, “X-ray-based attenuation correction for positron emission tomography/computed tomography scanners,” *Semin. Nucl. Med.*, 33(3): 166–179, 2003. doi: 10.1053/snuc.2003.127307.
- [5] A. J. Reader, G. Corda, A. Mehranian, et al, “Deep Learning for PET Image Reconstruction,” *IEEE Trans. Rad. Plasma Med. Sci.*, 5(1): 1–25, Aug. 2020, doi: 10.1109/trpms.2020.3014786.
- [6] T. Wang et al., “Machine learning in quantitative PET: A review of attenuation correction and low-count image reconstruction methods,” *Phys. Med.*, 76: 294–306, 2020. doi: https://doi.org/10.1016/j.ejmp.2020.07.028.
- [7] F. Liu, H. Jang, R. Kijowski, T. Bradshaw, and A. B. McMillan, “Deep learning MR imaging-based attenuation correction for PET/MR imaging,” *Radiology*, 286(2): 676–684, 2018. doi: 10.1148/radiol.2017170700
- [8] I. Häggström, C. R. Schmidlein, G. Campanella, and T. J. Fuchs, “DeepPET: A deep encoder–decoder network for directly solving the PET image reconstruction inverse problem,” *Med. Imag. Anal.*, 54: 253–262, 2019, doi: 10.1016/j.media.2019.03.013.
- [9] N. Rufó Rafael de la Cruz, J. Ortuño Fisac, and G. Kontaxakis, “Deep learning methodologies for brain image reconstruction in positron emission tomography,” in *Actas del XL Congreso Anual de la Soc. Esp. Ing. Biomédica*, Valladolid, Nov. 2022, pp. 337–340.
- [10] C.A. Coccosco, V. Kollokian, R.K.-S. Kwan, and A.C. Evans, “BrainWeb: Online Interface to a 3D MRI Simulated Brain Database,” *Neuroimage*, 5(4): S425, 1997.
- [11] I. Mérida et al., “CERMEP-IDB-MRXFDG: a database of 37 normal adult human brain [18F] FDG PET, T1 and FLAIR MRI, and CT images available for research,” *EJNMMI Research*, 11(1): 91, 2021, doi: 10.1186/s13550-021-00830-6.
- [12] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, “Image Quality Assessment: From Error Visibility to Structural Similarity,” *IEEE Trans. Imag. Proc.*, 13(4): 600–612, April 2004, doi: 10.1109/TIP.2003.81

Análisis cuantitativo del corazón en imágenes 4DCT en pacientes con taquicardia ventricular recurrente o refractaria

I. Hernández-Abad¹, A.P. Seiffert¹, E.J. Gómez^{1,2}, E. Cabello Murillo³, G. Pozo-Rodríguez³, P. Sánchez-González^{1,2}

¹ Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España;

ignacio.hernandez.abad@alumnos.upm.es, {ap.seiffert, enriquejavier.gomez, p.sanchez}@upm.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, España

³ Servicio de Radiofísica Hospitalaria, Hospital Universitario 12 de Octubre, Madrid, España; {eduardo.cabello, gustavo.pozo}@salud.madrid.org

Resumen

La taquicardia ventricular es un tipo de arritmia causado por señales eléctricas anormales que impiden el correcto funcionamiento de los ventrículos. Uno de los tratamientos más habituales para tratar esta patología consiste en realizar una ablación cardíaca. Cuando no se corrige la taquicardia ventricular mediante la acción de fármacos o la ablación por radiofrecuencia, a esta se considera refractaria o recurrente. La ablación cardíaca por radioterapia consiste en depositar una dosis de radiación de aproximadamente 25 Gy en una única sesión de radioterapia en la región arritmogénica previamente delimitada. En este trabajo se estudia imágenes de pacientes con taquicardia ventricular tratados con ablación cardíaca mediante radioterapia. Se desarrolla una metodología para la extracción de 107 características cuantificables de las imágenes 4DCT de los pacientes con cardiopatías, para posteriormente observar si existen diferencias estadísticamente significativas con las extraídas a partir de una región de interés semejante de un grupo control. Para segmentar la región de interés en los pacientes sin cardiopatías se implementa un algoritmo que realiza dicha segmentación mediante el empleo de las estructuras anatómicas de los pacientes tratados con ablación cardíaca como referencia. Para el análisis estadístico se realiza una doble reducción de dimensionalidad para obtener las características estadísticamente significativas como potenciales biomarcadores. Tras el análisis de las regiones de interés distintas y del corazón completo, las características que muestran mayor potencial como biomarcador son las pertenecientes a la matriz de coocurrencia de niveles de grises.

1. Introducción

Un corazón con un ritmo cardíaco caótico no permite el adecuado llenado y/o vaciado de las cámaras del corazón, lo que puede causar como consecuencia una insuficiencia en la distribución de la sangre en los pulmones o en el resto de los órganos del cuerpo [1]. Se considera taquicardia ventricular (TV) cuando se halla ≥ 3 latidos ventriculares consecutivos a una frecuencia ≥ 120 latidos por minuto [2]. La TV se presenta en un ECG con repeticiones del complejo QRS muy próximas entre sí, lo que indica que los ventrículos se están contrayendo de forma continua y no se realiza el correcto llenado de estos.

La TV está causada por una interrupción en los impulsos eléctricos normales que controlan el bombeo rítmico del corazón. Uno de los motivos por los que se producen dichas interrupciones puede estar relacionado con anomalías

del corazón que causan un proceso de cicatrización del tejido cardíaco. Esta anomalía es conocida como enfermedad cardíaca estructural y la causa más común es un ataque cardíaco previo [2]. Otra de las causas más comunes es debido a una enfermedad de las arterias coronarias que causa un flujo deficiente hacia el músculo cardíaco [3].

Se considera que la TV en un paciente es recurrente o refractaria cuando esta no se corrige mediante la desfibrilación. Un procedimiento para el tratamiento de esta patología es la ablación con catéter. Consiste en introducir un catéter en el corazón a través de las venas de alguna de las extremidades o del cuello, según la región del corazón a tratar, y producir tejido cicatrizante que impida la propagación eléctrica en ciertas zonas mediante la aplicación de frío extremo (crioablación) o energía de radiofrecuencia a través de la punta del catéter [2, 4].

Una alternativa terapéutica es la ablación cardíaca mediante radioterapia externa. Es una técnica cuya práctica clínica se remonta a 2017 [5] y para la que se requiere una tecnología específica, ya que el objetivo que persigue es depositar en el área arritmogénica una dosis de radiación muy elevada (en torno a 25 Gy). Las ventajas que tiene este procedimiento respecto al previamente descrito es que su efecto no se reduce únicamente a la superficie de la región de interés, sino que puede llegar a niveles más profundos. Esto tiene como consecuencia el aumento de la probabilidad de éxito, además de que se reduce la probabilidad de ciertas complicaciones con las infecciones al no ser invasiva y evita someter al paciente a nuevas intervenciones que pueden durar varias horas. Los resultados de esta técnica son prometedores hasta el momento en cuanto a supervivencia (superior a los 2 años en muchos casos), aunque existe un déficit de resultados de toxicidad tardía tanto por el estado de los pacientes como por el número todavía pequeño de los que se someten a este tipo de tratamiento [5–7].

El objetivo de este estudio consiste en el desarrollo de una metodología para la extracción de 107 características cuantitativas de las ROI de las imágenes 4DCT tanto de los pacientes con cardiopatías como de los sujetos control. Se estudian las diferencias estadísticamente significativas entre ambos grupos para identificar potenciales biomarcadores.

2. Materiales y Métodos

2.1. Materiales

Un total de 3 pacientes (todos varones y con una edad promedio de $76,5 \pm 3,5$ años) diagnosticados con TV recurrente o refractaria se han incluido en el estudio. Sus correspondientes imágenes 4DCT disponen de una segmentación previa de la ROI realizada por el servicio de Oncología Radioterápica en colaboración con el servicio de Cardiología del Hospital Universitario 12 de Octubre de Madrid. A su vez, se dispone de un total de 50 pacientes control que se dividen en dos grupos en función de sus volúmenes cardíacos. El grupo de menor tamaño dispone de los pacientes cuyo tamaño está comprendido entre los 400 y 600 cm³, siendo un total de 26 pacientes (17 mujeres y 9 varones y con una edad promedio de $65,6 \pm 11,6$ años). Por otro lado, el grupo de mayor tamaño dispone de corazones comprendidos entre los 1000 y 1600 cm³ y suman un total de 27 pacientes (25 varones y 2 mujeres y con una edad promedio de $75,2 \pm 8,9$ años). A su vez, se hace una distinción en función de la posición anatómica de las ROI, dependiendo de si esta se sitúa próxima a la arteria coronaria, a la circunfleja o al estómago. Todos los sujetos disponen de una segmentación del corazón realizada por el personal del Servicio de Oncología Radioterápica del Hospital Universitario 12 de Octubre. Se ha empleado el mismo protocolo de imágenes 4DCT para todos los pacientes de la base de datos, empleando como imagen el promediado de las 10 fases respiratorias. Según la Comisión de Investigación del Instituto de Investigación Hospital 12 de Octubre (i+12) el estudio no requiere la aprobación por el comité de ética del hospital.

Se emplea el entorno de desarrollo *Microsoft Visual Studio 2022*, que soporta proyectos escritos en lenguaje Python (se ha recurrido a la versión *Python 3.9*). Para la extracción de las características cuantitativas se emplea *Pyradiomics* [8] y para el análisis estadístico se emplean *pandas* [9] y *seaborn* [10].

2.2. Métodos

La metodología de este trabajo se divide en tres pasos: (1) segmentación de las ROI, (2) extracción de características, y (3) análisis estadístico.

Segmentación de las ROI. Para la extracción de las características cuantitativas de las imágenes 4DCT anonimizadas de los pacientes con cardiopatía se emplea la segmentación denominada *Planning Target Volume (PTV)*, realizada para la planificación de la radioterapia como región de interés (ROI). Para la realización de dicha segmentación, se emplea un mapa cardíaco por impedancias. Dicha técnica consiste en medir el gradiente entre pares de electrodos y posteriormente calcular las correspondientes interpolaciones para obtener un mapa electrofisiológico. A partir de dicho mapa se puede observar la región arritmogénica y definir en consecuencia el PTV. Se carga las imágenes 4DCT y sus estructuras asociadas a partir de los ficheros DICOM de los pacientes. En el caso de los pacientes con cardiopatías, la segmentación de la ROI se ha realizado previamente por el personal del servicio de Cardiología. En el caso de los sujetos control, se desarrolla un algoritmo de

segmentación que emplea las estructuras de los corazones tanto de estos como de los pacientes tratados con radioablación para adaptar las ROI de estos últimos a la anatomía cardíaca de los pacientes control. Para ello, se desarrolla un algoritmo que, en primer lugar, asigna unas coordenadas a los vóxeles pertenecientes a la ROI de los pacientes con cardiopatía en función de la posición que ocupan estos dentro de la segmentación del corazón, como se observa en la *Figura 1*. Posteriormente, calcula para cada vóxel la posición correspondiente en el corazón del sujeto control. Finalmente, se ejecuta un algoritmo de rellenado de huecos para obtener una única ROI sin espacios sin segmentar en su interior como se puede observar en la *Figura 2*.

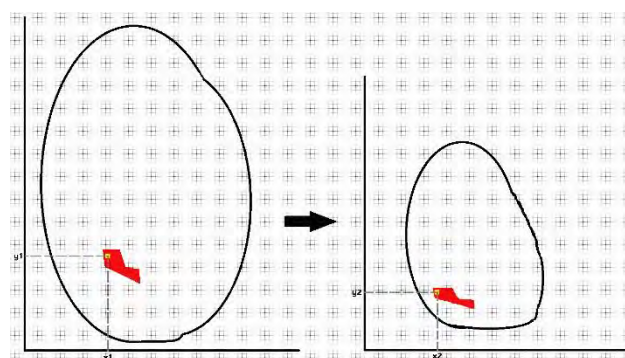


Figura 1. Ilustración esquemática del funcionamiento del algoritmo de transformación de la ROI.

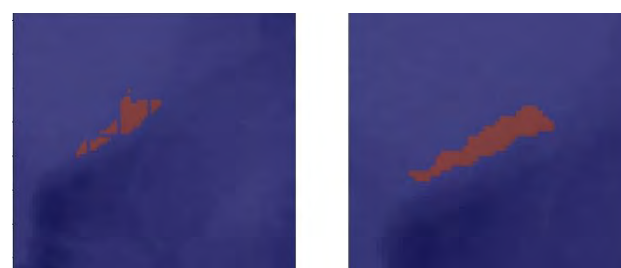


Figura 2. ROI segmentada en un sujeto control antes (izquierda) y después (derecha) de aplicar el algoritmo de rellenado de huecos.

Extracción de características. Se extraen un total de 107 características cuantitativas de las imágenes de cada uno de los pacientes de la base de datos. Estas características se agrupan en 7 clases distintas y otorgan información sobre la forma, el histograma y la textura del tejido de la ROI. En primer lugar, está las clases de las características de primer orden y de forma, que dan información sobre el histograma y la distribución espacial de la ROI respectivamente. Las otras 5 clases de características corresponden a las que se extraen de 5 matrices calculadas a partir de la propia imagen y que dan información de las intensidades de los vóxeles y su distribución espacial de forma conjunta, describiendo así la textura del tejido. Estas 5 matrices son: de coocurrencia (GLCM), de longitud de recorrido (GLRLM), de tamaño de zona (GLSZM), de diferencia de tonos (NGTDM) y de dependencia (GLDM). La descripción detallada de las características extraídas individuales está disponible en el sitio web de *Pyradiomics* (<https://pyradiomics.readthedocs.io/en/latest/>). En la configuración previa a la extracción, se han asignado contenedores de ancho fijo de 25 niveles de intensidad (*binWidth*

= 25) y se realiza un remuestreo para obtener vóxeles isotrópicos con dimensiones (1 1 1) a partir de una interpolación de los vóxeles con tamaño original (1 1 2) siguiendo las indicaciones de la *Image Biomarker Standardisation Initiative (IBSI)* [11].

Análisis estadístico. Se realizan dos reducciones de dimensionalidad consecutivas para disminuir el número de características que puedan emplearse como potenciales biomarcadores. En primer lugar, se calcula el *Z-score* de las características de los sujetos con cardiopatía, empleando las correspondientes características de los sujetos control como distribución normal. Se realiza el cálculo del *Z-score* de cada característica para que todas ellas estén representadas en las mismas unidades y de tal modo, permitir una mejor interpretación de los resultados. En esta primera reducción se seleccionan las características que cumplan un $p < 0,01$. En la segunda reducción se calcula una matriz de correlación a partir de las características previamente seleccionadas, donde cada elemento corresponde al coeficiente de correlación de Pearson para cada par de características. Se considera que un par de características están correlacionadas cuando el valor absoluto del coeficiente supera el 0,7. A partir de dicha matriz de correlación, se seleccionan en primer lugar las características que no presenten correlación con ninguna otra de las previamente seleccionadas y posteriormente se seleccionan en orden ascendente según el número de características que tengan correlacionadas. Se debe tener en cuenta que la diagonal principal de la matriz muestra la autocorrelación y por lo tanto su valor es siempre máximo. De este modo, se consigue obtener un número reducido de características estadísticamente significativas respecto a los sujetos control y sin información redundante. Se realiza un análisis estadístico para cada una de las tres posibles regiones de las ROI, una por cada paciente de la ablación cardíaca, a estudiar y otro para los corazones completos para poder comparar los resultados.

3. Resultados y Discusión

Los resultados del algoritmo de segmentación de la ROI en los sujetos control se ejemplifican en la siguiente figura. Como se puede observar en la *Figura 3*, la ROI calculada por el algoritmo muestra similitudes en la forma y la localización a la segmentación original, aunque se aprecian diferencias debidas a la adaptación a la anatomía propia del paciente.

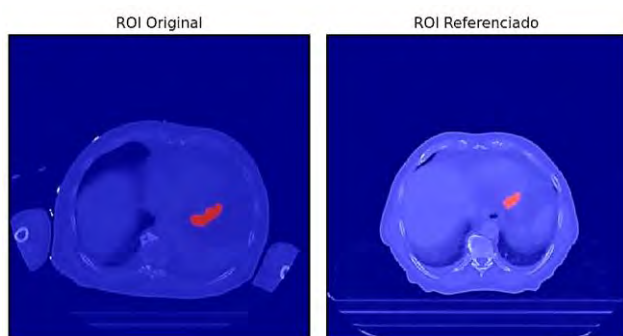


Figura 3. Cortes axiales de imágenes CT y ROI superpuesta de un paciente tratado con radioablación y de un paciente control con una ROI referenciada.

En cuanto a los resultados del análisis estadístico, se diferencia según la región anatómica donde se encuentra la ROI. En el análisis de la región próxima a la arteria coronaria, tras ambas reducciones de dimensionalidad, se seleccionan 4 características como potenciales biomarcadores, correspondiendo a la forma, la GLCM y la GLSZM (Tabla 1). En cuanto a la región próxima a la arteria circunfleja, 3 características han sido seleccionadas como potenciales biomarcadores, las cuales pertenecen a la clase de características de primer orden y GLCM (Tabla 2). En la región próxima al estómago no se selecciona ninguna característica debido a que, tras la primera reducción, ninguna es estadísticamente significativa. Finalmente se realiza un análisis empleando el corazón completo como ROI y se obtienen 7 potenciales biomarcadores, de los cuales 1 pertenece a la clase de forma, 2 a la GLCM, 1 a la GLRLM, 1 a la GLSZM y 1 a la NGTDM (Tabla 3). Se observa en la *Figura 4* las matrices de correlación de las ROIs próximas a las arterias y en la *Figura 5* la matriz correspondiente al análisis de los corazones completos.

Característica	Z-score
Diámetro máximo 2D (plano coronal)	3.14
GLSZM GLN	2.90
GLSZM SZN	3.12

Tabla 1. Valores de Z-score de las características seleccionadas de la región próxima a la arteria coronaria.

Característica	Z-score
Oblicuidad	2.58
GLCM IDMN	2.70
GLCM MCC	-3.97

Tabla 2. Valores de Z-score de las características seleccionadas de la región próxima a la arteria circunfleja.

Característica	Z-score
Esfericidad	-3.48
GLCM IDM	-2.80
GLCM MCC	-2.75
GLRLM RE	-3.55
GLSZM GLN	3.87
NGTDM Complejidad	3.58

Tabla 3. Valores de Z-score de las características seleccionadas de todo el corazón.

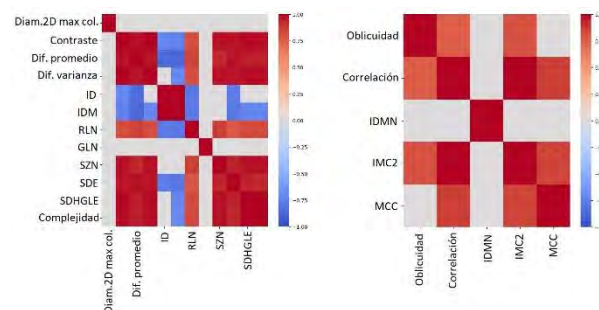


Figura 4. Matrices de correlación de las regiones próximas a la arteria coronaria (izquierda) y circunfleja (derecha).

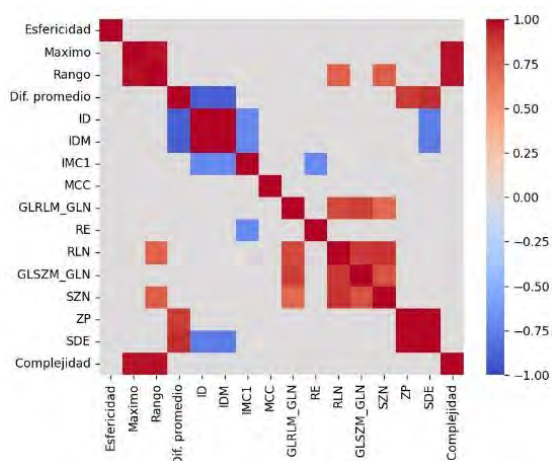


Figura 5. Matriz de correlación del corazón completo.

Como se puede observar, tanto en los análisis de las regiones próximas a ambas arterias como en el del corazón completo, existen numerosos potenciales biomarcadores que pertenecen a la clase GLCM. Esto puede ser indicativo del potencial que tiene esta clase de características para ser empleadas como biomarcadores en modelos de predicción. Por otro lado, el hecho de que el análisis de la región próxima al estómago no permita seleccionar ningún potencial biomarcador puede ser a causa de la región que se estudia o de las especificaciones del propio paciente.

Se debe considerar que debido a lo poco implementada que está la técnica de radioablación actualmente en la vía clínica habitual, la base de datos disponible para la extracción de características de pacientes con TV recurrente o refractaria es muy reducida. Aunque la metodología descrita permite la identificación de potenciales biomarcadores en sujetos independientes, se requiere una base de datos mayor para la unificación de los potenciales biomarcadores. Si se logra unificar los potenciales biomarcadores independientemente de la región, se puede entrenar un modelo de *machine learning* para la predicción de pacientes con probabilidad de sufrir TV.

Aunque la metodología desarrollada ha sido diseñada para la extracción de características cuantitativas del corazón, esta puede ser empleada para el análisis de diferentes órganos. Se debe tener en cuenta que, si se desea realizar la extracción de múltiples ROIs simultáneamente, se debe de modificar el algoritmo.

4. Conclusiones

La TV recurrente o refractaria es una patología que afecta de manera considerable la calidad de vida de los pacientes y puede llegar a comprometer su supervivencia. La ablación cardíaca por radioterapia externa presenta una reducción notable y prácticamente instantánea de los episodios de TV tras la finalización del tratamiento, y también la reducción o eliminación de fármacos (muchos de ellos hepatotóxicos). La extracción de características cuantitativas tiene el potencial de convertirse en una poderosa herramienta para la ayuda en la decisión clínica por parte de los cardiólogos. Así, en este trabajo se identifican potenciales biomarcadores de imagen para determinar a los pacientes candidatos a ser tratados por ablación cardíaca por radioterapia.

Agradecimientos

Este trabajo se ha realizado dentro del marco de la comunidad EELISA Health in the City.

Referencias

- [1] Mayo Clinic. Ventricular tachycardia. Symptoms & causes. 2022. <https://www.mayoclinic.org/diseases-conditions/ventricular-tachycardia/symptoms-causes/syc-20355138> (Consultada: Sep 2023).
- [2] Mitchell LB. Taquicardia ventricular (TV). <https://www.msmanuals.com/es-es/professional/trastornos-cardiovasculares/arritmias-cardiacas-especificas/taquicardia-ventricular-tv> (Consultada: Sep 2023).
- [3] Sanz Salvo J, et al. Taquicardia ventricular incesante como manifestación de isquemia miocárdica. *Revista Española de Cardiología*, vol 55, no 2, 2002, pp 193–199.
- [4] Mayo Clinic. Ventricular tachycardia. Diagnosis & treatment. 2022. <https://www.mayoclinic.org/diseases-conditions/ventricular-tachycardia/diagnosis-treatment/drc-20355144> (Consultada: Sep 2023).
- [5] Cuculich PS, et al. Noninvasive Cardiac Radiation for Ablation of Ventricular Tachycardia. *New England Journal of Medicine*, vol 377, no 24, 2017, pp 2325–2336, (ISSN: 0028-4793).
- [6] Martí-Almor J, et al. Ablación no invasiva de taquicardia ventricular con radioterapia estereotáctica en un paciente con miocardiopatía arritmogénica de ventrículo derecho. *Revista Española de Cardiología*, vol 73, no 1, 2020, pp 98–99, (ISSN: 0300-8932).
- [7] Lydiard S, et al. A Review of Cardiac Radioablation (CR) for Arrhythmias: Procedures, Technology, and Future Opportunities. *International Journal of Radiation Oncology*Biophysics*Physics*, vol 109, no 3, 2021, pp 783–800, (ISSN: 0360-3016).
- [8] Van Griethuysen JJM, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer research*, vol 77, no 21, 2017, pp e104–e107, (ISSN: 1538-7445).
- [9] McKinney W. Data Structures for Statistical Computing in Python. in *Proceedings of the 9th Python in Science Conference*, 2010, pp 56–61.
- [10] Waskom ML. seaborn: statistical data visualization. *Journal of Open Source Software*, vol 6, no 60, 2021, p 3021, (ISSN: 2475-9066).
- [11] Zwanenburg A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*, vol 295, no 2, 2020, pp 328–338, (ISSN: 1527-1315).

Deep Learning-Based Generation of Synthetic CT from MR Images for Craniosynostosis Planning

C. Santos Mayo¹, L. Cubero¹, A. Aguado del Hoyo^{2,4}, S. Ochandiano^{3,4}, J. Pascau^{1,4}

¹ Departamento de Bioingeniería, Universidad Carlos III de Madrid, Madrid, Spain

² Departamento de Radiología, Hospital General Universitario Gregorio Marañón, Madrid, Spain

³ Servicio de Cirugía Oral y Maxilofacial, Hospital General Universitario Gregorio Marañón, Madrid, Spain

⁴ Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain

Abstract

Craniosynostosis is a rare congenital defect caused by the premature fusion of one or more cranial sutures. This untimely cranial ossification hinders correct brain development. Its clinical diagnosis and treatment planning usually rely on Computed Tomography (CT), a potentially harmful imaging technique for young infants. It is with the intent of avoiding the use of ionizing radiation in this clinical pipeline that this work studies how feasible it is to resort to alternative non-detrimental imaging techniques such as Magnetic Resonance Imaging (MRI). We evaluate the performance of neural network generators trained on Generative Adversarial Networks in the MRI-to-CT translation task. We train nine generative models on 25 paired MR-CT medical scans, and validate and test their performance on 8 and 4 paired images, respectively. The results are promising both from qualitative and quantitative standpoints, particularly those of the models trained directly on 3D data. Results demonstrate that it is feasible to generate reliable and accurate synthetic CT scans from MR images with the proposed framework, opening up the possibility of harnessing the benefits of non-ionizing techniques to drive craniosynostosis diagnosis and treatment planning.

1. Introduction

1.1. Craniosynostosis

Craniosynostosis is a rare congenital defect caused by the premature fusion of one or more cranial sutures [1]. Formed during embryonic development, these sutures unite the cranial bones and allow for skull deformation during birth and postnatal calvarial expansion, as caused by natural brain development [2].

Untimely suture ossification leads to a growth restriction in the hardened direction and a consequent compensatory expansion perpendicular to it, ultimately hindering correct brain development [2-3]. The clinical effects of craniosynostosis include intracranial pressure, facial deformations, and motor and vision impairments [3].

Since craniosynostosis hinders proper neonatal development, a prompt and accurate diagnosis is essential to guide any active surgical correction.

Being an osseous-related pathology, the golden standard for its medical diagnosis relies on Computed Tomography (CT), which allows to assemble a precise volumetric model of the patient's skull. Afterwards, the defined cranial structure is compared to that of healthy patients to quantify the

malformation type, degree and location in order to best guide the consequent surgical treatment [3].

In most cases, the standard corrective procedure is open cranial vault remodelling, in which the surgeon detaches the affected cranial bones, remodels them externally and finally implants them back into the patient [3].

1.2. Clinical complication and study motivation

Computed Tomography uses ionizing radiation, which is especially detrimental to young infants who present a developing brain still not fully protected from external factors. Studies estimate that these patients receive three times the amount of radiation a normal adult does, ultimately increasing their predisposition to neural damage. Moreover, craniosynostosis patients undergo repetitive CT scanning throughout the initial diagnosis and postoperative follow-up, raising the medical concern during this process.

On the other hand, other safer imaging modalities, such as Magnetic Resonance Imaging (MRI), hold no clinical relevance in the diagnosis of craniosynostosis. MRI does not reveal osseous structures due to their inherent lack of fluid contents.

This study proposes a solution to this medical quandary by effectively designing a bridge between the two imaging modalities. More specifically, our aim is to substitute CT scans with synthetic CT scans generated from MR images. To this end, we have trained and optimized nine generative models on three Generative Adversarial Network frameworks using a database of 37 MR-CT paired head scans. In doing so, we demonstrate how it is possible to integrate these generative techniques in the clinical pipeline for craniosynostosis to harness the benefits that MRI's non-ionizing radiation presents and still obtain an accurate model of the patient's skull to drive diagnosis and treatment.

2. Materials and Methods

2.1. Database and processing pipeline

Due to the irrelevance that MRI data still presents in the clinic for craniosynostosis, a multi-modality MR-CT database from these patients was unavailable at the time of the study.

Instead, the CERMEP-IDB-MRXFDG [4] database was used, retaining paired head MR-CT scans of 37 healthy adults. Additional pre-processing was done in terms of MNI

frame cleaning, histogram clipping, surrounding elements removal, volume enclosing to satisfy network requirements and histogram rescaling into 8-bit resolution for standardization. After pre-processing, scans presented 256x256x150 pixel image dimensions with an isotropic 1x1x1 mm image resolution.

The database was finally divided into a train-validation-test (25-8-4) configuration for model training.

2.2. Generative Adversarial Networks

Generative Adversarial Networks (GANs) [5] are a Deep Learning methodology that allows to manufacture generative models. Generative models are designed to learn a real dataset’s distribution. By learning the ground truth distributions of *what is*, **Generators** offer the possibility to create synthetic data points that, even if not pertaining to the real dataset, are still able to match its distribution and approximate real instances. GANs drive the Generator training by confronting the model with a secondary convolutional neural network, a **Discriminator**, whose goal is to distinguish between the real data samples from the actual dataset and the fake instances produced by the Generator [5].

During GAN training, the Generator and Discriminator are pitted against one another in a mini-max two-player game setting where they are updated one at a time (Figure 1). Through training, generators become better at creating convincing fake samples, whereas discriminators improve in telling these artificial creations apart from real data.

In this study, we compared the performance of three GAN frameworks:

- **Pix2Pix** [6] and **3DcGAN** [7] model image translation by taking insights from paired datasets in 2D and 3D, respectively.
- **CycleGAN** [8] implements a dual and cyclic translation to ensure robustness when training on unpaired datasets. This study only contemplates the 2D version of this framework.

In order to satisfy network requirements, for models trained on the Pix2Pix and CycleGAN frameworks, CT images were sliced and translation was modelled in a 2D axial domain. The final 3D synthetic CT was reconstructed by stacking the resulting 2D axial synthetic CT slices. For models trained on the 3DcGAN configuration, translation was directly modelled in the volumetric domain and scans were kept intact. Table 1 depicts the parameters used during training.

Training parameter	Selection
GAN loss	Pix2Pix: cBCE+ λ_1 L1 CycleGAN: LS+ λ_2 Cyclic+ λ_3 Idt 3DcGAN: cBCE+ λ_1 L1+ λ_2 Edge
λ_1	100
λ_2	10
λ_3	0.5
Minimizer	Step Gradient Descent
Optimizer	Adam Solver
β_1	0.5
β_2	0.999
Initial Learning Rate	2×10^{-4}
N epochs + decay	2D: 100 + 200 3D: 3000 + 3000
Learning Rate Decay	Linear
Batch size	1
Data Augmentation	2D: Resize and Crop + Rotation 3D: R&C + Intensity Rescaling

Table 1. Network training parameters.

2.3. Model Architecture Comparison

We compared the performance of two generator architectures.

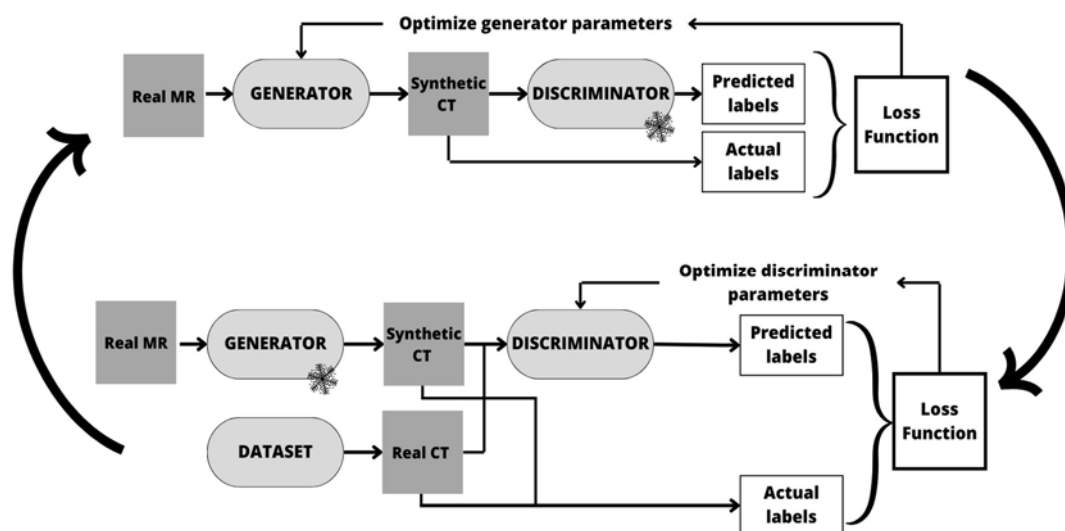


Figure 1. Conditional GAN training approach for the MRI-to-CT translation task.

- **U-Net** embodies an encoder-decoder architecture with long skip connections between symmetrical blocks. It concatenates contractive and expanding paths, each formed by 8 convolutional blocks connected by upsampling/downsampling operations, to fuse information on *what is* with *where it is*.
- **Res-Net** is a generator with 2 encoding/decoding blocks enclosing 9 residual learning blocks. For the 3D models, the generator was adapted as described in [7].

The discriminator implemented was a **Patch-GAN** [6] that, unlike traditional discriminators that classify an entire image as real or synthetic, constricts the categorization problem locally by looking only at smaller image patches to later average its response. By doing so, it pushes the generator to capture both low and high-frequency components [6] and, therefore, aids in accurately defining the bone-soft tissue boundary. The 3D implementation was as depicted in [7].

2.4. Evaluation

GANs drive generator training through network confrontation. In this configuration, overfitting and underfitting are not always evaluated attending a strict training loss convergence criterion. Instead, it is usually recommended to periodically generate synthetic samples to assess their quality as well as model performance.

Training performance was evaluated by comparing the ground-truth CT scans on the validation dataset to their homologous generated synthetic CT images, attending three metrics:

- **Mean Absolute Error (MAE):** computed as the average pixel-wise absolute difference. Unitless since scans were rescaled during processing.
- **Dice Similarity Coefficient (DSC):** measured as the overlap between skull segmentations, built from the ground-truth and synthetic CT volumes. It is a unitless metric bound in the [0-1] range. The closer to the upper limit, the better the model performance.
- **Hausdorff Distance 95th percentile (HD95):** defined as the 95th percentile in the surface distance between skull segmentations. Measured in millimetres (mm), a low value indicates good model performance.

Models were evaluated every 10 training epochs for 2D networks, and every 500 iterations for 3D ones. Finally, the models saved on the best iterations were compared on the testing set.

In total there were 9 models trained:

- 4 (2 U-Net + 2 Res-Net) on Pix2Pix
- 3 (2 U-Net + 1 Res-Net) on CycleGAN
- 2 (0 U-Net + 2 Res-Net) on 3DcGAN

3. Results

The performances on the validation set of the best-iteration models are depicted in Table 2. The five most accurate models were used for the final qualitative (Figure 2) and quantitative (Figure 3) evaluation on the test set.

MAE and DSC-based ANOVA testing confirmed final model performance heterogeneity (p-values < 0.05). All 2D models showed similar performances. U-Net generator behaved slightly better, but there was no significant difference with respect to the other two 2D models (p-values > 0.1). On the contrary, the two 3D models showed no significant differences but generated better predictions than 2D networks. This can be affirmed from a qualitative comparison (Figure 2) as well as quantitatively from the MAE, DSC and HD95 evaluation metrics. For these 3D networks, DSC score was superior to 0.93 whereas MAE and HD95 distances were inferior when compared to the other 2D models.

4. Discussion

This study proposes a GAN-based framework for head MR-to-CT mapping to explore the feasibility of CT substitution in the clinical pipeline of craniosynostosis. In the early stages of our study, a database of newborn patients (under two years of age) was gathered from Hospital Gregorio Marañón. However, the scarcity and heterogeneity of these data rendered it invalid for further model development and testing. As a proof-of-concept, this work demonstrates that the head MR-to-CT translation task is feasible to a high degree of accuracy on adult patients. All trained models achieved high accuracy, especially those in the volumetric domain. 3D networks output synthetic CTs with fewer blurriness and stair-step artefacts, and ultimately better

N	GAN	Generator	D. A.	MAE	DSC	HD95 (mm)
1*	Pix2Pix	U-Net	No	10.08±0.35	0.88±0.00	2.36±0.07
2	Pix2Pix	U-Net	Yes	11.11±0.34	0.85±0.01	3.19±0.12
3*	Pix2Pix	Res-Net	No	9.77±0.33	0.88±0.01	2.36±0.10
4	Pix2Pix	Res-Net	Yes	10.64±0.37	0.87±0.01	2.54±0.09
5	CycleGAN	U-Net	No	15.18±0.54	0.80±0.00	3.43±0.13
6	CycleGAN	U-Net	Yes	14.10±0.34	0.82±0.01	3.74±0.13
7*	CycleGAN	Res-Net	No	9.50±0.25	0.87±0.01	2.45±0.14
8*	3DcGAN	Res-Net	No	7.12±0.43	0.93±0.00	1.41±0.00
9*	3DcGAN	Res-Net	Yes	6.94±0.33	0.93±0.00	1.40±0.07

Table 2. Best-iteration model performances on the validation set based on MAE, DSC and HD95 metrics. Models marked with * were kept for final comparison on the testing set. N: Experiment number. D.A.: Data Augmentation.

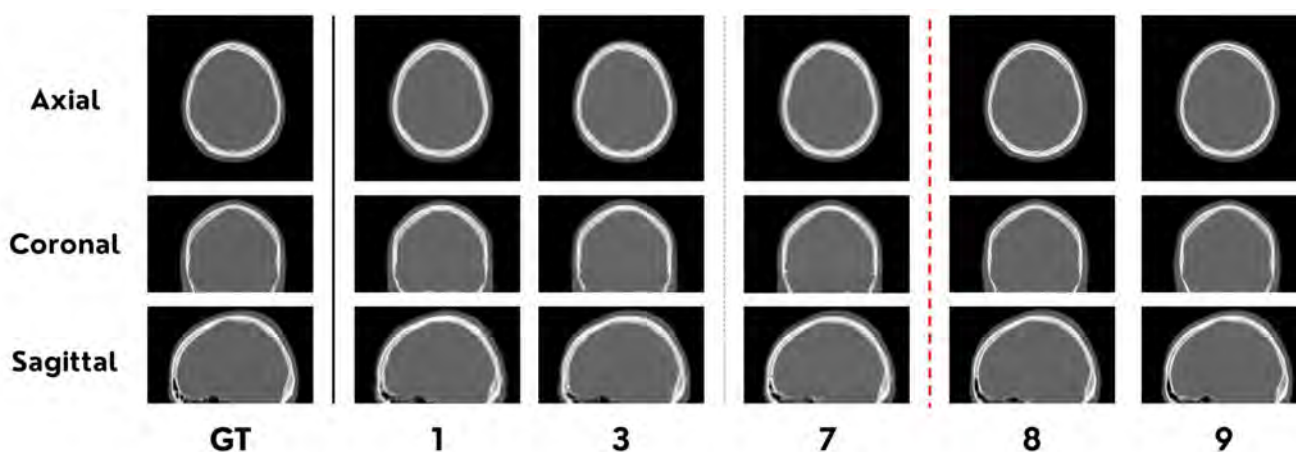


Figure 2. Axial, coronal and sagittal visual comparisons between models.

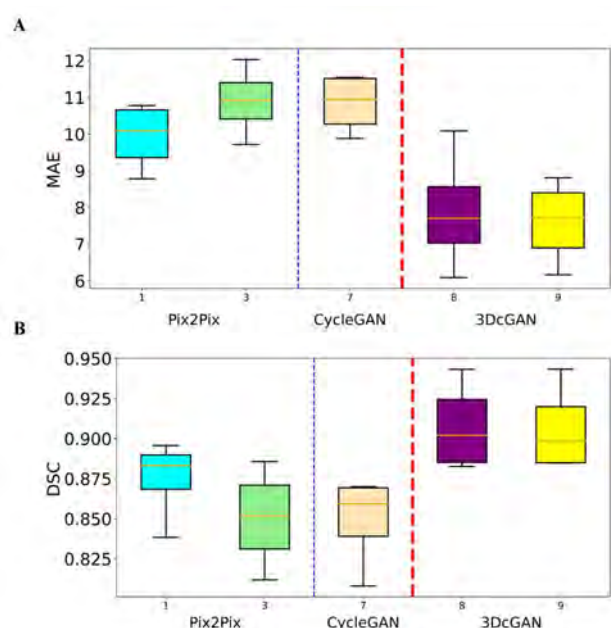


Figure 3. Quantitative metric comparison between models attending to A. MAE; B. DSC.

delineate the cranial surface along all image dimensions, resulting in closer approximations to the real CT images. While implementing our framework in the clinic is still not feasible since training still needs to be done on a large craniosynostosis-specific database, this work showcases how the concept could someday become a reality. Future work will revolve around the creation of such craniosynostosis-specific database to use to re-train the best-performance models.

Another limitation of the study is that the synthetic CT images do not represent the full Hounsfield unit intensity range. This is not necessarily a major obstacle in the craniosynostosis pipeline since it relies primarily on the spatial resolution of the images, which has been completely preserved. Additionally, as the database comprised data only from healthy individuals, translating this framework to pathological cases requires further and cautious analysis.

Overall, this study shows promising evidence of the possibilities for limiting the use of ionizing radiation in the clinic for craniosynostosis diagnosis and treatment planning.

Acknowledgements

CERMED database has © Copyright CERMED – Imagerie du vivant, www.cermep.fr and Hospices Civils de Lyon. All rights reserved.

Research supported by projects PI122/00601 and AC20/00102 (Ministerio de Ciencia, Innovación y Universidades, Instituto de Salud Carlos III, Asociación Española Contra el Cáncer and European Regional Development Fund “Una manera de hacer Europa”), project PerPlanRT (under the frame of ERA PerMed), TED2021-129392B-I00 and TED2021-132200B-I00 (MCIN/AEI/10.13039/501100011033 and European Union “NextGenerationEU”/PRTR).

References

- [1] B. Slater et al., “Cranial sutures: A brief review,” *Plastic and reconstructive surgery*, vol. 121, 170e–8e, May 2008. doi: 10.1097/01.prs.0000304441.99483.97.
- [2] N. Kajdic, P. Spazzapan, and T. Velnar, “Craniosynostosis - recognition, clinical characteristics, and treatment,” eng, *Bosnian journal of basic medical sciences*, vol. 18, no. 2, 2018.
- [3] D. García Mato, *Optimization of craniosynostosis surgery: Virtual planning, intraoperative 3d photography and surgical navigation*, eng, 20.
- [4] I. Mérida et al., “Cermep-idb-mrxfdg: A database of 37 normal adult human brain [18f]fdg pet, t1 and flair mri, and ct images available for research,” eng, *EJNMMI research*, vol. 11, no. 1, pp. 91–91, 2021.
- [5] I. Goodfellow et al., *Generative Adversarial Networks*, vol. 63, no. 11. NEW YORK: Assoc Computing Machinery, 2020, pp. 139–144. doi: 10.1145/3422622.
- [6] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” in *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*, 2017.
- [7] H. Liu, M. K. Sigona, T. J. Manuel, L. M. Chen, B. M. Dawant, and C. F. Caskey, “Evaluation of Synthetically Generated CT for use in Transcranial Focused Ultrasound Procedures,” arXiv.org, 2022.
- [8] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *Computer Vision (ICCV), 2017 IEEE International Conference on*, 2017.

DeepCeres: red neuronal artificial para la segmentación de lóbulos del cerebelo a partir de imágenes multimodales de RM de alta resolución

Sergio Morell Ortega ¹, Marina Ruiz ¹, Pierrick Coupé ² y J.V Manjón ¹

¹ Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas (ITACA), Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, España

² CNRS, Univ. Bordeaux, Bordeaux INP, LABRI, UMR5800, PICTURA, F-33400 Talence, Francia
Email: sermoor1@teleco.upv.es

Resumen

En el presente trabajo se ha diseñado e implementado una red neuronal artificial (RNA) volumétrica para la segmentación de lóbulos cerebelares a partir de imágenes multimodales (T1 y T2) de resonancia magnética nuclear de alta resolución (0.5 mm³). La herramienta permite mejorar la precisión en la estimación de volúmenes de subestructuras cerebelares (lóbulos) a partir de imágenes de RM de alta resolución, aumentando la capacidad de herramientas actuales (CERES, RASCAL, etc) que funcionan a resolución estándar (1 mm³) y no pueden medir con precisión la intrincada estructura cerebelar (con grosores corticales submilimétricos en algunas áreas).

1. Introducción

El cerebelo es la parte del encéfalo situada en la la región posterior de la cabeza, entre el cerebro y el tronco encefálico. Aunque sólo representa entorno al 10 % de la masa total del encéfalo, contiene aproximadamente el 50 % de las neuronas del cerebro, gracias al elevado grado de empaquetamiento neuronales en la superficie de los lóbulos, las subdivisiones anatómicas del mismo [1].

La era moderna en el estudio del cerebelo comenzó hace 120 años, cuando Santiago Ramón y Cajal publicó su primer trabajo. Este trabajo ha sido continuado en los últimos años, y tanto la comprensión de la función normal del cerebelo como de sus patologías, se ha consolidado como un foco activo en el campo de la neurociencia [2]. Durante décadas, se ha tenido la noción de que el cerebelo estaba exclusivamente involucrado en el control motor, pero en los últimos años la neurociencia entorno a esta estructura ha sufrido un cambio de paradigma. Las observaciones neuroanatómicas de imágenes cerebrales han recontextualizado el cerebelo como un nodo crítico en los circuitos neuronales distribuidos que sirven al comportamiento. La idea de que el cerebelo modula el pensamiento y la emoción de la misma manera que modula el control motor, avanza en la comprensión de los mecanismos de la cognición, y abre nuevas oportunidades terapéuticas en neurología conductual y neuropsiquiatría [3]. Por lo tanto, existe la necesidad de disponer de nuevos métodos que permitan el estudio objetivo y cuantitativo del cerebelo, dada la complejidad del problema derivada de la baja resolución de la imágenes en la clínica o la falta de herramientas automáticas precisas, entre otros factores.

El presente trabajo presenta una solución automática, robusta y eficiente para el problema de la segmentación del cerebelo, facilitando así el progreso en la comprensión

en profundidad de esta estructura. Para ello se hace uso de algoritmos basados en redes neuronales convolucionales, que en los últimos años han demostrado una gran aplicabilidad en el campo del análisis de imagen médica, complementando e incluso mejorando los resultados de algoritmos basados en enfoques clásicos [4].

2. Materiales

En este trabajo se han usado 75 casos de la base de datos del *Human Connectome Project* con imágenes de resonancia magnética (RM) multimodales T1 y T2 de una resolución de 0,7 mm³.

Para poder generar las segmentaciones de los 75 casos se ha empleado el *software* CERES [5], teniendo en cuenta que este método era el mejor según la comparativa del estado del arte [6]. Como resultado se generó mapas de segmentación con 28 etiquetas, siguiendo el protocolo de segmentación de Schmahmann [7] recogido en [6].

2.1. Generación de los datos de entrenamiento

Para generar las segmentaciones de los casos se ha llevado a cabo el siguiente preprocesado: eliminación de ruido, corrección de inhomogeneidad en el espacio nativo, registro afín al espacio MNI152 (a una resolución de 0,5 mm³), corrección de inhomogeneidad en el espacio MNI, recorte local del área del cerebelo y normalización de intensidad. Teniendo en cuenta que el *software* de segmentación CERES funciona a una resolución de 1mm³ se utilizó un decimado basado en *strides* para la descomposición y posterior reconstrucción de las segmentaciones, tal y como se observa en la Figura 1.

2.2. Corrección manual de la materia blanca

Tras procesar las imágenes se observó un subsegmentación de la sustancia blanca dentro de los lóbulos del cerebelo. En consecuencia se realizó un proceso de mejora de la segmentación de dicha estructura por parte de un anotador experto.

El resultado se puede observar en la Figura 2, donde se representa la sustancia blanca para ambos hemisferios, también conocida como árbol de vida por los anatomistas. Para aumentar el tamaño muestral se realiza el volteado horizontal de los casos, generando 150 casos. De los cuales 130 se usaron en la partición de entrenamiento, 10 en la de validación y 10 en test.

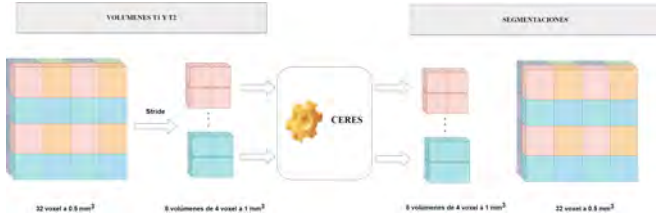


Figura 1: Método para generar las segmentaciones de la librería con CERES.

3. Métodos

3.1. Diseño de las arquitecturas de *Deep Learning*

Teniendo en cuenta las limitaciones computacionales propias al tratar con datos volumétricos de alta dimensionalidad (tensores de entrada de 252x156x162 voxels), con dos canales (T1 y T2) y con una alta densidad de etiquetas (27), se enfocó el diseño en una arquitectura en cascada. Tal y como se puede ver en la Figura 3 primero se realiza una segmentación del cerebelo en hemisferio derecho, izquierdo y fondo. Esta segmentación sirve como máscara cerebelo-fondo para los datos de entrada de la segunda red, facilitándole la atención intra-cerebelar.

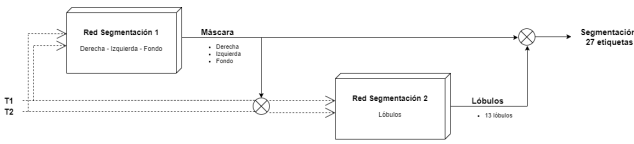


Figura 3: Arquitectura de segmentación.

En cuanto a las redes empleadas en las dos etapas, se ha experimentado con dos topologías:

- La clásica U-net [8], ampliamente validada por la comunidad científica en tareas de segmentación.
- Una adaptación de la anterior realizada en el grupo de investigación donde se sustituye la parte del *encoder* por sucesivas capas de *average pooling*. La arquitectura que hemos denominado DPN (*Deep Piramidal Network*) libera aproximadamente la mitad de memoria que una U-net completa, y funciona usando un enfoque *coarse-to-fine*, empezando con una imagen de una resolución 1/8 de la original y va refinando el resultado a medida que avanza el *decoder*.

Ambas dos topologías han sido entrenadas en las dos tareas (red de derecha-izquierda y lóbulos) para minimizar la siguiente función de pérdida

$$\mathcal{L} = \log \left(\frac{(1 - \sum_{c=1}^C Dice(y_c, \hat{y}_c)) + BCE(y, \hat{y})}{2} \right) \quad (1)$$

Donde C son las clases (etiquetas), \hat{y}_c es el tensor de salida para la clase c e \hat{y}_c es el tensor de salida para esa misma

clase. Dice es el coeficiente empleado para medir la similitud entre las segmentaciones (1 cuando son idénticas y 0 al contrario), y BCE es la entropía binaria cruzada. Adicionalmente, el operador logaritmo ha permitido aliviar los problemas de desvanecimiento del gradiente y acelerar la convergencia durante el entrenamiento.

3.2. Introducción de atlas como canal adicional

Hasta el advenimiento de las técnicas de aprendizaje profundo la mayoría de los métodos (como CERES) estaban basados en uso de un atlas. Estos métodos se caracterizaban por su robustez e interpretabilidad, a costa de una precisión limitada y con un coste computacional elevado. Actualmente, los métodos basados en aprendizaje profundo permiten obtener precisiones más elevadas con un mayor rendimiento temporal y computacional, pero en ocasiones carecen de la robustez de los métodos pasados.

Debido a esto se propone incorporar una presegmentación auxiliar como una información *a priori* generada mediante métodos basados en atlas. De esta forma el modelo cuenta con una información previa que le facilita la comprensión semántica de las estructuras del cerebelo.

Con el objetivo de aumentar la robustez en la generación de la segmentación auxiliar, se emplea un enfoque multi-atlas, donde se realiza un registro no lineal de cada imagen de la librería de entrenamiento (casos de HCP con sus T1 y correspondientes segmentaciones) a la imagen a segmentar. Posteriormente se aplican los registros a las segmentaciones, creándose así una librería específica de segmentaciones. Finalmente, se fusionan siguiendo el algoritmo de *Weighted Majority Voting*, donde para cada píxel se asigna la etiqueta resultado de pesar las etiquetas de cada caso de la librería por la similitud del nivel de intensidad de la T1 con respecto a la T1 del casos a segmentar.

Cabe destacar que el registro no lineal con técnicas habituales tiene un elevado coste temporal, así como la fusión de las etiquetas. Para la tarea de registro se ha empleado una red basada en aprendizaje profundo (una adaptación de *voxelmorph*[9]), y para la fusión de las etiquetas se ha implementado un versión multiproceso en *OpenMP* integrable en Python mediante la API *Numpy-C*. La construcción de la segmentación auxiliar *a priori* se podría resumir en la Figura 4.

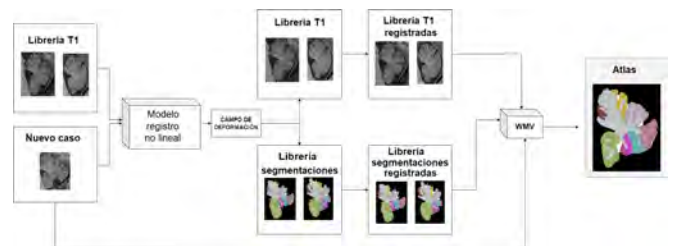


Figura 4: Método para la construcción de la segmentación auxiliar (multi-atlas)

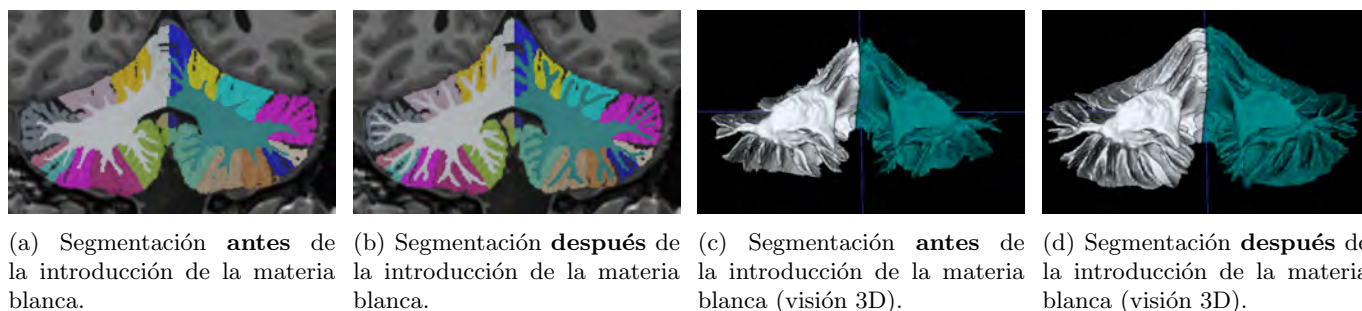


Figura 2: Vista coronal de las segmentaciones antes de la corrección manual y después.

4. Resultados y discusión

4.1. Resultados generales

Todos los experimentos se realizaron en una plataforma DGX1 de Nvidia con 8 GPUs V100 con 32 GB de memoria cada una en un entorno Linux (Ubuntu 19). Se utilizó la versión 1.15 de Tensorflow y la versión 2.2.4 de keras. Adicionalmente se realizaron los tests en equipos con GPUs con 16 GB de memoria para validar su aplicabilidad en equipos con menos prestaciones computacionales.

Se analizó el impacto general de las distintas arquitecturas UNET (3 433 473 parámetros) vs DPN (696 177 parámetros) o el uso de información *a priori* (con o sin atlas). En la Tabla 1 se recogen los valores del índice DICE de las variantes estudiadas y su comparación con el método de referencia CERES.

Nombre	Dice medio	Dice Cerebelo
CERES	0.7729 ± 0.0427	0.9377 ± 0.0090
2 DPN	0.9248 ± 0.0367	0.9887 ± 0.0028
2 Unet	0.9295 ± 0.0438	0.9900 ± 0.0022
2 DPN multi-atlas	0.9342 ± 0.0342*	0.9917 ± 0.0015
2 Unet multi-atlas	0.9347 ± 0.0339*	0.9903 ± 0.0020

Tabla 1: Análisis de las arquitecturas propuestas. (Resaltado el mejor de los resultados en cada caso en negrita. * Diferencias significativas para un test Wilcoxon $p < 0,05$)

En primer lugar, es posible observar la mejora sustancial con respecto al índice Dice del método CERES. No obstante, esta comparativa no es completa, pues los datos de empleados para el *test* fueron distintos. Sin embargo, da una idea orientativa de la ganancia obtenida tras la mejora de la librería, la introducción de la potenciación en T2 y la disponibilidad de imágenes de alta resolución.

Es posible ver en la Tabla 1 como la introducción del atlas ha tenido un impacto importante a nivel de lóbulos y de cerebelo completo, tanto para los modelos basados en la arquitectura DPN como en la Unet. En ambos ca-

sos, conceptualmente la red se especializa en la extracción semántica de los detalles finos de los lóbulos, sin necesidad de extraer características desde cero, además de acelerar la convergencia de los entrenamientos.

4.2. Impacto de la multimodalidad y alta resolución

Tras la selección de los modelos con atlas como aquellos con mejor desempeño, se ha comprobado el impacto que tiene el uso de imágenes de alta resolución ($0,5 \text{ mm}^3$ de resolución espacial) y la introducción de la imagen potenciada en T2.

Para ello se han entrenado los modelos de nuevo siguiendo la misma estrategia que los originales, pero en este caso eliminando el canal de la imagen potenciada en T2 o diezmando los volúmenes de entrada para reducir la resolución. Los resultados se pueden ver en la Tabla 2.

Resolución, Multimodalidad	Tipo de red	Dice medio
BR, T1	2 DPN	0.9117 ± 0.0473
BR, T1	2 UNET	0.9079 ± 0.0520
BR, T1+T2	2 DPN	0.9138 ± 0.0468
BR, T1+T2	2 UNET	0.9088 ± 0.0468
AR, T1	2 DPN	0.9239 ± 0.0353
AR, T1	2 UNET	0.9276 ± 0.0390
AR, T1 + T2	2 DPN	0.9332 ± 0.0439
AR, T1 + T2	2 UNET	0.9345 ± 0.0493

Tabla 2: Impacto de la resolución y la multimodalidad para los dos mejores modelos con atlas específico (para las 2 redes). AR: Alta Resolución (0.5 mm^3) y BR: Baja Resolución (1 mm^3). (* Mejores resultados resaltados en negrita).

Es posible observar como el aumento de resolución tiene un gran impacto sobre el índice Dice a nivel de lóbulos (0.9345 vs 0.9088 en UNET y 0.9332 vs 0.9138 en DPN). Cabe recordar la intrincada estructura del cerebelo, donde existe un elevado grado de empaquetamiento de tejidos y una superficie con un alto grado de textura, la folia, con grosores corticales submilimétricos. Esto provoca problemas de volumen parcial en imágenes de resolución estándar (1 mm^3) donde distintos tejidos aportan información sobre un mismo voxel, impidiendo su clasificación independiente. En alta resolución los problemas de volumen parcial se reducen. Por otra parte, pese a

que la ganancia de introducir la T2 es menor (0.9345 vs 0.9276 en UNET y 0.9332 vs 0.9239 en DPN), se observó un impacto positivo ya que permite reducir algunos falsos positivos y negativos al dar información de otra modalidad para una misma estructura anatómica, como podría ser la sustancia blanca y el líquido (vasos sanguíneos) que brillan de forma similar en T1 y distinta en T2.

Finalmente, con respecto a la disponibilidad en la práctica clínica de datos multimodales y de alta resolución, cabe destacar que datos de esta calidad son escasos. No obstante, y pese a que los modelos propuestos presentan rendimientos superiores a otros métodos incluso a una resolución estándar (1 mm³) se está trabajando en el diseño de métodos basados en aprendizaje profundo para sintetizar volúmenes T2 y aumentar de resolución de los volúmenes T1 y T2 usando técnicas de superresolución.

4.3. Ensemble de modelos

Analizando los resultados de las dos arquitecturas usadas, vimos que algunas estructuras son segmentadas mejor por un modelo en comparación con el otro, por lo que se propone combinación de ambas arquitecturas para mejorar resultado en términos de índice Dice. En consecuencia se estudió la generación de un *ensemble* de modelos para la segmentación del cerebelo. **Como resultado, el modelo dual combinado presenta un índice dice a nivel de lóbulo de 0.9403 ± 0.0296, en comparación con el mejor modelo individual (Unet con atlas) con un Dice de 0.9347 ± 0.00339** (* Diferencias en las medias significativas para un test Wilconxon $p < 0,05$).

4.4. Comparativa con los métodos del estado del arte

A modo de conclusión es posible realizar una comparativa semicuantitativa del método desarrollado y el estado del arte. Esta comparativa no puede ser totalmente cuantitativa puesto que los datos sobre los que se ha realizado son distintos, pero permite tener una noción del lugar que ocuparía la herramienta descrita en el panorama actual.

Como se puede apreciar en la Tabla 3 el método propuesto es muy superior (0.94) al segundo de la comparativa (0.87) lo que pone de manifiesto lo competitivo de la solución propuesta.

	Acapulco	Cerebnet	Ceres	Método propuesto
Dice Medio	0.77	0.87	0.7729	0.9403

Tabla 3: Comparativa con el resto de métodos del estado del arte.

5. Conclusiones

La fusión de los enfoques clásicos basados en atlas y los de aprendizaje profundo ha demostrado ser efectiva

para mejorar los resultados en ambas arquitecturas. Esta combinación ha dado lugar a modelos más precisos, robustos y eficientes, incluso superando a modelos más potentes pero sin la información previa. Además, se han explorado alternativas a la red UNET clásica, como la DPN, que ha mostrado buenos resultados con muchos menos parámetros (entorno a 1/5 parte), lo que reduce el riesgo de sobreajuste y mejora la generalización. La combinación de varios modelos, como UNET y DPN, ha proporcionado los mejores resultados, reduciendo la varianza del error de clasificación en la segmentación del cerebelo. Finalmente, la comparación con otros métodos del estado del arte, tanto clásicos como basados en *Deep Learning*, muestra que el enfoque propuesto es altamente competitivo, con índices Dice promedio considerablemente altos (0.87 vs 0.94), a pesar de las limitaciones de los datos y protocolos utilizados.

Agradecimientos

Este trabajo ha sido desarrollado gracias al proyecto PID2020-118608RB-I00 del Ministerio de Ciencia e Innovación de España.

Referencias

- [1] Martin I Sereno et al. The human cerebellum has almost 80 % of the surface area of the neocortex. *Proceedings of the National Academy of Sciences of the United States of America*, 117(32):19538–19543, aug 2020.
- [2] Handbook of the Cerebellum and Cerebellar Disorders. *Handbook of the Cerebellum and Cerebellar Disorders*, 2022.
- [3] Jeremy D. Schmahmann et al. The Theory and Neuroscience of Cerebellar Cognition. <https://doi.org/10.1146/annurev-neuro-070918-050258>, 42:337–364, jul 2019.
- [4] Ana Barragán-Montero et al. Artificial intelligence and machine learning for medical imaging: A technology review. *Physica Medica*, 83:242–256, mar 2021.
- [5] Jose E. Romero et al. CERES: A new cerebellum lobule segmentation method. *NeuroImage*, 147:916–924, feb 2017.
- [6] Aaron Carass et al. Comparing fully automated state-of-the-art cerebellum parcellation from magnetic resonance images. *NeuroImage*, 183:150–172, dec 2018.
- [7] JD Schmahmann et al. *MRI atlas of the human cerebellum*. 2000.
- [8] Olaf Ronneberger et al. U-net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science*, 9351:234–241, 2015.
- [9] Guha Balakrishnan et al. VoxelMorph: A Learning Framework for Deformable Medical Image Registration. *IEEE Transactions on Medical Imaging*, 38(8):1788–1800, sep 2018.

DeepThalamus: Red neuronal para la segmentación de núcleos talámicos a partir de imágenes de RM de alta resolución

Marina Ruiz Pérez ¹, Sergio Morell Ortega ¹, Pierrick Coupé ², José V Manjón ¹

¹ Instituto de Aplicaciones de las Tecnologías de la Información y de las Comunicaciones Avanzadas (ITACA), Universitat Politècnica de València, Camino de Vera s/n, 46022, Valencia, Spain

² CNRS, Univ. Bordeaux, Bordeaux INP, LABRI, UMR5800, PICTURA, F-33400 Talence, France
Email: mruiper@etsii.upv.es

Resumen

En este trabajo se presenta una red neuronal volumétrica para la segmentación de los núcleos talámicos a partir de imágenes de RM. Las herramientas actuales funcionan a resolución estándar (1 mm^3) y usan datos monomodales. En este trabajo se emplean imágenes multimodales (T1, T2 y WMn) y de alta resolución (0.5 mm^3) con el objetivo de mejorar la precisión en la estimación de volúmenes de los núcleos del tálamo cerebral. Para ello se creó una base de datos de tálamos segmentados manualmente y se diseñó la arquitectura de la red neuronal. Los resultados obtenidos mejoran el estado del arte y permitirán obtener medidas más precisas y una detección más temprana de patrones patológicos.

1. Introducción

El tálamo es una estructura cerebral que se ve implicada en patologías neurológicas como el Parkinson o la depresión. De ahí la importancia de su análisis anatómico.

Esta región cerebral tiene forma ovoide y está ubicada en la parte dorsal del diencefalo, mide unos 3 cm de largo y 1.5 cm de ancho y recibe información procedente de sensores de movimiento periféricos, del cerebelo y de neuronas de la capa premotora y motora frontales.

También tiene un rol clave en el control de las emociones, a través del circuito de Papez [1].

El tálamo consta de varios núcleos que se proyectan a distintas áreas del cerebro y que se encuentran involucrados en diferentes actividades. Conocer las alteraciones de cada núcleo puede ayudar a comprender la fisiopatología de gran cantidad de enfermedades.

Esto se puede realizar gracias a las técnicas de imagen médica, específicamente la imagen de resonancia magnética nuclear (RMN) es, probablemente, la modalidad de imagen médica más versátil en el ámbito clínico. Dichas imágenes se pueden usar para la segmentación de los núcleos talámicos. Sin embargo, dichos núcleos no se diferencian bien en imágenes de resonancia magnética comunes, como las ponderadas en T1 o T2, pues el contraste intertalámico es intrínsecamente bajo.

Hay diferentes métodos para la segmentación automática de los núcleos talámicos. Uno de los más conocidos es el método incluido en la popular herramienta Freesurfer [2] y usa como imagen de entrada imágenes potenciadas en T1. Otra herramienta más reciente [3] es un algoritmo basado en multi-atlas que usa una librería de 20 imágenes MP-RAGE WMn (*Whittematter-nulled*), con

una resolución de 0.7 mm^3 con segmentaciones manuales de los 12 núcleos del tálamo.

En el trabajo actual se presenta un nuevo método de segmentación que trabaja a alta resolución y de forma multimodal usando técnicas de Deep learning.

2. Materiales

2.1. Dataset

En este proyecto se han empleado 3 datasets distintos. El primero es un conjunto de imágenes de RMN del *Human Connectome Project* (HCP). Este dataset está formado por imágenes de RMN tomadas en una máquina de 3T de 1200 sujetos sanos de entre 22 y 35 años. Se usaron imágenes de alta resolución (tamaño de la matriz= $260\times 311\times 260$ vóxeles, $\text{vóxel size}=0.7\text{ mm}^3$) potenciadas en T1 y T2 de 75 sujetos.

El segundo dataset se compone de 20 sujetos y consistían en imágenes potenciadas de WMn junto con sus segmentaciones manuales (tamaño de la matriz= $256\times 420\times 256$ vóxeles, $\text{vóxel size}=0.7\times 0.5\times 0.7\text{ mm}^3$).

El tercer dataset está formado por 55 sujetos con imágenes potenciadas en T1, T2 y WMn provenientes de un proyecto de la universidad Burdeos (tamaño de matriz de $256\times 376\times 368$ vóxeles y una resolución de 0.6 mm^3). Se muestra un resumen en la **Tabla 1**.

Dataset	Ponderaciones	Número de imágenes
HCP	T1 T2	75
Thomas [3]	WMn	20
Burdeos	T1 T2 WMn	55

Tabla 1. Datasets que componen el conjunto de datos con los que se ha trabajado en este proyecto.

3. Métodos

3.1. Arquitectura de la red neuronal

La red U-NET [4] es la más empleada en segmentación de imagen médica. Está compuesta de un codificador y un decodificador convolucional con conexiones residuales en cada nivel de resolución. Sin embargo, su alta complejidad puede producir problemas de *overfitting*. Además, su entrenamiento con un número limitado de ejemplos también reduce su capacidad de generalización. Por ello, se ha diseñado una nueva red más

simple, siguiendo un enfoque *coarse-to-fine*, con un menor número de parámetros, capaz de aprender de forma correcta las segmentaciones, minimizando los problemas de sobreajuste.

Hemos denominado a esta red como *Deep Pyramidal Network* (DPN) y está basada en un enfoque incremental. Utiliza versiones de la imagen de entrada de resolución 1 , $\frac{1}{2}$, $\frac{1}{4}$ y $\frac{1}{8}$ obtenidas con capas de *average pooling* como entrada.

Al tensor de entrada se le aplican 3 bloques de convolución con *ReLU* y *Batch Normalization*. Después, el tensor resultante se redimensiona para que coincida con las dimensiones de las características del input de $\frac{1}{4}$ de resolución para incorporar características de alta resolución. Tras ello, las concatena.

A este nuevo tensor se le aplica de nuevo el mismo proceso hasta que se llega a la resolución original. Este último tensor se procesa a través de una capa de activación *Softmax*. Cada convolución usa 32 filtros con un kernel de tamaño $3 \times 3 \times 3$. A diferencia de la red UNET, esta arquitectura no usa un modelo de cuello de botella sino un enfoque puramente incremental. Se muestra un esquema de la arquitectura en la **Figura 1**.



Figura 1. Esquema de la arquitectura DPN.

3.2. Preprocesado

Todas las imágenes deben pasar por una etapa de preprocesado para situarlas en un espacio geométrico y de intensidad estándar (lo cual facilita su análisis).

Esta etapa prepara los datos para su procesamiento, se compone de una fase de eliminación de ruido con el filtro *Spatially Adaptive Non local means* (SANLM) [5]. Seguido de la corrección de inhomogeneidad con el método N_4 [6] y, después, un registro afín al espacio MNI152 de 0.5 mm^3 de resolución (*Montreal Neurological institute*). Las imágenes resultantes tienen un tamaño de matriz estándar de $362 \times 434 \times 362$ vóxeles y una resolución de 0.5 mm^3 .

Finalmente, se recorta del área de interés para seleccionar solo el subvolumen que contiene el tálamo derecho e izquierdo para su análisis posterior (de tamaño $76 \times 91 \times 79$ vóxeles).

3.3. Creación de la librería de entrenamiento

Uno de los objetivos de este proyecto es la creación de

una librería multimodal (T1, T2 y WMn) de alta resolución para entrenar distintos modelos de segmentación basados en *deep learning*. Desafortunadamente, el único dataset disponible con segmentaciones manuales es el dataset creado en [3] el cual solo incluye imágenes del tipo WMn. El único dataset disponible con las tres potenciaciones es el de Burdeos. Sin embargo, la calidad de las imágenes T1 y T2 no es tan buena como la calidad de las mismas en el dataset HCP.

Por lo tanto, se decidió transferir las segmentaciones del dataset [3] al dataset HCP. Para ello, era necesario que dicho dataset tuviera imágenes de tipo WMn. Para obtenerlas se usó una técnica de síntesis de imagen que permite sintetizar modalidades no adquiridas a partir de otras [7] usando el dataset de Burdeos para el entrenamiento y aplicando la red entrenada en el dataset HCP. Finalmente, para la obtención de las segmentaciones en el dataset HCP se usó una red neuronal (tipo DPN) entrenada con las imágenes de Thomas [3] y aplicada a las imágenes de HCP.

Las segmentaciones se revisaron manualmente para eliminar los posibles errores de segmentación y se añadió una nueva etiqueta manualmente (el espacio intertalámico) para producir una segmentación compacta de todo el tálamo (todo ello usando la herramienta ITK-SNAP).

3.4. Entrenamiento de la red

Se usó la librería creada para el entrenamiento de la red propuesta. Los 150 tálamos (75 izquierdos y 75 derechos) se dividieron en 130 casos de entrenamiento, 10 de validación y 10 de test.

Los experimentos se realizaron usando tensorflow 1.15 y keras 2.2.4. La red propuesta tiene un número total de 1,034,355 parámetros.

El optimizador utilizado fue adamax y la función de pérdida fue la suma del *generalized dice loss* (GDL) [8] y la *categorical cross entropy* (todo ello en escala logarítmica para acelerar la convergencia).

4. Resultados y discusión

4.1. Resultados de la arquitectura DPN

Los resultados obtenidos para el tálamo completo y para cada una de las 12 etiquetas, así como el valor medio de todas ellas puede verse en la **Tabla 2**.

Núcleo	DICE	Vóxeles
Tálamo completo	0.9819 ± 0.0030	
AVN	0.9196 ± 0.0312	1136
VAN	0.9273 ± 0.0249	3274
VLAN	0.9176 ± 0.0274	827
VLPN	0.9573 ± 0.0097	8832
VPLN	0.9375 ± 0.0130	4053
PN	0.9724 ± 0.0052	15487
LGN	0.9129 ± 0.0275	1093
MGN	0.9437 ± 0.0147	835
CN	0.9495 ± 0.0135	1580
MN	0.9719 ± 0.0065	1422
HN	0.9136 ± 0.0368	7980
MT	0.8777 ± 0.0826	308
IS	0.9109 ± 0.0124	22322
Media	0.9317 ± 0.0401	

Tabla 2. Resultados del índice DICE para los diferentes núcleos que componen el tálamo para la arquitectura DPN así como los volúmenes medios de cada núcleo (en vóxeles).

Estos resultados se han obtenido para las imágenes de entrada T1 y WMn.

Se obtuvo un valor medio de 0.9317 para los núcleos, siendo el núcleo con mayor DICE el *Pulvinar Nucleus* (PN) con un valor de 0.9724 y el que tiene el menor es el *Thalamus Mammillothalamic Tract* (MT), con un valor de 0.8777, siendo este de los núcleos más pequeños. En la **Figura 2**, se puede ver una imagen con las segmentaciones obtenidas de los núcleos talámicos.

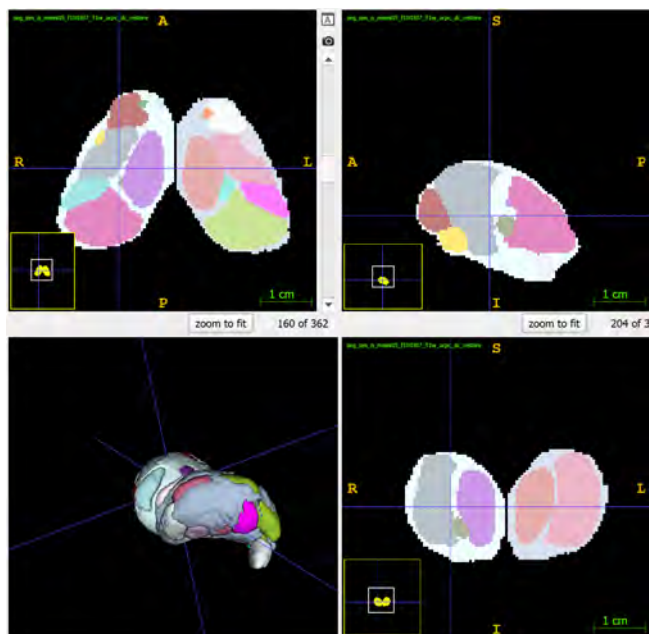


Figura 2. Ejemplo de dos tálamos segmentados (26 etiquetas).

En las segmentaciones, algunos núcleos ocupan menos de 300 vóxeles, mientras que otros ocupan más de 23500. Por esa razón, se utilizó el GDL [8] pero, a diferencia de la versión original, los pesos no se calcularon para cada caso sino para toda la población (media del volumen de

la población). Se comparó la red DPN propuesta con una red de tipo UNET mostrando que la calidad de la red propuesta es muy superior (0.9317 vs 0.9253).

4.2. Resultados multimodalidad

Se analizó el impacto del uso de las distintas modalidades como datos de entrada a la red. Se muestran los resultados en la **Tabla 3**. Se observó que la información multimodal de la imagen ponderada en T1 junto con la imagen WMn es la que da mejores resultados. El experimento realizado con solo la imagen WMn ya alcanza un elevado valor de DICE, acorde con ya lo estudiado [3]. Esta imagen es la que ofrece más contraste entre los núcleos del tálamo, pero junto con la información de la T1 este índice de similitud se eleva aún más. Con esta prueba se valida una de las hipótesis iniciales, que era que la entrada a la red de imágenes multimodales aportaba más información que una única modalidad. Destaca cómo la modalidad T2 empeora los resultados ya sea sola o cuando se combina con las otras modalidades. Esto puede ser debido a que este tipo de ponderación distingue bien el tejido cortical del líquido, pero ve al tálamo como una unidad donde es muy difícil distinguir los núcleos intratálamicas.

Modalidad	DICE
T1	0.8986 ± 0.0532
T2	0.8539 ± 0.0879
WMn	0.9264 ± 0.0456
T1 + T2	0.9016 ± 0.0559
T1 + WMn	0.9317 ± 0.0401
T2 + WMn	0.9212 ± 0.0447
T1 + WMn + T2	0.9290 ± 0.0571

Tabla 3. Resultados del DICE para diferentes ponderaciones de MRI arquitectura DPN.

4.3. Resultados resolución

Para comprobar la hipótesis de que una mayor resolución genera una mejor segmentación, se probó el mejor modelo obtenido con la arquitectura DPN con baja resolución (1 mm³).

El resultado comparativo se muestra en la **Tabla 4**. Se puede ver la diferencia en el valor de Dice alcanzado, donde queda demostrado que es inferior a baja resolución, validando una de las hipótesis.

DPN resolución	DICE
DPN alta resolución (0.5 mm ³)	0.9317 ± 0.0401
DPN baja resolución (1 mm ³)	0.9109 ± 0.0633

Tabla 4. DICE para la comparativa a baja y alta resolución de la arquitectura DPN.

4.4. Comparación con el estado del arte

Una comparación directa del método propuesto con los métodos del estado del arte no es posible. Por ejemplo, Freesurfer [1] usa imágenes de 1 mm³ y su protocolo de segmentación es totalmente distinto. El método más similar al método propuesto es Thomas [3] que consiste

en un algoritmo basado en multi-atlas con imágenes MP-RAGE donde se anula la contribución de materia blanca (WMn). Utiliza 20 imágenes MP-RAGE WMn, con una resolución de 0.7 mm^3 , con segmentaciones manuales de los 12 núcleos del tálamo realizadas por un neurorradiólogo experto. Para generar la segmentación, parte de un multi-atlas previo compuesto por estas 20 imágenes WMn, que se registran de forma no lineal y promedian. En la **Tabla 5** se pueden ver los resultados, tanto para el tálamo completo como para todos los núcleos, el método propuesto en este trabajo tiene un índice de DICE más elevado.

Nótese sin embargo que los datasets utilizados son distintos (20 imágenes del datase Thomas vs 10 imágenes de HCP) y que sus resoluciones son distintas (0.7 mm^3 para Thomas y 0.5 mm^3 para HCP). Por último, nótese también que en la comparativa no hemos incluido la etiqueta de espacio intertalámico ya esta no existe en el método de Thomas. Aún, con todas estas diferencias, se puede observar que la diferencia entre los métodos es muy notable, lo que sugiere que el método propuesto es muy competitivo.

Núcleo	Thomas	Propuesto
Tálamo completo	0.92	0.9819
AVN	0.80	0.9196
VAN	0.71	0.9273
VLAN	0.67	0.9176
VLPN	0.79	0.9573
VPLN	0.69	0.9375
PN	0.86	0.9724
LGN	0.75	0.9129
MGN	0.74	0.9437
CN	0.75	0.9495
MN	0.85	0.9719
HN	0.73	0.9136
MT	0.70	0.8777
Media	0.7533	0.9334

Tabla 5. Resultados del índice DICE para los diferentes núcleos que componen el tálamo para THOMAS y el método propuesto.

5. Conclusiones

En este trabajo se ha propuesto un nuevo método de segmentación de núcleos del tálamo basado en deep learning que utiliza imágenes de RMN de alta resolución multimodales. Se ha generado una herramienta de segmentación que proporciona un índice Dice de 0.9317 de valor medio de los núcleos del tálamo y de 0.9819 para el tálamo completo. Mediante los experimentos realizados y los resultados obtenidos se han podido validar las hipótesis iniciales. Es decir, se ha creado una librería multimodal de imágenes de alta resolución etiquetadas y se ha demostrado que el aumento de la resolución y el uso de imágenes multimodales mejoran los resultados de la segmentación.

Por otro lado, la herramienta propuesta ha sido comparada con un método del estado del arte (THOMAS)

obteniendo mejores resultados.

La red desarrollada se incluirá en una herramienta online que hemos llamado DeepThalamus que se pondrá a disposición de la comunidad científica a través de nuestra plataforma volBrain (volbrain.net).

Agradecimientos

Este trabajo ha sido desarrollado gracias al proyecto PID2020-118608RB-I00 del Ministerio de Ciencia e Innovación de España (AEI/10.13039/501100011033)

Referencias

- [1] J. P. Aggleton et al. Thalamic pathology and memory loss in early alzheimer’s disease: moving the focus from the medial temporal lobe to papez circuit. *Brain*, 139(7):1877–1890, 2016.
- [2] Juan Eugenio Iglesias et al. A probabilistic atlas of the human thalamic nuclei combining ex vivo mri and histology. *Neuroimage*, 183:314–326, 2018.
- [3] J.H. Su et al. Thalamus optimized multi atlas segmentation (thomas): fast, fully automated segmentation of thalamic nuclei from structural mri. *Neuroimage*, 194:272–282, 2019.
- [4] Olaf Ronneberger et al. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.
- [5] José V Manjón et al. Adaptive non-local means denoising of mr images with spatially varying noise levels. *Journal of Magnetic Resonance Imaging*, 31(1):192–203, 2010.
- [6] Nicholas J Tustison et al. N4itk: improved n3 bias correction. *IEEE transactions on medical imaging*, 29(6):1310–1320, 2010.
- [7] José V Manjón et al. Deep learning based mri contrast synthesis using full volume prediction using full volume prediction. *Biomedical Physics & Engineering Express*, 8(1):015013, 2021.
- [8] Carole H. Sudre et al. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings 3*, pages 240–248. Springer, 2017.

Biomecánica / Ingeniería de Rehabilitación

Miércoles 22 9:30 - 11:00
Aula Sebastián Ferignan

Análisis del riesgo de rotura de aorta aneurismática con comportamiento anisotrópico

L. García Sánchez^{1,2} P. Navas² J.M. Goicolea³

¹ ETSI de Telecomunicación, Universidad Politécnica de Madrid, laura.gsanchez@alumnos.upm.es

² Grupo de Mecánica Computacional. ETSI Caminos, Universidad Politécnica de Madrid, pedro.navas@upm.es

³ Grupo de Mecánica Computacional. ETSI Caminos, Universidad Politécnica de Madrid, jose.goicolea@upm.es

Resumen

El objetivo principal radica en evaluar las propiedades mecánicas de la pared aórtica en presencia de aneurismas. Para ello, se lleva a cabo la simulación de una geometría de aorta sana y otra aneurismática, tratando de identificar el riesgo de rotura y la necesidad de intervención quirúrgica. Esta información servirá como ayuda a la decisión médica en el tratamiento de aneurismas aórticos.

1. Introducción

Un aneurisma aórtico es una dilatación anormal de la pared de la arteria aorta. Los localizados en la parte ascendente suelen originarse debido a una necrosis, que provoca la fragmentación de elastina y la deposición de colágeno. Esto debilita las paredes aórticas, disminuyendo su resistencia mecánica y aumentando el riesgo de rotura, lo que puede ser potencialmente mortal. [1]

Las paredes aórticas se han modelado con materiales hiperelásticos, pues estos definen de forma adecuada el comportamiento de los materiales blandos. El reclutamiento de las fibras en las paredes arteriales determina su comportamiento anisotrópico, lo que implica que exhiben propiedades mecánicas dependientes de la dirección en función del alineamiento de las fibras. Es por esto que se ha seleccionado el material Holzapfel-Gasser-Ogden (GOH), que se ajusta satisfactoriamente a este comportamiento característico del tejido arterial. Por un lado, tiene en cuenta la propiedad elástica de las paredes definida por la elastina y, por otro lado, el comportamiento anisotrópico dado por el colágeno. En su definición, se describen dos familias de fibras que se encuentran en el plano e_1, e_2 de los ejes materiales locales, formando un ángulo $\pm\gamma$ con respecto a e_1 . El valor del ángulo $\pm\gamma$ determina el comportamiento anisotrópico del material. [2],[3],[4],[5]

Para realizar el análisis mecánico de forma computacional, se realiza una aproximación numérica de los materiales utilizados por el método de Elementos Finitos.

2. Materiales y métodos

2.1. Segmentación y Suavizado

La segmentación de imágenes se realiza en imágenes de Tomografía Computarizada. Estas imágenes son procesadas con el software *3D Slicer*, que utiliza herramientas como *píncel* y *seeds* para la marcación manual de la región de interés. Además, se realiza un suavizado a través del mismo programa y el software *MeshLab*.

2.2. Parches NURBS y mallado

Tras esto, se procede a la preparación de la geometría. Se emplean parches NURBS, con el objetivo de definir la forma tridimensional de la aorta y estructurar la malla a través de la interpolación de puntos de control y ecuaciones algebraicas. Para ello, se emplea el software *SpaceClaim* de *ANSYS*, que permite realizar cortes, extrusiones y ajustes en la geometría para obtener un objeto 3D con el espesor necesario.

En el siguiente paso se genera una malla tridimensional utilizando el programa *GiD* a partir de un archivo **.iges*. La malla, compuesta por elementos tetraédricos de 4 nodos, tiene un tamaño de 5 mm de promedio, totalizando 17.333 elementos para la aorta sana y 18.662 para la aneurismática.

Para trabajar con *FEBio*, se importan las mallas y geometrías mediante archivos **.feb* tras realizar una conversión del **.msh* de *GiD* gracias a una subrutina propia de *MatLab*.

3. Verificación

En este apartado, se detallan los distintos modelos realizados con el fin de analizar y validar el comportamiento estructural bajo distintas configuraciones.

3.1. Ensayo de Tracción Uniaxial

Se realizó un ensayo de tracción uniaxial en *FEBio* para evaluar el material anisotrópico antes de simular la geometría aórtica completa. Se diseñaron dos probetas, una con las fibras orientadas longitudinalmente y otras orientadas circunferencialmente. Se impuso un desplazamiento en la dirección longitudinal y se estudiaron las tensiones en función del alargamiento, obteniendo resultados similares a las investigaciones previas de Claudio García Herrera [2]. Se compararon mallas hexaédricas y tetraédricas, con resultados idénticos.

3.2. Ensayo Cilindro sometido a Presión Interna

En este caso, se emplea una geometría cilíndrica definida en la Tesis Doctoral de Claudio García Herrera [2]. Después de generar la malla, se establece un sistema de referencia local para cada elemento utilizando la función “*set axes from curvature*” de *FEBio*. Estos ejes son cruciales para definir las direcciones de las fibras de un material anisotrópico en una geometría curva. Tras ello, se configura el material Holzapfel-Gasser-Ogden, empleando un ángulo de 65° . *FEBio* define el

ángulo como la inclinación de las fibras con respecto al vector e_1 del sistema de referencia, que corresponde al vector circunferencial del cilindro. Por consiguiente, en este ensayo se ha adoptado un ángulo complementario a los 25° definidos por Claudio, para ser coherentes con la convención empleada en el software.

Tras esto, se aplica una carga superficial de 2 kPa en la cara interna del cilindro. El resultado obtenido es una curva de presión-alargamiento que muestra una tendencia similar a la de Claudio. Además, se verificó que los resultados eran consistentes con una malla tetraédrica.

3.2. Ensayo Cilindro Curvo sometido a Presión Interna

En este ensayo, la geometría empleada es un cilindro curvo. En este caso, la herramienta “*set axes from curvature*” detecta una curvatura incorrecta. Para solventarlo, se divide la geometría en distintas partes a las que se les aplica individualmente la función.

Por otro lado, el material empleado se define a partir de los parámetros de la Tabla 1, que corresponden a los datos obtenidos por Claudio García Herrera [2] para la aorta sana y aneurismática de pacientes mayores de 65 años del material GOH (Tabla 1).

	c (KPa)	k_1 (KPa)	k_2	γ ($^\circ$)
<i>Sana</i>	29,737	460,002	4,000	45,03
<i>Aneurismática</i>	15,105	172,368	0,593	42,79

Tabla 1. Valores del material GOH para las aortas.

Finalmente, se aplicó una carga superficial en la pared interna para simular la presión del flujo sanguíneo, utilizando una presión sistólica de 15 kPa. Los modelos quedan reflejados en la Figura 1.

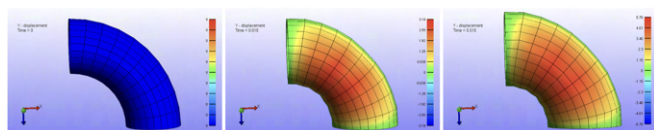


Figura 1: Mapa de color de desplazamientos en Y en un tubo curvo (a) en reposo, y tras aplicar una presión en el tejido aórtico (b) sano y (c) aneurismático.

El resultado puede verse en la Figura 2. Llama la atención que, a medida que aumenta la presión, las diferencias de alargamiento entre los dos materiales se hacen más notables. Esto se debe a que, a bajas deformaciones, la elastina domina el comportamiento del material antes de que las fibras de colágeno se orienten. Además, ambos ensayos muestran que la curva de la aorta aneurismática está por debajo de la otra, lo que indica que, al aplicar la misma presión, las deformaciones son mayores en la aorta aneurismática, lo que aumenta el riesgo de rotura. Estos resultados son coherentes al compararlos con los obtenidos en la Tesis de Claudio [2]; y también lo fueron los realizados con la malla tetraédrica.

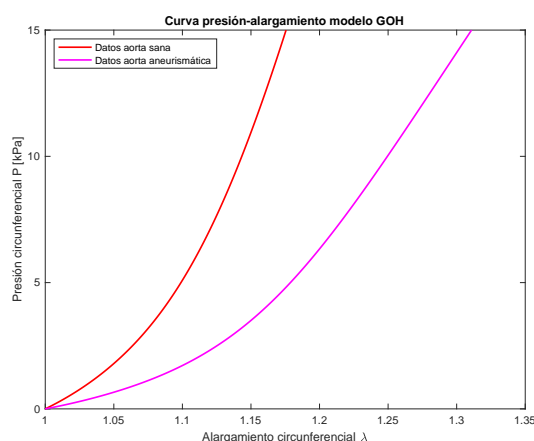


Figura 2: Resultados del ensayo de presión interna mediante una malla hexaédrica.

4. Aplicación al caso de aortas reales: modelo y resultados.

En este apartado se definen los tres ensayos realizados en las geometrías aórticas reales: en condiciones fisiológicas y dos en caso de accidente. Tras el proceso definido en la Sección 2, se definen los ejes de referencia locales en cada elemento del material, aplicando la función “*set axes from curvature*” a cada parte de la aorta. En las Figuras 3 y 4 la dirección roja corresponde a la dirección circunferencial de la arteria, la verde representa la dirección longitudinal y la azul es la radial.

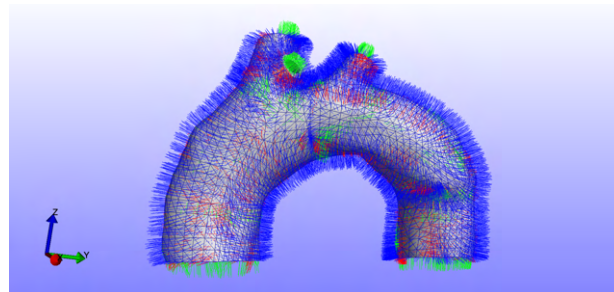


Figura 3: Sistema de referencia local de la aorta sana.

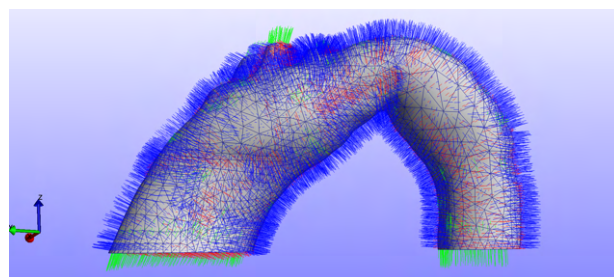


Figura 4: Sistema de referencia local de la aorta aneurismática.

El material seleccionado es el GOH con los datos de la Tabla 1. Se fijan los extremos en las direcciones X , Y y Z , manteniendo libres los troncos supra-aórticos. Se aplicó una carga superficial en la pared interna de \approx

5 kPa, diferencia entre presión sistólica y diastólica, ya que las imágenes de TC se obtuvieron en diástole.

Para la segunda simulación, se define un desplazamiento en Z del tronco braquiocefálico, aplicado tras el presurizado completo. Como añadido, se realiza una tercera simulación en la que se aumenta en un 50 % el desplazamiento anterior, para así poder comparar la respuesta del material ante muy grandes deformaciones.

4.1. Resultados del Modelo en Condiciones Fisiológicas

Se representa un mapa de colores que muestra el “*Max Shear Stress*”, que son las tensiones máximas calculadas a partir de las tensiones principales. Este criterio permite localizar los puntos en los que la arteria corre más riesgo de rotura. Ambos modelos de arteria coinciden en que el punto más propenso es cercano al nacimiento de los troncos, como se representa en las Figuras 5 y 6.

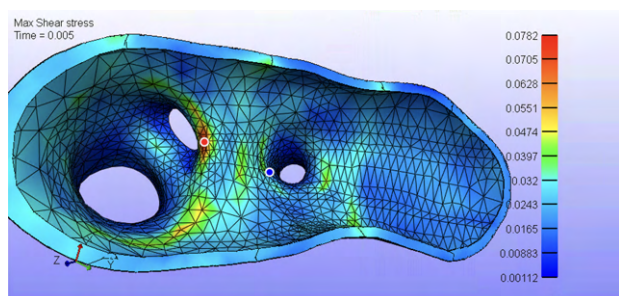


Figura 5: Punto de tensión máxima del modelo de aorta sana.

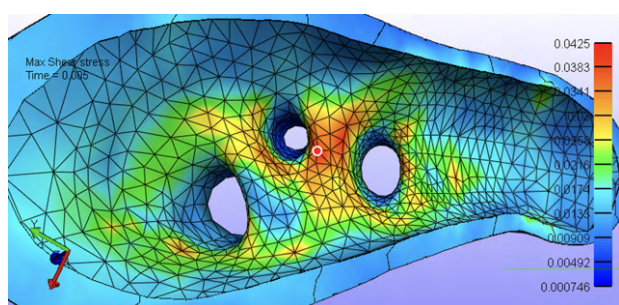


Figura 6: Punto de tensión máxima del modelo de aorta aneurismática.

Sin embargo, según se indica en la Tabla 2, las tensiones de rotura se definen en dirección circunferencial. Por ello, se emplea la relación $\sigma_\theta = \mathbf{a} \cdot \boldsymbol{\sigma} \cdot \mathbf{a}^T$, para obtener la gráfica 7 y el índice de rotura ($\mathbf{IR} = \frac{\sigma_{m\acute{a}x}}{\sigma_{rot}}$).

	$\sigma_{rot} (kPa)$
Aorta Sana	968 ± 215
Aorta Aneurismática	937 ± 219

Tabla 2. Tensiones circunferenciales de rotura de aortas en personas mayores de 65 años [2].

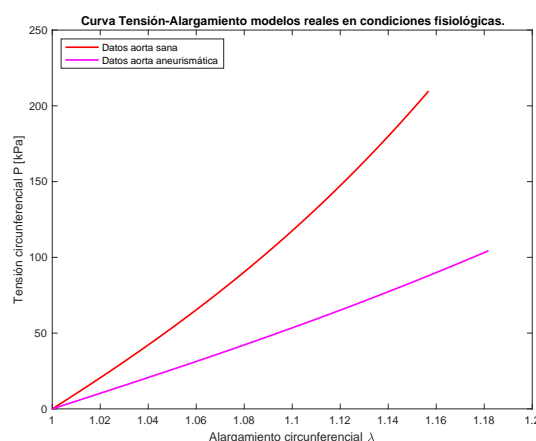


Figura 7: Curvas tensión - alargamiento circunferencial de los modelos de aortas en condiciones fisiológicas.

Los resultados obtenidos son coherentes con lo esperado, pues, ante una misma tensión, el material que simula las paredes aneurismáticas sufre mayores alargamientos.

El índice de rotura para la aorta sana es 0.21669 y para la aneurismática 0.1114. Ambos valores indican que ningún modelo sufren riesgo de rotura.

4.2. Resultados del Modelo en Condiciones Traumáticas

En este apartado, de nuevo se observa que los puntos críticos de estudio se encuentran en los troncos supra-aórticos, como puede verse en las Figuras 8 y 9.

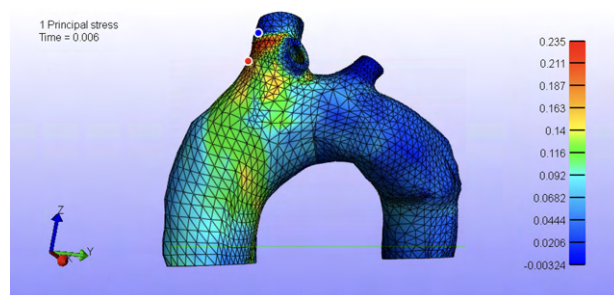


Figura 8: Tensiones principales máximas del modelo de aorta sana en condiciones traumáticas.

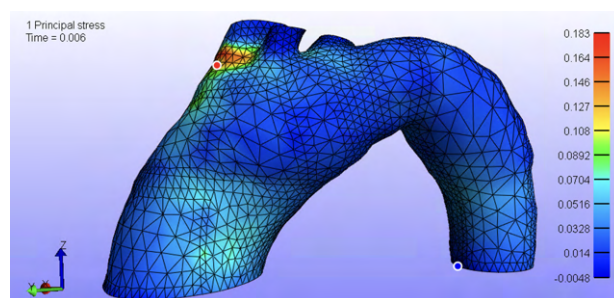


Figura 9: Tensiones principales máximas del modelo de aorta aneurismática en condiciones traumáticas.

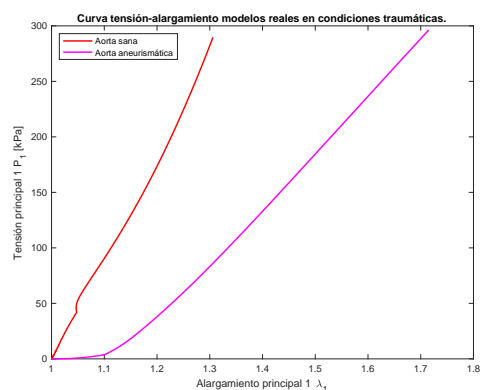


Figura 10: Curvas tensión - alargamiento circunferencial de los modelos de aortas en condiciones traumáticas.

En la Figura 10 se observa que la aorta aneurismática experimenta mayores alargamientos que la sana ante las mismas tensiones. También se puede identificar claramente el cambio en las curvas que indica la transición entre la presurización y el desplazamiento del tronco, debido a la redistribución de tensiones en respuesta a condiciones traumáticas.

Bajo condiciones traumáticas, los índices de rotura se incrementan a 0.2982 para la aorta sana y 0.2878 para la aneurismática debido al aumento de las tensiones en las paredes arteriales, lo que resulta en un mayor riesgo de ruptura en contraste con las condiciones fisiológicas.

Tras esto, se aumenta en un 50 % el desplazamiento en Z del tronco braquiocefálico para poder comparar los efectos mecánicos de las direcciones de las fibras. Los resultados mostrados en la Figura 11 nuevamente muestran coherencia, ya que la aorta sana tiene una mayor capacidad para soportar tensiones sin experimentar alargamientos excesivamente altos.

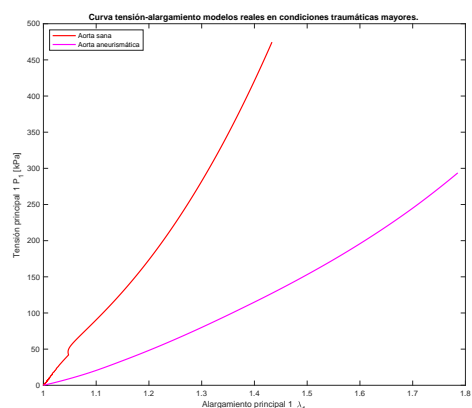


Figura 11: Curvas tensión - alargamiento circunferencial en condiciones traumáticas mayores.

Se obtiene un índice de rotura de 0.429 y 0.446 para la aorta sana y aneurismática respectivamente. Cabe destacar que, aunque hayan aumentado ambos índices,

el riesgo de rotura de la aorta aneurismática supera al de la aorta sana cuando, hasta el momento, se había mantenido por debajo. Esto demuestra la influencia del comportamiento de las fibras de colágeno ante altas tensiones. La estructura de las fibras de colágeno se ve comprometida en la aorta aneurismática, provocando una disminución de la capacidad de las paredes arteriales para soportar cargas y tensiones.

5. Conclusiones

En este estudio, se ha confirmado la correlación entre la orientación de las fibras de colágeno y las propiedades mecánicas de las paredes aórticas. Además, ante tensiones elevadas, es más evidente la pérdida de estas propiedades debido al deterioro de la capacidad de reclutamiento del colágeno. Por ello, se confirma que el riesgo de rotura no depende solo del diámetro de la arteria. Esto tiene importantes implicaciones para la toma de decisiones clínicas.

Además, también se concluye que las direcciones principales asociadas a las deformaciones reflejan el cambio en las direcciones de las fibras en función de las condiciones a las que se somete la aorta. En la Tabla 3, se recogen los ángulos obtenidos respecto a la dirección circunferencial de la proyección del autovector asociado a la deformación principal 1 en el plano que forman las direcciones circunferencial y longitudinal del elemento de máxima deformación de cada modelo.

	Fisiológico	Trauma 1	Trauma 2
<i>Sana</i>	6,45°	60,21°	82,15°
<i>Aneurismática</i>	5,63°	66,81°	79,86°

Tabla 3. Ángulos respecto a la circunferencial.

Cabe destacar que el ángulo incrementa en las condiciones de accidente respecto a las fisiológicas. Esto se debe a que se produce un desplazamiento longitudinal en el que las fibras se alinean en dicha dirección. Además, también aumenta cuando se realiza el ensayo con un desplazamiento 50 % mayor. Esto refleja que el alineamiento de las fibras de colágeno en las paredes aórticas es progresivo.

Referencias

- [1] A. Poduri, “Cellular Mechanisms of Ascending Aortic Aneurysms.” *New Approaches to Aortic Diseases from Valve to Abdominal Bifurcation*, pp. 79–84, 2019.
- [2] C. García Herrera, “Comportamiento mecánico de la aorta ascendente: caracterización experimental y simulación numérica”, doctoral thesis, E.T.S.I. Caminos, Canales y Puertos (UPM), 2008.
- [3] T. Gasser, R. W. Ogden, y G. Holzapfel, “Hyperelastic modelling of arterial layers with distributed collagen fibre orientations”, *Journal of the Royal Society Interface*, vol. 3, pp. 15–35, February 2006.
- [4] H. Fehervary, L. Maes, J. Vastmans, G. Kloosterman, y N. Famaey, “How to implement user-defined fiber-reinforced hyperelastic materials in finite element software”, *Journal of The Mechanical Behavior of Biomedical Materials*, vol. 110, October 2020.
- [5] FEBio User Manual, “4.1.2.8 Holzapfel-Gasser-Ogden”, [Online].

Clasificación de movimientos del tren superior con visión artificial y aprendizaje profundo en adulto sano y paciente pediátrico con parálisis cerebral

C. De Gregorio Bermejo¹, A. Romo Marín¹, B. De la Calle²,

C. Simón-Martínez³, M. Martínez-Zarzuela¹

¹ Departamento de Teoría de la Señal y Comunicaciones e Ingeniería Telemática, Universidad de Valladolid, Valladolid, España, {carla.gregorio, angela.romo }@estudiantes.uva.es, mario.martinez@uva.es

² Servicio de Rehabilitación, Hospital Universitario Río Hortega, Valladolid, España, bcalleg@saludcastillayleon.es

³ Institute of Informatics, University of Applied Sciences Western Switzerland (HES-SO) Valais-Wallis, Sierre, Switzerland, cristina.simon@hevs.ch

Resumen

Las escalas clínicas estándar se utilizan para valorar de forma cualitativa la funcionalidad del miembro superior en pacientes pediátricos con parálisis cerebral. Una alternativa para valorar cuantitativamente dichos pacientes es el empleo de ángulos articulares obtenidos mediante el uso de sistemas de adquisición con sensores inerciales o visión artificial. En este trabajo se ha adquirido una base de datos de movimientos de miembro superior empleando sensores XSens y software de visión OpenCap en adulto sano y paciente pediátrico, seleccionando las actividades en base a las escalas estándar empleadas por el equipo rehabilitador. Se han estudiado redes de aprendizaje profundo para segmentar de forma automática los movimientos y las fases de un movimiento más complejo, combinación de movimientos unimanuales y bimanuales.

Los resultados muestran que es posible reconocer movimientos de la vida diaria y que es posible segmentar las fases de una actividad de miembro superior tanto en sujetos sanos como patológicos, obteniendo una F1-score superior al 80%. Estos resultados sugieren que las redes neuronales desarrolladas son capaces de clasificar y segmentar los movimientos propuestos de miembro superior en sujeto sano y paciente pediátrico con parálisis cerebral.

1. Introducción

1.1. Parálisis cerebral

La parálisis cerebral es la causa más frecuente de discapacidad infantil, afectando a 2-3/1000 recién nacidos y hasta 40-100/1000 niños prematuros [1]. Se describe como un conjunto de trastornos que afectan al desarrollo del movimiento y postura del paciente. Se caracteriza por una alteración de la coordinación, así como de la regulación del tono muscular, debido a la espasticidad de los músculos [1] [2]. Resulta de una lesión en el cerebro en desarrollo que suele ocurrir a una edad temprana. No se considera un trastorno progresivo, sin embargo, a medida que el cerebro madura la clínica puede cambiar [2] [3].

Podemos diferenciar numerosos tipos de parálisis cerebral, no obstante, nos centraremos en la parálisis cerebral espástica, siendo el tipo más común entre los pacientes pediátricos, afectando al 80% de ellos [4]. Se caracteriza por presentar los músculos rígidos, causando movimientos abruptos. Dentro de este tipo, dependiendo de los miembros afectados, diferenciamos otros subtipos,

destacando entre ellos, la hemiplejía o hemiparesia, la cual afecta a un lado del cuerpo.

La investigación y clínica sobre el análisis del movimiento en estos pacientes se centra fundamentalmente en la marcha humana. Sin embargo, el 83% de los niños con parálisis cerebral presentan afectación del miembro superior [1] [5]. En estudios científicos, para la valoración objetiva de los pacientes se emplean escalas estándar de valoración como The Melbourne Assessment y The Assisting Hand Assessment. La primera evalúa de forma unilateral las extremidades de los niños, constando de un total de 14 tareas, que se puntúan de 3 a 5 puntos, siendo 5 el máximo y 3 un deterioro más severo [6]. Para la segunda escala, se valoran tareas bimanuales en niños que presentan una discapacidad unilateral. El objetivo es medir el uso efectivo de la mano afectada paralelamente con la mano asistente [7]. Una limitación de las escalas anteriores es que son dependientes de la experiencia del examinador y no se realizan mediciones articulares precisas durante su aplicación. En otros estudios, se emplean sistemas con multicámara y marcadores (Vicon u Optitrack), que representan el *gold standard* en análisis biomecánicos [8]. Sin embargo, estos sistemas apenas se utilizan en la práctica clínica debido a su alto precio, su baja disponibilidad y los altos requerimientos de espacio y de especialización para su utilización.

Por este motivo, es interesante la búsqueda de una alternativa a las escalas y los sistemas anteriores para evaluar dichos pacientes y que permita simplificar y objetivar las valoraciones. Este trabajo de investigación se enmarca en el estudio de sistemas de reconocimiento de movimientos con adquisición de movimiento asequibles que pueden emplearse para instrumentalizar estas escalas de valoración con Inteligencia Artificial. En nuestro estudio empleamos sensores inerciales y sistemas de visión artificial de bajo coste para adquirir una base de datos de movimientos de miembro superior y realizar una posterior clasificación automática de movimientos y segmentación de las fases de un movimiento complejo a través de un sistema de aprendizaje profundo.

2. Materiales y métodos

2.1. Sensores inerciales: XSens Awinda

Los sensores inerciales son pequeños dispositivos capaces de medir orientaciones combinando la información procedente de un acelerómetro, un giróscopo y un magnetómetro [9]. Hemos empleado el sistema XSens Awinda, compuesto por hasta 17 unidades de sensores inerciales. El movimiento de los miembros superiores se captura con 11 sensores. Estos se colocan a través de una camiseta especial y un juego de correas de velcro. En este estudio se colocaron sensores en la parte posterior de la cabeza, en el esternón, en ambas escapulas, en la pelvis a la altura de la cresta ilíaca, en la parte superior del brazo, en el antebrazo y en las manos [9] [10].

La calibración del sistema requiere conocer al menos la altura y la longitud del pie del sujeto. Posteriormente, se procede a la calibración. Es un proceso guiado por el propio programa, comenzando en N-pose o pose neutral, es decir, la persona de pie y erguida, con las piernas juntas y los brazos pegados al cuerpo. Después, el programa pide al sujeto que camine durante 10-15 segundos, hasta que solicite volver a la pose neutral, finalizando en dicha posición.

2.2. Visión artificial: OpenCap

OpenCap es un software de código abierto y bajo coste, desarrollado por la Universidad de Stanford. Su funcionalidad consiste en estimar la cinemática 3D y la dinámica del movimiento humano, a partir de videos capturados por varias cámaras [12]. En las grabaciones con *OpenCap* se ha empleado dos dispositivos tipo iPad con sistema operativo iOS. Dichos iPads se dispusieron en unos trípodes a 30-45 grados respecto a la línea en la que se ubicaba el sujeto a adquirir. El sistema permite emparejar las cámaras escaneando un código QR y su calibración empleando un tablero de ajedrez. Posteriormente, se introducen las medidas del sujeto: la altura, el peso y la edad. Nuevamente, se utilizará una pose para calibrar al sujeto de interés. En este caso, el sujeto se dispone de pie, erguido, con los pies paralelos separados y los brazos sutilmente abiertos formando un triángulo con el tronco.

2.3. Base de datos

Se ha adquirido una base de datos compuesta por un total de 17 sujetos controles, de edades entre 7-22 años y 8 sujetos pediátricos con parálisis cerebral, de edades de 5-16 años. Se recopilaron los siguientes datos: la edad, el género, la altura, el peso, el número de pie, y el pie y la mano dominante y de cada uno de los participantes.

Los criterios de inclusión de los sujetos pediátricos fueron un diagnóstico previo de parálisis cerebral y poseer la capacidad mínima para agarrar activamente un objeto. Los criterios de exclusión fueron haber sido sometidos a una cirugía previa de miembro superior o a inyecciones de toxina botulínica-A, un mes antes de las grabaciones.

El estudio se realizó con el dictamen favorable del comité de ética de la investigación con medicamentos CEIM ÁREA DE SALUD VALLADOLID ESTE en febrero 2023 para los sujetos controles y del comité de ética de la

investigación con medicamentos CEIM ÁREA DE SALUD VALLADOLID OESTE en marzo 2023 para los sujetos pediátricos con parálisis cerebral. Cada sujeto, así como los padres de los niños recibieron información oral y escrita sobre el estudio antes de participar. Además, todos los datos personales fueron tratados según lo dispuesto en el Reglamento (UE) 2016/679, de abril, General de protección de Datos Personales y su normativa de desarrollo. Las adquisiciones tuvieron lugar en la Escuela Técnica Superior de Ingenieros de Telecomunicaciones de la Universidad de Valladolid y en el Hospital Universitario Río Hortega.

3. Resultados y discusión

Se realizaron dos bloques de pruebas: un primer bloque de reconocimiento de actividades de la vida diaria de los pacientes y un segundo bloque de segmentación de fases del movimiento: “*Open/Close a bottle and pour water*”, un movimiento complejo de miembro superior, que combina movimientos unimanuales y bimanuales.

3.1. Bloque I: Reconocimiento de actividades de la vida diaria

En primer lugar, la selección de movimientos se llevó a cabo apoyándonos en las escalas estándar que emplean los fisioterapeutas para evaluar de forma cualitativa la funcionalidad del miembro superior en los niños. Se quiso incluir tanto movimientos unimanuales como bimanuales. En los movimientos unimanuales, la escala Melbourne basada en una selección de 14 tareas que suponen al niño alcanzar, agarrar, soltar y manipular objetos simples. Entre ellos seleccionamos una actividad simple “*Reach to grasp and approach*” (RTG), buscando la extensión y flexión del brazo, y una más compleja “*Turn pages*” (BOOK), para la pronación y supinación del miembro superior. En los movimientos bimanuales, utilizamos la escala Assisting Hand, que evalúa de forma paralela la mano afectada con la mano asistente. En este caso, adaptamos la actividad “*Put on a hat*” (HAT), para visualizar la simetría entre ambos brazos. Por último, la actividad “*Open/Close a bottle and pour water*” (BOTTLE), suma de movimientos unimanuales y bimanuales. Las capturas fueron grabadas a 60 Hercios (Hz). Tras las grabaciones, se acondicionaron los archivos proporcionados por el software de *OpenCap* y se generaron ventanas temporales de 60 muestras con un solapamiento de 30, es decir, en tamaños de ventana de un segundo. Las secuencias temporales incluyeron un total de 10 ángulos articulares de los brazos: “*arm_flex_r*”, “*arm_flex_l*”, “*arm_add_r*”, “*arm_add_l*”, “*arm_rot_r*”, “*arm_rot_l*”, “*elbow_flex_r*”, “*elbow_flex_l*”, “*pro_sup_r*” y “*pro_sup_l*”.

En los entrenamientos se utilizaron redes neuronales convolucionales, muy empleadas para el reconocimiento de movimiento humano y las técnicas de regularización: *DropOut* y *Batch Normalization*, con el objetivo de evitar el *overfitting*. Tras numerosas pruebas, observamos que la red propuesta en la Tabla 1 es capaz de reconocer los movimientos de los sujetos controles, obteniendo un F1-score de 0.858. Podemos ver con más detalle en la matriz

de confusión de la Figura 1.A que la red reconoce bien las actividades propuestas. Sin embargo, en los entrenamientos juntando todos los sujetos, se obtuvo un F1-score de 0.753, como observamos en la Tabla 1. La métrica en este caso es inferior, dado que las actividades etiquetadas como “BOOK” y “BOTTLE”, suponían cierta dificultad a los niños, al ser movimientos más complejos. Esto se ve reflejado en la Figura 1.B, en la que la red genera una alta tasa de falsos negativos en dichas etiquetas.

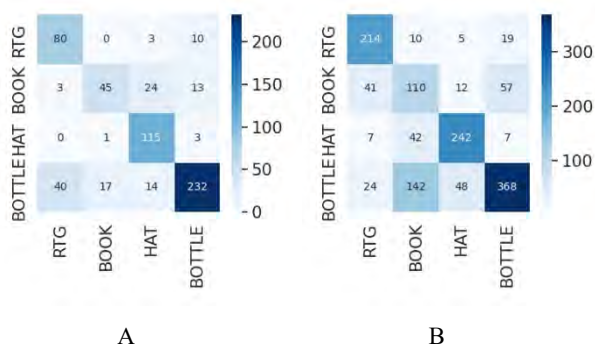


Figura 1. Matrices de confusión de los resultados del Bloque I (A. Sujetos controles y B. Todos los sujetos)

3.2. Bloque II: Segmentación de fases del movimiento: “Open/Close a bottle and pour water”

Este bloque de pruebas se centra en la actividad: “Open/Close a bottle and pour water”, una actividad compleja de miembro superior, que consta de fases unimanuales y bimanuales, consiste en agarrar una botella de agua, abrirla, verter agua en un vaso, cerrar la botella y volver a dejarla sobre la mesa. En primer lugar, se fijaron las fases de dicha actividad, diferenciando entre:

- Alejar/Acercar: el brazo se acerca o se aleja hacia la botella.
- Transferencia: de la botella de una mano a otra.
- Sujetar: la botella con una mano mientras con la otra abre o cierra la botella.
- Abrir/Cerrar: con una mano mientras con la otra sujeta la botella.
- Verter: consiste en echar agua con una mano en el vaso.

En un primer momento se consideraron 5 fases, sin embargo, durante los entrenamientos, se eliminó la fase de “Transferencia”. Durante las grabaciones con los sujetos pediátricos se observó, que los niños se sentían más seguros vertiendo agua con dos manos, por lo que se incluyó la etiqueta “2Manos”, para los entrenamientos con los sujetos pediátricos. El etiquetado de los videos se realizó mediante una interfaz desarrollada para dicho propósito, que facilitaba también la adecuada visualización de los videos. Las etiquetas se anexionaban a los .mot proporcionados por *OpenCap*. Posteriormente, el acondicionamiento de archivos es muy similar al Bloque I de pruebas: los archivos .mot se transforman a .csv y se dividen en ventanas de tamaño 30 muestras con un

solapamiento de 15, al ser fases de actividades, se consideraron tamaños de ventana más pequeños. Nuevamente utilizamos los 10 ángulos articulares del brazo y añadimos los pertenecientes a la espalda: “lumbar_extension”, “lumbar_bending” y “lumbar_rotation”, debido a cierta inclinación de los sujetos a la hora de coger la botella y el vaso de la mesa.

Se realizaron numerosos experimentos variando el tamaño de ventana, buscando el tamaño óptimo para las fases de esta actividad. Demostrando que los mejores resultados los proporcionaba las ventanas de tamaño 45 con un solapamiento de 30 muestras. Por otra parte, quisimos comprobar la importancia de los ángulos de las lumbares, verificando que se conseguían mejores resultados sin considerar dichos ángulos articulares. En cuanto a los entrenamientos utilizando los sujetos pediátricos, se consideraron las etiquetas de “Sujetar”, “Abrir/Cerrar” y “2Manos”, obteniendo los resultados mostrados en la Figura 2.A. En los entrenamientos teniendo en cuenta todos los sujetos capturados, obtuvimos buenos resultados, los podemos visualizar en la Tabla 1.B, verificando que la red propuesta es capaz de reconocer el movimiento de verter, realizándose con una mano o con las dos manos.

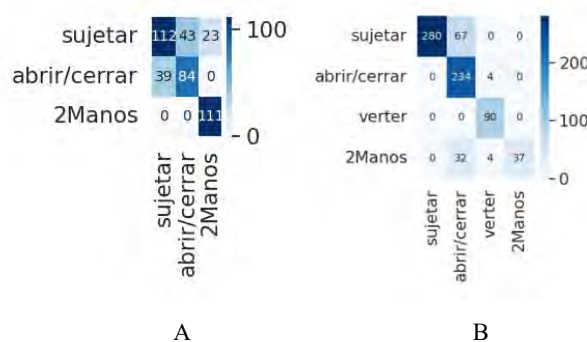


Figura 2. Matrices de confusión de los resultados del Bloque II (A. Sujetos pediátricos y B. Todos los sujetos)

4. Conclusiones

En este estudio se ha alcanzado el objetivo de reconocer movimientos de miembro superior relacionados con la vida diaria en pacientes pediátricos con parálisis cerebral, empleando dispositivos de adquisición de vídeo y redes neuronales de aprendizaje profundo.

En un primer bloque de experimentos, se consiguió reconocer actividades de la vida diaria con un F1-score de 87% para los sujetos control y un 75% considerando a todos los sujetos. En un segundo bloque de experimentos, se probaron redes neuronales para la segmentación automática de las fases de un movimiento más complejo, que comprendía fases unimanuales y bimanuales. En las pruebas realizadas, se consiguió un F1-score de 83%. Este resultado demuestra el potencial de sistemas de vídeo y de la red neuronal propuesta para clasificar los movimientos tanto con sujetos controles como sujetos pediátricos con parálisis cerebral.

Pruebas	Tipo de sujeto	Red neuronal			Resultados obtenidos set de test			
		Nº de capas	Parámetros	Técnicas de regular.	Sensitivity	Specificity	Accuracy	F1-score
BLOQUE I: RECONOCIMIENTO DE ACTIVIDADES	CONTROL	7 CNN: 16, 32, 64, 128, 256, 128, 64 + GAP + Dense 64	1.215.976	DropOut+ Batch Normalization + Regular. L1 y L2	0.853	0.957	0.876	0.858
	CONTROL Y PEDIÁTRICO	5 CNN: 16, 32, 64, 128, 256 + GAP + Dense 128, 64	738.556	DropOut+ Batch Normalization + Regular. L1 y L2	0.771	0.869	0.745	0.753
BLOQUE II: SEGMENTACIÓN DE FASES	CONTROL Y PEDIÁTRICO	5 CNN: 16, 32, 64, 128, 256 + GAP + Dense 128, 64	738.556	DropOut+ Batch Normalization + Regular. L1 y L2	0.824	0.948	0.855	0.836

Tabla 1. Resultados obtenidos, redes empleadas y sujetos utilizados en los experimentos.

Una de las principales limitaciones de este estudio, es el reducido número de sujetos patológicos en la base de datos. Adicionalmente, algunos resultados obtenidos sugieren que sería beneficioso incluir más sujetos adultos sanos en la muestra. Debido a lo anterior, se está trabajando ya en aumentar la base de datos y en incorporar adquisiciones de sujetos pediátricos sanos, con la finalidad de poder realizar estudios comparativos entre sujetos pediátricos sanos y patológicos.

Agradecimientos

Esta investigación ha sido financiada parcialmente por proyectos del Ministerio de Ciencia e Innovación (1245150A-100), la Gerencia Regional de Salud de Castilla y León (GRS 2670/A/22) y la Consejería de Empleo e Industria de Castilla y León (INVESTUN/21/VA/0003). C. Simón-Martínez cuenta con el apoyo del acuerdo de subvención Marie Sklodowska-Curie (890641).

Referencias

[1] Elena Beani et al, «Actigraph assessment for measuring upper limb activity in unilateral cerebral palsy | SpringerLink», 2019. <https://link.springer.com/article/10.1186/s12984-019-0499-7> (accedido 24 de marzo de 2023).

[2] S. D. Benedetto *et al.*, «Evaluation of the factors that impact upper limb coordination in children with cerebral palsy: A narrative review», *Retos*, vol. 48, pp. 470-480, mar. 2023, doi: 10.47197/retos.v48.96756.

[3] M. J. Menor-Rodríguez, M. Sevilla Martín, J. C. Sánchez-García, M. Montiel-Troya, J. Cortés-Martín, y R. Rodríguez-Blanque, «Role and Effects of Hippotherapy in the Treatment of Children with Cerebral Palsy: A Systematic Review of the Literature», *J. Clin. Med.*, vol. 10, n.º 12, p. 2589, jun. 2021, doi: 10.3390/jcm10122589.

[4] P. P. Argüelles, «Parálisis cerebral infantil», 2012.

[5] M. Cacioppo *et al.*, «A new child-friendly 3D bimanual protocol to assess upper limb movement in children with unilateral cerebral palsy: Development and validation», *J.*

Electromyogr. Kinesiol., vol. 55, p. 102481, dic. 2020, doi: 10.1016/j.jelekin.2020.102481.

[6] H. Park *et al.*, «Relationship between the more-affected upper limb function and daily activity performance in children with cerebral palsy: a cross-sectional study», *BMC Pediatr.*, vol. 21, p. 459, oct. 2021, doi: 10.1186/s12887-021-02927-2.

[7] P. Casares, «Escala AHA (Assisting Hand Assessment) para valoración acciones bimanuales espontáneas», *INEAVA*, 2015. <https://www.ineava.es/blog/infantil/escala-aha-assisting-hand-assessment-para-valoracion-acciones-bimanuales-espontaneas> (accedido 27 de marzo de 2023).

[8] A. M. Valevicius, P. Y. Jun, J. S. Hebert, y A. H. Vette, «Use of optical motion capture for the analysis of normative upper body kinematics during functional upper limb tasks: A systematic review», *J. Electromyogr. Kinesiol.*, vol. 40, pp. 1-15, jun. 2018, doi: 10.1016/j.jelekin.2018.02.011.

[9] J. Castellanos-Ruiz, L. M. Montealegre-Mesa, B. D. Martínez-Toro, J. J. Gallo-Serna, y O. Almanza-Fuentes, «Uso de sensores inerciales en fisioterapia: Una aproximación a procesos de evaluación del movimiento humano», *Univ. Salud*, vol. 23, n.º 1, pp. 55-63, dic. 2020, doi: 10.22267/rus.212301.214.

[10] Movella, «Wireless Human Motion tracker for real-time applications», 2021. <https://www.movella.com/products/wearables/xsens-mtw-awinda> (accedido 14 de abril de 2023).

[11] Movella, «Sensor Placement in Xsens Awinda System», 17 de enero de 2022. https://base.xsens.com/s/article/Sensor-Placement-in-Xsens-Awinda-System?language=en_US (accedido 14 de abril de 2023).

[12] S. D. Uhlrich *et al.*, «OpenCap: 3D human movement dynamics from smartphone videos». bioRxiv, p. 2022.07.07.499061, 10 de julio de 2022. doi: 10.1101/2022.07.07.499061.

Desarrollo de nuevos efectores finales para robots asistivos en la rehabilitación de miembro superior

J. Campello ¹, G.J. Garcia-Gomez ¹, J.L. Ramon ¹, A. Ubeda ¹, C.A. Jara ¹

¹ Human Robotics Group, Universidad de Alicante
andres.ubeda@ua.es

Resumen

La mayoría de los sistemas de rehabilitación robótica de miembro superior presentan mangos fijos que impiden la movilidad de la mano y requieren de una fuerza de agarre suficiente generando mucha incomodidad en pacientes con espasticidad y rango articular de muñeca limitado. En este artículo se presentan distintos tipos de efector final que tratan de reducir las limitaciones de los efectores convencionales. Para ello, se describen las condiciones de diseño y la evaluación preliminar de su uso por parte de usuarios sin limitaciones motoras a partir de la realización de distintos movimientos sencillos con diferente nivel de fuerza. Las respuestas de los usuarios muestran una preferencia por el mango ergonómico y un efector dotado de una articulación esférica que permite mayor movilidad de la muñeca.

1. Introducción

La pérdida súbita de la función motora del miembro superior es una condición clínica que generalmente es causada por daño cerebral, trauma a nivel de la columna u otras enfermedades neuromusculares. La recuperación funcional es variable y específica de cada paciente y depende de la gravedad de la lesión y del tipo de patología. En el caso de los pacientes con Daño Cerebral Adquirido (DCA) [1], que en su mayoría sufren un ictus, la recuperación es crucial durante la fase subaguda, que se produce entre uno y seis meses después del incidente.

Durante los últimos años, la rehabilitación convencional ha experimentado un importante desarrollo en herramientas complementarias para aumentar la eficacia de la terapia. Una de las principales formas de proporcionar una mejor y más rápida recuperación es el uso de dispositivos robóticos [2]. En el caso de la rehabilitación de miembros superiores, los robots de efector final han proporcionado una forma de aumentar la intensidad y la repetibilidad de los protocolos de rehabilitación [3].

En los últimos años, los robots colaborativos se están utilizando ampliamente para la rehabilitación, particularmente para pacientes con discapacidades en las extremidades superiores [4]. Estos robots están especialmente preparados para la interacción humano-robot, pudiendo asistir el movimiento en diferentes modalidades dependiendo de la movilidad de las extremidades (pasiva, activa, activo-asistida, etc). En estos casos, la interacción humano-robot está mediada por herramientas de efectores finales. El punto de conexión del efector final del robot está conectado a la extremidad del paciente y puede guiarlo por un camino fijo o aplicar un control asistido según sea necesario para la terapia de rehabilitación [5]. Estos sistemas suelen utilizar mangos, agarrados por el paciente como efectores finales robóticos, para generar

un movimiento de la extremidad en el espacio. Sin embargo, en la mayoría de los casos, estos efectores finales no permiten la movilidad de la mano ni el agarre, lo que reduce las posibilidades de los ejercicios de rehabilitación [6]. Este hecho tiene especial relevancia durante la interacción entre los usuarios con daño cerebral y el robot ya que, a menudo, estos pacientes tienen una fuerza de agarre disminuida o un grado de espasticidad suficiente como para impedir el agarre en el plano sagital por una elevada pronación de la muñeca. Por este motivo, en este artículo se proponen una serie de nuevos efectores finales para poder ser empleados por robots manipuladores tipo colaborativo con el objetivo de obtener herramientas terminales más ergonómicas y flexibles, adaptadas específicamente a las necesidades del paciente para apoyar la rehabilitación de su miembro superior.

2. Desarrollo de prototipos

Los dos prototipos presentados buscan principalmente disminuir el esfuerzo del paciente en el agarre del efector final y garantizar la movilidad de la muñeca en los distintos planos de movimiento. A continuación se describen los aspectos principales de diseño de ambos prototipos.

2.1. Estructura principal

Los prototipos se componen de PLA (ácido poliático), un material termoplástico utilizado comúnmente para la impresión 3D debido a sus propiedades como una temperatura de fusión baja o una buena resistencia. La estructura se divide en tres partes principales: la primera es la “cabeza” del efector, donde encaja el mango y se produce el agarre con la mano por parte del usuario. La segunda parte es el “cuerpo” del efector donde, con la ayuda de varias correas se permite la colocación y sujeción del efector a la parte del antebrazo más próxima a la muñeca. La última parte es el “cuello” del efector, que conecta la “cabeza” con el “cuerpo” y proporciona robustez.

2.2. Primer prototipo

El objetivo del primer prototipo es diseñar la unión del “cuello” con el “cuerpo” del efector para que estén conectados de alguna forma que les permita rotar circularmente cierta distancia (ver Figura 1). De esta manera, el agarre no tiene por qué ser exclusivamente vertical, lo cual es bueno para personas que no puedan mantener una posición de este tipo por mucho tiempo y les sea más cómodo inclinar un poco su agarre, como en el ca-

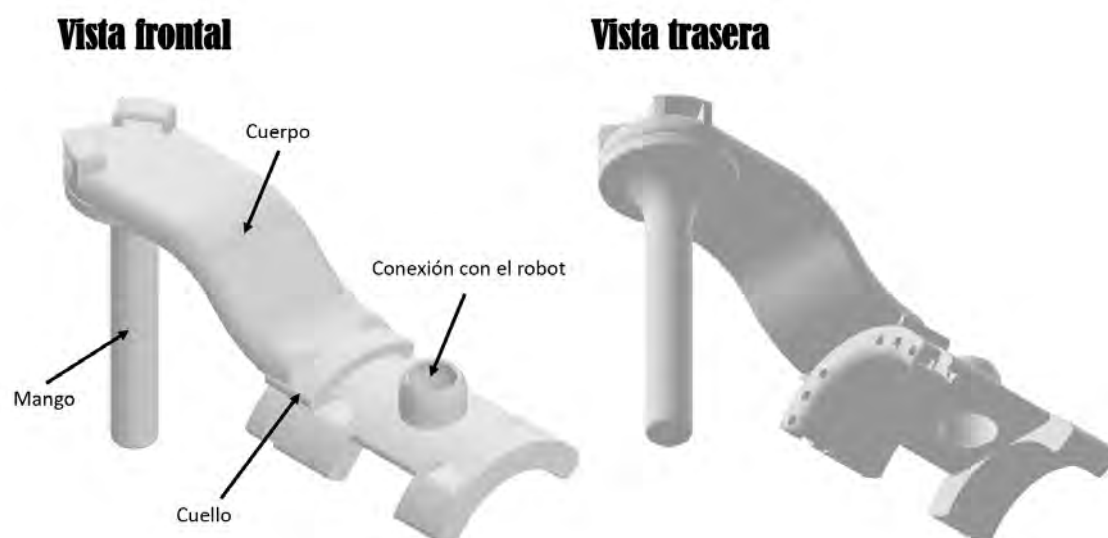


Figura 1. Primer prototipo con conexión de riel.

so de los pacientes que una pronosupinación de muñeca leve.

Para permitir un movimiento rotatorio circular se ha implementado un sistema de hendiduras para dar lugar a un riel complementario. Es necesario partir de una porción circular para que la rotación sea posible. Por último, se han implementado unas varillas para, además de permitir el movimiento rotacional, poder fijar la posición de la “cabeza” en el centro de forma vertical, ladeado hacia la derecha o hacia la izquierda.

2.3. Segundo prototipo

El objetivo del segundo prototipo es diseñar la la unión del “cuello” con el “cuerpo” del efector de forma que se simule el movimiento de la propia muñeca (ver Figura 2). Mediante una “replica” de una articulación esférica es posible dotar al modelo un movimiento libre, y en cualquier orientación del espacio. Este movimiento tridimensional tiene sus limitaciones, aunque es más que suficiente para pacientes que puedan tener algún tipo de parálisis, atrofia o debilidad en la mano o una excesiva pronosupinación por espasticidad.

Para permitir la rotación de una superficie esférica sobre una esfera, es necesario que esta cubra más de la mitad de la misma, de lo contrario, se podrían desacoplar fácilmente durante la realización de algún movimiento. Se mantiene el sistema de hendiduras, en este caso, para dar lugar a un saliente complementario con forma esférica.

3. Validación de prototipos

Una vez finalizado el diseño e impresión de los diferentes prototipos de efector final propuestos, se procede a su evaluación. Para ello se han seleccionado movimientos simples relacionados con el hombro. Las pruebas se han

llevado a cabo inicialmente en 7 usuarios sanos (con edades de 21.0 ± 1.1 años), utilizando un robot UR3 como herramienta. El robot UR3 (Universal Robots) es un robot manipulador colaborativo ideal para la interacción cercana con humanos y, por tanto, adecuado para tareas de rehabilitación (Figura 3).

Para realizar los ejercicios de manera adecuada, es necesario bloquear ciertas articulaciones del robot para permitir un recorrido continuo y un movimiento fluido. Los ejercicios que se utilizarán en las pruebas incluyen: rotación lateral del hombro (E1), abducción y aducción del hombro (E2), flexión y extensión del hombro (E3) y flexión y extensión de bíceps (E4). Todos los ejercicios se han planteado para realizarse con el codo flexionado y según el usuario se podrán ejecutar de pie o sentados. Los agarres que se utilizarán en las pruebas incluyen: prototipo 1 fijo en el centro (P1C), prototipo 1 fijo ladeado hacia la derecha (P1D), prototipo 1 fijo ladeado hacia la izquierda (P1I), prototipo 1 con rotación circular (P1R) y prototipo 2 libre (P2). Mediante estos ejercicios el objetivo es recopilar datos cualitativos y cuantitativos.

3.1. Evaluación cualitativa

Antes de comenzar las pruebas el usuario deberá elegir entre mango genérico y mango ergonómico en función de su comodidad y adaptabilidad a la mano. Esta elección es importante pues todos los ejercicios propuestos anteriormente se realizarán con dicho mango. A continuación, se evaluará la adaptabilidad de los prototipos (y sus diferentes posiciones) a los movimientos, es decir, se comprobará si los usuarios tienen dificultades o no para realizar los ejercicios. La adaptabilidad se marca cualitativamente del 1 al 5 donde 1 es “Muy Fácil”, 2 es “Fácil”, 3 es “Complicado”, 4 es “Difícil” y 5 es



Figura 2. Segundo prototipo con conexión esférica.

“Muy difícil”. Por último, se se proporciona al usuario un cuestionario para valorar aspectos como comodidad, robustez, adaptabilidad y fragilidad, entre los prototipos que ha utilizado.

3.2. Evaluación cuantitativa

Durante el transcurso de las pruebas se evaluará cuantitativamente la fuerza en Newtons soportada por los prototipos en cada dirección, mientras se mantiene una posición fija con el robot. De forma más específica, se va a medir: la fuerza vertical hacia arriba (F1), la fuerza vertical hacia abajo (F2), la fuerza horizontal hacia la derecha (F3) y la fuerza horizontal hacia la izquierda (F4).

4. Resultados y discusión

Tras realizar las pruebas con varios usuarios se ha obtenido que, en primer lugar, la mayoría han elegido el mango ergonómico, pues coinciden en que es el que mejor se adapta a la forma de la mano y los dedos. A continuación, se han obtenido resultados en cuanto a adaptabilidad y fuerza.

	P1C	P1D	P1I	P1R	P2
E1	2	2	2	2	1
E2	1	1	1	1	1
E3	2	2	2	1	1
E4	3	3	3	3	2

Tabla 1. Adaptabilidad media de los prototipos a los movimientos (dificultad del 1 al 5)

En cuanto a la adaptabilidad media de los prototipos al movimiento o dificultad para realizarlos, debido a que

es una medida plenamente subjetiva puede estar influenciada por diversos factores (Tabla 1). Es decir, la dificultad del ejercicio, además de depender del propio prototipo y de su posición, también lo hará del propio usuario, del robot, de su condición física etc. Sin embargo, al observar los valores medios de la tabla se puede diferenciar cierto patrón. Los prototipos se adaptan mejor al ejercicio número 2 (abducción y aducción de hombro). Todo lo contrario, sucede con el ejercicio número 4 (flexión y extensión de bíceps). Esto se debe a las restricciones en el rango de movimiento que aparecen durante la realización de cada uno de los ejercicios. Es por ello por lo que, el prototipo número 2, al simular el movimiento de la propia muñeca y permitir un recorrido mucho más libre, consigue eliminar estas limitaciones y aumentar el rango de movimiento de los ejercicios, lo que se traduce directamente en una mayor adaptabilidad y por tanto una mayor facilidad para realizarlos.

	P1C	P1D	P1I	P1R	P2
F1	83,52	74,79	79,72	70,83	73,25
F2	72,73	82,33	77,40	66,30	64,13
F3	49,60	45,69	47,68	44,00	46,85
F4	52,15	43,43	46,21	46,58	42,14

Tabla 2. Fuerza media soportada por los prototipos en cada dirección (Newtons)

En cuanto a fuerza media soportada por los prototipos en cada dirección podemos observar que los movimientos horizontales aplican menos fuerza que los movimientos verticales (Tabla 2). La fuerza ejercida varía en función del usuario, sin embargo, podemos observar que es independiente del prototipo o de su posición. Es posible que uno se pueda sentir más robusto que otro, sin embargo,



Figura 3. Ejemplo de agarre y manipulación del efector (arriba). Entorno real de experimentación (abajo).

la fuerza soportada es aproximadamente la misma en cada dirección para cada tipo de agarre. Cabe mencionar que la fuerza medida no es la fuerza máxima que pueda aplicar cada usuario, pues no es el objetivo del prototipo. Se ha medido pues una fuerza convencional para realizar acciones y movimientos de la vida cotidiana.

Por último, se les pasó a los usuarios una pequeña encuesta de consideraciones finales. Para ellos, el prototipo más frágil o inseguro oscila entre el prototipo número 1 con posición fija (cualquiera de ellas) y el prototipo número 1 con rotación circular. En conclusión, podemos determinar que el prototipo número 1 se siente más tosco. Por otra parte, los usuarios coinciden en que el prototipo número 2 es el más robusto o seguro, cómodo o ergonómico y el que mejor se adapta a los ejercicios, pues es con el que menos cuesta hacerlos permitiendo un amplio rango de movimientos.

5. Conclusiones

En este artículo se han presentado dos prototipos de efector final para robótica de rehabilitación que permiten resolver la rigidez de los dispositivos habituales de agarre de este tipo de tecnologías. Los resultados de evaluación preliminar sugieren una mejora en la ergonomía y el rango de movimiento, siendo mayor en el uso de un prototipo con conexión esférica, que puede beneficiar sobre todo a pacientes con problemas específicos de espasticidad y pérdida de fuerza. En el futuro se pretende evaluar estos prototipos con usuarios con daño cerebral mediante la realización de tareas asistidas por el robot

de efector final.

Agradecimientos

Esta publicación es parte del proyecto de I+D+i PID2022-139105OB-I00, financiado por MCIN/AEI/10.13039/501100011033.

Referencias

- [1] L. Goldman *et al.* Understanding Acquired Brain Injury: A Review. *Biomedicines*, 9:10, 2022.
- [2] E. Claffin *et al.* Emerging Treatments for Motor Rehabilitation After Stroke. *The Neurohospitalist*, 5:77–88, 2015.
- [3] S.H. Lee *et al.* Comparisons between end-effector and exoskeleton rehabilitation robots regarding upper extremity function among chronic stroke patients with moderate-to-severe upper limb impairment. *Scientific Reports*, 10:1806, 2020.
- [4] G. Chiriatti *et al.* A Framework for the Study of Human-Robot Collaboration in Rehabilitation Practices. *Advances in Service and Industrial Robotics*, 1:190–198, 2020.
- [5] F. Chien *et al.* Robot-assisted therapy for upper-limb rehabilitation in subacute stroke patients: A systematic review and meta-analysis. *Brain and Behavior*, 10:1742, 2020.
- [6] F. Molteni *et al.* Exoskeleton and End-Effector Robots for Upper and Lower Limbs Rehabilitation: Narrative Review. *Innovations Influencing Physical Medicine and Rehabilitation*, 10:174–188, 2018.

Diseño de un sistema integral de atención domiciliaria con robot móvil asistencial

J.A. Vera Repullo¹, D. Parras-Burgos², F. Ortiz Zaragoza¹, J. Roca González¹

¹ Departamento automática, ingeniería eléctrica y tecnología electrónica, Universidad Politécnica de Cartagena, Cartagena, España, {jose.vera, francisco.ortiz, jroca.gonzalez}@upct.es

² Departamento de Estructuras, Construcción y Expresión Gráfica, Universidad Politécnica de Cartagena, Cartagena, España, {dolores.parras}@upct.es

Resumen

En esta comunicación se describe el diseño de un sistema integral de atención domiciliaria que consta de un pequeño robot móvil asistencial para monitorizar el bienestar de las personas mayores que viven solas en su hogar y sugerirle actividades para mejorar su estado de ánimo. Todo ello integrado en un ambiente doméstico de asistencia en el hogar (AAL), que incorpora, además del robot, un conjunto de sensores domóticos de bajo coste. El sistema integral de atención domiciliaria se ha integrado haciendo uso de ROS, de tecnologías IoT tales como Node-RED, y de la plataforma domótica Home-Assistant.

1. Introducción

El envejecimiento de la población tiene un gran impacto en la sociedad provocando diferentes problemas de salud que merman las capacidades y dificultan la situación de las personas [1]. Según un estudio del 2020 “Indicadores estadísticos básicos. *Informes envejecimiento en red*” [2], hay 9.057.193 de personas mayores en España, lo que representa el 19,3% de la población total, resultando cada vez más difícil garantizar la ayuda de los familiares [3], por lo que los Estados miembros se enfrentan a desafíos comunes en materia de cuidados de larga duración.

Además, la situación sufrida en las residencias españolas durante la pandemia del COVID-19, ha convertido el problema de la atención a las personas dependientes en una prioridad urgente. En este sentido, varios expertos se han pronunciado a favor de un cambio en el modelo de atención a la dependencia en España, señalando entre las prioridades "la disponibilidad de una atención integral e integrada en el domicilio" [4]. Así, Shin y Viron [5] aluden a la necesidad de una estrecha vigilancia de la salud física y mental para detectar los primeros signos de desadaptación ante una posible amenaza. Marbán Gallego et al. [6] aluden a que durante la COVID-19 se produjo un desbordamiento de la capacidad de respuesta asistencial en las residencias de ancianos, provocando una situación dramática. Por lo tanto, la protección de las personas mayores en España debe afrontar un escenario disruptivo que implique una atención centrada en la persona mediante la promoción de viviendas adaptadas.

Por otro lado, se requiere también un esfuerzo de integración para adaptar las tecnologías disponibles al servicio de las personas mayores o en situación de dependencia de forma usable, accesible y transparente a

través del Ambient Assisted Living (AAL), la IA, la robótica asistencial y el diseño de interfaces adaptadas individualmente. Esto puede permitir el acceso a los servicios con independencia de los conocimientos tecnológicos de los individuos y facilitar la evaluación de las terapias por parte de los profesionales de la salud, teniendo en cuenta los comentarios de los usuarios a distancia. Además, deben crearse conjuntos de datos anonimizados, que podrían publicarse adecuadamente utilizando formatos de datos abiertos para permitir a la comunidad investigadora diseñar estudios para mejorar y crear estrategias de entrenamiento activo, para aumentar la actividad mental y física de las personas mayores basándose en datos reales.

Los sistemas AAL inteligentes se utilizan para ayudar no sólo a las personas mayores, sino también a sus robots de asistencia. Un ejemplo de simbiosis entre los AAL y la robótica se muestra en el proyecto ENRICHME de H2020, en el que los sensores ambientales proporcionan cierta información de actividad al robot de asistencia. Otro ejemplo es el proyecto japonés de larga duración RobotTown y la habitación resultante con estructura de información [7], que contiene varios sensores distribuidos que sirven de apoyo a los robots de asistencia y ejemplifica la tecnología del Internet de los Robots y las Cosas (IoRT). Algunos autores [8,9] consideran que el AAL será el futuro en el cuidado de las personas mayores, por ello, se requiere la revisión pormenorizada de los dispositivos que en la actualidad se usan tanto para la determinación de estado de salud física y mental de las personas mayores, como para su estimulación física y cognitiva.

En esta comunicación se presenta el diseño, desarrollo y fabricación de un sistema integral de atención domiciliaria con robot móvil asistencial compatible con una plataforma de entorno doméstico ubicuo, AAL (Ambient Assisted Living) con computación y almacenamiento local, en hogares reales, para proporcionar, a las personas mayores, técnicas de psicoterapia, psicoeducación y estimulación cognitiva con seguimiento remoto, procurando la mejora de su estado emocional y aptitudes físicas.

2. Materiales y métodos

Aunque el estado del arte revela que existen numerosos esfuerzos de investigación centrados en el desarrollo de sistemas de automatización del hogar y de cuidado de ancianos [8,10-13] que incluyen robots de servicio en el

contexto de la asistencia domiciliar, actualmente es difícil encontrar soluciones cercanas al mercado para la asistencia remota de ancianos en el hogar. Estas soluciones no incluyen robots asistentes y Ambient Assisted Living (AAL) asequibles, con capacidad para analizar el entorno, determinar el estado emocional de los ancianos e interactuar con los usuarios de forma natural, adaptándose a ellos según sus roles (pacientes ancianos, familiares o cuidadores), entre otras características. El sistema domótico que garantiza el AAL en este proyecto pretende ser lo menos intrusivo posible y se buscan, siempre que sea posible, soluciones comerciales de bajo coste [14].

El sistema propuesto puede estructurarse en tres bloques funcionales:

a) Robot móvil asistencial:

El robot móvil asistencial navega de manera autónoma por la casa para atender las necesidades del usuario o para sugerirle alguna actividad según su estado anímico.

Para el diseño y modelado 3D de la carcasa del robot móvil asistencial se ha utilizado el software SolidWorks Student 2021. La estructura del robot móvil permite movimientos verticales para regular su altura y adecuarse al usuario (Fig. 1), y movimientos de la cabeza mediante giros rotacionales en varios sentidos (Fig. 2).



Figura 1. Movimiento vertical del cuerpo del robot móvil asistencial.



Figura 2. Movimientos de la cabeza del robot móvil asistencial.

Además, para el funcionamiento e integración del robot móvil con el resto del sistema, se han seleccionado los siguientes componentes:

- Alexa Echo Show 8: es un asistente virtual controlado fabricado por Amazon a partir de febrero de 2020. Sus dimensiones son de 200,4x135,9x99,1 mm y su masa es de 1037 g.

- Base móvil: La base móvil o robótica es la IClebo de Kobuki, la cual ofrece la posibilidad de montar encima plataformas según una estructura modular para introducir el hardware extra. Esta base móvil es muy popular en tareas de educación o investigación en el mundo de la robótica. Asimismo, está pensada para la autonomía en sí misma ya que desde ella se proporciona la alimentación a partes del sistema (sensor y ordenador). La base tiene dos motores de DC tipo brushed controlados por PWM. Son motores alimentados a 12V.
- Cámara: La Orbbec Astra es una cámara 3D con color VGA que está altamente recomendada para uso en algoritmos de visión por computador como, por ejemplo, el reconocimiento facial o gestual, la medida tridimensional, la percepción del ambiente y la reconstrucción tridimensional de mapas. Esta cámara se caracteriza por proporcionar una capacidad de respuesta de alta gama, una medición de profundidad, gradientes suaves y contornos precisos.
- Sensor láser: El sensor Hokuyo UST-10LX es un sensor láser 2D de pequeñas dimensiones. Se caracteriza por poseer una buena precisión y una alta velocidad en la detección de obstáculos y en la localización de robots autónomos. Este sensor es capaz de escanear con una apertura de hasta 270 grados con una precisión milimétrica hasta unos 10 metros de distancia. Por último, es necesario destacar su reducido consumo de energía.
- Ordenador: En este proyecto, se ha utilizado un NUC. Este es un potente mini-ordenador de reducidas dimensiones, el cual es utilizado para cualquier tipo de propósito. Además, es un sistema cuya placa base es completamente personalizable en relación a la memoria, el procesador y el almacenamiento. Lo cual posibilita la creación de un ordenador adaptado a las necesidades existentes.
- Actuador lineal: Un actuador lineal se encarga de convertir el movimiento rotatorio de un motor en un movimiento lineal. En este proyecto se ha utilizado un actuador lineal de 150 mm de la marca DCHOUSE.
- Servomotor: Un servomotor es un tipo de motor que puede ser controlado de manera precisa en aspectos como la posición angular, la aceleración y la velocidad. Para la realización de este proyecto han sido necesarios 4 servomotores RDS3115.

b) Ecosistema de sensores domóticos:

Se denomina sensor a todo dispositivo capaz de detectar acciones o estímulos externos y responder en consecuencia. El ecosistema está formado por todos los sensores que nos proporcionan datos útiles para determinar los diferentes estados de la persona o que supongan un riesgo para la misma. Muchos de los sensores tienen su propio sistema informático que filtra sus datos y gestionan

la entrega. Se trata de un conjunto de sensores, tipo IOT, tanto comerciales como de diseño propio para monitorizar el estilo de vida del usuario. Como el sistema es escalable en función de las necesidades de cada usuario, se selecciona el despliegue mínimo necesario de sensores y actuadores para los diferentes casos de uso.

De los datos de los sensores y haciendo uso de algoritmos, se reconocen las actividades cotidianas con el fin de obtener patrones de comportamiento para cada individuo. Las principales actividades que se pueden monitorizar para que, con el apoyo del equipo de psicología del proyecto, proponer estrategias de coaching emocional, son las siguientes:

- Horas de sueño, actividad física en el interior de la vivienda, veces que va al baño, tiempo que permanece en estancias. Para ello se utilizan pulseras de actividad como Empatica E4 y sensores de movimiento como Mi Motion Sensor de Xiaomi.
- Tiempo sentado ante el televisor o en cama mediante sensores de asiento sencillos con interfaz electrónica desarrollada en el proyecto.
- Alteración de hábitos o estado de ánimo por cambios de sensación térmica y nivel lumínico mediante sensores de temperatura y luminosidad.
- Actividades de limpieza, organización, uso del frigorífico, etc.: mediante detectores de apertura, sensores de consumo eléctrico, etc.
- Niveles de estrés y actividad y otros parámetros biométricos que completen la estimación de estado de ánimo con pulsera médica Empatica E4.

Todos los sensores comerciales usan el protocolo Zigbee, por tanto, para poder realizar una gestión de los mismos de manera local desde Home Assistant, se usó la puerta de enlace genérica CC2531. En un principio se propuso también la opción de usar la puerta de enlace Conbee II pero se descartó por motivos de estabilidad. Por otro lado, para los sensores de diseño propio, se utiliza un microcontrolador de bajo coste ESP32 para publicar la información de los mismos mediante el protocolo de comunicación MQTT a un broker Mosquitto, montado sobre la misma Raspberry Pi como Home Assistant.

c) Unidad central home-assistant:

Añadiendo un nivel superior de interacción robot-entorno, mediante el desarrollo de *Skills*, se puede acceder también a funcionalidades mediante interfaz hablados lo que da al sistema una flexibilidad total. Se cuenta con un sistema de control que se comunica con la vivienda mediante conexión WIFI. Una vez conectado el dispositivo de Alexa, el usuario puede integrar acciones del robot móvil y de la vivienda que interaccionen con el usuario (Fig. 3). Alexa se utiliza como interfaz de voz entre el sistema y el usuario, gracias a que se ha desarrollado una interfaz software con el controlador de la vivienda usando protocolo MQTT. Pero a su vez, también se tiene la funcionalidad típica de un altavoz inteligente a disposición del usuario (recordatorios, citas, poner música, etc).

Las sugerencias de coaching emocional que se proponen tienen que ver generalmente con actividades lúdicas, como sugerir realizar ejercicio cuando se detecta que la persona lleva demasiado tiempo sentada, llamar a un familiar si se detecta estrés o cambio en estado de ánimo, poner música, incrementar la luminosidad de la vivienda, etc. Siempre se respeta la privacidad de la persona y se pregunta si se quiere realizar la actividad, no se obliga.

El sistema se ha probado en laboratorio, con voluntarios de la Universidad de Mayores de la UPCT recibiendo sus impresiones a través de cuestionarios y haciendo preguntas directas y recogiendo sus comentarios.



Figura 3. Esquema del sistema completo.

3. Resultados

Se ha desarrollado un robot móvil de acompañamiento (Fig. 4) junto a un sistema de Ambient Assisted Living integrado en el hogar, los algoritmos de inteligencia artificial y los sensores biométricos incorporados en la pulsera médica. Con todos estos dispositivos se puede monitorizar la actividad de la persona en casa, estimar su estado de ánimo, y sugerirle actividades de coaching emocional que puedan ayudar a estimular y mejorar dicho estado anímico.

Los resultados de laboratorio con usuarios voluntarios registran una buena aceptación generalizada del sistema, y también recogen sugerencias de mejora, como más robustez en el diseño del robot, que permanezca suficientemente alejado cuando los usuarios se estén moviendo o más sugerencias de coaching emocional.

Aunque el proyecto finalizó a mediados de 2023, actualmente se está probando en fase de proyecto piloto en un apartamento adaptado del centro de día de la Fundación Poncemar en Lorca (Murcia, España) y se ha concedido financiación para desarrollar un proyecto en colaboración público-privada con empresas interesadas en los resultados de esta investigación.

Para el desarrollo de las plataformas robóticas se ha utilizado ROS y algoritmos propios como los descritos en apartados anteriores para un mapeo y una navegación autónoma eficaz en un entorno real de una casa. En cuanto al sistema domótico, se ha testado de manera más

prolongada y se ha comenzado con la integración en un entorno de pruebas real (Fig. 5).



Figura 4. Primer prototipo del robot móvil asistencial.



Figura 5. AAL + instalaciones del laboratorio de robótica en UPCT. Escenario de prueba "Inicio del día": (1) robot Echo: Base Turtlebot impresa en 3D con Echo 8 y Alexa en la cabeza; (2) sujeto vistiendo la Empatica E4; (3) detector de presencia en cama; (4) detector de presencia en silla; (5) interfaz de usuario AAL; (6) simulación de detección de estado de atención al ver la TV; (7) simulador de otros elementos domóticos. Referencia (Barber, 2022).

4. Conclusiones

En esta comunicación se ha presentado el desarrollo de un sistema integral sobre plataformas abiertas, asequible y configurable en función de las necesidades de cada usuario que genera un ecosistema AAL con computación y almacenamiento ubicuo, haciendo uso de la nube solo como respaldo y como alojamiento de la máquina de aprendizaje. Con esto se consigue que la vivienda funcione correctamente, aunque caiga la conexión a Internet.

El sistema propuesto supone un cambio en los sistemas de ayudas técnicas domiciliarias, debido a que permite el reconocimiento de actividades domésticas del usuario. De este modo, puede detectarse un cambio en el comportamiento habitual del usuario, que permitirá detectar desde cambios en el estado emocional hasta deterioros en las capacidades cognitivas de la persona.

Agradecimientos

El proyecto HIMTAE, subproyecto Robwell (referencia RTI2018-095599-A-C22) ha sido financiado por el Programa Estatal de Investigación, Desarrollo e Innovación Orientada a los Retos de la Sociedad, en el marco del Plan Estatal de Investigación Científica y Técnica y de Innovación 2013-2016.

El proyecto piloto ADDIM-Poncemar (Asistencia Domiciliaria Integral para Mayores) financiado por la "Fundación Integra Digital" de la Región de Murcia se está desarrollando en la actualidad en los Centros de Día de la Fundación Poncemar, en el campus de la Salud de Lorca, Universidad de Murcia. Agradecemos su colaboración e implicación en la prueba de las soluciones presentadas en esta comunicación.

Referencias

- Luo, M.; Guo, L.; Yu, M.; Jiang, W.; Wang, H. The psychological and mental impact of coronavirus disease 2019 (COVID-19) on medical staff and general public – A systematic review and meta-analysis. *Psychiatry Research* **2020**, *291*, 113190. doi: <https://doi.org/10.1016/j.psychres.2020.113190>
- Pérez, J.; Abellán, A.; Aceituno, P.; Ramiro, D. Un perfil de las personas mayores en España. Indicadores estadísticos básicos. *Informes envejecimiento en red* **2020**, 11-39. doi:
- Comisión Europea. Libro verde sobre el envejecimiento. Fomentar la solidaridad y la responsabilidad entre generaciones. Bruselas, 2021.
- Davey, V. Situación en España de la evaluación de sistemas de atención a personas mayores en situación de dependencia. *Informes envejecimiento en red* **2021**, *28*, 18. doi:
- Shinn, A.K.; Viron, M. Perspectives on the COVID-19 pandemic and individuals with serious mental illness. *The Journal of Clinical Psychiatry* **2020**, *81*, 14205. doi: 10.4088/JCP.20com13412
- Gallego, V.M.; Codorniu, J.M.; Cabrero, G.R. El impacto de la Covid-19 en la población mayor dependiente en España con especial referencia al sector residencial. *J Ciència Saúde Coletiva* **2021**, *26*, 159-168. doi: 10.1590/1413-81232020261.33872020
- Hasegawa, T.; Murakami, K.; Kurazume, R.; Senta, Y.; Kimuro, Y.; Ienaga, T. In *Robot town project: Sensory data management and interaction with robot of intelligent environment for daily life*, Proceedings of the International Conference on Ubiquitous Robots and Ambient Intelligence, Pohang, Korea, 2007; p 369373.
- Cicirelli, G.; Marani, R.; Petitti, A.; Milella, A.; D'Orazio, T. Ambient Assisted Living: A Review of Technologies, Methodologies and Future Perspectives for Healthy Aging of Population. **2021**, *21*, 3549. doi: 10.3390/s21103549
- King, P.H. Smart Home Technologies and Services for Geriatric Rehabilitation. *IEEE Pulse* **2022**, *13*, 35-36. doi: 10.1109/MPULS.2022.3175383
- EnrichMe - Collaborative Projects | PAL Robotics [Online]. <https://pal-robotics.com/collaborative-projects/enrichme/> (Available: 26/09/2023),
- Abdollahi, H.; Mahoor, M.H.; Zandie, R.; Siewierski, J.; Qualls, S.H. Artificial Emotional Intelligence in Socially Assistive Robots for Older Adults: A Pilot Study. *IEEE Transactions on Affective Computing* **2023**, *14*, 2020-2032. doi: 10.1109/TAFFC.2022.3143803
- Di Nuovo, A.; Broz, F.; Wang, N.; Belpaeme, T.; Cangelosi, A.; Jones, R.; Esposito, R.; Cavallo, F.; Dario, P. The multi-modal interface of Robot-Era multi-robot services tailored for the elderly. *Intelligent Service Robotics* **2018**, *11*, 109-126. doi: 10.1007/s11370-017-0237-6
- Koutli, M.; Theologou, N.; Tryferidis, A.; Tzovaras, D. In *Abnormal Behavior Detection for Elderly People Living Alone Leveraging IoT Sensors*, 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), 28-30 Oct. 2019, 2019; pp 922-926.
- Calatrava, F.M.; Ortiz, F.J.; Vera, J.A.; Roca, J.; Jiménez, M.; Martínez, O. In *Heterogeneous system for monitoring daily activity at home and the well-being of the elderly*, Proceedings of the XLII Conference on Automatic: Minutes Book, Castellón (España), 2021; Castellón (España), pp 632-639.

Digitalización del test de Tinetti mediante técnicas de Visión Artificial

J.D. Blasco García¹, G. García López², M. Jiménez Muñoz², J.A. López Riquelme², J.J. Feliu Batlle², N. Pavón Pulido², M.T. Herrero¹

¹ Neurociencia Clínica y Experimental (NiCE). Facultad de Medicina. Instituto de Investigación en Envejecimiento. IMIB-Pascual Parrilla. Universidad de Murcia, Murcia, España, {jesusdamian.blascog, mtherrer@um.es}

² Departamento de Automática, Ingeniería Eléctrica y Tecnología Electrónica, Universidad Politécnica de Cartagena, Cartagena, España, {gabriel.garcia,marta.jimenez@edu.upct.es, jantonio.lopez,jorge.feliu,nieves.pavon@upct.es}

Resumen

En este trabajo se presenta un método, basado en Visión Artificial y técnicas de Deep Learning, que ayuda a los profesionales de la Salud, en el ámbito de la Rehabilitación, la Fisioterapia, la Terapia Ocupacional y la Geriátrica, a almacenar, monitorizar y evaluar digitalmente los resultados de la aplicación del Test de Tinetti en pacientes mayores, de forma no invasiva y sin necesidad de utilizar accesorios electrónicos vestibles.

La técnica propuesta utiliza un sensor RGB-D de bajo coste para adquirir secuencias de imágenes con información de profundidad asociada. Cada imagen se utiliza como entrada de la solución MediaPipe Pose Landmarker que detecta puntos de referencia del cuerpo humano proporcionando la ubicación relativa de las principales articulaciones del mismo. Dichos puntos se proyectan sobre el sistema de referencia del sensor RGB-D utilizando los datos que éste proporciona acerca de la profundidad de cada píxel, lo que permite calcular una serie de medidas útiles para analizar, matemáticamente, las capacidades motoras y de equilibrio de los pacientes, cuando realizan las pruebas habituales que se especifican en el test de Tinetti.

El prototipo se ha validado, primero, en un entorno de laboratorio y, posteriormente, con pacientes reales en una residencia de mayores.

Los resultados de cada prueba se presentan de forma gráfica, de manera que, a través de la observación de dichas representaciones, los profesionales puedan evaluar la evolución motora de los pacientes, con el objeto de ayudar en la prevención de caídas.

1. Motivación

El envejecimiento de la población [1] ha incrementado notablemente el volumen de pacientes tanto en Atención Primaria, como en ciertas especialidades relacionadas con las dolencias propias de la edad avanzada, lo que supone un desafío para los sistemas de Salud en términos de gestión de recursos y un gran problema económico y social.

Uno de los problemas que afectan principalmente a las personas mayores se relaciona con la pérdida de las capacidades motoras por diferentes causas: la edad en sí, o enfermedades como la artrosis, la artritis, la sarcopenia, la osteoporosis o el deterioro cognitivo, entre otras. Aunque los problemas de movilidad pueden afectar a toda la población, están presentes, principalmente, en los mayores de 60 años y se incluyen dentro de la sintomatología de los síndromes geriátricos. Dichos problemas suelen ser origen de caídas que, en ocasiones, son mortales, o conllevan la pérdida total o parcial de movilidad y/o la afectación

cognitiva, cuando se produce un traumatismo en la cabeza. Es por ello, que la Organización Mundial de la Salud (OMS), resalta que “las estrategias preventivas deben hacer hincapié en (...) la priorización de la investigación relacionada con las caídas y el establecimiento de políticas eficaces para reducir los riesgos”, y considera que los trabajos de prevención en mayores deben incluir, entre otros, “intervenciones multifactoriales, como evaluaciones individuales del riesgo de caídas, seguidas de intervenciones y derivaciones médicas en función de los riesgos detectados” [2]. Nótese que los problemas motores afectan a casi el 20% de los individuos mayores de 65 años. A partir de los 75 años, aproximadamente el 50% de los ancianos tiene problemas para salir de su casa y un 20% se halla confinado en su domicilio.

En el contexto de la prevención de caídas, existen tests [3] y métodos de cribado para evaluar el riesgo mediante la observación de la postura, el equilibrio y la marcha. En su mayoría se basan en puntuar determinadas posturas o movimientos observando, subjetivamente, la ejecución de los mismos. Para unificar [4] y mejorar la precisión de estas pruebas se han realizado intentos de usar sensores de e-Health y herramientas digitales que toman datos de los movimientos del usuario-paciente y permiten cuantificar objetivamente una serie de parámetros. Se han probado sensores de diverso coste y precisión [5-9], obteniéndose resultados relativamente significativos, tras la realización de pruebas en entornos reales. En cuanto a las técnicas de reconocimiento de actividad humana o HAR (Human Activity Recognition), muy pocas se aplican para evaluar el riesgo de caída de personas, y menos si se trata de analizar automáticamente los parámetros relacionados con el equilibrio, la postura o la marcha definidos en los métodos documentados en el estado del arte (por ejemplo, el test de Tinetti [10]), y que son ampliamente utilizados por los profesionales de la Salud. A nivel de dispositivos comercializados para este fin, existen aplicaciones para dispositivos móviles inteligentes que permiten registrar las observaciones de los terapeutas, pero estas observaciones las hace el profesional y simplemente las registra, de manera digital, de una forma similar a realizar un apunte en una agenda.

El trabajo que se presenta explora una vía novedosa para ayudar a los terapeutas a predecir el riesgo de caídas a través de los tests habitualmente usados, pero digitalizando

el proceso mediante un sistema de Visión Artificial y la aplicación de técnicas de Deep Learning, para llevar a cabo las evaluaciones a través de una herramienta que analiza matemáticamente la evolución del equilibrio y de la actividad motora de los usuarios-pacientes. De forma resumida este trabajo se organiza como se indica a continuación: en la Sección 2 se detallan las arquitecturas hardware y software del sistema diseñado, la Sección 3 describe las pruebas realizadas y los resultados obtenidos y, finalmente, en la Sección 4 se analizan las ventajas y limitaciones del sistema y se esbozan los trabajos futuros a realizar para mejorar el prototipo desarrollado.

2. Diseño del sistema

El sistema desarrollado permite que un terapeuta proponga y evalúe ítems del test de Tinetti, mediante la captura de movimiento del usuario-paciente a través del sensor RGB-D mencionado. Concretamente, el test de Tinetti se define como un examen de valoración geriátrica ampliamente utilizado para evaluar el equilibrio y la marcha en adultos mayores, que consta de dos partes: una evaluación del equilibrio (9 posibles ítems) y una evaluación de la marcha (7 posibles ítems). Un ejemplo de ítem relacionado con el equilibrio es “Soporte bípedo (de pie)”, donde la persona debe pararse con los pies juntos, sin ayuda, y mantener la postura durante 10 segundos. Un ejemplo de ítem relacionado con la marcha es “Longitud del paso”, donde se evalúa la longitud de los pasos durante la marcha.

El sistema, que se compone de un conjunto de procesos ejecutados como nodos ROS (Robotic Operating System), en un ordenador personal convencional al que se ha conectado un sensor RGB-D (modelo Orbecc Astra Pro), permite llevar a cabo las siguientes tareas:

- Seleccionar un determinado ejercicio relacionado con el ítem correspondiente del test.
- Iniciar la captura de movimiento para dicho ejercicio usando el sensor RGB-D.
- Parar la captura de movimiento.
- Evaluar automáticamente, a lo largo del tiempo, la posición de los puntos de referencia del cuerpo humano que proporciona la solución MediaPipe Pose Landmarker, tomando la imagen RGB proporcionada por el sensor RGB-D como entrada y calculando la distancia de dichos puntos de referencia al sensor a partir de la imagen de profundidad obtenida del mismo.
- Calcular un conjunto de características a partir de la secuencia de puntos de referencia mencionados, expresados en el sistema de referencia del sensor RGB-D que permiten estimar, objetivamente, la capacidad del usuario-paciente para permanecer en equilibrio o describir una trayectoria concreta durante la marcha, según el ítem del test de Tinetti que deba realizar.

2.1. Características evaluadas.

Las características que la aplicación permite medir automáticamente (ver Figura 1), a partir del conjunto de

puntos de referencia del cuerpo humano obtenidos del par de imágenes proporcionados por el sensor RGB-D, son:

- Trayectoria: Camino descrito por el punto medio de la cadera proyectado sobre el suelo.
- Balanceo del tronco: Definido por el ángulo de inclinación de los hombros.
- Separación de brazos: Definido por los ángulos de los brazos y el tronco, uno para cada lado del cuerpo.
- Soporte de brazos: Definido por los ángulos formados por los segmentos que unen hombro, codo y muñeca. Útil para determinar si el usuario-paciente usa los brazos cuando se sienta o se levanta.
- Separación de tobillos: Definido por la distancia horizontal entre los puntos del tobillo.
- Zancada: Definido por la diferencia de profundidad entre las articulaciones del tobillo.
- Inclinación de tronco: Calculada como los ángulos entre hombro, cadera y rodilla de ambos lados del cuerpo.

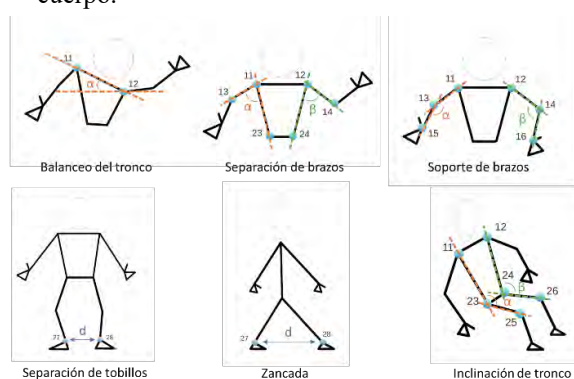


Figura 1. Características estimadas por la aplicación.

2.2. Relación de características con ítems del test de Tinetti.

La Tabla 1 muestra la relación existente entre las características estimadas de forma automática por la aplicación y los diferentes ítems del test de Tinetti (ver Tabla 2).

Ítem	Trayectoria	Balanceo de tronco	Separación de brazos	Soporte de brazos	Separación de tobillos	Zancada	Inclinación del tronco
1							x
2				x			x
3							x
4		x			x		
5		x			x		
6		x			x		
7		x			x		
8	x	x			x	x	
9		x		x			x
10	x						
11						x	
12						x	
13						x	
14	x						
15	x	x	x				
16					x		

Tabla 1. Relación entre características e ítems del test de Tinetti.

Por ejemplo, el balanceo del tronco y la separación de tobillos se utilizan para valorar varios ítems. El terapeuta observa los resultados de las dos características y asigna una puntuación a partir de los resultados gráficos mostrados por la aplicación.

ITEM	Significado		
1	Equilibrio sentado	9	Sentarse
2	Levantarse de la silla	10	Inicio de la marcha
3	Intentos de levantarse de la silla	11	Longitud y altura del paso
4	Equilibrio de pie inmediato (5'')	12	Separación de pies
5	Equilibrio de pie	13	Simetría del paso
6	Equilibrio al ser empujado	14	Continuidad del paso
7	Equilibrio con ojos cerrados	15	Trayectoria
8	Giro de 360 grados	16	Postura caminando

Tabla 2. Significado de cada ítem del test de Tinetti.

Además, puede compararse diferentes ítems del test aplicado en diferentes momentos, lo que facilita el seguimiento del usuario-paciente a lo largo del tiempo, observando si el riesgo de caída aumenta, disminuye o se mantiene.

3. Resultados.

El sistema se ha validado tanto en un entorno de laboratorio controlado, como en un entorno real. En el primer caso, los propios investigadores han realizado movimientos naturales y forzados para comprobar que el sistema puede estimar las características necesarias utilizadas para evaluar cada uno de los ítems del test de Tinetti. La Figura 2 muestra un ejemplo de representación gráfica del balanceo del tronco llevada a cabo por uno de los investigadores. Nótese que, si el usuario camina erguido y sin balancear el tronco, la señal oscila poco ya que la inclinación de la línea de los hombros apenas cambia y, por tanto, el ángulo de balanceo es estable; sin embargo, si durante la marcha se producen alteraciones significativas en dicho ángulo, se aprecian, claramente, los máximos.

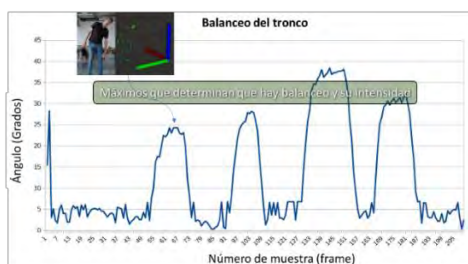


Figura 2. Balanceo del tronco a lo largo del tiempo. Se mide el grado de inclinación de la línea de los hombros.

En el segundo caso, se seleccionó un conjunto de 6 usuarios-pacientes de una residencia de mayores (3 hombres y 3 mujeres), con diferentes patologías y capacidad motora. La prueba realizada fue supervisada por un fisioterapeuta. Dado que estas pruebas pretendían comprobar la eficacia del sistema de captura de movimientos y la usabilidad de la aplicación, sólo se solicitó a los participantes que realizaran 3 ejercicios:

- Sentarse/levantarse de una silla: Era interesante medir el soporte de brazos, entre otras características.
- Caminar en línea recta: Además de la trayectoria, se midió la posición de los brazos durante el trayecto.
- Mantener equilibrio (de pie, ojos abiertos).
- Mantener equilibrio (de pie, ojos cerrados).

El tiempo requerido por la mayoría de los usuarios-pacientes que participaron en la prueba fue de, aproximadamente, 20 minutos en total, ya que no es necesario ninguna preparación adicional de los sujetos (como colocación de dispositivos vestibles u otros que puedan interferir en la vida diaria de los residentes). Sólo

se requiere determinar, adecuadamente, la posición del sensor RGB-D y suficiente espacio y distancia para que éste pueda capturar la imagen de la porción de cuerpo del usuario-paciente necesaria, mientras realiza los ejercicios de marcha o equilibrio propuestos.

La Figura 3 muestra un instante de la prueba “Levantarse/sentarse” realizada por un residente. La Figura 4 muestra otro instante de una prueba de equilibrio realizada por otro residente.

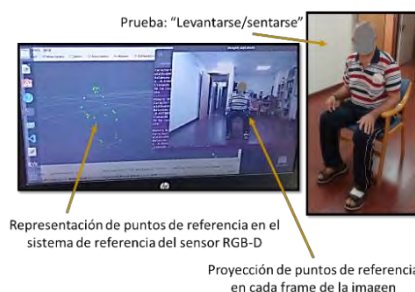


Figura 3. Aspecto de la aplicación mientras se realiza la prueba “Levantarse/sentarse”.

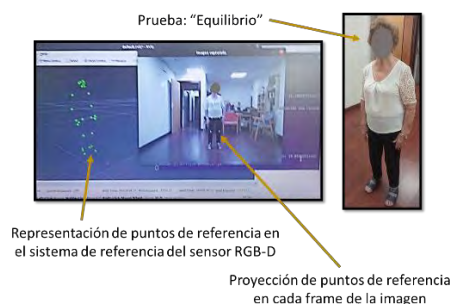


Figura 4. Aspecto de la aplicación mientras se realiza una prueba de equilibrio.

La Figura 5 muestra la gráfica correspondiente a la medida de la característica “Inclinación de tronco”, realizada cuando se pide al usuario-paciente que se levante y se sienta en una silla varias veces. Se observa la evolución del ángulo del tronco y se obtiene una señal con características de escalón. En el nivel alto (ángulos superiores a 160°) el sujeto estaría de pie, mientras que en nivel bajo de la señal (ángulos entre 120° y 140°) estaría sentado. El número de “escalones” se corresponde con el número de repeticiones de la prueba (pasar de sentado a en pie y viceversa), y el ancho de la señal da una idea del tiempo empleado para realizar el proceso. La Figura 6 muestra la trayectoria seguida mientras se evalúa la marcha (caminar recto 4 metros hacia la cámara). La Figura 7 muestra la separación de los pies y la posición de los brazos mientras se realiza la prueba de permanecer en equilibrio con los ojos cerrados. Se observa que el usuario-paciente no necesitó mover los brazos para mantener el equilibrio y, además, la posición de los pies fue correcta (juntos y estables).

4. Conclusiones y trabajo futuro.

A la vista de los resultados descritos, se ha diseñado un sistema útil para ayudar a los profesionales de la Salud a monitorizar y evaluar el riesgo de caídas en personas mayores, mediante la aplicación del conocido test de Tinetti, utilizando una herramienta informática.

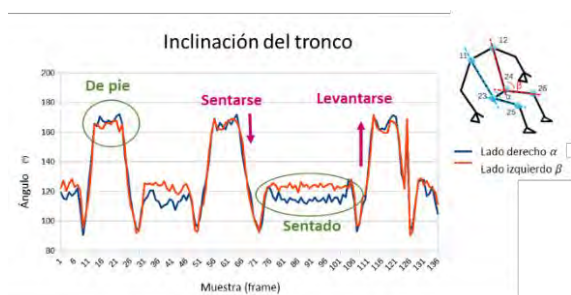


Figura 5. Proceso de levantarse/sentarse.

Trayectoria (Proyección en el plano del suelo)

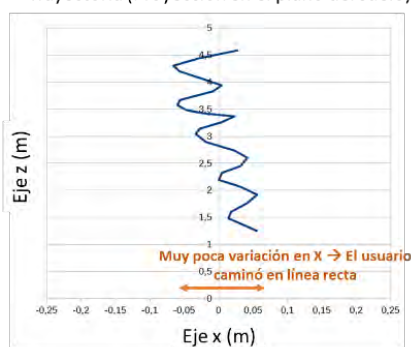


Figura 6. Trayectoria proyectada en el suelo.

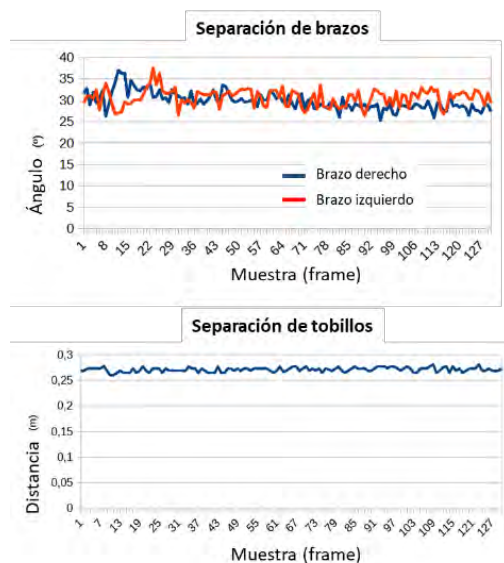


Figura 7. Medida del equilibrio con los ojos cerrados.

Dicha herramienta es capaz de capturar información sobre la capacidad motora y de equilibrio de los usuarios-pacientes, calcular una serie de características alineadas con los diferentes ítems del test, y representar gráficamente los resultados de cada prueba, permitiendo que el terapeuta pueda almacenarlos digitalmente, evaluarlos y compararlos.

Aunque existen sistemas de detección de caídas basados en el uso de dispositivos electrónicos vestibles, la aplicación propuesta es novedosa, ya que se usa un sensor RGB-D para realizar la adquisición de datos, lo que resulta más cómodo para el usuario, que no necesita portar ningún elemento adicional, y se mejora el coste económico del sistema, ya que se puede usar un único sensor de bajo coste para todos los usuarios-pacientes.

El sistema es aún un prototipo y se está trabajando en su mejora. Se está diseñando una plataforma en la Nube que permite la administración remota de la información, y un conjunto de módulos software, que extenderá las capacidades del sistema para que se pueda estimar automáticamente la puntuación del test de Tinetti, y otros similares, y predecir el nivel de riesgo de caída y la evolución del mismo, ayudando, así, a los profesionales de la Salud en la tarea de monitorización y seguimiento remoto de los usuarios-pacientes.

Agradecimientos

Este trabajo ha sido parcialmente financiado a través del proyecto JUNO+ (TED2021-130942B-C21 / TED2021-130942A-C22), Ministerio de Ciencia e Innovación, y a través de la Fundación Séneca (Región de Murcia), beca 21593/FPI/21. Asimismo, cabe resaltar nuestro agradecimiento a la residencia de mayores “Edad Dorada Mensajeros de la Paz” de San Pedro del Pinatar (Murcia), por su desinteresada colaboración.

Referencias

- [1] 10 datos sobre el envejecimiento y la salud. <https://www.who.int/features/factfiles/ageing/es/> Visitada el 30/09/2020.
- [2] Organización Mundial de la Salud. Caídas. Enero 2018; <http://www.who.int/mediacentre/factsheets/fs344/es/> Visitada el 03/12/2019.
- [3] Barry, E., et al. (2014). Is the timed up and go test a useful predictor of risk of falls in community dwelling older adults: A systematic review and meta- analysis. BMC Geriatrics, 14(1) doi:10.1186/1471-2318-14-14.
- [4] Mayhew, A. J., et al. (2023). Normative values for grip strength, gait speed, timed up and go, single leg balance, and chair rise derived from the canadian longitudinal study on ageing. Age and Ageing, 52(4) doi:10.1093/ageing/afad054
- [5] Vera-Remartínez, et al. (2023). Validity and reliability of an android device for the assessment of fall risk in older adult inmates. Nursing Open, 10(5), 2904-2911. doi:10.1002/nop.2.1532.
- [6] Lv, X., et al. (2023). Information collection system for fall detection of stroke patients under cascade algorithm in the context of multi-modal information fusion and e-health. Expert Systems, 40(4) doi:10.1111/exsy.12809.
- [7] Morat, T., et al. (2023). Evaluation of a novel technology-supported fall prevention intervention – study protocol of a multi-centre randomised controlled trial in older adults at increased risk of falls. BMC Geriatrics, 23(1) doi:10.1186/s12877-023-03810-8.
- [8] Yun, S., et al. (2023). Fine-grained motion recognition in at-home fitness monitoring with smartwatch: A comparative analysis of explainable deep neural networks. Healthcare (Switzerland), 11(7) doi:10.3390/healthcare11070940
- [9] Massimo W. Rivolta, et al., Evaluation of the Tinetti score and fall risk assessment via accelerometry-based movement analysis, Artificial Intelligence in Medicine, Vol. 95, 2019, pp. 38-47, <https://doi.org/10.1016/j.artmed.2018.08.005>
- [10] Michel Raïche, et al., Screening older adults at risk of falling with the Tinetti balance scale, The Lancet, Volume 356, Issue 9234, 2000, Pages 1001-1002, [https://doi.org/10.1016/S0140-6736\(00\)02695-7](https://doi.org/10.1016/S0140-6736(00)02695-7)

Telemedicina y e-Salud

Miércoles 22 9:30 - 11:00

Aula PB3

Diseño e implementación de una aplicación basada en procesos para la gestión de consentimientos de pacientes

J.A. García Linares, I. Román Martínez¹, J. Calvillo Arbizu^{1,2}, L.J. Reina Tosina^{2,3}

¹ Departamento de Ingeniería Telemática, Universidad de Sevilla, Sevilla, España, {iroman, jcalvillo}@us.es

² Grupo de Ingeniería Biomédica, Universidad de Sevilla, Sevilla, España

³ Departamento de Teoría de la Señal y Comunicaciones, Universidad de Sevilla, Sevilla, España, jreina@us.es

Resumen

En este trabajo se presenta la primera versión de una aplicación para la gestión de consentimientos de pacientes. La solución está planteada bajo el paradigma de gestión de procesos, de modo que se han diseñado y definido procesos, utilizando el estándar BPMN, tanto para la solicitud de consentimientos por parte del profesional sanitario como para la aceptación o rechazo de los mismos, por parte del paciente. Por otro lado, se han considerado requisitos de reutilización e interoperatividad, de modo que la solución facilita la personalización de los consentimientos, que están basados en recursos del estándar FHIR.

1. Motivación y objetivo

Las entidades sanitarias deben cumplir con las regulaciones del RGPD (Reglamento General de Protección de Datos) [1] al recopilar y procesar los datos de los pacientes. Este reglamento utiliza los consentimientos como recurso fundamental para proteger la privacidad y los derechos de los individuos en lo que respecta a sus datos sanitarios.

Dentro del dominio sanitario, se puede definir “consentimiento” como el permiso que un paciente otorga de manera voluntaria e informada para que sus datos de salud sean recopilados, procesados y/o utilizados por entidades proveedoras de servicios sanitarios o terceros involucrados en la provisión de dichos servicios.

Cuando se recopilan datos de salud sensibles para fines de tratamiento médico o de investigación, se espera que el consentimiento sea explícito. Esto significa que los pacientes deben otorgar su permiso de manera activa y específica, generalmente mediante una acción afirmativa clara, como marcar una casilla de verificación o firmar un formulario de consentimiento.

Aunque, bajo ciertas condiciones, los consentimientos puedan ser implícitos, siempre es de obligado cumplimiento el principio de transparencia, por el que los pacientes deben estar informados, de forma concisa, inteligible y con un lenguaje claro y sencillo, sobre el uso de sus datos. Además, los pacientes deben poder revocar sus consentimientos en cualquier momento. Por lo que las organizaciones deben asegurar que proporcionan a los pacientes mecanismos que permitan al paciente ejercitar estos derechos.

Toda esta gestión de consentimientos debe ser asumida por las organizaciones sanitarias, que deben diseñar procedimientos específicos para ello. Estos deben estar

optimizados, especialmente al considerar que en muchas ocasiones estarán vinculados a procedimientos asistenciales y por tanto influirán en la eficiencia de estos.

El objetivo de este trabajo es realizar una primera identificación de los procesos más elementales en la gestión de consentimientos explícitos y desarrollar una aplicación que permita ejecutarlos. La solución debe proporcionar facilidades para el seguimiento, análisis y mejora continua de los procesos diseñados y para facilitar su integración con otros procesos de la organización. Para ello la aplicación se implementa utilizando tecnologías de gestión de procesos empresariales [2], concretamente se desarrolla una “Aplicación de negocio jBPM” [3].

En este contexto particular, la adopción de procesos de negocio resulta especialmente beneficiosa para:

- La identificación y corrección proactiva de brechas de seguridad. El seguimiento en tiempo real y el análisis de registros históricos de los procesos en los que se manejan los consentimientos facilitará identificar posibles brechas de seguridad o incumplimientos en la protección de la privacidad. Las facilidades para modificar los procesos aceleran la adopción de las medidas correctivas oportunas.
- Facilitar la transparencia y la identificación de responsabilidades en la gestión de datos sanitarios, al mejorar la trazabilidad y el rastreo del flujo de datos.
- Facilitar la adaptabilidad y agilidad ante los cambios normativos o legales relacionados con la protección de datos, permitiendo actualizar y ajustar los procesos de forma oportuna para garantizar el cumplimiento normativo y mantener la confidencialidad de la información.

2. Paradigmas y tecnologías

2.1. La gestión de procesos

La gestión de procesos empresariales (BPM, *Business Process Management*) [2] es una disciplina empresarial que permite mejorar la eficiencia de las organizaciones a través de la gestión sistemática de sus procesos de negocio. La Figura 1 recoge de forma resumida el marco operacional de BPM que incluye, entre otros:

- La especificación y formalización de procesos: las notaciones normalizadas, como BPMN [4], independizan la especificación de procesos de la plataforma en la que estos se utilizan. Normalmente

las soluciones técnicas proporcionan facilidades con interfaces gráficas que permiten la especificación de procesos con pocos requisitos de codificación.

- La ejecución y el seguimiento de instancias de procesos: ofreciendo a los supervisores datos de las ejecuciones en tiempo real y de los históricos. Esto permite monitorizar el estado de cualquier ejecución de un proceso en cualquier instante y realizar análisis que faciliten la mejora continua de los procesos. Estas tareas se engloban en la disciplina denominada *Business Activity Monitoring* (BAM) [5-8]



Figura 1. Marco operacional de la gestión de procesos empresariales

Aplicando este paradigma para los procesos relacionados con el manejo de consentimientos en una organización sanitaria, se busca alcanzar una mayor eficiencia, seguridad y responsabilidad en la gestión de los consentimientos. Por otro lado, se facilita la mejora continua de los procesos implicados, adaptándolos a las necesidades cambiantes y optimizándolos, tras el análisis del histórico de sus ejecuciones.

El proyecto se ha desarrollado utilizando como marco de trabajo, jBPM [9][10]. En jBPM los procesos de negocio son un tipo particular de “activo de empresa”. Todos los activos de empresa manejados en un proyecto forman lo que se conoce como “base de conocimiento” (KB, *knowledge base*). Además de la definición de procesos se pueden incluir también en la KB: reglas de negocio, especificaciones de argumentos y respuestas de las tareas o tipos de datos, entre otros activos.

2.2. Uso de estándares

Los procesos de atención sanitaria actuales a menudo implican la colaboración de distintas administraciones, tanto a nivel intra como interhospitalario. Esta circunstancia entraña retos de integración, administrativos y técnicos. Si las soluciones técnicas utilizan estándares se facilita su interoperatividad con otras, y por tanto la ejecución de procesos en los que participen diversas organizaciones será más fácil.

La aplicación desarrollada puede integrarse con cualquier *backend* de persistencia conforme a la API REST definida

en FHIR [11] ya que la información utilizada para la gestión de consentimientos sigue el formato de los recursos de este estándar. En concreto, se han utilizado los recursos *Questionnaire* [12] y *QuestionnaireResponse* [13] para la persistencia de las plantillas de consentimientos, de las solicitudes de consentimientos para los pacientes y de las respuestas de estos (los consentimientos).

3. Resultados

3.1. Objetivos

Los objetivos específicos de este proyecto han sido:

1. Identificar y modelar los procesos más elementales para la solicitud y concesión/denegación de consentimientos.
2. Implementar una aplicación autónoma que incluyera tanto la interfaz web para la ejecución de las tareas humanas implicadas en estos procesos (de pacientes o profesionales) como el motor de procesos.
3. Incorporar en los procesos tareas que automaticen la comunicación con servidores FHIR que actúen de *backend* de persistencia para los recursos implicados en la gestión de consentimientos.
4. Adoptar un diseño modular en la aplicación e incorporar una API REST que actúe de fachada de la API del motor de procesos, simplificando su uso para así facilitar el desarrollo de aplicaciones cliente.
5. Incorporar un mecanismo simple de control de acceso a la aplicación.

Se ha diseñado e implementado una “aplicación de negocios” [3] autónoma, que incluye tanto la interfaz de usuario para la ejecución de las tareas humanas implicadas en los procesos como el propio motor de procesos.

Se ha comenzado trabajando con los procesos más elementales:

- Creación de solicitudes de consentimientos explícitos por parte de los profesionales sanitarios.
- Recogida de las respuestas de los pacientes a estas solicitudes, con la otorgación o denegación del consentimiento.

3.2. Diseño de los procesos

Los desarrolladores de jBPM proporcionan un “*workbench*” denominado “*Business central*” (BC). Consiste en un servidor, con una consola web de gestión, que abarca la gestión de todo el ciclo de vida de un proceso. Permite especificar los activos de empresa y desplegarlos en contenedores o motores de procesos, gestionar las tareas humanas y los contenedores, monitorizar las instancias de proceso en ejecución y los históricos, ofreciendo además interfaces de configuración, repositorio de artefactos, etc. Sin embargo, dado que en este proyecto se requería una aplicación autónoma, sólo se utiliza BC para la edición de la base de conocimiento, que posteriormente se importa en el proyecto y se despliega en el motor embebido en la aplicación.

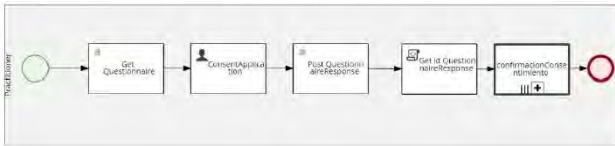


Figura 2. Proceso para la solicitud de consentimientos

La Figura 2 presenta el proceso principal, iniciado por un profesional sanitario para solicitar consentimientos a pacientes. En la primera actividad se recupera del *backend*, un servidor FHIR, lo que se ha denominado el “metacuestionario”, que está almacenado como un recurso *Questionnaire*. La siguiente tarea es una tarea humana que se asignará al mismo profesional que inició la instancia del proceso. Se le ofrece un formulario, basado en la información del metacuestionario, para personalizar el consentimiento que se quiere recoger. Esto hace que la aplicación desarrollada sea muy flexible, ya que configurar los consentimientos no implica recodificación. En esta tarea también se especifica a quién/es va dirigida la solicitud de consentimiento. La salida es un *QuestionnaireResponse*, que contiene la información del consentimiento, y la lista de destinatarios de la solicitud de consentimiento. Finalmente, se persiste en el *backend* esta respuesta y se ejecutan tantas instancias del subproceso *confirmacionConsentimiento* como usuarios a los que vaya destinada la solicitud (Figura 3). En este subproceso cada paciente procede a confirmar o denegar el consentimiento solicitado, y estas respuestas se persisten en el *backend* como nuevos recursos *QuestionnaireResponse*.

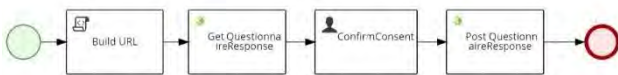


Figura 3. Proceso para la concesión o denegación de consentimiento por parte de los pacientes

Una vez especificados los procesos, la KB se importa en el proyecto de la aplicación y no es necesario volver a utilizar BC.

3.3. Casos de uso

La aplicación se ha desarrollado con Spring [14], y se ha utilizado el motor de plantillas Thymeleaf [15] para el desarrollo del *frontend*. Utiliza la librería HapiFHIR [16] para el manejo de los recursos FHIR. Esta primera versión utiliza un acceso basado en login y password, soportado por el marco de seguridad de Spring. La Figura 4 muestra la arquitectura tecnológica de la solución desarrollada.

Se han desarrollado los siguientes casos de uso:

- Generales
 - Registrar nuevo usuario
 - Inicio de sesión
- Específicos de profesionales sanitarios
 - Solicitud de nuevo consentimiento
 - Listar solicitudes enviadas
 - Detallar una solicitud enviada
- Específicos de pacientes
 - Ver solicitudes recibidas
 - Confirmar/Denegar consentimiento

En el caso de uso “Solicitud de un nuevo consentimiento” el profesional sanitario configura el consentimiento que quiere solicitar, selecciona los pacientes a los que desea pedir consentimiento y se envía una solicitud a cada uno de ellos. Para ello la aplicación comienza arrancando una instancia del proceso principal (Figura 2). La Figura 5 muestra el formulario que se presenta al profesional sanitario en la primera tarea humana del mismo. Podemos ver que el metacuestionario que se utiliza para la configuración del consentimiento es muy abierto y flexible. En cualquier caso, si se identificaran nuevos requisitos, la modificación no implica recodificación de la aplicación, ya que sólo sería necesario editar el recurso *Questionnaire* en el que está basado.

Al final de este caso de uso se crean tantas instancias del proceso para la concesión/denegación (Figura 3) como pacientes se hayan seleccionado.

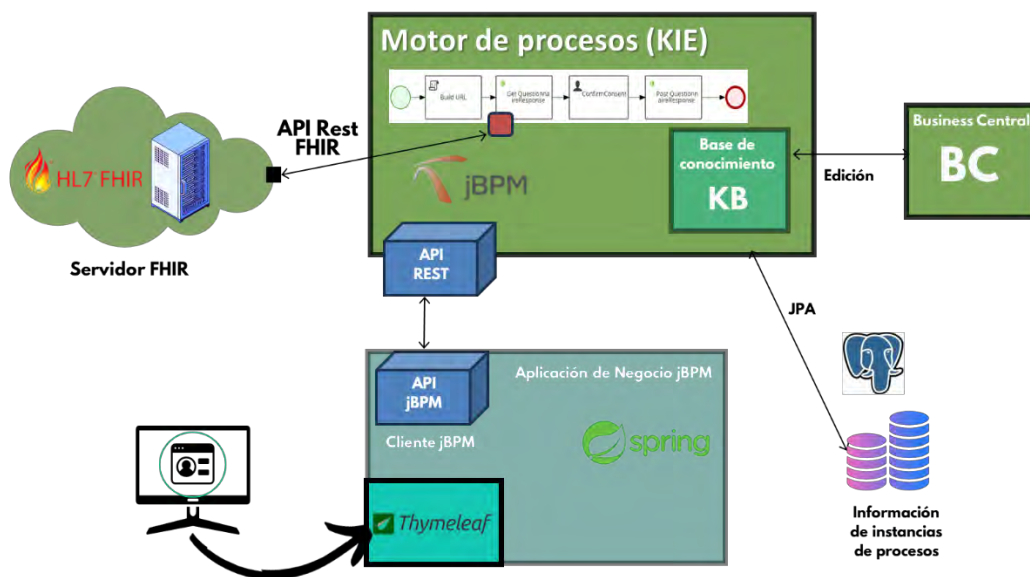


Figura 4. Arquitectura tecnológica de la solución

Figura 5. Formulario para la solicitud de consentimientos

El caso de uso “Confirmar/Denegar consentimiento” permite al paciente leer los detalles del consentimiento que se le solicita y aceptarlo o rechazarlo. El resto de casos de uso están relacionados con el registro y autenticación de usuarios y con las consultas de consentimientos.

4. Conclusiones

La aplicación presentada en este trabajo es el resultado final de un TFG no becado, y representa la versión inicial de un servicio, más complejo, para la gestión de consentimientos. La aplicación gestiona los dos procesos más elementales de la gestión de consentimientos, la solicitud y la autorización.

Esta primera versión nos ha permitido conocer la arquitectura de las aplicaciones de negocio basadas en jBPM y tener una visión general de las amplias posibilidades que este paradigma aporta. Pero es evidente que no es una solución final; los procesos ya desarrollados necesitan ser refinados, hay procesos que aún no se han considerado y, por supuesto, sería necesaria su integración con los otros servicios de la organización que decidiera utilizarla, gestión de usuarios, identificación, control de acceso, persistencia, etc...

La aportación de esta aplicación es principalmente metodológica. La adopción del paradigma BPM en las organizaciones sanitarias puede tener grandes beneficios, pero es compleja. La incorporación de aplicaciones autónomas, enfocadas en procesos concretos pero con interfaces abiertas que permitan que estos se reutilicen como subprocesos en otros más complejos, gestionados en otras aplicaciones, puede facilitar la adopción de BPM. En lugar de tener un único servicio de gestión de procesos, complejo y por tanto difícil de desarrollar, probar, mantener... se puede diseñar un ecosistema de servicios especializados (en la gestión de determinados procesos, en funciones de análisis, etc.) que puedan cooperar, y que irá creciendo al ritmo que la organización pueda asumir. La aplicación desarrollada es un ejemplo de uno de estos servicios, su incorporación no interrumpe el trabajo en la organización, pero la prepara para su evolución hacia el uso de BPM.

Referencias

- [1] Regulation (EU) 2016/679. General Data Protection Regulation. Disponible online: <http://data.consilium.europa.eu/doc/document/ST-5419-2016-INIT/en/pdf> (Consultada: septiembre 2023).
- [2] Mathias Weske. Business Process Management: Concepts, Languages, Architectures. Springer-Verlag Berlin Heidelberg 978-3-540-73522-9 Published: 19 September 2007 doi: 10.1007/978-3-540-73522-9
- [3] jBPM business applications. Documentación disponible online <https://www.jbpm.org/businessapps/gettingStarted.html> (Consultada: septiembre 2023):
- [4] Business Process Model and Notation (BPMN). Version 2.0. Enero 2011. Disponible online. <http://www.omg.org/spec/BPMN/2.0> (Consultada: septiembre 2023):
- [5] Schmidt, W. (2013). Business Activity Monitoring (BAM). Chapter 5 In: Rausch, P., Sheta, A., Ayesh, A. (eds) Business Intelligence and Performance Management. Advanced Information and Knowledge Processing. Springer, London. https://doi.org/10.1007/978-1-4471-4866-1_15
- [6] J. -P. Friedenstab, C. Janiesch, M. Matzner and O. Muller, "Extending BPMN for Business Activity Monitoring," *2012 45th Hawaii International Conference on System Sciences*, Maui, HI, USA, 2012, pp. 4158-4167, doi: 10.1109/HICSS.2012.276.
- [7] J. G. Kang and K. H. Han, "A Business Activity Monitoring System Supporting Real-Time Business Performance Management," *2008 Third International Conference on Convergence and Hybrid Information Technology*, Busan, Korea (South), 2008, pp. 473-478, doi: 10.1109/ICCIT.2008.224.
- [8] Costello, C., Molloy, O. (2009). Building a Process Performance Model for Business Activity Monitoring. In: Wojtkowski, W., Wojtkowski, G., Lang, M., Conboy, K., Barry, C. (eds) Information Systems Development. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-68772-8_19
- [9] Fiorini, & Gopalakrishnan, A. V. (2015). Mastering jBPM6 design, build, and deploy business process-centric applications using the cutting-edge jBPM technology stack. Packt Publishing.
- [10] De Maio, Salatino, M., Aliverti, E., & Dal Monte, E. (2014). *jBPM6 developer guide: learn about the components, tooling, and integration points that are part of the JBoss Business Process Management (BPM) framework* (Third edition.). Packt Publishing.
- [11] Especificación de la API REST de FHIR, disponible online <https://www.hl7.org/fhir/http.html> (Consultada: septiembre 2023):
- [12] Recurso FHIR “Questionnaire” release 5. Detalles disponibles online <https://www.hl7.org/fhir/questionnaire.html> (Consultada: septiembre 2023)
- Recurso FHIR “QuestionnaireResponse” release 5. Detalles disponibles online <https://www.hl7.org/fhir/questionnaire-response.html> (Consultada: septiembre 2023)
- [13] Página web del Framework Spring. <https://spring.io/> (Consultada: septiembre 2023)
- [14] Página web del motor de plantillas Thymeleaf <https://www.thymeleaf.org/> (Consultada: septiembre 2023)
- [15] Página web del motor de HapiFhir <https://hapifhir.io/> (Consultada: septiembre 2023)

EMMA: una nueva plataforma para la monitorización y acompañamiento de personas con endometriosis

F.J.Gárate¹, Emanuela Spagnolo², B.González¹, Ana López², P.Sanchez-González^{1,3}, E.J. Gómez^{1,3}, Alicia Hernández²

¹ Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España; fgarate/p.sanchez@upm.es

² Unidad de Endometriosis, Hospital Universitario La Paz, Madrid, España

³ Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Madrid, España

Resumen

La endometriosis es la causa más común de dolor pélvico crónico y afecta al 10%-20% de las personas que menstrúan en edad fértil. A pesar del tratamiento clínico, muchas de estas pacientes no logran controlar el dolor, con la consecuencia de restricciones más o menos graves en la vida laboral, social, familiar de pareja y sexual. Este trabajo presenta EMMA, una plataforma tecnológica que incluye una aplicación móvil para pacientes con endometriosis y una aplicación web para profesionales clínicos. El objetivo de EMMA es mejorar el seguimiento clínico de la enfermedad, entendiendo mejor los patrones y desencadenantes asociados a la endometriosis, identificando tendencias y evaluando la eficacia de los tratamientos, así como, ofreciendo soporte a las personas que padecen la enfermedad.

1. Motivación

La endometriosis es una enfermedad crónica en la que el endometrio crece fuera del útero, siendo las localizaciones más frecuentes el peritoneo pélvico o los ovarios, aunque también puede encontrarse en el sistema gastrointestinal o urinario [1]. Se estima que esta patología afecta al 10-15% de las personas que menstrúan en edad reproductiva, aunque si se considera de forma aislada la población de pacientes jóvenes con dolor pélvico las cifras oscilan del 50-60% [2], mientras que entre las pacientes que presentan subfertilidad el porcentaje de endometriosis ronda el 25-50% [3].

Algunos de los síntomas más comunes de la endometriosis son el dolor pélvico intenso durante el período menstrual (dismenorrea), dolor durante las relaciones sexuales (dispareunia), sangrado menstrual abundante o irregular, dolor al orinar o defecar, fatiga y dificultad para concebir o infertilidad [4], [5].

Además, la endometriosis, como las enfermedades crónicas, tiene un impacto significativo en la salud psicológica y el bienestar emocional de las personas que las padecen. La carga física constante, el dolor persistente e incapacitante, las limitaciones en la vida diaria y la incertidumbre sobre el futuro pueden desencadenar una variedad de respuestas psicológicas negativas, como el estrés, la ansiedad, la depresión y una disminución de la calidad de vida de las pacientes [6], [7].

El Observatorio Global por la eHealth (GOe) de la Organización Mundial de la Salud (OMS) define el término *mobile health* (mHealth) o salud móvil como una práctica médica que se apoya en las tecnologías móviles [8]. Las tecnologías móviles desempeñan un papel cada vez más importante en el acompañamiento de las personas que viven con enfermedades crónicas. Estas herramientas

digitales ofrecen una serie de beneficios, mejoran el autocuidado y la gestión de la enfermedad, proporcionan información y educación sobre la enfermedad mejorando la calidad de vida de las pacientes.

Así, las aplicaciones móviles y las plataformas en línea ofrecen acceso a recursos educativos, consejos de estilo de vida saludable y pautas para el tratamiento de la enfermedad. De esta forma, se fomenta la implicación de las pacientes al brindarles conocimientos y herramientas para comprender mejor su condición y tomar decisiones informadas sobre su cuidado [10], [11]. Además, las tecnologías móviles facilitan el seguimiento y la monitorización de los síntomas, el progreso y la adherencia al tratamiento. Estas herramientas permiten a las pacientes registrar y realizar un seguimiento de sus síntomas, medicamentos o citas médicas. Esto proporciona una visión más completa de su salud y permite a los profesionales sanitarios tomar decisiones más informadas sobre el manejo de la enfermedad [12].

En este trabajo se describe el diseño y desarrollo de una plataforma mHealth de monitorización de síntomas, medicamentos y evolución clínica de personas que padecen endometriosis. La plataforma, EMMA, permite a los profesionales clínicos facilitar a los pacientes información relacionada con la enfermedad de una forma regulada y supervisada, abordando el acompañamiento de las personas con endometriosis de una manera integral.

Se han encontrado aplicaciones cuyo único objetivo es la recogida de datos en forma de cuestionarios para su posterior estudio. También se han encontrado algunas plataformas a las que a este seguimiento se les suma proporcionar consejos y recomendaciones de estilo de vida. Incluso algunas de ellas están vinculadas a sistema de salud de su país correspondiente (Endo-App [13] y Luna [14] en Alemania y Francia respectivamente) y cuentan con el Certificado Europeo (CE) de dispositivo médico. Sin embargo, en España no se ha encontrado ninguna aplicación que ofrezca una plataforma de apoyo y monitorización a personas con endometriosis que esté vinculada y supervisada por profesionales clínicos pertenecientes a uno o varios centros médicos.

2. Diseño de la plataforma EMMA

El diseño de la plataforma se realizó en colaboración entre la Unidad de Endometriosis del Hospital Universitario La Paz (HULP) y el Grupo de Bioingeniería y Telemedicina (GBT) de la Universidad Politécnica de Madrid (UPM). Tras un detallado estudio del problema clínico, una revisión del estado del arte y diferentes sesiones de co-diseño, se llegó a la solución de una plataforma tecnológica formada por:

- Una **aplicación móvil** con dos funcionalidades principales: (1) recogida de datos a partir de cuestionarios, tanto para la realización de un seguimiento regular de la sintomatología como el estado anímico de la paciente. (2) recomendaciones clínicas y noticias destacadas. Estas recomendaciones y noticias estarán supervisadas por los profesionales clínicos.
- Una **aplicación web** en la que los profesionales clínicos puedan consultar los datos recogidos en la aplicación móvil de una forma concisa y gráfica. Además, podrán escribir, verificar y actualizar la sección de recomendaciones que se visualizará en la aplicación móvil.

2.1. Requisitos funcionales y no funcionales

Toda la información recopilada en las reuniones de co-diseño fue formalizada mediante la especificación de requisitos, donde se recogió de una manera precisa y estandarizada el comportamiento de la plataforma y las funcionalidades que debe proveer. La tabla 1 recoge un resumen de los requisitos funcionales identificados.

Funcionalidades
<i>Aplicación Web</i>
Autenticación.
Gestión de clínicas.
Gestión de usuarios.
Gestión de pacientes.
Gestión de consejos.
Visualización de datos
<i>Aplicación Móvil</i>
Autenticación.
Registro cuestionario.
Visualización consejos / recomendaciones

Tabla 1. Requisitos funcionales

2.2. Arquitectura de la plataforma

La arquitectura de la plataforma se muestra en la figura 2. Se pueden diferenciar tres aplicaciones diferentes: la aplicación móvil, la aplicación web y la API de comunicación.

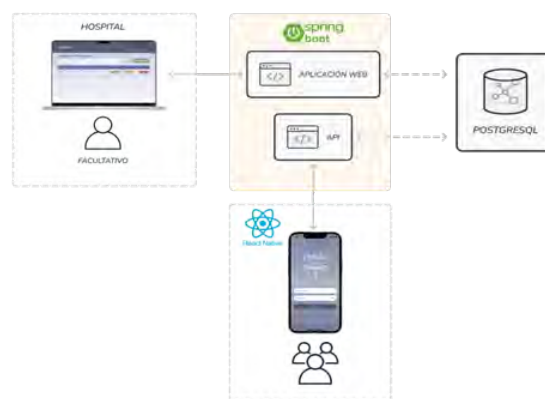


Figura 1. Arquitectura de la plataforma.

Los pacientes mediante la aplicación móvil podrán rellenar cuestionarios y recibir información sobre su patología, contribuyendo así a su autonomía y permitiendo gestionar su salud de forma activa. La app se desarrolla en React Native pudiendo ser instalada en dispositivos Android e iOS.

Desde la aplicación web el personal clínico podrá gestionar a las pacientes, ver su evolución y recuperar datos que le permitan realizar un tratamiento más eficaz y personalizado. Por otro lado, también será posible generar consejos con bases científicas y adaptados a las pacientes con endometriosis. La aplicación Web se desarrolla en java utilizando el *framework* Spring MVC; para la vista se emplea Thymeleaf, Bootstrap y Chartjs para la elaboración de las gráficas.

La API REST de comunicación actuará de puente entre las aplicaciones web y móvil, haciendo posible el intercambio de datos. Los resultados de los cuestionarios son enviados desde la aplicación móvil a la API de comunicación, donde serán almacenados en una base de datos PostgreSQL para su posterior consulta desde la aplicación web. También será el mecanismo utilizado para proporcionar información de interés a las pacientes. Al igual que la aplicación web la API se desarrolla en Java utilizando en este caso el *framework* Spring boot.

Para la transferencia de datos se ha seguido el estándar HL7 FHIR v5 referente al recurso cuestionario [15] que proporciona un control detallado sobre el orden, la presentación, la fraseología y la agrupación para permitir una recopilación de datos coherente y consistente.

3. Desarrollo de la plataforma EMMA

3.1. Aplicación móvil

Inicio de sesión

Para iniciar sesión en la aplicación móvil es necesario introducir unas credenciales acceso, las cuales deben ser obtenidas previamente en una consulta médica en la consulta de endometriosis del HULP. De esta manera, se asegura un acceso restringido a personas que han sido diagnosticadas previamente con endometriosis y son pacientes de La Paz.



Figura 2. Pantalla de inicio de sesión de la aplicación móvil y ejemplo de credenciales de acceso generadas.

Cuestionarios

Tras el primer inicio de sesión, la aplicación redirige a la paciente al cuestionario inicial, que recoge información clínica esencial para categorizar de forma demográfica a la paciente.

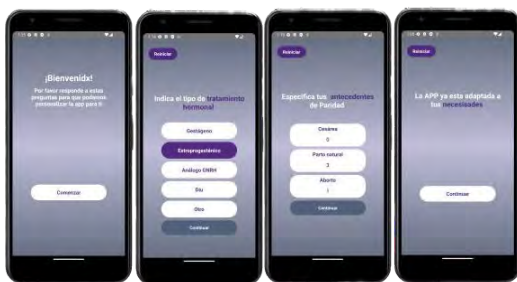


Figura 3. Pantallas del cuestionario inicial.

El cuestionario periódico consta de una recopilación de preguntas que han de ser respondidas de forma diaria, o en su defecto, con regularidad. Este cuestionario ha sido diseñado con la intención de estudiar la evolución de la sintomatología, necesidad de analgesia o cumplimiento del tratamiento de las pacientes.

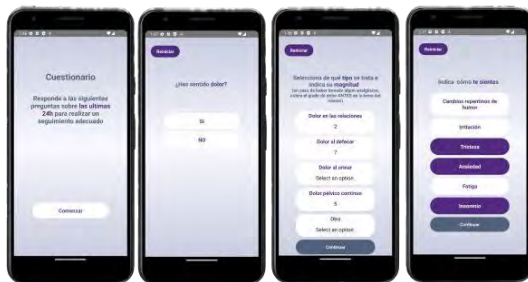


Figura 4. Pantallas del cuestionario periódico.

Consejos y recomendaciones

Los consejos y recomendaciones están relacionados con temas de interés vinculados a la endometriosis, desde recomendaciones sobre nutrición o deporte, hasta ejercicios de meditación o yoga.



Figura 5. Pantallas de consejos recomendaciones.

3.2. Aplicación web

Inicio

Desde la pantalla de inicio el usuario facultativo puede ver la evolución en el tiempo del número de pacientes registrados en la plataforma, así como los cuestionarios completados dentro de la clínica en la que este asociado.



Figura 6. Pantalla de inicio.

Visualización de datos

El usuario facultativo puede consultar los datos de las pacientes registradas en la plataforma, en esta pantalla se pueden consultar las respuestas del cuestionario inicial, ver los días en los que el paciente respondió el cuestionario periódico, pudiendo ver las respuestas de ese día en concreto.



Figura 7. Pantalla de visualización de datos.

El gráfico de intensidad del dolor muestra los datos de los distintos tipos de dolor que afectan a la paciente, su intensidad, en gris a modo de columna, la necesidad de toma de analgésicos, así como la intensidad del dolor percibida tras la toma del analgésico. También es posible consultar los días en los que la paciente ha presentado sangrado, trastornos gastrointestinales o anímicos.

Gestión de consejos y recomendaciones

Los usuarios facultativos pueden gestionar los consejos y recomendaciones a los que pueden acceder las pacientes desde la aplicación móvil. Gracias a esta utilidad las pacientes con endometriosis pueden acceder a una fuente fiable de información y tomar una actitud informada en lo relativo a sus estilos de vida.

Figura 8. Pestaña de creación de consejos recomendaciones

Validación

Este estudio ha tenido una duración de un mes, las pacientes fueron reclutadas en la consulta del HULP. Como criterios de inclusión se debía estar diagnosticada con endometriosis, disponer de un dispositivo móvil Android con acceso a *Google Play Store* y tener un nivel intermedio de comprensión lectora del castellano, finalmente, se ha contado con la participación de nueve pacientes. De las pacientes que instalaron la plataforma, un 20% no llegaron a entrar, mientras que, de las pacientes que entraron en la plataforma entorno al 40% abandonó su uso a la semana de descargarla y un 30% accedió a ella de forma regular.

4. Conclusiones y líneas futuras

Las aplicaciones móviles han demostrado ser herramientas poderosas en el seguimiento y acompañamiento de personas con endometriosis. La recopilación de datos confiables, precisos y completos sobre la endometriosis sigue siendo un desafío constante en el campo de la investigación médica. Este trabajo trata de mejorar la recopilación de datos y avanzar en la comprensión y el tratamiento de esta compleja enfermedad.

La plataforma desarrollada, EMMA, cumple con las especificaciones definidas y tanto las pacientes como los facultativos respondieron positivamente a la aplicación. Estas afirmaciones se basan en la realización de una validación técnica de la robustez, eficiencia, compatibilidad y funcionalidad, demostrando que el software desarrollado tiene la calidad suficiente para dar viabilidad técnica a la plataforma tecnológica *mHealth*.

Se contempla la integración con el sistema informáticos del centro a fin de recopilar distintos eventos clínicos, que actualmente son solicitados mediante el cuestionario inicial, así como incluir la información obtenida mediante los cuestionarios en la historia clínica del paciente.

Este estudio, tanto por su duración como por el número de participantes, tiene como objetivo ser un primer paso que permita recopilar información para futuros estudios. Se ha comprobado que la plataforma tecnológica puede dar soporte al estudio, ya que no se han producido incidencias técnicas reseñables durante la duración del mismo, la tasa de abandono puede ser debida a distintos factores. Por un lado, la no disponibilidad de aplicación en el *App Store de Apple* ha dejado fuera un segmento de la población que se presume podría ser más proclive al uso de soluciones digitales, y por otro, la metodología de la intervención seguida en este trabajo requiere de gran implicación por parte de los pacientes para completar los cuestionarios de forma diaria y la compensación en forma de consejos o recomendaciones ha sido pequeña.

Los resultados obtenidos evidencian la necesidad de elaborar una nueva intervención, con cuestionarios, periodicidad y recomendaciones personalizadas a las necesidades de los pacientes, que reduzca la tasa de abandono.

Agradecimientos

Agradecer a las pacientes y profesionales clínicos que han participado y han hecho posible este trabajo de investigación. Este trabajo se ha realizado dentro del marco de la comunidad EELISA Health in the City.

Referencias

- [1] G. A. J. Dunselman et al., “ESHRE guideline: management of women with endometriosis,”
- [2] D. P. Goldstein, C. De, and S. Jean Emansi, “Adolescent Endometriosis,” 1980.
- [3] “Endometriosis and infertility”, doi: 10.1016/j.fertnstert.2004.
- [4] N. P. Johnson and L. Hummelshoj, “Consensus on current management of endometriosis,” *Human Reproduction*, vol. 28, no. 6, pp. 1552–1568, 2013, doi: 10.1093/humrep/det050.
- [5] Bulun SE. Endometriosis. *N Engl J Med*. 2009 Jan 15;360(3):268-79. doi: 10.1056/NEJMra0804690. PMID: 19144942.
- [6] L. Culley et al., “The social and psychological impact of endometriosis on women’s lives: A critical narrative review,” *Hum Reprod Update*, vol. 19, no. 6, pp. 625–639, Nov. 2013, doi: 10.1093/humupd/dmt027.
- [7] A. L. Hsu and P. Stratton, “Invasive and Noninvasive Methods for the Diagnosis of Endometriosis.” [Online]. Available: www.clinicalobgyn.com
- [8] S. Ryu, “Book Review: *mHealth: New Horizons for Health through Mobile Technologies: Based on the Findings of the Second Global Survey on eHealth (Global Observatory for eHealth Series, Volume 3)*,” *Healthc Inform Res*, vol. 18, no. 3, p. 231, 2012, doi: 10.4258/hir.2012.18.3.231.
- [9] Brancozzi, A. (2021). *Mobile cross-platform gesture- guided visual pain tracking for endometriosis (Dissertation)*. Retrieved from <https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-294331>
- [10] J. A. Cafazzo, M. Casselman, N. Hamming, D. K. Katzman, and M. R. Palmert, “Design of an *mHealth* App for the Self-management of Adolescent Type 1 Diabetes: A Pilot Study,” *J Med Internet Res*, vol. 14, no. 3, p. e70, May 2012, doi: 10.2196/jmir.2058.
- [11] C. Laloo, L. A. Jibb, J. Rivera, A. Agarwal, and J. N. Stinson, ““There’s a Pain App for That,”” *Clin J Pain*, vol. 31, no. 6, pp. 557–563, Jun. 2015, doi: 10.1097/AJP.000000000000171.
- [12] M. F. Alwashmi et al., “Perceptions of Patients Regarding Mobile Health Interventions for the Management of Chronic Obstructive Pulmonary Disease: Mixed Methods Study,” *JMIR Mhealth Uhealth*, vol. 8, no. 7, p. e17409, Jul. 2020, doi: 10.2196/17409.
- [13] “Endo-App - Aplicaciones en Google Play.” <https://play.google.com/store/apps/details?id=app.endometriose.prod&hl=es&gl=US> (accessed Jun. 07, 2023).
- [14] “Luna - Aplicaciones en Google Play”, Accessed: Jun. 17, 2023. [Online]. Available: https://play.google.com/store/apps/details?id=com.hdsi.luna&hl=es_419&gl=US
- [15] “HL7 FHIR v5 Resource Questionnaire.” <https://www.hl7.org/fhir/questionnaire.html> (accessed Jun. 07, 2023).

Arquitectura de Microservicios para Ensayos Clínicos en Pérdida de Peso con Plataformas mHealth

A. Cobo Sánchez de Rojas^{1,2}, F. Fernández Martín¹, G. García Sáez^{1,2}, J. Tapia Galisteo^{1,2}, C. González Blanco^{2,3}, R. Corcoy Pla^{2,3}, E. Ruiz Moreno^{4,5}, B. Pérez Gómez^{4,5}, M. Pollán Santamaría^{4,5}, O. Núñez^{4,5}, M.E. Hernando Pérez^{1,2}

¹ Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España, antonio.cobo@ctb.upm.es.

² CIBER de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, España

³ Servicio de Endocrinología y Nutrición, Hospital de la Santa Creu i Sant Pau, Barcelona, España

⁴ CIBER de Epidemiología y Salud Pública, Instituto de Salud Carlos III, Madrid, España

⁵ Centro Nacional de Epidemiología, Instituto de Salud Carlos III, Madrid, España

Resumen

Las terapias contra el sobrepeso y la obesidad se basan en la modificación del estilo de vida sin embargo la adherencia es baja porque los estilos saludables son difíciles de mantener en el tiempo. Las funcionalidades adecuadas para que las aplicaciones móviles (mHealth) faciliten el cambio permanecen sin identificar. El presente trabajo presenta una arquitectura mHealth de microservicios para estudios clínicos con los que identificarlas. La metodología de diseño y desarrollo incluyó historias de usuario, Domain-Driven Design (DDD) y Scrum. Los microservicios identificados, su definición, y la organización de nuestra arquitectura son suficientes para realizar estudios clínicos en pérdida de peso con plataformas mHealth. Entre los trabajos futuros más inmediatos destaca la conexión de la presente arquitectura con servicios de gestión de estudios clínicos de uso común como RedCap.

1. Introducción

La obesidad y el sobrepeso son dos enfermedades crónicas caracterizadas por un exceso de acumulación de tejido adiposo en el cuerpo [1]. Ambas incrementan el riesgo de muerte y de desarrollar algunas comorbilidades como la diabetes tipo-2, la depresión, la enfermedad cerebrovascular, y algunas neoplasias [2]. Están consideradas por la OMS como una epidemia global con una prevalencia del 13 % en el caso de la obesidad y del 39 % en el del sobrepeso [3]. En particular, las cifras correspondientes a Europa prácticamente duplican las estimaciones a nivel mundial con el 60 % de los adultos y casi 1 de cada 3 niños afectados de sobrepeso u obesidad [4].

Aunque pueden estar originadas por factores biológicos como enfermedades de origen endocrino, la causa fundamental es un desequilibrio entre las calorías ingeridas y las que se gastan [5]. El tratamiento se basa en la adopción de hábitos alimentarios saludables y el incremento de los niveles de actividad física [5]. Sin embargo, las terapias tradicionales adolecen de una baja adherencia [6] ya que la promesa de beneficios futuros a cambio de renunciar a placeres inmediatos es un incentivo muy débil para el cerebro [7].

Se necesitan nuevos enfoques efectivos en el tiempo y económicamente viables. Las aplicaciones móviles (mHealth) son una alternativa prometedora. Sin embargo, las apps relacionadas con este problema clínico disponibles en el mercado suelen ser herramientas de autogestión; no cuentan con estrategias adecuadas de

motivación y soporte psicológico para los pacientes, ni con herramientas de seguimiento para los profesionales clínicos. Para solventar estas carencias se necesitan plataformas que sirvan como soporte a la realización de estudios clínicos que aporten evidencias científicas que respalden que las funcionalidades implementadas alcanzan su máximo potencial [8].

En la actualidad, existen herramientas de propósito general para administrar estudios clínicos que permiten recoger datos mediante cuestionarios. Sin embargo, la tendencia en el ámbito de la pérdida de peso es buscar herramientas de recogida menos intrusivas (como chatbots y sensores vestibles). Las arquitecturas de microservicios, en las que se centra este trabajo, tienen el potencial de integrar componentes heterogéneos de software, y pueden facilitar la integración de nuevas funcionalidades con las de plataformas ya existentes.

Las arquitecturas de microservicios se contraponen a las arquitecturas monolíticas. Todos los componentes de un monolito están conectados y se ejecutan en un único proceso. Por contra, los microservicios son unidades autónomas con responsabilidades bien delimitadas; se ejecutan de manera independiente unos de otros; y se comunican mediante interfaces bien definidas. Los microservicios tienen una gran capacidad de escalado, lo que permitiría dar servicio a un gran número de usuarios y convertir la plataforma en una fuente de datos que permitiera sortear las carencias de los estudios de epidemiología nutricional [9].

En el presente trabajo se presenta el proceso de diseño de una arquitectura mHealth de microservicios para la realización a estudios clínicos en pérdida de peso. Se comenzó por analizar el dominio de aplicación e identificar las funcionalidades necesarias. Se definieron los microservicios necesarios y se implementó la arquitectura para probarla con un diseño concreto de estudio clínico.

2. Metodología

Para el diseño de la arquitectura de microservicios se ha seguido la metodología descrita por Newman (2019) [10].

2.1. Especificación funcional

Los requisitos funcionales de las plataformas mHealth para estudios clínicos en pérdida de peso se han espe-

cificado usando el formato conocido como “historias de usuario”. Las historias de usuario son descripciones cortas de funcionalidades o requisitos desde la perspectiva del usuario final. Un ejemplo de historia de usuario es: “como investigador clínico quiero mover un estudio clínico del estado de diseño al estado de producción (activo)”.

Este método de especificación ya ha sido validado para el desarrollo ágil de aplicaciones médicas de soporte a la decisión [11]. Las historias de usuario son un medio más flexible y efectivo que otros métodos, como los casos de uso que se usan en el Unified Modeling Language (UML), ya que permiten la comunicación clara y concisa entre los desarrolladores y las partes interesadas y añadir niveles de detalle de forma iterativa si es necesario.

2.2. Domain-Driven Design (DDD)

DDD es un proceso de diseño de software que pone el énfasis en comprender y modelar las particularidades del entorno empresarial en el que el sistema va a funcionar. En el caso del presente trabajo, se trata de un entorno de investigación clínica en aplicaciones mHealth para perder peso. Los desarrolladores y los médicos y nutricionistas del equipo de trabajo identificaron los diferentes dominios de dicho entorno mediante un proceso iterativo. Se comenzó por establecer un lenguaje ubicuo, es decir, un vocabulario común entre desarrolladores y especialistas. Después, las funcionalidades se agruparon por dominios con responsabilidades diferenciadas para que las fronteras entre cada dominio y los demás quedaran claramente definidas. De esta manera, cada uno de los dominios pudo implementarse con un microservicio diferente e independiente con su propia base de datos, lógica de negocio, e interfaz de usuario.

2.3. Definición de la arquitectura de referencia

La Tabla 1 muestra los principios de diseño que se siguieron durante la definición de la arquitectura de referencia.

Nombre
Separación de funcionalidades
Desacoplamiento de componentes
Comunicación entre microservicios
Gestión distribuida de los datos
Escalabilidad y Alta disponibilidad
Gestión descentralizada
Tolerancia a errores
Migración incremental

Tabla 1. Principios de diseño de la arquitectura de referencia

2.4. Implementación de los microservicios

La implementación consistió en realizar un desarrollo software. Las metodologías ágiles se vienen utilizando como método preferente de desarrollo software [12]. Éstas consisten en un proceso de desarrollo incremental e iterativo en el que en cada iteración se libera un producto completamente operativo con funcionalidades adicionales respecto de la iteración anterior. En concreto, se siguió la metodología Scrum.

Los requisitos del sistema completo (las historias de usuario) se organizaron por orden de prioridad en una lista de tareas pendientes, también conocida como backlog. Las historias de usuario inicialmente definidas eran demasiado grandes como para implementarlas en una sola iteración (lo que en Scrum se conoce como sprint); así pues, las más prioritarias se descompusieron en historias de usuario más manejables. Durante la implementación de los microservicios se utilizaron sprints de una y dos semanas de duración.

2.5. Prueba de los servicios

Se realizaron pruebas unitarias y pruebas de integración para comprobar que la lógica de negocio de cada microservicio ejecutaba las operaciones correctas con los resultados esperados, que los servicios se comunicaban entre sí conforme a lo esperado, y que la información persistente se almacenaba correctamente en la base de datos correspondiente. Los tests se realizaron haciendo uso de Postman; una herramienta de soporte al desarrollo de APIs.

3. Resultados y discusión

3.1. Especificación funcional

Las historias de usuario aptas para ser implementadas en un sprint son muy numerosas; así que no está justificado definir las historias de usuario en una jerarquía con tres niveles. Debido a las limitaciones de espacio del presente documento, en la Figura 1 sólo se muestran los dos primeros; estando diferenciados en la parte superior los requisitos para profesionales (investigadores y clínicos) de los requisitos para particulares (los pacientes y la población general) en la parte inferior.

Level 1	Manage professional account	Visit a patient	Conduct clinical studies	Analyze data
	Register into the system	Manage patients	Design a study	Get data set
	Log in the system	Visualize patient's dashboard	Manage invitations	...
	Configure account	Prescribe care plan	Launch a study	
Level 2			Recruit participants	
			Monitor a study	
			Finish a study	
Level 1	Manage individual account	Receive assistance	Participate in a study	
	Register into the system	Visualize my dashboard	Manage participation	
	Log in the system	Enter data		
Level 2	Configure account	Receive feedback		
		Access educational material		

Figura 1. Primeros dos niveles de la jerarquía de requisitos.

3.2. Domain-Driven Design (DDD)

Los requisitos se dividieron en dos grandes grupos. Por un lado, aquellos que cubren las tareas administrativas; y por el otro, aquellos que cubren las tareas de recolección de datos y procesamiento. Las tareas administrativas incluyen la creación y gestión de estudios clínicos y la administración de las cuentas de usuario (tanto profesionales como particulares). Por otro lado, las tareas de recolección y procesamiento acarrearán la siguiente serie de procesos: entrada de datos, descomposición nutricional, cálculo de indicadores agregados con significado clínico, y finalmente, visualización, análisis estadístico y reacción automática frente a los datos agregados. La

Figura 2 muestra los microservicios definidos de acuerdo a las responsabilidades descritas.

3.3. Arquitectura de referencia

La Figura 3 muestra la arquitectura de la plataforma. Está compuesta por diferentes microservicios empaquetados en contenedores para que puedan desplegarse independientemente unos de otros. De esta forma, las instituciones de investigadores y clínicos pueden optar por alojar los datos de sus estudios en los servidores centrales del proveedor de servicio o en sus propios servidores.

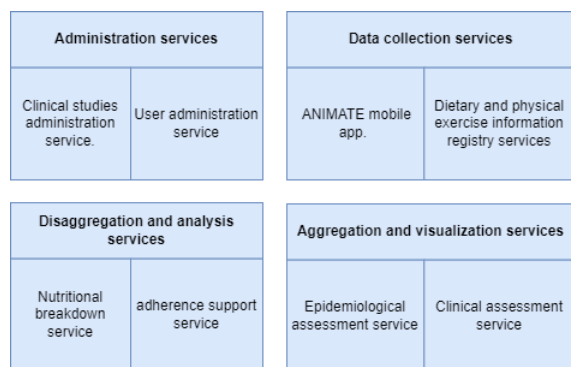


Figura 2. Microservicios resultantes del proceso de Domain-Driven Design.

Durante la definición de un estudio, los investigadores indicarán la opción adecuada (central, externa o archivo) y los datos de configuración correspondientes, de acuerdo con las instrucciones de sus administradores de sistemas.

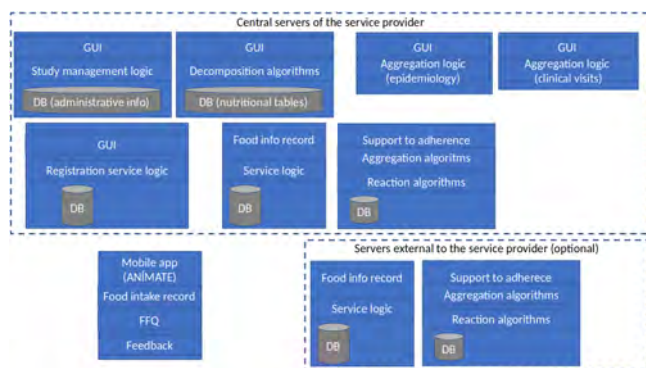


Figura 3. Arquitectura de referencia.

- Administración de estudios clínicos: gestiona y coordina los estudios clínicos realizados en la plataforma. Permite que un profesional diseñe y administre estudios que pueden estar en 3 estados: 'diseño' (se pueden hacer cambios a los parámetros de diseño como la duración), 'activo' (el diseño permanece inalterable y se pueden añadir participantes), y 'terminado' (no se permiten modificaciones de ningún tipo y los participantes quedan libera-

dos).

- Administración de usuarios: Permite que los usuarios se registren en el sistema y se encarga de la autenticación de los mismos cuando quieren ingresar en la plataforma.
- Aplicación móvil: permite a los pacientes y a los participantes en un estudio interactuar con el sistema. En concreto, permite el registro de datos de la ingesta, de ejercicio físico, del peso y de las variables psicológicas; permite consultar el progreso del tratamiento; y es el canal a través del cuál los pacientes reciben los mensajes de realimentación del servicio de soporte a la adherencia (recordatorios, alarmas, notificaciones, recomendaciones, felicitaciones y mensajes educativos).
- Algoritmos de descomposición: este servicio utiliza como información de entrada la información de ingesta recogida por la aplicación móvil y la descompone en su contenido nutricional en función de las tablas nutricionales seleccionadas.
- Registro de datos de estilo de vida: este servicio recibe la información de recogida por la aplicación móvil y la almacena en una base de datos. En el caso de la información de ingesta, hace uso del servicio de descomposición nutricional para obtener la información nutricional de los datos de ingesta de los pacientes y los almacena en una base de datos.
- Soporte a la adherencia: este servicio consulta la información almacenada por el servicio de registro de datos y la procesa con varios algoritmos de análisis inteligente que reaccionan frente al comportamiento y la evolución del paciente generando mensajes de realimentación del servicio (recordatorios, alarmas, notificaciones, recomendaciones, felicitaciones y mensajes educativos)
- Servicios de agregación y visualización: los epidemiólogos y profesionales clínicos utilizan estos servicios para generar y visualizar informes de análisis nutricionales e informes de progreso.

3.4. Implementación de los microservicios

Cada uno de los servicios implementados consiste en un servicio Web a través del cuál el microservicio se comunica con otros microservicios, una lógica de aplicación que ejecuta las tareas de las que el microservicio es responsable, y una base de datos que almacena los datos pertenecientes al dominio de negocio que el microservicio representa. Los servicios con los que los usuarios finales interactúan directamente también incorporan una interfaz de usuario.

Para la comunicación entre microservicios se utilizó el protocolo HTTPS junto con el formato JSON para el intercambio de datos.

Se priorizó la implementación de los servicios de gestión de estudios clínicos, administración de usuarios, registro de datos, y soporte a la adherencia.

La recogida y descomposición de información alimentaria requiere información detallada sobre productos disponibles en los supermercados, ingredientes de diferentes recetas, así como los valores nutricionales de todos ellos. Esta información es complicada y cara de conseguir. Para solventar ese problema, se integró en la plataforma una aplicación de terceros (FatSecret). Esta aplicación hizo una doble función; por un lado hizo las veces de una implementación digital de registro dietético mediante la cual los pacientes pueden reportar sus ingestas desde su smartphone. Por otro lado, hizo las veces del servicio de descomposición nutricional de nuestra arquitectura.

El servicio de registro de datos se implementó para que recuperara tanto la información de las ingestas como la de su descomposición nutricional mediante consultas a la API de FatSecret y almacenara los resultados en su propia base de datos. La lógica de aplicación tanto del servicio de registro como del de soporte a la adherencia servicios está implementada en Java, mientras que sus servicios Web de comunicación están implementados con la librería Python Fast API y con nodejs, respectivamente. Los servicios Web se programaron en un lenguaje diferente buscando velocidad y versatilidad. Las bases de datos de ambos servicios se implementaron en MySQL 8.0.32.

Tanto la lógica de aplicación como el servicio Web del servicio de administración de estudios clínicos se implementaron en Python 3.11.2. La base de datos también se implementó en MySQL 8.0.32. Finalmente, se utilizó el framework de desarrollo IONIC 5.4.10 para implementar una interfaz de usuario temporal. En la actualidad, RedCap es una herramienta en auge para administrar estudios clínicos que permite recoger datos mediante cuestionarios; y se ha comenzado a explorar la posibilidad de integración del servicio de administración de estudios con RedCap.

El servicio de administración de usuarios consiste fundamentalmente en un servicio de registro y autenticación. Tanto su lógica de aplicación como su servicio Web se implementaron en Python 3.11.2 y se apoya en Firebase 11.24.12. Finalmente, se utilizó el framework de desarrollo IONIC 5.4.10 para implementar la interfaz de usuario.

3.5. Prueba de los servicios

Se realizaron las siguientes pruebas unitarias; todas ellas con resultados satisfactorios.

- Creación de estudios clínicos y modificación de sus parámetros.
- Recuperación y almacenamiento de datos alimentarios y ejercicio físico.
- Registro de usuarios en el sistema.

Se realizaron las siguientes pruebas de integración; todas ellas con resultados satisfactorios.

- Ciclo de vida de un estudio clínico.
- Ciclo de vida de un participante dentro de un estudio clínico.

- Bucle de registro de datos, evaluación de estilo de vida, generación de realimentación y generación de informes.

4. Conclusiones

Los microservicios identificados, así como su definición, y la organización de los mismos descrita en la arquitectura de referencia son suficientes para realizar estudios clínicos en pérdida de peso con plataformas mHealth. Entre los trabajos futuros más inmediatos destaca la conexión de la presente arquitectura con servicios de gestión de estudios clínicos de uso común como RedCap.

Agradecimientos

Esta investigación fue financiada por CIBER - Consorcio Centro de Investigación Biomédica en Red (CB06/01/0051), el proyecto ANÍMATE 2 (CIBER-BBN early stage intramural projects 2022) y el proyecto ALEVINT (CIBERESP-BBN collaboration projects 2022), Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación.

Referencias

- [1] A. Engin. *The Definition and Prevalence of Obesity and Metabolic Syndrome*. Springer International Publisher, 2017.
- [2] C. Andolfi *et ál*. Epidemiology of Obesity and Associated Comorbidities. *J Laparoendosc Adv Surg Tech A*, 28:919–924, 2018.
- [3] W.P.T. James. WHO recognition of the global obesity epidemic. *Int J Obes*, 32:S120–S126, 2008.
- [4] Who european regional obesity report 2022. Technical report, WHO Regional Office for Europe, 2022.
- [5] S.M. Green. Obesity: prevalence, causes, health risks and treatment. *Br J Nurs*, 6:1181–1185, 1997.
- [6] F.L. Greenway. Physiological adaptations to weight loss and factors favouring weight regain. *Int J Obes*, 39:1188–1196, 2015.
- [7] E. Papiés. The Psychology of Desire and Implications for Healthy Hydration. *Ann Nutr Metab*, 76:31–36, 2021.
- [8] V. Alcántara-Aragón *et ál*. Web Support for Weight-Loss Interventions: PREDIRCAM2 Clinical Trial Baseline Characteristics and Preliminary Results. *Diabetes Technol Ther*, 20:380–385, 2018.
- [9] S. del Pozo de la Calle *et ál*. Sources of information on food consumption in Spain and Europe. *Nutr Hosp*, 31:29–37, 2015.
- [10] S. Newman. *Building Microservices, 2nd ed*. O'Reilly Media, 2021.
- [11] V. Kannan *et ál*. User stories as lightweight requirements for agile clinical decision support development. *J Am Med Inform Assoc*, 26:1344–1354, 2019.
- [12] P. Kokol. Agile Software Development in Healthcare: A Synthetic Scoping Review. *Applied Sciences*, 12:9462, 2022.

Herramienta de salud digital IMIDoc para la monitorización de pacientes con enfermedades inflamatorias inmunomediadas

J.M. Iniesta-Chamorro^{1*}, D. Benavent², M. Novella-Navarro³, C. Plasencia-Rodríguez³, A. Balsa³, E.J. Gómez^{1,4}

¹ Biomedical Engineering and Telemedicine Centre, ETSIT, Center for Biomedical Technology, Universidad Politécnica de Madrid, Spain, *jm.iniesta@upm.es

² Hospital Universitario de Bellvitge, Servicio de Reumatología, Barcelona, Spain

³ Hospital Universitario La Paz, IdiPaz, Servicio de Reumatología, Madrid, Spain

⁴ Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, CIBER, Madrid, Spain

Resumen

Las enfermedades reumáticas y musculoesqueléticas (ERM) son un grupo heterogéneo de enfermedades crónicas que afectan principalmente a articulaciones, huesos, músculos, tendones y ligamentos, aunque también pueden afectar a otros órganos y sistemas. La creciente importancia de estas enfermedades se debe a su prevalencia y al impacto que tienen en la calidad de vida de los pacientes, en particular, enfermedades inflamatorias inmunomediadas (IMID) como la artritis reumatoide (AR) y la espondiloartritis (EspA) producen gran discapacidad funcional y cursan con brotes periódicos, por lo que es importante un seguimiento y monitorización estrecha de la enfermedad. En la actualidad, la salud digital es una realidad y su uso en el campo de la reumatología requiere de un modelo asistencial que incorpore las tecnologías de la información y comunicaciones (TICs) y la implementación efectiva de la digitalización en la práctica clínica. En este trabajo se aborda el diseño y desarrollo de la herramienta IMIDoc para el seguimiento clínico, educación y empoderamiento de los pacientes con IMIDs favoreciendo una gestión más efectiva de la salud y eficiente que los métodos de atención tradicionales. Esta herramienta será evaluada en un estudio clínico multicéntrico en el que participarán más de 200 pacientes y equipos clínicos de 5 hospitales.

1. Introducción y objetivo

Las enfermedades reumáticas y musculoesqueléticas (ERM), un grupo heterogéneo de enfermedades que afectan a articulaciones, huesos, músculos, tendones y ligamentos, tienen cada vez más importancia debido a su prevalencia e impacto en la vida de los pacientes. Entre cientos de miles y millones de personas en todo el mundo se ven afectadas por enfermedades como la artritis reumatoide (AR) y la espondiloartritis (EspA), con una prevalencia mundial estimada del 0,46% (intervalo de confianza [IC] del 95%: 0,39-0,54) y del 0,2%-1,6%, respectivamente [1]. Estas enfermedades se caracterizan por síntomas que incluyen dolor y/o inflamación que pueden presentarse con reagudizaciones que requieren una evaluación profesional y modificaciones del tratamiento. Estos brotes pueden repercutir negativamente en la calidad de vida de los pacientes, por lo que es esencial identificarlos y actuar lo antes posible. El tratamiento clínico de estas enfermedades requiere un estrecho seguimiento de los síntomas físicos y psicológicos, lo que no siempre es factible en un entorno ambulatorio debido a las limitaciones de tiempo y la carga asistencial [2]. Por estas razones, la telemedicina y las tecnologías de salud

móvil (mHealth) representan una nueva oportunidad para la gestión clínica de los pacientes crónicos.

Las mediciones o resultados de salud reportados por el paciente electrónicamente (electronic Patient Reported Outcomes, ePROs) pueden obtenerse mediante soluciones como la mHealth, lo que apoya la detección de los síntomas que aparecen entre las visitas presenciales y permite el seguimiento de síntomas que los reumatólogos podrían pasar por alto durante las visitas debido a las limitaciones de tiempo comentadas [3].

Las soluciones digitales promueven la autogestión de la salud, proporcionando información a los pacientes sobre la enfermedad, gestión de síntomas, etc. Sin embargo, tanto el desarrollo como la implantación de las tecnologías de sanidad móvil presentan varios retos y riesgos. Algunas publicaciones recientes han puesto de manifiesto que la calidad de la mayoría de las aplicaciones reumatológicas no es elevada, lo que conlleva varios riesgos importantes [4]. Esta situación pone de manifiesto la necesidad de una cierta estandarización en su desarrollo, con el fin de garantizar la seguridad del paciente y mejorar la utilidad de estas tecnologías. En este sentido, la Alianza Europea de Asociaciones de Reumatología (EULAR) publicó directrices para el desarrollo, la evaluación y la implementación de futuras soluciones mHealth que ayuden a la autogestión de personas que viven con las ERM [5].

Con la finalidad de proporcionar un seguimiento cercano y una intervención personalizada de los pacientes con ERM se diseñó el Modelo de Atención Mixta (MAM) [6]. El MAM es un modelo de atención aplicable a la práctica clínica que incorpora el uso de las tecnologías de la información y la comunicación (TIC) en el seguimiento de los pacientes. Estas características empoderan al paciente y permiten una monitorización más estrecha que los métodos tradicionales de seguimiento de pacientes, favoreciendo una gestión de la salud más eficaz y eficiente.

En este trabajo se describe la solución clínica digital IMIDoc, adaptada al modelo MAM y a las necesidades de los profesionales sanitarios y pacientes con AR y EspA. La evaluación de la implementación del modelo MAM en la práctica clínica a través de la herramienta IMIDoc se realizará en un estudio clínico multicéntrico en el que participarán 360 pacientes con AR y EspA de cinco hospitales españoles: Hospital Universitario La Paz

(HULP), Hospital Universitario Vall d'Hebron (HUVH), Hospital Universitario Belvitge (HUB), Hospital Universitario Infanta Leonor (HUIL) y Hospital Universitario Ramón y Cajal (HURyC). En cada centro se incluirán, emparejados por edad y sexo, 40 pacientes con AR y 40 con EspA en el grupo experimental; y 20 pacientes con AR y 20 con EspA en el grupo control que serán tratados en la Unidad de Terapia Complejas (UTC) según el método tradicional. Entre los objetivos del estudio, están el determinar factores asociados a la adherencia y al seguimiento mediante el MAM, evaluar la aceptación, usabilidad y satisfacción de la solución digital IMIDoc, la comparación con el seguimiento convencional, etc.

2. Materiales y métodos

Para llevar a cabo el diseño y desarrollo de IMIDoc se ha seguido un enfoque de co-creación en el que han trabajado de forma colaborativa e interdisciplinar equipos clínicos, técnicos, académicos y pacientes. El proceso de co-creación se estructuró en dos fases principales: fase de planificación y fase de desarrollo.

En la fase inicial, de planificación, el equipo clínico revisó meticulosamente sus experiencias previas e identificó las necesidades insatisfechas prevalentes. Para garantizar un diseño centrado en el usuario, la fase inicial consistió en la revisión de entrevistas previamente realizadas a clínicos y pacientes [6]. Los perfiles de usuario se catalogaron principalmente en dos categorías: profesionales sanitarios y pacientes. Para los profesionales sanitarios, se establecieron funcionalidades como la supervisión de los datos del paciente y la respuesta a las alertas generadas por los registros de los pacientes. Por su parte, para los pacientes, se establecieron funcionalidades para el registro de datos, recepción de notificaciones y acceso a contenido informativo. Esta fase culminó con un primer borrador de la especificación de requisitos.

En la fase del desarrollo tomaron participación los equipos técnicos de software e ingeniería biomédica, que se adhirieron a metodologías ágiles. Estas metodologías consisten en el desarrollo iterativo que aplicadas en entornos sanitarios pueden mejorar el desarrollo de software y reforzar la cooperación entre los profesionales clínicos, equipos técnicos, y pacientes. Esto apoyaría la implantación de la salud digital, promoviendo un acceso más equitativo de la asistencia sanitaria [7]. En la planificación del desarrollo de IMIDoc se establecieron iteraciones o “sprints” donde cada uno de los cuales culminaba en una reunión de recapitulación para evaluar los progresos, comprobar las nuevas funcionalidades, y recalibrar acciones si era necesario.

Como punto de partida para diseñar la arquitectura de la plataforma IMIDoc, se ha considerado el protocolo MAM y los agentes que conforman la intervención clínica, que combina el seguimiento presencial y el telemático (Fig. 1).

La plataforma se compone de una aplicación para pacientes, una aplicación para los profesionales sanitarios, y una infraestructura tecnológica que soporta la comunicación, procesamiento y almacenamiento de la información. Los pacientes pueden registrar incidencias,

ePROs relacionados con la enfermedad en la solución digital y, a su vez la plataforma mediante notificaciones informará de las tareas pendientes y disponibilidad de nuevo contenido de interés. Los datos registrados por los pacientes quedan disponibles para el seguimiento clínico a través de una aplicación para los profesionales sanitarios. Los clínicos digitales pueden contactar y resolver las incidencias de forma telemática. Si no es posible resolver la incidencia telemáticamente, el clínico digital informa al clínico presencial y se contacta con el paciente para gestionar una visita presencial.

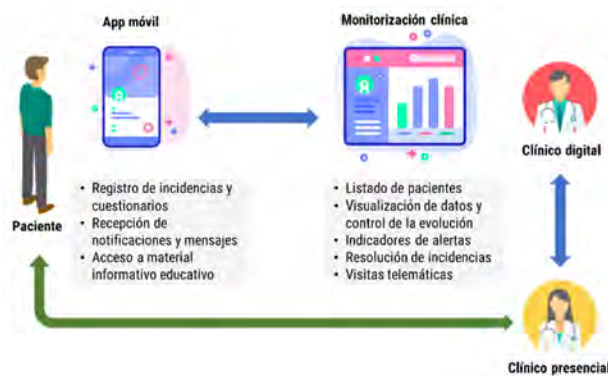


Figura 1. Esquema del protocolo MAM

Durante el proceso de diseño y desarrollo de las aplicaciones y de la infraestructura que soporte las funcionalidades de las mismas, fue necesario evaluar posibles tecnologías, tanto nativas, Web e híbridas, así como también diferentes esquemas para el despliegue de servicios. Para ello, se consideró el perfil de los usuarios, tipos de dispositivos y sistemas operativos, alcance del estudio clínico, ventajas y desventajas asociadas, etc.

Tras alcanzar un alto nivel de desarrollo de IMIDoc se planificó una fase de pruebas o piloto con una duración de 4 semanas en el que participaron un grupo de médicos y pacientes. Antes del iniciar el período de pruebas, se proporcionó a todos los participantes una guía completa en la que se detallaban las funcionalidades de las aplicaciones, las acciones básicas a probar, los métodos de introducción de datos y los pasos para la resolución de problemas. Durante las pruebas se identificaron errores, aspectos mejorables, así como también opiniones y sugerencias mediante formularios y entrevistas informales, con el objetivo de depurar el sistema y alcanzar la pertinencia de las funciones y la experiencia general.

3. Resultados

En la fase de ideación de plataforma se identificaron aspectos y características relevantes que la solución digital debía incorporar. Una de las áreas más valoradas por los pacientes fue la inclusión de información y contenidos relacionados con la enfermedad con el objetivo de apoyar la autogestión de la salud. Por su parte, los facultativos incidieron en la importancia de componentes para la monitorización, así como su fiabilidad y rendimiento. Estos comentarios fueron fundamentales para orientar el proceso de análisis y especificación de requisitos con el objetivo de conseguir un diseño orientado a una mayor

adherencia y usabilidad para los pacientes y una implementación ágil y robusta para los médicos.

El proceso de desarrollo iterativo se llevó a cabo a lo largo de 15 sprints que abarcaron desde enero a agosto de 2023, cada uno de ellos intercalado por una reunión quincenal de actualización y validación. Las reuniones iniciales sentaron las bases centrándose en la arquitectura del servicio y en funcionalidades básicas como el inicio de sesión y el registro. Las reuniones posteriores se orientaron a mejorar la interfaz de usuario, desarrollar flujos de trabajo en la solución digital e integrar diseños de cuestionarios. La fase intermedia hizo hincapié en el perfeccionamiento de las funciones de comunicación, el desarrollo de recursos y medicación, y la gestión de incidencias. También se integraron herramientas de visualización de datos como gráficos históricos y pantallas de evolución. Las reuniones finales se centraron en la depuración del sistema, concretamente en la mensajería y la gestión de errores, antes de realizar una revisión final exhaustiva.

El proceso colaborativo e interdisciplinar permitió definir los instrumentos a implementar para la recogida de los datos de interés, así como diseñar y desarrollar las aplicaciones para los usuarios: app móvil para pacientes con AR o EspA, y aplicación Web para los facultativos.

3.1. Conjunto de instrumentos y variables

De acuerdo con los objetivos del estudio clínico en el que se evaluará la implementación del modelo MAM en la práctica clínica a través de IMIDoc, en la tabla 1 se especifican las variables clínicas y ePROs que se capturarán y que conforman la monitorización y seguimiento de los pacientes con AR y EspA.

Comunes	<p><u>Sociodemográficas y de tratamiento</u> Edad, Sexo, hábitos tóxicos, fecha de diagnóstico, fecha de inicio del tratamiento, medicación, etc.</p> <p><u>ePROs</u> Valoración global enfermedad (VGP), Escala visual dolor (EVA), Número de articulaciones dolorosas (NAD) y tumefactas (NAT)</p> <p><u>Incidencias</u> Brotos, infección, medicación, otros motivos</p> <p><u>Adherencia</u> Cuestionario (Compliance Questionnaire on Rheumatology, CQR)</p>
AR	<p><u>ePROs</u> Índice de la actividad de la enfermedad (DAS28), Cuestionario de evaluación de la salud (HAQ)</p>
EspA	<p><u>ePROs</u> Cuestionarios de la actividad de la enfermedad (Bath Ankylosing Spondylitis Disease Activity, BASDAI; Ankylosing Spondylitis Disease Activity Score, ASDAS), Cuestionario del estado de salud (Assessment of SpondyloArthritis International Society Health Index, ASAS-HI), Recuento de entesis dolorosas</p>

Tabla 1. Variables clínicas y ePROs recogidos por IMIDoc

Para analizar el uso de IMIDoc se han definido e incluido variables que recogerán el número de interacciones con las aplicaciones por parte de los usuarios, tiempos medios de uso de los escenarios y/o apartados, tiempos entre las

acciones clínicas, como el tiempo transcurrido hasta el contactar con los pacientes ante incidencias, etc.

Adicionalmente, se han implementado instrumentos para evaluar la usabilidad de las aplicaciones, satisfacción y experiencia de los usuarios: Escala de Usabilidad del Sistema (SUS) y el indicador Net-Promoter Score (NPS).

3.2. Aplicaciones de IMIDoc

El desarrollo de las aplicaciones se ha basado en el patrón Modelo-Vista-Controlador y se han utilizado tecnologías Web (HTML, CSS, TypeScript y Angular). La app móvil para pacientes es del tipo híbrida y se ha desplegado para dispositivos Android e iOS. Para la implementación de los módulos y componentes del “backend” se ha utilizado un esquema basado en contenedores desplegados en la nube con Amazon Web Services (AWS). El backend incluye una API Web que responde a las peticiones de las aplicaciones, de un sistema de gestión de bases de datos PostgreSQL para el almacenamiento de la información. Por último, se han utilizado otros servicios como Google Analytics para registrar métricas de uso y Firebase para almacenar parte del contenido multimedia de las aplicaciones.

La app desarrollada (Figura 2), para pacientes con AR o EspA, incluye funcionalidades para que puedan registrar la actividad de su enfermedad, así como algunos otros dominios potencialmente afectados en su enfermedad.

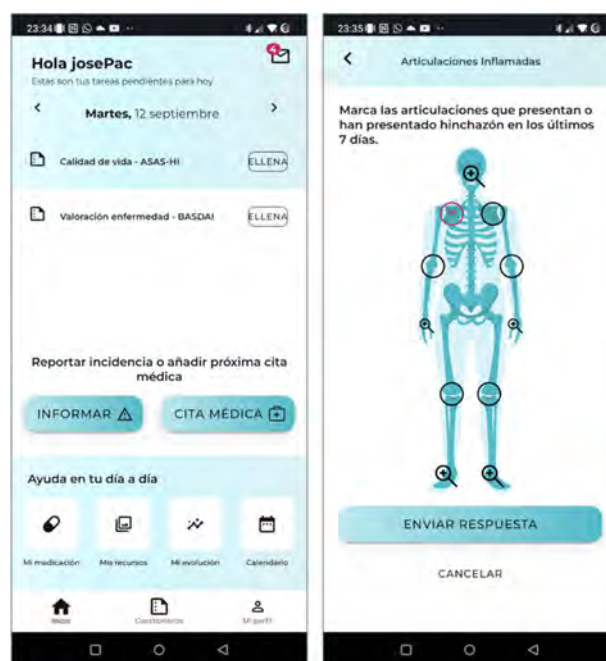


Figura 2. App móvil IMIDoc para pacientes

Se incorporaron mecanismos para la gestión de la medicación como recordatorios de los horarios en un calendario personalizado, que también puede incluir las fechas de consulta. Además, se incluyeron contenidos educativos para guiar a los pacientes en la comprensión de su enfermedad y las opciones de tratamiento.

Por su parte, la aplicación Web destinada a los profesionales sanitarios, principalmente reumatólogos clínicos, incluye los ePROs registrados por los pacientes e indicadores de las alertas o incidencias que permiten a los

clínicos responder a los cambios significativos en la actividad de la enfermedad, resolver incidencias y registrar visitas telemáticas y presenciales. En la Figura 3 se muestra una imagen de la pantalla de visualización de las incidencias registradas por un paciente.

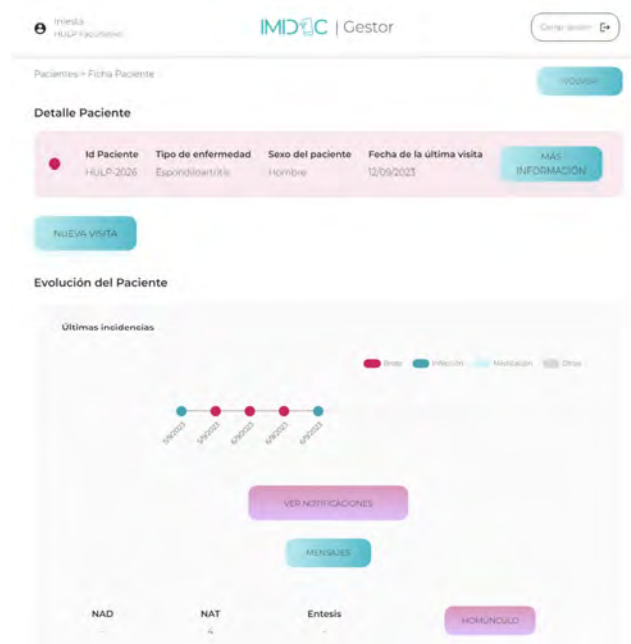


Figura 3. Aplicación Web IMIDoc para facultativos

4. Discusión y conclusiones

Los resultados obtenidos ponen de manifiesto la utilidad de adoptar un enfoque estructurado de co-creación para el desarrollo de una solución clínica digital como IMIDoc, adaptada a la práctica de la reumatología, con el objetivo de alcanzar elevados niveles de usabilidad, adherencia y satisfacción por parte de los usuarios.

Las experiencias previas, evaluadas a través del estudio Digireuma [6], sirvieron de base para la evaluación inicial de las necesidades, y los comentarios de médicos y pacientes fueron decisivos para dar forma a las interfaces y especificar las funcionalidades de IMIDoc.

La utilización de una metodología ágil ha resultado ser un mecanismo especialmente valioso para todos los involucrados, permitiendo canalizar las aportaciones de los participantes de iterativa e incremental, garantizando que la aplicación desarrollada se ajuste estrechamente al dinámico e intrincado panorama de las necesidades médicas. Al dividir el ciclo de vida del desarrollo de software en iteraciones más pequeñas, la metodología ágil, no sólo ha permitido minimizar los riesgos potenciales, sino que también ha facilitado la adaptación a los cambios emergentes. En esencia, se ha fomentado un entorno de colaboración en el que los equipos interdisciplinarios, incluidos médicos, desarrolladores de software e ingenieros biomédicos, han podido trabajar en sinergia para producir una solución adaptada a un escenario clínico real. Este trabajo aporta otra experiencia positiva y eficaz en el entorno clínico reumatológico.

Aunque la fase piloto fue una experiencia positiva, se detectaron errores y capturaron observaciones que sugieren posibilidades futuras de mejora evolutiva. No obstante, la mayoría de las sugerencias y la totalidad de los errores fueron abordados durante la fase de pruebas.

El objetivo principal de IMIDoc es dar respuesta a una necesidad no satisfecha en el cuidado de pacientes con AR y EspA, proporcionando una monitorización estrecha y continuada. Las prácticas clínicas convencionales se caracterizan por consultas médicas discretas y programadas, lo que deja lagunas en la atención al paciente cuando más se necesita una atención inmediata.

A través de un proceso estructurado de co-creación, se ha desarrollado IMIDoc, que se postula como una solución de salud digital innovadora en reumatología, para apoyar la autogestión de la salud y favorecer la mejora de la calidad de vida de pacientes con AR y EspA. El siguiente paso es llevar a cabo el estudio clínico multicéntrico planteado con más de 300 pacientes en la práctica clínica.

Agradecimientos

Los autores agradecen a todos los profesionales sanitarios, investigadores y técnicos que colaboraron en el desarrollo de IMIDoc y a los pacientes que participaron en las pruebas técnicas de validación. Al proyecto PI22/00777 “Hacia la Telemonitorización en Enfermedades Inflammatorias Inmunomediadas: Implementación de un Modelo Asistencial Mixto (MAM) en una Unidad de Terapias Complejas”, financiado por Fondo de Investigaciones Sanitarias (FIS) del Instituto de Salud Carlos III (ISCIII) y cofinanciado con Fondos Estructurales de la UE.

Referencias

- [1] Almutairi K, Nossent J, Preen D, Keen H, Inderjeeth C. The global prevalence of rheumatoid arthritis: a meta-analysis based on a systematic review. *Rheumatol Int.*; 41(5):863-877, 2021. doi: 10.1007/s00296-020-04731-0.
- [2] Woolf AD, Erwin J, March L. The need to address the burden of musculoskeletal conditions. *Best Pract Res Clin Rheumatol.*; 26(2):183–224, 2012. doi: 10.1016/j.berh.2012.03.005
- [3] Shelton J, Casey S, Puhl N, Buckingham J, Yacyshyn E. Electronic patient-reported outcome measures using mobile health technology in rheumatology: A scoping review., 2021. <https://doi.org/10.1371/journal.pone.0253615>
- [4] Luo D., Wang P., Lu F, Elias J, Sparks J, Lee Y. Mobile Apps for Individuals With Rheumatoid Arthritis: A Systematic Review. *JCR: Journal of Clinical Rheumatology* 25(3):p 133-141, 2019. DOI: 10.1097/RHU.0000000000000800
- [5] Najm A, Nikiphorou E, Kostine M, et al. EULAR points to consider for the development, evaluation and implementation of mobile health applications aiding self-management in people living with rheumatic and musculoskeletal diseases. *RMD Open* 2019;5:e001014. doi: 10.1136/rmdopen-2019-001014
- [6] Benavent D., Fernández-Luque L., Núñez-Benjumea F.J., Navarro-Compán V., Sanz-Jardón M., Novella-Navarro M., et al. Monitoring chronic inflammatory musculoskeletal diseases mixing virtual and face-to-face assessments— Results of the digireuma study. *PLOS Digit Health* 1(12): e0000157. <https://doi.org/10.1371/journal.pdig.0000157>
- [7] Kokol, P. Agile Software Development in Healthcare: A Synthetic Scoping Review. *Appl. Sci.* 2022, 12, 9462. <https://doi.org/10.3390/app12199462>

Interoperabilidad de Bluetooth Low Energy con ISO/IEEE X73-PHD: estudio de propuestas basadas en encapsulación de datos

M. López García, D. Naranjo Hernández¹, J. Reina Tosina¹

¹Grupo de Ingeniería Biomédica, Depto. de Teoría de la Señal y Comunicaciones, Universidad de Sevilla, Sevilla, España

Resumen

La familia de estándares ISO/IEEE 11073 (X73) tiene como objetivo garantizar la interoperabilidad entre los dispositivos de salud personal. A pesar de la popularidad de Bluetooth Low Energy (BLE) como tecnología de comunicación en dispositivos biomédicos, no existe una propuesta consensuada que permita la compatibilidad entre ambos estándares. En este trabajo se propone la encapsulación de datos X73 para su envío a través de BLE como forma de facilitar la interoperabilidad. Se presentan dos alternativas que han sido evaluadas en el contexto de la especialización 11073-10420 e implementadas sobre la placa evaluación de un dispositivo BLE comercial. Los resultados muestran la viabilidad de ambas alternativas y se discuten en términos del volumen de información que es necesario transmitir y el consumo que representa para la interfaz radio.

1. Introducción

La demanda de dispositivos de salud personal (*personal health device*, PHD) se encuentra en constante aumento debido al auge de la salud conectada y su relación con los nuevos paradigmas de la medicina [1]. Los sensores biomédicos desempeñan un papel fundamental en la visión actual del cuidado de la salud. Sin embargo, la falta de interoperabilidad entre los PHD continúa siendo una barrera para el éxito de los sistemas de salud conectada [2,3].

Con ese propósito, la Organización Internacional de Normalización (ISO) y el Instituto de Ingenieros Eléctricos y Electrónicos (IEEE) desarrollaron la familia de estándares ISO/IEEE 11073, también referida como X73-PHD. Estos estándares facilitan la comunicación entre el sensor biomédico y el dispositivo gestor de la información, proveyendo un marco para la interoperabilidad a través de una sintaxis común para el correcto flujo de los datos médicos. Adicionalmente, se necesita una tecnología de comunicación entre sensor biomédico y gestor, preferiblemente inalámbrica. Actualmente, la más popular es *Bluetooth Low Energy* (BLE), ampliamente utilizada en entornos IoT [4].

A pesar del fuerte respaldo de BLE por la industria, la compatibilidad con X73 es un problema no resuelto, siendo necesaria la adaptación de sus protocolos. BLE aparece a partir de la versión 4.0 de *Bluetooth*, centrada en la eficiencia energética y, en consecuencia, disminuyendo enormemente el consumo de las versiones anteriores. Esta reducción de consumo aumentó el número de dispositivos que implementan BLE. Sin embargo, la mejora en el consumo implicó a su vez una reducción de los servicios

ofrecidos [5], y esto llevó a que X73 y BLE no sean compatibles, por lo que el fabricante es el encargado de compatibilizar ambos estándares.

En este respecto, hay distintas publicaciones que aportan soluciones parciales [6,7], pero ni son definitivas ni están detalladas.

En este trabajo se propone un adaptador X73 que encapsula los datos a BLE según dos variantes, una encapsulación directa haciendo uso de la APDU del modelo de comunicación y otra usando XML, lenguaje ampliamente utilizado en sistemas de información sanitarios [3]. Como prueba de concepto, se analiza la viabilidad de ambas propuestas para transmitir la medida de un dispositivo de bioimpedancia basado en [8], sobre una placa de evaluación BLE. Para tal fin, se considerarán los siguientes objetivos específicos:

- Estudio de los parámetros de una medida de bioimpedancia para su posterior adaptación a ISO/IEEE 11073 en una placa de evaluación BLE. Se considerará dentro de este objetivo la extensión de la especialización 11073-10420 (*Body composition analyzer*, BCA) para enriquecer el contenido de los datos intercambiados.
- Encapsulación de los datos X73 para su envío a través de BLE proponiendo dos variantes para solucionar la incompatibilidad.
- Diseño de una aplicación móvil para la desencapsulación de los datos recibidos.

2. Materiales y métodos

ISO/IEEE 11073 es una familia de estándares compuesta principalmente por la norma 11073-10101, relativa a la nomenclatura; los estándares 11073-104zz, dedicados a las especializaciones de distintos PHD, y 11073-20601 *Optimal Exchange Protocol*, como base de los estándares. X73 se basa en una comunicación punto a punto entre un agente, dispositivo que recopila los datos, generalmente un sensor biomédico, y un gestor, dispositivo que recibe la información y la gestiona.

X73-PHD cuenta con tres conceptos principales: *domain information model* (DIM), modelo de servicio (SM) y modelo de comunicación (CM). El DIM caracteriza los datos recopilados por el agente en un conjunto de objetos, el SM provee las instrucciones necesarias para el intercambio de información entre agente y gestor, y el CM establece el formato de la comunicación punto a punto

mediante el intercambio de mensajes y una máquina de estados. Además, el CM codifica la información del DIM a notación binaria, generalmente haciendo uso de las Reglas de Codificación Médica (MDER). Para la definición de un sistema compatible con X73 es obligatorio seguir el CM y opcionalmente el DIM y el SM [9]. En este trabajo, se utilizan DIM, SM y CM con el fin de conseguir un desempeño más fiel a los estándares X73-PHD.

Entre las especializaciones de X73 se encuentra 11073-10420 (BCA), cuyo modelo de datos contempla tres tipos de parámetros: de composición corporal, relativos a la bioimpedancia y método de bioimpedancia. Son más de 20 los parámetros contemplados en este trabajo, mientras que la especialización BCA solo cuenta con 10. Los parámetros añadidos amplían algunas variables de interés para el estudio de la composición corporal, detallan la técnica de la bioimpedancia multifrecuencia y añaden características sobre el modelado de la bioimpedancia. Los parámetros de nueva creación deben ser definidos según las reglas del DIM.

En este trabajo se hará uso de un adaptador X73 para codificar los datos del sensor biomédico a la notación adecuada. Siguiendo la nomenclatura estandarizada, como agente X73 se utilizará la placa de evaluación de un dispositivo BLE comercial. Específicamente, el componente seleccionado es el integrado BGM220P, de Silicon Labs, que incorpora Bluetooth 5.2, tiene una potencia de transmisión máxima de 8 dBm y una sensibilidad de -98.9 dBm a 1 Mbit/s. El entorno de desarrollo utilizado es Simplicity Studio, desarrollado por la misma compañía para los dispositivos BLE. Su lenguaje de programación es C/C++ y ofrece diversas herramientas para el análisis y la viabilidad de las conexiones Bluetooth, pudiendo variar la potencia de transmisión o el intervalo de conexión, parámetros determinantes del consumo y la latencia.

Por su parte, como gestor X73 se empleará un teléfono móvil Android, el entorno para el desarrollo de la aplicación es Android Studio y el lenguaje de programación elegido es Kotlin. Para el desarrollo de la aplicación también se ha recurrido al software EFR Connect de Silicon Labs, que implementa un escáner BLE que permite conectar sencillamente el móvil con otros dispositivos BLE que se estén anunciando. Como se describe en la Figura 1, placa de evaluación y móvil establecen una conexión BLE y, posteriormente, intercambian la información a través del perfil de atributo genérico (GATT) de Bluetooth. Este perfil define la forma en la que los dispositivos BLE intercambian la información.



Figura 1. Diseño del sistema.

3. Resultados

3.1. Definición del DIM

Se consideran tres tipos de objetos presentes en este trabajo: MDS (*Medical Device System*), numérico y enumerado. El primero describe el contexto de la medida, el segundo es un objeto de valor numérico, como el peso, y el tercero asigna una codificación a cada valor posible del objeto.

En la Tabla 1 se detallan los diferentes atributos del objeto MDS definidos en este trabajo, entre los que se destacan: *Handle*, identificador del objeto; *System-Type-Spec-List*, especialización; y *Dev-Configuration-Id*, que muestra si la configuración es estándar (no añade parámetros nuevos a la especialización) o extendida (contiene parámetros adicionales).

Atributo MDS	Valor del atributo
Handle	0
System-Type-Spec-List	{MDC_DEV_SPEC_PROFILE_BCA, 2}
System-Model	{“Fabricante”, “Modelo”}
System-Id	EUI-64
Dev-Configuration-Id	Configuración extendida: 0x4000-0x7FFF

Tabla 1. Objeto MDS.

En la Tabla 2 se muestran los atributos de los objetos numéricos y enumerados para el parámetro “peso”. Los corchetes presentes en la tabla indican los valores numéricos de los atributos. *Type* define el identificador del objeto [10,11], *Metric-Spec-Small* las características de la medida (periódica o aperiódica, manual o calculada, etc. [10]) y *Unit-Code* la unidad. A los identificadores de los parámetros no contemplados en la especialización 10420 se les debe asignar un valor de 2 octetos entre 61440 y 65535, pues son los que están reservados para uso privado [10].

3.2. Modelo de comunicación (CM)

Se basa en una serie de mensajes que accionan el comportamiento de la máquina de estados. Los mensajes se identifican mediante su APDU y tienen que recibir respuesta, con el objetivo de ser robusto ante los errores.

Los mensajes más importantes son los que se muestran en la Figura 2. Esos mensajes son de asociación, configuración, envío de los atributos MDS, envío de datos y desasociación. Los mensajes son los contenedores de los objetos, que se establecen a partir de las instrucciones del modelo de servicio (SM) GET, SET, EVENT REPORT y ACTION [12].

3.3. Adaptación a BLE

La adaptación se realiza en C, como recomienda la norma 11073-20601, y siguiendo las reglas MDER y la sintaxis ASN.1. Es necesaria la creación de librerías para cambios de tipos, codificación, encapsulación y nomenclatura X73.

Atributo	Valor del atributo
Handle	1
Type	MDC_PART_SCADA [2] MDC_MASS_BODY_ACTUAL [57664]
Metric-Spec-Small	mss-avail-intermittent, mss-avail-stored-data, mss-upd-aperiodic, mssmsmt-aperiodic, mss-acc-agent-initiated, mss-cat-calculation [0xF0, 0x42]
Unit-Code	MDC_DIM_KILO_G [1731]
Absolute-Time-Stamp	Marca temporal de la medida en formato ss/aa/mm/dd hh:mm:ss.mm. 8 octetos.
Simple-Nu-ObservedValue	Valor observado en formato FLOAT_TYPE. 4 octetos.

Tabla 2. Objeto peso.

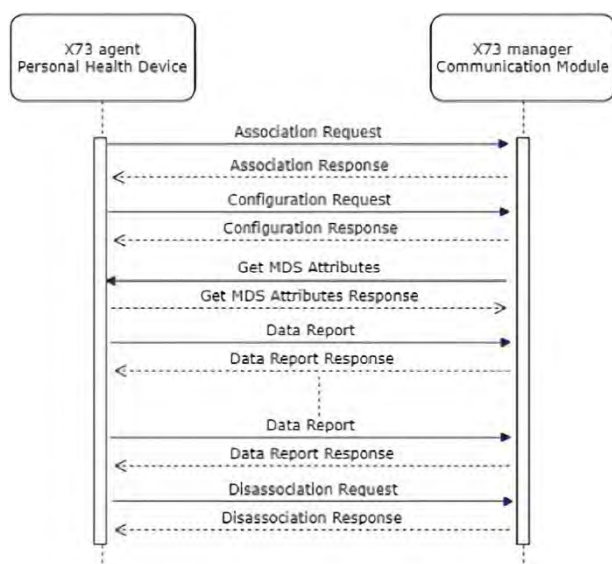


Figura 2. Diagrama de comunicación X73.

La primera de las variantes tratadas en el trabajo supone la generación de los mensajes mediante el uso de las APDU para conformar paquetes para el envío de datos a través del perfil GATT. Para la definición de los mensajes se introduce una lista de configuración que incorpora los parámetros del sensor biomédico, y a través de funciones creadas en este trabajo, se adapta automáticamente a la notación X73. Todo ello permite adaptar dispositivos que no incorporan el CM de X73, pudiendo servir adicionalmente como pasarela X73.

Posteriormente, se necesita la creación y adaptación de las primitivas del SM para transformar la notación X73 en los mensajes del modelo de comunicación. Una vez armado el mensaje, se necesita fragmentarlo en paquetes debido a que el tamaño máximo de los mensajes (MTU) en GATT es 250 octetos. Para rearmar el mensaje se utiliza la APDU para la identificación y la *metadata* del mensaje para conocer la longitud del mensaje completo y de los

fragmentos. El envío de la información se realiza mediante las notificaciones especificadas en el GATT.

La segunda variante consiste en el uso de ficheros XML para realizar la adaptación a BLE, como en [7], con la salvedad de que en este caso se realiza mediante el uso del perfil GATT. En particular, el fichero XML describe la estructura de datos de una característica (contenedor de atributos de GATT). Cada XML corresponde con un mensaje X73 que contiene los parámetros en el formato requerido por la norma 11073-20601. De esta forma, sólo se transmiten los parámetros variables durante la medida y los demás se interpretan según las normas del XML.

3.4. Validación y discusión

Haciendo uso de las librerías BLE de Kotlin, la recepción de la información es sencilla, y la desencapsulación supone el proceso inverso al descrito en el sub-apartado anterior. En la Figura 3 se muestra el mensaje de petición de asociación capturado desde el móvil. En la aplicación móvil se valida que todos los mensajes se reciben y se interpretan correctamente. En la figura se observa que la secuencia hexadecimal es el contenido del mensaje en sí, y el resto es el procesamiento realizado por el gestor para interpretar el mensaje.

También se aprecia en la Figura 3 que la potencia recibida por el móvil es -57 dBm, para una distancia de 0.5 m y 0 dBm de potencia transmitida por la placa de evaluación. Téngase en cuenta que es necesario que la potencia recibida sea superior a -85 dBm para tener una conexión BLE estable y segura.

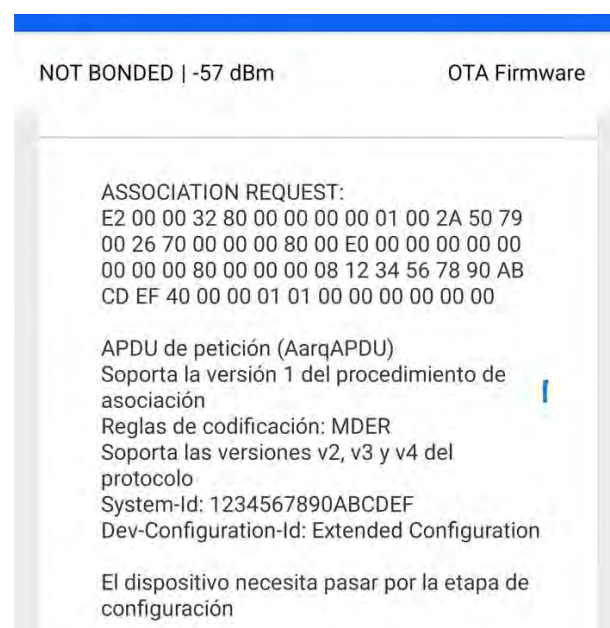


Figura 3. Mensaje de petición de asociación.

Una de las diferencias principales entre ambas variantes consiste en el volumen de la información transmitida. La segunda variante reduce considerablemente la cantidad de octetos a transmitir debido a que sólo se lleva a cabo el envío de la información variable. La primera variante transmite el triple de octetos, lo que supone aumentar drásticamente el tiempo de transmisión y empeorar el

rendimiento en condiciones de bajo nivel de señal frente a los errores.

El consumo energético se ve influenciado principalmente por dos factores, el tiempo de transmisión y la potencia de transmisión. Un aumento significativo del consumo energético es altamente negativo en los PHD alimentados por batería porque supone el acortamiento de su vida útil. Por lo tanto, optimizar los dos parámetros mencionados es necesario para prolongar la autonomía de la batería y, en consecuencia, crear mejores productos. La segunda variante arroja resultados mucho más favorables sobre el tiempo de transmisión. En cuanto a la potencia de transmisión, depende de la configuración de la placa de evaluación. En este caso, para una distancia de 0.5 m, son aceptables las potencias transmitidas superiores a -10 dBm.

Por otro lado, la primera variante es adaptable a cualquier PHD presente en las especializaciones X73 de forma directa y genérica. No es necesaria la generación de nuevo código, pues tras la introducción de los parámetros del sensor biomédico se adapta automáticamente a la norma X73. Adicionalmente, puede servir para adaptar dispositivos no compatibles con el CM de X73. Ninguna de estas funcionalidades es posible en la segunda variante debido a las limitaciones del código XML. Cada fichero XML contiene un mensaje específico para el PHD en uso, aunque se puede tomar como referencia para otros PHD y no entrañar excesivas dificultades añadidas.

4. Conclusiones

La falta de interoperabilidad lastra en gran parte el avance de los sensores biomédicos para su uso masivo. X73-PHD trata de solucionar dicho problema, pero su implementación se enfrenta a varias dificultades. Tras el estudio extendido de los estándares, es coherente suponer que la complejidad de X73 frena su uso para muchos fabricantes, siendo necesaria una renovación de la infraestructura o una actualización de la infraestructura actual.

Otra problemática es la falta de consideración de tecnología de comunicación en los estándares X73-PHD, lo que dificulta una implementación homogénea de los fabricantes y promueve las soluciones propietarias. Sería recomendable la inclusión en los estándares de algunas características de los medios de transporte de información más populares.

En este trabajo se ha presentado un caso de uso validado para la integración de ISO/IEEE 11073 con *Bluetooth Low Energy*, proponiendo distintas soluciones para estandarizar los PHD. Además de garantizar la interoperabilidad de los dispositivos de salud personal mediante X73, se ofrecen dos variantes para ello, una que explora la automatización de estos sistemas y la transmisión de información X73 genérica que adapte sistemas incompatibles, y otra que estudia un lenguaje muy utilizado en sistemas de información sanitarios y lo adapta optimizando parámetros importantes en las conexiones BLE. Es importante incorporar la inclusión de sistemas incompatibles con X73 debido a que la mayoría de los fabricantes optan por soluciones propietarias.

Agradecimientos

Este trabajo ha sido financiado, en parte, por la Junta de Andalucía, a través el proyecto P18-TPJ-3074, y en parte por iGlucó Tech SL a través del contrato FIUS 43671004.

Referencias

- [1] Li J, Caryon P. Health Care 4.0: A vision for smart and connected health care. *IISE Transactions on Healthcare Systems Engineering*, vol 11, 2021 (ISSN 2472-5579).
- [2] Torab A, Samad T, Jodati A, Rezaei P. Interoperability of heterogeneous health information systems: a systematic literature review. *BMC Medical Informatics and Decision Making*, vol 23, 2023 (ISSN 1472-6947).
- [3] Cuevas D, García JP, Bravo, M, López R, Reyna MA, Zermeno NA, González ML. ECG standards and formats for interoperability between mhealth and healthcare information systems: a scoping review. *International Journal of Environmental Research and Public Health*, vol 19, 2022 (ISSN: 1660-4601).
- [4] Darroudi SM, Gomez C, Crowcroft J. Bluetooth Low Energy mesh networks: a standards perspective. *IEEE Communications Magazine*, vol. 58, no. 4, pp. 95-101, 2020 (ISSN: 1558-1896).
- [5] Cristobal A, Torrado A, Rodriguez C, Hernandez JA, Luaces M, Borromeo S. Implementation of ISO/IEEE 11073 PHD SpO2 and ECG device specializations over Bluetooth HDP following health care profile for smart living. *Sensors*, vol 22, 2022 (ISSN: 1424-8220).
- [6] Schmidt M, Obermaisser R. Middleware for the integration of Bluetooth LE devices based on MQTT and ISO/IEEE 11073. *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2017, pp. 1-4.
- [7] Kang HW, Kim CM, Koh SJ. ISO/IEEE 11073-based healthcare services over IoT platform using 6LoWPAN and BLE: architecture and experimentation. *IEEE Int. Conf. on Networking and App*, 2016, pp. 313-318.
- [8] Naranjo D, Reina J, Roa LM, et al. Smart bioimpedance spectroscopy device for body composition estimation. *Sensors*, vol 20(1), 70, 2020 (ISSN: 1424-8220).
- [9] Badawi HF, Laamarti F, El Saddik A. ISO/IEEE 11073 personal health device (X73-PHD) standards compliant systems: a systematic literature review. *IEEE Access*, vol 7, 2019, pp. 3062-3073 (ISSN: 2169-3536).
- [10] López M. Desarrollo de un sistema para la gestión del procesado de información y las comunicaciones en sensores biomédicos. Trabajo Fin de Grado, *Universidad de Sevilla*, 2023.
- [11] Huang ZY, Wang Y, Wang L. ISO/IEEE 11073 treadmill interoperability framework and its test method: Design and implementation. *JMIR Med Inform*, 2020 (ISSN 2291-9694).
- [12] Laamarti F, Badawi HF, Ding Y, Arafsha F, Hafidh B, Saddik A. An ISO/IEEE 11073 standardized digital twin framework for health and well-being in smart cities. *IEEE Access*, vol. 8, 2020, pp. 105950-105961 (ISSN: 2169-3536).

Bioinstrumentación

Miércoles 22 15:00 - 16:30
Salón de Grados

Piezoelectric paint-based force sensor for hand rehabilitation

I. Payo Gutiérrez¹, J. Fernández Almenar¹, D. Rodríguez Rosa¹, S. Juárez Pérez¹, E. López Dolado², A. de los Reyes Guzmán³

¹Escuela de Ingeniería Industrial y Aeroespacial, Universidad de Castilla-La Mancha, Toledo, Spain, {ismael.payo,javier.fernandez,david.rosa,segio.juarez}@uclm.es

²Clinical Chief Rehabilitation Department Hospital Nacional de Paraplégicos, Toledo, Spain, lamidolado@gmail.com

³Biomechanics and Technical Aids Unit Department of Physical Medicine and Rehabilitation, Hospital Nacional de Paraplégicos, Toledo, Spain, adlos@sescam.jccm.es

Abstract

A novel hand rehabilitation device for spinal cord injury patients has been designed with piezoelectric paint-based cantilever beam force sensors. The piezoelectric paint is made of piezoelectric powder homogeneously distributed in a polymer binder (water-based acrylic). The proposed device is constituted by 16 cantilever beam force sensors activated by push-buttons which allows us to measure the applied force with the fingers of the hand helping in the recovery of the patient. The sensors were calibrated, obtaining a linear relationship between the electric signal and the applied force and a sensitivity about 8 V/N. The proposed method has been designed to improve the hand-eye coordination when the patient applies a controlled force with the fingers. Two exercises were proposed for tracking different force profiles and demonstrate the usefulness of the device.

1. Introduction

The use of high technology in rehabilitation processes is a topic very common in the scientific community in the last years. A special attention receives the hand rehabilitation because its important role in the humans' life, that is, the humans use the hands to relate to the environment that round them, pointing, touching, manipulating, etc. The advantages and disadvantages of using advanced technology in hand rehabilitation are exposed in [1], where it is concluded that advanced technology can enrich treatment and help patients who are unable to visit the clinic regularly. However, advanced technology is not superior to traditional treatment and cannot replace the occupational therapist. A complete guide of hand rehabilitation with traditional treatment can be consulted in [2]. Therefore, Advanced technology can help to patients when the use of traditional treatments is not easy to carry out due to different causes, as e.g., mobility problem of the patient to visit the clinic and occupational therapist shortage.

Different methods with advanced technology have been proposed in the literature for hand rehabilitation. The simplest classification that can be done is to divide them in passive and active methods. The words passive and active refer to the willingness of the subject to perform the movement, that is, a method is named passive when a device causes the movement of the hand without the subject's intervention. On the contrary, a method is named active when the subject voluntarily causes the movement of the hand. Depending on the degree of in-

jury of the subject, one of the methods is used or both. Passive methods are used for rehabilitation of paralysed fingers. These methods require, generally, complex robotic devices to move the hand and the fingers, as shown in [3]. Examples of these methods can be consulted in [4,5], where the fingers of the hand are moved by devices driven by controlled motors, and in [6,7], where the fingers are moved by exoskeletons driven by shape memory alloy actuators.

Active methods are used when the subject can move the hand and fingers, albeit with some difficulty. Applications of augmented reality, virtual reality and virtual games have been proposed in the last years to improve the rehabilitation process of the hand, as remarked in [8], versus traditional methods, as the nine hole peg test of finger dexterity [9]. Some particular examples of this type of virtual methods can be consulted in [10–12]. Other active methods use sensed gloves to measure the movement and pressure applied by the fingers when an object is grasped. Some examples can be consulted in [13,14].

Finally, some devices are designed for active and passive methods according to the necessity of the subject. These devices are the most complex of all because they require two modes of operation. Some examples can be consulted in [15,16].

In this study, a novel active method is proposed for hand rehabilitation based on push-buttons which measure the force applied by the fingers. This device can be used by spinal cord injury patients as trainer to improve the hand-eye coordination. So far, no similar experiences have been reported. This work is an attempt to objectify measurement using a physical rehabilitation device.

2. Proposed force sensor

Piezoelectric paint spread on a flexible beam is used to design the proposed force sensor. Typical resistive strain gauges might also be used but the authors are interested in designing novel sensors with different technology but similar performance. Detailed information about piezoelectric paints used as strain sensors can be consulted in [17]. Figure 1 shows a schematic of the proposed sensor. The force applied by the finger on the push-button is transmitted to the free end of a preloaded flexible beam that operates as a spring. The applied force is measured by the piezoelectric paint sensor placed on the

clamped end of the beam. When the force is released, the push-button goes up pushed by the flexible beam. The piezoelectric paint is sandwiched between two electrodes, as shown in Fig. 1. The bottom electrode is the beam because it is made of spring steel (good electrical conductor); and the upper electrode is made of silver paint. The wires are attached to the electrodes by means of an adhesive copper film. The piezoelectric paint is constituted by lead zirconate titanate (PZT) powder mixed with an acrylic lacquer in a ratio of 3:2. Because PZT is a ferroelectric material, the resultant piezoelectric paint must be poled with a high voltage (5 kV/mm) before to use.

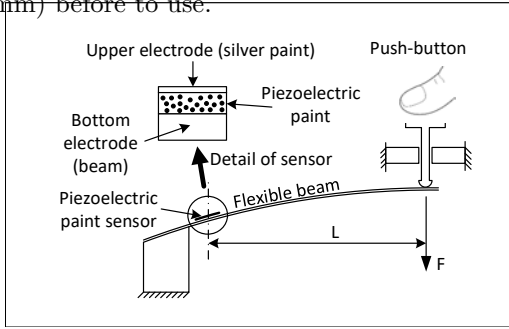


Figure 1. Schematic of the proposed sensor.

The electrical charge generated by the piezoelectric paint can be related to the applied force with a linear relation. On the one hand, the strain ε at the clamped end of the beam (where the piezoelectric paint is placed) and the applied force F can be related by the following equation, according to the classical theory of the deflection of beams [18]

$$\varepsilon = \frac{tL}{2EI}F \quad (1)$$

where t is the thickness of the beam, L is the distance between the piezoelectric paint sensor and the application point of the force, E is the Young's modulus and I is the area moment of inertia of the beam section. On the other hand, the electrical charge generated by the piezoelectric paint can be related to the strain by the following equation, according to [17]

$$q = \frac{d_{13}A E_p}{1 - \nu} \varepsilon \quad (2)$$

where d_{13} is the piezoelectric constant of the piezoelectric paint, A is the area of the electrode, E_p is the Young's modulus of the piezoelectric paint, and ν is the Poisson's ratio. From (1) and (2), the electric charge can be related to the applied force by the following equation

$$q = kF \quad (3)$$

where $k = tLAd_{13}E_p/(2(1 - \nu)EI)$.

3. Experimental validation

3.1. Experimental prototype

Figure 2 shows two photographs of the designed prototype, plant and elevation views. A matrix of 16 (4x4)

sensors, as described in the previous section, is used as rehabilitation device. 16 beams of spring steel (length = 240 mm, width = 12 mm, thickness = 0.5 mm) have been used as flexible beams. The Young's modulus and the dimensions of the beam determine its mechanical stiffness and therefore the required applied force. In this case, the dimensions of the beam have been chosen so that it has low mechanical stiffness and thus the patient has to apply small force when pressing the button. In the final design, the beams are covered by a cover that allows the subject to use the device in a comfortable way, but the authors have preferred to show the inside of the device (Fig. 2) for a better understanding of its operation. The piezoelectric paint sensor was made following the same procedure used in [17]. The electrical charge generated by each sensor was measured by a charge amplifier as shown in Fig. 3. The output voltage of the amplifier v_o is related to the electrical charge generated by the piezoelectric paint q with the following equation

$$v_o = \frac{q}{C} \quad (4)$$

where C is the capacitance of the capacitor. This type of amplifier allows us to read both static and dynamic measurements. A variable resistor (potentiometer) is used to compensate the drift and offset effects in the output voltage. The op-amp model used is TI TL05 and the values of the resistors and the capacitor are the following: $R = R_2 = 1 \text{ M}\Omega$, $R_1 = 10 \text{ k}\Omega$, $R_p = 56 \text{ k}\Omega$, and $C = 1.8 \text{ nF}$. The signal of each sensor was acquired and processed by a data acquisition board NI 6024E and a computer using Simulink Desktop real-time of Matlab software. The data sample frequency was chosen to be 1kHz. Figure 4 shows a schematic of the data acquisition and processing system.

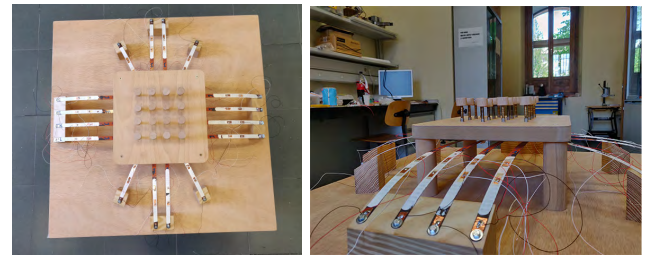


Figure 2. Plan view of the prototype (on the left). Elevation view of the prototype (on the right).

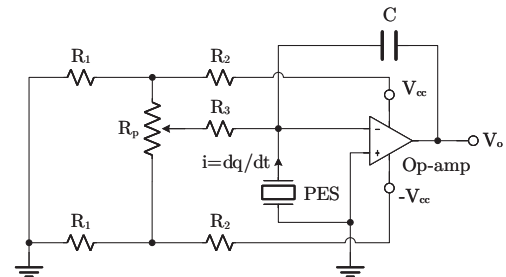


Figure 3. Signal conditioning circuit.

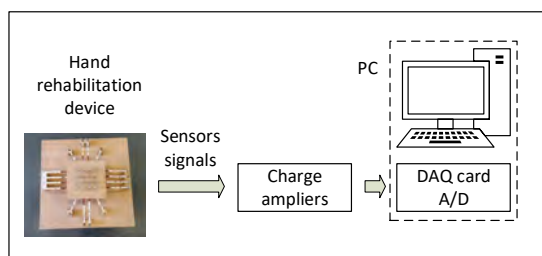


Figure 4. Schematic of the data acquisition and processing system.

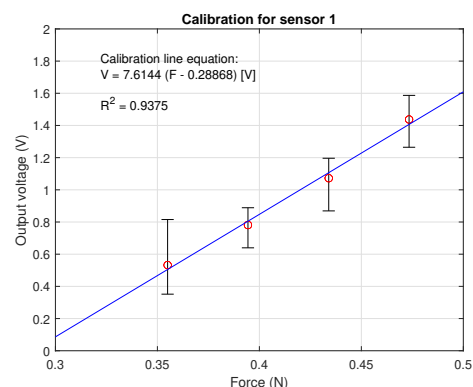


Figure 5. Calibration of sensor 1.

3.2. Experimental results

Before testing the proposed device, every push-button (sensor) must be calibrated to achieve the voltage-force curve. Notice that this relation should be linear according to (3) and (4). Four weights were used to calibrate each sensor: 0.355 N, 0.394 N, 0.434 N and 0.473 N. Five tests were performed with every weight. Figure 5 shows the result obtained with the sensor 1. Maximum, minimum and mean value are plotted for each weight. The regression least-square line is also plotted. The relationship between the electric signal and the applied force is linear and the sensitivity is about 8 V/N. Notice that the device is designed to measure small forces. After calibration tasks, the device is ready to be used.

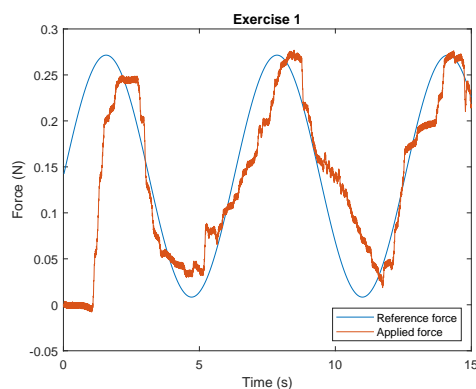


Figure 6. Result of the exercise 1.

Beside to design typical rehabilitation tasks based on repetitive finger movements reading forces applied by the fingers of the hand, two rehabilitation exercises were designed to observe the hand-eye coordination: 1) Tracking an oscillating reference force with a finger. Figure 6 shows the result of the test performed by a healthy person. A reference force (blue signal), type sine function, varying between 0-0.27 N with a frequency of 7 Hz, was chosen for testing. The force applied by the finger is plotted in orange. Notice that the designed application operates in real-time, that is, the user sees in the screen of a computer how the signal changes in real-time when pressing and releasing the push-button trying to track the reference signal; 2) Tracking a reference force, type step function, with the index finger of each hand. Figure 7 shows the result of the test also performed by the same person. In this exercise there are two reference forces, one by each finger. The references change according to step functions. Notice that this exercise is more difficult than the previous exercise because it requires a good coordination between both fingers. It is observed the reaction time of the subject (delay time of the applied force regarding the reference).

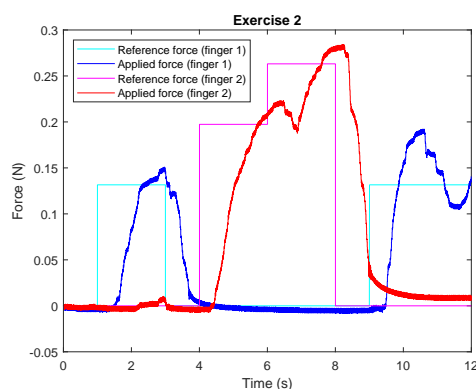


Figure 7. Result of the exercise 2.

4. Conclusions

A new active method for hand rehabilitation has been proposed in this study for recovery of spinal cord injury patients. The proposed device allows us to measure accurately the force applied by the fingers of the hand on push-buttons. Piezoelectric paint-based force sensors are used for this purpose. Beside to measure the force applied by the fingers in traditional exercises, this device allows the patient to train the hand-eye coordination, so important in handling tasks. In addition, the device gives useful information to physicians to evaluate the recovery of the patient.

A first prototype has been designed to demonstrate the feasibility of the idea, but this prototype must be finished and tested with patients who will feedback the pros and cons of the method. The device has been designed to be scaled and adapted to the patient in a second design stage.

As a final conclusion, this work modifies a physical rehabilitation device by including electronics to allow the quantification of variables of interest related to manipulation. Although no patient trials have been performed so far, these will be performed in the near future on patients with cervical spinal cord injury.

Acknowledgment

This study was supported in part by the Junta de Comunidades de Castilla-La Mancha, Spain, under Contract SBPLY/21/180501/000135 and in part by the European Social Fund. This study also forms part of the Advanced Materials programme and was supported by MCIN with funding from European Union NextGenerationEU (PRTR-C17.I1) and by Junta de Comunidades de Castilla-La Mancha.

References

- [1] Yafi Levanon. The advantages and disadvantages of using high technology in hand rehabilitation. *Journal of Hand Therapy*, 26(2):179–183, 2013.
- [2] Susan Weiss and Nancy Falkenstein, editors. *Hand Rehabilitation (Second Edition)*. Mosby, Saint Louis, 2005.
- [3] Zan Yue, Xue Zhang, and Jing Wang. Hand rehabilitation robotics on poststroke motor recovery. *Behavioural Neurology*, 2017:1–20, 11 2017.
- [4] Stefan Hesse, H Kuhlmann, J Wilk, C Tomelleri, and Stephen Kirker. A new electromechanical trainer for sensorimotor rehabilitation of paralysed fingers: A case series in chronic and acute stroke patients. *Journal of neuroengineering and rehabilitation*, 5:21, 10 2008.
- [5] Husam Abdulkareem and Géza Husi. Design and development of continuous passive motion (cpm) for fingers and wrist grounded-exoskeleton rehabilitation system. *Applied Sciences*, 11:815, 01 2021.
- [6] C. Yeow, A. Baisch, S. Talbot, and C. Walsh. Cable-driven finger exercise device with extension return springs for recreating standard therapy exercises. *Journal of Medical Devices-Transactions of The Asme*, 8:014502, 2014.
- [7] Abdul Ab Rahim, Andreas Lachmann, Cheng Low, and Adam Mohd Amin. Sma actuated finger exoskeleton device for rehabilitation of acute paresis patient. *Applied Mechanics and Materials*, 773-774:883–887, 07 2015.
- [8] Margarida F. Pereira, Cosima Prahm, Jonas Kolbensschlag, Eva Oliveira, and Nuno F. Rodrigues. Application of ar and vr in hand rehabilitation: A systematic review. *Journal of Biomedical Informatics*, 111:103584, 2020.
- [9] Virgil Mathiowetz, Karen Weber, Nancy Kashman, and Gloria Volland. Adult norms for the nine hole peg test of finger dexterity. *The Occupational Therapy Journal of Research*, 5(1):24–38, 1985.
- [10] Jia Liu, Jianhui Mei, Xiaorui Zhang, Xiong Lu, and Jing Huang. Augmented reality-based training system for hand rehabilitation. *Multimedia Tools and Applications*, 76:1–21, 07 2017.
- [11] Fernando Carneiro, Rafael Tavares, Jose Rodrigues, Paulo Abreu, and Maria Restivo. A gamified approach for hand rehabilitation device. *International Journal of Online Engineering (iJOE)*, 14:179, 01 2018.
- [12] Jakub Szczechowicz and Marek Pieniażek. Rehabilitation of patients after traumatic injuries to the hand and during the course of upper limb enthesopathic syndromes using personal electronic devices a preliminary report. *Rehabilitacja Medyczna*, 22:49–53, 06 2018.
- [13] Diana Cotoros, Mihaela Baritz, and Anca Stanciu. Influence of carpal tunnel syndrome induced by vibratory actions upon fingers mobility. In *Acoustics Vibration of Mechanical Structures II*, volume 801 of *Applied Mechanics and Materials*, pages 284–289. Trans Tech Publications Ltd, 11 2015.
- [14] Rafael Tavares, Paulo Abreu, and Manuel Quintas. Data acquisition glove for hand movement impairment rehabilitation. *International Journal of Online Engineering (iJOE)*, 12:52, 04 2016.
- [15] B. Birch, E. Haslam, I. Heerah, N. Dechev, and E. J. Park. Design of a continuous passive and active motion device for hand rehabilitation. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4306–4309, 2008.
- [16] Daoming Wang, Yakun Wang, Bin Zi, Zixiang Cao, and Huafeng Ding. Development of an active and passive finger rehabilitation robot using pneumatic muscle and magnetorheological damper. *Mechanism and Machine Theory*, 147:103762, 2020.
- [17] Ismael Payo and Jack Michael Hale. Dynamic characterization of piezoelectric paint sensors under biaxial strain. *Sensors and Actuators A: Physical*, 163(1):150–158, 2010.
- [18] J. M. Gere and S. Timoshenko. *Mechanics of Materials*. PWS Publishing Company, second edition edition, 1997.

Optimización de un sistema de seguimiento y monitorización de órganos trasplantados durante la fase de transporte

S. Pérez Jiménez¹, D. Moreno Blanco¹, F. J. Rubio², E. J. Gómez^{1,3}, M. Quintana², P. Sánchez González^{1,4}

¹ Grupo de Bioingeniería y Telemedicina, ETSIT Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, España, {sandra.perezj}@alumnos.upm.es, {diego.morenob, p.sanchez, enriquejavier.gomez}@upm.es

² Centro Avanzado de Simulación y Entrenamiento Clínico del Instituto de Investigación del Hospital Universitario de La Paz, 28046 Madrid, España

³ Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Madrid, España

Resumen

La medicina moderna ha reconocido el trasplante de órganos y tejidos como uno de los avances más destacados y prometedores. La evolución del procedimiento ha logrado aumentar los índices de supervivencia y la esperanza de vida. Sin embargo, la demanda de tejidos y órganos a trasplantar supera ampliamente la oferta disponible. Debido a esta creciente necesidad, en los últimos años se han reconocido los problemas inherentes al proceso de trasplante. Entre ellos, se destacan los daños sufridos durante la fase de preservación y la falta de seguimiento del órgano durante la fase de transporte. Estas negligencias se asocian a pérdidas de órganos plenamente funcionales y un impacto en la disponibilidad de órganos para trasplante. Por esta razón, en este trabajo se proponen diversas soluciones con el objetivo de mejorar el proceso de transporte y preservación. Con tal fin, se desarrolla un sistema capaz de adquirir las variables de estado del órgano y establecer comunicación con el personal clínico asociado al trasplante. El sistema consta de una nevera, que mantiene el órgano en frío, que incorpora un dispositivo que evalúa la temperatura interna del órgano, lo abastece de líquido de preservación, detecta su geolocalización y verifica un correcto sellado. El software del dispositivo se ha diseñado para comunicarse con un servidor web creado para almacenar y visualizar la información recabada. Se ha diseñado una base de datos y una plataforma web de acceso restringido para permitir un conocimiento completo del estado del órgano de manera universal por parte del personal clínico.

1. Introducción

El concepto de trasplante engloba la sustitución de un tejido dañado por uno funcional, a fin de mejorar la calidad de vida y recuperar el estado de salud. Se trata de un procedimiento sumamente complejo y que abarca distintas etapas e intermediarios [1], [2]. En la actualidad, los avances de las técnicas quirúrgicas, la aplicación de la medicina inmunosupresora y el cuidado en la compatibilidad entre donante y receptor ha posicionado al trasplante como una técnica exitosa. En el 2021, se trasplantaron 144302 órganos, según la *Global Observatory on Donation and Transplantation* [3].

Sin embargo, la disponibilidad de órganos se encuentra limitada por la alta demanda y un ritmo de crecimiento desorbitado. Por ejemplo, en 2022 en EE. UU., las cifras de pacientes en lista de espera para un trasplante de riñón

fueron de 88901 personas frente a los 25499 trasplantes realizados. En función a estas cifras se asocian 17 muertes diarias evitables [4].

Para mejorar los resultados de trasplante, se deben enfocar en tres áreas. Primero, una correcta coordinación del proceso, logrado a través de organismos como la ONT en España, que ha posicionado España como el país con mayor número de donantes por millón de habitantes, siendo el modelo de gestión un referente a nivel mundial [5], [6]. Segundo, una monitorización constante del estado del órgano, que permita esclarecer la toma de decisiones e informar sobre situaciones que puedan poner en riesgo el tejido. Por último, dado que los requerimientos fisiológicos del órgano tras la extracción son muy concretos y susceptibles, resulta imperativo un sistema de preservación que lo mantenga protegido y nutrido.

La técnica más utilizada es el almacenamiento estático en frío (AFE), que mantiene el órgano a una temperatura entre los 0 a los 4°C [7]. Sin embargo, la técnica tradicional demuestra fallas que podrían solucionarse introduciendo un sistema de perfusión, lo que se conoce como preservación dinámica [8].

En este trabajo se presenta un sistema de monitorización del estado del órgano durante el proceso de transporte. El sistema permite la comunicación entre los hospitales asociados a una ruta de trasplante, mediante el envío de la información adquirida sobre el órgano. Además, se trata de un diseño con preservación tradicional y dinámica, al integrar un sistema de perfusión a la nevera de trasplante. Este trabajo se ha desarrollado a partir de un proyecto previo que incluía las funciones de monitorización de la temperatura del órgano, la posición y el tiempo de preservación.

2. Diseño

El diseño del sistema optimizado se ha realizado de forma multidisciplinar entre el Grupo de Bioingeniería y Telemedicina de la Universidad Politécnica de Madrid y con el Centro Avanzado de Simulación y Entrenamiento Clínico del Instituto de Investigación del Hospital Universitario de La Paz en Madrid.

El diseño de la arquitectura del sistema, mostrado en la Figura 2, muestra como se compone de cuatro elementos diferentes: una aplicación móvil, el dispositivo de monitorización, el servidor bróker de protocolo MQTT y un servidor dedicado que sustenta una base de datos y un portal web de acceso restringido.

2.1. El dispositivo de monitorización

Para la monitorización del estado del órgano, se han enumerado como más importantes las variables de la temperatura, la posición geográfica y el tiempo de preservación. De esta forma, el dispositivo de monitorización se ha centrado en un microcontrolador Arduino MKR Wifi 1010, un modelo con modulo Wifi y Bluetooth (BLE) incluidos. A esta placa Arduino, se han añadido dos sensores de temperatura (sumergibles en líquido), un modulo GPS, un sensor resistivo de presión y un zumbador.

A esta estructura, se le ha incorporado un sistema de perfusión para la preservación del órgano. El control de activación de la bomba se ha realizado a través de un relé conectado al Arduino. El circuito esquematizado y completo del sistema se muestra en la Figura 1.

2.2. Comunicación para el envío de variables con el servidor web dedicado

Se ha dispuesto un servidor web para el almacenamiento de las métricas realizadas por los sensores de cada nevera, correspondientes a un trasplante concreto. Para el envío de datos al servidor, se ha utilizado el protocolo MQTT, que consiste en un modelo publicación/suscripción a través de un bróker de mensajería [9]. Para el envío de datos por parte del Arduino, se necesita dotar al controlador de una red Wifi, por lo que se ha desarrollado una aplicación móvil que utiliza la tecnología Bluetooth para configurar esta conexión [10].

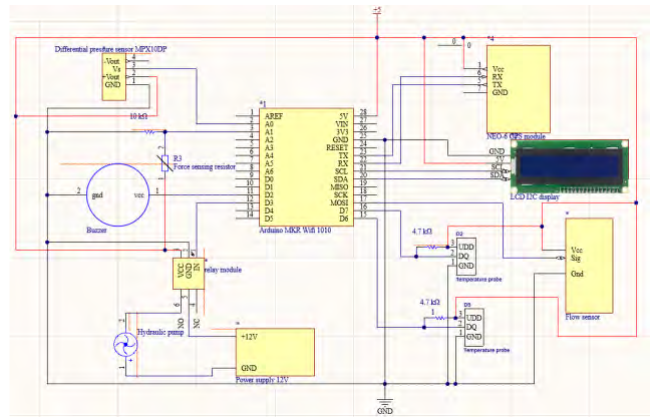


Figura 1. Esquemático del circuito completo de adquisición de variables de estado del órgano y control de la bomba de flujo.

3. Almacenamiento y visualización de la información

Se ha diseñado un servidor web que permite el almacenamiento y la visualización de los datos recopilados. La principal función de la plataforma es la gestión de las rutas de trasplante, que ha sido limitada su acceso al personal clínico registrado. La plataforma permite modificar un traslado de un órgano identificando: la nevera de trasplante, el hospital donante, el hospital trasplantador y la fecha. De tal manera, en paralelo, se almacenan y configuran las neveras disponibles y los hospitales asociados. La información que recaba el dispositivo de monitorización va asociado a los identificadores de las neveras.

Se diferencian para la configuración de la plataforma dos roles de usuario: el usuario administrativo y el usuario clínico. El primero puede acceder a la gestión de usuarios y hospitales asociados a la plataforma, mientras que el segundo configura los trasplantes a realizar y las neveras disponibles en el sistema (Figura 3)

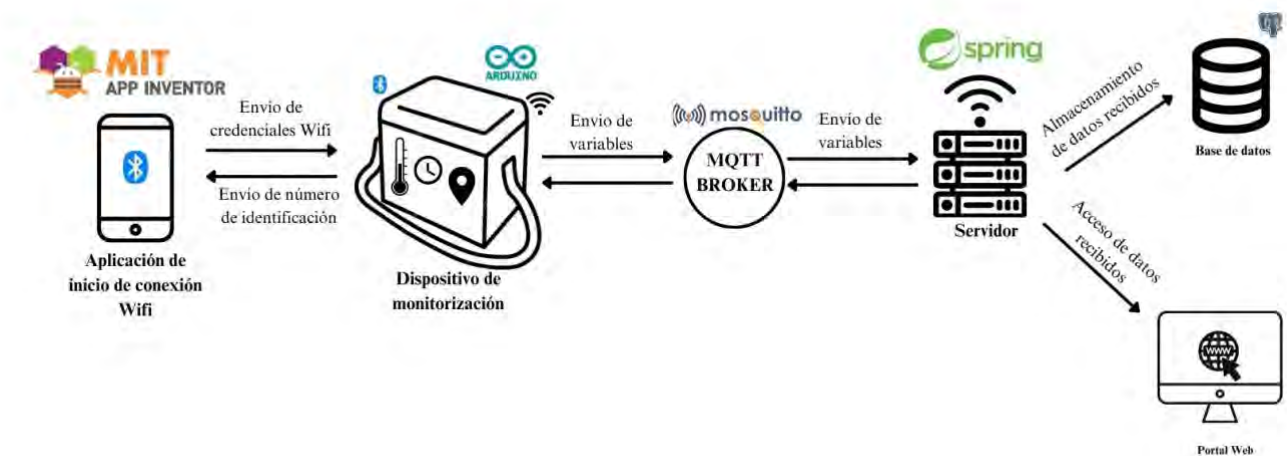


Figura 2. Diagrama de arquitectura del simulador.

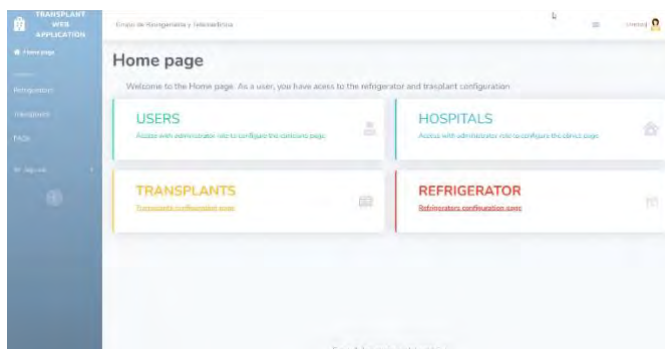


Figura 3. Pantalla de inicio de la página web para clínicos.

Desde el rol de clínico, se accede a la visualización en tiempo real de la información de trasplante. Para ello, existe en “Transplants” la opción “View para ver los datos de cada ruta de trasplante. La información también se almacena en la tabla “Information”, enviada como un objeto JSON a través de un mensaje MQTT desde el Arduino

3.1. Validación técnica y clínica

Por un lado, para la comprobación del funcionamiento correcto y esperado del sistema se han diseñados dos pruebas de validación técnica, una para la evaluación de los métodos de preservación estática y dinámica en situación hipotérmica y otra para la evaluación de la comunicación del servidor y la nevera durante un simulacro de trasplante. Se pretende evaluar las funcionalidades asociadas con el envío correcto y rápido de los datos, el almacenamiento completo de la información y el funcionamiento continuo y controlado del sistema de perfusión

Por otro lado, con el objetivo de evaluar la experiencia del usuario durante la utilización del sistema, se ha realizado una validación clínica con el personal médico adjunto del servicio de cirugía y trasplante de IdiPaz a través de la escala de evaluación *System Usability Scale* (SUS) [11]. Se ha recogido la experiencia tanto con el dispositivo de monitorización como con el portal web de gestión de trasplante a través de un cuestionario, que ha evaluado aspectos como la frecuencia de uso esperada, la facilidad de uso, la autonomía, la consistencia y la seguridad del sistema.

4. Resultados y discusión

El resultado final del sistema formado por la nevera de trasplante y el circuito de adquisición de valores y de perfusión se muestra en la Figura 4.

4.1. Validación técnica

La validación técnica se ha llevado a cabo con un segmento de intestino de cerdo de 40 cm para realizar una evaluación del sistema de perfusión. Se han medido las temperaturas del interior del órgano y del líquido de preservación en dos casos distintos: con el sistema de perfusión desactivado (simulando una preservación tradicional estática en frío) y con el sistema de perfusión activado.



Figura 4. A la izquierda, la nevera con el circuito integrado en una caja de protección. A la derecha, la circuitería del sistema de monitorización y perfusión.

Para ambos escenarios, se ha mantenido el órgano durante 4 horas. Durante este tiempo, se ha evaluado la respuesta de la plataforma web, observando los valores mostrados cada diez segundos sobre la temperatura, la ratio de flujo y la presión de líquido. Una vez finalizado en tiempo, se han recogido las temperaturas de la base de datos y se han mostrado con respecto al tiempo en minutos. De esta manera, se logró observar la variación de temperatura de cada método tal como se muestra en las Figura 5 y Figura 6.

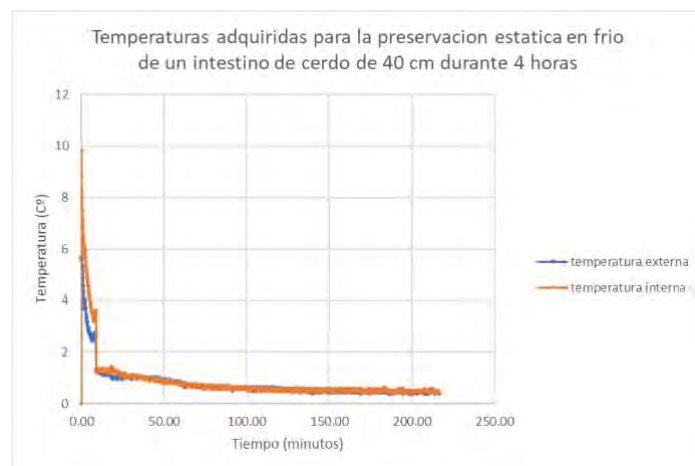


Figura 5. Representación de la evolución de la temperatura durante el proceso de preservación estática.



Figura 6. Representación de la evolución de la temperatura durante el proceso de preservación dinámica

Según los resultados obtenidos para los experimentos de monitorización de la temperatura, se ha destacado una

diferencia apreciable en el uso del sistema de perfusión y las medidas de la temperatura interna del órgano.

Por un lado, en la preservación estática el intestino ha alcanzado los 0.44°C de temperatura interna mínima, mientras que en la preservación dinámica el valor mínimo es de 1.37°C. Estos datos indican una optimización del sistema tradicional a través de la utilización de un sistema de perfusión. El órgano se ha mantenido más alejado a los 0°C, asociados a daños tisulares por congelación del tejido.

Por otro lado, cuando se ha complementado con la preservación dinámica, la información del sensor interno de temperatura, cuya media detectada es de 1.63°C, es distante a la temperatura externa, de media 0.84°C.

En cuanto a la segunda prueba, se ha realizado un transporte de la nevera desde el Hospital Universitario de La Paz hasta el Hospital Universitario 12 de Octubre. Con esta prueba técnica, se ha corroborado un seguimiento en tiempo real de una ruta de trasplante. Se ha detectado que, durante un tramo de túnel, al perder la señal de red móvil que alimentaba la comunicación *Wifi* del Arduino, los datos no se enviaron. Sin embargo, la recuperación de la señal promovía un nuevo envío correcto de la información sin necesidad de reiniciar el microcontrolador.

4.2. Validación clínica

Los resultados de la prueba SUS dotaron al sistema de una puntuación del 82.8, situándolo por encima de la media (68 puntos sobre 100) y en un rango de uso aceptable. Esta puntuación corresponde a un sistema afable con el usuario, de fácil utilización, rápida adaptación a la interfaz y con una integración de las funciones adecuada que lo posiciona como una opción segura y de alto rendimiento.

5. Conclusiones

En vistas a los resultados obtenidos en las validaciones, se puede concluir que se trata de un sistema innovador con mejoras frente al método actual de transporte de órganos.

Por un lado, el sistema de perfusión ofrece mejoras para el mantenimiento de un órgano a trasplantar, dado que se han obtenido unas temperaturas más seguras y alejadas del rango de lesión tisular. Sin embargo, el sistema actual está limitado a un caudal muy pequeño para órganos humanos de mayor tamaño que la muestra.

Por otro lado, el sistema de monitorización es un componente fundamental que logra obtener las variables más relevantes sobre el estado del órgano. Asimismo, el almacenamiento de la información en una base de datos permite la evaluación de las condiciones de transporte y un apoyo en la toma de decisiones, almacenando los momentos exactos de bajadas de temperaturas extremas y de aperturas de la nevera indebidas.

En líneas generales, los resultados de las validaciones han sido óptimos y se han enumerado como líneas futuras la mejora del diseño del sistema de perfusión, la integración con sistemas de trazabilidad hospitalaria, una alternativa ante limitaciones de red, la incorporación de un sistema de

detección de choques y movimientos bruscos y un refuerzo en la protección del circuito.

Referencias

- [1] “Organ Donation and Transplantation: How it works.” <https://my.clevelandclinic.org/health/articles/11750-organ-donation-and-transplantation> (accessed Sep. 07, 2023).
- [2] J.M. Dueñas Jurado, “Protocolos clínicos de actuación ante el proceso de donación y extracción de órganos y tejidos para trasplante en donación en muerte encefálica,” *Cuad. med. forense*, vol. 21, Jun. 2015.
- [3] “count one - GODT.” <https://www.transplant-observatory.org/contador1/> (accessed Sep. 07, 2023).
- [4] “Estadísticas sobre la donación de órganos | donaciondeorganos.gov.” <https://donaciondeorganos.gov/conocer/2n8u/estadisticas-sobre-la-donacion-de-%C3%B3rganos> (accessed Sep. 07, 2023).
- [5] “Donantes de órganos por millón de habitantes.” <https://www.epdata.es/donates-millon-habitante/dbd1d35e-3d7d-493d-a379-0c4b2d763db2> (accessed Sep. 08, 2023).
- [6] “Balance de actividad de donación y trasplante de la ONT 2022 – Organización Nacional de Trasplantes.” <https://www.ont.es/2023/02/10/balanceont20230117/> (accessed Sep. 08, 2023).
- [7] L. Jing, L. Yao, M. Zhao, L. P. Peng, and M. Liu, “Organ preservation: from the past to the future,” *Acta Pharmacologica Sinica* 2018 39:5, vol. 39, no. 5, pp. 845–857, Mar. 2018, doi: 10.1038/aps.2017.182.
- [8] M. H. J. Maathuis, H. G. D. Leuvenink, and R. J. Ploeg, “Perspectives in organ preservation,” *Transplantation*, vol. 83, no. 10, pp. 1289–1298, May 2007, doi: 10.1097/01.TP.0000265586.66475.CC.
- [9] HiveMQ, “MQTT & MQTT 5 Essentials A comprehensive overview of MQTT facts and features for beginners and experts alike eBook.”
- [10] “Arduino Client for MQTT.” <https://pubsubclient.knolleary.net/api> (accessed Jun. 14, 2023).
- [11] A. S. for P. Affairs, “System Usability Scale (SUS),” Sep. 2013.

Development of a stand-alone device to simultaneously record cardiocography biosignals in an open format for testing methods of assessing fetal heart rate variability

LM. López García¹, LF. Memba Massoko^{2,3}, M. Martínez Sober³

C. Sanz Martí⁴, I. Moscardó Chafer¹, A. Antolí Francés¹

¹ Department of Obstetrics and Gynecology, Hospital Lluís Alcanyís, València, Spain, lopez_luigar@gva.es

²FISABIO-Health Department Xàtiva-Ontinyent, València, Spain, dsxo_xativaperinatal@gva.es

³ Electronics Engineering Department, ETSE- Universitat de València, Spain, marcelino.martinez@uv.es

⁴ Department of Obstetrics and Gynecology, Hospital Verge dels Liris, Alacant, Spain, csanzma@msn.com

Abstract

Fetal hypoxia/acidemia recognition improves with computerized analysis of biosignals collected from cardiocography (CTG), particularly the assesment of short-term variability (STV) of fetal heart rate (FHR). Several methods to compute STV have been described with varied performance results according to acquisition method, sampling and storage rates and algorithm definition. Dawes-Redman algorithm (STV16) is the most widely reproduced in available commercial systems. However, it shows a low correlation with the beat-to-beat variation determined from fetal electrocardiographic signal (fECG). STV240 algorithm has been introduced in an attempt to approximate STV assesment to real beat-to-beat variation. There is no comparison in the literature between these two algorithms, taking as gold standard variability obtained from ECG tracing. With a view to providing reliable records for the standardization and comparison of STV algorithms, most notably, STV16 and STV240, we have designed, assembled and developed a stand-alone device well able to connect with different CTG machines and collect simultaneously biomedical signals of interest, particularly FHR, uterine activity and fECG, from the standard monitor outputs. It generates a file in an open format that allows assesment of computerized parameters of CTG. By means of R-R instantaneous variation from fECG as a reference we have found no agreement by Intraclass Correlation Coefficient between STV16 and STV240, neither with STV calculated from fECG. Nevertheless, the last two correlated closely. Standardisation of algorithms, interoperability and research in computerized CTG need to be provided with simultaneous recordings of biosignals involved, including the ECG raw signal. STV16 and STV240 require individualised normal ranges.

1. Introduction

Fetal monitoring by cardiocography (CTG) is the most widely used technique to identify fetal hypoxia/academia, its diagnostic performance improving with computerized analysis of biosignals collected, particularly with the assesment of short-term variability (STV) of fetal heart rate (FHR). Methods to compute STV have different performance results according to algorithm definition, sampling and storage rates and acquisition method: assesment from an external transducer that catches the

mechanical activity of fetal heart by a Doppler ultrasound system or internal monitoring by means an electrode attached to the fetus that captures fetal electrocardiographic –ECG- signal.

In time domain, STV16 of Dawes-Redman (the most widely reproduced) and STV240, described by Kouskouti et al., are algorithms introduced in commercial systems. In order to calculate STV they average every minute values of FHR acquired from external monitoring in time intervals of 3.75 sec (STV16) or 250 msec (STV240) [1]. It must be noted that the comparison of algorithms needs simultaneous recordings of biosignals, including direct fetal ECG, if we are to set the actual value of variability [2].

2. Objectives

- To design, assemble and develop an autonomous electronic device to allow for the collection of biosignals from different cardiocograph models. The device collects simultaneously biomedical signals of interest (particularly fetal heart rate, uterine activity and fetal ECG signal) from the standard outputs of the monitor and generates a file in an open format that allows the calculation of computerized parameters of CTG monitoring.
- To establish the accuracy of STV16 and STV240 algorithms assesed from external Doppler ultrasound CTG, taking as gold standard the beat-to-beat variability obtained from the fetal ECG signal.

3. Material and methods

3.1. Implementation of the digitizing device (DD)

3.1.1 System Functional Description

We have developed an energy and operationally autonomous device –it not require a connection to a computer for its operation- able to acquire biosignals of interest by the standard outputs of different CTG machines. Corometrics 250™ CTG monitor was the

equipment employed for the present project. In figure 1 we represent the functional block diagram of the device.

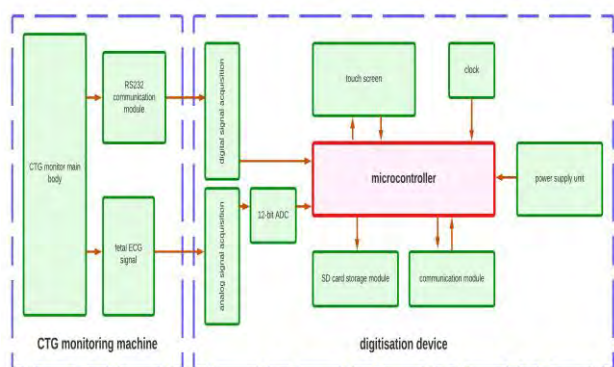


Figure 1. Functional block diagram

Input:

-*Digital signal:* Corometrics 250™ (fetal monitor –FM-, in this case as Data Terminal Equipment –DTE-) has three RS232 Serial Communication Protocol ports. In response to a request command coming from the host digitisation device (as Data Communications Equipment –DCE-) the FM starts the sending of data blocks every 990 to 1100 milliseconds, each containing four values for each channel of FHR (called HR1 and HR2, which enable monitoring of twin gestations) and four values of the abdominal wall pressure collected by the external pressure transducer that assesses uterine activity (UA), as well as other maternal parameters. FM estimates FHR from the Doppler signal obtained by the external US transducer or from the fetal ECG signal when internal monitoring is performed. In either case, FHR figures are sampled at 4 Hz and have a resolution of 1 bpm. AU is sampled at 4 Hz and has a resolution of 1 mmHg. A converter module RS-232/TTL enables the microcontroller to receive the data blocks.

-*Analog signal:* Corometrics 250™ has an analog output for maternal ECG and fetal ECG signals. Fetal ECG signal has a bandwidth of 100 Hz with an amplitude of 10 V/mV. A sample rate of 1 kHz was chosen, which meets the requirement of $2 \cdot F_{max} \leq F_s$. The ADS1015 12 bits allows for a ADC converter resolution of 0.8 mV.

Interactions: three consecutive menus operated on the touch screen provide access to the device's functions: the first menu allows us to start the telematic updating process of the time and software, the next screen requesting identification of the pregnant woman and enabling access to the menu of specific functionalities: starting a new case, starting a new file, sending the request command.

Data handling and storage: raw data from data blocks are extracted, decoded and stored. Each monitoring generates a folder on the memory card containing four files:

- cabecera_FicheroAnalogico.dat and cabecera_FicheroDigital.dat contains metadata about date and time of registration, resolution, and sampling rate (Hz), among others.

- ficheroAnalogico.dat contains the voltage value of the fetal ECG signal digitalized at a sampling rate of 1 kHz.
- ficheroDigital.dat contains the information of the signal channels provided by FM. FHR1 (HR1) is the calculation source of heart rate from external transducer or internal monitoring; FHR2 (HR2) comes from a second external US transducer; Uterine activity (UA) is obtained through the pressure transducer applied on the maternal abdomen; Maternal NIBP is a non-invasive blood pressure measurement; Maternal SpO2 represents maternal oxygen saturation level and MHR is maternal heart rate.



Figure 2. Digitizing device. Main screen

Output

Figure 2 shows the device operating in simulation mode.

Display of digital data on screen: data extracted from blocks are displayed on screen: HR1, HR2, MHR, UA and SpO2. Additional information is also shown: identification number of the case, time and battery charge level. An additional button allows access to the menu to start a new case, a new file or send the request command.

Display on screen of analog signal of fetal ECG: it is displayed in refreshing or still mode. Complementary buttons allow us to select the scale and the number of samples displayed.

3.1.2 System Technical Specification

In figure 3 we represent the wiring diagram.

Microcontroller: the microcontroller used is the ESP32, chosen by its ultra-low power consumption, integrated Wi-Fi and its capability to receive updates by FOTA (Firmware Over-The-Air).

Touchscreen: ILI9341 LCD TFT is a 3.2-inch screen that integrates a resistive touch sensor, controlled by the XPT2046 chip that increases its accuracy.

Powering: an 18650 battery and battery charging module provides the device with an autonomy of between 6-10 hours of use.

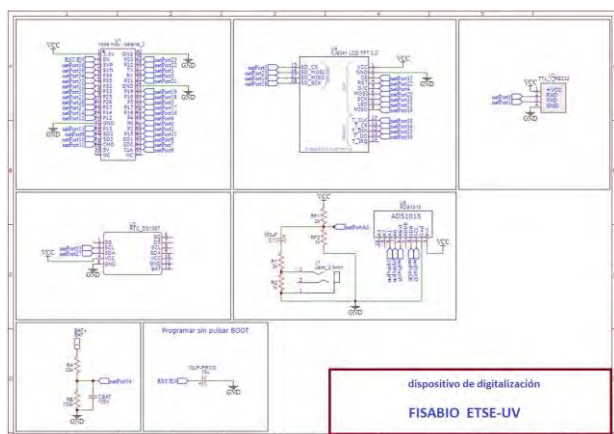


Figure 3. Wiring diagram

Microcontroller programming: the IDE (Development Environment) is the Arduino IDE V 2.0.0.

Testing: a DYNATECH-NEVADA 215A Patient Simulator and a digital oscilloscope were used to establish the accuracy of the fetal ECG signal recording. The integrity of the data recorded from the data blocks issued by the FM was checked by means of a commercial RS 232 recorder (AirDrive RS-232 Recorder) connected in series between FM and DD.

System updating: DD can connect to a Wi-Fi network to check if an update is available, download and install new firmware and restart automatically.

3.2. Comparison of methods of STV assessment

3.2.1 Population

Ten consecutive CTG records, digitized with the DD in single pregnancies who underwent internal monitoring on clinical indication during June-July 2023 were included. Digitalization was carried out from the start of internal monitoring until delivery. The duration ranged between 10 and 260 minutes.

For the purpose of comparing the STV16 and STV240 algorithms when an external Doppler monitoring is used we studied joined together the fragments from the ten cases where the twin mode was employed (external and internal monitoring on the same pregnant woman). It allows us to collect simultaneously FHR from external Doppler monitoring, FHR from internal monitoring, uterine activity and fetal ECG signal. A total of 52 minutes of recording time was used for this purpose. STV16 and STV240 groups were compared with STV values from ECG signal (ecgSTV).

Instantaneous variability of FHR from ECG signal was calculated from the time that elapsed between R peaks of consecutive waves. Matlab findpeaks function, which identifies values and locations of local maxima, was employed to find them. These values were averaged by the minute to obtain a format in accordance with those achieved with STV16 and STV240 algorithms. Calculations were made on the entire record without removing fragments with episodic changes from the

baseline. Artefacts were ruled out by disregarding instantaneous FHR variation values greater than 25 bpm.

Institutional review board approval was obtained.

3.2.2 Statistics:

Mean values of STV16, STV240 and ecgSTV groups were compared by paired samples t-test. The agreement between series was evaluated by means of Intraclass Correlation Coefficients (ICC, McGraw-Wong 1996). The A-1 version shows degree of absolute agreement and C-1 the degree of consistency among measurements. Also, Bland-Altman plots were depicted. For all statistical tests, a significant level of 0.05 was pre-defined. Calculations were computed in MATLAB R2023b.

4. Results

The DD developed is able to acquire FHR traces obtained from internal and external monitoring together with the fetal raw ECG signal as a model in order to have adequate data to calculate computerized parameters. Therefore, it allows us to contrast the agreement of the algorithms used to estimate STV (STV16 and STV240) from external monitoring with that obtained from the fetal ECG raw signal.

Figure 4 shows a cardiotocographic monitoring comparing FHR depiction obtained from the fetal ECG signal with the frames of FHR provided by the FM from the external US monitoring and the internal monitoring. US Doppler monitoring trace is slightly delayed in relation to the internal monitoring trace because of the methodological extraction of the former (autocorrelation technique on Doppler envelope) which tends to average the durations of successive cycles.

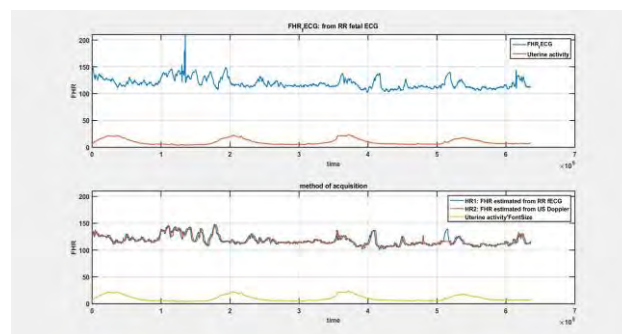


Figure 4. Comparison of traces acquired from ECG signal (top), US Doppler and internal monitoring (bottom). Time in msec

Figure 5 shows instantaneous STV as calculated from the ECG signal and then compared with this same STV averaged by the minute and with the by the minute-STV extracted by means of STV16 and STV240 algorithms. It represents HR1 and HR2 channels in a trace of ten minutes of a case with normal perinatal outcome, corresponding to the tracing in figure 4.

Mean values of STV16, STV240, ecgSTV groups from US Doppler signal acquisition were: 9.85 (+/-8.62), 2.22 (+/- 5.57), 3.68 (+/- 2.25) msec., respectively. Paired samples t-test shows no difference between STV240 and

stvECG. Nevertheless, STV16 is significantly higher than STV240 and stvECG. Our results are similar to those described in the literatura [1].

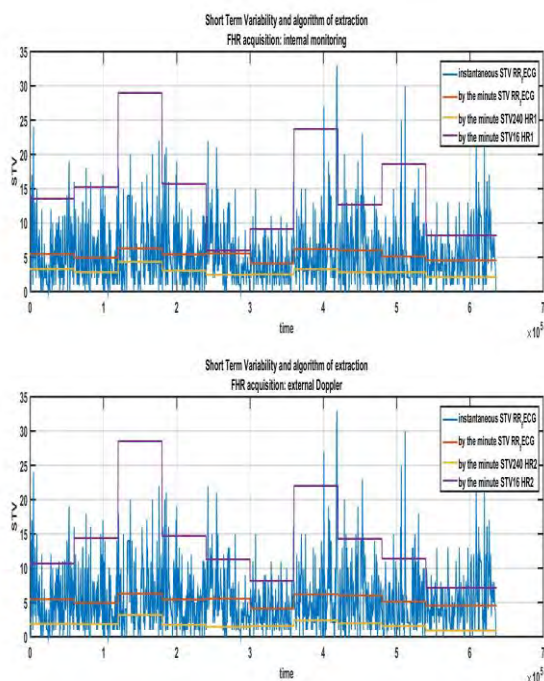


Figure 5. Comparison of STV (msec.) from: instantaneous variability (fetal ECG signal), average by the minute and STV16 and STV240 algorithms: top internal monitoring, bottom down US Doppler

Tables 1 and 2 describe the level of reliability of STV measurements carried out by the STV16 and STV240 algorithms with regard to values calculated from the ECG by Intraclass Correlation Coefficients. All C-1 ICCs become significant for $p=0.05$ (all series show consistency between them). A-1 ICC shows total agreement between STV240 and ecgSTV but does not become significant between STV16 and the other two: STV16 overestimates ecgSTV (also STV240) as shown in the Bland-Altman plots in Figure 6 and the difference increases for high STV figures.

Consistency	ICC C-1	confidence interval 95%	p
ECGstv vs STV240	0.33	0.04-0.57	0.01
ECGstv vs STV16	0.36	0.07-0.59	<0.01
STV240 vs STV16	0.65	0.44-0.79	<0.01

Tabla 1. Consistency among measurements

Agreement	ICC A-1.	confidence interval 95%	p
ECGstv vs STV240	0.32	0.04-0.56	0.01
ECGstv vs STV16	0.24	-0.05-0.51	0.06
STV240 vs STV16	0.42	-0.09-0.73	0.07

Tabla 2. Agreement among measurements

5. Conclusions

STV16 is a well established algorithm evaluating short term variability. It has proven its effectiveness in adverse perinatal outcome recognition in predelivery assesment of fetal wellbeing but not during labour. Its design was empirically based on the basis of computational and storage capacity in the devices available at the time of its implementation and not on a pathophysiological basis such as fetal electrophysiological activity [3].

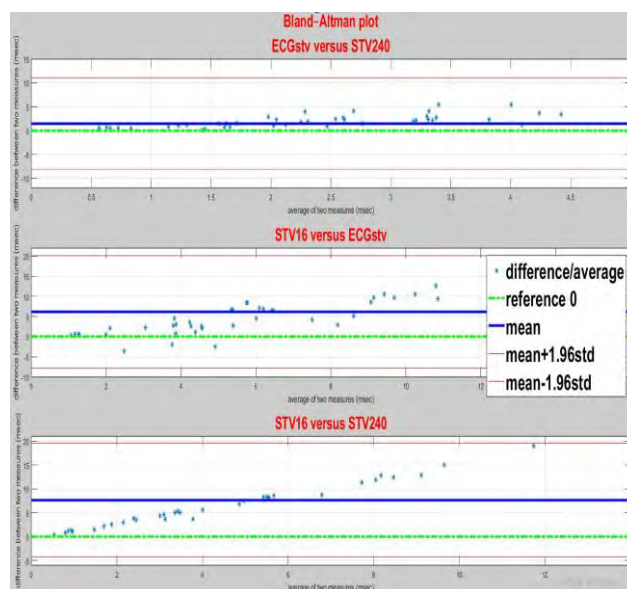


Figure 6. Bland-Altman plots: top ECGstv versus STV240, middle ECGstv versus STV16, bottom STV240 versus STV16

Improving assisted interpretation in cardiocography needs to explore new methods of computerized analysis and machine learning models [4]. For that purpose, it is essential to have suitable records with raw signal/traces similar to those provided by the device developed.

Acknowledges

This work has been funded by the Fundación para el fomento de la investigación sanitaria y biomédica de la comunitat valenciana (FISABIO UGP-21-045).

References

- [1] Kouskouti et al. Validation of a new algorithm for the short-term variation of the fetal heart rate: an antepartum prospective study. *J Perinat Med.*, vol 46(6), pp 599-604 2018 (doi: 10.1515/jpm-2017-0035. PMID: 28672744)
- [2] Jezewski J et al. Comparison of doppler ultrasound and direct electrocardiography acquisition techniques for quantification of fetal heart rate variability. *IEEE Trans Biomed Eng.*, vol 53(5), pp 855-64, 2006 (doi: 10.1109/TBME.2005.863945. PMID: 16686408).
- [3] Jones GD et al. Computerized Analysis of Antepartum Cardiotocography: A Review. *Maternal-Fetal Medicine*, vol 4(2), pp 130-140, 2022. DOI: 10.1097/FM9.0000000000000141
- [4] Aeberhard JL et al. Artificial intelligence and machine learning in cardiocography: A scoping review. *Eur. J. Obstet. Gynecol. Reprod. Biol.* vol 281, pp 54-62, 2023. PMID: 36535071 DOI: 10.1016/j.ejogrb.2022.12.008

Prototipo de un sistema integral de nariz-lengua-ojo electrónicos para la detección de enfermedades en muestras de orina

A. Ruiz-Canales¹, D. Parras-Burgos², B. Miras Cabrera³, G. Martínez-Muñoz¹, A. Conesa-Celdrán¹, A. de la Rica-Martínez⁴, M.J. Oates¹

¹ Dpto. de Ingeniería, Universidad Miguel Hernández de Elche, Alicante, España, acanales@umh.es; gemma.veinte@gmail.com; agusconesaceldran@gmail.com; moates990@gmail.com

² Dpto. de Estructuras, Construcción y Expresión Gráfica, Universidad Politécnica de Cartagena, Cartagena, España, dolores.parras@upct.es

³ Dpto. de Física Aplicada y Tecnología Naval, Universidad Politécnica de Cartagena, Cartagena, España, baltasar.miras@upct.es

⁴ Hospital General Universitario de Elche, Elche, Alicante, España, albadelarica@gmail.com

Resumen

Tradicionalmente la nariz, la lengua y el ojo electrónicos han sido tecnologías costosas y al alcance de muy pocos usuarios. Estos dispositivos cuentan con múltiples utilidades, desde el control de calidad en el ámbito industrial hasta el análisis de diferentes muestras en el ámbito médico. En este sentido, surge la necesidad de desarrollar un prototipo capaz de integrar estas tres tecnologías, que sea de bajo coste y completamente funcional. Para el desarrollo de este prototipo se analizaron los tres sistemas en profundidad y se definieron diferentes requerimientos para su diseño: ergonómico, compacto, intuitivo, de fácil limpieza, y con aspectos destacados de estética y funcionalidad. Además, el prototipo desarrollado puede ser capaz de analizar muestras en modo discreto y en modo continuo con la parte de nariz electrónica, tiene alimentación autónoma mediante células fotovoltaicas, conectividad mediante salidas USB y una tarjeta de conexión inalámbrica GPRS. El primer prototipo se ha fabricado mediante impresión 3D por FDM (Fused Deposition Modeling) en material PET-G, asegurándose de esta forma un bajo coste de fabricación. Junto con la descripción del dispositivo global se muestra un ejemplo de la metodología de análisis de datos posteriores empleada en nariz electrónica, que es genérica para las tres tecnologías citadas. Los resultados muestran ser prometedores y precisan de un mayor entrenamiento de muestras, junto con un avance en la automatización de todos los procedimientos estadísticos y analíticos.

1. Introducción

El uso de la nariz, la lengua y el ojo electrónicos tiene múltiples utilidades en diferentes ámbitos, por ejemplo, el control de calidad y el análisis de muestras en diferentes industrias como la alimentaria [1]. En el ámbito clínico, ya es posible utilizar la nariz y la lengua electrónica como sistema para la detección precoz de enfermedades a partir de los compuestos orgánicos volátiles (*Volatile Organic Compounds* en inglés, VOCs) liberados por el cuerpo humano demostrando, por ejemplo, su eficacia a partir de muestras de orina en la detección y el diagnóstico de enfermedades como la diabetes y las infecciones bacterianas, el cáncer de próstata, etc. [2-4]. Tradicionalmente, este tipo de tecnologías son costosas y al alcance de muy pocos usuarios, pero con los avances mostrados en recientes publicaciones se muestran

alternativas más asequibles y versátiles [5-8]. Por otro lado, en el mercado actual se pueden encontrar dispositivos monofuncionales con una única tecnología, pero no se ofrece la posibilidad de albergar en un mismo dispositivo los tres sistemas.

La nariz electrónica es un dispositivo sensible a los VOCs. Mediante sensores de gases comerciales se genera una señal electrónica relacionada con un conjunto de compuestos orgánicos volátiles, dependiente del tipo de sensor. En este sentido, este dispositivo se ha empleado también en el estudio de la detección de enfermedades en orina mediante nariz electrónica por el grupo de trabajo del Departamento de Ingeniería de la Universidad Miguel Hernández de Elche (UMH) y del Centro de Investigación e Innovación Agroalimentario y Agroambiental (CIAGRO).

Por otro lado, el ojo electrónico es un microfotofluorómetro a partir de diodos LED dopados. En el caso específico de esta comunicación se ha probado con un prototipo de 6 fuentes, con 8 sistemas de detección con una pantalla RGB LCD de 320x240 píxeles. Este dispositivo genera señales en función de la longitud de onda dependiendo de la sustancia a analizar, el cual es posible emplearlo en la detección de enfermedades en orina.

Finalmente, las lenguas electrónicas son sistemas de análisis basados en potenciostatos. Un potenciostato es un dispositivo electrónico requerido para controlar una celda de tres electrodos y ejecutar la mayoría de los experimentos electroanalíticos. Un bipotenciostato y un polipotenciostato son capaces de controlar dos electrodos de trabajo y más de dos electrodos de trabajo, respectivamente. El sistema funciona manteniendo el potencial del electrodo de trabajo a un nivel constante con respecto al potencial del electrodo de referencia mediante el ajuste de la corriente en un electrodo auxiliar. Incluye un sistema de tres electrodos para las investigaciones de los mecanismos de reacción relacionados con la química redox y otros fenómenos químicos. Permite detectar solutos específicos en una muestra líquida. El grupo de trabajo de la UMH está empleando lenguas electrónicas comerciales y de bajo coste, aunque de momento no ha desarrollado

mucho esta tecnología y se ha centrado más en la nariz y el ojo electrónicos [9].

En este sentido, surge la necesidad de desarrollar en un único dispositivo un sistema que pueda albergar estos tres órganos electrónicos y que cumpla con determinados requerimientos de diseño, tales como: que sea un diseño ergonómico, intuitivo en su utilización, de bajo coste, posibilidad de alimentación autónoma, etc. En este trabajo se muestra el desarrollo de un prototipo con estas tres tecnologías, fabricado por impresión 3D, consiguiendo un producto versátil y asequible.

2. Materiales y métodos

Para el desarrollo de este trabajo se parte de tres órganos electrónicos ya construidos en fase prototipo (nariz, lengua y ojo electrónico). La nariz electrónica está compuesta por sensores de óxido de metal MQ o MOQ. La lengua electrónica es de tipo voltamétrico con 3 electrodos alojados en un cilindro con relleno de epoxi, electrodo de referencia Ag|AgCl/KCl_{sat} y contraelectrodo de platino. El ojo electrónico es un fotofluorómetro de bajo coste formado por 6 fuentes y 8 detectores.

Tras un estudio de los elementos y componentes que conforman los órganos electrónicos y tras una fase creativa para el diseño de la apariencia estética del producto, se modela mediante el software de diseño asistido por ordenador SolidWorks 2020 (Dassault Systemes) un dispositivo que cumple con diferentes tipos de requerimientos:

- Diseño de un producto ergonómico y compacto.
- Aspecto atractivo y de diseño minimalista.
- Fácil manejo e intuitivo en su utilización.
- Bajo coste de fabricación.
- Un dispositivo de fácil limpieza.
- Unión de las tres tecnologías en un mismo dispositivo.
- Posibilidad de medir con la nariz electrónica muestras discretas y continuas.
- Posibilidad de alimentación autónoma mediante placas fotovoltaicas.
- Un diseño con distintos módulos que permitan utilizar cada una de las tecnologías.
- Facilidad de agarre y transporte.
- Sistema de protección para la manipulación de sustancias agresivas.

Para la alimentación autónoma del sistema se cuenta con 9 placas solares de 35x65 mm de 5v y 30mA, otorgando una potencia por placa de 0,15 W, teniendo en total 1,35 W. Estas placas pueden cargar una batería con unas dimensiones máximas de 90x70x100 mm de tamaño, siendo esta de 12V 4Ah como máximo.

3. Resultados y discusión

3.1 Diseño del prototipo

El diseño final resultante es un dispositivo compacto, versátil y accesible a las tres tecnologías por separado (Fig. 1) y con diferentes elementos de conectividad, como pantallas frontales para el control y visualización de información, tarjeta de conexión inalámbrica GPRS, y conexiones USB en la parte trasera (Fig. 2). Además, el

dispositivo permite acceder al interior (Fig. 3a) para la reparación y mantenimiento de los componentes electrónicos, y dispone de una cubierta de protección que permite su fácil transporte (Fig. 3b).



Figura 1. Disposición de muestras en: a) la nariz electrónica, b) la lengua electrónica, y c) el ojo electrónico.

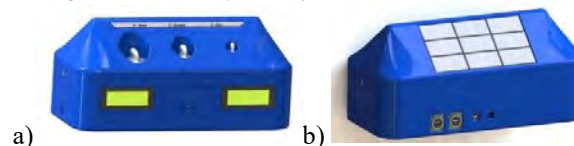


Figura 2. Modelo 3D del prototipo: a) Vista frontal, b) Vista trasera.

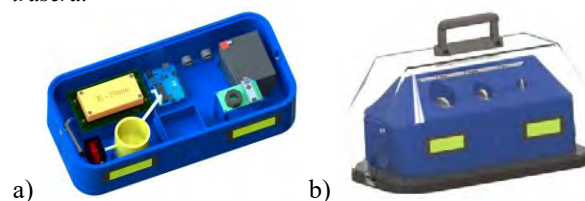


Figura 3. Modelo 3D del prototipo: a) Interior del dispositivo, b) cubierta de protección y transporte.

El primer prototipo (Fig. 4), se fabricó en Tereftalato de polietileno modificado con glicol (PETG), mediante impresión 3D por deposición de material fundido (FDM). El relleno de las piezas para la fabricación de los distintos componentes ha sido variable entre el 12% y el 50% dependiendo de la ubicación de la pieza y los esfuerzos que esta debe soportar, y con espesores de 1 mm de pared. Para la unión y fijación de los elementos se han empleado principalmente uniones roscadas mediante tornillos autorroscantes y pegamento específico para este tipo de plástico. El coste aproximado de la fabricación del prototipo completo ha sido de 450 €.



Figura 4. Imagen del prototipo fabricado por FDM.

3.2 Análisis estadístico de los resultados de las muestras de la nariz electrónica

La metodología estadística descrita a continuación se usa en las tres tecnologías (nariz, lengua y ojo electrónico). Específicamente se ha hecho un análisis de su aplicación en la detección de la infección de orina. En este estudio se llevó a cabo un análisis de 203 muestras de orina infectada y no infectada. Además de medir con el prototipo descrito, se analizaron los datos con medidas analíticas estándar para la orina (urocultivo).

El urocultivo se realizó cuantitativamente utilizando bucles calibrados de 1 μ L. Cada colonia aislada correspondía a 1.000 UFC/mL. La incubación se realizó a 35-37°C en atmósfera enriquecida de CO₂ en base de agar sangre y a la misma temperatura en atmósfera de aerobiosis en agar cistina-lactosa-deficiente en electrolitos (agar CLED). La

interpretación de los resultados del urocultivo se realizó según los criterios originales propuestos por Kass, considerando infecciones urinarias aquellas que presentaban en el cultivo ≥ 100.000 unidades formadoras de colonias (UFC). Estos datos analíticos se contrastaron con medida simultánea de nariz electrónica.

Con estos datos se realizó, en primer lugar, un análisis de componentes principales (PCA). Con este tipo de análisis se consigue reducir el número de variables representativas de la muestra a 2 y así se puede representar cada muestra en un gráfico 2D. La distancia y distribución espacial de los puntos en esta gráfica indica posibles relaciones y agrupaciones de las muestras. En segundo lugar, se entrenaron modelos basados en máquinas de vector de soporte (*Support Vector Machine* en inglés, SVM) con el objetivo de poder obtener una herramienta para una clasificación rápida de los datos. En estos modelos se entrenaron las muestras incluyéndolas en cada una de las 2 categorías; infectados y no infectados.

Se realizó, además, un segundo análisis para estudiar la capacidad de clasificación de las diferentes infecciones presentes en el grupo de muestras infectadas. Se detectaron un total de 21 infecciones diferentes en las muestras analizadas. Solamente 4 de estas infecciones tenían un número de muestras superior a 10, por lo cual, solamente éstas han sido utilizadas para realizar un primer modelo. Se presenta en las siguientes figuras un resumen del modelo y los datos utilizados, así como los resultados de la predicción en una matriz de confusión. Ambos modelos estaban basados en SVM, y se usaron un 85% de los datos de todos los análisis realizados para entrenarlos y el 15% restante para realizar una validación del modelo. La clasificación genérica de las muestras es de un total de 203, las muestras infectadas fueron 106, y las muestras no infectadas, 97.

De las 21 infecciones identificadas se escogieron 5 porque mostraron diferencias significativas: *E. coli*, *Klebsiella pneumoniae*, *Enterococcus faecalis* y *Proteus mirabilis*. Se puede ver en la Figura 5 una gráfica donde se representan los dos primeros coeficientes principales (PCs) que incluyen un 69% de la variabilidad de los datos. No se ve una diferenciación visual clara de ambos grupos control, representado en amarillo, e infectado, representado en lila. El algoritmo *k-means* basado en PCs es capaz de discernir ambos grupos con un 51,26% de precisión. Este algoritmo clasifica de forma no reforzada todo el conjunto introducido, es decir es capaz de clasificar 100 del total de las muestras de forma correcta. Además, visualmente se puede observar que, dentro de los grupos, el infectado presenta una mayor variabilidad entre muestras que el control.

Usando modelos basados en SVM se consiguió mejorar la precisión de la clasificación aumentándola a un 75% de precisión. Para ello se entrenaron usando de forma aleatoria un 85% de los datos para entrenamiento y un 15% para validación. Es decir, el modelo se evalúa con un total de 31 muestras de las cuales 23 clasifica de forma correcta. Esto se puede ver en la matriz de confusión mostrada en la Figura 6. Conforme aumente el número de muestras se podrá tener un conjunto de evaluación de mayor tamaño, de forma que tendremos resultados más fiables. La etiqueta 0 representa muestras infectadas y la 1 no infectadas. Se

puede observar que el modelo presenta una mayor precisión clasificando muestras no infectadas. Mientras que el 72% de las muestras infectadas fueron clasificadas correctamente, el 82% de las no infectadas se clasificaron como tal.

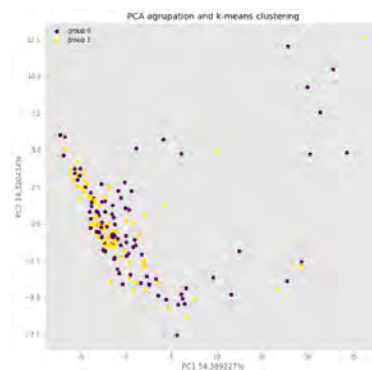


Figura 5. Distribución de los datos experimentales. Total: 203, infectados 106, y no infectados 97.

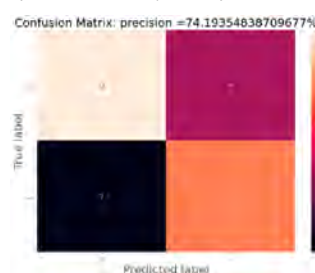


Figura 6. Matriz de confusión inicial.

En la Figura 7 se muestra un gráfico con los 4 tipos de muestras marcadas en diferentes colores. No se ve una distribución de las muestras clara de *E. coli* y *Klebsiella*, pero se puede agrupar todas las demás de forma visual según las diagonales que forman.

Mediante los modelos basados en SVM se consigue mejorar la precisión de la clasificación aumentándola a un 81%. Para ello se han entrenado usando de forma aleatoria un 85% de los datos para entrenamiento y un 15% para validación. Es decir, el modelo se evaluó con un total de 11 muestras de las cuales 9 se clasificó de forma correcta. Esto se puede ver en la matriz de confusión mostrada en la Figura 8. Conforme aumente el número de muestras se puede tener un conjunto de evaluación de mayor tamaño, de forma que se tendrán resultados más fiables, debido a que el número de muestras de testeo es muy bajo no se pueden sacar conclusiones firmes. El modelo muestra aun así una buena capacidad de clasificación y una buena distribución entre las muestras.

Estos resultados, además de ser debidos a un entrenamiento parcializado de los modelos, puede implicar una relación médica subyacente. En el gráfico que arroja el análisis PCA de la Figura 7, se puede ver que las muestras infectadas se agrupan de forma más compacta de lo que lo hacen las no infectadas. Esto se puede entender a que estos datos son más similares entre sí, es decir, presentan menor variabilidad.

Por otro lado, se observa que las muestras no infectadas están mucho más dispersas en el gráfico, lo que significa que presentan mucha mayor variabilidad en las medidas. Esto podría afectar al modelo haciendo que sea más capaz de clasificar las infectadas al poder relacionarlas mejor como un mismo grupo.

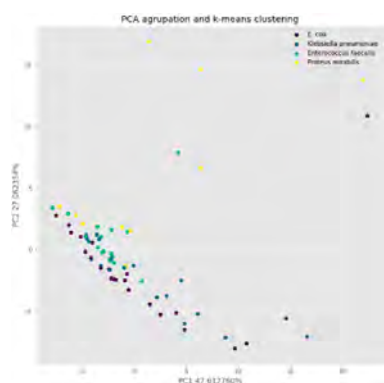


Figura 7. Análisis PCA posterior usando modelos SVM. Se aprecia una diferenciación de cada tipo de infección.

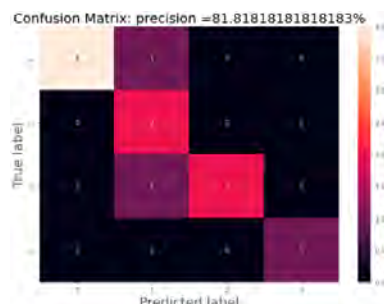


Figura 8. Matriz de confusión final.

3.3 Empleo del dispositivo para diagnóstico rápido y de bajo coste

A pesar de que los resultados para diferenciar el agente etiológico que produce la infección no son muy sólidos hasta el momento, representan un buen precedente para aumentar la precisión de esta prueba diagnóstica. Los resultados obtenidos probablemente puedan mejorarse aumentando el número de muestras medidas para el entrenamiento del sistema. Por lo tanto, es necesario realizar más investigaciones sobre este tema. Además, la producción de ciertos COV es característica de las bacterias [10] por lo que es muy probable que el perfeccionamiento de esta tecnología permita en el futuro diagnosticar otras patologías infecciosas midiendo la emisión de estos compuestos en diferentes fluidos biológicos, como el pleural, líquidos, lavados broncoalveolares o líquidos articulares. La investigación en este campo es necesaria para encontrar nuevas líneas de diagnóstico que aceleren la optimización de los tratamientos antimicrobianos.

La alta prevalencia de infecciones del tracto urinario sitúa a esta patología como un buen objetivo para mejorar el sistema sanitario. La identificación de herramientas diagnósticas nuevas y precisas capaces de seleccionar a los pacientes que necesitan tratamiento antibiótico sigue siendo una cuestión fundamental en el ámbito sanitario. El porcentaje económico que representa para cualquier sistema sanitario hace pensar que la inversión en avances encaminados a optimizar el manejo terapéutico de esta enfermedad tendrá un impacto positivo en poco tiempo.

4. Conclusiones

En este estudio se ha partido de tres tecnologías independientes e innovadoras como son la nariz, la lengua y el ojo electrónico, llevando a cabo el diseño y desarrollo

de un sistema que es capaz de integrar en un solo dispositivo las funcionalidades de las tres tecnologías. Los requerimientos de diseño definidos al comienzo del proyecto se han cumplido, consiguiendo un producto ergonómico, compacto y minimalista, intuitivo en su utilización, y con posibilidad de alimentación autónoma. El coste de fabricación obtenido es asequible y abre diferentes opciones de desarrollo para la obtención de prototipos de bajo coste para el uso de este tipo de tecnologías con un dispositivo multifuncional.

Los resultados del ejemplo mostrado de análisis de datos posteriores, genérico para las tres tecnologías citadas, han mostrado ser prometedores para la detección de infección de orina y precisan de un mayor entrenamiento de muestras, junto con un avance en la automatización de todos los procedimientos estadísticos y analíticos.

Agradecimientos

Los autores agradecen a la empresa TELENATURA EBT S.L. el apoyo tecnológico prestado en la realización de los experimentos y del prototipo descrito.

Referencias

- Oates, M.J.; Fox, P.; Sanchez-Rodriguez, L.; Carbonell-Barrachina, Á.A.; Ruiz-Canales, A. DFT based classification of olive oil type using a sinusoidally heated, low cost electronic nose. *Computers and Electronics in Agriculture* **2018**, *155*, 348-358. doi: <https://doi.org/10.1016/j.compag.2018.10.026>
- Bax, C.; Capelli, L.; Grizzi, F.; Prudenza, S.; Taverna, G.; Ieee. In *A novel approach for the non-invasive diagnosis of prostate cancer based on urine odour analysis*, IEEE International Symposium on Olfaction and Electronic Nose (ISOEN), Aveiro, PORTUGAL, May 29-Jun 01, 2022; Aveiro, PORTUGAL, 2022.
- Sabeel, T.M.A.; Eluwa, S.E.; FauzanKhairiCheHarun; Sabeel, S.M.A. In *Detection of volatile compounds in urine using an electronic nose instrument*, International Conference on Computer, Electrical and Electronics Engineering (ICCEEE), Khartoum, SUDAN, Aug 26-28, 2013; Khartoum, SUDAN, 2013.
- Saidi, T.; Moufid, M.; Zaim, O.; El Bari, N.; Ionescu, R.; Bouchikhi, B. In *Artificial sensory systems combined with UV-Vis Spectrophotometry as a robust approach for VOCs analysis of human urine and exhaled breath*, 26th Anniversary World Congress on Biosensors (Biosensors), Gothenburg, SWEDEN, May 25-27, 2016; Gothenburg, SWEDEN, 2017; pp 256-257.
- Abu-Khalaf, N.; Hmidat, M. Visible/Near Infrared (VIS/NIR) spectroscopy as an optical sensor for evaluating olive oil quality. *Computers and Electronics in Agriculture* **2020**, *173*, 105445. doi: <https://doi.org/10.1016/j.compag.2020.105445>
- Apetrei, C.; Apetrei, I.M.; Villanueva, S.; de Saja, J.A.; Gutierrez-Rosales, F.; Rodriguez-Mendez, M.L. Combination of an e-nose, an e-tongue and an e-eye for the characterisation of olive oils with different degree of bitterness. *Analytica Chimica Acta* **2010**, *663*, 91-97. doi: <https://doi.org/10.1016/j.aca.2010.01.034>
- Oates, M.J.; Abu-Khalaf, N.; Molina-Cabrera, C.; Ruiz-Canales, A.; Ramos, J.; Bahder, B.W. Detection of Lethal Bronzing Disease in Cabbage Palms (Sabal palmetto) Using a Low-Cost Electronic Nose. **2020**, *10*, 188. doi: <https://doi.org/10.3390/bios10110188>
- Sanaeifar, A.; ZakiDizaji, H.; Jafari, A.; Guardia, M.d.I. Early detection of contamination and defect in foodstuffs by electronic nose: A review. *TrAC Trends in Analytical Chemistry* **2017**, *97*, 257-271. doi: <https://doi.org/10.1016/j.trac.2017.09.014>
- Pascual, L.; Campos, I.; Vivancos, J.-L.; Quintás, G.; Loras, A.; Martínez-Bisbal, M.C.; Martínez-Mañez, R.; Boronat, F.; Ruiz-Cerdá, J.L. Detection of prostate cancer using a voltammetric electronic tongue. *Royal Society of Chemistry* **2016**, *141*, 4562-4567. doi: <https://doi.org/10.1039/C6AN01044J>
- Guernion, N.; Ratcliffe, N.M.; Spencer-Phillips, P.T.N.; Howe, R.A. Identifying bacteria in human urine: current practice and the potential for rapid, near-patient diagnosis by sensing volatile organic compounds. *Clinical Chemistry and Laboratory Medicine Journal* **2005**, *39*, 893-906. doi: <https://doi.org/10.1515/CCLM.2001.146>

Modelado y Simulación

Miércoles 22 15:00 - 16:30

Paraninfo

Rol de la dirección de fibras y heterogeneidades iónicas en la simulación de arritmias auriculares

J. Barrios Álvarez de Arcaya¹, M. Termenón-Rivas¹, G. Romitti¹, R. Sebastián¹, A. Liberos¹, M. Rodrigo¹

¹ CoMMLab, Universitat de València, Valencia, España, miguel.rodrigo@uv.es

Resumen

Los gemelos digitales se plantean como herramientas útiles para la simulación del comportamiento fibrilar de pacientes con fibrilación auricular (FA), en las que ensayar potenciales terapias. No obstante, el grado de complejidad de estas simulaciones para su uso clínico es incierto. Para mejorarlas se incluyen dos características del sustrato auricular, como son la dirección de fibras (DF) y las heterogeneidades iónicas (HI).

Se utilizó un modelo 3D anatómico de las aurículas donde se incluyó tanto DF como HI correspondientes a los diferentes tipos de tejido auricular. En este modelo, se simuló en múltiples protocolos de estimulación para el inicio de FA, incluyendo y eliminando DF y HI. Posteriormente, se midió y caracterizó el número y tipo de arritmias generadas para evaluar el efecto independiente de la DF y las HI en estas simulaciones.

En promedio, se generaron patrones arrítmicos en el 37% de las simulaciones, siendo un 4% reentradas funcionales y un 33% reentradas anatómicas, principalmente localizadas en el septo y en la cava inferior respectivamente. La DF tiene un rol muy importante al conseguir un 51% de reentradas si esta se encuentra sobre el modelo frente al 24% si no son incluidas. Las HI no tienen un efecto notorio, ya que los cambios al incluirlas fueron de 39% vs 38% cuando no se incluyó.

La inclusión de DF en simulaciones detalladas es necesaria para conseguir comportamientos reentrantes similares a los observados en la práctica clínica y lograr una mayor similitud del modelo digital con el caso real de cada paciente.

1. Introducción

Las arritmias cardíacas son una de las principales causas de mortalidad en países desarrollados. De entre ellas, la arritmia que más se produce es la fibrilación auricular (FA). Sin embargo, los mecanismos por los que se induce y mantiene esta patología no se comprenden totalmente. Con el objetivo de comprender mejor esta patología se usan modelos computacionales [1]. En estos modelos se pueden simular diferentes anatomías y diferentes escenarios, como la generación de patrones reentrantes, para obtener información sobre los mecanismos de inicio y mantenimiento de la arritmia. Además, gracias a estas simulaciones se puede realizar un diagnóstico y tratamiento óptimo de cada arritmia específica.

La cantidad de detalle de estos modelos computacionales es, por otra parte, un aspecto todavía sin determinar a la hora de reproducir comportamientos reentrantes realistas en anatomías específicas de pacientes. Un factor a incluir es la dirección de fibras (DF) en la que cada tejido muestra mayor conductividad, por lo que es necesario definir para cada tejido su difusión en las direcciones longitudinal y transversal [2]. Otro potencial factor a tener en cuenta son

las heterogeneidades iónicas (HI), ya que no en todos los tejidos hay las mismas concentraciones ni canales iónicos y, consecuentemente, iguales corrientes, lo que provoca distintos comportamientos a nivel electrofisiológico.

En este estudio se analiza la inclusión de dirección de fibras (DF) y heterogeneidades iónicas (HI) sobre una anatomía en condiciones de fibrilación auricular crónica, y así ver el rol de cada uno de estos factores para simular comportamientos arrítmicos.

2. Materiales y métodos

2.1. Modelo anatómico heterogéneo

Como modelo anatómico se usó la anatomía auricular propuesta por Krueger et al. [3], conformada por las dos aurículas, izquierda y derecha, en la que se incluye la dirección de fibras de cada nodo del modelo y se diferencian distintos tipos de tejidos auriculares, entre los que diferenciamos, aurícula derecha (AD), aurícula izquierda (AI), nodo sinusal (NS), crista terminalis (CT), músculos pectíneos (PM), haz de Bachmann (HB), istmo cavo tricúspideo (ICT), orejuela derecha (OD), orejuela izquierda (OI), venas cavas (VC), venas pulmonares (VP), válvula tricúspide (VT), válvula mitral (VM), fosa oval (FO), tal y como se muestra en el panel A de la Figura 1.

Cada región auricular se diferencia por su anisotropía: cociente entre la difusión longitudinal y la transversal de cada tipo de tejido cuyos valores se muestran en la Tabla 1.

Región	Difusión transversal ($\mu\text{m}^2/\text{ms}$)	Ratio difusión longitudinal/transversal
AD, AI	0,12	3,75
NS	0,44	1
CT	0,12	6,56
MP	0,05	23,25
ITC,VC,OI,OD	0,12	1
VM,VT,FO	0,08	17

Tabla 1: Difusiones transversales y ratios de difusión

Por otro lado, las HI se introdujeron variando los valores de las corrientes iónicas para cada tejido. Estos valores fueron obtenidos teniendo en cuenta varios artículos [4-10]. En la Tabla 2 se muestran los factores por los que se

multiplica a las diferentes corrientes iónicas de cada tejido, teniendo en cuenta como modelo basal el tejido de la aurícula izquierda.

Región	g_{to}	g_{CaL}	g_{Kr}	g_{Kl}	g_{Ks}
AD	1	1	0,625	1	1
AI, NS, HB	1	1	1	1	1
CT	1,35	1,6	0,9	1	1
MP	1,05	0,95	0,9	1	1
ICT, OD, FO	1	1	0,7906	1	1
OI	0,65	1,05	2,75	1	1
VM, VT	1,05	0,65	3	1	1
PV	1,35	0,4	2	0,7	1

Tabla 2. Heterogeneidades iónicas relativas

Se generaron cuatro modelos distintos combinando la inclusión y ausencia de DF y HI. En los casos con ausencia de DF, los ratios de difusión se mantuvieron en el valor unidad, eliminando por tanto cualquier dirección de propagación dominante y haciendo el modelo isotrópico. En los casos con ausencia de HI, se usó para todo el tejido el perfil iónico de la aurícula izquierda.

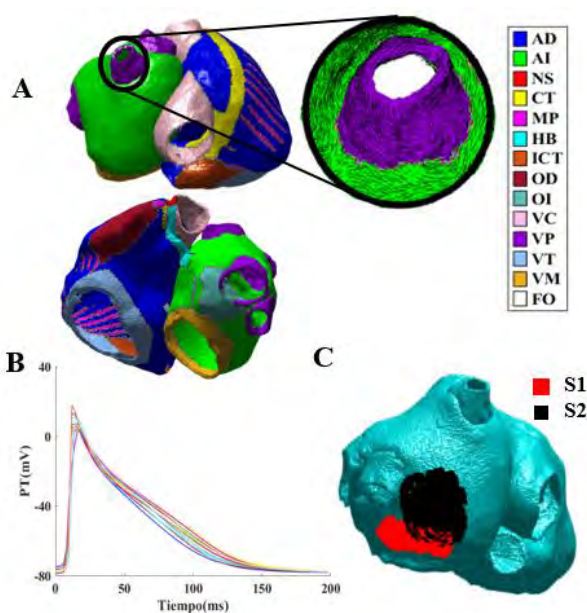


Figura 1. A) Anatomía auricular con las distintas regiones. B) Potenciales transmembrana de las regiones auriculares. C) Zonas de estimulación del protocolo S1-S2.

2.2. Simulaciones

Para las simulaciones se hizo uso de un Solver propio donde se resuelven, mediante el uso de GPUs, las ecuaciones diferenciales del modelo celular de Koivumäki [11, 12] para todos los nodos del modelo anatómico.

Este Solver incluye también la modificación del modelo iónico para incluir el efecto del remodelado eléctrico auricular provocado por la progresión de FA. Para ello, se introdujo en todos los modelos un valor de difusión eléctrica correspondiente al 75% del valor basal, y un valor de remodelado eléctrico correspondiente a una aurícula bajo FA persistente (remodelado del 75%, siendo 100% el remodelado de una FA crónica). Esta última condición implicó un incremento en la conductancia de I_{kl} del 80%, en el valor máximo de I_{NCX} del 37,5%, en la ratio entre PLB y SERCA de un 13%, en la sensibilidad de RyR al $[Ca^{2+}]_{SR}$ del 75%, y una reducción de expresión de SERCA en 12%, en la conductancia de I_{to} del 33%, en la conductancia de I_{kur} del 16%, en la conductancia de I_{CaL} del 44%, y en la ratio de SLN a SERCA de 30%.

Para provocar actividad reentrante en esta anatomía se hizo uso del protocolo S1-S2 desde las regiones marcadas en el panel C de la Figura 1. Se hizo un barrido de valores de S1 y S2 para evaluar la arritmicidad del tejido bajo diferentes ritmos. En el caso de los tiempos de S1, se varió entre 150 y 200 ms a intervalos de 10 ms, mientras que el S2 se varió entre el 60% y el 80% del valor de S1 a intervalos de 4 ms. En total se realizaron 228 simulaciones.

En las simulaciones, en primer lugar se estabilizó el modelo con una batería de 20 pulsos espaciados S1 ms. Tras los 20 pulsos, se aplicó el S2 y se mantuvo la simulación durante 10s, tras el que se analizó si había actividad reentrante, así como la naturaleza de esta, si era anatómica o funcional, y su localización.

3. Resultados

Las distintas combinaciones de S1 y S2 en presencia y ausencia de DF y HI resultaron en distintos patrones arrítmicos. Se observaron reentradas anatómicas en venas pulmonares, vena cava inferior y válvula mitral, así como reentradas funcionales en la pared posterior de la aurícula izquierda y en el septo. También se observaron muchos casos en los que el protocolo de estimulación no produjo ningún patrón fibrilar. En la Figura 2 se pueden ver ejemplos de una reentrada funcional en la pared posterior de la aurícula en el panel A, y dos reentradas anatómicas, una correspondiente a la vena cava inferior en el panel B y otra en las venas pulmonares izquierdas en el panel C, bajo diferentes condiciones del modelo.

En la Figura 3 se muestran los resultados de las simulaciones para cada combinación de condiciones del modelo, añadiendo o suprimiendo DF y HI, y para los distintos valores de S1 y S2. Las celdas coloreadas en negro son celdas no simuladas, las rojas son casos en los que no se producía reentrada, las celdas verdes más oscuras representan casos reentrantes anatómicos, y las verdes más claras casos reentrantes funcionales.

Para las simulaciones con DF y HI, todas las reentradas anatómicas con S1 de 170 y 180ms se dan en la vena cava inferior, menos el caso de S1=180ms y S2=136ms, donde la reentrada se da en la vena pulmonar superior izquierda (VPSI). Las reentradas funcionales en esas mismas condiciones se dan todas entre la pared posterior de la aurícula izquierda y el septo, tal y como se muestra en el

ejemplo del panel A de la Figura 2. En las 4 combinaciones de DF y HI, todas las reentradas anatómicas con un S1 de 150 ms se tratan de *flutter* típicos alrededor de la válvula mitral.

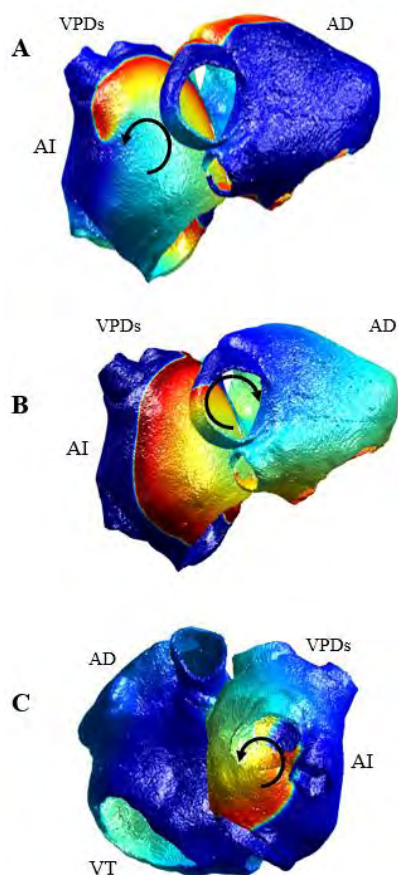


Figura 2. A) Reentrada funcional en la pared posterior de la AI en caso con DF y HI y valores de S1=190 y S2=142 ms. B) Reentrada anatómica en la vena cava inferior en caso con DF y sin HI y valores de S1=170 y S2=132 ms. C) Reentrada anatómica en venas pulmonares izquierdas sin DF ni HI y valores de S1=190 y S2=142 ms.

En la situación con DF y sin HI, las reentradas de S1=170 y 180ms, comparten a su vez localización, como se puede ver en el panel B de la Figura 2, ejemplificando la reentrada anatómica en la vena cava inferior con los valores de S1=170 y S2=132 ms. La reentrada anatómica con S1=190 y S2=136 ms se encuentra el VPSI.

Si se elimina DF pero se conserva HI, todas las reentradas son anatómicas alrededor de las dos venas pulmonares izquierdas. Por último, eliminando tanto DF como HI, las reentradas anatómicas que no son las de S1=150ms se dan en VPSI, mientras que las reentradas funcionales se dan entre la pared posterior de la aurícula izquierda y el septo.

La tabla 3 muestra un resumen de estos resultados. La cantidad de reentradas es mayor en casos con DF, mostrando patrones arrítmicos en el 51% de los casos para ambas situaciones con esta característica implementada en el modelo, con o sin heterogeneidades. Si esta característica no se aplica, los casos arrítmicos disminuyen hasta un 24% en el caso con HI y un 23% sin HI.

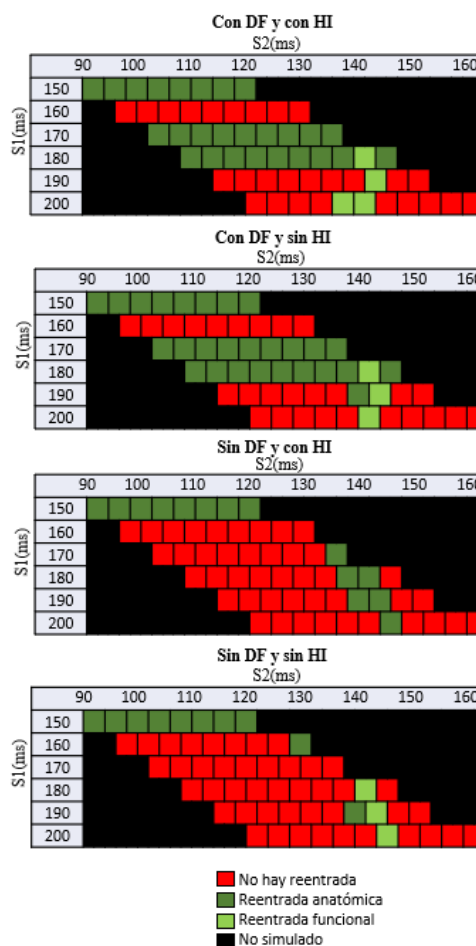


Figura 3. Comportamiento arritmico en las diferentes simulaciones realizadas

Con respecto al tipo de patrones arrítmicos iniciados, la situación con DF y HI es el caso en el que más patrones fibrilares (reentradas funcionales) se genera, aunque con poca diferencia respecto al mismo caso suprimiendo HI. Si en el modelo se elimina DF, pero sí que se mantienen las HI, todas las reentradas son anatómicas. Por último, en el caso sin DF ni HI, solo el 17% de los casos totales son anatómicas, mientras que el porcentaje de ritmos fibrilares es del 5%.

	Casos totales arrítmicos (%)	Casos reentrantes anatómicos (%)	Casos reentrantes funcionales (%)
DF/HI	51	44	7
DF/No HI	51	47	5
No DF/HI	24	24	0
No DF/No HI	23	17	5

Tabla 3. Resultados de las simulaciones.

4. Conclusiones y discusión

En este trabajo se ha visto la influencia de 2 parámetros incluidos en las simulaciones biofísicas realistas, como son la DF y HI, que han permitido generar patrones arrítmicos similares a los observados en la práctica clínica, como son

el *flutter* típico y las reentradas funcionales cerca de las venas pulmonares. Bajo las mismas condiciones de simulación, la inclusión de DF en el modelo anatómico provoca que el número de casos arrítmicos sea significativamente mayor, teniendo un 51% de casos reentrantes en los modelos con esta característica, cuando en los modelos sin DF los casos reentrantes no superan el 25%. Por otro lado, la inclusión o no inclusión de HI en el modelo no genera una gran diferencia en la cantidad de casos arrítmicos, aunque sí hay diferencia en el tipo de reentrada generado.

El efecto de la HI, tal y como se menciona en el artículo de Calvo et al [7], influencia el movimiento de las reentradas funcionales, atrayéndolas hacia las venas pulmonares. Esto se puede ver sobre todo en el caso sin DF y con HI, donde el efecto de HI es más visible, y es que no hay ninguna reentrada funcional. Las reentradas pueden comenzar siendo funcionales, pero la influencia de las HI hace que se aproximen a las venas pulmonares y una vez transformada la reentrada de funcional a anatómica esta permanece como este último tipo. Las venas pulmonares actúan entonces como centro de la reentrada durante el resto del tiempo. Mientras que en el caso sin DF ni HI, la reentrada generada es estática y por lo tanto, resulta en una reentrada funcional. Si aplicamos DF además de las HI, se mitiga el efecto de las HI, ya que en esas simulaciones sí que se pueden encontrar reentradas funcionales.

En vista de estos resultados, la inclusión de DF en los gemelos digitales es necesaria en los modelos anatómicos para simular la actividad reentrante; sin embargo, las HI no tienen un papel tan importante y, por tanto, deberían incluirse solo en casos en los que fuese altamente necesario, como la simulación de reentradas anatómicas específicas.

En cuanto a las limitaciones de este estudio, se ha realizado con tan solo una zona de S1 y S2 y en tan solo una anatomía. En el futuro se replicará en anatomías específicas de pacientes y utilizando diferentes protocolos de estimulación.

Agradecimientos

Este trabajo fue financiado por Subvención de la Generalitat Valenciana AICO/2021/318 (Consolidables 2021) y por la Subvención PID2020-114291RB-I00, financiada por MCIN/ 10.13039/501100011033 y por “ERDF A way of making Europe”.

Referencias

- [1]. O. Dössel, M. W. Krueger, F. M. Weber, M. Wilhelms and G. Seemann, "Computational modeling of the human atrial anatomy and electrophysiology," *Med Biol Eng Comput*, vol. 50, (8), pp. 773-799, 2012. (ISSN: 1741-0444)
- [2]. Jalife, J., M. Delmar, J. Anumonwo, O. Berenfeld and J. Kalifa, "Basic Cardiac Electrophysiology for the Clinician," 2009. (ISBN:9781444316940)
- [3]. M. W. Krueger, G. Seemann, K. Rhode, D. U. J. Keller, C. Schilling, A. Arujuna, J. Gill, M. D. O'Neill, R. Razavi and O. Dössel, "Personalization of Atrial Anatomy and Electrophysiology as a Basis for Clinical Modeling of Radio-Frequency Ablation of Atrial Fibrillation," *Tmi*, vol. 32, (1), pp. 73-84, 2013. (ISSN 1558-254X)
- [4]. G. Schram, M. Pourrier, P. Melnyk and S. Nattel, "Differential Distribution of Cardiac Ion Channel Expression as a Basis for Regional Specialization in Electrical Function," *Circulation Research*, vol. 90, (9), pp. 939-950, 2002. (ISSN: 1524-4571)
- [5]. J. R. Ehrlich, T. Cha, L. Zhang, D. Chartier, P. Melnyk, S. H. Hohnloser and S. Nattel, "Cellular electrophysiology of canine pulmonary vein cardiomyocytes: action potential and ionic current properties," *The Journal of Physiology*, vol. 551, (3), pp. 801-813, 2003 (ISSN:1469-7793)
- [6]. A. Dasí, A. Roy, R. Sachetto, J. Camps, A. Bueno-Orovio and B. Rodriguez, "In-silico drug trials for precision medicine in atrial fibrillation: From ionic mechanisms to electrocardiogram-based predictions in structurally-healthy human atria," *Frontiers in Physiology*, vol. 13, pp. 966046, 2022 (ISSN: 1664-042X)
- [7]. C. Calvo, M. Deo, S. Zlochiver, J. Millet and O. Berenfeld, "Attraction of Rotors to the Pulmonary Veins in Paroxysmal Atrial Fibrillation: A Modeling Study," *Biophysical Journal*, vol. 106, (8), pp. 1811-1821, 2014. (ISSN: 1542-0086)
- [8]. Sánchez, C., A. Bueno-Orovio, E. Pueyo and B. Rodríguez, "Atrial Fibrillation Dynamics and Ionic Block Effects in Six Heterogeneous Human 3D Virtual Atria with Distinct Repolarization Dynamics," *Front. Bioeng. Biotechnol.*, vol. 5, 2017. (ISSN: 1664-042X)
- [9]. G. Seemann, C. Höper, F. B. Sachse, O. Dössel, A. V. Holden and H. Zhang, "Heterogeneous three-dimensional anatomical and electrophysiological model of human atria," *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical, and Engineering Sciences*, vol. 364, (1843), pp. 1465-1481, 2006 (ISSN: 1364503X)
- [10]. Aslanidi, O. V., T. D. Butters, C. X. Ren, G. Rycroft and Henggui Zhang, "Electrophysiological models for the heterogeneous canine atria: Computational platform for studying rapid atrial arrhythmias," 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2011. (ISSN: 2375-7477)
- [11]. J. T. Koivumäki, T. Korhonen and P. Tavi, "Impact of Sarcoplasmic Reticulum Calcium Release on Calcium Dynamics and Action Potential Morphology in Human Atrial Myocytes: A Computational Study," *PLoS Computational Biology*, vol. 7, (1), pp. e1001067, 2011. (ISSN: 1553-7358)
- [12]. J. T. Koivumäki, R. B. Clark, D. Belke, C. Kondo, P. W. M. Fedak, M. M. C. Maleckar and W. R. Giles, "Na⁺ current expression in human atrial myofibroblasts: identity and functional roles," *Frontiers in Physiology*, vol. 5, pp. 275, 2014 (ISSN: 1664-042X)

Metodología para la obtención de modelos 3D de cavidades nasales con parámetros funcionales especificados

L. Picazo Selva ^{1*}, M. Burgos Olmos ¹, D. Sanz Prieto ², A. Pérez Ramos ³, F. Esteban Ortega ⁴, I. M. García ¹, M. Bastir ²

¹ Dpto. de Ingeniería Térmica y de Fluidos, Universidad Politécnica de Cartagena, Murcia, Spain

² Grupo de Paleoantropología, Departamento de Paleobiología, Museo Nacional de Ciencias Naturales – Centro Superior de Investigaciones Científicas, Madrid, España.

³ Dpto. de Ecología y Geología, Facultad de Ciencias, Universidad de Málaga, 29071, Málaga, Spain

⁴ Servicio Andaluz de Salud, Sevilla, Spain. Hospital Universitario Virgen del Rocío, Servicio de Otorrinolaringología.

Resumen

Se presenta una metodología novedosa para la obtención de modelos tridimensionales de cavidades nasales con parámetros funcionales especificados utilizando el software Flowgy (www.flowgy.com). La metodología parte de la cavidad nasal patológica de un paciente y es capaz de reconstruir su geometría de manera que cumpla con unos parámetros funcionales establecidos previamente. Esto permite generar un modelo 3D que puede servir de guía al cirujano para realizar la cirugía nasal.

La metodología se basa en la combinación de técnicas de mecánica de fluidos computacional (CFD) y cirugía virtual (CV) implementadas en Flowgy, y las técnicas de morfometría geométrica 3D (MG-3D). Se describen en detalle los fundamentos técnicos de cada etapa: obtención del modelo 3D, cálculo de parámetros con CFD, selección de la cavidad de referencia, transformación con MG-3D y cirugía virtual.

Los resultados muestran que es posible obtener un modelo 3D de la cavidad nasal que cumple con los parámetros funcionales deseados. Esto supone una guía objetiva para el cirujano, mejorando el diagnóstico y la cirugía nasal.

Patente solicitada con Número de solicitud: P202330235.

1. Introducción

La cirugía nasal para la corrección de obstrucciones y problemas respiratorios se ha basado tradicionalmente en la experiencia y criterio subjetivo del cirujano [1]. En los últimos años, el desarrollo de potentes técnicas computacionales ha permitido modelar y simular de forma objetiva el comportamiento de la cavidad nasal [2], abriendo nuevas posibilidades para asistir al cirujano.

Diversos estudios previos han explorado el uso de técnicas como la mecánica de fluidos computacional (CFD) [3], la cirugía virtual [4] o la morfometría geométrica 3D [5] para caracterizar y modificar modelos 3D de cavidades nasales. Sin embargo, la novedad del presente trabajo es la combinación integrada de estas tres técnicas en una metodología unificada implementada en el software Flowgy [6], que permite obtener modelos 3D personalizados de cavidades nasales con unos requisitos funcionales especificados. Estos modelos sirven de guía objetiva al cirujano sobre los cambios anatómicos necesarios para mejorar la funcionalidad respiratoria del paciente.

A continuación, se describe en detalle la metodología propuesta, explicando los fundamentos técnicos de cada

etapa. Los resultados muestran ejemplos de su aplicación a distintos modelos de cavidades nasales.

El procedimiento descrito en este artículo está bajo revisión en una solicitud de patente nacional española con Número de solicitud: P202330235. Cabe señalar que la información presentada en este artículo se adhiere a todos los protocolos relacionados con la protección de patentes y no contiene ningún material propietario o patentado.

2. Metodología

El flujo de trabajo completo implementado en Flowgy consta de las siguientes etapas:

1. Obtención del modelo 3D de la cavidad nasal patológica

El primer paso es la adquisición mediante tomografía computarizada (TC) de imágenes médicas del paciente a tratar. Estas imágenes permiten reconstruir la geometría 3D de la cavidad nasal mediante técnicas de segmentación. El modelo 3D resultante debe tener una resolución suficiente para representar fielmente las características anatómicas relevantes para el posterior análisis.

2. Procesado del modelo 3D e identificación de landmarks

Una vez obtenido el modelo 3D, el software Flowgy realiza una serie de preprocesados sobre la malla poligonal para prepararla para las siguientes etapas. Esto incluye operaciones como eliminación de ruido, suavizado, reparación de huecos e igualación topológica.

A continuación, se identifican puntos anatómicos clave (landmarks) sobre la superficie 3D, como se muestra en la Figura 1. Estos puntos son necesarios para los análisis de morfometría geométrica 3D posteriores.

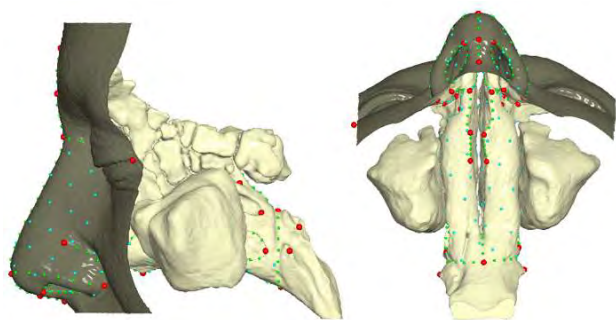


Figura 1. Configuración de puntos anatómicos. Los puntos rojos identifican posiciones anatómicas fijas, los puntos verdes cuantifican curvas y los puntos azules superficies.

3. Cálculo de parámetros funcionales y geométricos

En esta etapa se simula numéricamente el flujo de aire en el interior de la cavidad nasal mediante técnicas de dinámica de fluidos computacional (CFD). Esto permite calcular parámetros funcionales críticos como la resistencia al flujo o la simetría entre ambas fosas nasales.

Además, se calculan parámetros geométricos globales del modelo 3D como dimensiones, áreas de sección transversal, volúmenes o ángulos nasofaríngeos.

4. Selección de la cavidad nasal de referencia

A partir de los parámetros funcionales objetivos especificados por el médico para el paciente, se selecciona la cavidad nasal de referencia más adecuada de una base de datos poblacional. Esta cavidad de referencia presenta unos parámetros funcionales similares a los deseados y unos parámetros geométricos compatibles con la anatomía del paciente.

5. Transformación mediante morfometría geométrica 3D

La transformación morfométrica 3D se realiza mediante el algoritmo de Thin Plate Spline (TPS) [7], [8], que deforma el modelo patológico para adaptarlo al modelo de referencia seleccionado utilizando los puntos anatómicos (landmarks).

Así se obtienen la forma y geometría nasal que cumplan con los requisitos funcionales especificados.

6. Cirugía virtual y postproceso

El proceso de transformación morfométrica puede requerir ajustes adicionales para lograr el más alto nivel de precisión y funcionalidad del modelo. Estos ajustes se pueden realizar utilizando herramientas de cirugía virtual, permitiendo ajustes localizados en el modelo [9], [10]. Este paso asegura que el modelo no solo se adhiere a las restricciones geométricas sino también a los parámetros funcionales, mejorando así su aplicabilidad clínica.

Finalmente, se realiza un postproceso del modelo para garantizar la validez topológica y poder exportarlo para su fabricación mediante impresión 3D.

Nuestra metodología se centra en proporcionar modelos 3D altamente personalizados de las cavidades nasales,

teniendo en cuenta la variabilidad inherente en las estructuras anatómicas entre individuos. Reconocemos que esta variabilidad plantea desafíos en la estandarización de parámetros funcionales. Para superar este obstáculo, estamos trabajando en el desarrollo y la implementación de parámetros adimensionales de resistencia y simetría de flujo. Estos parámetros están siendo analizados en diferentes poblaciones y géneros para establecer un conjunto de métricas que puedan servir como un valioso punto de referencia tanto para clínicos como para investigadores.

3. Resultados

La metodología implementada en Flowgy se ha validado generando modelos 3D de cavidades nasales que cumplen unos requisitos funcionales objetivo. En primer lugar, se realizó un estudio de sensibilidad deformando el modelo 3D de un individuo enfermo sano a uno enfermo y viceversa empleando el conjunto completo de puntos de referencia medidos en cada individuo. Como se puede observar en la Figura 2, en ambos casos los modelos deformados hacia las configuraciones deseada presentan valores de resistencia bilateral y simetría del flujo similares a los modelos objetivo. Estos resultados indican que la deformación mediante TPS es una metodología válida para alcanzar determinados requisitos funcionales.

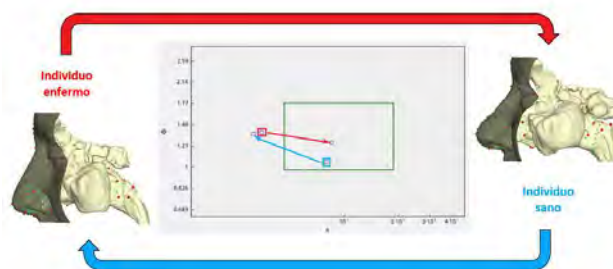


Figura 2. Transformaciones recíprocas mediante TPS entre ambos individuos. El interior del rectángulo verde representa un área del gráfico definidos por los valores de resistencia bilateral y asimetría del flujo donde existe una alta probabilidad de encontrar individuos sanos.

Tras la anterior validación, se realizó un caso problema para el cual al paciente enfermo se le ensancharon los orificios nasales estrechos empleando únicamente un subconjunto de puntos de referencia para la región de la nariz externa de un individuo sano. Tras la deformación mediante TPS, se puede observar cómo el individuo presenta valores de resistencia bilateral y simetría del flujo considerados sanos, por lo que la transformación realizada se presenta como una alternativa a elegir por el cirujano a la hora de operar al paciente (Figura 3).

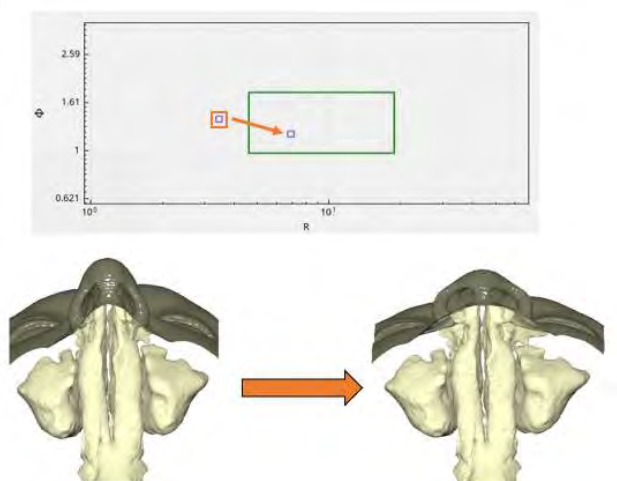


Figura 3. Transformaciones mediante TPS de un individuo enfermo hacia una configuración sana.

Por tanto, la metodología propuesta permite de forma robusta y reproducible diseñar modelos 3D personalizados de cavidades nasales que cumplan unos parámetros funcionales específicos, lo cual es de enorme utilidad para guiar al cirujano en la cirugía real.

4. Discusión

La metodología presentada en este trabajo constituye un avance respecto al estado del arte, ya que representa la primera integración de técnicas tan diversas como CFD, cirugía virtual y morfometría geométrica 3D para solucionar el problema clínico de la cirugía nasal funcional.

Frente a otros estudios previos centrados en una única técnica, la combinación planteada de las tres permite un flujo de trabajo completo que cubre todas las etapas necesarias para obtener modelos 3D funcionales de cavidades nasales.

Además, la implementación de toda la metodología en un único software (Flowgy) facilita enormemente su uso por personal clínico, superando las dificultades que supone utilizar herramientas aisladas.

En cuanto a limitaciones, la precisión final del modelo obtenido depende en gran medida de la resolución de las imágenes TC originales. Además, es necesario validar más exhaustivamente la técnica con estudios clínicos a gran escala.

5. Conclusiones

- Se ha presentado una novedosa metodología implementada en Flowgy que integra técnicas de CFD, cirugía virtual y morfometría geométrica para generar modelos 3D de cavidades nasales funcionales.
- La metodología permite de forma reproducible obtener modelos que cumplan unos requisitos funcionales especificados por el médico.
- Estos modelos sirven de guía objetiva al cirujano sobre los cambios necesarios para mejorar la funcionalidad respiratoria del paciente.

- La implementación en Flowgy facilita enormemente su aplicación clínica.

La metodología descrita constituye un desarrollo propio de los autores, estando protegida mediante una solicitud de patente nacional e internacional pendiente de concesión.

6. Financiación

Este trabajo ha sido patrocinado por el Ministerio de Ciencia e Innovación de España:

- La Subvención PID2019-105097RB-I00 a M. Burgos. está financiada por CIN/AEI/10.13039/501100011033 del Ministerio de Ciencia e Innovación de España y la Unión Europea.
- La subvención PID2020-115854GB-I00 a M. Bastir está financiada por MCIN/AEI/10.13039/501100011033 del Ministerio de Ciencia e Innovación de España y la Unión Europea.

7. Protección de la propiedad intelectual

La metodología presentada en este artículo está protegida por una solicitud de patente nacional ante la Oficina Española de Patentes y Marcas con Número de solicitud: P202330235 y título "PROCEDIMIENTO PARA LA OBTENCIÓN DE MODELOS TRIDIMENSIONALES DE CAVIDADES NASALES CON PARÁMETROS FUNCIONALES ESPECIFICADOS".

Referencias

- [1] T. L. Smith et al., «Medical therapy vs surgery for chronic rhinosinusitis: a prospective, multi-institutional study with 1-year follow-up», *Int. Forum Allergy Rhinol.*, vol. 3, n.o 1, pp. 4-9, ene. 2013, doi: 10.1002/alr.21065.
- [2] O. James Ayodele, A. Ebenezer Oluwatosin, O. Christian Taiwo, y A. Adebukola Dare, «Computational Fluid Dynamics Modeling in Respiratory Airways Obstruction: &Current Applications and Prospects&», *Int. J. Biomed. Sci. Eng.*, vol. 9, n.o 2, p. 16, 2021, doi: 10.11648/j.ijbse.20210902.12.
- [3] J. Xi y P. W. Longest, «Numerical predictions of submicrometer aerosol deposition in the nasal cavity using a novel drift flux approach», *Int. J. Heat Mass Transf.*, vol. 51, n.o 23, pp. 5562-5577, nov. 2008, doi: 10.1016/j.ijheatmasstransfer.2008.04.037.
- [4] P. Claes, M. Walters, y J. Clement, «Improved facial outcome assessment using a 3D anthropometric mask», *Int. J. Oral Maxillofac. Surg.*, vol. 41, n.o 3, pp. 324-330, mar. 2012, doi: 10.1016/j.ijom.2011.10.019.
- [5] M. V. Toro Ibacache, G. Manriquez Soto, y I. Suazo Galdames, «Morfometría Geométrica y el Estudio de las Formas Biológicas: De la Morfología Descriptiva a la Morfología Cuantitativa», *Int. J. Morphol.*, vol. 28, n.o

4, pp. 977-990, dic. 2010, doi: 10.4067/S0717-95022010000400001.

[6] «Flowgy The most affordable, fastest and efficient nasal surgery». <https://www.flowgy.com/>, consultado en octubre, 2023.

[7] W. Keller y A. Borkowski, «Thin plate spline interpolation», *J. Geod.*, vol. 93, n.o 9, pp. 1251-1269, sep. 2019, doi: 10.1007/s00190-019-01240-2.

[8] P. Gunz, P. Mitteroecker, y F. L. Bookstein, «Semilandmarks in Three Dimensions», en *Modern Morphometrics in Physical Anthropology*, D. E. Slice, Ed., en *Developments in Primatology: Progress and Prospects*. New York: Kluwer Academic Publishers-Plenum Publishers, 2005, pp. 73-98. doi: 10.1007/0-387-27614-9_3.

[9] M. A. Burgos, E. Sanmiguel-Rojas, C. del Pino, M. A. Sevilla-García, y F. Esteban-Ortega, «New CFD tools to evaluate nasal airflow», *Eur. Arch. Otorhinolaryngol.*, vol. 274, n.o 8, pp. 3121-3128, may 2017, doi: 10.1007/s00405-017-4611-y.

[10] M. A. Burgos, E. Sanmiguel-Rojas, N. Singh, y F. Esteban-Ortega, «DigBody: A new 3D modeling tool for nasal virtual surgery», *Comput. Biol. Med.*, vol. 98, pp. 118-125, ene. 2018, doi: 10.1016/j.compbiomed.2018.05.016.

Modelado computacional de la ablación de tejidos biológicos por electroporación irreversible utilizando campo eléctrico pulsado de alta intensidad

F. Estevez Labori¹, A. González Suárez^{1*}

¹ Universidad Internacional de Valencia, Valencia, España, festevezlabori@alumnos.viu.es; ana.gonzalez.s@professor.universidadviu.com *Autor de correspondencia

Resumen

La ablación por electroporación irreversible (IRE, Irreversible Electroporation) utilizando campo eléctrico pulsado de alta intensidad es un método relativamente nuevo para el tratamiento de patologías del sistema cardiovascular y el cáncer irreseccable, entre otros. Esta técnica utiliza pulsos ultracortos de alto voltaje para provocar la permeabilización de la membrana celular causando así la muerte dirigida del tejido diana. A diferencia de otras técnicas de ablación tradicionales, como la ablación por radiofrecuencia, la IRE se presenta como una alternativa de ablación no térmica reduciendo así las complicaciones ocasionadas por el calentamiento. Una patología donde la aplicación de la IRE podría ser particularmente ventajosa es en el tratamiento de la fibrilación auricular (AF, Atrial Fibrillation), arritmia más común en la práctica clínica. En el presente trabajo se realizó un estudio comparativo por modelado computacional del tratamiento de la AF con IRE por dos vías de acceso diferente: la vía endovascular (endocárdica) y la vía torácica (epicárdica). Se pudo comprobar que la vía epicárdica permite una ablación más efectiva del área de interés (grasa epicárdica) con menores daños secundarios en estructuras cercanas. En este caso, la ablación no deseada del miocardio y la muerte celular por efectos térmicos fueron mínimas. De igual manera se pudieron determinar valores preferibles de los parámetros de ablación a aplicar a fin de optimizar la efectividad y la seguridad de la técnica.

1. Introducción y objetivo

La ablación de tejidos biológicos es un procedimiento médico en el cual se aplica cierta cantidad de energía a un tejido diana para extirpar o destruir dicho tejido con fines terapéuticos o paliativos. Una técnica que está recibiendo mucha atención en los últimos años, y que se propone como una alternativa no térmica a los métodos más tradicionales como la ablación por radiofrecuencia, es la ablación por electroporación irreversible (IRE, *Irreversible Electroporation*) [1].

Cuando un tejido se somete a un campo eléctrico de pulsos ultracortos de gran intensidad se forman poros en la membrana de las células del tejido (electroporación). Si la intensidad y la duración del campo eléctrico aplicado no supera cierto umbral, alrededor de 375 V/cm para miocitos cardíacos [2], la electroporación es reversible y la célula podrá recuperar nuevamente la integridad de la membrana, sus funciones, y mantener una alta viabilidad, y esto tiene mucha utilidad en quimioterapia y terapia génica. Sin embargo, si la intensidad del campo eléctrico supera

valores alrededor de 1000 V/cm, la magnitud de la permeabilización de la membrana es tal que provoca una alteración muy significativa de la homeostasis celular de la cual la célula no se podrá recuperar. En este caso el proceso es irreversible, conduce a la muerte celular, y se denominada electroporación irreversible (IRE) [3].

Aunque la IRE ya se ha utilizado ampliamente en la resección de tumores, una patología donde la IRE podría utilizarse también es en el tratamiento de la fibrilación auricular (AF, *Atrial Fibrillation*). En la AF, ondas eléctricas desorganizadas que generalmente se originan en las raíces de las venas pulmonares, interfieren con los impulsos eléctricos ordenados originados en el nódulo sinoauricular provocando una activación irregular de los latidos del corazón. La IRE permite estripar los plexos ganglionares que se encuentran embebidos en la grasa del epicardio, y que se sospecha sean el contribuyente principal en la inducción y el mantenimiento de estos impulsos nervios indeseados [4].

Entre las varias ventajas tiene la ablación por IRE se encuentran; la selectividad tisular, la corta duración del tratamiento, márgenes de la zona de ablación bien definidos, ausencia de necrosis inducida por daño térmico y la posibilidad de monitoreo en tiempo real. Como una técnica de naturaleza básicamente no térmica, la IRE no presenta los riesgos de producir daño tisular por calentamiento excesivo. Sin embargo, el carácter relativamente nuevo esta técnica, hace que todavía no exista un consenso sobre cuáles son los mejores parámetros que garanticen la mayor efectividad y seguridad del tratamiento.

En este contexto, el objetivo de este trabajo fue el estudio mediante modelado computacional de la ablación por IRE en el tratamiento de la AF. Se realizó un estudio comparativo entre dos vías de acceso diferentes, la vía torácica y la vía intracardiaca, para identificar cuál de estas propiciaría un mejor resultado de tratamiento, en términos de efectividad y seguridad. A fin de determinar cuáles serían los parámetros de tratamiento más apropiados (potencial eléctrico aplicado, número de pulsos, etc.) se tuvieron en cuenta la variabilidad de las dimensiones anatómicas y diferentes rangos de los parámetros de tratamiento.

2. Materiales y métodos

Para el análisis del tratamiento de la AF por IRE, se estudiaron dos vías de acceso diferentes: la vía endovascular (enfoque endocárdico), Figura 1 A, en la cual el catéter se introduce por vía endovenosa, y la vía torácica (enfoque epicárdico), Figura 1 B, en la cual la zona a tratar se alcanza a través de apertura en la cavidad torácica. Para el enfoque epicárdico, se consideró que el electrodo estaría irrigando continuamente la zona a tratar con solución salina para mejorar el contacto eléctrico.

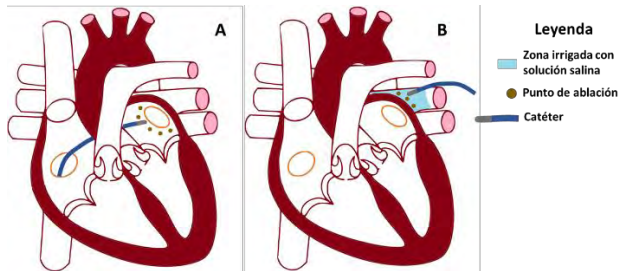


Figura 1. Tratamiento de la arritmia cardiaca AF utilizando IRE. A) enfoque endocárdico. B) enfoque epicárdico. Diagrama modificado de acuerdo con la licencia Creative Commons Attribution-Share Alike 3.0

2.1. Solución de la simulación computacional

La simulación computacional del enfoque endocárdico y epicárdico, se realizó mediante el método de elementos finitos utilizando el software COMSOL Multiphysics (COMSOL, Burlington, MA, USA)..

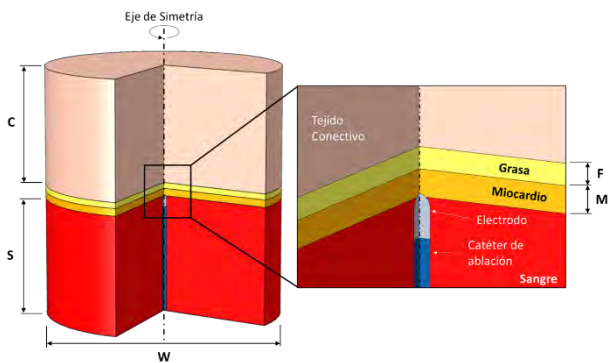


Figura 2. Modelo geométrico 2D con simetría axial del enfoque endocárdico. Dimensiones estudiadas: (C) 40 mm; (F) 1, 1.5 y 2 mm; (M) 1.5, 2, 2.5 mm (S) 40 mm; (W) 80 mm

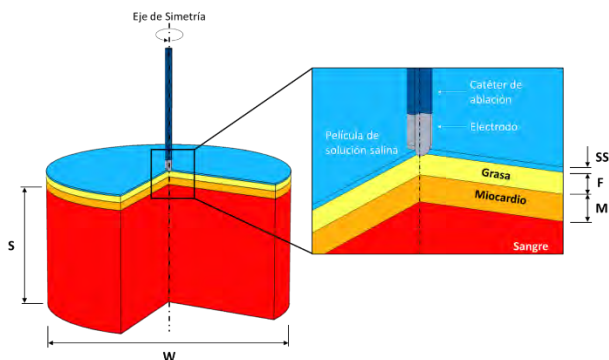


Figura 3. Modelo geométrico 2D con simetría del enfoque epicárdico. Dimensiones estudiadas: (F) 1, 1.5 y 2 mm; (M) 1.5, 2, 2.5 mm (S) 40 mm; (SS) 0.5 mm; (W) 80 mm

Se resolvieron los problemas eléctrico y térmico acoplados utilizando modelos geométricos 2D con simetría axial de dominio limitado como muestran las Figura 2 y Figura 3. Las propiedades físicas de los materiales y tejidos utilizados se basaron en estudios previos [5].

La Figura 4 muestra el perfil de cadenas de pulsos de potencial eléctrico utilizado para la modelación de la ablación por IRE. El potencial de pulso (Φ) se estudió a tres niveles diferentes: 1000, 1500 y 2000 V. La duración de cada pulso (t_p) fue de 100 μ s y la separación entre pulsos fue de 1s. Se estudiaron cadenas de 30 y 50 pulsos totales (N_p).

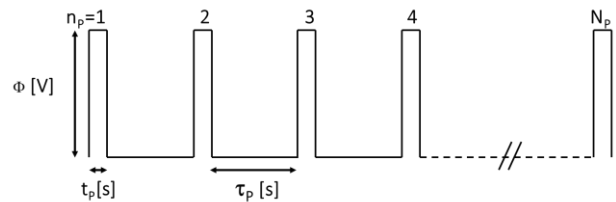


Figura 4. Estructura de las cadenas de pulsos de potencial eléctrico efector utilizados. Φ [V] es el potencial eléctrico de los pulsos, (t_p) [s] es la duración de cada pulso, (n_p) es el número de pulso, N_p es el número total de pulsos y τ_p [s] es la separación entre pulsos

Para el cálculo de la conductividad eléctrica dinámica de los tejidos en función del campo eléctrico y la temperatura $\sigma(\mathbf{E}, T)$, se utilizó la ecuación [6]:

$$\sigma(\mathbf{E}, T) = \sigma_0 \cdot (1 + A \cdot flc2hs(\mathbf{E} - \mathbf{E}_{del}, \mathbf{E}_{range})) + \alpha \cdot (T - T_0)$$

El cálculo del daño térmico en el tejido debido a efectos térmicos se basó en la relación de Arrhenius [7]:

$$\Omega(t) = \ln \left\{ \frac{c(0)}{c(t)} \right\} = \int_0^t A e^{\frac{-\Delta E_a}{RT(t)}} dt$$

Un valor de $\Omega = 4,6$ equivale a un 99% de probabilidad teórica de muerte celular, $\Omega = 1$ equivale a un 63%.

2.2. Condiciones iniciales y de frontera

La Figura 5 y Figura 6 muestra las condiciones iniciales y de frontera utilizadas para le solución de los problemas eléctrico y térmico acoplado.

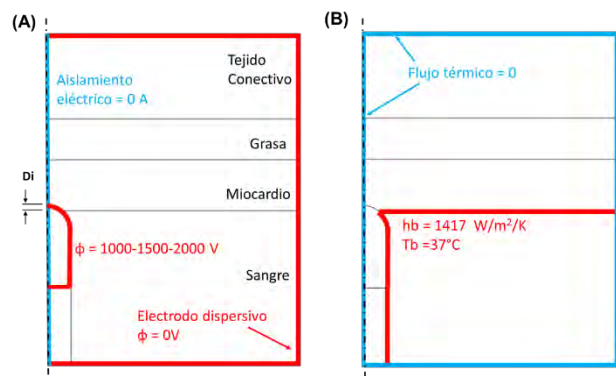


Figura 5. Condiciones de frontera para el modelo del enfoque endocárdico: A) problema eléctrico. B) problema térmico.

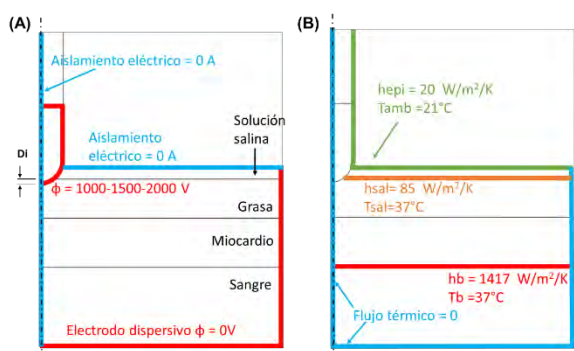


Figura 6. Condiciones de frontera para el modelo del enfoque epicárdico: A) problema eléctrico. B) problema térmico.

2.3. Evaluación de los resultados de la simulación

Se realizó análisis gráfico 2D y 3D de las superficies y contornos de temperatura, campo eléctrico e índice de muerte celular para visualizar la magnitud y distribución de estas variables en las diferentes zonas de los tejidos.

Para medir la efectividad del tratamiento, se calculó el volumen de la zona electroporada ($E \geq 1000$ V/cm) en grasa y en miocardio [4]. Para medir la seguridad del tratamiento, se calculó el volumen de la zona de tejido, tanto en grasa como en miocardio, donde se produce una mortalidad celular mayor de 63% ($\Omega \geq 1$) [7].

2.4. Análisis estadístico

El análisis comparativo entre los valores medios obtenidos se realizó mediante prueba t-pareada, mientras que la relación entre variables estudió por análisis de regresión lineal, utilizando el software estadístico SPSS 29 IBM.

3. Resultados y discusión

3.1. Efectividad de la electroporación irreversible

Para evaluar la efectividad del tratamiento, se calculó el volumen de grasa electroporada por IRE. La Figura 7 muestra la relación entre el volumen de grasa electroporada ($E \geq 1000$ V/cm) y los valores de voltaje de pulso aplicado.

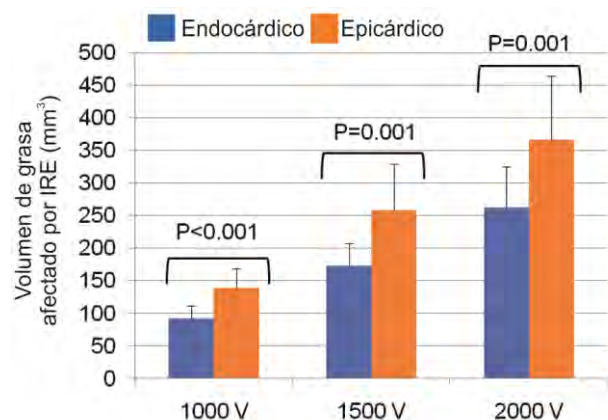


Figura 7. Volumen de grasa afectado por la IRE en el enfoque endocárdico y epicárdico para diferentes valores de voltaje.

Como se observa en la figura, independientemente del potencial aplicado, el volumen de ablación fue mayor cuando se utilizó el enfoque epicárdico ($P \leq 0.05$). Un

mayor potencial eléctrico de pulso se correlacionó con un volumen mayor de grasa electroporada ($P < 0.001$)

3.2. Seguridad: Evaluación del daño tisular por electroporación irreversible indeseada en el miocardio

La seguridad del tratamiento se evaluó calculando el volumen de zona del miocardio afectado de manera no deseada por la IRE. La Figura 8 muestra la relación entre el volumen de ablación en el miocardio y el potencial eléctrico de pulso aplicado.

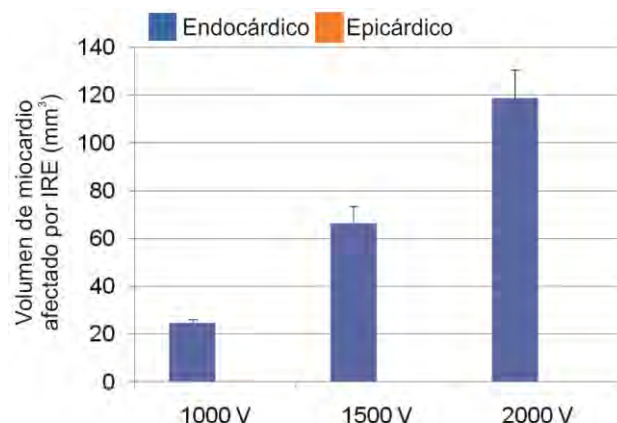


Figura 8. Volumen de miocardio afectado IRE en el enfoque endocárdico y epicárdico, para diferentes valores de voltaje.

Afectación a nivel del miocardio solo se produjo en el enfoque endocárdico, y esta aumenta a medida que se incrementa el potencial de pulso eléctrico aplicado.

3.3. Seguridad: Evaluación del daño térmico en grasa y miocardio

La evaluación del daño térmico en los tejidos se basó en el cálculo del volumen de la zona donde el índice de mortalidad $\Omega \geq 1$. La Figura 9 muestra la relación entre el volumen de la zona afectada y las características de los pulsos eléctricos aplicados (voltaje y número de pulsos).

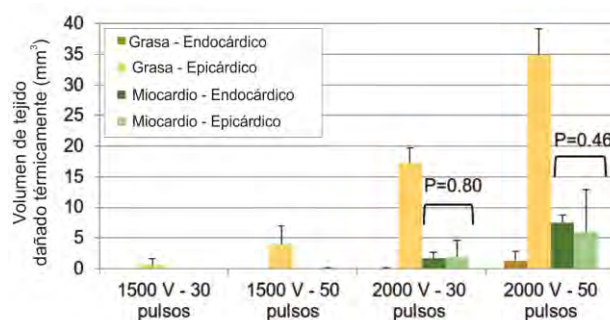


Figura 9. Volumen de tejido dañado térmicamente durante el tratamiento por IRE, para diferentes valores de voltaje y número de pulsos.

Con pulsos de 1000 V, no se observó daño térmico en ningún caso. Si se utilizan pulsos de 1500 V, tampoco se produjo daño térmico en el enfoque endocárdico, pero en el enfoque epicárdico, ya se comienza a observar algún daño térmico a nivel de la grasa, tanto a 30 como a 50 pulsos. Con pulsos de 2000 V, el daño térmico en la grasa fue más marcado en el enfoque epicárdico, particularmente

cuando se aplican cadenas de 50 pulsos. A nivel del miocardio se produce algún daño tanto usando enfoque epicárdico como endocárdico, pero las diferencias no fueron estadísticamente significativas ($P > 0.05$).

3.4. Distribución de temperatura y del daño térmico en los tejidos al final del tratamiento

La Figura 10 muestra los resultados del análisis de la distribución de temperatura y daño térmico en los tejidos a diferentes valores de voltaje y número de pulsos.

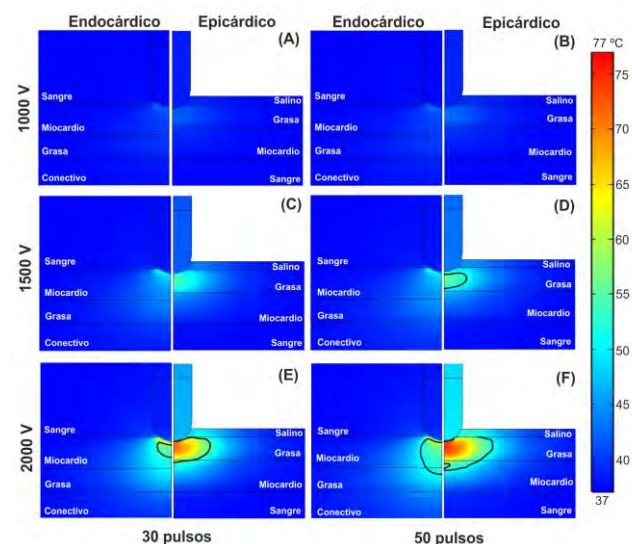


Figura 10. Distribución de temperatura obtenida para condiciones anatómicas de espesor de grasa 1.5 mm y de miocardio de 2 mm. Se utilizaron cadenas de 30 y 50 pulsos a diferentes valores de voltaje. La línea continua negra representa el contorno de daño térmico $\Omega = 1$.

Con pulsos de 1000 V, la zona de calentamiento se localiza fundamentalmente en las áreas cercanas al electrodo y se mantiene a niveles inferiores a 40-45 °C, tanto para el enfoque endocárdico como epicárdico, independiente del número de pulsos aplicado (paneles A, B respectivamente). No se observa muerte celular significativa ($\Omega \geq 1$).

Utilizando pulsos de 1500 V, ya se comienza a observar una zona de calentamiento más alejada del electrodo. Este calentamiento es mayor en la grasa, y cuando se utiliza el enfoque epicárdico y 50 pulsos (panel D). Se observa una zona de muerte celular limitada a la grasa (panel D).

Con pulsos de 2000 V, Figura 10 paneles E y F, se observan zonas de calentamiento significativo y áreas de daño térmico importantes, independientemente del número de pulsos eléctricos aplicados, tanto en el enfoque epicárdico como endocárdico. Sin embargo, el daño ocurre en sitios diferentes. En el enfoque endocárdico, el daño se observa principalmente en el miocardio, mientras que en el enfoque epicárdico, es en la grasa, con mínima afectación del miocardio. El número de pulsos también afecta al impacto (paneles E y F). La condición más crítica se observa con la aplicación de 50 pulsos utilizando el enfoque epicárdico.

4. Conclusiones

La vía de acceso epicárdica fue más efectiva que la endocárdica ya que se obtienen volúmenes de campo

eléctrico mayores en la grasa con menor afectación de tejidos aledaños. Los parámetros más apropiados para el tratamiento fueron la vía de acceso del enfoque epicárdico, y utilizar cadenas de 50 pulsos de 1000 V.

Agradecimientos

Quiero agradecer primero que nada a la Dra. Ana González Suárez por su ayuda extraordinaria durante la preparación de mi TFM y a la VIU por su excelente programa y por seleccionar TFM y financiar mi asistencia a este Congreso.

Proyecto financiado por el Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación, Fondo Europeo de Desarrollo Regional (MCIN/AEI/10.13039/501100011033/FEDER,UE) con la Referencia PID2022-136273OA-C33.

Referencias

- [1] A. González-Suárez, R. M. Irastorza, S. Deane, B. O'Brien, M. O'Halloran y A. Elahi, «Full torso and limited-domain computer models for epicardial pulsed electric field ablation,» *Computer Methods and Programs in Biomedicine*, vol. 221, pp. 2-10, 2022.
- [2] I. Kaminska, M. Kotulska, A. Stecka, J. Saczko, M. D. Zalesinska, T. Wysocka, A. Choromanska, N. Skolucka, R. Nowicki, J. Marczak y J. Kulbacka, «Electroporation-induced changes in normal immature rat myoblasts (H9C2),» *Gen. Physiol. Biophys.*, vol. 31, n° 1, pp. 19-25, 2012.
- [3] M. C. Sheehan y G. Srimathveeravalli, «Pulsed electric fields,» de *Principles and Technologies for Electromagnetic Energy Based Therapies*, Oxford, Elsevier Academic Press, 2022, pp. 71-106.
- [4] S. Avazzadeh, S. McBride, B. O'Brien, K. Coffey, A. Elahi, M. O'Halloran, A. Soo y L. R. Quinlan, «Ganglionated Plexi Ablation for the Treatment of Atrial Fibrillation,» *J Clin Med*, vol. 9, n° 10, p. 3081, 2020.
- [5] P. Hasgall, F. Di Gennaro, C. Baumgartner, E. Neufeld, B. Lloyd, M. Gosselin, D. Payne, A. Klingenböck y N. Kuster, «It Is Database for Thermal and Electromagnetic Parameters of Biological Tissues; Version 4.1.» 22 February 2022. [En línea]. Available: <https://itis.swiss/virtual-population/tissue-properties/overview/>. [Último acceso: 11 November 2022].
- [6] Y. Zhao, S. Zheng, N. Beitel-White, H. Liu, C. Yao y R. Davalos, «Development of a Multi-Pulse Conductivity Model for Liver Tissue Treated With Pulsed Electric Fields,» *Front. Bioeng. Biotechnol.*, vol. 396, p. 8, 2020.
- [7] A. González-Suárez, J. Pérez, R. M. Irastorza, A. D'Avila y E. Berjano, «Computer modeling of radiofrequency cardiac ablation: 30 years of bioengineering research,» *Computer Methods and Programs in Biomedicine*, vol. 214, p. 106546, 2022.

Computer modeling of radiofrequency thermocoagulation (RF-TC) using the recording intracerebral electrodes implanted for stereo electroencephalography (SEEG) monitoring

S. Collavini^{1,2}, J.J. Pérez³, E. Berjano³, M. Fernández-Corazza⁴, S. Oddo², R. Irastorza^{1,5,6}

¹ Institute of Engineering and Agronomy, National University Arturo Jauretche, Buenos Aires, Argentina, santiagocollavini@gmail.com

² Neurosciences and Complex Systems Unit (EnyS), CONICET, Hosp. “El Cruce N. Kirchner”, National University A. Jauretche (UNAJ), Buenos Aires, Argentina, silviaoddo4@gmail.com

³ BioMIT, Departamento de Ingeniería Electrónica, Universitat Politècnica de València, València, Spain, jjperez@eln.upv.es, eberjano@eln.upv.es

⁴ Research Institute of Electronics, Control and Signal Processing (LEICI), National University of La Plata-CONICET, La Plata, Argentina, marianof.corazza@ing.unlp.edu.ar

⁵ Instituto de Física de Líquidos y Sistemas Biológicos (IFLySiB CONICET La Plata), La Plata, Argentina, rirastorza@iflysib.unlp.edu.ar

⁶ Grupo de Materiales Granulares, Departamento de Ingeniería Mecánica, Facultad Regional La Plata, Universidad Tecnológica Nacional, La Plata, Argentina

Resumen

Our objective was to use computational modeling to provide a full and comprehensible physical description of the electrical and thermal behavior during electroencephalography-guided radiofrequency thermocoagulation (SEEG-RF-TC), a minimally invasive procedure to destroy epileptic foci in the brain. We used the Finite Element Method to solve a coupled electrical-thermal problem and obtain the temperature distributions in the tissue for a constant power thermocoagulation. We found that 5 W provokes an impedance roll-off (abrupt increase) after approximately 10 s, and that this is related to temperatures of 100 °C in the surroundings of the contacts. Due to the internal resistance associated with the wires of the electrode, self-heating provoked a temperature increase of up to 3.3 °C throughout the device. Our findings should alert clinicians who are guided by impedance during SEEG-RF-TC regarding overheating associated with roll-off. The self-heating (approximately 3 °C) occurring along the electrode should also be taken into account as it could provoke thermal side-effects.

1. Introduction

Epilepsy is a neurological pathology characterized by recurrent seizures resulting from abnormal electrical activity in groups of neuronal populations. To keep seizures under control, the most common treatment consists in the administration of drugs. Unfortunately, 30% of these patients do not respond to pharmacological treatment. In 50% of these refractory cases, tissue resection results in the best option [1]. For cortical resection to be effective, the epileptogenic zone (EZ) must be correctly identified [2], hence the importance of locating it correctly. Brain mapping procedures are used to correctly identify the EZ. Recording of brain activity by intracranial electroencephalography (iEEG) is the preferred method for identifying the location and extent

of the EZ, as it records electrical signals directly from the cortical surface (electro-corticography, ECoG) or over deep brain regions (stereo-electroencephalography, SEEG).

Stereo electroencephalography (SEEG) is a minimally invasive mapping technique aimed at localizing EZ in the brain by inserting very thin electrodes (< 1.3 mm diameter) into deep places in the brain (see Fig. 1A). Once the electrodes have located the area suspected of being the EZ, it is possible to apply radiofrequency (RF) power between them to achieve destruction of said area by thermal necrosis. This technique is called SEEG-guided RF thermocoagulation (SEEG-guided RF-TC) [3]. It is essentially a bipolar RF ablation technique that is clinically performed by applying constant power (values around 5 W) for 10–40 s [4]. The terms TC and ablation are synonymous in this context. Although it is not common in the field of RF ablation, in the context of SEEG-guided RF-TC, each of the metal electrodes that make up the measurement probe is called contact, and said probe is called an electrode. The application of RF causes a rise in temperature in the tissue adjacent to the energized contacts, creating ablation volume of a more or less ellipsoidal shape, and although the possibility of increasing the size of the ablation zone by applying RF between separate electrode contacts has been studied [5], it is usual to apply power between two adjacent contacts of the same electrode [6].

There are different recommendations regarding the power-duration protocol to be used. While some groups use fixed parameters (e.g. 5 W for 10–30 s), others have proposed progressively increasing the power until observing how the current and voltage drop spontaneously [6], which from an electrical point of view corresponds with a drastic increase of impedance (called

impedance roll-off). Some clinical groups suggest that this technique avoids unintentionally creating excessively large ablation zones, and that it limits tissue temperature to values between 78 and 82 °C [6]. However, our knowledge about RF ablation does not suggest that, but rather that the roll-off occurs when a zone of tissue completely desiccated surrounds one of the contacts, preventing the passage of more electrical current, and even more, this zone of desiccation is associated with tissue temperatures close to 100 °C [7]. This is corroborated by the fact that the patient or neurosurgeon can "hear the coagulation" before the ablation zone reaches its maximum size [6]. What is actually heard is an audible pop, which corresponds to the creation of steam within the tissue. Our objective was to use computational modeling to provide a full and comprehensible physical description of the electrical and thermal behavior during SEEG-guided RF-TC. To our knowledge, this is the first in-silico study of this technique.

2. Materials and methods

2.1. Description of the model

A two-dimensional was built assuming a 0.8 mm diameter Microdeep electrode (DIXI Medical, Chaudfontaine, France) completely surrounded by homogeneous tissue. Hence, we assumed rotational symmetry around the electrode axis (see Fig. 1B). Although this electrode model has several contacts, we only considered two contiguous contacts, with length of 2 mm and separation of 1.5 mm. The SEEG electrodes are made of flexible plastic and have an internal lumen (0.15 mm diameter), through which a guide-wire or other deeper mapping electrode can be inserted. The contacts inside the brain are connected to external contacts by ultra-fine connecting wires. We previously measured the internal resistance of these wires R_{int} (both at DC and 500 kHz) of several commercial electrodes, and we found values of 100–200 Ω , which is by no means negligible compared to the tissue resistance R_{tissue} . This has a direct impact on the real power applied to the neural tissue (P_{tissue}), as a function of the power delivered by the device (P_{set}), which can be formulated as:

$$P_{tissue} = P_{set} \frac{R_{tissue}}{R_{tissue} + 2 \cdot R_{int}} \quad (1)$$

The effectively applied power is hence reduced depending on the relationship between R_{tissue} and $2 \cdot R_{int}$. We simulated programmed power (P_{set}) of 5 W during 30 s or until impedance roll-off (stated as a drastic impedance raise above the initial value), whichever happens sooner. We also included a linear ramp during the first 5 s. Hence, the voltage and current were modified each time-step to keep power constant at the value of P_{set} . This value was separately applied to tissue and internal wires. While the value of P_{tissue} was set on the two contacts to be directly delivered to the neural tissue, the dissipated power in $2 \cdot R_{int}$ was applied uniformly along a line (see Fig. 1) to mimic the power loss associated with the resistance of the internal wires and to assess the self-

heating of the electrode. The contacts had 50 μm thickness and the simulations considered that the internal lumen was empty.

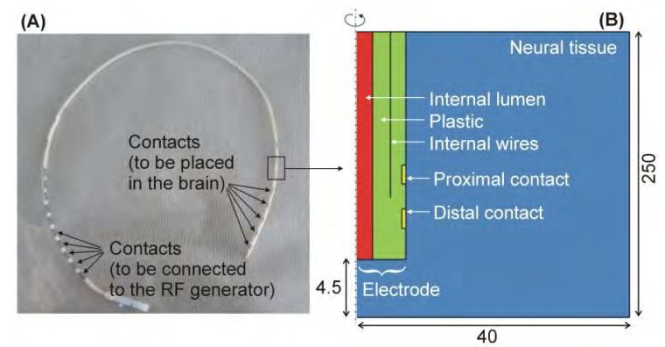


Figura 1.A: SEEG electrode. B: Model geometry (out of scale). Outer dimensions in mm.

2.2. Governing equations

The Ansys Multiphysics and FEniCSx programs were used to solve the electro-thermal coupled problem using the Finite Element Method. Laplace's equation is used to obtain the electric field distribution and current density, and from them, the power density deposited in the tissue (W/m^3). This value was coupled with the Bioheat Equation to calculate the temperature distribution in tissue and in the electrode [8]. The blood perfusion term in the tissue was set to zero once the temperature reached 55 °C, since this value was considered to be lethal threshold [8]. Initial and outer boundaries temperature was the body temperature (37 °C).

2.3. Material properties

The following properties were used for the materials of the model: electrical conductivity (σ), thermal conductivity (k), density (ρ) and specific heat (c). The contacts were assumed to be platinum-iridium with $\sigma = 4.6 \times 10^6 \text{ S/m}$, $k = 71 \text{ W/m}\cdot\text{K}$, $\rho = 21,500 \text{ kg/m}^3$ and $c = 132 \text{ J/kg}\cdot\text{K}$ [9]. The plastic part of the electrode was assumed to be polyurethane with $\sigma = 10^{-5} \text{ S/m}$, $k = 23 \text{ W/m}\cdot\text{K}$, $\rho = 1440 \text{ kg/m}^3$ and $c = 1050 \text{ J/kg}\cdot\text{K}$ [9]. The neural tissue was assumed to be gray matter with $\sigma = 0.152 \text{ S/m}$ (at 500 kHz), $k = 0.55 \text{ W/m}\cdot\text{K}$, $\rho = 1045 \text{ kg/m}^3$, $c = 3696 \text{ J/kg}\cdot\text{K}$ and blood perfusion rate of 0.0133 s^{-1} [10]. Electrical conductivity in the tissue is increased by +1.5%/°C up to 100 °C. Above 100 °C, the σ value is dropped two orders of magnitude, and it remains constant (even if the temperature goes back to lower values than 100 °C). Thermal conductivity was assumed to be constant. In order to model the latent heat associated with the phase change (liquid-gas), the product $\rho \times c$ was increased $\times 400$ between 99 and 100 °C. Above 100 °C, it returned to the initial value.

3. Results

Figure 2 shows the impedance progress during RF ablation. We found that R_{tissue} initially decreased from 1326 to 765 Ω , reaching a minimum at ~ 5.7 s, followed

by a gradual increase until reaching 1000 Ω at 10.3 s, and a drastic increase from that time on (impedance roll-off).

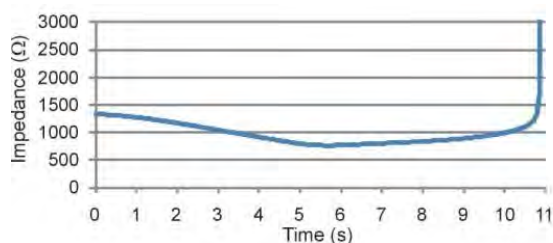


Figura 2. Impedance progress during RF ablation.

Figure 3 shows the progress of the temperature distributions after the impedance roll-off. It is noticeable a mark edge effect which provokes a overheating at tissue next to the contact edges. The ablation zone was progressively growing from these two hot spots. The impedance roll-off approximately coincided with the 100 °C zone covering one of the two electrodes (proximal contact) at 10.8 s, which coincides with the instant when the roll-off occurs. The coagulation zone size (assessed using the 55 °C isoline, shown in Fig. 3) was 3.8 mm of minor diameter and 10.0 of major diameter.

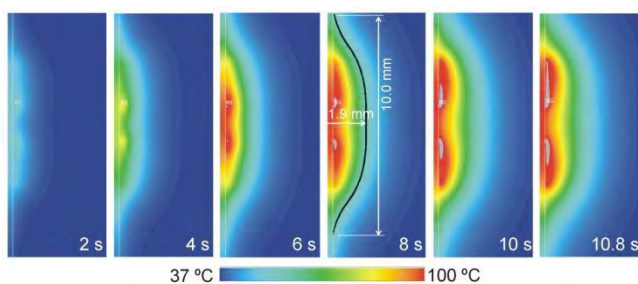


Figura 3. Progress of the temperature distribution during RF ablation (the gray area corresponds to temperatures above 100 °C).

The self-heating in the electrode was minimal but not zero. Figure 4 shows the progress of the temperature at a point along the line and sufficiently far from the contacts, which allows to illustrate the self-heating due to the internally dissipated power into the electrode. The maximum value was 40.3 °C, reached around second 9.

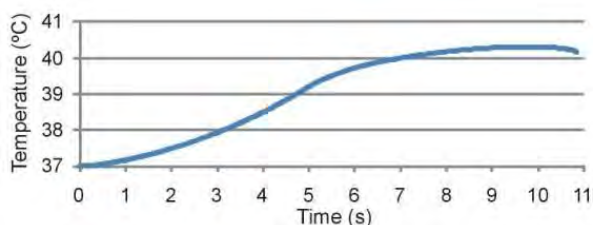


Figura 4. Progress of the temperature into the electrode at a point far from the contacts, thus reflecting self-heating.

4. Discussion

This computational study explores the electro-thermal behavior of SEEG electrodes when their contacts are energized with RF to create thermal lesions and destroy the epileptic focus. Two reasons led us to conduct this

study. First, to describe in mathematical terms the relationship between the impedance evolution and the temperature distribution in the tissue. This is relevant because clinically the impedance roll-off is associated with a self-control mechanism that prevents the size of the coagulation zone from growing above a limit [6], but at the same time it is ignored that the maximum temperature possibly exceeds 100 °C. In this regard, our results suggest that the abrupt increase in impedance is associated with the cessation of current due to the fact that one of the contacts has been completely surrounded by an area of dehydrated tissue, and this is the result of temperatures associated with a phase change from liquid to gas. Clinicians should keep this in mind when suggesting impedance-guided protocols.

And second, we are also interested in obtaining an estimate of the heating of the catheter caused by the internal resistance of the wires that connect the external contacts with the internal ones. This issue has also been ignored in the clinical literature, and should not be, since the resistance of some models of SEEG electrodes have a not insignificant internal resistance compared to tissue resistance (e.g. 400 Ω compared to 700–1000 Ω). Knowing this fact is important because the power programmed by the user in the RF generator is really distributed between the internal wires and the tissue, and might have hence an impact on predicting the coagulation zone size, and even pose a risk due to self-heating of the electrode that may be in contact with subcutaneous tissue on its path to the brain.

The following limitations have to be pointed out: 1) the inner technical details of the electrode were ignored (due to not knowing exactly the constituent materials or their geometry) and consequently the inner wires were modeled as a thin metal cylinder extending to the midpoint of the contacts; 2) only homogeneous tissue was considered, ignoring whether the pathological tissue (object of ablation) could show different electrical or thermal properties; and 3) a perfect contact (no electrical or thermal resistance) was assumed between the metallic contacts and the tissue, favoring the deposition of RF power. We do not know at this time if due to the extremely small size of the electrodes, the gap between contact and tissue caused by the mechanical insertion of the introducer needle could have any impact in terms of dissipating RF power and/or heat.

5. Conclusions

This work presents the first computational model on SEEG-guided RF-TC using two contiguous contacts, and taking the resistance of the internal wires into account. The results describe the close relation between the impedance progress, including the abrupt increase (roll-off) and the temperature distributions in the tissue, specially the appearance of overheating points (100 °C). The findings also provide an estimate of the self-heating provoked by the resistance of the internal wires connecting the contacts. These findings provide an explanation based on biophysical laws about the electrical and thermal behavior of tissue during this procedure.

Future work will be developed to validate our computer results with experiments based on ex vivo tissue.

Funding information

Spanish Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación, Fondo Europeo de Desarrollo Regional (Grant PID2022-136273OB-C31 funded by MCIN / AEI/ 10.13039/501100011033 / FEDER, UE)

References

- [1] J. Engel Jr. Research on the human brain in an epilepsy surgery setting. *Epilepsy Research*, vol. 32, no. 1-2, pp. 1-11, 1998.
- [2] H.O. Lüders, I. Najm, D. Nair, P. Widdess-Walsh, W. Bingman. The epileptogenic zone: general principles. *Epileptic Disorders*, vol. 8, no. 2, pp. 1-9, 2006.
- [3] P. Bourdillon, B. Devaux, A.S. Job-Chapron, J. Isnard. SEEG-guided radiofrequency thermocoagulation. *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 48, no. 1, pp. 59-64, Feb. 2018. doi: 10.1016/j.neucli.2017.11.011.
- [4] M. Guenot, J. Isnard. La thermocoagulation multiple de foyers épileptogènes guidée par la SEEG (thermo-SEEG). [Multiple SEEG-guided RF-thermolesions of epileptogenic foci], *Neurochirurgie*, vol. 54, no. 3, pp. 441-447, May 2008.
- [5] M.D. Staudt, S. Maturu, J.P. Miller. Radiofrequency Energy and Electrode Proximity Influences Stereoelectroencephalography-Guided Radiofrequency Thermocoagulation Lesion Size: An In Vitro Study with Clinical Correlation. *Operative Neurosurgery*, vol. 15, no. 4, pp. 461-469, Oct. 2018. doi: 10.1093/ons/oxx291.
- [6] P. Bourdillon, J. Isnard, H. Catenoix, A. Montavont, S. Rheims, P. Ryvlin, K. Ostrowsky-Coste, F. Manguiere, M. Guénot. Stereo-electro-encephalography-Guided Radiofrequency Thermocoagulation: From In Vitro and In Vivo Data to Technical Guidelines. *World Neurosurgery*, vol. 94, pp. 73-79, Oct. 2016. doi: 10.1016/j.wneu.2016.06.095.
- [7] M. Trujillo, J. Alba, E. Berjano. Relationship between roll-off occurrence and spatial distribution of dehydrated tissue during RF ablation with cooled electrodes. *International Journal of Hyperthermia*, vol. 28, no. 1, pp. 62-68, 2012. doi: 10.3109/02656736.2011.631076.
- [8] E.J. Berjano. Theoretical modeling for radiofrequency ablation: state-of-the-art and challenges for the future. *Biomedical Engineering Online*, vol. 5, p. 24, Apr. 2006. doi: 10.1186/1475-925X-5-24.
- [9] J.J. Pérez, E. Ewertowska, E. Berjano. Computer Modeling for Radiofrequency Bipolar Ablation Inside Ducts and Vessels: Relation Between Pullback Speed and Impedance Progress. *Lasers in Surgery and Medicine*, vol. 52, no. 9, pp. 897-906, Nov. 2020. doi: 10.1002/lsm.23230.
- [10] P.A. Hasgall et al. IT'IS Database for thermal and electromagnetic parameters of biological tissues. Version 4.1, Feb. 22, 2022, DOI: 10.13099/VIP21000-04-1.itis.swiss/database.

Validación de un algoritmo para el cálculo automático de la distancia inter lesión en la ablación por catéter de radiofrecuencia de Fibrilación Auricular

F. Setién Dodero¹, F. Fanjul Vélez¹, J. L. Arce Diego¹

¹ Grupo de Ingeniería Biomédica, Departamento de Tecnología Electrónica e Ingeniería de Sistemas y Automática (TEISA), Universidad de Cantabria, Av. de los Castros 46, 39005 Santander, España

fanjul@unican.es, arcedj@unican.es

Resumen

La Fibrilación Auricular (FA) es un trastorno del ritmo cardíaco caracterizado por contracciones auriculares rápidas e irregulares, que puede aumentar el riesgo de accidente cerebrovascular (ACV) y disminuir la calidad de vida de los pacientes. Una de las técnicas principales para tratar la FA es la ablación con catéter de RF, que implica aislar eléctricamente las venas pulmonares del resto de la aurícula, a base de lesiones puntuales rodeando las venas. Todavía existe discusión en la comunidad sobre cuál es la distancia entre lesiones óptima para mejorar los resultados a largo plazo de la ablación de FA. Se ha desarrollado una herramienta en Python que, partiendo de los datos del procedimiento de ablación de FA, encuentra la secuencia óptima de ablaciones que rodea a las venas pulmonares y de esta forma puede calcular la distancia entre toda la secuencia de lesiones. El algoritmo automatizado demostró ser efectivo en la mayoría de los casos y en casi la totalidad de los casos de forma semiautomática. El trabajo deja disponible una herramienta para la comunidad que puede ayudar a la hora de optimizar la ablación de FA. En el futuro se podría mejorar el algoritmo para que fuera 100% automático, aunque ahora mismo ya es de utilidad y existen varios estudios clínicos que están en marcha utilizando esta herramienta.

1. Introducción

1.1. La Fibrilación auricular

La fibrilación auricular (FA) es un trastorno del ritmo cardíaco caracterizado por contracciones rápidas e irregulares. En condiciones normales, el corazón late de manera regular y coordinada, con un patrón de contracciones rítmicas que permiten que la sangre circule eficientemente por todo el cuerpo. Sin embargo, en la FA, las señales eléctricas que controlan el ritmo cardíaco se vuelven caóticas, lo que resulta en una actividad auricular desorganizada y rápida. Durante un episodio de FA, las aurículas no se contraen de manera efectiva, lo que puede llevar a la formación de coágulos sanguíneos en el corazón. Estos coágulos pueden desprenderse y llegar a órganos importantes como el cerebro, donde pueden causar un accidente cerebrovascular (ACV) o ictus.

La FA afecta a un 4,4% de la población mayor de 40 años y a un 8,5% de la población mayor de 60 años, multiplica entre 1,5 y 3,5 el riesgo de muerte y es la causa principal de entre un 20%-30% de los ICTUS. Existe tratamiento farmacológico, pero este fracasa hasta en un 80% de los casos, en consecuencia, para evitar accidentes vasculares

se anti-coagula a los pacientes, lo que a su vez resulta complicado y no está exento de riesgos. Aunque todavía se desconocen las causas fisiológicas exactas que la desencadenan, se conocen dos principios sobre la enfermedad:

- Cuando la aurícula entra en FA pierde contractilidad, lo que con el tiempo produce dilatación auricular, esta dilatación modifica las propiedades eléctricas del tejido cardíaco favoreciendo que la FA se mantenga[1]
- Haïssaguerre [2] descubrió que con frecuencia los disparadores originales de la FA se encuentran en los ostiums de las venas pulmonares y eso dio inicio a la ablación con catéter de la FA. Se sabe también que hay otros focos extra pulmonares, aunque de menor relevancia [3]

En una situación normal, el ritmo del corazón lo marca el nodo sinusal, que está compuesto de unas células especializadas para tal propósito. Después la actividad eléctrica se propaga por todo el corazón proporcionando un latido sincronizado. En ocasiones, puede haber células localizadas en cualquier parte del miocardio, que se exciten y produzcan un latido de forma similar al propio nodo sinusal. En los ostiums de las venas pulmonares se produce una transición de tejido entre el cardíaco y el venoso, y es precisamente allí donde es frecuente que se generen estos latidos ectópicos que pueden arrancar un episodio de FA.

1.2. Ablación con catéter

El objetivo actual del tratamiento no farmacológico consiste en el aislamiento eléctrico de las venas pulmonares del resto de la aurícula. Existen diferentes técnicas para realizar este aislamiento y la ablación con catéter de radiofrecuencia (RF) es una de las más habituales. También se puede realizar con balones de crioablación, balones láser, de forma quirúrgica o con catéteres de electroporación. En la ablación por catéter se introduce un catéter de poliuretano y mallado de acero en la vena femoral y se accede hasta al corazón. Se necesita realizar punción transeptal, para poder pasar del lado derecho del corazón al lado izquierdo. Los catéteres se pueden deflectar para moverse libremente por la cavidad, la punta de 3,5mm de platino permite aplicar energía de radiofrecuencia sobre el tejido. La ablación se produce de forma térmica, cuando

la corriente de radiofrecuencia atraviesa el tejido cardiaco, se calienta produciéndose muerte celular cuando el tejido supera los 50°C de temperatura [4]. Es necesario realizar varias aplicaciones de forma precisa y circular para aislar la vena pulmonar del resto de la aurícula.

1.3. Navegadores

El movimiento del catéter dentro del corazón del paciente tradicionalmente se realizaba apoyado por fluoroscopia, una técnica que permite ver en tiempo real la posición del catéter utilizando rayos-X. La principal desventaja de esta técnica, más allá de utilizar radiaciones ionizantes, es que solo permite visualizar una imagen en dos dimensiones.

La ablación de los ostiums de las venas pulmonares en la aurícula izquierda requiere de una visualización más precisa. Para este propósito y otros se desarrollaron los Navegadores electro-anatómicos (Figura 1). En la punta de los catéteres se sitúa un sensor de posicionamiento magnético, que permite tanto la visualización del catéter como la reconstrucción de mapas del interior cardiaco con información eléctrica del miocardio. Durante todas las aplicaciones de RF se graban la posición, la fuerza, temperatura y demás parámetros de la ablación.



Figura 1: Mapa de una aurícula izquierda con el navegador Carto3 y las lesiones de radiofrecuencia rodeando a las venas pulmonares.

2. Antecedentes y Motivación

El Haddad M, Taghji P, Philips T, et al. [5] demostraron que las lesiones en la ablación circunferencial de la FA debían ser contiguas para lograr un aislamiento duradero, Mattias Duytschaever [6] propuso y validó el protocolo CLOSE como método para lograr la ablación efectiva y duradera de las venas pulmonares. Esta estrategia se ha visto confirmada en 17 centros en el estudio VISTAX [7]. En este estudio se sacaron del protocolo aquellos pacientes que no cumplían el criterio de >6mm de contigüidad entre lesiones próximas. Desde entonces ha habido poca discusión sobre el tamaño adecuado de la distancia inter-lesión y otros parámetros de la ablación. Hay muy pocos estudios que hayan analizado si una distancia inter-lesión de >5mm puede ser más adecuada que los 6mm del protocolo CLOSE. El procedimiento de evaluación de la distancia inter-lesión automático y accesible para toda la comunidad permitirá analizar grupos grandes de pacientes y posiblemente dar luz a esta cuestión.

2.1. Definición del problema

Aunque la medida de la distancia entre dos lesiones es una operación trivial, no lo es tanto determinar cuál es la siguiente lesión con la que se debe medir la distancia. En una ablación circunferencial el electrofisiólogo trata de realizar una secuencia ordenada de lesiones, pero lamentablemente, el procedimiento es complejo, y muy

pocas veces sucede que todas las ablaciones sigan un orden. Además, en muchas ocasiones se producen zonas donde se repiten las ablaciones por falta de estabilidad o el camino se desvía de la trayectoria esperada. Para poder medir la distancia interlesión necesitamos determinar una secuencia de lesiones contiguas que rodea la vena con saltos tales que se minimice la distancia inter lesión (Figura 2). Lo que implica ignorar aquellas lesiones fuera de la secuencia de distancias mínimas.



Figura 2: Secuencia de lesiones buscada, los números indican el orden en que fueron realizadas las quemaduras.

3. Objetivos

El objetivo de este trabajo es validar un sistema automático de medida de la distancia inter-lesión. Este sistema requiere de un algoritmo que encuentre la secuencia de ablaciones contiguas de forma circular y permitirá estudiar el efecto de la distancia inter-lesión en el éxito de la ablación de FA.

4. Métodos

A continuación se describe la metodología seguida en el trabajo, desde la selección de pacientes hasta la descripción de la herramienta implementada.

4.1. Selección de pacientes y ablación

Se utilizaron datos provenientes de dos Hospitales diferentes: en el primer Hospital se analizaron 227 ablaciones de FA tanto paroxística como persistente. Los procedimientos se realizaron según protocolo CLOSE, ablaciones circunferenciales con una distancia interlesión objetivo de 6mm, en unos pocos pacientes se realizaron ablaciones adicionales del istmo cavo-tricúspide y de la vena cava superior; en el segundo Hospital se analizaron 450 ablaciones de FA tanto paroxística como persistente. En este grupo de ablaciones se juntaron diversas técnicas de ablación:

- Protocolo CLOSE con una distancia interlesión objetivo de 6mm.
- Aislamiento circunferencial de las venas con la técnica de *dragging* (desplazamiento continuo del catéter). En este caso se etiqueta una lesión nueva cada vez que el desplazamiento supera los 4mm.
- Líneas adicionales, en función del paciente se añade a la ablación circunferencial una o varias líneas como: línea de techo, línea anterior (entre la válvula mitral y la vena superior derecha) y línea de box posterior (entre las venas inferiores)

Se extrajeron los datos del Navegador Carto3, que proporciona la posición y el valor de la lesión en cada punto en un fichero separado por comas (csv).

4.2. Herramienta construida

Para medir la distancia interlesión se realizó un software en Python con las conocidas librerías de computación científica y aprendizaje automático: Pandas, Numpy, Sklearn, Scipy y Matplotlib. Para crear una interfaz semi automática se utilizó la librería Tkinter. La herramienta construida ejecutaba la siguiente secuencia de pasos:

- Eliminar lesiones sobrantes (opcional o semi-automática), se pueden eliminar aquellas lesiones que vayan a dificultar los siguientes pasos.
- Agrupación (automática o semi-automática), se separan las dos coronas y las posibles líneas.
- Búsqueda de la secuencia (automática o semi-automática)
- Cálculo de la distancia interlesión (automática)

4.3. Agrupación de las coronas

La ablación de FA como mínimo consiste en realizar dos coronas, una izquierda y otra derecha, además de otras posibles líneas. El primer paso es separar las dos coronas y las posibles líneas. Para esta labor utilizamos el algoritmo Kmeans, que en el caso de dos coronas funciona correctamente un 100% de las veces. Cuando existen líneas adicionales no siempre es así dificultando los pasos posteriores (Figura 3).

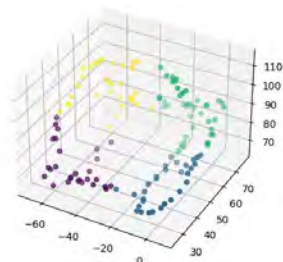


Figura 3: Kmeans no es capaz de diferenciar las líneas de las coronas

4.4. Buscando la secuencia

4.4.1 El camino más corto

El problema de la búsqueda del camino más corto entre dos puntos es un problema muy estudiado en teoría de grafos y existen varios algoritmos que lo solucionan como Algoritmo de Dijkstra, Algoritmo de Bellman - Ford, Algoritmo de Floyd - Warshall o el Algoritmo de Johnson [8]–[9].

Sin embargo, el problema del camino más corto circular no es un problema clásico de la computación. Existen autores como M. J. Atallah [10] que sugieren soluciones eficientes con un cambio a coordenadas polares.

Una solución más sencilla, es ejecutar cualquier algoritmo del camino más corto dos veces, una para el viaje de ida y otra para la vuelta de forma que se genere un camino circular. Esta es la solución que hemos adoptado.

4.4.2 La función distancia

Si directamente creamos un grafo con todas las lesiones y utilizamos la distancia geométrica, los algoritmos de

camino más corto no resolverán nuestro problema. La distancia más corta es siempre la línea recta, pero nosotros queremos buscar la distancia más corta pasando por encima de todas las lesiones, para obtener así una secuencia (Figura 4).

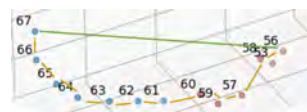


Figura 4: Camino más corto la línea recta en verde, camino buscado en naranja.

La modificación que realizamos es elevar a un exponente la distancia geométrica, para que de esta forma la suma de distancias pequeñas sea más corta que un solo salto grande.

Después de realizar varias pruebas afinando el algoritmo se decidió utilizar el algoritmo de Floyd–Warshall con la siguiente definición de distancia:

- Las distancias menores de 3mm, se consideran 0mm. Esto evita excesivos saltos cuando se acumulan las lesiones por falta de estabilidad.
- La distancia en general es la distancia lineal elevada al 5 exponente. Subir el exponente más no tenía ningún efecto y con exponentes menores en ocasiones “saltaba lesiones”.
- Si la distancia es mayor de un límite (12mm), se considera infinita. Distancias mayores no tenían sentido en el problema y de esta forma reducíamos el trabajo del algoritmo.

4.4.3 La selección de los puntos iniciales

Para crear un camino circular realizamos la siguiente secuencia:

1. Escogemos un punto inicial y un punto final.
2. Buscamos la secuencia del punto inicial al final.
3. Borrarnos del grafo todos los puntos (lesiones) por los que hemos pasado.
4. Buscamos la secuencia del punto final a la inicial.

En general casi cualquiera pareja de puntos iniciales y finales sirve para el propósito. Solo no son adecuados, aquellos puntos que van a quedar fuera de la secuencia óptima circular. En las primeras pruebas probamos con los dos puntos más alejados de la corona con el objetivo de balancear el número de puntos en cada semi-trayecto. Resultó que con mucha frecuencia los puntos más alejados entraban dentro de los pocos que no estaban en la secuencia óptima. Intentamos buscar alternativas como puntos que estaban a un porcentaje de la distancia máxima, pero finalmente no fuimos capaces de encontrar una solución mejor que establecer los puntos iniciales de forma aleatoria. En caso de fallo, se puede reintentar el algoritmo de forma semiautomática.

5. Resultados

Se realizaron sucesivas ejecuciones del algoritmo, en una primera ejecución se procesaron los datos de forma automática, supervisados por un operador solo para

comprobar el nivel de éxito alcanzado. En el caso de que el funcionamiento no fuera correcto, se realizaba una segunda ejecución de forma semiautomática. En esta segunda ejecución, se podían borrar lesiones o establecer los puntos iniciales, final y algunos intermedios para la búsqueda del camino circular (Tabla 1).

5.1. Validación del algoritmo

En el 94% de los casos del grupo 1 y el 73% de los casos del grupo 2, el algoritmo realizó correctamente la segmentación de ambos círculos de ablación y determinó la secuencia de las lesiones de ablación de forma totalmente automática. Solo en tres de los casos (0,4%) no fue posible encontrar la secuencia de lesiones ni siquiera de forma semiautomática. Si bien es cierto que la calidad de las lesiones de estos casos, básicamente una nube de puntos, les excluiría de cualquier estudio clínico con mucha probabilidad.

Grupo	Éxito automático	Éxito semi-automático	Fallo	Total
Grupo 1	214	12	1	227
Grupo 2	328	120	2	450

Tabla 1: Resultados del algoritmo

5.2. Análisis de los errores

Tan solo tres orígenes explicaban la mayor parte (99,6%) del fallo del algoritmo automático:

1. La selección incorrecta de los puntos iniciales.
2. Las líneas de ablación contiguas a las coronas
3. Agrupaciones de lesiones o bucles, habitualmente en la parte inferior de las venas.

Todos los problemas se pueden mitigar con facilidad en la ejecución semiautomática. El primero de los problemas es trivial, y basta con ejecutar de nuevo el algoritmo y dejar que escoja otros puntos iniciales. Para evitar que el camino siga líneas de ablación, basta con no seleccionar ningún punto inicial en las líneas. Se puede repetir la ejecución hasta que funcione correctamente o seleccionar los puntos directamente. Finalmente, para evitar los bucles en las aglomeraciones de lesiones, es necesario marcar puntos intermedios y de esta forma evitar que el algoritmo entre y salga de estas creando un lazo (Figura 5).

6. Conclusión

Este trabajo muestra el proceso para utilizar un algoritmo semi-automático para medir la distancia interlesión de un procedimiento de ablación de FA y además lo deja disponible para la comunidad [11]. Este dato es importante para optimizar el procedimiento de ablación con catéter de RF, dado que aún no existe todavía un valor óptimo para la eficacia y la eficiencia. Aún no es posible utilizarlo de forma 100% automática, aunque en la mayoría de las ocasiones el operador no tendrá nada más que aceptar los resultados.



Figura 5: La secuencia esperada (naranja) se ve frustrada por una acumulación de lesiones (verde), sucede con cierta frecuencia en la parte inferior de las venas.

Agradecimientos

Este Trabajo ha sido parcialmente financiado por el proyecto del Plan Nacional de I+D+i “Cribado diagnóstico de microorganismos mediante microscopia avanzada e inteligencia artificial en patologías humana” (PID2021-127691OB-I00), del Ministerio de Ciencia e Innovación, cofinanciado con fondos FEDER.

Referencias

- [1] cardiofamilia.org, «Fisiopatología de la fibrilación auricular». Cátedra de Terapias Avanzadas en Patología Cardiovascular.
- [2] D. C. Shah, M. Haïssaguerre, y P. Jaïs, «Catheter ablation of pulmonary vein foci for atrial fibrillation: PV foci ablation for atrial fibrillation.», *Thorac Cardiovasc Surg*, vol. 47 Suppl 3, pp. 352-356, ago. 1999, doi: 10.1055/s-2007-1013198.
- [3] W.-S. Lin *et al.*, «Catheter ablation of paroxysmal atrial fibrillation initiated by non-pulmonary vein ectopy.», *Circulation*, vol. 107, n.º 25, pp. 3176-3183, jul. 2003, doi: 10.1161/01.CIR.0000074206.52056.2D.
- [4] E. H. David, «Biophysics and Pathophysiology of Radiofrequency Lesion Formation», en *Catheter Ablation of Cardiac Arrhythmias*, 2019.
- [5] M. El Haddad *et al.*, «Determinants of Acute and Late Pulmonary Vein Reconnection in Contact Force-Guided Pulmonary Vein Isolation: Identifying the Weakest Link in the Ablation Chain.», *Circ Arrhythm Electrophysiol*, vol. 10, n.º 4, p. e004867, abr. 2017, doi: 10.1161/CIRCEP.116.004867.
- [6] P. Taghji *et al.*, «Evaluation of a Strategy Aiming to Enclose the Pulmonary Veins With Contiguous and Optimized Radiofrequency Lesions in Paroxysmal Atrial Fibrillation: A Pilot Study.», *JACC Clin Electrophysiol*, vol. 4, n.º 1, pp. 99-108, ene. 2018, doi: 10.1016/j.jacep.2017.06.023.
- [7] M. Duytschaever *et al.*, «Standardized pulmonary vein isolation workflow to enclose veins with contiguous lesions: the multicentre VISTAX trial.», *Europace*, vol. 22, n.º 11, pp. 1645-1652, nov. 2020, doi: 10.1093/europace/eaal157.
- [8] R. Bellman, «ON A ROUTING PROBLEM», *Quarterly of Applied Mathematics*, vol. 16, pp. 87-90, 1958.
- [9] E. W. Dijkstra, «A note on two problems in connexion with graphs», *Numerische Mathematik*, vol. 1, n.º 1, pp. 269-271, dic. 1959, doi: 10.1007/BF01386390.
- [10] M. J. Atallah, D. Z. Chen, y D. T. Lee, «An optimal algorithm for shortest paths on weighted interval and circular-arc graphs, with applications», en *Algorithms—ESA '93*, T. Lengauer, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 1993, pp. 13-24.
- [11] «Herramientas para investigación con los datos del navegador Carto3». <https://github.com/nomine555/cartools>

Perturbación eléctrica y térmica provocada por sensores de temperatura en estudios preclínicos de ablación cardiaca por radiofrecuencia

L. Cuenca Dacal¹, E. Berjano¹, A. González-Suárez^{2,3}, J.J. Pérez¹

¹ BioMIT, Department of Electronic Engineering, Universitat Politècnica de València, Valencia, España, luis.cuencadacal@gmail.com, eberjano@eln.upv.es, jjperez@eln.upv.es

² Translational Medical Device Lab, School of Medicine, University of Galway, Ireland, ana.gonzalezsuarez@universityofgalway.ie

³ Universidad Internacional de Valencia, Valencia, España

Resumen

La ablación cardiaca por radiofrecuencia (RF) es una terapia de gran relevancia para tratar las arritmias cardiacas. Con la RF se consigue calentar un tejido de forma muy precisa, y generar un daño irreversible él, impidiendo la activación de esas células, y permitiendo eliminar focos arrítmicos y aislar estructuras. El diseño y desarrollo de esta tecnología requiere realizar ensayos preclínicos. Frecuentemente, se emplean ensayos ex vivo con muestras de tejido biológico para determinar el tamaño y forma de las lesiones y las temperaturas alcanzadas. El objetivo de este estudio es averiguar si los sensores de temperatura posicionados en el interior del tejido durante dichos ensayos producen alguna perturbación. Se construyó un modelo computacional 3D incluyendo una porción de tejido cardiaco, y se analizaron las temperaturas alcanzadas. Se modelaron dos tipos de sensores de temperatura (termopares y fibra óptica) en distintas localizaciones. Mientras que las fibras ópticas apenas provocan perturbaciones (desviaciones menores de 0.6°C), la presencia de los termopares altera de forma importante la medida (diferencias de hasta 9.78°C). Sugerimos que la causa es el propio material metálico de los cables del termopar, que actúa como disipador térmico. Pese a todo, el tamaño y forma de la lesión inducida por RF no parece estar afectado por la presencia de cualquiera de los dos sensores.

1. Introducción

La ablación de tejidos biológicos para el tratamiento de distintas patologías consiste en generar un daño permanente en el tejido mediante diferentes fuentes de energía, tales como radiofrecuencia (RF), microondas, láser, crioablación, etc.). Esto permite alcanzar una temperatura ablativa, durante un tiempo determinado en un lugar preciso del tejido objetivo, lo que induce una lesión irreversible. Una técnica ablativa de gran relevancia actualmente es la ablación cardiaca por RF para el tratamiento de las arritmias cardiacas. Se lleva a cabo introduciendo un catéter que en su extremo distal dispone de un electrodo, y con el que se establece una corriente en RF entre este y el electrodo dispersivo adherido normalmente en forma de parche a la espalda del paciente [1]. Los estudios preclínicos se realizan para determinar la seguridad y eficacia de los nuevos diseños y desarrollos [2]. Consisten en replicar las condiciones de la práctica clínica en el laboratorio utilizando muestras de

tejido ex vivo [3,4]. El resultado permite conocer el tamaño de la lesión y las temperaturas intratisulares. Estas últimas se miden con pequeños sensores de fibra óptica o termopares (0.3/0.4 mm de diámetro) situados normalmente muy próximos al catéter de ablación [5-7], lo que podría alterar la temperatura registrada. El objetivo de este estudio fue usar modelado computacional para evaluar el impacto perturbador que producen estos sensores de temperatura en el tejido sometido a ablación por RF en el contexto de estudios preclínicos de ablación cardiaca.

2. Métodos

2.1. Modelado de la geometría

Se ha elaborado un modelo computacional 3D que incluye un catéter de ablación de 4 French de diámetro con un electrodo distal de 4 mm de longitud situado de forma perpendicular al tejido cardiaco [8] y rodeado de sangre. Por simetría, se consideró solo la mitad del volumen a estudiar (Figura 1). El electrodo dispersivo se ha modelado como una condición de contorno eléctrica de 0 V en las caras exteriores del modelo [9].

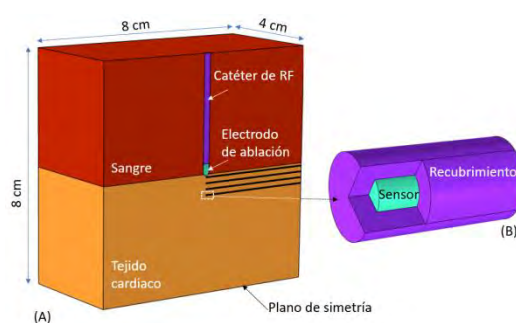


Figura 1. (A) Geometría del modelo para el caso de sensores en posición perpendicular al electrodo de ablación. (B) Detalle del extremo de un sensor de temperatura

Se han considerado dos sensores típicamente utilizados en los estudios preclínicos: termopar tipo T modelo IT-21 de Physitemp Instruments (Clifton, NJ, USA) y fibra óptica modelo ODiSI 6000 de Luna (Roanoke, VA, USA). Los termopares se han modelado como cilindros internos sólidos de diámetro 0.2 mm recubiertos con un aislante hasta alcanzar un diámetro exterior de 0.4 mm y

un recubrimiento en el extremo de 0.1 mm de grosor (ver Fig. 1B). Los sensores de fibra óptica están modelados como cilindros sólidos de 0.15 mm de diámetro recubiertos con aislante hasta alcanzar un diámetro externo de 0.3 mm y mismo espesor en el extremo. Estas medidas están inspiradas en los sensores empleados en estudios preclínicos previos [10-12].

2.2. Distribución espacial de los sensores

Se han analizado tres distribuciones espaciales de sensores de temperatura idénticas a aquellas reportadas en la literatura. En el primer caso los sensores se situaron paralelos al eje del catéter, a distancias de 1, 2, 3 y 4 mm del borde del catéter, y se analizó el efecto de los sensores cuando estos se encuentran a 1, 3, 5 y 7 mm de profundidad desde la superficie del tejido en contacto con la sangre [5] (ver Figura 2).

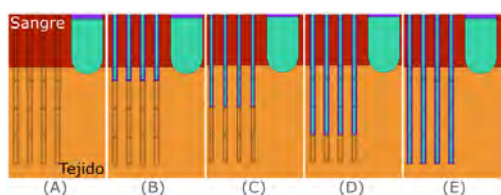


Figura 2. Modelos con los sensores paralelos al eje del catéter a profundidades de 1, 3, 5 y 7 mm (B-E) y sin sensores (A).

En el segundo caso los sensores de temperatura se situaron perpendiculares al eje del catéter, con sus extremos a profundidades de 1, 3, 5 y 7 mm [4] (ver Figura 3).

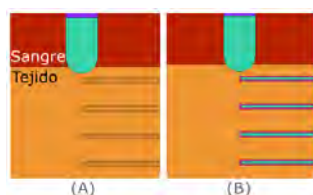


Figura 3. Modelo con sensores perpendiculares al catéter a profundidades de 1, 3, 5 y 7 mm (B) y sin sensores (A).

En el tercer caso los sensores se situaron paralelos al catéter, alineados, sin separación entre ellos, y a 3, 5 y 7 mm de profundidad [6] (Figura 4).



Figura 4. Modelo con sensores paralelos al catéter y sin separación.

2.3. Propiedades de los materiales y del tejido

Las propiedades del tejido se han tomado de la base de datos IT'IS Foundation para una frecuencia de 500 kHz [10]. Las propiedades del catéter provienen de Pérez *et al* [11]. Las propiedades térmicas de la sangre no se tuvieron en cuenta porque el problema térmico no se ha resuelto en el subdominio del fluido. En su lugar se emplearon condiciones de contorno que modelan la refrigeración en

el endocardio. La conductividad del tejido aumenta con la temperatura un 1.5%/°C. Las propiedades del interior del termopar se han simplificado como el promedio de las propiedades del cobre / constantán [12], mientras que las propiedades de la fibra óptica son el promedio de las propiedades encontradas en la base de datos usando las palabras clave “optic”, “quartz”, y “fiber” [12]. Todas las propiedades los materiales están resumidas en la Tabla 1.

Material	σ (S/m)	k (W/m·K)	ρ (kg/m ³)	c (J/kg·K)
Electrodo/Pt-Ir	4.6×10^6	71	21500	132
Catéter/Poliuretano	10^{-5}	23	1440	1050
Sangre	0.748	-	-	-
Miocardio	0.541	0.56	1081	3686
Fibra óptica	5×10^9	2	2210	750
Poliimida	1×10^8	0.12	1440	1090
Cobre-Constantán	3.03×10^7	206.4	8915	389
Teflón	1×10^{-16}	0.3	2162	1350

Tabla 1. Características físicas de los elementos utilizados en el modelo a 500 kHz y evaluados a 37 °C

2.4. Ecuaciones y condiciones de frontera

Se utilizó el software COMSOL Multiphysics (COMSOL, Burlington, MA, USA) para resolver el problema electro-térmico mediante el método de los elementos finitos. La ecuación para el problema térmico fue la ecuación del biocalor (1), donde ρ es densidad (kg/m³); c es calor específico (J/kg·K), T es temperatura (°C), t es tiempo (s), k es conductividad térmica (W/m·K), Q_{RF} es calor generado por radiofrecuencia (W/m³), Q_p es calor disipado a través de la sangre (W/m³), Q_{met} es calor generado por el metabolismo (W/m³). Q_{met} y Q_p son despreciables en este contexto [1]. En el problema eléctrico se emplea una aproximación cuasiestática, donde el vector campo eléctrico \vec{E} se obtiene de la ecuación (2), donde Φ es el potencial eléctrico. Esta ecuación se obtiene de la ecuación de Laplace (3), donde σ es conductividad eléctrica (S/m). El calor generado por RF se obtiene por tanto de la ecuación (4).

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T) + Q_{RF} + Q_p + Q_{met} \quad (1)$$

$$\vec{E} = -\nabla \Phi \quad (2)$$

$$\nabla \cdot (\sigma(T) \nabla \Phi) = 0 \quad (3)$$

$$Q_{RF} = \sigma |\vec{E}|^2 \quad (4)$$

No se ha considerado la vaporización de agua porque la temperatura del tejido siempre se mantuvo por debajo de 100 °C. Se ha utilizado una potencia constante de 6 W (equivalente a 12 W en el volumen completo). Se han establecido flujos de calor convectivos en las superficies tejido-sangre y electrodo-sangre con coeficientes de 610 W/m²·K y 3346 W/m²·K, respectivamente, así como en las superficies termopar-sangre (8120 W/m²·K) y fibra óptica-sangre (9470 W/m²·K) [13].

3. Resultados

3.1. Sensores paralelos al eje del catéter

Se observan importantes diferencias en la temperatura registrada en el modelo sin sensores y con termopares. En general, la diferencia es menor cuando aumenta la distancia al electrodo. Cuando los sensores se insertan a poca profundidad, la lectura de los termopares es menor que en el modelo sin sensores, llegando a subestimar la medida en 9.28 °C tras 10 s de aplicación en el termopar más próximo al catéter cuando este se inserta 1mm en el tejido. Sin embargo, en las simulaciones en las que los sensores se encuentran a mayor profundidad, las lecturas son sobreestimadas, alcanzando una diferencia de 1.54 °C en el termopar más próximo al catéter cuando se inserta a la profundidad de 7 mm (ver Figuras 5 y 6).

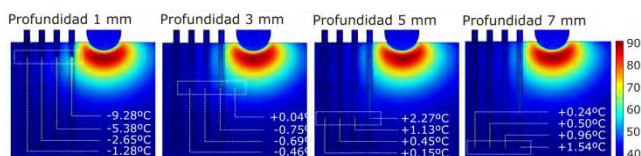


Figura 5. Mapa de temperaturas en el caso con termopares paralelos al catéter separados, y diferencia de registros con respecto al modelo sin sensores (escala en °C).

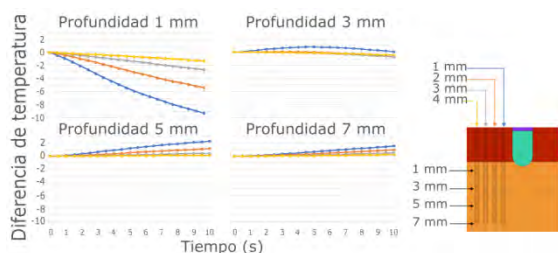


Figura 6. Diferencia entre temperatura en modelos con y sin termopares paralelos al catéter.

Para el caso de las fibras ópticas, las diferencias son mucho menores. En este caso, las medidas siempre son subestimadas, y la máxima desviación se produce con los sensores a 1 mm de profundidad, en el caso del sensor más próximo al catéter, tras 10 s de aplicación de corriente (ver Figuras 7 y 8).

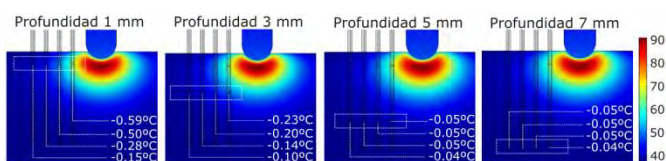


Figura 7. Mapa de temperaturas en simulaciones con sensores de fibra óptica paralelos al catéter, y diferencia de registros con respecto al modelo sin sensores.

3.2. Sensores perpendiculares al eje del catéter

La presencia de termopares situados de manera perpendicular al eje del catéter altera considerablemente la temperatura en su entorno, subestimando sensiblemente la temperatura, con una diferencia de hasta 9.78 °C en el termopar más próximo al catéter. Sin embargo, los sensores de fibra óptica subestiman mucho menos la temperatura, con una diferencia máxima de 0.32 °C (ver Figuras 9 y 10).

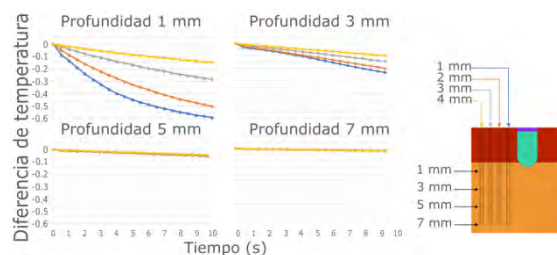


Figura 8. Diferencia de temperaturas en simulaciones con y sin sensores de fibra óptica paralelos al catéter.

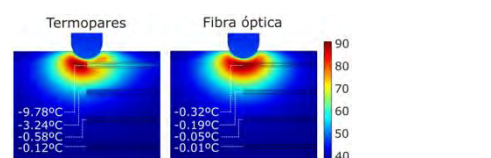


Figura 9. Mapa de temperaturas en simulaciones con sensores de temperatura perpendiculares al catéter y diferencia de registros con respecto al modelo sin sensores

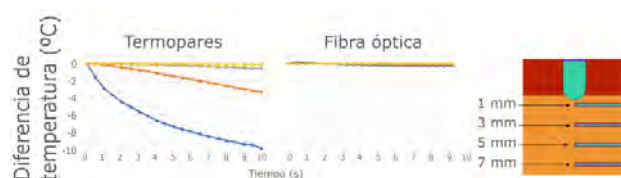


Figura 10. Diferencia entre temperatura en modelos con sensores perpendiculares al catéter

3.3. Sensores paralelos al eje del catéter sin separación

A pesar de que uno de los sensores está muy próximo al catéter, en el caso de que los sensores se encuentran paralelos al eje del catéter y sin separación entre ellos, la alteración de la temperatura es solo de 0.8 °C para el termopar, y de 0.25 °C para la fibra óptica. El mapa de temperatura sí que parece considerablemente deformado por la presencia de los termopares (ver Figura 10).

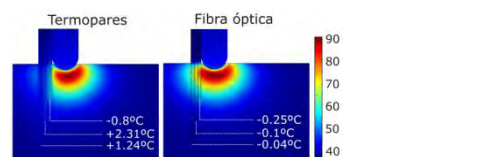


Figura 11. Mapa de temperaturas en simulaciones con termopares paralelos al catéter sin separación, y diferencia de registros con respecto al modelo sin sensores.

4. Discusión

Se ha desarrollado un modelo computacional para comprobar el impacto que supone insertar pequeños sensores de temperatura en el tejido durante estudios preclínicos de ablación cardiaca por RF. Se ha comprobado que la presencia de termopares altera más la temperatura que cuando se emplean sensores de fibra óptica. Esta alteración aumenta con la proximidad al catéter y con el tiempo de aplicación. La explicación que hemos encontrado para el mejor desempeño de la fibra óptica en este tipo de aplicación es que conducen menos calor por no tener elementos metálicos los cuales

favorecen la transmisión del calor (la conductividad térmica del cuarzo es 100 veces menor que la del cobre-constantán). Aunque las lecturas de temperatura con termopares varían significativamente con respecto a los modelos sin sensores, no lo hace así el tamaño del volumen de lesión valorado mediante la isoterma de 55 °C, que en general se mantiene similar y únicamente alterado en el entorno muy próximo a los sensores. En el diseño de ensayos experimentales, o en el potencial desarrollo de sensores para medir temperatura en tejido biológico, sería de gran interés considerar las perturbaciones que introduce el sensor, ya que la desviación en la medida de la temperatura observada puede conducir a errores en las conclusiones de los experimentos. Estas conclusiones podrían ser también valiosas para otras aplicaciones de ablación térmica en diferentes tejidos, no solo la ablación cardiaca por RF. Asimismo, la frecuencia usada no es el factor determinante en la variación de la medida, sino el nivel de temperatura en el punto de medición. Por ese motivo, las conclusiones serían también válidas para aplicaciones a frecuencias diferentes de 500 kHz dentro del rango de RF.

Como limitaciones, debemos resaltar que en la práctica clínica es posible que las condiciones sean diferentes a las planteadas en este estudio (diseño de electrodo, posiciones relativas, flujo de sangre, potencia, etc.). Tampoco se han considerado las posibles interferencias eléctricas entre la corriente en RF y los cables de los termopares. Por último, este estudio consideró solo tejido homogéneo. El tejido biológico contiene en realidad grasa, tejido conectivo, además de miocardio. Sin embargo, nuestro estudio se centró en los efectos que provocan los sensores en ensayos preclínicos de banco, donde los fragmentos suelen ser bastante homogéneos. Asimismo, la capacidad de evacuar el calor por parte del sensor es relativamente independiente del tejido, y considerar un tejido más complejo conduciría posiblemente a las mismas conclusiones.

5. Conclusiones

Los hallazgos sugieren que la presencia de termopares altera las medidas de temperatura del tejido, aunque no afectan al tamaño y forma de la lesión. La fibra óptica provoca menor perturbación térmica debido posiblemente a que no contiene materiales metálicos y por tanto disipan menos el calor. En el caso de que se realice un ensayo ex vivo con termopares o termistores, el análisis de un modelo computacional adaptado a ese ensayo sería de gran utilidad.

6. Agradecimientos

Proyecto financiado por el Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación, Fondo Europeo de Desarrollo Regional (MCIN/AEI/10.13039/501100011033/FEDER,UE) con referencias PID2022-136273OB-C31 y PID2022-136273OA-C33.

Referencias

[1] González-Suárez, A.; Pérez, J.J.; Irastorza, R.M.; D'Avila, A.; Berjano, E. Computer modeling of radiofrequency

cardiac ablation: 30 years of bioengineering research. *Comput. Methods Programs Biomed.* 2022, 214, 106546.

- [2] Hsu, S.S.; Hoh, L.; Rosenbaum, R.M.; Rosen, A.; Walinsky, P.; Greenspon, A.J. A method for the in vitro testing of cardiac ablation catheters. *IEEE Trans. Microw. Theory Tech.* 1996, 44, 1841–1847.
- [3] Rossmann, C.; Motamarry, A.; Panescu, D.; Haemmerich, D. Computer simulations of an irrigated radiofrequency cardiac ablation catheter and experimental validation by infrared imaging. *Int. J. Hyperth.* 2021, 38, 1149–1163.
- [4] Demazumder, D.; Mirotnik, M.S.; Schwartzman, D. Comparison of irrigated electrode designs for radiofrequency ablation of myocardium. *J. Interv. Card. Electrophysiol.* 2001, 5, 391–400.
- [5] Petersen, H.H.; Chen, X.; Pietersen, A.; Svendsen, J.H.; Haunsø, S. Tissue temperatures and lesion size during irrigated tip catheter radiofrequency ablation: An in vitro comparison of temperature-controlled irrigated tip ablation, power-controlled irrigated tip ablation, and standard temperature-controlled ablation. *Pacing Clin. Electrophysiol.* 2000, 23, 8–17.
- [6] Nakagawa, H.; Ikeda, A.; Sharma, T.; Govari, A.; Ashton, J.; Maffre, J.; Lifshitz, A.; Fuimaono, K.; Yokoyama, K.; Wittkampf, F.H.M.; et al. Comparison of in vivo Tissue Temperature Profile and Lesion Geometry for Radiofrequency Ablation with High Power-Short Duration and Moderate Power-Moderate Duration: Effects of Thermal Latency and Contact Force on Lesion Formation. *Circ. Arrhythm. Electrophysiol.* 2021, 14, e009899.
- [7] Kooroor, P.; Daly, M.P.; Pouliopoulos, J.; Byth, K.; Dewsnap, B.I.; Eipper, V.E.; Yung, T.; Uther, J.F.; Ross, D.L. Comparison of radiofrequency ablation in normal versus scarred myocardium. *J. Cardiovasc. Electrophysiol.* 2006, 17, 80–86.
- [8] Pérez, J.J.; Gonzalez-Suárez, A.; Maher, T.; Nakagawa, H.; D'Avila, A.; Berjano, E. Relationship between luminal esophageal temperature and volume of esophageal injury during RF ablation: In silico study comparing low power-moderate duration vs. high power-short duration. *J. Cardiovasc. Electrophysiol.* 2022, 33, 220–230.
- [9] Irastorza, R.M.; Gonzalez-Suarez, A.; Pérez, J.J.; Berjano, E. Differences in applied electrical power between full thorax models and limited-domain models for RF cardiac ablation. *Int. J. Hyperth.* 2020, 37, 677–687.
- [10] Hasgall, P.A.; Di Gennaro, F.; Baumgartner, C.; Neufeld, E.; Lloyd, B.; Gosselin, M.C.; Payne, D.; Klingensböck, A.; Kuster, N. *IT'IS Database for Thermal and Electromagnetic Parameters of Biological Tissues, Version 4.1; IT'IS Foundation: Zürich, Switzerland, 2022*
- [11] Pérez, J.J.; Ewertowska, E.; Berjano, E. Computer modeling for radiofrequency bipolar ablation inside ducts and vessels: Relation between pullback speed and impedance progress. *Lasers Surg. Med.* 2020, 52, 897–906
- [12] Material Property Database. Available online: www.matweb.com.
- [13] Suárez, A.G.; Hornero, F.; Berjano, E.J. Mathematical modeling of epicardial RF ablation of atrial tissue with overlying epicardial fat. *Open Biomed. Eng. J.* 2010, 4, 47–55.

Nanomedicina

Miércoles 22 15:00 - 16:30
Aula Sebastián Ferignan

Laser re-irradiation of palladium nanoparticles for antibacterial applications

A. M. Vilas¹, M. Fernández-Arias¹, M. Bountinguiza^{1,2}, D. Rodríguez³, F. Arias-González⁴, P. Pou-Álvarez¹, A. Riveiro^{1,2}, J. Gil⁴ and J. Pou^{1,2}

¹LaserON Research Group, CINTECX, School of Engineering, University of Vigo, Lagoas-Marcosende, E-36310, Spain; anamaria.vilas@uvigo.gal (A.M.V); monfernandez@uvigo.gal (M.F.-A); mohamed@uvigo.es (M.B); ppou@uvigo.gal (P.P.-Á), ariveiro@uvigo.es (A.R); jpou@uvigo.es (J.P)

²Galicia Sur Health Research Institute (IIS Galicia Sur), SERGAS-UVIGO, E-36310 Vigo, Spain

³Biomaterials, Biomechanics and Tissue Engineering Group, Materials Science and Metallurgical Department, UPC-Barcelona TECH, 08930 Barcelona, Spain; daniel.rodriiguez.rius@upc.edu

⁴School of Dentistry, Universitat Internacional de Catalunya, 08017 Barcelona, Spain; farias@uic.es (F.A.-G), xavier.gil@uic.cat (J.G)

Summary

Bacterial resistant infections represent the second leading cause of mortality worldwide, being responsible for almost the 14% of the global deaths according to figures of 2019.

Nanoparticles of noble metals have received special attention because of their outstanding bactericidal activity, that is intrinsically linked to their physico-chemical properties, an asset that makes them an appealing alternative to combat resistant microorganisms. So as to optimize their applications within the biomedical field, it becomes relevant to develop and employ means to influence and act upon nanoparticles' physico-chemical characteristics.

*In this research, palladium nanoparticles were synthesized via laser ablation. A nanosecond Nd:YVO₄ laser operating at 532 nm was employed to ablate a palladium target immersed in deionized water in order to obtain the colloidal suspension. Then, the previous suspension was passed twice through a re-irradiation system to narrow the size and size distribution of palladium nanoparticles. Bacterial assays with *Staphylococcus aureus* revealed an effect of size on the bactericidal behaviour of palladium nanoparticles.*

1. Introduction and objectives

Antimicrobial resistance (AMR) was believed to be under control in the past decades, but since 2010 it has been generating more and more concern, as the inadequate use of antimicrobials has caused an increasing rate of resistance development and there is a shortage of new efficient drugs to combat these superbacteria [1].

Given the drug resistance on the rise, it is estimated that the burden of deaths from AMR could snowball to 10 million lives per year by 2050, above cancer, diabetes and traffic accidents, at a cumulative cost to global economic output of 100 trillion USD, if no action is taken. This implies that by 2050, one person would die every three seconds and each person in the world today would cost 10,000 USD. Those studies and statistics pointed to an increasing concern about antimicrobial resistance growth and spreading and convey despairing predictions for the future. For this reason, further research and development of novel treatments is required to create opportunities to combat this global health problem [2,3].

Ancient civilizations were aware of the beneficial properties of metals for health, such as silver (Ag), copper (Cu), gold (Au), etc., and thus, employed them for

therapeutic purposes. Recent advances in the field of nanotechnology have grown an increasing interest in metallic nanoparticles (NPs) as a promising tool to treat infectious diseases, so that their physico-chemical properties determine to a large extent their bactericidal behaviour [4–7]. Among metallic NPs, silver NPs have received special attention due to its outstanding bactericidal activity at concentrations that are not cytotoxic to human cells [8–10]. However, several reports have already warned about the potential development of bacterial resistance to silver NPs after repeated exposure. Hence, it is of major importance to investigate on different metallic materials as effective as silver [11].

Laser ablation permits to synthesize nanomaterials without any chemical precursor or reaction that may produce any toxic by-products, which becomes determining when the objective of those nanomaterials is to be employed in biomedical environments. In addition, it presents remarkable advantages such as low cost, high productivity, good stability and a certain degree of both size and shape control, by adjusting processing parameters [12–14]. The close relationship between the properties of NPs and their size is perhaps the most important quality of this new family of materials and underlines the importance of finding ways to affect these dimensional parameters in order to optimise their use in different biomedical applications.

In this work, we designed a system to re-irradiate colloidal suspensions of NPs in order to decrease their size and narrow their size distribution. We studied the system requirements and defined its desirable performance. Then, a prototype was manufactured and used in laboratory assays with palladium NPs, to verify its desired functioning. Bactericidal assays with *Staphylococcus aureus* were performed with original and re-irradiated colloidal suspensions in order to analyse the influence of size on their antibacterial behaviour.

2. Materials and methods

2.1. System design and manufacture

Prior to any experimental procedure, it was necessary to define several requirements for the re-irradiation system. The system must ensure that the laser beam interacts in the

same way with all the NPs that pass through it, in order to homogeneously re-irradiate all of them. In addition, it should integrate two storage tanks (one for the starting solution and another for the re-irradiated solution) with their own agitation systems, to maintain a homogeneous concentration inside and prevent nanoparticles in solution from agglomerating or depositing at the bottom of the tanks. Both storage tanks should be connected by a thin duct that carries the initial solution from its tank through the capillary, where it will be re-irradiated by the laser beam, to the remaining tank; preferably without any additional pumping element.

Once the requirements were defined, the experimental system was designed and built accordingly.

2.2. Nanoparticle synthesis

To obtain the Pd nanoparticles, a nanosecond Nd:YVO₄ laser operating at 532 nm was used to ablate a palladium foil of 99.9% of purity (Alfa Aesar, Massachusetts, USA) immersed in deionized water.

The main operational parameters of the laser are detailed in Table 1.

Table 1. Processing conditions used for the experimental assays

Operation frequency (kHz)	20
Maximum intensity (A)	27.7
Average output power (W)	5
Pulse width (μs)	1
Scanning speed (mm/s)	50

After the laser ablation process, the colloidal suspension was twice re-irradiated employing the prototype of the fragmentation system (see Figure 1). For the experimental assays, two capillaries of different section were employed (0.5 and 1.0 mm). The obtained samples are defined in Table 2. The re-irradiation process was carried out using the same laser source with the same processing parameters used to obtain the initial colloidal solution (Table 1).

Table 2. Identification of the NPs produced

Sample identification	Capillary width (mm)	Obtention
a0	---	As ablated NPs
a1	0.5	NPs after one re-irradiation
a2	0.5	NPs after two re-irradiations
b0	---	As ablated NPs
b1	1.0	NPs after one re-irradiation
b2	1.0	NPs after two re-irradiations

2.3. Nanoparticles' characterization

Transmission electron microscopy (TEM)

After the experimental assays, drops of each sample were deposited on Formvar/Carbon supported copper grids with 400 mesh of size (TedPella Inc., Redding, CA, USA) and left to dry before their analysis. Then, nanoparticle size, morphology and dispersion were characterized by transmission electron microscopy (TEM) using JEOL JEM 1010 equipment (JEOL, Akishima, Japan). Size distribution was studied on the basis of TEM micrographs, employing the software ImageJ to measure the diameter of 400 nanoparticles for each sample.

High transmission electron microscopy (HRTEM)

The palladium nanoparticles were also observed by high resolution transmission electron microscopy (HRTEM) using JEOL JEM 2010F FEG (JEOL, Akishima, Japan), so as to study their crystalline structure.

X-Ray diffraction (XRD)

Drops of each sample were also placed on glass sample holders to acquire a layer of the proper thickness to analyse via X-Ray diffraction (XRD), in order to examine the crystalline structure of nanoparticles. With that purpose, the Pananalytical X'Pert Pro X-ray diffractometer (Malvern Panalytical, Malvern, UK), using monochromated Cu-K α radiation ($\lambda = 1.54 \text{ \AA}$) over the 30–130° 2 θ range with step size of 0.026° was used. The XRD spectra of each sample was compared to that of the precursor material (palladium foil) and the ICDD-JCPDS database was employed to identify the crystalline phases.

UV-Vis spectroscopy

To conclude the physico-chemical characterization, the UV-VIS spectra of the samples were measured to study their stability and optical properties. The spectrophotometer HP 8452 (Hewlett Packard, Palo Alto, CA, USA) was used to perform the measurements within the range 190 – 820 nm.

Bactericidal tests

The bacterial adhesion to non-treated titanium samples and samples decorated with the nanoparticles was studied with *Staphylococcus aureus* (Colección Española de Cultivos Tipo (CECT) 435, Spain) cultured in BHI broth (Scharlab SL, Sentmenat, Spain).

Control and treated samples were immersed in ethanol and distilled water for 15min each and put in a 24-multiwell plate (Nunc, USA) with 1ml bacterial suspension at 37C for 2h. The adhered bacteria were collected seeded onto agar plates supplemented with BHI medium. The agar plates were then incubated at 37°C for 24h and the CFUs counted.

3. Results and discussion

Figure 1 presents the re-irradiation system's prototype positioned on the laser processing station.

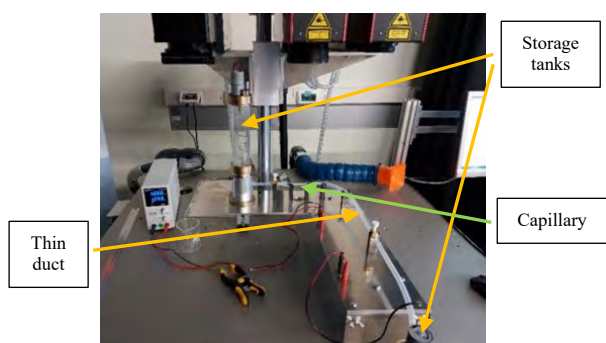


Figure 1. System prototype positioned on the laser processing station

The analysis of the morphology and size distribution of NPs synthesized by laser ablation and re-irradiated on the system was performed using different TEM micrographs. Figure 2 shows TEM images of Pd nanoparticles as produced and re-irradiated two times along to their corresponding size distribution histograms. Histograms were made by measuring the diameter of 400 NPs per sample.

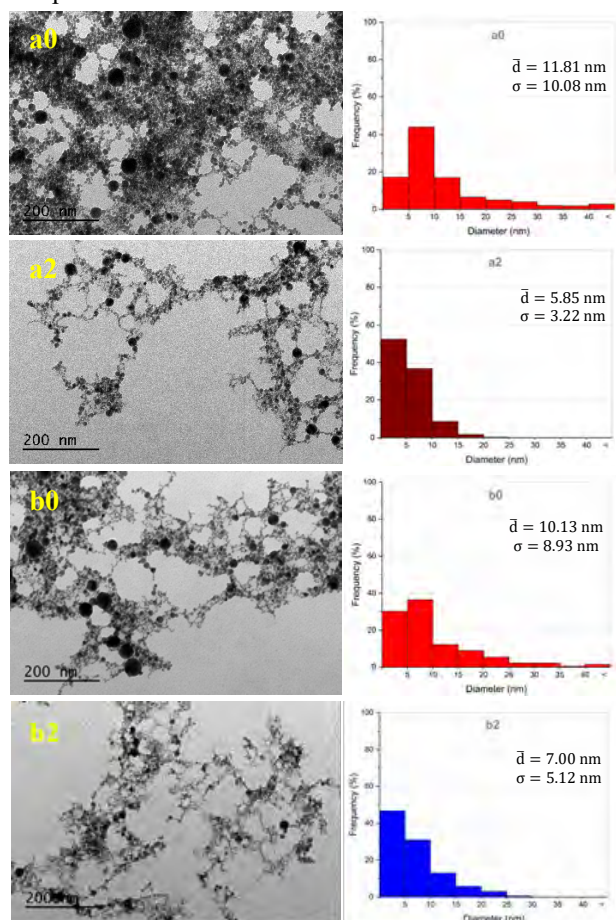


Figure 2. TEM micrographs and size distribution of Pd NPs as obtained (a0 and b0) and re-irradiated two (a2 and b2)

It can be observed that all Pd NPs obtained are spherical with a certain tendency to agglomeration, forming chain structures (Figure 2). The re-irradiation process does not seem to alter NPs' morphology, since all micrographs show NPs with rounded shape forming clusters. However, it indeed affects size and dispersion: the mean diameter decreases more than a 50% and a clear influence of capillary thickness can be noted.

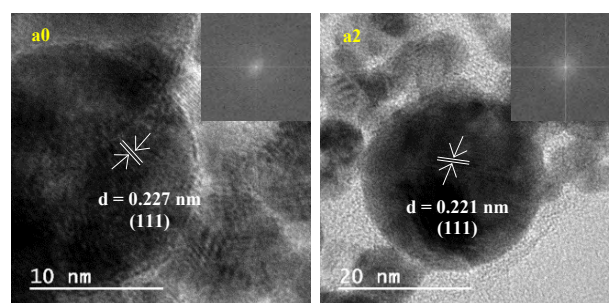


Figure 3. HRTEM of as produced Pd NPs (a0) and re-irradiated two times (a2)

All the particles obtained, the as-ablated as well as the re-irradiated ones, are crystalline and pure Pd. Figure 3 shows HRTEM of samples a0 and a2, with clear lattice fringes and their corresponding Fast Fourier Transform (FFT). The measured interplanar distances from the FFT correspond to the family planes of cubic Pd (JCPDS-ICDD ref. 00-005-0681). It is important to highlight that composition and crystalline structure do not seem to be significantly modified by the re-irradiation process.

The XRD analysis confirmed the elemental nature of the Pd NPs (Figure 4). All NPs are crystalline and pure palladium with characteristic diffraction peaks at 40.12° (111), 46.66° (200), 68.08° (220) and 82.09° (311).

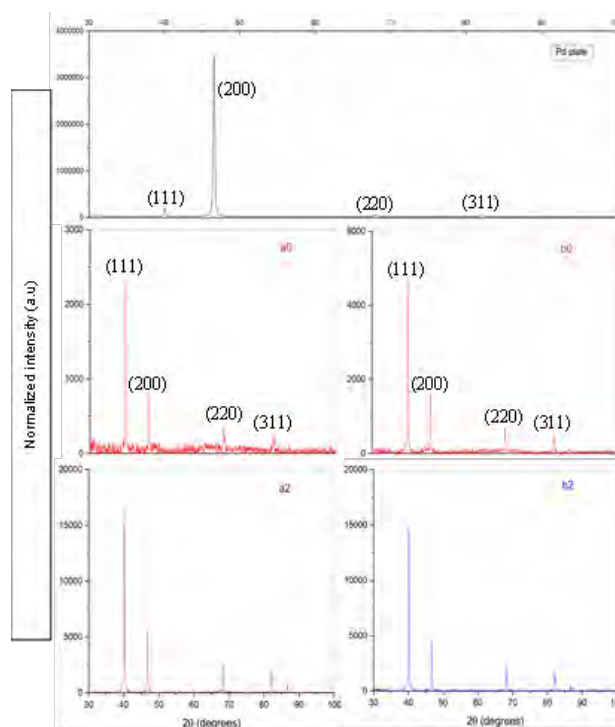


Figure 4. XRD of the palladium plate and palladium nanoparticles: as produced (a0 and b0) and re-irradiated two times (a2 and b2)

To conclude with the characterization, the UV-VIS spectra of the colloidal suspensions show a peak at approximately 190-200 nm (Figure 5). The narrowing in the re-irradiated solutions is indicative of a restricted size distribution, which is in agreement with TEM images.

The results from the bactericidal test are depicted in Figure 6. Regarding the graph, a smooth effect of particle size could be evidenced, as smaller NPs present higher

antibacterial activity against *S. aureus*. Sample a2 (Table 2) exhibits the most reduced mean diameter and size dispersion, that are linked to an improved bactericidal effect. This could be due to the fact that smaller size implies higher specific surface and hence, greater reactivity and a more intense interaction with bacteria.

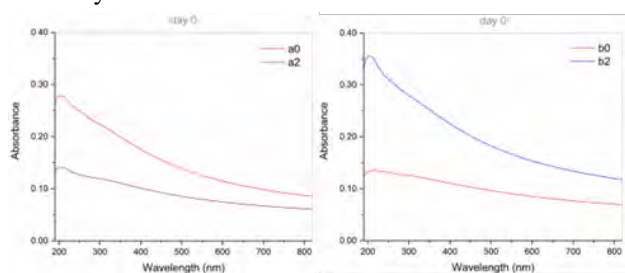


Figure 5. UV-VIS spectra of the palladium NPs: as produced (a0 and b0) and re-irradiated two times (a2 and b2)

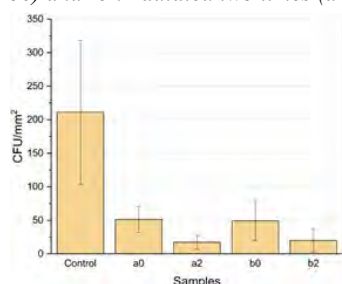


Figure 6. Bactericidal effect of Pd NPs: as produced (a0 and b0) and re-irradiated two times (a2 and b2). Values presented as mean \pm SD

4. Conclusions

Spherical Pd NPs with crystalline structure were synthesized via laser ablation technique using a nanosecond Nd:YVO₄ laser operating at 532 nm. A fragmentation system was designed and manufactured to re-irradiate the nanoparticle suspensions in order to reduce their mean size and size dispersion. Bacterial adhesion assays with *Staphylococcus aureus* were performed to analyse the effect of particle size on antimicrobial effectiveness of Pd NPs. The results acquired seem to indicate that size plays a significant role on bactericidal activity, as smaller Pd NPs with higher specific surface present increased activity against *S. aureus*, but more tests are required to completely understand their bactericidal and cytotoxic behaviour prior to their application as therapeutic treatment in humans.

5. Acknowledgements

This work was partially supported by the Government of Spain (PID2020-117900RB-I00 (MCI/AEI/FEDER, UE), EQC2018-004315-P (AEI/FEDER UE), FPU21/04488), Xunta de Galicia (ED431C 2019/23) and University of Vigo (Margarita Salas Postdoc Fellowship Program, M. Fernández Arias & P. Pou-Álvarez).

6. References

- [1] The Expert Panel on the Potential Socio-Economic Impacts of Antimicrobial Resistance in Canada, When Antibiotics Fail (2019). <https://cca-reports.ca/reports/the-potential-socio-economic-impacts-of-antimicrobial-resistance-in-canada/> (Accessed April 2022).
- [2] G. Annunziato, Strategies to overcome antimicrobial resistance (AMR) making use of non-essential target inhibitors: A review, *Int J Mol Sci.* 2019, vol. 20, pp 5844.
- [3] N. Gilbert, Antibiotic resistance marching across Europe, *Nature.* 2011, vol 22.
- [4] A.J. Huh, Y.J. Kwon, “Nanoantibiotics”: A new paradigm for treating infectious diseases using nanomaterials in the antibiotics resistant era, *Journal of Controlled Release.* 2011, vol 156, pp 128–145.
- [5] X. Zhu, A.F. Radovic-Moreno, J. Wu, R. Langer, J. Shi, Nanomedicine in the management of microbial infection - Overview and perspectives, *Nano Today.* 2014, vol 9, pp 478–498.
- [6] Saima Muzammil, Sumreen Hayat, M. Fakhar-e-Alam, Bilal Aslam et al. Nanoantibiotics: Future nanotechnologies to combat antibiotic resistance. *Frontiers In Bioscience,* 2018, vol 10, pp 352-374,
- [7] A. Karnwal, G. Kumar, G. Pant, K. Hossain, A. Ahmad, M.B. Alshammari, Perspectives on Usage of Functional Nanomaterials in Antimicrobial Therapy for Antibiotic-Resistant Bacterial Infections, *ACS Omega.* 2023, vol 8, pp 13492-13508.
- [8] M. Rai, A. Yadav, A. Gade, Silver nanoparticles as a new generation of antimicrobials, *Biotechnol Adv.* 2009, vol 27, pp 76–83.
- [9] S.B. Yaqoob, R. Adnan, R.M. Rameez Khan, M. Rashid, Gold, Silver, and Palladium Nanoparticles: A Chemical Tool for Biomedical Applications, *Front Chem.* 2020, vol 8, pp 376.
- [10] M. Yamada, M. Foote, T.W. Prow, Therapeutic gold, silver, and platinum nanoparticles, *Wiley Interdiscip Rev Nanomed Nanobiotechnol.* 2015, vol 7, pp 428–445.
- [11] A. Panáček, L. Kvítek, M. Smékalová, R. Večeřová, M. Kolář, M. Röderová, F. Dyčka, M. Šebela, R. Prucek, O. Tomanec, R. Zbořil, Bacterial resistance to silver nanoparticles and how to overcome it, *Nat Nanotechnol.* 2018, vol 13, pp 65–71.
- [12] M.B. Larosi, J. del V. García, A.R. Rodríguez, Laser Synthesis of Nanomaterials, *Nanomaterials.* 2022, vol 12, pp 2903.
- [13] E. Fazio, B. Gökce, A. De Giacomo, M. Meneghetti, et al. Nanoparticles engineering by pulsed laser ablation in liquids: Concepts and applications, *Nanomaterials.* 2020, vol 10, pp 1–50.
- [14] V. Amendola, M. Meneghetti, Laser ablation synthesis in solution and size manipulation of noble metal nanoparticles, *Physical Chemistry Chemical Physics.* 2009, vol 11, pp 3805–3821.

Improving the efficacy of magnetic nanoparticle-mediated hyperthermia using trapezoidal pulsed electromagnetic fields as an in vitro anticancer treatment in uveal melanoma and glioblastoma multiforme cell lines

Lilia Souiade^{1,*}, Javier Domingo-Diez¹, Cesar Alcaide¹, Milagros Ramos^{1,2}, José Javier Serrano Olmedo^{1,2}

¹Center for Biomedical Technology (CTB), Universidad Politécnica de Madrid (UPM), 28040, Madrid, Spain,

²Centro de Investigación Biomédica en Red para Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Spain

Summary

Magnetic hyperthermia (MHT) is a novel cancer treatment that utilizes magnetic nanoparticles (MNPs) to generate localized heat under low frequency alternating magnetic field (AMF). Recently, trapezoidal pulsed alternating magnetic fields (TPAMF) have proven their capability in enhancing the efficiency of heating in MHT as compared to the sinusoidal one.. Our study aims to compare TPAMF waveform's killing effect against the sinusoidal waveform in B16F10 and CT2A cell cultures to determine more efficient waveforms in causing cell death. For that purpose, we used SPIONs (superparamagnetic iron oxide nanoparticles) at 1 and 4 mg/ml with four different AMF waveforms: three non-sinusoidal, namely trapezoidal (TP), almost-square (TS) and triangular (TR), and the sinusoidal signal (SN). Among the explored cases, the non-harmonic signals demonstrated better efficacy than the SN one as MHT treatment. After MHT treatments, B16F10 cells incubated with SPIONs at 1 mg/ml, reached cell death rates of 31±2% and 17±5%, when TP and TS signals were applied. Whereas cell death rates reached by SN was only 0% by both TP and TS signals. In CT2A cells preincubated with SPIONs at 4 mg/ml, MHT treatments using TP and TS signals induced a cell death rate of 46±15% and 35±16%, respectively, and the MHT treatments using SN produced 14±2% and 4±2% compared to TP and TS. Based on these results, we can conclude that the application of non-harmonic signals enhances MHT treatment efficiency against tumor cells.

1. Introduction

The cancer is still the leading cause of death all over the world. In 2020, there were nearly 19.3 million new cancer cases and 10.0 million deaths per year. In 2040, it is expected to be 28.4 million cases[1]. Many studies have been directed towards hyperthermia treatment against cancer using magnetic fields, radiofrequency, and microwaves[2,3]. MHT has several advantages over traditional cancer therapies, it is less harmful to healthy cells due to the local heat release in the tumoral area after the administration of magnetic nanoparticles (MNPs). MHT can be used in combination with other treatments, such as immunotherapy[4], chemotherapy or radiation therapy, to enhance their efficacy[5,6]. The most widely

used MNPs for MHT are iron oxide nanoparticles of nanoscopic size, so that they magnetically behave as a superparamagnetic material for what they are called SPION. Previously reported preclinical hyperthermic studies have showed promising results and significant damage to cancer cells after treatment where conventional sinusoidal waveforms. The idea of enhancing heating efficiency by optimizing the signal slope has recently emerged. In[7], the heating efficiency enhancement in MHT has been experimentally explored using non-harmonic waveforms to approve that their application to excite the nano-heat agents optimizes the heating power performance, where TP and TS signals showed a higher heat generation performance compared to SN signal. The experimental results showed superior heat generation by TS that achieved over 71% of the thermal power generation efficiency compared to SN, and 45% by TP. Therefore, the optimization of hyperthermia by applying non-sinusoidal signals could allow to overcome the issues posed in both preclinical and clinical studies and permit NPs to reach their optimal alignment to optimizes the killing effect on tumor cells and the MNP internalization process. The present work was carried out to achieve actionable results to create a new way to fight cancer and enhance MHT-based anticancer treatments. For this purpose, research was conducted on B16F10 (melanoma) and CT2A (glioblastoma) subjected to different waveforms (non-harmonic waveforms against the sinusoidal waveforms) generated by an enhanced AMF generator.

2. Materials and Methods

2.1. Experimental device

The MHT device was developed in our laboratory[8], and is capable of generating non-harmonic waveforms, specifically these types: Trapezoidal-Square-TS, Trapezoidal-TP, Trapezoidal-Triangular-TT and Triangular-TR, with frequencies ranging from 100 kHz to 1 MHz, plus the SN signal. The AMF is generated through a solenoid copper coil surrounding a holder (PLA) maintained at 37°C using a water pump.

2.2. Magnetic nanoparticles

In our study, we worked with SPIONs (10.6 nm). The synthesis method of APS-SPIONs has been previously described in[9].

2.3. Cell culture

The B16F10 and the CT2A cancer cell lines were cultivated in DMEM supplemented with 10% heat-inactivated FBS, 2 mM of glutamine, 100 units/ml antibiotics of penicillin and 100 µg/ml of streptomycin, and non-essential amino-acids. The cell lines were incubated in a humidified atmosphere with 5% CO₂ at 37°C, until achieving a density around 85 to 95 % of confluence.

2.4. Experimental protocolo of in vitro Magnetic hyperthermia

To assess the MHT effect by exciting nanoparticles under an AMF, B16F10 and CT2A cells were grown on P35. The protocol involves seeding of $1,6 \times 10^4$ cells/cm² and 3×10^4 cells/cm² for B16F10 and CT2A, respectively. The cell culture dishes were divided into four groups: cells without any treatment (-AMF-MNP), cells exposed to AMF without MNP (+AMF-MNP), cells incubated with MNP (-AMF+MNP) and cells incubated with MNP and exposed to AMF (+AMF+MNP). Later, the APS-SPION were added to the culture dishes (-AMF+MNP) and (+AMF+MNP) at 1 mg/ml and 4 mg/ml, incubated for 24h and 2 h respectively. The dishes (+AMF-MNP) and (+AMF+MNP) were placed with the solenoid coil under the AC external magnetic field set at TS/SN for 2.14mT/200 kHz, TP/SN for 2.78mT/300kHz and TR/SN for 3.42mT/300kHz within an exposition period of 30 min. All experiments were initiated at 37°C.

2.5. Post Magnetic hyperthermia Calcein-AM/PI staining

Cell viability was evaluated by Calcein-AM and Propidium Iodide (PI) staining after various hyperthermic treatments applied to the B16F10 and CT2A cell lines. The cells were incubated for 15-20 minutes with Calcein-AM and PI. Observations and images acquisitions were carried out using the Fluorescence Microscope (Leica DMI 3000B, Germany). To illustrate the results obtained, the GraphPad Prism and ImageJ programs were implemented.

2.6. Cytotoxicity studies

The cytotoxicity of MNPs at 1 and 5 mg for 6h, 24h on B16F10 and CT2A cells was investigated using the XTT assay[10]. The assay was repeated in triplicate.

2.7. Statistical análisis

The data were statistically analyzed using the software package GraphPad Prism Version 8.0.1 for Windows. Data were tested by Shapiro–Wilk test for normality and presented as mean ± SEM (n=3). The differences between data groups were calculated using the ANOVA test, being considered statistical significance for $p < 0.05$.

2.8. Intracellular localization of Superparamagnetic nanoparticles

To determine the localization of APS-SPION an antibody targeted against lysosomes, anti-CD63 antibody (Abcam, USA), was used. B16F10 and CT2A cells were grown on coverslips at a density of $1,6 \times 10^4$ cells/cm² and 3×10^4 cells/cm², respectively, and incubated with 10.6 nm nanoparticles at 1 mg/ml for 24 h. Then, the cells were rinsed 3 times with PBS and fixed using 4% paraformaldehyde (Merck, Sigma Aldrich). The fixed cells were washed 3 times in PBS and blocked for 1 h with a PBS solution containing 10% horse serum (Gibco/Life Technologies, USA) and 0.25% Triton X-100 (Sigma Aldrich, USA). Later, the cells were incubated overnight at 4°C with an anti-CD63 antibody diluted to 1:500 in PBS solution with 0.25% Triton and 1% horse serum. After removing the solution with the primary antibody and rinsing, samples were incubated with anti-rabbit Alexa Fluor™ 488 secondary antibody (Abcam, USA) diluted to 1:500 in PBS solution with 0.25% Triton and 1% horse serum for 30 min at room temperature. Finally, the coverslips were washed with distilled H₂O, dried, and mounted with ProLong (Invitrogen, USA).

3. Results

3.1 In vitro cytotoxicity of APS-SPION

The cell lines B16F10 and CT2A exhibited practically similar responses to the presence of MNPs, although CT2A cells showed a lower cell viability when incubated with MNPs at 5 mg/ml for 24h and 48h compared to B16F10 cells. XTT results showed a non-significant decrease in cell viability when both cell lines were incubated with MNPs at 1 mg/ml, during all time periods tested. However, when cells were incubated with the APS-SPION at 5 mg/ml, a significant reduction in cell viability was observed after 24h and 48h of incubation in both cell lines. Whilst cell viability remains over 90% within 6h.

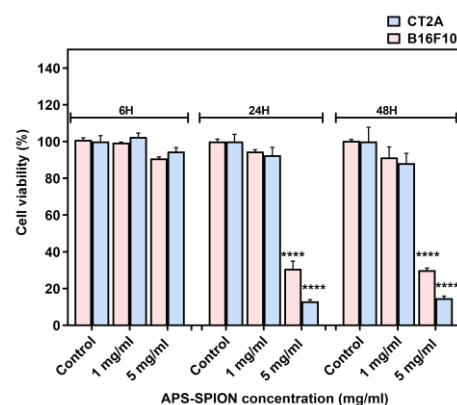


Figure 1: B16F10 and CT2A cell viability evaluated by XTT assay after incubation with APS-SPION at 1 and 5 mg/ml for 6h, 24h and 48h.

3.2 Cell uptake of APS-SPION

To assess the uptake of the nanoparticles by B16F10 and CT2A cells in our study, cells were preincubated with 10.6 nm APS-SPION for 24h and evaluated by

microscopy. APS-SPION accumulations were detected under bright field microscopy (Figure 2.A). Nanoparticles were observed outside and also inside the cells as black cytoplasmic spots of distinct sizes (arrows in Figure 7.A and D). Figure 2 shows the intracellular distribution of APS-SPION visualized in cells with a strong labeling for lysosomes (Figure 2.B and E), as determined by immunostaining with the anti-CD63 green fluorescent antibody, that specifically stains the lysosomes. The resulting image confirmed the presence of APS-SPION within the cancer cells.

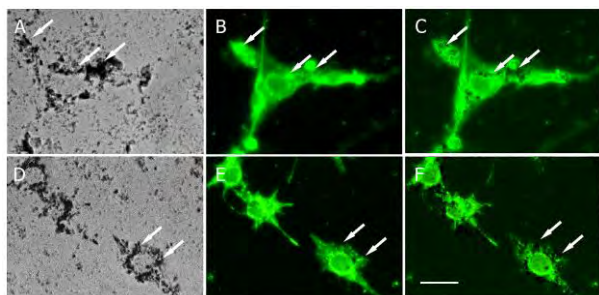


Figure 2: Intracellular localization of MNPs in B16F10 (A-C) and CT2A (D-F) cells. Cells were incubated with MNPs for 24h and then stained with an anti-CD63 to mark lysosomes. Photomicrographs show MNPs in bright field microscopy as black dots (A and D) and lysosomes by fluorescence microscopy, marked in green (B and E). Note that cells with a strong lysosome labeling show large accumulations of MNPs (arrows). Panels C and F show the merged images from panels A-B and D-E, respectively. Scale bar: 20 μ m.

3.3 Safety of the AMF applicator

In order to demonstrate that the cell damage during MHT treatment is only due to the power generated by the APS-SPION exposed to an AMF, a PLA sample holder connected to the water pump was designed to keep the temperature of the cell cultures stable at 37°C during AMF exposure. The cells were cultured in P35 and placed inside the holder. The temperature of the culture dishes was measured with an optical temperature probe immersed in the medium. A second optical probe was placed within a beaker of water at room temperature. The temperature of the sample remains stable around 37°C during the AMF exposure as shown in Figure 3.A, proving to be adequate thermal isolation. Viability of the cells exposed to AMF in the absence of MNPs, was assessed using Calcein-AM/PI staining assay and compared to their corresponding controls (cells without AMF exposure). The results from Calcein-AM/PI assay (Figure 3.B and 3.C) show that the exposure to AMF did not affect cell viability in any of the signals analyzed.

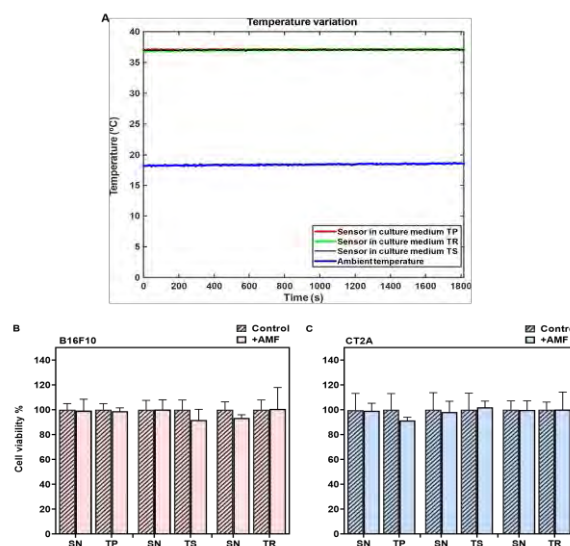


Figure 3: Safety of the MHT applicator using different signals. (A) Data of temperature measurement during hyperthermia treatment without APS-SPION. The red, black, and green curves represent the cell culture medium temperature for different non-harmonic signals. The blue curve represents the room temperature. (B) B16F10 and (C) CT2A cell viability evaluated by Calcein/PI 24h after different waveform AMF application for 30 min (+AMF) compared to cells not exposed to AMF (control).

3.4 Evaluation of the in vitro Magnetic hyperthermia treatment

The application of AMF using non-sinusoidal waveforms (TP, TS, and TR) produced in both cell lines a greater decrease in cell viability compared to the application of SN waveform. The Figure 4 represents the mortality produced by the coaction of SPION and different AMF signals, which are non-sinusoidal (TP, TS, TR) sinusoidal (SN) applied at the same frequency and amplitude, on B16F10 and CT2A. Cell mortality was evaluated 24h after MHT treatment when cells were incubated with MNPs at 1 mg/ml (Figure 4.A) and directly after treatment when MNPs were incubated at 4 mg/ml (Figure 4.B). The cell death achieved by TP, TS and TR AMF application had a superior impact among all the cases tested than the cell death produced by the corresponding SN waveform. TP and TS waveforms significantly reduced cell viability in B16F10 cells preincubated with MNPs at 1 mg/ml (Figure 4.A). Exceptionally for TR-AMF MHT treatment 24h later on B16F10 at 1mg/ml, which gave nearly equal rate of mortality with SN-AMF at 300 kHz/3.42mT. Comparatively, the cell mortality decrease observed directly after TP-AMF and TS-AMF exposure on CT2A preincubated at 4 mg/ml was highly significant over SN-AMF opposed to cell viability decrease 24h after MHT treatment in same cells preincubated with MNPs at 1mg/ml. In summary, promising results have been achieved by the applicability of the non-sinusoidal signals (TP, TS, and TR) at the best conditions (4 mg/ml over 2 h of incubation with cells) where cell mortality has been directly evaluated after exposure to AMF.

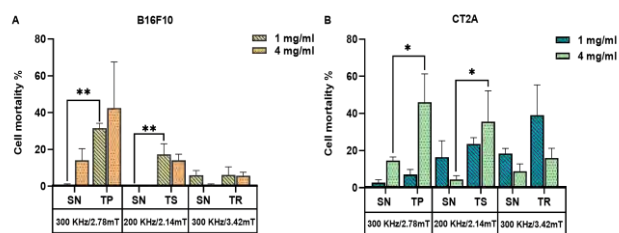


Figure 4: Cell death rates of B16F10 (A) and CT2A (B) cells pre-incubated with MNPs at 1mg/ml and 4mg/ml and exposed for 30 min to AMFs using different waveforms (TP/SN, TS/SN, TR/SN) at the indicated parameters. To evaluate cell mortality cells were stained with Calcein-AM/PI and fluorescence microscopy images were quantified using Image J software. Data represents the mean \pm SEM of three independent experiments ($n=3$). ANOVA, post hoc Fisher's LSD test; * $p<0.05$, ** $p<0.01$.

4. Discussion

Our study demonstrates an enhancement of the efficiency of MHT for killing cancer cells by non-harmonic signals. The TP, TS and TR are more effective and capable of enhancing the efficiency of MHT against tumoral cells, these results support the previous published studies by [11,12,7]. This is coherent with another experimental study that has revealed the capability of the non-sinusoidal waveforms to enhance the energy dissipated by MNPs excited by a non-sinusoidal signal as compared to the energy dissipated by MNPs excited by a conventional signal[7]. The application of these new waveforms was found to increase heat production efficiency to above 71% by the TS and 45% by the TP over the SN signal, while the TR signal trailed behind in increasing the heat efficiency and was less powerful than SN. For the in vitro tests, non-sinusoidal waveforms applicability in MHT was found to increase cell death significantly on B16F10 cell line to $31 \pm 2\%$ by TP-AMF and to $17 \pm 5\%$ by TS-AMF at 1mg/ml and 30 min of exposition to AMF. Besides, the cell death was significantly increased to $46 \pm 15\%$ in CT2A by TP-AMF for 30 min of exposition at 4 mg/ml. The TP and TS have both a strong effect on CT2A at 4 mg/ml and 30 min of exposure to AMF.

5. Conclusion

The conventional sinusoidal signals were exceptionally applied in hyperthermic research against cancer. The present study demonstrates the excellent therapeutic efficacy of non-sinusoidal signals (TP/TS/TR) for enhancing MHT treatment against cancer cells. The coaction of new non-harmonic signals in the presence of SPIONs can create an improvement for cancer treatment by producing higher impact inside cells over the harmonic one. Our results revealed that MHT using non-sinusoidal waveforms increase the tumor cell death compared to the sinusoidal signal. The synergetic effect of AMF and MNPs is the responsible of cell death without a perceptible increase in temperature to the range of mild hyperthermia inside the culture medium. We conclude that hyperthermia therapies using non-harmonic

waveforms offer substantial advantages over the conventional harmonic waveforms.

Referencias

- [1] Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* 2021;71(3):209-249. doi:10.3322/caac.21660
- [2] Mahmoudi K, Bouras A, Bozec D, Ivkov R, Hadjipanayis C. Magnetic hyperthermia therapy for the treatment of glioblastoma: a review of the therapy's history, efficacy and application in humans. *Int J Hyperth.* 2018;34(8):1316-1328. doi:10.1080/02656736.2018.1430867
- [3] Paulides MM, Dobsicek Trefna H, Curto S, Rodrigues DB. Recent technological advancements in radiofrequency- and microwave-mediated hyperthermia for enhancing drug delivery. *Adv Drug Deliv Rev.* 2020;163-164:3-18. doi:10.1016/j.addr.2020.03.004
- [4] Zhang Y, Gao X, Yan B, et al. Enhancement of CD8+ T-Cell-Mediated Tumor Immunotherapy via Magnetic Hyperthermia. *ChemMedChem.* 2022;17(2). doi:10.1002/cmdc.202100656
- [5] Dunne M, Regenold M, Allen C. Hyperthermia can alter tumor physiology and improve chemo- and radiotherapy efficacy. *Adv Drug Deliv Rev.* 2020;163-164:98-124. doi:10.1016/j.addr.2020.07.007
- [6] Peiravi M, Eslami H, Ansari M, Zare-Zardini H. Magnetic hyperthermia: Potentials and limitations. *J Indian Chem Soc.* 2022;99(1). doi:10.1016/j.jics.2021.100269
- [7] Zeinoun M, Domingo-Diez J, Rodriguez-Garcia M, et al. Enhancing magnetic hyperthermia nanoparticle heating efficiency with non-sinusoidal alternating magnetic field waveforms. *Nanomaterials.* 2021;11(12). doi:10.3390/nano11123240
- [8] Zeinoun M, Serrano D, Medina PT, Garcia O, Vasic M, Serrano-Olmedo JJ. Configurable High-Frequency Alternating Magnetic Field Generator for Nanomedical Magnetic Hyperthermia Applications. *IEEE Access.* 2021;9:105805-105816. doi:10.1109/ACCESS.2021.3099428
- [9] Luengo Y, Nardecchia S, Morales MP, Serrano MC. Different cell responses induced by exposure to maghemite nanoparticles. *Nanoscale.* 2013;5(23):11428-11437. doi:10.1039/c3nr02148c
- [10] Berridge M V., Herst PM, Tan AS. Tetrazolium dyes as tools in cell biology: New insights into their cellular reduction. *Biotechnol Annu Rev.* 2005;11(SUPPL.):127-152. doi:10.1016/S1387-2656(05)11004-7
- [11] Allia P, Barrera G, Tiberto P. Nonharmonic Driving Fields for Enhancement of Nanoparticle Heating Efficiency in Magnetic Hyperthermia. *Phys Rev Appl.* 2019;12(3). doi:10.1103/PhysRevApplied.12.034041
- [12] Barrera G, Allia P, Tiberto P. Fine tuning and optimization of magnetic hyperthermia treatments using versatile trapezoidal driving-field waveforms. *Nanoscale Adv.* 2020;2(10):4652-4664. doi:10.1039/d0na00358a

Effectiveness of gold nanorods of different sizes in photothermal therapy to eliminate melanoma and glioblastoma cells

J. Domingo-Diez¹, L. Souiade¹, Marta Sánchez¹, Vanesa Manzaneda², Andrés Guerrero², Lorena Marrodán Bretón¹, Isidoro Martínez Ramírez¹, Carmen Ramírez-Catillejo¹, J.J. Serrano Olmedo^{1,3,4}, M. Ramos^{1,3,4}

¹ Center for Biomedical Technology (CTB), Universidad Politécnica de Madrid (UPM), 28040, Madrid, Spain

² Departamento de Química Física, Universidad Complutense de Madrid, Avenida Complutense s/n, 28040, Madrid, Spain

³ Centro de Investigación Biomédica en Red para Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud, Carlos III, Spain

⁴ Departamento de Tecnología Fotónica y Bioingeniería. ETSI Telecomunicaciones, Universidad Politécnica de Madrid, 28040 Madrid, Spain

Summary

Gold nanorods are the most commonly used nanoparticle in photothermal therapy (PTT) due to their efficiency to convert the light into heat. This study aimed to investigate the efficacy of gold nanorods of different sizes (large and small) in eliminating two types of cancer cells: glioblastoma and melanoma cells. After the cytotoxic profiles of the nanoparticles were established, PTT was applied to the melanoma and glioblastoma cells with a high efficiency and mortality rate. Several methods were used to evaluate the efficiency of PTT including fluorescence, confocal or dark field microscopy, biochemical analysis, and flow cytometry. To determine cell viability and the type of cell death triggered by PTT, Calcein-propidium iodide, Annexin-V staining and dehydrogenase activity assays were performed. Our findings revealed that after PTT was applied to the cancer cells, the main cell death was apoptosis. This is advantageous as the presence of apoptotic cells can stimulate antitumoral immunity in vivo. Considering the high efficacy of these gold nanorods in PTT, large nanoparticles could be useful for biofunctionalization purposes. Large nanorods offer a greater surface area for attaching biomolecules, thereby promoting high sensitivity and specificity in recognizing target cancer cells. Additionally, large nanoparticles could also be beneficial for theragnostic applications, involving both therapy and diagnosis, due to their superior detection sensitivity.

1. Introduction

Cancer is the leading cause of almost 10 million deaths worldwide, and it is estimated that in the future, the number of new cancer cases worldwide will exceed the 19.3 million recorded in 2020, and the number of deaths will exceed 10 million [1]. Traditional treatments for cancer include surgery, chemotherapy and radiotherapy [2]. However, these treatments often lead to significant side effects. Due to the heterogeneous nature of cancer and the associated risks of current treatments, there is a need for the development of new therapeutic strategies. Innovative approaches, such as hyperthermia, are being explored and can synergize with conventional methods [3], [4].

Nanoparticle-mediated Photothermal Therapy (PTT) is a non-invasive treatment that involves increasing the temperature of the target tissue, selectively killing cancer cells. This is achieved by using nanoparticles (NPs) as mediators that convert light into heat upon laser exposure. The elevated temperature in the tumor microenvironment directly affects cancer cells, triggering necrotic and apoptotic pathways. The extent of these pathways depends on the temperature reached during the irradiation procedure (typically ranging from 41–45 °C) [5]. NPs have a tendency to passively accumulate in the tumor due to the enhanced permeability and retention effect caused by the leaky and disorganized tumor vasculature [6]. PTT also promotes the dilation of leaky tumor blood vessels, increasing their permeability and making tumor tissues more susceptible to temperature changes compared to healthy tissues [7], [8].

Gold nanorods (GNRs) are characterized by their remarkably high light-to-heat conversion efficiency, making them the most commonly used photothermal agents in PTT. The optical properties of GNRs are influenced by their size and shape, and they are governed by the phenomenon of localized surface plasmon resonance (LSPR). LSPR occurs when GNRs are exposed to light radiation, causing the plasmon to oscillate on the surface of the GNRs [9]. A widely used molecule for GNR biofunctionalization is polyethylene glycol (PEG), which extends the blood circulation time of GNRs and enhances their biocompatibility [10]. When GNRs are not biofunctionalized with PEG, their surfaces are highly recognized by circulating blood plasma proteins, which can alter the properties of the GNRs. This recognition often leads to aggregation, rendering them less effective in PTT [11]

Small GNRs (SGNRs), with approximate sizes of 40 x 10nm, are the most commonly used GNRs for PTT. However, larger GNRs may possess better optical and thermal properties than the smaller ones, making them useful for biological applications. Due to their increased

optical cross sections, large GNRs (LGNRs) with a size 40 x 65 nm exhibit greater backscattering than conventional SGNRs. Additionally, their larger surface area allows for the combination with a greater number of biomolecules, facilitating the specific targeting of cancer cells [12].

In the present study, we evaluate the effectiveness of two PEG-coated GNRs with different sizes in PTT and compare them to commercially available GNRs. The effects of LGNRs and SGNRs were evaluated on glioblastoma (CT2A) and melanoma (B16F10) cell lines undergoing PTT, while characterizing the phenomena of apoptosis and necrosis that occur.

2. Materials and Methods

2.1. Synthesis and characterization of GNRs

GNRs of different sizes with LSPR at 800 nm were prepared using a seeded growth method with some modifications as described in [13].

Functionalization of GNRs. Typically, PEG-SH (15 mg) was added under stirring to 5 mL of a freshly prepared aqueous suspension of GNRs (2 mM of Au0, 1 mM CTAB). After 1 h, the excess of free PEG-SH was removed by one centrifugation cycle (8000 rpm, 30 min). Then, the precipitate was redispersed in 5 mL of 10 mM Buffer (EBSS).

Commercial GNRs (CGNRs) with an average length and diameter of 41 x 10 nm were obtained from Nanopartz (CP12-10-808-3KPA-PBS-50-1, Nanopartz Inc, Colorado, USA), and were used for making some comparisons and to demonstrate the high quality of the synthesized GNRs.

2.2. Cell Culture

Murine glioma (CT2A) and murine melanoma (B16F10) cells were maintained in DMEM (Gibco, USA) supplemented with 10 % heat inactivated fetal bovine serum (FBS), 2mM Glutamine (Gibco, USA), and 100 units/mL penicillin and 100 µg/mL streptomycin. Cell lines were maintained at 37°C in 5% CO₂ and 95% air in a humidified atmosphere.

2.3. Cell Viability Assays

To assess cell viability following GNR incubation and laser irradiation, two methods were employed: the XTT assay [14] and calcein/propidium iodide staining.

2.4. GNR Uptake by Cells

CT2A and B16F10 cells were seeded on a cell-treated glass coverslip with a diameter of 10 mm and placed into a well in a 24-well plate. Cells were seeded at a density of 5x10⁴ cells per well. Twenty-four hours after seeding, GNRs (50 µg/mL) were added to the cells and incubated for an additional 24h. Cells were then rinsed in PBS 1x to remove unbound GNRs and fixed with 4% paraformaldehyde for 15 minutes at Room temperature. Subsequently, the cells were stained with Phalloidin-FITC (1:1000; Sigma-Aldrich) for 3h at RT in the dark. Nuclei were counterstained with Hoechst (1:500; Invitrogen). Coverslips were mounted on a glass slide with ProLongTM Glass Antifade Mountant (Life Technologies Corporation,

Eugene, OR, USA). Images were captured using a confocal microscope (Laser Scanning Confocal Stellaris Falcon; Leica Microsystem, Germany). Intracellular GNRs were visualized with a confocal reflectance mode.

2.5. Photothermal Therapy (PTT)

For PTT experiments, CT2A and B16F10 cancer cells were seeded at 7 x 10³ cells/well in a p96 multi-well microplate (Fisher Scientific, Waltham, MA, USA) in a total volume of 120 µL. After 24 h, initial DMEM with 10% FBS was removed, and GNRs at a concentration of 2 µg/mL, were added to the cells in DMEM-PR supplemented with 1% FBS. The cells were then incubated with the GNRs for an additional 24 h. Following the incubation period, the cells were washed with PBS to remove any GNRs that were not internalized by the cells. Cell viability was assessed using two different assays: XTT and Calcein/PI.

2.6. Flow Cytometry

The apoptosis induced in CT2A and B16F10 cells by PTT treatments was evaluated using flow cytometry analysis at 24 h after the treatments. Annexin-V (Bionova Científica) was used to stain apoptotic cells, while PI (Sigma-Aldrich) was used to stain necrotic cells.

Therefore, apoptotic cells were defined as Annexin-V+/PI- and necrotic cells as Annexin-V+/PI+ and Annexin-V-/PI+ cells. Ten thousand events were acquired and analyzed using the Flowjo software.

2.7. Statistical Analysis

The results are shown as the mean ± standard error of the mean from three to four experiments. The data were analyzed by single factor analysis of variance followed by the post hoc Tukey's honestly significant difference test. A significance level of P< 0.05 was chosen. GraphPad Prism version 9 (GraphPad Prism Software, San Diego, CA, USA) was used for all statistical tests.

3. Results

Characterization of GNRs

The efficacy of the synthesized large and small GNRs (LGNRs and SGNRs) for PTT was demonstrated by comparing them with commercial 10x41nm GNRs (CGNRs). The ability of GNRs to convert light into heat was initially determined by analyzing the increase in temperature produced when increasing concentrations of the different types of GNRs were irradiated with an 808 nm laser at 4.5 W for 10 min (Table 1)

Concentration (µg/mL)	Temperature Increment (°C)		
	CGNRs	SGNRs	LGNRs
10	17.5	17.7	25.6
5	12.0	9.8	18.9
2	7.3	8.7	14.5
1	7.2	5.5	10.7

Table 1: Temperature increments obtained when GNRs at the specified concentrations were irradiated by the 808nm laser at 4.5W for 10 min.

3.1. GNR uptake by the cells

Internalization of GNRs into living cells has been extensively reported [15]. The uptake of GNRs by cells is dependent on several factors such as size, shape, charge, surface chemistry, and cell type [16]. GNRs were incubated for 24h in both cancer cells, glioblastoma and melanoma cells at 2 $\mu\text{g}/\text{mL}$. The presence of GNRs inside the cells was detected by confocal microscopy using the property of these GNRs to reflect the incident light. The orthogonal projections made in an intermediate section of the cell body showed the presence of GNRs inside both cell types (red spots in Figure 4), demonstrating that LGNRs and SGNRs were internalized by both cell types.

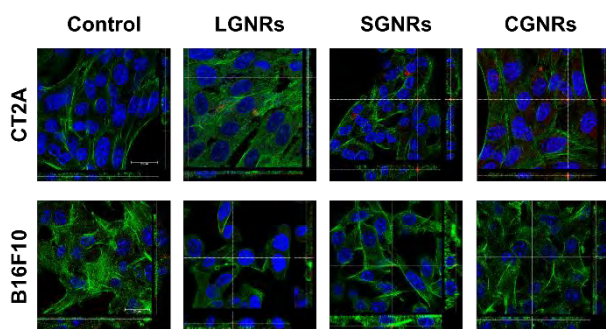


Figure 1: Confocal laser scanning microscopy images of CT2A and B16F10 cells incubated with LGNRs, SGNRs and CGNRs at 2 $\mu\text{g}/\text{ml}$ for 24 h. After fixation, cells were stained with phalloidin (green) and the nuclei were counterstained with Hoechst (blue). GNRs were detected by reflectance (red). Images show that all types of GNRs (red) were effectively internalized into the cytosol of CT2A and B16F10 cells. Scale bar: 30 μm .

3.2. Photothermal therapy at 37°C

PTT was applied to CT2A and B16F10 cells. GNRs were incubated with the cells for 24h. After the incubation period, the cells were rinsed with PBS to remove any excess or non- the unbound GNRs, the experimental conditions better mimic the situation in vivo, where GNRs would be exposed to cells and subsequently washed away by the surrounding fluid or systemic circulation. After the cells were incubated with the GNRs for 24 h and rinsed with PBS, they were subjected to laser irradiation at a wavelength of 808 nm at 4.5 W for 10 min. The PTT treatments were performed at RT and at 37 $^{\circ}\text{C}$, to better simulate physiological conditions and assess the impact of PTT treatments in a more relevant environment. After 24 h of laser irradiation, the cell viability by XTT was evaluated.

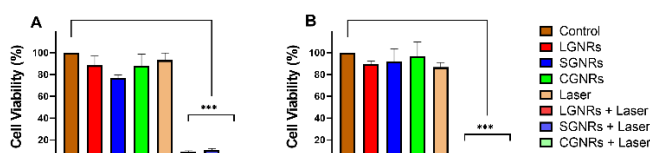


Figure 2: CT2A (A) and B16F10 (B) cell viability evaluated by the XTT assay after PTT treatments performed irradiating cells preincubated with LGNRs, SGNRs and CGNRs at 2 $\mu\text{g}/\text{ml}$ for 24 h and then irradiated with an 808 nm laser at 4.5 W for 10 min at 37 $^{\circ}$. *** $p < 0.001$ versus control.

3.3. Cell death evaluation by Flow cytometry

After hyperthermia treatments, apoptosis and necrosis can occur sequentially or simultaneously, depending on the number of NPs internalized by cells and the amount of energy dissipated by these NPs after irradiation [17]. It is important to note that necrosis often leads to the release of intracellular components into the extracellular medium, which can trigger inflammatory responses. Therefore, apoptotic cell death is preferred as a mechanism to eliminate cancer cells. Cell death produced in CT2A cells preincubated with LGNRs and SGNRs 24 h after laser application occurred preferentially by apoptosis rather than necrosis (Figure 7A). Similar necrotic rates were observed in CT2A cells 24 h after laser irradiation in the presence of LGNRs and SGNRs (Figure 7A). In B16F10 cells, apoptosis was also found to be the predominant mechanism of cell death 24 h after laser irradiation. The apoptotic pathway was the predominant method of eliminating these cells when using both types of GNRs, although LGNRs produced a slightly higher percentage of necrotic cells compared to SGNRs. (Figure 7B).

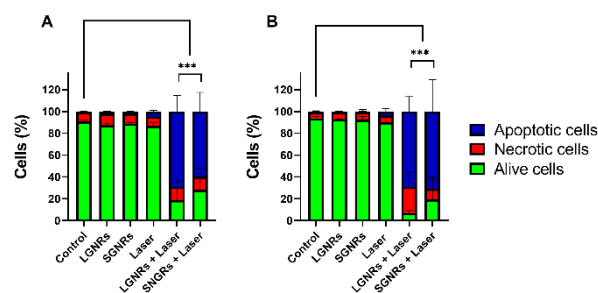


Figure 3: Analysis of cell death by flow cytometry (Annexin V and PI staining) in CT2A (A) and B16F10 (B) cells 24 h after PTT treatments. For PTT treatments, cells were preincubated with LGNRs and SGNRs at 2 $\mu\text{g}/\text{ml}$ for 24 h and then irradiated with an 808 nm laser at 4.5 W for 30 min at RT. Apoptotic cells were defined as Annexin-V+/PI- cells and necrotic cells as Annexin-V-/PI+ and Annexin-V+/PI+ cells. *** $p < 0.001$ (between apoptotic cell groups versus control).

4. Discussion

Localized physical treatments, such as nanoparticle-based PTT, utilize hyperthermia to induce damage and destroy cancer cells and tumor tissues. Unlike other methods, like chemotherapy, PTT does not typically lead to the development of resistance in cells [18] By using controlled heating via light-induced hyperthermia, NP-mediated-PTT can effectively target and eliminate cancer cells without the risk of resistance development.

In this study, LGNRs and SGNRs were internalized by both, glioblastoma and melanoma cells, as demonstrated by confocal fluorescence microscopy (Figure 4). It was observed that the binding and uptake of GNRs exhibited slight variations depending on the cell type (as shown in Figure 5). B16F10 cells exhibited a slightly higher uptake of GNRs compared to CT2A cells. However, no significant differences in uptake were observed between LGNRs and SGNRs in both cell lines. These results are consistent with

previous findings that have described variations in the internalization ability of particles in different cell lines [19].

Regarding the PTT at 37 °C, almost all the cancer cells, in both glioblastoma and melanoma cell lines, were eliminated after being irradiated at 4.5W for 10 minutes. There were no differences between the three gold nanoparticles used in PTT.

Apoptosis and necrosis are the two most frequently observed cell death pathways following PTT treatments. The processes of necrosis and apoptosis in response to PTT depend on the power of the laser and the temperature attained within the cancerous tissues during irradiation [5] with low-energy irradiation inducing apoptosis rather than necrosis. The effectiveness of LGNRs and SGNRs in promoting apoptotic cell death rather than necrosis (Figure 3) demonstrated that these nanoparticles are optimal for PTT treatments.

5. Conclusion

In this study, we have demonstrated the effectiveness of GNRs of two different sizes in promoting cell death through PTT in glioblastoma and melanoma cell lines. The use of GNR-based PTT treatments at a physiological temperature of 37 °C resulted in nearly complete efficacy, leading to the elimination of glioblastoma cells and, notably, melanoma cells. This is particularly significant, considering that melanoma is known for its high resistance to conventional treatments like chemotherapy. Moreover, we found that the primary cell death pathway activated by these GNR-mediated PTT treatments is apoptosis, which is favorable as the presence of apoptotic cells can stimulate antitumoral immunity *in vivo*. Given the high efficacy of both types of GNRs in PTT, LGNRs can be useful for biofunctionalization due to their higher surface area, which can be useful for attaching biomolecules and promoting high sensitivity and specificity in recognizing target cancer cells or tissues.

References

- Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN. Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249.
- Vines, J.B.; Yoon, J.-H.; Ryu, N.-E.; Lim, D.-J.; Park, H. Gold Nanoparticles for Photothermal Cancer Therapy. *Front. Chem.* **2019**, *7*, 167.
- Wei, G.; Wang, Y.; Yang, G.; Wang, Y.; Ju, R. Recent progress in nanomedicine for enhanced cancer chemotherapy. *Theranostics*, **2021**, *11*, 6370–6392.
- Kong, C.; Chen, X. Combined Photodynamic and Photothermal Therapy and Immunotherapy for Cancer Treatment: A Review. *Int. J. Nanomed.* **2022**, *17*, 6427–6446.
- Gupta, N.; Malviya, R. Understanding and advancement in gold nanoparticle targeted photothermal therapy of cancer. *Biochim., Et Biophys. Acta BBA-Rev. Cancer* **2021**, 1875, 188532.
- Riley, R.S.; Day, E.S. Gold nanoparticle-mediated photothermal therapy: Applications and opportunities for multimodal cancer treatment. *WIREs Nanomed. Nanobiotechnol.* **2017**, *9*, e1449.
- Taylor, M.L.; Wilson, R.E.; Amrhein, K.D.; Huang, X. Gold Nanorod-Assisted Photothermal Therapy and Improvement Strategies. *Bioengineering* **2022**, *9*, 200.
- Hou, Y.; Yang, X.; Liu, R.; Zhao, D.; Guo, C.; Zhu, A.; Wen, M.; Liu, Z.; Qu, G.; Meng, H. Pathological Mechanism of Photodynamic Therapy and Photothermal Therapy Based on Nanoparticles. *Int. J. Nanomed.* **2020**, *15*, 6827–6838.
- Ahmad, T.; Sarwar, R.; Iqbal, A.; Bashir, U.; Farooq, U.; Halim, S.A.; Khan, A.; Al-Harrasi, A. Recent advances in combinatorial cancer therapy via multifunctionalized gold nanoparticles. *Nanomedicine* **2020**, *15*, 1221–1237.
- Thi, T.T.H.; Pilkington, E.H.; Nguyen, D.H.; Lee, J.S.; Park, K.D.; Truong, N.P. The Importance of Poly (Ethylene Glycol) Alternatives for Overcoming PEG Immunogenicity in Drug Delivery and Bioconjugation. *Polymers* **2020**, *12*, 298.
- Dobrovolskaia, M.A.; Patri, A.K.; Zheng, J.; Clogston, J.D.; Ayub, N.; Aggarwal, P.; Neun, B.W.; Hall, J.B.; McNeil, S.E. Interaction of colloidal gold nanoparticles with human blood: Effects on particle size and analysis of plasma protein binding profiles. *Nanomed. Nanotechnol. Biol. Med.* **2009**, *5*, 106–117.
- SoRelle, E.D.; Liba, O.; Hussain, Z.; Gambhir, M.; de la Zerda, A. Biofunctionalization of Large Gold Nanorods Realizes Ultrahigh-Sensitivity Optical Imaging Agents. *Langmuir* **2015**, *31*, 12339–12347.
- González-Rubio, S.; Salgado, C.; Manzaneda-González, V.; Muñoz-Úbeda, M.; Ahijado-Guzmán, R.; Natale, P.; Almedio-Vedia, V.G.; Junquera, E.; Barcina, J.O.; Ferrer, I.; et al. Tunable gold nanorod/NAO conjugates for selective drug delivery in mitochondria-targeted cancer therapy. *Nanoscale* **2022**, *14*, 8028–8040.
- Berridge, M.V.; Herst, P.M.; Tan, A.S. Tetrazolium dyes as tools in cell biology: New insights into their cellular reduction. *Biotechnol. Annu. Rev.* **2005**, *11*, 127–152.
- Fernandez Cabada, T.; Sanchez Lopez de Pablo, C.; Martinez Serrano, A.; del Pozo Guerrero, F.; Serrano Olmedo, J.J.; Ramos Gomez, M. Induction of cell death in a glioblastoma line by hyperthermic therapy based on gold nanorods. *Int. J. Nanomed.* **2012**, *7*, 1511–1523.
- Robinson, R.; Gerlach, W.; Ghandehari, H. Comparative effect of gold nanorods and nanocages for prostate tumor hyperthermia. *J. Control. Release* **2015**, *220*, 245–252.
- Qin, Z.; Bischof, J.C. Thermophysical and biological responses of gold nanoparticle laser heating. *Chem. Soc. Rev.* **2011**, *41*, 1191–1217.
- Chu, K.F.; Dupuy, D.E. Thermal ablation of tumours: Biological mechanisms and advances in therapy. *Nat. Rev. Cancer* **2014**, *14*, 199–208.
- White, B.E.; White, M.K.; Alsudani, Z.A.N.; Watanabe, F.; Biris, A.S.; Ali, N. Cellular Uptake of Gold Nanorods in Breast Cancer Cell Lines. *Nanomaterials* **2022**, *12*, 937.

Funding: This study was partially funded by the Ministerio de Ciencia, Innovación y Universidades of Spain, ref.: PGC2018-097531-B-I00 and PDC2022-133028-I00, funded by the European Union-NextGenerationEU.

Realidad Virtual

Miércoles 22 15:00 - 16:30
Aula PB5

HC3DU: Un nuevo plugin para ofrecer retroalimentación háptica en simulaciones de Realidad Virtual

M. Rodríguez-Matesanz¹, E.J. Gómez^{1,2}, P. Sánchez-González^{1,2}

¹ Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España, {manuel.rodriguez.matesanz, enriquejavier.gomez, p.sanchez}@upm.es,

² Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, 28029 Madrid, Spain

Resumen

La sensación háptica es la capacidad de percibir la información táctil referente a la textura, calor y presión de un objeto con la piel, así como la fuerza ejercida sobre dicho objeto. La retroalimentación de fuerzas en Realidad Virtual tiene aún que explorarse debido a los altos costes de los dispositivos hápticos y a la escasez de software que permita la implementación de forma sencilla en los motores de desarrollo más actuales. En este trabajo se propone Haptic Collisions 3D for Unity (HC3DU), un plugin para Unity Engine basado en el framework Chai3D que permite retroalimentar al usuario con sensación de tacto mediante la colisión entre los objetos virtuales y el manejador del dispositivo Falcon Novint en simulaciones de Realidad Virtual.

1. Introducción

La palabra “háptico” deriva del griego *haptesthai*, que significa “tocar” o “relacionado con la sensación de tacto” [1]. Por tanto, la sensación háptica es la capacidad de percibir la información táctil referente a la textura, calor y presión de un objeto con la piel, así como de sentir la resistencia de los objetos ante la fuerza ejercida por la piel [2]. Mediante el uso de hardware conocido como “interfaces hápticas”, se puede ofrecer al usuario la información sobre el peso, inercia, dureza y suavidad de la geometría de estos objetos [3]. Estas propiedades influyen en la percepción final de los elementos del entorno con los que el usuario interactúa, favoreciendo la inmersión y sensación de presencia en los entornos en entornos virtuales [4]. Para obtener esta sensación se suele hacer uso de dispositivos de tecnologías sensitivas (*wearables*) [5]. Sin embargo, un estudio reciente sobre productos comerciales que proveen retroalimentación háptica para Realidad Virtual (RV) concluye que muy pocos de éstos acaban siendo disponibles para el público general y con ellos, las aplicaciones desarrolladas que hacen uso de este hardware [6]. Por otro lado, los dispositivos hápticos (DH) de escritorio que aplican pequeñas fuerzas a través de una conexión mecánica como por ejemplo un lápiz óptico, o un manejador en la mano del usuario, trataron de solucionar esto mediante la comercialización de dispositivos de bajo coste (por debajo de los 250\$ por unidad) y de fácil acceso para el público general, tales como el Falcon Novint o Phantom OMNI [7]. Esto ha permitido el desarrollo de APIs, aplicaciones y herramientas centradas en la formación médica que aprovechen de esta tecnología para dar una mayor sensación de realismo y reducir la curva de aprendizaje [8]. Así, se beneficia a los estudiantes y

profesionales, que requieren de un proceso adecuado y extenso de aprendizaje antes de poder realizar procedimientos quirúrgicos correctamente. Los avances en poder de computación han permitido que las simulaciones basadas nuevas tecnologías de Realidad Extendida (RE): Realidad Aumentada (RA), Realidad Mixta (RM) y RV; sirvan como apoyo en el entrenamiento de habilidades médicas con escenarios y comportamientos más complejos mientras se recogen datos objetivos sobre la realización [9]. A esto se le suma la reducción de costes y su consecuente popularización de los dispositivos de RV en el campo del entretenimiento en un ámbito general, que ha fomentado el interés en el desarrollo de aplicaciones didácticas en la medicina con simuladores quirúrgicos empleados en neurocirugía, cirugía laparoscópica o endoscópica; simuladores para la adquisición de destrezas técnicas específicas como la reanimación cardio pulmonar (RCP) o la práctica de otras emergencias sanitarias y habilidades no técnicas.

Este artículo presenta un plugin multiplataforma para Unity Engine que permite ofrecer retroalimentación háptica en las simulaciones de RV que hagan uso de los DH de bajo coste creadas con este motor para Windows, MacOS o Linux.

2. Renderizado Háptico

El renderizado háptico es el proceso que permite mostrar sensación de fuerza y tacto mediante un DH [10]. Los algoritmos del renderizado háptico funcionan mediante la detección de la posición y rotación del efector final o mango del DH en el espacio de trabajo de este. Esta matriz de transformación se traslada al espacio virtual y se comprueba la detección de colisiones del mango virtual con el resto de los objetos del entorno háptico (EH). Posteriormente, se procesa una respuesta ante estas colisiones y se ofrece una vibración y/o resistencia en el DH, ofreciendo la mencionada retroalimentación háptica [11]. De la misma manera que en el renderizado gráfico donde la imagen se compone de un modelo basado en la posición de una cámara virtual y su proyección en el espacio, el procesamiento háptico de una fuerza se basa en un modelo basado en la posición de la herramienta (el mango virtual en el EH) y sus interacciones. Es importante destacar que para que el computo de las colisiones del EH se realice de forma estable y efectiva, la frecuencia del bucle háptico debe mantenerse por encima de los 1000 Hz [12]. Esto supone un problema en algunos casos pues, un

buen algoritmo de renderizado háptico que permita una retroalimentación háptica realista se acabará enfrentando a un alto consumo de memoria y uso la CPU. Varias APIs hápticas multiplataforma de escritorio (Windows, MacOS y Linux) de alto nivel han sido desarrolladas intentando solventar de la mejor forma este problema [13]. De estas APIs, y al igual que otros desarrollos de plugins previos como el de *HPGE* por parte de Nicolo Balzarotti y Gabriel Baud-Bovy [14], se ha escogido Chai3D para el desarrollo de HC3DU pues ofrece soporte a múltiples dispositivos mediante el SDK de Force Dimensions, entre ellos el ya mencionado Falcon Novint; es gratuito y de código abierto (con licencia BSD); y utiliza una estructura similar a los motores de desarrollo utilizados para las generación de simulaciones de RV como Unity Engine.

3. Metodología

Se ha implementado en dos módulos separados: El módulo de físicas, que se encarga del renderizado háptico y procesa la respuesta a las colisiones del manejador del dispositivo con los objetos de la escena 3D y el módulo gráfico, que se encarga de organizar en Unity Engine las estructura y visualización de los modelos 3D y la aplicación general. El módulo de físicas se ha implementado compilado del framework multiplataforma Chai3D (C++) y todas sus dependencias en una librería dinámica, de forma que se puedan llamar a sus métodos desde cualquier aplicación (no solo desde Unity Engine). Por otro lado, el módulo gráfico se compone de una serie de scripts en C# que interactúan con los componentes de Unity y a su vez, con los del módulo de físicas para que trabajen como uno, aunque desarrollen funciones independientes.

3.1. Características generales

HC3DU está diseñado como un plugin que se pueda utilizar de forma *plug-and-play* en cualquier aplicación desarrollada con Unity Engine. Una vez importado como UnityPackage (paquete de Asset de Unity), el usuario solo debe añadir los componentes necesarios a los objetos que desee tocar retroalimentación háptica o sentir sus propiedades táctiles (suavidad de la geometría, nivel de textura, fricción y dureza). Los objetos interactuables hápticamente (Touchables en inglés) pueden establecer desde el propio editor de Unity cómo definen las cajas de colisión (en inglés Axis Aligned Bounding Boxes o AABB) según se adapten mejor a las necesidades del usuario. Por un lado, puede seleccionar primitivas simples como Caja, Esfera, Cilindro, Plano o Toroide. Sin embargo, si desea una mayor aproximación en las colisiones con la malla 3D del objeto, pueden utilizarse los vértices de esta o de una combinación de submallas para crear una malla de colisión personalizada a costa de más memoria y uso de CPU en el procesamiento. Al modificar los valores de estos objetos, se establecen automáticamente en el renderizado háptico, por lo que los cambios son instantáneos.

Las texturas hápticas son independientes a la textura renderizada por el motor gráfico. Además de los valores de las propiedades un objeto “tocable”, la vibración aplicada al manejador del DH se procesa a partir de una textura de color o en caso de querer especificarlo, de una textura de

normales que defina el mapeado del relieve de la geometría. Estas texturas deberían coincidir con las texturas aplicadas al material gráfico para no causar disonancia ojo-tacto. Para facilitar al usuario establecer la textura háptica a un objeto virtual, se ha habilitado una sección donde solo debe arrastrar la imagen que desea utilizar. Se ha denominado “Albedo” ya que actúa como textura difusa. Si el usuario no asigna un mapa de normales, se utiliza el Albedo (Figura 1).

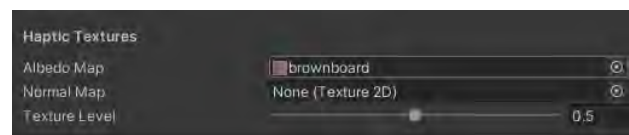


Figura 1. Texturas hápticas en Touchable Objects

Además, cualquier objeto de la escena puede ser controlable por el DH. Solo debe añadirse al objeto deseado el script de tipo Controlable (*ControllableObject*) o sobrescribir la clase con un script personalizado que herede de ésta. Sin embargo, para aplicar la retroalimentación de fuerza en el propio mecanismo del DH, el usuario debe tener al menos una herramienta háptica (*HapticTool*) añadida a un objeto de la escena. Toda la configuración se inicializa por defecto según el objeto al que ha sido añadido el script, aunque los valores del tamaño, espacio de trabajo, y propiedades táctiles de la herramienta pueden cambiarse en cualquier momento desde el inspector del editor de Unity (Figura 2).

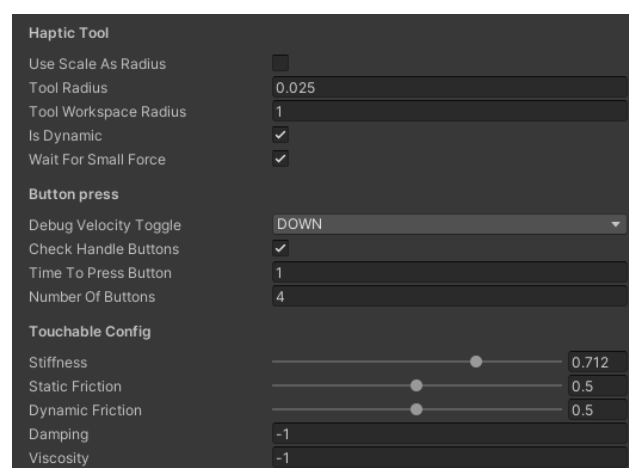


Figura 2. Propiedades de la herramienta háptica (HapticTool)

3.2. Integración del módulo gráfico y háptico

Cuando HC3DU crea un objeto en el módulo gráfico (Unity), se crea también un objeto virtual en el mundo de Chai3D con las mismas características y las almacena en un diccionario con un identificador personalizado. Esta duplicación permite mantener los dos módulos separados, pero poder actualizar los valores de uno u otro cuando es necesario, manteniéndolos en hilos separados a través de algoritmos de *multi-threading*. El usuario puede escoger qué módulo es el predominante, actualizando los valores del otro módulo a partir de los de éste. De esta forma, se

puede utilizar el motor de físicas de Unity (*PhysX* de Nvidia o *Havok Game Dynamics* según esté configurado) y sus sólidos rígidos que establecen posiciones; o que sea Chai3D el que establezca las posiciones y éstas se apliquen a las matrices de transformación de Unity. Esto ya supone una diferencia con respecto a soluciones previas como HPGE (Sección 2), que únicamente trasladan las posiciones de Unity a Chai3D, por lo que no reacciona de vuelta en caso de, por ejemplo, deformaciones de la malla a través de las colisiones. Tal y como se puede ver en la Figura 4, Unity utiliza un sistema de coordenadas basado en la regla de la mano izquierda con eje Y-Arriba, mientras que Chai3D hace uso de un sistema de coordenadas basado en la mano derecha con Z-Arriba. Lo que implica que se debe hacer una transformación de coordenadas de las posiciones, rotaciones y escalas; entre ambos módulos cada vez que se actualizan los valores.

Los objetos tocables por defecto están diseñados para escenas estáticas donde éstos no se mueven y, por tanto, que son más fáciles de procesar ya que únicamente dependen del movimiento de la herramienta háptica contra estos objetos y el procesamiento de la colisión. Sin embargo, HC3DU ofrece soporte para objetos en movimiento tanto animados como no animados. Los objetos en movimiento no animados son aquellos que no varían la posición local de los vértices de la malla 3D, pero sí la posición global del objeto en el espacio. Al no variar la malla de colisión, el computo de la colisión se hace comprobando los movimientos lineales de la AABB de la herramienta háptica con las de los objetos en movimiento. Finalmente, aquellos objetos animados son los que aplican variación de posición de los vértices de la malla, modificando, por tanto, las posiciones y rotaciones de los triángulos que componen su malla. En estos casos, HC3DU permite reconstruir en tiempo real la malla de colisión, aunque esto no es recomendable en mallas de alto poligonaje pues provoca picos de uso de CPU y bajadas de frames tanto en el bucle del módulo gráfico (que habitualmente mantiene una frecuencia de 60Hz) como en el háptico (de 1000Hz). Lo recomendable en los casos de modelos 3D con miles de polígonos es utilizar mallas de colisión diezmadas para el módulo háptico. Es decir, utilizar mallas que mantengan el volumen a costa de la suavidad en la geometría (por el número reducido de polígonos) y que sean más rápidas de procesar tal y como se ve en la Figura 3. Otra opción es utilizar primitivas simples para el módulo háptico (cilindros o cajas) mientras que el módulo háptico renderiza visualmente las mallas de alta definición.

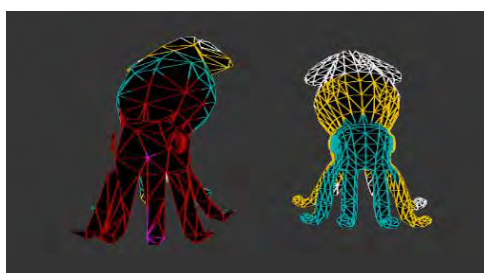


Figura 3. Malla optimizada de colisión con 150 polígonos (izquierda) vs. Malla sin optimizar con 3500 polígonos (derecha)

4. Resultados

HC3DU está diseñado para que, aunque hay dos módulos que funcionan a frecuencias diferentes de forma independiente en hilos paralelos, el conjunto actúe como uno a la hora de sincronizar los valores necesarios en los objetos, siguiendo el flujo de trabajo mostrado en la Figura 4.

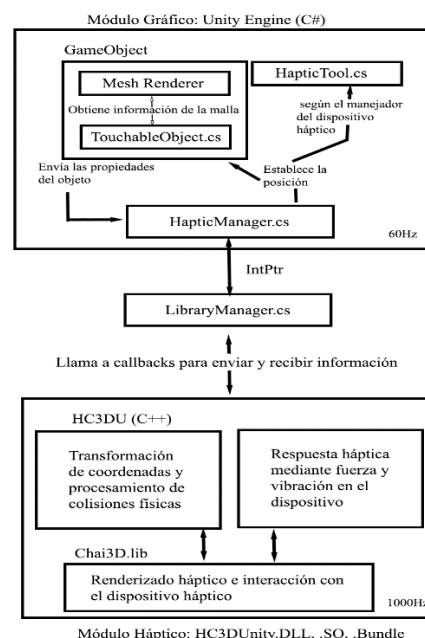


Figura 4. Flujo entre Unity y Chai3D a través de HC3DU

El gestor *HapticManager.cs* se ha utilizado como centro del servicio y es accesible por todos los *TouchableObjects* a través de una referencia estática. Este gestor llama otro gestor genérico llamado *LibraryManager.cs*, que hace de intermediario entre ambos módulos. Su función es, por un lado, alocar en memoria el puntero perteneciente a la librería y guardar la referencia a su puntero en memoria. El puntero permitirá acceder posteriormente a cualquier método de la librería de C++ a través de *callbacks* (punteros a función) a bajo nivel. De esta forma, cualquier objeto puede acceder utilizando la API de alto nivel de Unity y acceder mediante estos gestores, a toda la API de bajo nivel encargada de la simulación de físicas en el renderizado háptico.

Para poner a prueba el funcionamiento de HC3DU, se ha creado un entorno virtual en el que el usuario debe controlar mediante el dispositivo háptico Falcon Novint un endoscopio que se mueve y rota alrededor de una rodilla (Figura 5). El usuario debe encontrar varios tumores de color verde en el menor tiempo posible. El botón central del manejador cambia el modo de control de movimiento a rotación y viceversa, permitiendo los dos controles de forma sencilla. Una vez se acerca al tumor, debe esperar unos instantes y éste desaparece, ofreciendo una puntuación por el objetivo cumplido. Se termina el ejercicio cuando se encuentran todos los tumores que aparecen aleatoriamente. Al tocar los huesos se hace uso tanto de la retroalimentación de fuerza como táctil

mediante una textura háptica. Además, para comprobar la eficacia con RV inmersiva, se hace uso del plugin OpenXR que permite renderizar la imagen de Unity en headsets de RV que estén conectados al PC. En este caso, Meta Quest 2 y HTC Vive Pro.

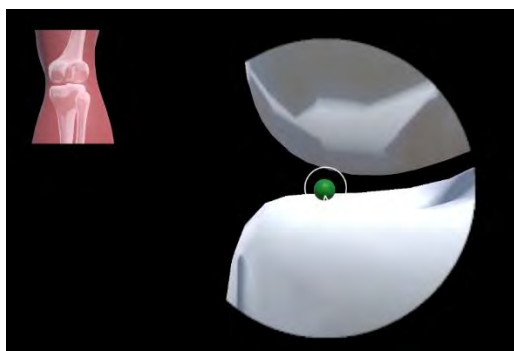


Figura 5. Ejercicio de ejemplo con HC3DU

5. Conclusiones

Este trabajo ha consistido en desarrollar un plugin que se pueda utilizar de manera sencilla e intuitiva en simulaciones de Realidad Virtual desarrolladas en Unity Engine para ofrecer retroalimentación háptica con el dispositivo Falcon Novint. Haptic Collisions 3D for Unity trabaja haciendo uso del framework Chai3D como librería y sin dependencias externas, que permite controlar el dispositivo y permitirle mostrar las fuerzas procesadas en el manejador a través de la resistencia y vibración de la parte mecánica del dispositivo.

Se puede concluir que se ha conseguido con éxito el desarrollo del plugin, que puede ser utilizado siempre que se disponga de un ordenador al que conectar el Falcon pues únicamente está disponible para plataformas de escritorio. Sin embargo, permitir reconocer formas, geometría y propiedades táctiles de objetos virtuales, ofrece un incremento del nivel de realismo e inmersión al añadir un sentido más a las simulaciones de Realidad Virtual, ya acostumbradas a los sentidos audiovisuales.

Referencias

- [1] B. Hannaford and A. M. Okamura, “Haptics,” *Springer Handbooks*, pp. 1063–1084, 2016, doi: 10.1007/978-3-319-32552-1_42/FIGURES/11.
- [2] L. Rincon-Gonzalez, J. P. Warren, D. M. Meller, and S. Helms Tillery, “Haptic interaction of touch and proprioception: Implications for neuroprosthetics,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, no. 5, pp. 490–500, Oct. 2011, doi: 10.1109/TNSRE.2011.2166808.
- [3] “(PDF) Haptic Feedback for Virtual Reality.” https://www.researchgate.net/publication/2356993_Haptic_Feedback_for_Virtual_Reality (accessed Jun. 14, 2023).
- [4] J. Kreimeier *et al.*, “Evaluation of Dierent Types of Haptic Feedback Influencing the Task-based Presence and Performance in Virtual Reality,” *ACM International Conference Proceeding Series*, vol. 19, pp. 289–298, Jun. 2019, doi: 10.1145/3316782.3321536.
- [5] C. Pacchierotti, S. Sinclair, M. Solazzi, A. Frisoli, V. Hayward, and D. Prattichizzo, “Wearable haptic systems for the fingertip and the hand: Taxonomy, review, and perspectives,” *IEEE Trans Haptics*, vol. 10, no. 4, pp. 580–600, 2017, doi: 10.1109/TOH.2017.2689006.
- [6] H. Borgmann and WFB Wirtschaftsförderung Bremen GmbH Abt. Messe Bremen, “Actuator 18 16th International Conference on New Actuators, 25-27 June 2018, Bremen, Germany : conference proceedings”, Accessed: Jun. 14, 2023. [Online]. Available: https://kuleuven.limo.libis.be/discovery/fulldisplay/lirias1992235/32KUL_KUL:Lirias
- [7] S. Martin and N. Hillier, “Characterisation of the Novint Falcon Haptic Device for Application as a Robot Manipulator”.
- [8] D. Escobar-Castillejos, J. Noguez, L. Neri, A. Magana, and B. Benes, “A Review of Simulators with Haptic Devices for Medical Training,” *J Med Syst*, vol. 40, no. 4, pp. 1–22, Apr. 2016, doi: 10.1007/S10916-016-0459-8/TABLES/2.
- [9] J. Pottle, “Virtual reality and the transformation of medical education,” *Future Healthc J*, vol. 6, no. 3, p. 181, Oct. 2019, doi: 10.7861/FHJ.2019-0036.
- [10] M. A. Srinivasan and C. Basdogan, “Haptics in virtual environments: Taxonomy, research status, and challenges,” *Comput Graph*, vol. 21, no. 4, pp. 393–404, Jul. 1997, doi: 10.1016/S0097-8493(97)00030-7.
- [11] “(PDF) Haptic Rendering in Virtual Environments.” https://www.researchgate.net/publication/2380464_Haptic_Rendering_in_Virtual_Environments (accessed Jun. 14, 2023).
- [12] S. D. Laycock and A. M. Day, “A Survey of Haptic Rendering Techniques,” *Computer Graphics Forum*, vol. 26, no. 1, pp. 50–65, Mar. 2007, doi: 10.1111/J.1467-8659.2007.00945.X.
- [13] P. K. Supervised and P. Kmoch, “Overview of current developments in haptic APIs”.
- [14] N. Balzarotti and G. Baud-Bovy, “HPGE: An haptic plugin for game engines,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11385 LNCS, pp. 330–339, 2019, doi: 10.1007/978-3-030-11548-7_31/FIGURES/4.

Diseño e implementación de un planificador quirúrgico para cirugía colorrectal junto con un visualizador basado en realidad mixta

C. Lobato Gómez ¹, L. Salazar Carrasco ¹, J.A. Sánchez Margallo ^{1*}, F.M. Sánchez Margallo ¹

¹ Centro de Cirugía de Mínima Invasión Jesús Usón, {clobato, lsalazar, jasanchez, msanchez}@ccmijesususon.com.

* jasanchez@ccmijesususon.com

Resumen

En el contexto de la cirugía actual, los planificadores quirúrgicos han emergido como herramientas esenciales para garantizar el éxito y la seguridad de los procedimientos quirúrgicos. La planificación quirúrgica es un proceso multifacético que involucra la evaluación exhaustiva de datos clínicos, imágenes radiológicas y la anatomía específica del paciente. Este trabajo pretende aportar una herramienta de segmentación semi-automática y planificación para cirugía colorrectal, junto con un visualizador de realidad mixta, que permita al cirujano optimizar los tiempos de segmentación y mejorar su experiencia durante el proceso de la planificación de los procedimientos quirúrgicos y la identificación de estructuras anatómicas y lesiones, con el fin último de mejorar la calidad asistencial.

1. Introducción

El cáncer colorrectal (CCR) ocupa el tercer lugar en cuanto a neoplasias malignas más frecuentemente diagnosticadas en todo el mundo y se sitúa como la cuarta causa más importante de fallecimientos debido al cáncer [1].

La cirugía sigue siendo la opción principal de tratamiento tanto para casos de cáncer colorrectal con fines curativos como para aquellos en los que se busca aliviar los síntomas en una etapa avanzada de la enfermedad. La reducción de riesgos y la mejora de los resultados son aspectos cruciales que los planificadores quirúrgicos aportan al proceso quirúrgico, por lo que se vuelven es-

pecialmente relevantes, dado que aproximadamente el 80 % de los casos diagnosticados de cáncer de colon requieren de ella. [2].

La tecnología actual, en lo referente al modelado 3D y la realidad aumentada, ha revolucionado la forma en que los cirujanos planifican y ejecutan procedimientos. Estas innovaciones proporcionan una visualización aún más detallada de las estructuras anatómicas, lo que facilita la personalización de enfoques quirúrgicos y la adaptación a las variaciones anatómicas individuales. Este trabajo aúna ambas técnicas con la intención de aportar una herramienta de utilidad que mejore el proceso de la planificación quirúrgica. Para ello, se presenta el diseño e implementación de un planificador quirúrgico para cirugía colorrectal, junto con un método de segmentación semi-automática del colon, así como un sistema de visualización e interacción con el modelo basado en realidad mixta.

2. Materiales y métodos

A continuación, vamos a presentar los diferentes componentes principales que forman parte del planificador quirúrgico, detallando la metodología empleada para su desarrollo.

2.1. Segmentación de imagen

La segmentación de imágenes médicas es un proceso que habitualmente se lleva a cabo de manera manual, lo que



Figura 1: Paso a paso del proceso de segmentación. Colocación automática de las semillas (A). Localización del gas externo al paciente (B). Localización del gas interno al paciente (C).

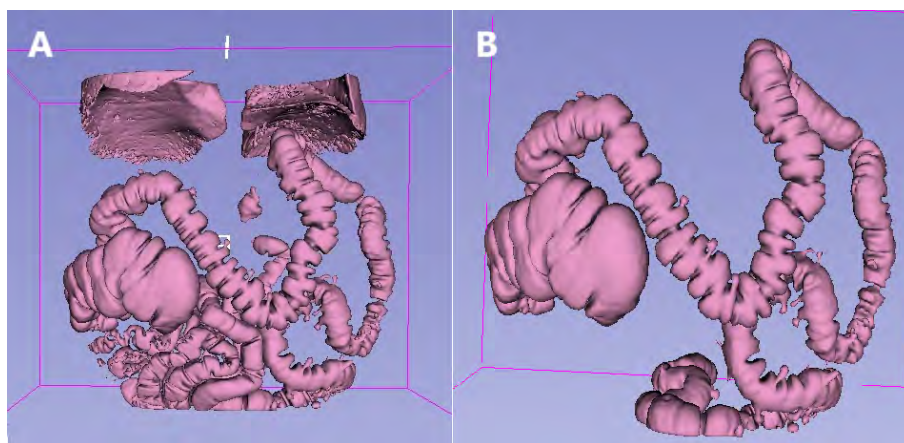


Figura 2: Antes y después del último proceso. Resultado tras el primer procesado, donde se puede apreciar que el threshold incluye aire de intestino delgado y pulmones. (A). Resultado final tras el segundo procesado con crecimiento de regiones (B).

implica una labor meticulosa y de larga duración.

Con el fin de mejorar la sencillez y eficiencia en el proceso de modelado, se ha creado una herramienta de segmentación semi-automática, en este caso enfocada en la estructura anatómica del colon. Para llevar a cabo la implementación de esta herramienta, se ha utilizado el plataforma de código abierto 3DSlicer (www.slicer.org), centrada en la visualización y el procesamiento de imágenes médicas. 3DSlicer permite la incorporación de herramientas mediante módulos programables en Python, en base a lo cual se ha desarrollado la presente herramienta.

Este enfoque de segmentación se basa en investigaciones previas [3], que emplearon diversos algoritmos de procesamiento de imágenes para generar modelos tridimensionales del colon a partir de imágenes de Tomografía Axial Computerizada (TAC) de pacientes. La contribución de este estudio radica en la implementación de varias técnicas presentes en la vanguardia de este campo en una herramienta de software de código abierto diseñada específicamente para su uso en cirugía colorrectal. Esta herramienta estará disponible para su descarga gratuita desde el repositorio de 3DSlicer.

La mayor dificultad en el procesamiento de las imágenes de TAC del abdomen radica en la fragmentación de la región anatómica correspondiente al colon. Esta fragmentación se debe a obstáculos causados por movimientos peristálticos y acumulación de heces, lo que complica la tarea de segmentación y provoca una distribución no uniforme del gas a lo largo del colon. En consecuencia, previo al estudio preoperatorio de TAC, se recomienda que el paciente siga un tratamiento específico, que incluye una dieta blanda para prevenir la acumulación de heces, así como la administración de un contraste yodado que haga que cualquier líquido residual sea radiopaco, lo que facilitará su procesamiento posterior.

Por otro lado, para mejorar la visualización, es conveniente insuflar gas en el colon del paciente durante la adquisición de las imágenes. Para este propósito, se uti-

liza dióxido de carbono (CO_2), ya que se ha demostrado que es menos molesto y presenta menos riesgos para el paciente [4]. Estas medidas combinadas permiten simplificar significativamente el proceso de segmentación.

La presencia de gas en el interior del colon juega un papel crucial al facilitar la diferenciación entre la luz del colon y sus paredes en la imagen de TAC. Sin embargo, es importante considerar que el tejido del colon presenta una densidad muy similar a la de los tejidos circundantes. Por lo tanto, para paliar este efecto, se aplica un umbral de aproximadamente -800 Unidades Hounsfield de modo que permita descartar los píxeles con densidades radiológicas más altas. Además, es necesario tener en cuenta que en el abdomen hay diversas cavidades que contienen gas aparte del colon, sin contar con el aire de la sala donde se tomó la muestra, exterior al paciente.

Por consiguiente, para poder eliminar el aire externos al colon, se ha implementado un algoritmo basado en Crecimiento de Regiones. En este algoritmo, se colocan de forma automática "semillas" en el cuerpo del paciente y en la región de aire que envuelve al mismo (Figura 1.A). A través de este proceso de crecimiento, las semillas se expanden ocupando los píxeles adyacentes que comparten similitud hasta que se encuentran con las semillas opuestas (referentes a otras estructuras anatómicas). En la Figura 1.B se aprecia cómo las semillas han crecido hasta abarcar todo el volumen de gas externo. Una vez que se han diferenciado estos dos espacios (aire exterior y paciente), es posible excluir el volumen de aire exterior con una máscara. Por último en la Figura 1.C se aplica un proceso de threshold para filtrar entre las distintas densidades.

Para refinar el proceso anterior y eliminar otras cavidades que puedan contener gas en la segmentación, se ha desarrollado un proceso con interrupciones. En este proceso, se solicita al usuario que coloque manualmente semillas adicionales de crecimiento a lo largo de varios puntos del colon. Estos puntos anatómicos son habitualmente conocidos por los profesionales médicos y fácilmente identificables en cualquiera de las vistas del

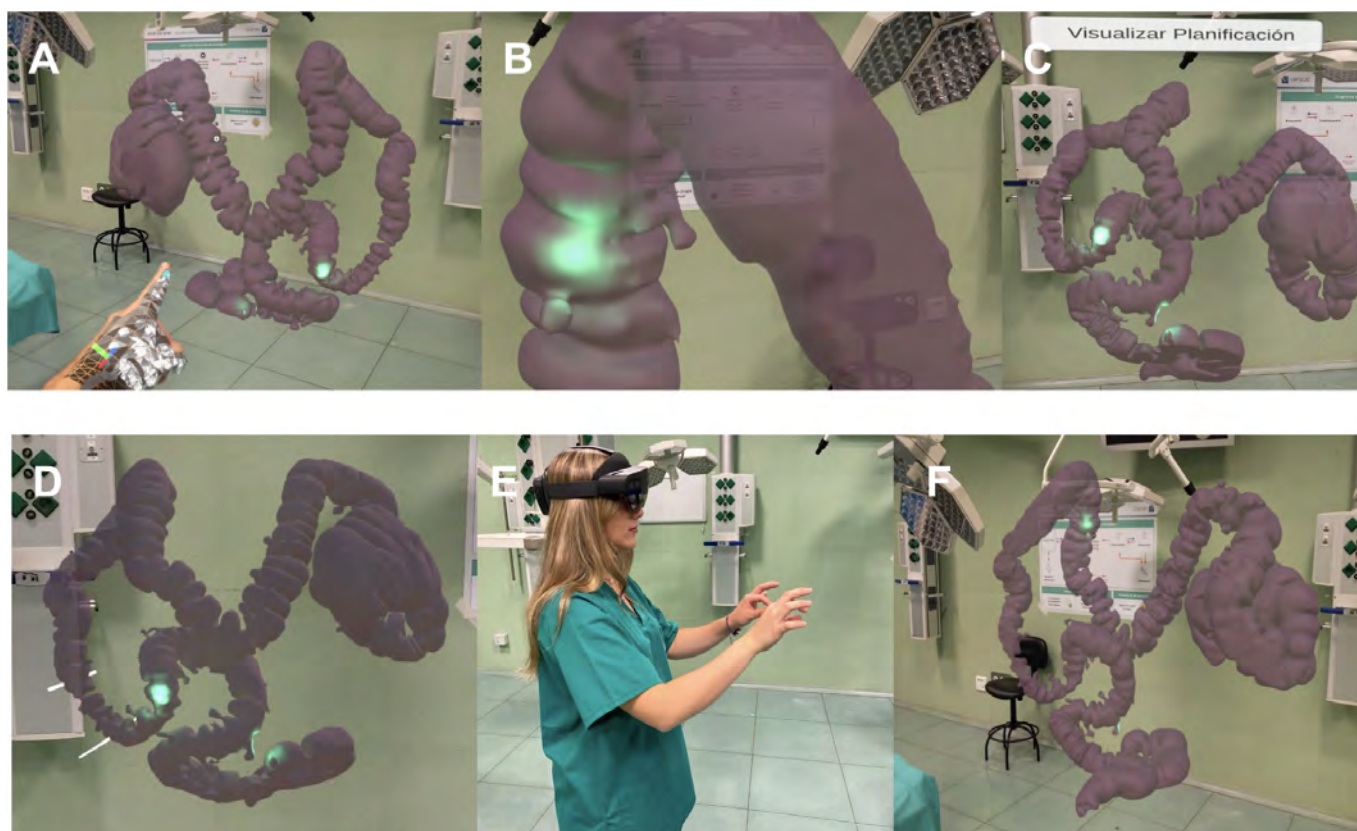


Figura 3: Visualización e interacción con el modelo de planificación quirúrgica mediante Realidad Mixta. Visualización del modelo de colon mediante el uso de las gafas HoloLens (A). Ampliación del modelo para remarcar uno de los puntos de luz donde existe la posibilidad de una lesión (B). Vista del botón "Visualizar Planificación" donde se carga el modelo (C). Posibilidad de realizar cortes (D). Usuario probando el sistema de visualización (E). Visualización completa del modelo (F).

planificador (axial, coronal o sagittal). Esto asegura que el crecimiento de regiones se realice correctamente en caso de que el colon tenga alguna posible obstrucción a lo largo de su recorrido.

Para validar la eficacia de la herramienta de segmentación, se han realizado pruebas utilizando 20 modelos de pacientes diferentes obtenidos de The Cancer Imaging Archive, una base de datos pública proporcionada por el National Cancer Institute (<https://www.cancer.gov/>).

2.2. Planificación quirúrgica

Una vez disponemos el modelo de colon segmentado del paciente, la planificación quirúrgica se plantea como una herramienta donde el cirujano puede añadir una serie de marcadores para resaltar puntos de interés y añadir segmentos para delimitar la zona de una posible resección y anastomosis durante la intervención. Para su posterior visualización en el sistema de realidad mixta, las coordenadas de estos elementos serán exportadas en un documento (formato JSON) junto al resultado de la segmentación.

Esta herramienta ha sido diseñada intentando simplificar al máximo la interfaz de usuario, procurando en todo momento que el proceso sea lo más intuitivo e interactivo posible para el cirujano. A través de una sencilla botonera, el cirujano podrá seleccionar puntos de interés clave

y trazar líneas de referencia que permitan esbozar estrategias quirúrgicas específicas para cada paciente. La flexibilidad de esta herramienta permite la personalización y adaptación a las lesiones y variaciones anatómicas individuales, lo que resulta esencial en la cirugía colorectal, donde la precisión es fundamental.

2.3. Visualizador basado en realidad mixta

Unity es una potente plataforma de desarrollo de videojuegos y aplicaciones de simulación tanto en 2D, 3D, como en realidad virtual/realidad aumentada (RV/RA). Esta herramienta proporciona un entorno de desarrollo integrado que abarca la creación, diseño y programación de aplicaciones interactivas. Una de las ventajas más destacables es que se trata de una multiplataforma donde se pueden exportar proyectos a una gran variedad de soportes. Gracias a la posibilidad de realizar tanto entornos en 2D como 3D ofrece un alto rendimiento y capacidades avanzadas de representación gráfica, incluido el ámbito clínico. Es la opción líder para el desarrollo de experiencias de RV y RA, ofreciendo soporte para diversas plataformas como; Oculus Rift, HTC Vive, Microsoft HoloLens y muchos otros.

Por lo tanto, se ha utilizado esta plataforma para la visualización e interacción con el modelo de planificación quirúrgica. Para ello, se ha programado un script que ex-

trae los datos de la planificación (archivo JSON). Estos datos hacen referencia a las coordenadas tridimensionales que nos permitirán ubicar el modelo de colon en el espacio, así como los marcadores (puntos de interés, planos de resección, etc.) indicados por el cirujano durante el proceso de planificación.

Para la visualización e interacción con el modelo de planificación se han empleado las gafas de Microsoft HoloLens. Este dispositivo nos permite combinar elementos virtuales con el entorno físico. Están equipadas con sensores avanzados, cámaras de profundidad y sensores de seguimiento ocular que permiten un control preciso del entorno y del usuario. Ofrecen capacidades de procesamiento de alto rendimiento y un sistema de visualización que superpone hologramas en el campo de visión del usuario.

La interacción entre Unity y HoloLens se logra a través del Microsoft Mixed Reality Toolkit (MRTK), una colección de scripts y recursos diseñados para simplificar el desarrollo de aplicaciones de RM en Unity. Se examinaron detenidamente las capacidades de MRTK para aprovechar las características únicas de HoloLens, como el seguimiento del entorno, los gestos de manos y la visualización de hologramas, con el fin de ofrecer una experiencia versátil, intuitiva y realista al usuario.

Una vez instalada en el dispositivo HoloLens la aplicación desarrollada Unity, el modelo de colon es cargado de forma automática, con el cual podrá interactuar (mover, rotar, cambiar de tamaño) el cirujano. A modo de simplificar la interfaz, sólo se muestra un botón que permite la visualización de las distintas marcas de la planificación quirúrgica definidas previamente. Se puede apreciar en la figura 3 las marcas luminosas que corresponden con los puntos de interés que haya reflejado el usuario durante la planificación. En la figura 3.D se puede apreciar en el colon transversal las marcas indicadas para una resección con la consecuente anastomosis. Se da la posibilidad al cirujano de mover las líneas de corte para poder desplazarlas o modificarlas en caso de querer actualizar la planificación previa.

3. Resultados y discusión

Se ha obtenido un plataforma de planificación quirúrgica basada en algoritmos para llevar a cabo una segmentación con un resultado positivo y eficaz en el 100 % de los 20 casos probados. El programa cuenta con una interfaz gráfica sencilla e intuitiva, lo que facilita un aprendizaje más rápido para los cirujanos a la hora de utilizar la herramienta por primera vez. La plataforma de planificación permite llevar a cabo, paso a paso, una segmentación y una planificación quirúrgica correcta. Sin embargo, hay que mencionar que con un dispositivo intel core i7 12700H y una tarjeta Nvidia RTX 3050 se requiere de un tiempo de ejecución notable, ya que las pruebas para los 20 modelos han arrojado una media de 250 segundos por modelo. Aún así, sigue siendo un tiempo considerablemente inferior que realizando la segmentación de forma tradicional, de media 1800 segundos, mejorando significativamente la sencillez y automatización

del proceso. En cuanto a la conexión 3DSlicer y Unity, se ha conseguido importar los datos de la planificación de manera satisfactoria, permitiendo una adecuada integración. Del mismo modo, utilizando las gafas HoloLens se puede observar el modelo de colon personalizado en forma de holograma, así como la información referente a la planificación quirúrgica previa a la intervención y con todo detalle anatómico.

4. Conclusiones

Se ha desarrollado un planificador quirúrgico para cirugía colorrectal intuitivo y el cual permite tanto segmentar la anatomía del colon del paciente como llevar a cabo la planificación de una manera guiada y sencilla. El resultado de la planificación personalizada de la cirugía puede ser visualizada en forma de holograma previo a la intervención, permitiendo al cirujano interactuar con el modelo 3D anatómico del paciente, junto con las anotaciones previas para planificar la intervención. Esta simbiosis entre técnicas de procesamiento de imagen médica y tecnología de visualización inmersiva permite realizar diagnósticos más precisos, tratamientos personalizados y una atención médica más eficiente. Todo ello, con el objetivo de mejorar la calidad y seguridad del proceso quirúrgico en la cirugía del colon, con la consecuente mejora de la calidad asistencial.

Agradecimientos

Este estudio ha sido cofinanciado por los fondos FEDER y la Junta de Extremadura (TA18023, IB20189), el Programa Erasmus+ de la Unión Europea (621668-EPP-1-2020-1-ES-EPPKA2-KA) y el Plan Complementario Biotecnología Aplicada a la Salud, cofinanciado por el Ministerio de Ciencia e Innovación con fondos de la Unión Europea NextGenerationEU, el Plan de Recuperación, Transformación y Resiliencia (PRTR-C17.I1) y el Programa Operativo FEDER Extremadura 2021-2027.

Referencias

- [1] Laversanne M. Soerjomataram I. Arnold M. Sierra M.S. Jemal A. Bray F. Global patterns and trends in colorectal cancer incidence and mortality. *Gut*, 489:683–691, 2017.
- [2] George J. Rockall T. Surgery for colorectal cancer. *Surg*, 2014(38(1)):32–37, 2020 Jan.
- [3] Wyatt C. L. Ge Y. Vining D. J. Automatic segmentation of the colon for virtual colonoscopy. *Computerized medical imaging and graphics*, (38(1)):24(1): 1–9, 2021.
- [4] Herráiz M. Colonoscopia con insuflación de dióxido de carbono: ¿lujo o necesidad? *Gastroenterología y Hepatología*, pages 36(1): 43–47.

Valoración de la Experiencia del Usuario en Entornos Virtuales de Rehabilitación de Miembros Superiores

V. Herrera Tirado ¹, J. Albusac Jiménez ¹, C. González Morcillo ¹, R. Perales Gómez ², L. Blanco Coloma ³, S. Ceruelo Abajo ⁴, V. Lozano Berrio ³, A. de los Reyes Guzmán ^{3,5}

¹ Departamento de Tecnologías y Sistemas de Información, UCLM, Ciudad Real, España, {Vanesa.Herrera,JavierAlonso.Albusac,Carlos.Gonzalez}@uclm.es.

² Unidad de Terapia Ocupacional. Hospital Nacional de Paraplégicos (SESCAM), Toledo, España, rpgomez@sescam.jccm.es

³ Unidad de Biomecánica y Ayudas Técnicas. Hospital Nacional de Paraplégicos (SESCAM), Toledo, España, lbcoloma@externas.sescam.jccm.es, {vlozanob,adlos}@sescam.jccm.es

⁴ Servicio de Rehabilitación. Hospital Nacional de Paraplégicos (SESCAM), Toledo, España, sceruelo@sescam.jccm.es

⁵ Unidad de Neurorrehabilitación, Biomecánica y Función Sensitivo-Motora (HNP-SESCAM), Unidad Asociada de I+D+I al CSIC

Resumen

Las aplicaciones de Realidad Virtual están en auge en el contexto de la rehabilitación de patologías neurológicas y es necesario abordar estudios para conocer el nivel de aceptación de los usuarios, en este caso pacientes con Lesión Medular. Se ha desarrollado un cuestionario exhaustivo formado por 40 preguntas para recoger el nivel de satisfacción y aceptación del usuario comparando la aplicación virtual del Box and Block en su versión inmersiva y no inmersiva. En el estudio participaron 6 sujetos sanos y 6 pacientes con LME. Se obtuvieron diferencias entre ambos entornos virtuales y la ausencia de efectos adversos relacionados con la RV se convierte en un factor fundamental para su aplicación en terapias de rehabilitación.

1. Introducción

En los últimos años, la Realidad Virtual (RV) ha experimentado un crecimiento significativo en su aplicación en diversidad de campos, incluida la medicina y la sanidad. Numerosos estudios se centran en contrastar y demostrar las ventajas de la RV en estas áreas [1, 2, 3]. En particular, dentro del área de la rehabilitación, la RV ofrece un conjunto de ventajas tanto para los especialistas como para los pacientes. En primer lugar, permite la creación de entornos personalizados y estimulantes que van más allá de la realidad cotidiana, brindando una rehabilitación más dinámica. Además, la capacidad de gamificar las terapias aumenta la motivación y el compromiso del paciente [4], reduciendo el riesgo de abandono. Por último, la RV facilita la medición precisa de los movimientos del paciente, lo que simplifica el análisis de su progreso a través de diferentes variables cinemáticas.

En este contexto, están emergiendo aplicaciones que hacen uso del seguimiento directo de las manos sin la necesidad de dispositivos externos. Además, algunos estudios han encontrado que los pacientes prefieren la interacción natural con las manos en entornos de RV en lugar de depender de controladores externos [5].

En concreto, dentro del ámbito de la rehabilitación de pacientes con Lesión de Médula Espinal (LME), la RV muestra ventajas prometedoras. Varios estudios han explorado el uso de la RV en rehabilitación, especialmente

en la recuperación de funciones motoras y la gestión del dolor. Sin embargo, y dado a que se trata de un ámbito de uso reciente, son pocos los estudios relacionados con la aceptación y la percepción del paciente en términos de usabilidad. Marika et al. [6] analizaron la usabilidad y aceptación de la RV en pacientes con dolor crónico, encontrando que la mayoría tuvo experiencias positivas, mostrando interés en su uso futuro. Además, la facilidad de uso de las aplicaciones y la calidad de la experiencia del usuario también influyen en la percepción general.

A partir de este escenario, surge el proyecto en el que se enmarca este trabajo, concebido especialmente para la rehabilitación de miembros superiores en pacientes con LME. Dicho proyecto tiene como objetivo el diseño y desarrollo de una plataforma en la que se despliega un conjunto de *exergames* (juegos serios para la realización de ejercicios físicos enfocados a la rehabilitación de miembros dañados) basados en RV. La plataforma no solo brinda la posibilidad de personalizar la experiencia según las necesidades individuales de cada paciente, sino que también recopila datos cinemáticos que resultan ser herramientas valiosas para la evaluación del progreso y el desempeño en los ejercicios de rehabilitación.

A raíz del desarrollo de esta plataforma, surge la necesidad de medir las sensaciones del usuario/paciente en su uso, así como dar respuesta a una pregunta que va más allá del análisis funcional de esta modalidad terapéutica: *¿se trata de una terapia aceptada por el paciente, que le resulte cómoda y que, además, le motive en su proceso de rehabilitación?*

2. Metodología

La metodología se basa en un cuestionario diseñado específicamente para registrar aspectos cruciales relacionados con la experiencia del paciente (LME y sano) en la utilización de aplicaciones de RV con fines terapéuticos.

2.1. Participantes

En el estudio participaron 6 pacientes con LME y 6 sujetos sanos como grupo de referencia. Los dos grupos eran homogéneos en edad. Todos los pacientes con LME presentaban lesión medular cervical entre los niveles C1-C7 y clasificación ASIA A-D, con función residual de

los miembros superiores que les permitiese interactuar con las aplicaciones virtuales planteadas. El estudio se realiza con la aprobación previa del Comité de Ética.

2.2. Entornos de Realidad Virtual y Adaptación del test Box and Block (BBT)

Para ambos entornos de RV se implementó una versión virtual del test clínico BBT [7]. Los participantes agarran, desplazan y sueltan los cubos virtuales con el uso de sus propias manos, conservando el mismo modo de interacción que con el BBT original. El primer entorno (no inmersivo), se basa en la interacción con elementos virtuales presentados en una pantalla, utilizando el dispositivo Leap Motion Controller (LMC). El segundo entorno (inmersivo) se desarrolló utilizando el dispositivo Meta Quest 2 (MQ2) que permite una experiencia de RV envolvente.

2.3. Cuestionario de evaluación

Se ha desarrollado un cuestionario exhaustivo formado por 40 preguntas para valorar el grado de usabilidad, la sensación de bienestar del usuario durante la realización de ejercicios, la interacción con los elementos virtuales y, finalmente, la percepción global. Para ser más precisos, este cuestionario consta de varias secciones, comenzando con el System Usability Scale (SUS) [8]. Se trata de un test estandarizado para medir la usabilidad global de los entornos virtuales a partir de 10 preguntas en escala Likert puntuadas de 1 a 5. La segunda sección está compuesta por el test estándar Virtual Reality Sickness Questionnaire (VRSQ) [9], diseñado para evaluar la satisfacción y la sensación de presencia en los entornos virtuales. La tercera sección del cuestionario incorpora 4 preguntas diseñadas para este estudio y que abordan aspectos claves de la interacción con los entornos de RV. Finalmente, la última sección consta de 2 preguntas que aportan una valoración global del sistema y de la aplicación utilizada. Estas dos últimas secciones se han diseñado específicamente para este estudio, por lo que no se trata de un test estandarizado.

3. Resultados

3.1. System Usability Scale

En esta sección, se presentan los resultados del System Usability Scale (SUS) como una evaluación de la usabilidad de sistemas y aplicaciones. SUS ofrece puntuaciones en una escala de 0 a 100, y en este estudio se incluyen calificaciones y percentiles [10] para una mejor comprensión.

Los resultados obtenidos (Tabla 1) reflejan una percepción positiva generalizada de usabilidad en ambos grupos, tanto en pacientes como en individuos sanos, independientemente del tipo de entorno de RV utilizado. No obstante, se observa una diferencia entre pacientes LME y personas sanas al evaluar la aplicación no inmersiva, obteniendo los pacientes una puntuación menor. En contraste, al analizar la aplicación inmersiva, se observó una convergencia significativa entre ambos grupos, indicando que experimentaron una percepción positiva similar en este entorno. Estos resultados se muestran en la Tabla 2, donde los promedios SUS otorgan una preferencia por el entorno inmersivo, en ambos grupos de participantes.

ID	VR	Punt.	Perc.	Calific.	Experiencia
P01	MQ2	85.0	60.42	A+	Ocasional
P01	LMC	47.5	12.5	F	Ocasional
P02	MQ2	92.5	85.42	A+	Ninguna
P02	LMC	75.0	37.5	B	Ninguna
P03	MQ2	72.5	29.17	C+	Ocasional
P03	LMC	27.5	4.17	F	Ninguna
P04	MQ2	85.0	60.42	A+	Ocasional
P04	LMC	85.0	60.42	A+	Ninguna
P05	MQ2	87.5	72.92	A+	Ocasional
P05	LMC	72.5	29.17	C+	Ocasional
P06	MQ2	92.5	85.42	A+	Ninguna
P06	LMC	80.0	45.83	A	Ninguna
S01	MQ2	90.0	79.17	A+	Ocasional
S01	LMC	62.5	20.83	D	Ninguna
S02	MQ2	87.5	72.92	A+	Ninguna
S02	LMC	72.5	29.17	C+	Ninguna
S03	MQ2	77.5	41.67	B+	Ocasional
S03	LMC	42.5	8.33	F	Ninguna
S04	MQ2	95.0	93.75	A+	Muy frec.
S04	LMC	85.0	60.42	A+	Muy frec.
S05	MQ2	95.0	93.75	A+	Muy frec.
S05	LMC	82.5	50.0	A	Ninguna
S06	MQ2	97.5	100.0	A+	Muy frec.
S06	LMC	55.0	16.67	D	Ocasional

Tabla 1. Puntuaciones obtenidas con SUS. De izquierda a derecha: identificador del participante (paciente-P, sano-S), dispositivo de RV utilizado (LMC - MQ2), puntuación total (sobre 100), percentil, calificación y experiencia previa con el dispositivo.

Grupo	Dispositivo de RV	Promedio SUS
Pacientes	LMC	64.58
Pacientes	MQ2	85.83
Sanos	LMC	66.67
Sanos	MQ2	90.42

Tabla 2. Promedio de puntuaciones SUS por grupo de participante y dispositivo de RV.

Con el fin de completar esta sección, se incluyó en el cuestionario una pregunta referente a la frecuencia de experiencia previa con este tipo de aplicaciones de RV. Los hallazgos revelaron patrones significativos. Aquellos participantes con experiencia previa calificada como «Muy frecuente» obtuvieron un promedio de 93.13 en la escala del SUS, lo que indica un alto grado de usabilidad del sistema (Tabla 3). Por otro lado, los individuos que carecían de experiencia previa obtuvieron un promedio de 72.73 en la misma escala. Estos resultados sugieren que una mayor familiaridad y adaptación con el entorno y dispositivo de RV podría estar relacionada con una percepción más favorable de la usabilidad del sistema. Sin embargo, es importante destacar que no se disponía de datos suficientes para evaluar la experiencia previa calificada como «Frecuente».

Para finalizar, se llevó a cabo un análisis estadístico, comparando las puntuaciones SUS entre los entornos LMC y MQ2. En el caso de pacientes LME, el test t (prueba estadística utilizada para determinar si existe

una diferencia significativa entre las medias de dos grupos) arrojó un valor de $-2,214$ sin diferencias significativas ($p=0.0512$). El resultado negativo sugiere que las puntuaciones de usabilidad para pacientes tienden a ser más altas en el entorno totalmente inmersivo con Meta Quest 2, en comparación con el entorno LMC.

Experiencia previa	Promedio puntuación SUS
Ninguna	72.73
Ocasional	74.72
Frecuente	NaN
Muy frecuente	93.13

Tabla 3. Promedio de las puntuaciones SUS según la experiencia previa.

4.2. Virtual Reality Sickness Questionnaire

VRSQ es un test de 9 preguntas utilizado para analizar la presencia y severidad de síntomas de mareo y malestar en entornos de RV. Esta versión reducida, basada en el cuestionario SSQ[11], encapsula dos dimensiones distintas de sintomatología y una valoración global, cada una valorada con un máximo de 100 puntos: oculomotor y desorientación. Los resultados del VRSQ permiten evaluar la tolerancia de los participantes a la RV.

En la figura 1 se observan los promedios de las puntuaciones individuales obtenidas para las dos dimensiones, las cuales fueron notablemente bajas, oscilando entre 0 y 25 en una escala de 0 a 100. Esto indica una percepción baja de malestar en relación con la experiencia en RV.

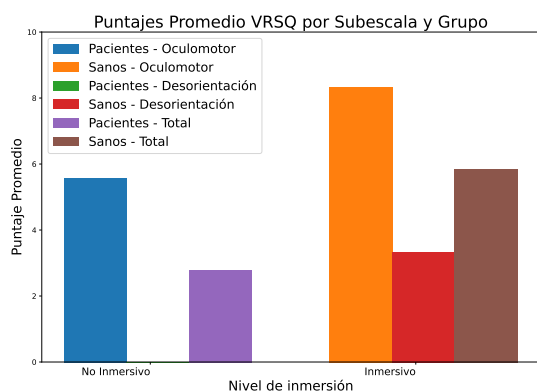


Figura 1: Comparación de puntuaciones VRSQ.

El análisis estadístico no reveló diferencias significativas entre los grupos. A pesar de esto, se observa una tendencia general de puntuaciones más bajas en los pacientes con LME en comparación con los sanos.

3.3. Interacción y gamificación

Esta sección del cuestionario explora aspectos vinculados con la interacción y la gamificación a través de 4 preguntas específicas: P1. La interacción con los elementos virtuales fue fluida y natural; P2. Me ha resultado fácil agarrar los objetos; P3. Me ha resultado fácil soltar los objetos; P4. Los elementos audiovisuales me han

ayudado a conocer el estado de la interacción; P5. Sentí sensación de presencia y conexión con el entorno virtual. Los participantes proporcionaron sus respuestas utilizando una escala tipo Likert de 1 a 5, siendo 1 la puntuación mínima.

En general, los pacientes experimentaron una mejora en la interacción con elementos virtuales en el entorno inmersivo en comparación con el no inmersivo. En este último, calificaron la interacción con un promedio de 2.67, mientras que en el inmersivo, obtuvieron un promedio de 4.67, indicando una mejoría significativa. En la facilidad para agarrar y soltar objetos también se observaron diferencias con un promedio de 3 en el LMC y 4.83 en el MQ2. Además, la sensación de presencia fue mucho más alta en el entorno inmersivo, con un promedio de 4.83 para los pacientes y 4.67 para los usuarios sanos (Tabla 4).

Pregunta	LMC-P	LMC-S	MQ2-P	MQ2-S
P1	2.67	3.00	4.67	4.33
P2	3.00	3.17	4.83	3.83
P3	3.00	3.67	4.00	3.67
P4	3.83	4.17	5.00	4.50
P5	2.50	3.33	4.83	4.67

Tabla 4. Promedios interacción en función de los participantes (P-paciente, S-sano) y dispositivo de RV.

Se examinó la percepción de la fluidez y naturalidad en la interacción con elementos virtuales, y los resultados revelaron diferencias notables ($t=-3,708$, $p=0,001$) entre la interacción en entornos no inmersivos e inmersivos. La magnitud negativa del test sugiere que la percepción de la interacción con elementos virtuales experimentó variaciones significativas entre los grupos evaluados. De manera paralela, la evaluación de la facilidad para agarrar objetos virtuales también evidenció diferencias significativas entre los grupos evaluados ($t=-2.121$, $p=0.045$). Esta observación apunta a la influencia en la percepción de la destreza al interactuar con elementos virtuales. Asimismo, se identificó diferencia significativa ($t=-4.844$, $p=7.687e-05$) en la sensación de presencia y conexión con el entorno virtual. Este resultado enfatiza el papel crucial de los entornos inmersivos en la ampliación de la sensación de presencia y conexión en la experiencia de realidad virtual. Una sensación de mayor presencia puede contribuir a una inmersión más profunda y una conexión más sólida con el entorno virtual. En resumen, los entornos inmersivos tienen un impacto significativo en la percepción de la fluidez, facilidad de agarre y sensación de presencia en la interacción con elementos virtuales. Los resultados respaldan la noción de que la inmersión puede potenciar positivamente la experiencia de interacción en aplicaciones de RV, permitiendo una realización más precisa de ejercicios y la posibilidad de mejorar los resultados esperados en la terapia.

3.4. Valoración global

En la última sección del cuestionario, se exploraron las percepciones de los participantes en relación con la valo-

ración global de la aplicación de RV en los dos entornos de inmersión. A los pacientes se les realizó 2 preguntas y su respuesta se registró en una escala de 1 al 5, donde 1 indicaba «Totalmente en desacuerdo» y 5 «Totalmente de acuerdo». Las preguntas planteadas fueron: P1. Me gustaría seguir utilizando aplicaciones de realidad virtual, enfocadas a la rehabilitación de los miembros superiores, con este dispositivo; P2. Creo que este tipo de rehabilitación puede ser un buen complemento a la rehabilitación convencional.

A la P1., los pacientes que utilizaron MQ2 obtuvieron un promedio de 5.0, indicando un alto nivel de aceptación y deseo de continuar utilizando estas aplicaciones. Sin embargo, este promedio baja sustancialmente (a un 3.33) cuando se trata de RV no inmersiva, lo cual sugiere una percepción algo menos entusiasta, pero aún positiva. En los pacientes sanos los promedios en ambos entornos fueron más similares. Se obtuvo, 4.67 en el uso de MQ2 y 4.17 con LMC (Tabla 5).

De manera similar, en la pregunta 2., los pacientes que utilizaron MQ2 nuevamente presentaron el promedio más alto, con un valor de 4.67, mientras que los pacientes LMC obtuvieron un promedio de 4.17. En sanos estos promedios se asemejan (Tabla 5).

	P1-LMC	P1-MQ2	P2-LMC	P2-MQ2
Paciente	3.33	5.00	3.83	5.00
Sano	4.17	4.67	4.33	4.67

Tabla 5. Promedios de la valoración global.

En los pacientes se encontraron diferencias significativas según el tipo de dispositivo de RV utilizado. Los pacientes que utilizaron MQ2 mostraron una mayor disposición a continuar usando estas aplicaciones en comparación con los que usaron LMC ($t=-2.5, p=0.031$). Sin embargo, en la pregunta sobre si este tipo de rehabilitación podría ser un buen complemento a la rehabilitación clásica, no se encontraron diferencias significativas ($t=-2.150, p=0.057$), lo que sugiere que ambas poblaciones analizadas compartieron opiniones similares en este aspecto.

Estos resultados sugieren que la RV inmersiva puede generar un mayor entusiasmo entre los pacientes, concretamente en este estudio en pacientes con LME.

4. Conclusiones

Los hallazgos encontrados animan a ampliar la muestra y sugieren un camino prometedor para el desarrollo de aplicaciones de RV enfocadas en la rehabilitación de miembros superiores, con un enfoque particular en la inmersión como factor clave para una experiencia satisfactoria y beneficiosa en pacientes con LME. Además, los resultados en el test VRSQ refuerzan la percepción general de que la RV es una tecnología segura y bien tolerada en entornos terapéuticos, lo que aumenta su atractivo y su prometedor futuro en el ámbito de la rehabilitación.

Agradecimientos

Proyectos PID2020-117361RB-C21 y PID2020-117361RB-C22 financiados por

MCIN/AEI/10.13039/501100011033.

Referencias

- [1] Anna Grochowska, Adam Wichniak, and Marek Jarema. Virtual reality – a valuable tool to advance treatment of mental disorders. *Archives of Psychiatry and Psychotherapy*, 21(1):65–73, March 2019.
- [2] Wanling Zhang, Dharendra Paudel, Rui Shi, Jie Liang, Jingwen Liu, Xiansheng Zeng, Yunfei Zhou, and Bin Zhang. Virtual Reality Exposure Therapy (VRET) for Anxiety Due to Fear of COVID-19 Infection: A Case Series. *Neuropsychiatric Disease and Treatment*, 16:2669–2675, December 2020.
- [3] Brian Mallari, Emily K. Spaeth, Henry Goh, and Benjamin S. Boyd. Virtual reality as an analgesic for acute and chronic pain in adults: a systematic review and meta-analysis. *Journal of Pain Research*, 12:2053–2085, 2019.
- [4] Florian Kern, Carla Winter, Dominik Gall, Ivo Käthner, Paul Pauli, and Marc Erich Latoschik. Immersive virtual reality and gamification within procedurally generated environments to increase motivation during gait rehabilitation. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 500–509, 2019.
- [5] M-Carmen Juan, Julen Elexpuru, Paulo Dias, Beatriz Sousa Santos, and Paula Amorim. Immersive virtual reality for upper limb rehabilitation: comparing hand and controller interaction. *Virtual Reality*, 27(2):1157–1171, June 2023.
- [6] Michelle Dy, Karla Olazo, Courtney R Lyles, Susan Lisler, Janice Weinberg, Catherine Lee, Marlon E Tarver, Amrita Saha, Kimberly Kontson, Roxana Araojo, Elizabeth Brown, and Urmimala Sarkar. Usability and acceptability of virtual reality for chronic pain management among diverse patients in a safety-net setting: a qualitative analysis. *JAMIA Open*, 6(3):o0ad050, Jul 2023.
- [7] Virgil Mathiowetz, Gloria Volland, Nancy Kashman, and Karen Weber. Adult norms for the box and block test of manual dexterity. *The American Journal of Occupational Therapy*, 39(6):386–391, 1985.
- [8] John Brooke. Sus: a “quick and dirty” usability. *Usability evaluation in industry*, 189(3):189–194, 1996.
- [9] Hyun K. Kim, Jaehyun Park, Yeongcheol Choi, and Mungyeong Choe. Virtual reality sickness questionnaire (vrsq): Motion sickness measurement index in a virtual reality environment. *Applied Ergonomics*, 69:66–73, 2018.
- [10] Jeff Sauro and James R. Lewis. *Quantifying the user experience: practical statistics for user research*. Elsevier, Morgan Kaufmann, Amsterdam Boston Heidelberg, 2nd edition edition, 2016.
- [11] Robert S Kennedy, Norman E Lane, Kevin S Berbaum, and Michael G Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.

M3Display: Sistema de realidad aumentada para la rehabilitación de la función motora del miembro superior

A. Cisnal¹, G. Alonso-Linaje¹, M. Veganzones¹, J.P. Turiel¹, J.C. Fraile¹

¹ Instituto de las Tecnologías Avanzadas de la Producción (ITAP), Universidad de Valladolid, Valladolid, España, cisnal@ieee.org.

Resumen

Se presenta un sistema basado en realidad aumentada para la rehabilitación de la función del miembro superior en pacientes con discapacidad motora. Este sistema utiliza la tecnología de realidad aumentada para combinar elementos virtuales con el mundo real, mejorando así la experiencia del usuario. El sistema se compone de una cámara que captura el entorno real a través de la reflexión de un espejo ubicado detrás de la pantalla de visualización. A través de un algoritmo de segmentación en tiempo real, las imágenes obtenidas mediante la cámara son analizadas para detectar las manos del paciente. Posteriormente, se realiza un procesamiento adicional que permite la creación del entorno de realidad aumentada mediante la integración de diversos elementos virtuales con las manos detectadas y la imagen original capturada. Además, incorpora una interfaz de usuario que permite una interacción sencilla e intuitiva con el sistema. Esta interfaz está compuesta por seis juegos serios diseñados específicamente para la rehabilitación del miembro superior. Este sistema ofrece una solución innovadora y prometedora para la rehabilitación de la función del miembro superior en pacientes con discapacidad motora. Al combinar la tecnología de realidad aumentada con juegos serios diseñados para la rehabilitación, se brinda a los pacientes una experiencia de rehabilitación más motivadora y efectiva.

1. Introducción

En el ámbito de la rehabilitación neuromotora, se ha observado un creciente interés en el uso de plataformas robóticas. Esta tecnología permite a los pacientes llevar a cabo su rehabilitación sin depender de una asistencia continua del personal médico, lo que amplía la accesibilidad a los programas de rehabilitación física. Esta cuestión es de vital importancia, especialmente ante la proyección de una reducción sustancial en la relación entre terapeutas y pacientes en las próximas décadas debido al envejecimiento de la población.

Numerosos estudios han demostrado que la rehabilitación basada en robots, en comparación con otros enfoques más tradicionales, mantiene a los pacientes motivados a través de la utilización de entornos de realidad virtual (RV) y realidad aumentada (RA) en combinación con técnicas de gamificación [1].

Sin embargo, el alto coste de los robots de rehabilitación dificulta su adquisición para el uso en el hogar por parte de los propios pacientes. Esto es una gran desventaja ya que se ha demostrado que una rehabilitación intensiva y temprana es de vital importancia para maximizar la recuperación motora funcional [2].

En este contexto, han surgido sistemas basados en RV y RA que permiten una rehabilitación motora sin necesi-

dad de un dispositivo robótico adicional. Estos sistemas potencialmente aumentarían la motivación del usuario, lo cual es un factor clave en el proceso de recuperación motora [3].

La Realidad Aumentada (RA) es el conjunto de tecnologías que permite al usuario visualizar el mundo real en combinación con objetos virtuales a través de una pantalla única. Para conseguir la superposición de elementos virtuales, un sistema de RA debe estar formado, por lo general, por los siguientes elementos: (1) Elemento de captura: dispositivo hardware que capta la imagen del mundo real; (2) Procesador: elemento hardware que combina la imagen del mundo real con la información virtual que debe superponer; (3) Software: programa informático específico que gestiona este proceso; (4) Dispositivo de visualización: dispositivo hardware que muestra los elementos del mundo real superpuestos con los elementos de la realidad aumentada; (5) Marcadores o detección de objetos: patrones visuales predefinidos del mundo real o algoritmos de detección de objetos que permite la superposición precisa de elementos virtuales en el entorno real.

En esta comunicación, se presenta M3Display, un sistema de RA diseñado específicamente para su utilización en terapias de rehabilitación del miembro superior. El sistema utiliza como dispositivo de visualización una pantalla, en lugar de gafas de RV utilizadas comúnmente en este campo [4, 5]. Esta elección se basa en las investigaciones que han demostrado que el uso de dispositivos voluminosos e invasivos, pueden generar estrés en los pacientes [6].

Por otro lado, estudios previos [7, 8] han propuesto sistemas de rehabilitación basados en RA que emplean marcadores (Figura 1). Aunque se demostró que la superposición de objetos virtuales genera una motivación positiva en los pacientes, se concluyó que la percepción visual de la profundidad y el sistema de seguimiento debían ser mejorados.

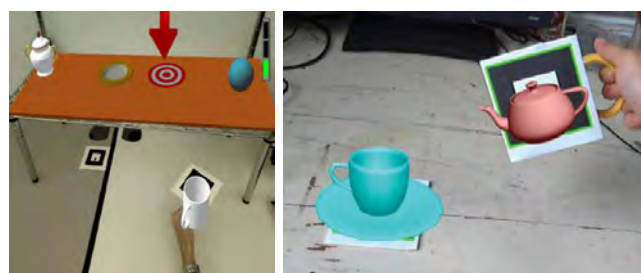


Figura 1: Sistemas de rehabilitación basados en RA que utilizan marcadores [7, 8].

Además, cabe destacar que la aplicación de marcadores en pacientes puede tener efectos adversos, dado que la mayoría de las personas que han experimentado un accidente cerebrovascular presentan espasticidad en la mano. Como alternativa, se propone el empleo de algoritmos de detección de objetos, los cuales, a pesar de ser computacionalmente más complejos y más susceptibles a errores, ofrecen la gran ventaja de prescindir de dichos marcadores.

El sistema de RA desarrollado se basa en un algoritmo de detección precisa de objetos, con un enfoque particular en la identificación de la mano. En la sección 2, se proporciona una descripción de los componentes de hardware que integran el sistema, así como su disposición espacial. En la sección 3, se explora el software desarrollado para la detección de la mano. Además, se presentan los juegos serios incorporados en el sistema, los cuales han sido diseñados para potenciar la eficacia y el compromiso del paciente durante las sesiones de rehabilitación. Por último, se lleva a cabo una breve discusión sobre el sistema desarrollado y se presentan las conclusiones derivadas de este estudio.

2. Hardware

El sistema se compone de una cámara Logitech HD Pro C920 (resolución de 1920 x 1080 a 30 fps y campo visual diagonal de 78°), un ordenador Intel Core i5 11400 CPU (6 núcleos, 2.6 GHz, 16GB de RAM) y una pantalla de 27" (Philips 273V7QDSB, resolución 1920 x 1080, @75 Hz). Esta pantalla es elegida por su peso ligero de 4 kg y su perfil de tan solo 45 mm de profundidad. Además, el sistema incluye un espejo de dimensiones 380 x 550 mm, un soporte para la pantalla y otros componentes auxiliares para la unión de los diferentes elementos. Estos componentes auxiliares pueden ser tanto comerciales como impresos en 3D para cumplir con los requisitos específicos del sistema.

La configuración del sistema se muestra en la Figura 2a. La pantalla se coloca en una posición perpendicular a la visión del paciente, situada entre este y sus manos. Un mecanismo permite la colocación del espejo detrás de la pantalla. En la parte trasera de la estructura se encuentra una cámara que captura el movimiento de las manos del paciente a través de la reflexión proporcionada por el espejo. Para asegurar la cámara en su posición, se emplea una estructura ubicada en la parte posterior de la pantalla.

Con el fin de garantizar una representación sin distorsiones de la mano en la pantalla, se requiere una colocación precisa de la cámara. Esta se posiciona estratégicamente en un ángulo de inclinación de 45 grados con respecto a la normal del espejo y se dirige hacia el centro del mismo (Figura 2b). De esta manera, se logra una superposición efectiva entre la línea de visión del paciente y la imagen capturada por la cámara. Como resultado, se genera la ilusión de que el brazo del paciente se extiende de manera ininterrumpida a través de la pantalla (Figura 2c).

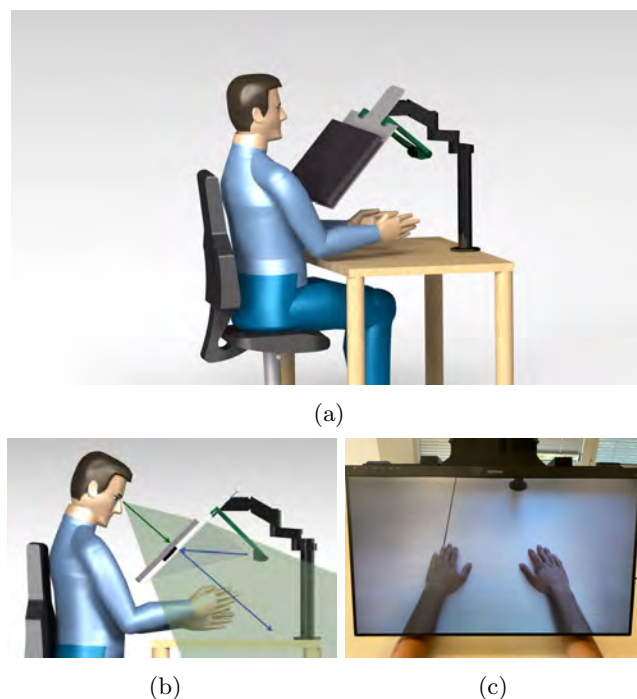


Figura 2: Estructura de M3Display. (a) Configuración del sistema; (b) Obtención de la imagen por medio del espejo y disposición de la cámara; (c) Prolongación de la mano de forma continua a través de la pantalla.

3. Software

El software consiste en una herramienta para la detección de manos en tiempo real y una serie de juegos en un entorno de RA reunidos en un interfaz de usuario. Estos dos módulos se presentan en los apartados 3.1 y 3.2 respectivamente.

3.1. Procesado de Imágenes

Las imágenes tomadas por la cámara se procesan inicialmente para detectar manos, seguido por un procesamiento adicional para construir el entorno de realidad aumentada a partir de ensamblar diferentes componentes del entorno virtual con las manos detectadas y la imagen original.

La detección de manos se realiza con una herramienta ligera que, a partir de los puntos de referencia obtenidos por MediaPipe Hands [9], obtiene una segmentación semántica precisa y computacionalmente sencilla. Esta herramienta permite segmentar manos hasta una velocidad de 90 fps, cumpliendo el requerimiento de funcionar en tiempo real impuesto por su uso en aplicaciones de RA.

La segmentación semántica se realiza en varias etapas, comenzando por la extracción de puntos de referencia por MediaPipe Hands [9], mostrados en la Figura 3b junto con el esqueleto que definen. La información extraída se procesa aplicando una serie de operaciones lógicas y morfológicas para transformar los puntos de referencia en un esqueleto de la estructura de la mano, y de ahí a una segmentación precisa. Partir del procesamiento de MediaPipe Hands permite conseguir gran precisión, y utilizar una secuencia de operaciones lógicas y

morfológicas produce un algoritmo computacionalmente ligero por encima de MediaPipe Hands, que es un software altamente optimizado.

Las operaciones lógicas y morfológicas aplicadas comienzan por dilataciones con kernels cuadrados de gran tamaño sobre el esqueleto de la estructura de la mano para conseguir una primera aproximación a la segmentación final. A continuación se realiza un muestreo del color en varios puntos cercanos al esqueleto de la mano para definir umbrales en el espacio de representación CIELAB (espacio de colores que ha permitido obtener la mayor precisión en el conjunto de datos de test tras probar varias alternativas) con los que distinguir puntos de la mano y del fondo, y así realizar un ajuste fino de las máscaras de segmentación obtenidas en el paso anterior. Un ejemplo de segmentación obtenida con esta herramienta se muestra en la Figura 3.

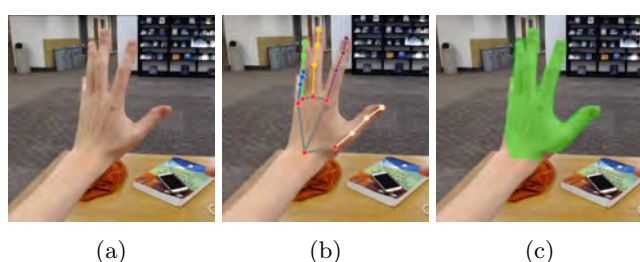


Figura 3: Etapas de la segmentación. (a) Imagen a segmentar; (b) Salida de MediaPipe Hands; (c) Segmentación obtenida.

Este algoritmo hace posible implementar juegos de RA controlados con una mano de forma sencilla. Nótese que el hecho de realizar segmentación semántica dificulta el utilizar más de una mano en los juegos. Es posible adaptar el algoritmo para realizar segmentación de instancias, ya que MediaPipe Hands distingue manos independientes, a costa de aumentar la carga computacional del algoritmo.

3.2. Interfaz de Usuario

El sistema cuenta con una colección de juegos de realidad aumentada que tienen como motor de segmentación el algoritmo mostrado en el apartado 3.1. Estos juegos se recogen en un interfaz de usuario (Figura 4) para facilitar el acceso a los mismos y permitir interrelacionar diferentes ejecuciones de los juegos mediante el uso de perfiles de usuario y registros de ejecuciones pasadas. Esto permite adaptar a cada paciente la dificultad de los juegos en función de su desempeño.

Los juegos desarrollados están orientados a la rehabilitación del miembro superior en personas con discapacidad motora. Por ello, todos los juegos se han diseñado de tal forma que la mano tome un papel protagonista y activo en el juego. De esta forma se involucra al paciente en la terapia induciendo que realice movimientos con su miembro superior de forma inadvertida. Esto se consigue utilizando la mano como elemento de interacción entre el paciente y el entorno virtual en el que se le introduce. Los juegos desarrollados (Figura 5) entran dentro de la categoría de juegos serios. El objetivo en estos juegos es

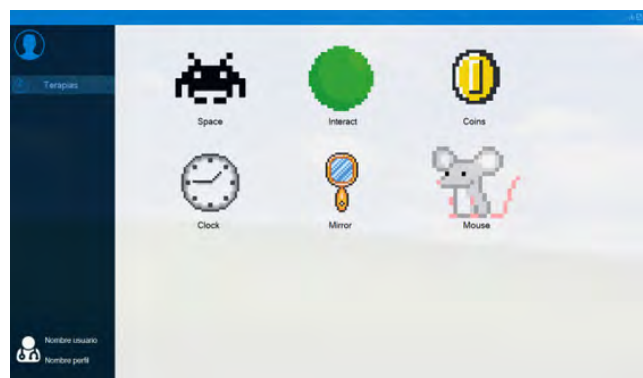


Figura 4: Interfaz de usuario de M3Display.

eliminar todos los objetivos que aparecen en la pantalla controlando una nave que dispara (Figura 5a), interactuar con una pelota virtual (Figura 5b), perseguir un ratón que huye de la mano introducida (Figura 5c), y alcanzar la posición del reloj indicada en cada instante (Figura 5d).

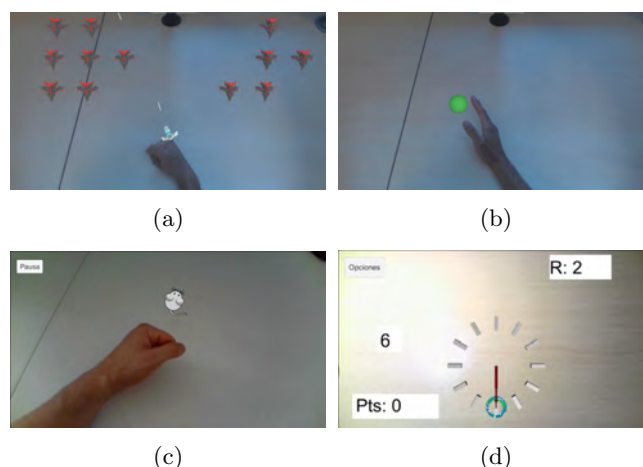


Figura 5: Juegos serios desarrollados. (a) Space; (b) Interact; (c) Mouse; (d) Clock.

4. Discusión y conclusiones

La presente comunicación se ha centrado en la descripción del sistema de RA utilizado en terapias de rehabilitación del miembro superior. Sin embargo, es importante destacar que este sistema posee una versatilidad significativa.

El sistema permite la incorporación de una plataforma robotizada (p. ej. RobHand [10] o M3Rob [11]) para asistir en el movimiento de la mano y/o muñeca del paciente. Esto es de especial importancia en la primera etapa de la rehabilitación cuando la movilidad del paciente es limitada. Con este fin, se desarrolló un algoritmo de segmentación semántica capaz de identificar los elementos mecatrónicos presentes en los entornos controlados de rehabilitación [12], permitiendo así eliminar la presencia del robot en la imagen y proporcionar al paciente una percepción visual alterna.

Por otro lado, el sistema también ofrece la posibilidad de llevar a cabo terapias en espejo. En lugar de utilizar un espejo colocado perpendicularmente entre ambas

manos, se posicionan las manos sobre una superficie y la pantalla muestra la imagen de la mano no afectada mientras la mano parética se sustituye por la imagen en espejo de la mano no afectada. Por ejemplo, se solicita a un paciente que realice movimientos simétricos con ambas manos. De esta manera, un paciente con una deficiencia en el lado izquierdo visualizará su mano derecha sana en el lado derecho de la pantalla, mientras que en el lado izquierdo de la pantalla verá reflejada su mano derecha. Esto crea la impresión de que la mano derecha no afectada, que se muestra reflejada en el lado izquierdo de la pantalla, es en realidad la mano izquierda funcional nuevamente. Esta idea ha sido planteada en estudios previos [5, 13, 14, 15].

El presente proyecto se encuentra actualmente en una etapa de desarrollo activo. Los próximos pasos a seguir comprenden la ampliación de la gama de juegos serios disponibles en el sistema, así como la realización de un análisis exhaustivo de riesgos y usabilidad. Posteriormente, se llevarán a cabo ensayos clínicos con el fin de evaluar la eficacia del dispositivo en cuestión.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia e Innovación, a través del proyecto RTC2019-007350-1 y por la empresa TICCYL Digital S.L.U. Además, este trabajo ha sido respaldado por el Ministerio de Ciencia, Innovación y Universidades, mediante el proyecto PID2019-111023RB-C33 y por la Junta de Castilla y León por las ayudas para financiar la contratación predoctoral de personal investigador, cofinanciadas por el Fondo Social Europeo (FSE). Por último, nos gustaría agradecer las ideas y sugerencias de los profesionales del Instituto de Rehabilitación Funcional (IRF) La Salle de Madrid.

Referencias

- [1] Omar Mubin, Fady Alnajjar, Nalini Jishtu, Belal Alsinglawi, and Abdullah Al Mahmud. Exoskeletons with virtual reality, augmented reality, and gamification for stroke patients' rehabilitation: Systematic review. *JMIR Rehabilitation and Assistive Technologies*, 6(2):1–11, 2019.
- [2] Franziska Herpich and Fred Rincon. Management of Acute Ischemic Stroke. *Critical Care Medicine*, 48(11):1654–1663, 11 2020.
- [3] Mónica S. Cameirão, Sergi Bermúdez I. Badia, and Paul F.M.J. Verschure. Virtual reality based upper extremity rehabilitation following stroke: A review. *Journal of Cyber Therapy and Rehabilitation*, 1(1):63–74, 2008.
- [4] Andrea Turolla, Mauro Dam, Laura Ventura, Paolo Tonin, Michela Agostini, Carla Zucconi, Pawel Kiper, Annachiara Cagnin, and Lamberto Piron. Virtual reality for the rehabilitation of the upper limb motor function after stroke: A prospective controlled trial. *Journal of NeuroEngineering and Rehabilitation*, 10(1):1–9, 2013.
- [5] Jörg Trojan, Martin Diers, Xaver Fuchs, Felix Bach, Robin Bekrater-Bodmann, Jens Foell, Sandra Kamping, Mariela Rance, Heiko Maaß, and Herta Flor. An augmented reality home-training system based on the mirror training and imagery approach. *Behavior Research Methods*, 46(3):634–640, 2014.
- [6] Guozheng Xu, Xiang Gao, Lizheng Pan, Sheng Chen, Qiang Wang, Bo Zhu, and Jinfei Li. Anxiety detection and training task adaptation in robot-assisted active stroke rehabilitation. *International Journal of Advanced Robotic Systems*, 15(6):1–18, 2018.
- [7] Wang Ying and Wang Aimin. Augmented reality based upper limb rehabilitation system. *ICEMI 2017 - Proceedings of IEEE 13th International Conference on Electronic Measurement and Instruments*, 2018-Janua:426–430, 2017.
- [8] Atif Alamri, Jongeun Cha, and Abdulmotaleb El Sadik. AR-REHAB: An augmented reality framework for poststroke-patient rehabilitation. *IEEE Transactions on Instrumentation and Measurement*, 59(10):2554–2563, 2010.
- [9] Fan Zhang, Valentin Bazarevsky, Andrey Vakunov, Andrei Tkachenka, George Sung, Chuo-Ling Chang, and Matthias Grundmann. MediaPipe Hands: On-device Real-time Hand Tracking, June 2020.
- [10] Ana Císnal, Víctor Lobo, Víctor Moreno, Juan C. Fraile, Rubén Alonso, and Javier P. Turiel. RobHand, un exoesqueleto de mano para la rehabilitación neuromotora aplicando terapias activas y pasivas. In *Actas de las XXXIX Jornadas de Automática*, volume 2, pages 34–41, Badajoz, 2018. Área de Ingeniería de Sistemas y Automática, Universidad de Extremadura.
- [11] Ana Císnal, Víctor Martínez-Cagigal, Gonzalo Alonso-Linaje, Selene Moreno-Calderón, Javier Pérez Turiel, Roberto Hornero, and Juan Carlos Marinero Fraile. An Overview of M3Rob, a Robotic Platform for Neuro-motor and Cognitive Rehabilitation Using Augmented Reality. In *XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, pages 180–183, Valladolid, 2022.
- [12] G. Sánchez-Brizuela, D. Sierra-Rodríguez, E. de la Fuente-López, J.C. Fraile, and J. Pérez-Turiel. Segmentación semántica de elementos mecánicos en entornos aumentados de rehabilitación. In Dpto. de Ingeniería de Sistemas y Automática, editor, *Jornadas de Robótica, Educación y Bioingeniería*, Málaga, 5 2022.
- [13] Simon Hoermann, Luara Ferreira dos Santos, Nadine Morkisch, Katrin Jettkowski, Moran Sillis, Hemakumar Devan, Parimala S. Kanagasabai, Henning Schmidt, Jörg Krüger, Christian Dohle, Holger Regenbrecht, Leigh Hale, and Nicholas J. Cutfield. Computerised mirror therapy with Augmented Reflection Technology for early stroke rehabilitation: clinical feasibility and integration as an adjunct therapy. *Disability and Rehabilitation*, 39(15):1503–1514, 2017.
- [14] Simon Hoermann, Leigh Hale, Stanley J. Winsler, and Holger Regenbrecht. Patient engagement and clinical feasibility of Augmented Reflection Technology for stroke rehabilitation. *International Journal on Disability and Human Development*, 13(3):355–360, 2014.
- [15] Holger Regenbrecht, Simon Hoermann, Claudia Ott, Lavell Muller, and Elizabeth Franz. Manipulating the experience of reality for rehabilitation applications. *Proceedings of the IEEE*, 102(2):170–184, 2014.

OTOVIRT: Un flujo de trabajo para el entrenamiento por simulación y el ensayo preoperatorio en cirugía virtual otológica

M. Lazo-Maestre¹, J. Mansilla-Gil¹, F. Ropero-Romero¹, C. Alonso-González¹, J. Ambrosiani-Fernández², J. Reina-Tosina³, S. Sánchez-Gómez¹, M. A. Callejón-Leblic^{1,3}

¹ Servicio de Otorrinolaringología, Hospital Universitario Virgen Macarena, Sevilla, España, {manu11235, mansillagiljorge}@gmail.com;

{cristina.alonso.sspa, francisco.ropero.sspa, serafin.sanchez.sspa}@juntadeandalucia.es;

² Departamento de Anatomía y Embriología Humana, Universidad de Sevilla, Sevilla ambrosiani@us.es

³ Grupo de Ingeniería Biomédica, Universidad de Sevilla, Sevilla, {mcallejon, jreina}@us.es;

Resumen

El desarrollo de intervenciones de cirugía otológica presenta el reto de trabajar con estructuras de muy compleja anatomía y gran variabilidad inter e intra-sujeto. El objetivo de este estudio es introducir un nuevo flujo de trabajo para la simulación virtual de cirugías del oído y la obtención sistemática de variables anatómicas y datos clínicos de los pacientes. El sistema propuesto está basado en el procesado de imágenes, el análisis de elementos finitos de los modelos y la simulación virtual; y ofrece un entorno en el que realizar el análisis pre y post-operatorio. Tras un registro inicial de imágenes de resonancia magnética (RM) y tomografía computerizada (TC), se segmentan una serie de estructuras de interés como el hueso temporal y la cóclea, entre otras, para generar los modelos tridimensionales de la anatomía del paciente. Se ha realizado para ello una armonización de herramientas basadas en software abierto que simplifican y optimizan la capacidad de apoyo al entrenamiento por parte del personal clínico. OTOVIRT ha sido validado en un entorno clínico con imágenes de seis pacientes candidatos a cirugía de implante coclear.

1. Introducción

Existen en el oído una serie de estructuras de interés de compleja anatomía que presentan un reto para el desarrollo de intervenciones quirúrgicas [1,3,4]. El desarrollo de herramientas de apoyo tiene potencial porque permite la planificación de intervenciones y ofrece información valiosa así como una mejora en el rendimiento [2,3]. Estas estructuras vitales necesitan ser preservadas, así que con el fin de incrementar la seguridad del paciente es cada vez más necesario el uso de nuevas herramientas que ayuden a planificar y controlar las intervenciones [4,5,6].

El uso de implantes cocleares (en adelante CI de *Cochlear Implant*) permite a los pacientes obtener un sistema de apoyo a la escucha en casos de pérdida severa de la audición [1,4,5]. Los pacientes susceptibles de recibir un CI primero deben someterse a una serie de pruebas que permitan valorar la viabilidad de la intervención [1], entre las que se incluyen RM y TC de cabeza y oído. Como el origen de la pérdida auditiva y su duración son clave en la

elaboración de este criterio, muchos pacientes pediátricos son implantados en el rango de 1 a 3 años de edad [1].

Existe una gran cantidad de herramientas de apoyo a la simulación, al análisis individual, poblacional y al entrenamiento quirúrgico [4,5,6]. A medida que la tecnología progresa, muchas herramientas tienden a integrarse y así es el caso de técnicas como la realidad virtual, el aprendizaje máquina o la tecnología háptica, que han aparecido con fuerza en el entorno clínico [2,3,4,5,6,7,8,9]. En el caso de las diseñadas para la asistencia quirúrgica son usualmente catalogadas por sus conceptos de diseño: aquellas que apoyan el estudio operatorio y las orientadas al entrenamiento [4,5].

Hay pruebas de la eficacia y la necesidad de simuladores y otras herramientas de apoyo a la cirugía otológica [4,5,6,7], por lo que resulta útil intentar integrarlas para obtener mayor experiencia en el desarrollo de estas complejas intervenciones. En este proyecto se ha diseñado un flujo de trabajo OTOVIRT que intenta semi-automatizar y guiar en el proceso de entrenamiento y de análisis pre y post operatorio. Además, este flujo contempla la creación de modelos tridimensionales de los pacientes que permitan la ejecución de simulaciones multifísicas. A continuación se irán detallando los aspectos principales.

2. Materiales y métodos

El flujo propuesto comprende una serie de etapas (Fig. 1):

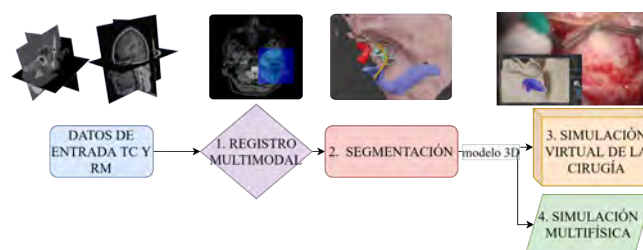


Figura 1. Diagrama de OTOVIRT. A partir de datos de entrada basados en imágenes TC y RM, se realiza un coregistro de ambas imágenes (1) y la segmentación de las estructuras de

interés (2). Posteriormente los modelos son utilizados para planificación pre-quirúrgica (3) y simulación multifísica (4).

2.1. Tipos de imágenes y coregistro multimodal

En el flujo propuesto se tiene a disposición el uso de los siguientes tipos de imagen médica:

- RM de tipología T1 y T2 de la cabeza completa del paciente, con una profundidad de corte de 0.65 mm. A partir de estas imágenes se realizan las correspondientes segmentaciones de la materia cerebral de cada paciente.
- TC del oído basadas en imágenes CBCT (*Cone Beam CT*), con una profundidad de corte de 0.18 mm. A partir de estas imágenes, se realizan las correspondientes segmentaciones de las estructuras del oído interno.

Estas dos imágenes son alineadas y solapadas mediante un proceso de registro multimodal que permite situar la información de una imagen en las coordenadas anatómicas de la otra [8]. Esto es especialmente útil en la segmentación de algunas estructuras complejas como el seno sigmoideo, la cual no se aprecia con suficiente resolución en ninguna de las dos pruebas de imagen. Sin embargo el seno sigmoideo se encuentra entre el cerebelo, que queda delimitado en RM y el hueso temporal, delimitado en TC, con lo cual el uso de ambas imágenes permite retener esta estructura con precisión.

2.2. Segmentación

2.2.1 Segmentación de las estructuras del oído interno

Para acometer la labor de segmentar las regiones de interés se ha desarrollado un módulo o plugin en el software 3D Slicer (<https://www.slicer.org/>) con secciones y herramientas adaptadas al trabajo con cada estructura anatómica de interés. En cada una se ha adaptado una técnica que resulte intuitiva para el cirujano y aproveche la anatomía para economizar el proceso, tratando de buscar ergonomía y reducir el tedio. El conjunto de secciones es el siguiente y puede observarse en la Fig. 2:

- Segmentación semi-automática de la cóclea: el usuario debe delimitar una ROI (*region of interest*) centrada en la cóclea. A continuación se realiza una umbralización seguida de una apertura morfológica para retener la cóclea.
- Segmentación del nervio facial: se señala una serie de puntos que serán interpolados con un spline bicúbico para trazar la ruta que sigue el nervio. Tras señalarla, se genera una estructura tubular que sigue la ruta definida y señala el nervio.
- Segmentación de la arteria carótida interna: se usa una herramienta igual a la anterior con valores preestablecidos con un diámetro mayor, cercano al de la estructura.
- Segmentación del hueso temporal: se usa primero una umbralización sencilla y luego otra limitada al volumen de la primera, que facilita retener las trabéculas del hueso irregular.
- Segmentación de estructuras irregulares con crecimiento de semillas: en el caso de estructuras de

compleja o irregular anatomía el seno sigmoideo o la cadena de huesecillos puede ser necesario una herramienta multipropósito que permita un mayor control manual pero de sencillo uso. Se han configurado macros para acceder a segmentos de interés y no interés para retener la zona con precisión.

- Medidas: además en la extensión realizada queda disponible el análisis de métricas como el volumen de las estructuras, los momentos de inercia, la elongación, los ejes principales, la longitud del ducto coclear y otras.

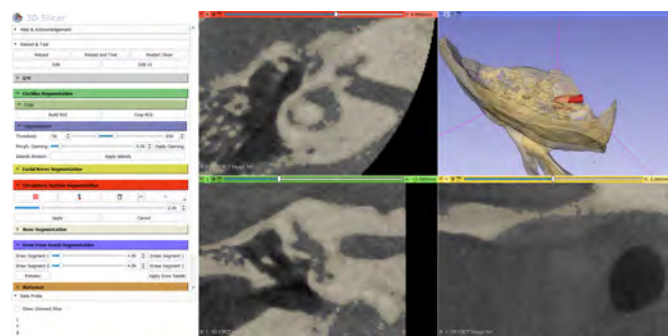


Figura 2. Módulo desarrollado en 3D Slicer para la segmentación semi-automática de las estructuras del oído interno y sus diferentes secciones.

2.2.2 Segmentación de los tejidos de la cabeza

A partir de RM de tipo T1 y T2 de la cabeza del paciente, mediante el software SimNIBS [10] basado en CAT12 (<https://github.com/simnibs/simnibs>), se pueden generar de forma automática una serie de estructuras de la cabeza, obteniendo mayor precisión en aquellas relacionadas con el tejido cerebral [11].

2.3. Simulación virtual de intervención

Finalmente se desarrolla la simulación de la cirugía virtual en el flujo de OTOVIRT, que está basada en Blender. Incluye: i) una ventana principal versátil que facilita un control preciso con rotaciones y traslaciones de los modelos y la cámara mostrados en el espacio, ii) un sistema de carga de modelos que automáticamente asigna colores e identificadores a las estructuras segmentadas, iii) un conjunto de herramientas de modelado que emulan el procedimiento de la manera más similar posible a la intervención real.

3. Simulación multifísica

Además, las simulaciones multifísicas pueden ser llevadas a cabo para poder predecir la magnitud y distribución de las corrientes eléctricas que fluyen a través de la cóclea del paciente. A partir de las estructuras segmentadas y las superficies triangulares obtenidas, se genera una malla volumétrica de tetraedros apta para la simulación mediante elementos finitos, haciendo uso de herramientas y toolboxes de libre acceso como iso2mesh (<https://iso2mesh.sourceforge.net/cgi-bin/index.cgi>) y CGAL (<https://www.cgal.org/>). Estos modelos pueden ayudarnos a lograr un mejor entendimiento del efecto de

diferentes estrategias de estimulación sobre predicciones de distribución de la corriente, facilitando así una optimización personalizada de los modos de estimulación del CI [12].

4. Resultados

En primer lugar, OTOVIRT ha sido probado con diferentes tipos de imágenes como CBCT, TC estándar, RM y μ TC. A continuación se muestra una tabla de resultados con pruebas realizadas con datos del repositorio *OpenEar* [13]. La Tabla 1 muestra el valor del índice Jaccard obtenido al valorar el solape entre las estructuras segmentadas y las consideradas verdad de referencia del repositorio mencionado. Se puede calcular el índice de Jaccard como:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad [14,15]$$

	Sistema cocleovestibular	Nervio facial	Cadena osicular
Sujeto 1	0,88	0,71	0,82
Sujeto 2	0,81	0,77	0,82

Tabla 1. Muestra de resultados obtenidos para la comparación con los segmentos de referencia del repositorio *OpenEar* para tres estructuras anatómicas del oído interno. Se usa el índice Jaccard.

Posteriormente, se prueba y valida OTOVIRT en un entorno clínico con imágenes de seis pacientes y con la ayuda de profesionales de la medicina y la cirugía. Se realiza la segmentación de las zonas de interés y se generan modelos de la cabeza (Fig. 3) y el oído (Fig. 4).

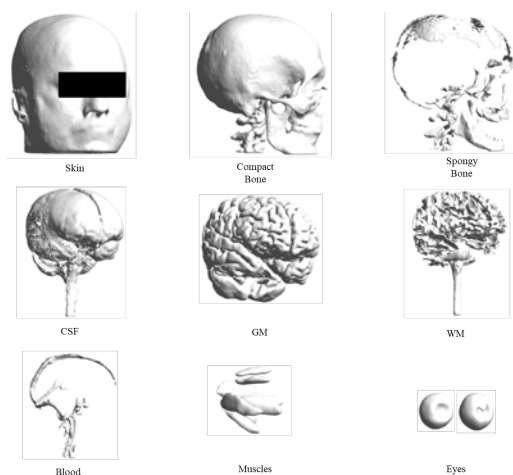


Figura 3. Ejemplos de las superficies de la cabeza generadas para un mismo sujeto.

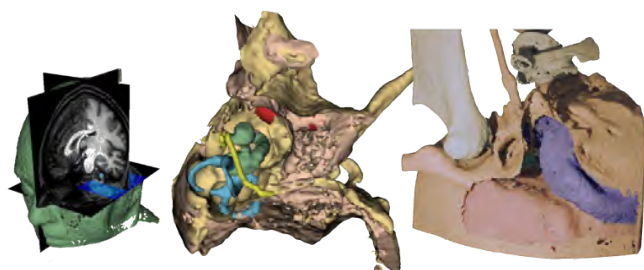


Figura 4. Muestras de los modelos obtenidos durante el proceso. De izquierda a derecha: modelo con el co-registro de RM y

TC, modelo con las segmentaciones situadas en el hueso temporal y modelo listo para la simulación virtual de la cirugía.

4.1. Análisis y validación de las estructuras

Se obtuvieron índices Jaccard superiores a 0.80 cuando se usó el flujo OTOVIRT para segmentar estructuras de interés como el sistema cocleovestibular y la cadena osicular a partir de imágenes abiertas del repositorio *OpenEar*. Estos valores nos permiten validar el procedimiento OTOVIRT en cuanto a su precisión con respecto a segmentaciones manuales de referencia de este repositorio. En el caso del nervio facial, cabe destacar que OTOVIRT semi-automatiza o simplifica su segmentación haciendo uso de un spline bicúbico con estructura tubular de diámetro constante, dando lugar a un índice Jaccard de 0.71 y 0.77 para sendos sujetos cuando se compara con la segmentación manual provista en el repositorio *OpenEar*. Como limitación, esta simplificación pierde precisión en las zonas en las que el nervio se engrosa y ensancha.

Posteriormente, se analizaron y validaron los modelos obtenidos a partir de imágenes de pacientes obtenidas en nuestro setup clínico. Tras la generación de las superficies de las diferentes estructuras de la cabeza y del oído interno, se procedió a realizar un análisis visual y a la elección de las superficies más precisas seleccionadas para el modelo final, todas ellas alineadas entre sí gracias al proceso previo de registro multimodal (Figura 4). Los modelos y simulaciones virtuales pre-quirúrgicas, según la validación realizada por personal experto en cirugía, demostraron su utilidad en el estudio de las estructuras e irregularidades anatómicas del paciente, mejorando así la eficacia de los resultados quirúrgicos y minimizando el riesgo de complicaciones durante la cirugía.

5. Discusión

La necesidad planteada en el proyecto OTOVIRT de simplificación del proceso de cirugía virtual aprovecha el progreso de la tecnología y la difusión de aplicaciones software gratuitas y frameworks abiertos como sustrato computacional para la creación de los escenarios virtuales 3D.

Una de las mejoras que introduce OTOVIRT se basa en la simplificación del proceso de ensayo pre-operatorio virtual de cara a su implantación como una fase más del proceso quirúrgico. Al proponer un flujo simple y adaptable se logra una mayor accesibilidad a herramientas que son clave en la mejora de los procedimientos. OTOVIRT facilita la homogeneización del trabajo, la comparación entre individuos y sienta las bases de sistemas automáticos al generar datos etiquetados como verdades de referencia que son clave para hacer eficaces estos sistemas.

La viabilidad de los resultados de OTOVIRT y su expansión e implantación en la práctica asistencial en la reorganización del proceso asistencial quirúrgico mediante el entrenamiento por simulación y el ensayo prequirúrgico han sido respaldadas por la Sociedad Española de Otorrinolaringología y Cirugía de Cabeza y Cuello (SEORL-CCC), que ha facilitado la difusión de los resultados de OTOVIRT en entrenamiento por simulación a

través de diferentes actividades formativas y científicas organizadas^[16,17,18,19].

Aunque los resultados preliminares con seis pacientes han sido muy positivos y acogidos con gran interés por la comunidad médica es necesario aún validar la herramienta OTOVIRT con un mayor número de pacientes, con el fin de obtener resultados experimentales sólidos y fiables que sustenten la validez y utilidad de la herramienta propuesta. Se espera que en el futuro, la obtención de métricas y el análisis de los modelos anatómicos permitan entrenar algoritmos de inteligencia artificial que ayuden a automatizar los procesos.

6. Conclusión

El flujo de trabajo OTOVIRT propuesto ofrece un entorno en el que realizar el análisis pre y post-operatorio. Tiene aplicabilidad clínica y capacidad de apoyo al entrenamiento quirúrgico. La validación del mismo con datos de imagen específicos de paciente se ha probado exitosa y provee al cirujano de un flujo de trabajo asequible y de fácil uso con el que entrenar y mejorar los resultados de la intervención mientras se mantiene la seguridad del paciente. El flujo propuesto en este proyecto OTOVIRT asienta las bases para futuros análisis multimodales y simulaciones multifísicas personalizadas que ayuden a entender las distribuciones de corrientes bajo distintas estrategias de estimulación del implante coclear.

Agradecimientos

Este trabajo ha sido financiado por el proyecto OTOVIRT (PIN-0097-2020): “Cirugía Virtual para el entrenamiento por simulación y el ensayo preoperatorio en cirugía otológica y cirugía endoscópica endonasal”.

Referencias

- [1] Deep NL, Dowling EM, Jethanamest D. “Cochlear Implantation: An Overview.” *Journal of Neurological Surgery: Part B, Skull Base, U.S. National Library of Medicine*. 2019;80(2):169-177. doi:[10.1055/s-0038-1669411](https://doi.org/10.1055/s-0038-1669411)
- [2] Van Der Meijden OAJ., Schijven MP. “The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: A current review”. *Surg Endosc.* 2009; 23, 1180–1190. doi:[10.1007/s00464-008-0298-x](https://doi.org/10.1007/s00464-008-0298-x).
- [3] Fang, TY. *et al.* “Evaluation of a haptics-based virtual reality temporal bone simulator for anatomy and surgery training.” *Computer methods and programs in biomedicine*. 2014; vol. 113,2 : 674-81. doi:[10.1016/j.cmpb.2013.11.005](https://doi.org/10.1016/j.cmpb.2013.11.005)
- [4] Okuda Y. *et al.* “The utility of simulation in medical education: What is the evidence?” *Mt Sinai J Med*. 2009;76(4):330-343. doi:[10.1002/msj.20127](https://doi.org/10.1002/msj.20127)
- [5] Wiet, Gregory J *et al.* “Otolologic Skills Training.” *Otolaryngologic clinics of North America*. 2017; vol. 50,5: 933-945. doi:<https://doi.org/10.1016/j.otc.2017.05.005>
- [6] Luv Javia and Ellen S. Deutsch. “A systematic review of simulators in otolaryngology” *Otolaryngol Head Neck Surg*. 2012;147(6):999-1011. doi:[10.1177/0194599812462007](https://doi.org/10.1177/0194599812462007)
- [7] Jiang Wang *et al.* “Fully automated segmentation in temporal bone CT with neural network: a preliminary assessment study”. *BMC Medical Imaging*. 2021; 21. issn: 14712342. doi:[10.1186/s12880-021-00698-x](https://doi.org/10.1186/s12880-021-00698-x).
- [8] El-Gamal F., El-Zahraa A. *et al.* “Current trends in medical image registration and fusion.” *Egyptian Informatics Journal*. 2016; 17: 99-124. doi: [10.1016/j.eij.2015.09.002](https://doi.org/10.1016/j.eij.2015.09.002).
- [9] Hu Y. *et al.* “Weakly-supervised convolutional neural networks for multimodal image registration”. *Medical Image Analysis*. 2018; 49 ,pp. 1–13. issn: 13618423. doi: [10.1016/j.media.2018.07.002](https://doi.org/10.1016/j.media.2018.07.002).
- [10] Guilherme BS. *et al.* “SimNIBS 2.1: A Comprehensive Pipeline for Individualized Electric Field Modelling for Transcranial Brain Stimulation”. *Brain and Human Body Modeling*. 2019; pp 3-25; doi:https://doi.org/10.1007/978-3-030-21293-3_1
- [11] Nielsen J.D. *et al.* “Automatic skull segmentation from MR images for realistic volume conductor models of the head: Assessment of the state-of-the-art”. *NeuroImage*. 2018; 174, pp. 587–598. issn: 1053-8119. doi: <https://doi.org/10.1016/j.neuroimage.2018.03.001>.
- [12] Callejón-Leblic MA., Fratter A., Ropero-Romero F., Reina-Tosina J., Sánchez-Gómez S. “A Full Head Model to Analyze Intra and Extra-cochlear Electric Currents under Different Cochlear Implant Stimulation Strategies”, 2023 *Brain & Human Body Modeling (BHBM) Conference*, AUG. 17-18, 2023, Boston.
- [13] Sieber, D., Erfurt, P., John, S., dos Santos, G. R., Schurzig, D., Sørensen, M. S., & Lenarz, T. “Data descriptor: The openEar library of 3D models of the human temporal bone based on computed tomography and micro-slicing”. *Scientific Data*. 2019;6. <https://doi.org/10.1038/sdata.2018.297>
- [14] Bertels, J. *et al.* “Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory and Practice”. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*. *MICCAI 2019*. Lecture Notes in Computer Science, vol 11765. Springer, Cham. https://doi.org/10.1007/978-3-030-32245-8_11
- [15] Bag S., Kumar SK., Tiwari MK., “An efficient recommendation generation using relevant Jaccard similarity”. *Information Sciences*. 2019; vol 483 pp-53-46. doi: <https://doi.org/10.1016/j.ins.2019.01.023>
- [16] Lazo-Maestre M., Mansilla-Gil J. *et al.* “OTOVIRT: Una herramienta para el modelado de imágenes médicas para el soporte al diagnóstico y la simulación de cirugías virtuales del oído”, *Póster en Jornadas Andaluzas de Bioingeniería (JABI)*, 3-5 Jul. 2023, Sevilla.
- [17] Sánchez-Gómez S. *et al.* “OTOVIRT: a workflow for semi-automatic modeling and segmentation of the temporal bone as a simulation tool for haptic virtual surgery”, *Póster en IFOS Dubai ENT World Congress 2023*. 17-21 Jan. 2023, Dubai.
- [18] Ropero-Romero F. *et al.* “A Modelling Workflow and Virtual Reconstruction of the Transmastoid Approach for Cochlear Implant Surgery in Children”, *Póster en XVI ESPCI, European Symposium on Pediatric Cochlear Implantation*. 31 May - 3 Jun. 2023. Rotterdam.
- [19] Callejón-Leblic MA. *et al.* “Implantlabs: una plataforma multimodal y base de datos para la adquisición y análisis de datos de investigación en audiología.”, *XIX CNA 2K23, Congreso Nacional de Audiología*. 15-17 Jun. 2023, Madrid.

Automatización para la evaluación ergonómica integrando las tecnologías de realidad virtual, captura de movimiento, electromiografía de superficie y uso de exoesqueleto

G. Salcedo Eugenio¹, D. Ojados González¹, I.J. Ibarra Berrocal¹, A. Macián Morales¹

¹Servicio de Apoyo a la Investigación Tecnológica (SAIT), Universidad Politécnica de Cartagena, Cartagena, España
{gustavo.salcedo, lola.ojados, isidro.ibarra, alvaro.macian}@upct.es

Resumen

Este proyecto trata el desarrollo de un sistema automático de evaluación ergonómica en tiempo real, lo que permitirá realizar un análisis biomecánico basado en los métodos más relevantes de evaluación ergonómica, para identificar riesgos de lesiones musculoesqueléticas de un trabajador durante la ejecución de su tarea.

Mediante una aplicación de realidad virtual se integran los movimientos de un trabajador en un modelo virtual utilizando la tecnología de captura de movimiento que permite obtener posiciones y ángulos de las articulaciones. Adicionalmente permite obtener parámetros biológicos de esfuerzos musculares mediante un sistema de electromiografía de superficie.

En la aplicación virtual se programarán los métodos de evaluación ergonómica (Rula, Reba y ecuación NIOSH, entre otros) y se visualizarán las alertas de las posturas forzadas, sobreesfuerzos y parámetros biológicos.

Este sistema facilitará la labor del técnico de prevención, proporcionando información objetiva, fiabilidad en las mediciones, se podrá utilizar en cualquier entorno y permitirá realizar una evaluación ergonómica en tiempo real y generar un informe final cuando finalice la evaluación.

1. Introducción

Las nuevas tecnologías abarcan casi todas las disciplinas, la ingeniería del factor humano puede aprovechar estas nuevas herramientas para implementarlas en alguno de sus métodos convencionales basados en técnicas de evaluación subjetiva. Actualmente, existen herramientas de medición que eliminan cualquier factor subjetivo y aportan datos objetivos para los análisis.

Un ejemplo de estas nuevas herramientas es la captura de movimiento, o mocap por sus siglas en inglés (Motion Capture), es una tecnología utilizada en la industria del entretenimiento y la animación para registrar y digitalizar los movimientos de objetos o seres vivos en el mundo real y luego aplicar esos movimientos a modelos digitales en 3D. Esto se logra mediante el uso de sensores, cámaras y software especializado que rastrean y registran los movimientos en tiempo real. (Ver Figura 1).

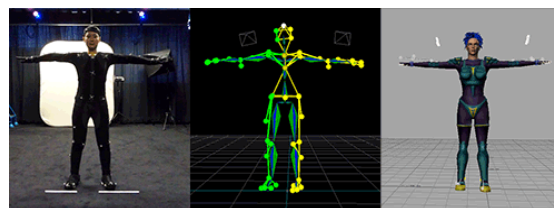


Figura 1. Mocap (Sistema óptico)

Fuente: <https://coatayork1.wordpress.com/2015/03/15/post-proceso-del-mocap/>

El Mocap Inercial es una tecnología de captura de movimiento que utiliza sensores inerciales, como acelerómetros y giroscopios, para registrar y rastrear movimientos en tiempo real. Estos sensores se colocan en el cuerpo o en objetos, eliminando la necesidad de marcadores externos. (Ver Figura 2).



Figura 2. Mocap (Sistema no óptico / Seguimiento inercial)

Fuente: <https://www.awn.com/news/mpc-puts-new-xsens-motion-capture-system-work>

Los movimientos digitalizados pueden analizarse en un escenario virtual utilizando los datos procesados que se capturaron y generando evaluaciones objetivas que podrían sustituir a los métodos actuales de observación para medir los ángulos posturales que son incluidos en un informe con una puntuación de riesgo.

Otra tecnología que se aprovecha en este proyecto es la Realidad Virtual (VR) concretamente la realidad virtual no inmersiva, se refiere a experiencias de realidad virtual que no requieren el uso de cascos o gafas de VR para bloquear completamente la vista del mundo real. En cambio, estas experiencias de VR se ejecutan en dispositivos de pantalla tradicionales, como computadoras, tabletas y teléfonos móviles.

Aunque la VR no inmersiva ofrece una experiencia menos envolvente en comparación con la VR inmersiva, sigue siendo una forma valiosa de experimentar contenido virtual y puede ser más accesible y económica. (Ver Figura 3).



Figura 3. Realidad Virtual no inmersiva

Fuente: <https://www.fisioterapia-online.com/articulos/realidad-virtual-una-alternativa-terapeutica-en-fisioterapia>

Además, se incluye la "electromiografía de superficie" (sEMG) que es una técnica médica utilizada para registrar la actividad eléctrica superficial de los músculos. A diferencia de la electromiografía de aguja, que implica la inserción de agujas delgadas en los músculos para medir la actividad eléctrica directamente, la sEMG coloca electrodos en la piel sobre los músculos que se están evaluando. (Ver Figura 4).

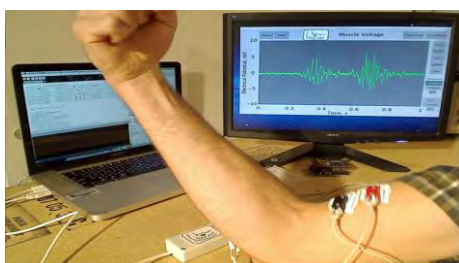


Figura 4. sEMG

Fuente: <https://fisioterapeutaadomicilio.es/electromiograma-superficie-biofeedback-muscular/>

Estos electrodos registran las señales eléctricas generadas por los músculos durante su contracción y relajación. La sEMG es menos invasiva que la EMG de aguja y se utiliza a menudo en estudios de investigación, evaluaciones de rendimiento muscular en deportes y fisioterapia.

En este proyecto se aplican las anteriores tecnologías a la ingeniería del factor humano, específicamente a la evaluación ergonómica. Los métodos de evaluación ergonómica permiten identificar y valorar los factores de riesgo presentes en los puestos de trabajo para, posteriormente, en base a los resultados obtenidos, plantear opciones de rediseño que reduzcan el riesgo y lo sitúen en niveles aceptables de exposición para el trabajador. La exposición al riesgo de un trabajador en un puesto de trabajo depende de la amplitud del riesgo al que se expone, de la frecuencia del riesgo y de su duración (las posturas forzadas, los levantamientos de carga, la repetitividad de movimientos, etc.). Dicha información es posible obtenerla mediante métodos de evaluación ergonómica.

Finalmente se incluirán pruebas con un exoesqueleto textil para soporte lumbar que está compuesto por una estructura flexible diseñada para ser usada en el cuerpo humano y proporcionar soporte, protección o mejora de la movilidad. A diferencia de los exoesqueletos tradicionales hechos de materiales rígidos como metal, los

exoesqueletos textiles utilizan tejidos y materiales flexibles para lograr sus objetivos. Incorporan elementos elásticos o compresivos para ayudar a los usuarios a moverse con mayor facilidad, reducir la fatiga muscular o brindar soporte a áreas específicas del cuerpo. (Ver Figura 5).



Figura 5. Exoesqueleto textil HELK

Fuente: <https://www.gogoa.eu/en/exoesqueleto-sociosanitario-helk>

El objetivo del proyecto es crear un sistema automático y portátil de evaluación ergonómica, basado en las tecnologías de realidad virtual, captura de movimiento y electromiografía de superficie. Integrar los principales métodos de evaluación ergonómica para generar un análisis biomecánico e identificar riesgos de lesiones musculoesqueléticas de un trabajador durante la ejecución de su tarea.

2. Metodología

La adquisición de datos y variables del trabajador se genera mediante tres tecnologías. La captura de movimiento (MoCap) se realiza utilizando un hardware inalámbrico con sensores inerciales (Traje Perception Neuron Pro) y mediante el software (Axis Neuron Pro), se captura en tiempo real los movimientos en un modelo virtual. (Ver Figuras 6 y 7).



Figura 6. Mocap (Perception Neuron Pro)

Fuente: https://www.aiuto-jp.co.jp/products/product_2459.php



Figura 7. Axis Neuron Pro

Los datos obtenidos se exportan en tiempo real y de forma sincronizada al software de realidad virtual Unity. Unity es un software esencial en la industria del

desarrollo de videojuegos y aplicaciones interactivas debido a su versatilidad, capacidad multiplataforma y amplia base de usuarios. (Ver Figura 8).



Figura 8. Software UNITY

En Unity se visualiza el modelo virtual y se programan los métodos más relevantes para realizar una evaluación ergonómica (REBA, RULA, NIOSH, OWAS y BIO-MEC).

El sistema visualiza alertas de los ángulos críticos de las articulaciones, los sobreesfuerzos de los músculos y la superación de los límites de los parámetros biológicos.

Por medio del sistema inalámbrico y portátil de sEMG Delsys Trigno se adquieren los datos eléctricos de los músculos que intervienen. (Ver Figura 9).



Figura 9. sEMG DELSYS Trigno Wireless

Fuente: https://www.researchgate.net/figure/DELSYS-Trigno-wireless-EMG-measurement-instrument_fig2_279278282

En el software EMGworks Acquisition se configuran, visualizan y guardan los datos de los sensores sEMG. (Ver Figura 10).

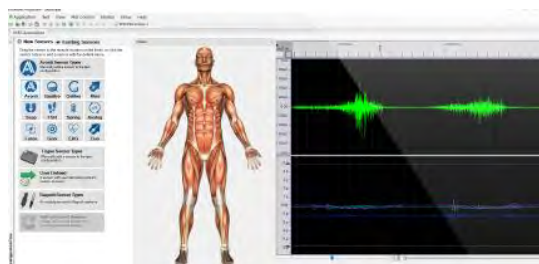


Figura 10. EMGworks Acquisition

Fuente: <https://www.medicaexpo.fr/prod/delsys/product-123919-989599.html>

3. Resultados

Se ha desarrollado un procedimiento de trabajo para la validación del uso de las tecnologías de Mocap y sEMG.

Se realizó un estudio piloto en una muestra de población de personal sanitario (Personal limpieza de hospital – Auxiliar de enfermería del Hospital General Universitario Santa Lucía “HGUSL”). Además, se incorporó el empleo de un exoesqueleto para comparar las posturas y esfuerzos realizados durante la jornada laboral usando o no este dispositivo. (Ver Figura 11).

Se toman datos en campo para realizar la evaluación ergonómica en tiempo real de las tareas de limpieza de pacientes, cambio de sábanas y acondicionamiento de camas en los ámbitos hospitalario.

Para la toma de datos en campo se selecciona una muestra de 5 trabajadores/as voluntarios/as que realicen las tareas indicadas anteriormente, primero sin ayuda del exoesqueleto y posteriormente con ayuda del exoesqueleto, durante al menos 5 horas de su jornada de trabajo.

El procedimiento de toma de datos se realizó colocando el traje Mocap y sensores sEMG en el trabajador en uno de los ensayos con una duración de 2 horas, un descanso de 1 hora para recarga de sensores y en el siguiente ensayo se incorporó el exoesqueleto durante 2 horas.

Se tomaron fotografías y captura de video, por lo que fue necesario la autorización de los implicados según lo indicado en la ley de protección de datos.



Figura 11. Mocap, sEMG y exoesqueleto durante el ensayo

El sistema de captura de movimiento (Mocap) permitió registrar y analizar los movimientos corporales durante la actividad laboral, brindando información precisa sobre las posturas adoptadas, la amplitud de movimiento y la interacción de los empleados con su entorno de trabajo. (Ver Figura 12).

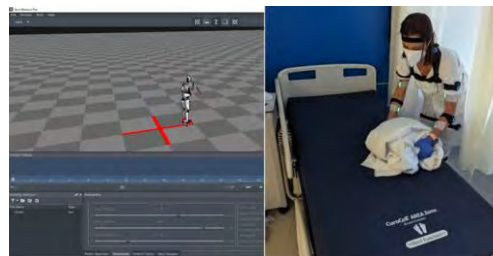


Figura 12. Voluntaria del HGUSL con Mocap y sEMG

Por otro lado, los equipos de electromiografía hicieron posible la medición de los esfuerzos musculares realizados por los trabajadores, registrando la actividad

eléctrica de los músculos y proporcionando datos objetivos sobre la carga física a la que se vieron sometidos durante sus labores diarias. (Ver Figura 13,14 y 15).



Figura 13. Voluntaria del HGUSL con sensores de sEMG

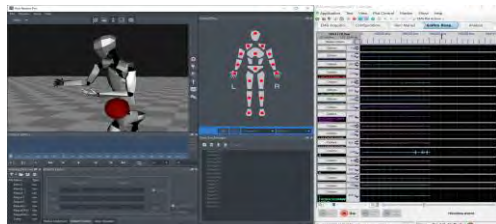


Figura 14. Adquisición de datos Mocap y sEMG

En el software Unity se realiza la integración de los sistemas Mocap, sEMG y métodos de evaluación ergonómica.



Figura 15. Integración de Mocap, sEMG y Sistema automático de evaluación ergonómica en Unity

Por medio de una encuesta se recogió la opinión de cada trabajador para evaluar la experiencia de uso del exoesqueleto. Se obtienen datos del perfil del entrevistado y opinión sobre el exoesqueleto.

4. Conclusiones

Este sistema facilita la labor del técnico de prevención, proporciona información objetiva, fiabilidad en las mediciones, se puede utilizar en cualquier entorno y permite aplicar los métodos ergonómicos más relevantes generando un informe final.

Es una herramienta tecnológica para analizar y evaluar las posturas de un trabajador, genera un modelo 3D en tiempo real, fácil de calibrar, configurable para percentiles de ambos géneros y los datos obtenidos se almacenan en una base de datos.

El exoesqueleto proporciona soporte adicional a las extremidades del cuerpo, reduciendo la carga sobre las articulaciones y mejorando la movilidad de los usuarios. Además, los trabajadores del sector sanitario pueden utilizar este exoesqueleto para reducir la fatiga muscular y

prevenir lesiones musculoesqueléticas al levantar y transportar pacientes.

Según los datos de la encuesta realizada a cada trabajador, el exoesqueleto es cómodo, ligero, fácil de utilizar, ajustable al cuerpo, permite realizar todos los movimientos necesarios para llevar a cabo el trabajo, no es transpirable, genera una sensación de calor en la zona lumbar y piernas. Respecto al traje Mocap, los sensores de las manos son incómodos a la hora de realizar su trabajo.

Posteriormente se realizará el análisis biomecánico basado en los métodos más relevantes de evaluación ergonómica, para identificar riesgos de lesiones o traumatismos musculoesqueléticos de un/a trabajador/a durante la ejecución de tareas propias de su puesto de trabajo.

La inclusión del exoesqueleto en las pruebas permitirá realizar evaluaciones y comparaciones adicionales en cuanto a las posturas adoptadas y los esfuerzos ejercidos por los empleados en sus tareas con y sin su ayuda. Este dispositivo, diseñado para asistir y mejorar la biomecánica del cuerpo humano, facilita información valiosa para la investigación.

Agradecimientos

Por último, agradecer el apoyo y compromiso al SEDIC (Servicio de Diseño Industrial y Cálculo Científico) de la UPCT y al Hospital General Universitario Santa Lucía.

Referencias

- [1] Balasubramanian, K. R. (2018). Fatigue Evaluation in Manual Handling Using Surface EMG and Ergonomic Design of Trolley. *Ergonomics International Journal*. 2. 10.23880/EOIJ-16000145.
- [2] Plantard, P., Shum, H., Multon, F. (2017). Usability of corrected Kinect measurement for ergonomic evaluation in constrained environment", *International Journal of Human Factors Modelling and Simulation*, 5(4), 338.
- [3] Bleser, G., Miezal, M., Christmann, C., Bleser, G., Taetz, B., Miezal, M., Christmann, C., Steffen, D. (2017). Development of an Inertial Motion Capture System for Clinical Application. De Gruyter Oldenbourg. [Online]. 16(2), 113–129, 2017. [date of reference October 25th of 2017]. Available at: <https://www.degruyter.com/downloadpdf/j/icom.2017.16.issue-2/icom-2017-0010/icom2017-0010.pdf>
- [4] Wu, J., Wang, Z., Raghuraman, S., Prabhakaran, S., Jafari, R. (2014). Demonstration abstract: Upper body motion capture system using inertial sensors. *IPSN-14 Proceedings of the 13th International Symposium on Information Processing in Sensor Networks*, pp. 351–352.
- [5] Jing, C., Lei, Y., Jiasun, D. et al. (1994). The application of surface electromyography in the assessment of ergonomic risk factors associated with manual lifting tasks. *J. Huazhong Univ. Sci. Technol. [Med. Sci.]* 24, 552–555. <https://doi.org/10.1007/BF02911352>

Premios SEIB-FENIN

Miércoles 22 16:30 - 19:00

Paraninfo

Segmentación de patrones de marcha en pacientes con parálisis cerebral mediante aprendizaje profundo y visión artificial

A. Romo Marín¹, C. De Gregorio Bermejo¹, B. De la Calle²,
C. Simón-Martínez³, M. Martínez-Zarzuela¹

¹ Departamento de Teoría de la Señal y Comunicaciones e Ingeniería Telemática, Universidad de Valladolid, Valladolid, España, {angela.romo,carla.gregorio}@estudiantes.uva.es, mario.martinez@uva.es

² Servicio de Rehabilitación, Hospital Universitario Río Hortega, Valladolid, España, bcalleg@saludcastillayleon.es

³ Institute of Informatics, University of Applied Sciences Western Switzerland (HES-SO) Valais-Wallis, Sierre, Switzerland, cristina.simon@hevs.ch

Resumen

La parálisis cerebral (PC) ha sido reconocida como la discapacidad motora más común en la infancia. La afectación de la marcha constituye una de sus principales limitaciones, por lo que los pacientes pueden beneficiarse de análisis clínicos de la marcha instrumentados (Instrumental Gait Analysis, IGA). De entre todos los dispositivos ofertados en el mercado, la visión artificial con cámaras de infrarrojo y marcadores corporales constituye el gold standard. Sin embargo, su elevado precio y la necesidad de amplios espacios y formación de profesionales dificultan su utilización. En este trabajo, se estudió el empleo de unidades de medición inercial (Xsens) y sistemas de video sin marcadores (OpenCap) para registrar patrones de marcha y segmentar sus fases empleando redes neuronales. Con ello se pretende demostrar el potencial de ambas técnicas, que al ser mucho más económicas y rápidas que la visión artificial, podrían extender el análisis instrumentado en el ámbito clínico.

1. Introducción

El deterioro de la marcha es una de las múltiples deficiencias motoras que condicionan la actividad diaria de los niños y niñas con parálisis cerebral [1,2]. Aunque esta condición es permanente, las limitaciones que esta produce evolucionan y progresan con el paso del tiempo, necesitando de un seguimiento constante y ajustes en el tratamiento. Los análisis de la marcha basados en la inspección visual pueden dar como resultado un análisis subjetivo y aislado, que dificulta el diagnóstico correcto, lo que ha provocado que muchos de los pacientes hayan tenido que someterse a intervenciones quirúrgicas innecesarias o excesivas [2].

El análisis instrumentado de la marcha se ayuda de dispositivos electrónicos para ofrecer un estudio mucho más completo y preciso del paciente [2,3] y permite analizar múltiples parámetros (espaciotemporales, cinemáticos y cinéticos, entre otros) que ayudan a identificar y comprender a qué se deben las distintas desviaciones respecto a la marcha no patológica ideal [3]. En el mercado podemos encontrar una amplia selección de sistemas. Entre las tecnologías disponibles, la visión artificial con multicámara representa el *gold standard* en análisis biomecánicos [4]. Sistemas como Vicon u Optitrack utilizan varias cámaras para detectar y reconstruir la trayectoria seguida por una serie de marcadores, pasivos o activos, localizados en puntos anatómicos concretos. No obstante, debido a su elevado coste, sus grandes dimensiones y la necesidad de expertos

para manejar el sistema, el uso de esta tecnología no está extendido en el ámbito clínico. En España solamente en torno a 8 hospitales cuentan con sistemas de este tipo para la evaluación de pacientes.

Por otro lado, en los últimos años ha crecido el interés por sistemas alternativos de video sin marcadores y basados en unidades de medida inercial (IMU) [4]. Estos últimos son dispositivos portátiles compuestos por acelerómetros, giroscopios y magnetómetros. Ambos tipos de sistemas, además de un precio más reducido, cuentan con otra serie de ventajas frente a los sistemas *gold standard*, los cuales son difícilmente portátiles, requieren de un elevado tiempo de colocación y limitan los datos recogidos a escasos pasos por ensayo. Concretamente, los dispositivos inerciales harían posible la adquisición continua de movimientos en lugares cotidianos para el paciente, evitando así entornos artificiales, como los de un laboratorio, que lo condicionan además psicológicamente y pueden influir en su marcha [1,5]. A pesar de todas las virtudes nombradas anteriormente, ninguno es aún un elemento cotidiano para el análisis de patrones de movimiento en el entorno médico. Una de las razones la constituye la no disponibilidad de modelos de acceso abierto y suficientemente validados capaces de segmentar automáticamente la señal de marcha en sus respectivas fases.

Trabajos previos estudian la viabilidad de sistemas alternativos al *gold standard* en combinación con redes neuronales de aprendizaje profundo para la detección automática de las fases de la marcha. Existen dos enfoques principales, en función del dato de entrada aportado a la red: modelos que emplean como *input* las imágenes de movimiento captadas por los sistemas de video [6] y modelos cuya entrada parte de la información espaciotemporal recogida por IMUs [7-9]. Dentro de estos últimos, la tendencia ha sido la de emplear como variables aceleraciones y orientaciones, es decir, un total de seis canales [7-9]. Esta estrategia puede derivar en una elevada dimensionalidad, en función del número de sensores inerciales usados, aumentando en consecuencia las posibilidades de overfitting, enlenteciendo el proceso de aprendizaje y forzando al uso de redes neuronales complejas y de gran profundidad. Por ello, en este estudio se ha optado por emplear como variables los ángulos articulares, bajo la hipótesis de que estas entradas aportan más información acerca del movimiento de marcha de los

distintos sujetos, a la vez que disminuyen la dimensionalidad de los datos.

Las contribuciones de esta investigación incluyen: 1) Elaboración de una base de datos multimodal para el análisis de la marcha con dispositivos asequibles, que incluye sujetos adultos sanos y pacientes pediátricos con parálisis cerebral. Esta base de datos incluye adquisiciones en vídeo con dos cámaras (OpenCap) y con IMUs (Xsens), 2) Desarrollo de una herramienta fácil de utilizar para el etiquetado manual de las fases de la marcha a partir de la información procedente de varias cámaras, 3) Estudio comparativo de rendimiento de redes neuronales para la segmentación de las fases de la marcha a partir de los ángulos articulares de los sujetos.

2. Materiales y métodos

2.1. Descripción de la base de datos

La adquisición de datos se realizó empleando simultáneamente tecnología de captura de vídeo y sensores vestibles. Los sujetos de estudio fueron voluntarios sanos, tanto adultos como niños, y pacientes pediátricos diagnosticados de parálisis cerebral. Los adultos fueron grabados en las instalaciones de la ETSI de Telecomunicaciones de la Universidad de Valladolid de acuerdo al protocolo con referencia PI 21-2341 v3.0 aprobado por el comité de ética de investigación con medicamentos CEIm ÁREA DE SALUD VALLADOLID ESTE. Los niños en el Hospital Universitario Río Hortega de Valladolid, de acuerdo al protocolo con referencia 23-PI038, con dictamen favorable del comité CEIm ÁREA DE SALUD VALLADOLID OESTE.

Para la adquisición de muestras con sensores vestibles se hizo uso de las unidades de medición inercial Xsens MVN Awinda®. Se emplearon un total de 17 IMU para el estudio de movimientos de: foot (2), lower-legs (2), upper-legs (2), pelvis (1), hands (2), lower-arms (2), upper-arms (2), shoulders (2), chest (1), head (1). En cuanto a la captura de datos mediante los sistemas de vídeo, se utilizaron 2 Apple iPad® y la plataforma de código abierto OpenCap.

La base de datos incluye un total de 25 sujetos. Más concretamente, esta se compone de 17 sujetos de control, de edades entre 7-22 años y, de 8 sujetos pediátricos con parálisis cerebral, de edades entre 5-16 años. Los siguientes datos fueron recopilados para cada uno de los sujetos: edad, peso, altura, género, número de pie y mano y pie dominante. Además, a cada uno de ellos se le pidió realizar todas o alguna de estas tres actividades: ‘marcha’, ‘marcha larga’ y/o ‘marcha con ortesis AFO’. Concretamente se adquirieron datos de:

- 14 sujetos adultos de control. A estos se les pidió el desarrollo de la actividad de ‘marcha’.
- 3 sujetos pediátricos de control, los cuales, ejecutaron las actividades de ‘marcha’ y ‘marcha larga’. La diferencia entre ambas actividades es que, en la última de ellas, se le pide al sujeto dar un total de 12 pasos ida y 12 pasos vuelta.
- 8 sujetos pediátricos con parálisis cerebral. A parte de las actividades de ‘marcha’ y ‘marcha larga’, a

aquellos que portaban una ortesis pie-tobillo se les tomó una captura adicional de ‘marcha con ortesis AFO’.

2.2. Etiquetado de la base de datos

Para el entrenamiento de la red neuronal se siguió un enfoque de aprendizaje supervisado y para facilitar el etiquetado de datos se desarrolló una interfaz gráfica en Python empleando la librería *tkinter*.

En esta interfaz, pueden visualizarse simultáneamente los vídeos registrados por las cámaras en OpenCap, de forma que hay dos ángulos de visión distintos desde los que observar el movimiento del sujeto. Los vídeos se recorren de manera sincronizada a través de *sliders* de desplazamiento. Con el *slider* superior se selecciona el *frame* que se corresponde con el inicio de una fase de la marcha, mientras que con el inferior se indica qué muestra se asocia con el final de dicha fase. Para etiquetar ese intervalo de *frames*, es suficiente con pulsar el botón asociado a esa fase de la marcha, de forma que, todas las muestras ubicadas entre ese rango de inicio y final quedan descritas y etiquetadas con un número asociado a esa fase de la marcha.

Se consideran las siguientes etiquetas:

- 0 (No movimiento). No se corresponde con ninguna fase de la marcha.
- 1 (Apoyo Inicial). Desde que el talón del pie de referencia toca el suelo hasta que el miembro contralateral está a punto de elevarse y dejar de estar en contacto con el suelo.
- 2 (Apoyo Medio). Asociada a las muestras en las que todo el peso recae sobre el pie de referencia, al haberse levantado del suelo el otro.
- 3 (Apoyo Final). Caracteriza a aquellas muestras que van desde que el pie contralateral hace contacto con el suelo mediante el talón hasta que los dedos del pie de referencia abandonan el piso.
- 4 (Swing). Cuando el pie de referencia se encuentra levantado.

Al inspeccionar la base de datos de sujetos pediátricos con parálisis cerebral, se observó que al finalizar la fase de swing y comenzar la de apoyo inicial estos pacientes no realizaban la flexión dorsal típica de este momento de la marcha, pisando el suelo bien con la punta o bien con toda la superficie del pie a la vez. Por ello, se decidió añadir dos etiquetas a mayores, las cuales fueron usadas en algunos de los entrenamientos:

- 5 (Swing Parálisis Cerebral). Cuando el pie de referencia se dispone a finalizar la fase de swing sin encontrarse realizando una flexión dorsal.
- 6 (Apoyo Inicial Parálisis Cerebral). Asociada a un inicio patológico del apoyo inicial, donde el pie se encuentra apoyado sobre el antepié o sobre la planta, en vez de sobre el talón.

2.3. Pre-procesado de la base de datos

Como paso previo a la etapa de pre-procesado, se generaron conjuntos de muestras en función de su

etiqueta., tanto para miembro inferior izquierdo como para miembro inferior derecho.

Este primer paso se continuó con el pre-procesado. Es importante apuntar que, en esta etapa, no se utilizaron aquellos archivos correspondientes a la etiqueta 0, es decir, aquella que se asociaba con la ausencia de movimientos. Se decidió llevar a cabo esta acción dado que el número de ejemplos disponibles era muy bajo y el objetivo del estudio no era reconocer la ausencia de actividad, sino segmentar las fases de la marcha. Se realizaron las siguientes operaciones:

- Selección de características. Para el entrenamiento de la red, en este primer estudio se emplearon únicamente los datos proporcionados por OpenCap. Este proporciona un total de 35 variables correspondientes a ángulos articulares tanto del tren superior como del tren inferior del cuerpo. El conjunto de variables empleado fue variando de unos entrenamientos a otros, a fin de comprobar cuáles eran más relevantes clínicamente.
- Segmentación. Cada uno de los archivos con las secuencias de marcha se segmentaron en ventanas temporales cuya longitud y porcentaje de solapamiento fue variando en los distintos entrenamientos.

2.4. Redes neuronales de aprendizaje profundo

Para llevar a cabo la tarea de clasificación se desarrollaron dos arquitecturas de red distintas. Por una parte, una red neuronal *fully-connected* tipo Deep Neural Network (DNN) y por otro lado una red convolucional (CNN) respectivamente.

El número de neuronas de la capa de entrada vino determinado por las dimensiones del dato de entrada, es decir: la longitud de la ventana (filas) y por los ángulos articulares considerados (columnas). El número de muestras y variables fue modificándose a lo largo de los distintos entrenamientos, a fin de optimizar el rendimiento de estos últimos. Para evitar situaciones de overfitting se aplicaron los métodos de regularización *Dropout*, *Early Stopping* y *Reduce Learning Rate on Plateau*. Además, se usó Adam como optimizador, con una tasa de aprendizaje con un valor de 10^{-3} , y *sparse categorical cross entropy* como función de pérdida, usada en tareas de clasificación multi-clase.

3. Resultados y discusión

3.1. Análisis de resultados en marcha de sujeto de control

En primer lugar, centrándonos en el análisis y segmentación de la marcha de sujetos de control, tanto el modelo de red DNN como CNN fueron capaces de conseguir unos óptimos resultados, incluso cuando el número de variables y parámetros entrenables era muy reducido, concretamente 3 y 17604, respectivamente. Así pues, teniendo en cuenta los valores en las métricas de los distintos tipos de entrenamiento, se podría afirmar que, con tan solo la información de la flexión rodilla (*Knee*, *K*), tobillo (*Ankle*, *A*) y cadera (*Hip*, *H*) del pie derecho o del

pie izquierdo (KAH(R) y KAH(L)), o bien con la flexión de rodilla y tobillo tanto del lado izquierdo como derecho (KA(RL)), bastaría para clasificar correctamente las fases de la marcha en este tipo de sujetos. Muestra de ello, son las métricas de rendimiento recogidas en la Tabla 1, obtenidas durante los correspondientes entrenamientos en los que se aplicó validación cruzada.

Variables articulares	SEN	ESP	ACC	F1-SCORE
KA(RL)	0.98	1	0.99	0.98
KAH(R)	0.98	1	0.99	0.99
KAH(L)	1	1	1	1

Tabla 1. Métricas de rendimiento del clasificador de fase en pie derecho con los conjuntos de variables articulares KA(RL), KAH(R) y KAH(L).

Esto contrasta con las complejas arquitecturas utilizadas en otros trabajos relacionados, como la red FCN-BiLSTM [7] o la CNN-GRU [8]. En algunos de estos estudios, como para la Gait-Net desarrollada por Song et al. [8], utilizaban como datos de entrada imágenes, en vez de ángulos articulares, lo cual puede ser la causa de que se requiriesen redes con un mayor número de parámetros entrenables. Asimismo, en otros trabajos donde se emplearon dispositivos inerciales, utilizaron como entrada datos de 6 canales, esto es, los 3 ejes del acelerómetro y los 3 ejes del giroscopio [6-7, 9]. Ante ello, la dimensionalidad es prácticamente igual de elevada que si se utilizaran las 35 variables proporcionadas por OpenCap, aumentando en consecuencia las posibilidades de overfitting, enlenteciendo, el proceso de aprendizaje y forzando al uso de redes neuronales complejas y de gran profundidad.

Utilizando datos de vídeo hemos comprobado como en sujeto de control es posible hacer una correcta segmentación de las fases de la marcha empleando un reducido número de ángulos articulares del plano sagital. Las métricas de rendimiento obtenidas por la red se muestran en la Tabla 1, con un F1-score igual o cercano a 1. Adicionalmente, este resultado da pie a pensar que es posible realizar un estudio de la marcha empleando un número limitado de sensores. De esta manera, se lograría reducir el coste y el tiempo empleado en la colocación de los mismos. Por consiguiente, nuestra hipótesis sería que, considerando los conjuntos de variables KA(RL), KAH(R) y KAH(L) como los necesarios, únicamente sería preciso usar un total de 4 o 6 sensores. En consecuencia, se reduciría en gran medida el número inicial de 17 IMUs utilizadas para la recopilación de nuestra base de datos. Esto puede verse con mayor claridad en la Figura 1.

3.2. Análisis de la marcha en paciente con PC

Al entrenar la red con marcha de sujeto de control y evaluar su rendimiento con marcha de sujeto con parálisis cerebral, los resultados obtenidos al clasificar la fase del pie derecho fueron notablemente superiores en comparación con el clasificador de fase de pie izquierdo, tal y como puede verse en la Tabla 2. Estos resultados son lógicos, teniendo en cuenta que la mayor parte de los sujetos con parálisis

cerebral, concretamente cinco, tenían afectación en el lado izquierdo y, que en los tres restantes, a pesar de tener afectación en el lado derecho, esta no era tan marcada como la vista para el lado izquierdo, donde se podía distinguir claramente el fenómeno de marcha en punta.

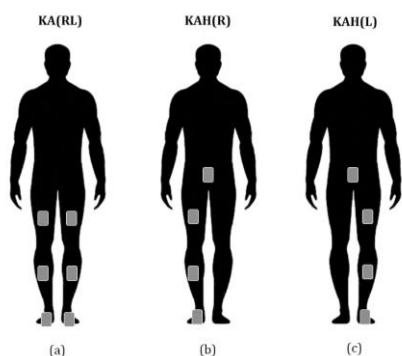


Figura 1. Sensores IMU necesarios para recolectar los conjuntos de variables (a) KA(RL), (b) KAH(R) y (c) KAH(L).

Clasificador	SEN	ESP	ACC	F1-SCORE
Derecho	0.98	0.99	0.99	0.97
Izquierdo	0.67	0.88	0.83	0.67

Tabla 2. Métricas de rendimiento del clasificador de fase en pie derecho y pie izquierdo para la variable articular KA(RL).

En cuanto al conjunto de variables disponibles, pudo verse como la información del tren superior del cuerpo dificultó el aprendizaje de la red, consecuencia de la inexistencia de un braceo fluido en los niños con parálisis cerebral. Además, igual que para la segmentación de marcha en controles, información sobre la flexión de cadera ayudaba a mejorar resultados, en esta ocasión, en la que en conjunto de test estaba conformado por sujetos con parálisis cerebral, ocurría todo lo contrario. Sin embargo, en aquellos entrenamientos en los que se empleó el conjunto de paciente pediátrico con PC tanto para entrenar como evaluar la red, añadir información sobre esta variable articular consigue mejorar los resultados. La razón de ello recae en que el movimiento de esta articulación difiere gran cantidad entre sujetos sanos y con parálisis cerebral, sin embargo, dentro de estos segundos, es bastante similar en todos.

Finalmente, añadir las etiquetas asociadas a alteraciones de marcha (5 y 6), ayudó a mejorar notablemente los resultados a la hora de clasificar la fase en la que se encontraba el pie izquierdo. Sin embargo, con ninguno de los conjuntos de variables usados fue posible lograr que las métricas de sensibilidad y F1-Score superasen el valor 0.5. La principal limitación y posible causa del bajo rendimiento de las redes con pacientes PC es el tamaño del *dataset* disponible. Nuestras redes solamente han podido extraer patrones y características de un total de 8 pacientes con PC, dificultando así su capacidad de aprendizaje.

4. Conclusiones

En este estudio se ha comprobado que, para la clasificación de las fases de la marcha en sujeto sano con información

de vídeo, estructuras simples de red (DNN, CNN) logran un buen rendimiento. Adicionalmente, dados los resultados obtenidos a partir de imágenes de vídeo, resulta factible pensar que es posible alcanzar una óptima tarea de segmentación empleando únicamente 4 o 6 sensores inerciales. De esta manera, se reduciría el precio, así como el tiempo requerido para la colocación de los mismos. Por otro lado, para segmentar la marcha en paciente con PC, habiendo entrenado previamente la red con sujeto sano se ha podido comprobar cómo es más adecuado no considerar las variables del tren superior del cuerpo ni de la cadera. En consecuencia, es necesario implementar modelos de clasificación entrenados de forma independiente con pacientes con esta condición, necesitando de la creación de nuevas etiquetas cada vez que se dé de algún tipo de desviación de la marcha.

Agradecimientos

Esta investigación ha sido financiada parcialmente por proyectos del Ministerio de Ciencia e Innovación (124515OA-100), la Gerencia Regional de Salud de Castilla y León (GRS 2670/A/22) y la Consejería de Empleo e Industria de Castilla y León (INVESTUN/21/VA/0003). C. Simón-Martínez cuenta con el apoyo del acuerdo de subvención Marie Skłodowska-Curie (890641).

Referencias

- [1] Bekteshi, S., Monbaliu, E., McIntyre, S., Saloojee, G., Hilberink, S. R., Tatishvili, N., & Dan, B. (2023). Towards functional improvement of motor disorders associated with cerebral palsy. *The Lancet Neurology*, 22(3), 229-243. [https://doi.org/10.1016/S1474-4422\(23\)00004-2](https://doi.org/10.1016/S1474-4422(23)00004-2).
- [2] Nowicki, P. D. (Ed.). (2020). *Orthopedic Care of Patients with Cerebral Palsy: A Clinical Guide to Evaluation and Management across the Lifespan*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-46574-2>.
- [3] Armand, S., Decoulon, G., & Bonnefoy-Mazure, A. (2016). Gait analysis in children with cerebral palsy. *EFORT Open Reviews*, 1(12), 448-460. <https://doi.org/10.1302/2058-5241.1.000052>.
- [4] Tunca, C., Pehlivan, N., Ak, N., Arnrich, B., Salur, G., & Ersoy, C. (2017). Inertial Sensor-Based Robust Gait Analysis in Non-Hospital Settings for Neurological Disorders. *Sensors*, 17(4), Article 4. <https://doi.org/10.3390/s17040825>.
- [5] Gao, S., Chen, J., Dai, Y., & Hu, B. (2022). *Wearable Systems Based Gait Monitoring and Analysis*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-97332-2>.
- [6] Song, C., Huang, Y., Huang, Y., Jia, N., & Wang, L. (2019). GaitNet: An end-to-end network for gait based human identification. *Pattern Recognition*, 96, 106988. <https://doi.org/10.1016/j.patcog.2019.106988>.
- [7] Rifaat, N., Ghosh, U. K., & Sayeed, A. (2022). Accurate gait recognition with inertial sensors using a new FCN-BiLSTM architecture. *Computers and Electrical Engineering*, 104, 108428. <https://doi.org/10.1016/j.compeleceng.2022.1084-8>.
- [8] Martindale, C. F., Christlein, V., Klumpp, P., & Eskofier, B. M. (2021). Wearables-based multi-task gait and activity segmentation using recurrent neural networks. *Neurocomputing*, 432, 250-261. <https://doi.org/10.1016/j.neucom.2020.08.079>.

Preliminary screening of polycaprolactone-polystyrene electrospun scaffolds for the formation of tumoroids

M. Pareja Jiménez¹, E. Casanova Batlle², J. Ciurana Gay², E. Polonio Alcalá^{1,*}, and T. Puig Miquel^{1,*}

¹ New Therapeutic Targets Laboratory (TargetsLab) – Oncology Unit, Medical Sciences Department, University of Girona, Girona, Spain; martaparjim@gmail.com, emma.polonio@udg.edu, teresa.puig@udg.edu

² Product, Process and Production Engineering Research Group (GREP), Department of Mechanical Engineering and Industrial Construction, University of Girona, Girona, Spain; enric.casanova@udg.edu, quim.ciurana@udg.edu

Abstract

Triple negative breast cancer (TNBC) subtype is the most aggressive and conventional chemotherapy is the standard treatment since lacks validated targeted therapies. Thus, there is an unmet need for improved in vitro research tools, including scaffolds for the formation of tumoroids. In this preliminary study, the main aim was to assess the tumoroid-forming potential of polycaprolactone (PCL)/polystyrene (PS) scaffolds. Six different polymer solutions, varying PCL/PS ratios, were formulated and PCL/PS scaffolds were manufactured using the electrospinning technique. PCL/PS structures were characterized by SEM imaging. Subsequently, a sterilization assay was performed to evaluate potential alterations derived from this process. Afterwards, TNBC MDA-MB-231 cells were cultured on PCL/PS structures for 10 days. The findings revealed that cells cultured on PCL 100/PS 0, PCL 80/PS 20, and PCL 60/PS 40 matrices exhibited the formation of tumoroids. In conclusion, PCL/PS scaffolds offer a valuable in vitro tool for tumoroid formation.

1. Introduction

Breast cancer is the foremost incidence cancer and the leading cause of cancer-related mortality among women worldwide [1]. Triple negative breast cancer (TNBC) subtype is characterized by the absence of the overexpression of estrogen, progesterone, and/or human epidermal growth factor receptor-2 (HER2) receptors [2]. This aggressive subtype predominantly affects young women and manifests a poor prognosis. Conventional treatment for TNBC is chemotherapy, often resulting in favorable outcomes. However, approximately 30% of patients experience relapse within five years, leaving them with no therapeutic options. Thus, the pursuit of novel biomarkers and therapeutic approaches for TNBC is essential.

Monolayer cell culture is a cost-effective system traditionally used by researchers. Nevertheless, this technique does not accurately mimic physiological conditions. To bridge this gap, researchers have developed different three-dimensional (3D) cell culture strategies [3,4]. Among these, electrospinning emerges as an additive manufacturing technique that enables the fabrication of fibers closely resembling the dimensions of the extracellular matrix. Additionally, this technique offers the advantage of using biomaterials [5].

Therefore, the main aim of this study was to conduct an initial screening to explore the ability of various electrospun scaffolds to form tumoroids using the TNBC MDA-MB-231 cell line.

2. Material and Methods

2.1. Chemicals and reagents

Polycaprolactone (PCL, Mn 80,000 g/mol), polystyrene (PS, Mw 192,000 g/mol), dichloromethane (DCM, ≥ 99% (GC)), paraformaldehyde, glutaraldehyde, sodium cacodylate, Triton™ X-100, and bovine serum albumin (BSA) (≥ 98.0%) were purchased from Sigma-Aldrich (St. Louis, MO, USA). N-N dimethylformamide (DMF; 99.9% GLR) and ethanol absolute (AGR) were obtained from Labkem, Labbox Labware S.L. (Barcelona, Spain). BSA Fraction V pH for Western blotting (> 96%) was purchased from PanReac AppliChem (Gatersleben, Germany). Dulbecco's Modification of Eagle's Medium (DMEM), penicillin/streptomycin (10,000 U/mL), and phosphate-buffered saline (PBS) were obtained from Corning (New York, USA). Fetal bovine serum (FBS) was purchased from HyClone (Logan, UT, USA). Trypsin (10×) was obtained from Dutscher (Bernolsheim, France). Rhodamine-phalloidin was purchased from Cytoskeleton Inc. (Denver, CO, USA) and 4,6-diamidino-2-phenylindole (DAPI) was obtained from BD Pharmingen (Franklin Lakes, NJ, USA).

2.2. The manufacturing of electrospun fibers

Different combinations of PCL and PS were dissolved in DCM-DMF (1:3) at a concentration of 15% (w/v) for 24 hours at 50 °C with continuous agitation. Scaffolds were manufactured using an electrospinning device (Spraybase, Dublin, Ireland). The polymeric solution was transferred to 20 mL syringe, connected via a polytetrafluoroethylene tube (inner diameter of 1 mm) to stainless steel 22G needle (inner diameter of 0.41 mm). The electrospinning process was controlled by the Syringe Pump Pro software (New Era Pump Systems; Farmingdale, NY, USA) (Table 1) and 1.5 mL of solution was ejected. These parameters were chosen to ensure the formation of the Taylor cone consistently throughout the electrospinning process. Moreover, environmental conditions, including temperature and humidity, remained stable, with temperatures ranging between 21 °C and 25°C and

humidity levels maintained at 52% to 55%. The resulting structures were kept at room temperature for at least 24 h for a correct solvent evaporation and cut into squares of 2.56 or 6.25 cm², depending on the specific experiment requirements.

Polymer Solution		Flow Rate (mL/h)	Voltage (kV)	Distance (cm)
PCL (%)	PS (%)			
100	0	1.5	10.0	9
80	20	3.0	10.5	10
60	40	1.5	8.4	10
40	60	3.0	9.2	10
20	80	1.0	11.1	10
0	100	3.0	16.4	10

Table 1. Electrospinning Parameters

2.3. Scaffold microarchitecture

Scaffolds, both non-sterilized and sterilized, underwent a carbon coating process using a K950 turbo evaporator (Emitech, Kent, UK). Observations were conducted using a Hitachi S4100 field emission scanning electron microscopy (SEM; Hitachi, Tokyo, Japan). Images were digitally captured by Quartz PCI software (Quartz, Vancouver, Canada). Fiber diameter and pore area were determined through MATLAB software (MathWorks; Natick, MA, USA). To elaborate, filament diameter was assessed by measuring the distance of the perpendicular distance between the top and bottom ends of the filaments. Pore area was calculated through manual delineation of pore shapes on the images. Eight filaments and three pores were randomly selected from each photo for analysis.

2.4. Sterilization assay

Two methodologies were employed for this assay: exposure to UV light and immersion in 70% ethanol. Five combinations of these methods were explored:

1. 30 minutes of UV light exposure, for each side.
2. 4 hours of immersion in 70% ethanol.
3. 4 hours of immersion in 70% ethanol combined with 30 minutes of UV light exposure, for each side.
4. 16 hours of immersion in 70% ethanol.
5. 16 hours of immersion in 70% ethanol combined with 30 minutes of UV light exposure, for each side.

Following the sterilization process, the tested PCL/PS scaffolds were washed two times with PBS and subsequently placed in non-adherent 6-well plates. Supplemented medium devoid of antibiotics was added into each well and incubated for 7 days to facilitate the detection of any potential contaminant. Non-sterilized scaffolds were used as positive controls.

2.5. Cell Line

The MDA-MB-231 human TNBC cell line was procured from the American Type Culture Collection (ATCC, Rockville, MD, USA). The cells were cultured routinely in DMEM supplemented with 10% FBS and 50 U/mL of penicillin/streptomycin. Cultures were maintained at 37°C in a controlled atmosphere with 5% CO₂. Regular monitoring was conducted, and mycoplasma contamination was systematically assessed to ensure cell line integrity.

2.6. Three-Dimensional Cell Culture

Sterilized scaffolds were placed in non-adherent cell culture 12-well plates, immersed in culture medium, and incubated for a minimum of 30 minutes to facilitate cell adhesion. The pertinent cell density was prepared in a 50 µL volume of medium. Subsequently, MDA-MB-231 cells were seeded on scaffolds following the procedure described elsewhere [6].

2.7. Evaluation of tumoroid formation

MDA-MB-231 cells were cultured on PCL/PS scaffolds for 10 days.

For immunofluorescence imaging, the samples were fixed with a 4% paraformaldehyde solution (w/v), permeabilized using 0.2% Triton™ X-100 (v/v), and blocked by a 3% BSA solution (w/v). Staining was performed using rhodamine-phalloidin (1:250) and DAPI (1:1000). Fluorescence was observed using an A1R confocal laser scanning microscope (CLSM; Nikon, Tokyo, Japan). Images were taken through Nikon NIS-Elements AR v4.10 software (Nikon).

For SEM analysis, the samples were fixed with a 2.5% glutaraldehyde solution (v/v) in 0.1 M sodium cacodylate buffer at pH 7.4, washed in 0.1 M sodium cacodylate buffer, and dehydrated using a graded series of ethanol concentrations (50%, 75%, 80%, 90%, 95%, and 100%). The 3D structures were dried using K850 CPD critical point dryer (Emitech) and coated with gold by K950 turbo evaporator (Emitech). Observations were made using a Hitachi S4100 field emission SEM (Hitachi, Tokyo, Japan).

3. Results & Discussion

3.1. The characterization of scaffolds

The microarchitecture of PCL/PS scaffolds was examined using SEM (Figure 1) to assess both pore area and fiber diameter (Table 2).

Examination of the PCL/PS matrices revealed their composition as fiber-rich structures. As illustrated in Figure 1, the fibers became rougher as the PS percentage increased. No beads were detected in any of the scaffolds. The PCL/PS scaffolds showed a similar pore area, ranging from approximately 20 to 56 µm². However, all structures displayed a notable diversity in pore size. Notably, the filaments in the PCL 100/PS 0 scaffolds were significantly thinner compared to those in the PCL 60/PS 40 ($p=0.0001$)

and PCL 40/PS 60 structures ($p=0.0001$). Similar findings were observed in the PCL 0/PS 100 matrices (60/40, $p=0.003$; 40/60, $p=0.002$).

PCL/PS (%)	Pore Area (μm^2)	Fiber Diameter (μm)
100/0	20.27± 6.32	0.82± 0.15
80/20	32.36± 9.53	1.17± 0.08
60/40	56.47± 18.91	1.48± 0.08
40/60	31.81± 5.99	1.73± 0.19
20/80	27.09± 10.82	1.15± 0.10
0/100	41.09± 12.37	0.85± 0.07

Table 2. Pore area and fiber diameter of PCL/PS scaffolds.

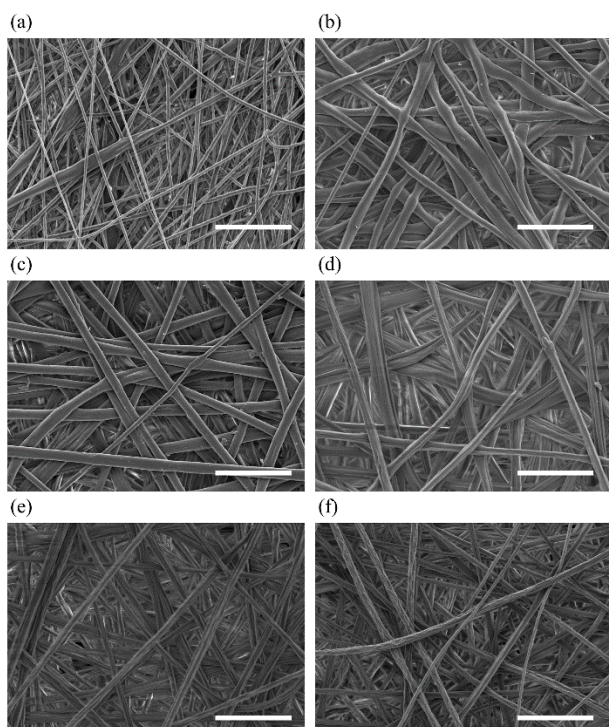


Figure 1. Scanning electronic microscopy (SEM) pictures from (a) PCL 100/PS 0 scaffolds, (b) PCL 80/PS 20 scaffolds, (c) PCL 60/PS 40 scaffolds, (d) PCL 40/PS 60 scaffolds, (e) PCL 20/PS 80 scaffolds, and (f) PCL 0/PS 100 scaffolds. Scale bar: 10 μm .

3.2. The selection of sterilization process

The selection of sterilization methods and their duration is a critical consideration, as it can potentially alter the scaffold [7]. In our study, we investigated two distinct sterilization approaches: UV light exposure and immersion in 70% ethanol. Specifically, we examined five combinations of these methods, as detailed in Section 2.

Our findings (Table 3) demonstrated that a 4-hour immersion in 70% ethanol alone (4h EtOH) and combined with 30 minutes of UV light exposure (for each side) (4h EtOH+UV) resulted in a contamination-free medium and represented the least invasive sterilization approaches. Afterwards, pore area (Table 4) and fiber diameter (Table

5) of 4h EtOH and 4h EtOH+UV were analyzed using SEM images.

PCL/PS (%) / Sterilization process	0/100	20/80	40/60	60/40	80/20	100/0
30 min UV	X	○	○	○	○	X
4 h EtOH	X	○	○	○	○	X
4 h EtOH + UV	X	○	○	○	○	X
16 h EtOH	X	○	○	○	○	X
16 h EtOH + UV	X	○	○	○	○	X
Control Positive	X	○	○	○	○	X

Table 3. Sterilization assay findings of PCL/PS scaffolds. X represents contaminated medium and ○ non-contaminated medium.

PCL/PS (%)	4h EtOH	4h EtOH + UV
100/0	39.71± 9.65	58.66± 26.77
80/20	40.45± 12.09	96.21± 23.08
60/40	64.65± 10.03	60.64± 17.10
40/60	125.56± 25.23	67.94± 24.97
20/80	97.04± 20.96	94.72± 17.64
0/100	95.94± 16.96	109.89± 40.11

Table 4. Pore area (μm^2) of PCL/PS scaffolds.

PCL/PS (%)	4h EtOH	4h EtOH + UV
100/0	0.67± 0.06	0.91± 0.11
80/20	1.17± 0.08	1.43± 0.11
60/40	1.34± 0.11	1.35± 0.11
40/60	1.90± 0.19	1.97± 0.08
20/80	1.08± 0.10	1.36± 0.18
0/100	1.27± 0.08	1.52± 0.14

Table 5. Filament diameter (μm) of PCL/PS scaffolds.

Regarding the pore area (Table 4), the 4h EtOH+UV sterilization process led to a significant increase in PCL 20/PS 80 structures ($p=0.021$), while the 4h EtOH sterilization induced significant enlargement in PCL 40/PS 60 scaffolds ($p=0.018$) compared to non-sterilized scaffolds (Figure 2).

Regarding the fiber diameter (Table 5), the 4 EtOH ($p=0.006$) and 4h EtOH+UV ($p=0.0001$) sterilization processes led to a significant increase in PCL 0/PS 100 structures compared to non-sterilized scaffolds.

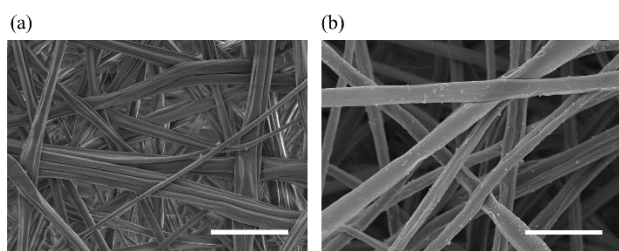


Figure 2. Scanning electron microscopy (SEM) pictures from PCL 40/PS 60 scaffolds (a) before and (b) after 4-hour immersion in 70% ethanol sterilization process. Scale bar: 10 μm .

3.3. Cell culture and tumoroid formation

Tumoroids are 3D cell aggregates designed to replicate the architecture and function of tumors, offering a more faithful representation of the tumor microenvironment [8]. Our findings demonstrated the formation of tumoroids in PCL 100/PS 0, PCL 80/PS 20, and PCL 60/PS 40 scaffolds (Figure 3). In contrast to the results of Abadi *et al.* [9], who suggested that rougher filament surfaces might enhance cell attachment, our research revealed the opposite. The fibers became rougher as the PS percentage increased, leading to a reduction in cell attachment. Interestingly, Girard *et al.* demonstrated the formation of tumoroids using mPEG-PLA-PLGA, which are polymer materials more inaccessible and expensive than those utilized in our study [10].

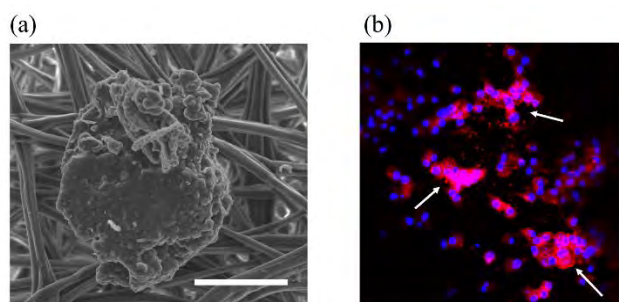


Figure 3. (a) Scanning electron microscopy (SEM) picture from PCL 60/PS 40 and (b) Immunofluorescence picture from PCL 100 /PS 0.

4. Conclusion

To sum up, PCL/PS scaffolds allow the formation of tumoroids *in vitro*, specifically PCL 100/PS 0, PCL 80/PS 20 and PCL 60/PS 40 matrices. Therefore, this study establishes the foundation for the further development and exploration of these polymeric structures as a tool for *in vitro* investigations of TNBC. This approach will permit a more faithful emulation of physiological conditions, enabling more precise analysis of drug responses and biomarker studies.

Acknowledgements

This work was funded by PLEC2021-007523 / AEI / 10.13039/501100011033 and by the European Union – NextGenerationEU. It was also supported by Spanish grant from Instituto de Salud Carlos III and co-funded by European Union ERDF/ESF, “A way to make Europe”/ “Investing in your future”(PI19/00372). The authors thank

the support of Catalan Government (2021SGR01589) and Oncolliga Foundation and RadicalSwim (OncoSwim). The authors thank Research Technical Services from the University of Girona.



References

- [1] World Health Organization (WHO). Global Cancer Observatory (GLOBOCAN). Cancer Today 2020. <https://gco.iarc.fr/today/home> (accessed September 14, 2023).
- [2] Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, et al. Breast cancer. *Nature Reviews Disease Primers*, vol 5, sup 1, 2019; pp 1–31.
- [3] Boucherit N, Gorvel L, Olive D. 3D Tumor Models and Their Use for the Testing of Immunotherapies. *Front Immunol*, vol 10, sup 11, 2020; pp 603640.
- [4] Hu S, Dasbiswas K, Guo Z, Tee YH, Thiagarajan V, Hersen P, et al. Long-range self-organization of cytoskeletal myosin II filament stacks. *Nat Cell Biol*, vol 19, sup 2, 2017, pp 133–41.
- [5] Reneker DH, Chun I. Nanometre diameter fibres of polymer, produced by electrospinning. *Nanotechnology*, vol 7, 1996, pp 216–23.
- [6] Polonio-Alcalá E, Rabionet M, Gallardo X, Angelats D, Ciurana J, Ruiz-Martínez S, et al. PLA electrospun scaffolds for three-dimensional triple-negative breast cancer cell culture. *Polymers (Basel)*, vol 11, sup 5, 2019, pp 916.
- [7] Polonio-Alcalá E, Casanova-Batlle E, Puig T, Ciurana J. The solvent chosen for the manufacturing of electrospun polycaprolactone scaffolds influences cell behavior of lung cancer cells. *Sci Rep*, vol 12, sup 1, 2022, pp 19440.
- [8] Xu H, Jiao D, Liu A, Wu K. Tumor organoids: applications in cancer modeling and potentials in precision medicine. *Journal of Hematology & Oncology*, vol 15, sup 1, 2022, 1–20.
- [9] Abadi FJH, Tehran MA, Zamani F, Nematollahi M, Mobarakeh LG, Nasr-Esfahani MH. Effect of nanoporous fibers on growth and proliferation of cells on electrospun poly (ϵ -caprolactone) scaffolds. *International Journal of Polymeric Materials and Polymeric Biomaterials*, vol 63, sup 2, 2013, pp 57–64.
- [10] Girard YK, Wang C, Ravi S, Howell MC, Mallela J, Alibrahim M, et al. A 3D fibrous scaffold inducing tumoroids: a platform for anticancer drug development. *PLoS One*, vol 8, sup 10, 2013, pp e75345.

Modelos basados en redes neuronales artificiales para la generación y predicción de series temporales de glucosa

R. González-Domínguez ¹, D. Chushig-Muzo ¹, C. Soguero-Ruíz ¹,

¹ Departamento de Teoría de la Señal y las Comunicaciones y Sistemas Telemáticos, Universidad Rey Juan Carlos
r.gonzalezdo.2019@alumnos.urjc.es, {david.chushig,cristina.soguero}@urjc.es

Resumen

La Diabetes Mellitus (DM) es una enfermedad crónica en constante aumento. Según la Organización Mundial de la Salud, el número de personas con DM aumentó de 108 millones en 1980 a 463 millones en 2023, y se estima 700 millones de nuevos casos en 2045. El uso de dispositivos medidores continuos de glucosa (CGM, del inglés *Continuous Glucose Monitoring*) ha permitido un mejor control de los niveles de glucosa en personas con DM. Este trabajo tiene dos objetivos principales. Primero, validar la efectividad de modelos basados en redes neuronales (ANN, del inglés *Artificial Neural Networks*) para la generación de series temporales sintéticas de glucosa, específicamente, *Conditional Probabilistic Auto-Regressive (CPAR)* y *DoppelGANger (DGAN)*. Segundo, evaluar tres modelos novedosos de ANNs para predecir valores de glucosa en diferentes ventanas temporales, en particular *Neural Basis Expansion Analysis Time Series Forecasting (N-BEATS)*, *Temporal Convolutional Network (TCN)* y *Neural Hierarchical interpolation for Time Series (N-HITS)*. Los experimentos fueron realizados considerando datos de glucosa de CGM de un paciente diabético real. Los resultados indican que los valores de glucosa creados con DGAN son más similares a los generados por CPAR, siendo más preciso y robusto. En la estimación de valores de glucosa, N-HITS presenta un rendimiento superior tanto en ventanas cortas como largas, alcanzado un $MSE=119.6961$ y $MRAE=0.0803$ en una ventana de predicción de 8 horas. Este modelo representa un avance en la predicción de valores de glucosa, ayudando en la prevención de episodios de hipoglucemia e hiperglucemia.

1. Introducción

Las enfermedades crónicas (ECs) suponen un gran desafío sanitario y socioeconómico a nivel mundial. Las ECs se caracterizan por un progreso lento y prolongado y según cifras de la Organización Mundial de la Salud son la causa del 71 % del número total de muertes cada año globalmente [1]. La diabetes tipo 1 (DT1) es una de las ECs más prevalentes, con una frecuencia de 5,9 por 10.000 personas globalmente [2]. La DT1 aumenta la concentración de glucosa en sangre causado por la destrucción de las células beta en el páncreas, conduciendo a una caída en la producción de insulina [2]. Para su control, los pacientes con DT1 requieren terapia con insulina exógena y un control continuo de glucosa.

Debido a los avances tecnológicos y recursos invertidos en investigación clínica han surgido los dispositivos *Continuous Glucose Monitoring (CGM)*, los cuales permiten un control continuo de glucosa, y permitiendo prevenir episodios de hipoglucemia/hiperglucemia [3]. Dada la gran cantidad de datos que proporcionan los dispositivos CGM, en los últimos años, ha habido un au-

mento significativo de estudios en el ámbito de la salud que emplean modelos de aprendizaje automático (ML, del inglés *Machine Learning*) y modelos basados en Redes Neuronales Artificiales (ANN, del inglés *Artificial Neural Networks*). A pesar de los beneficios de estos modelos, en la investigación clínica la disponibilidad de datos está sujeta a diversas regulaciones y aprobaciones por diversos comités éticos. Esto restringe sustancialmente el acceso a datos clínicos y obstaculiza el desarrollo de sistemas predictivos y de apoyo a la toma de decisiones clínicas. La generación de datos sintéticos es clave para abordar estos retos, y fomentar nuevas líneas de investigación. En los últimos años, los modelos basados en redes generativas adversarias (GANs, del inglés *Generative Adversarial Networks*) han demostrado ser eficientes para generar datos sintéticos, especialmente en aplicaciones de visión por computadora [4]. Varios autores han explorado el uso de GANs en el ámbito clínico [4], sin embargo pocos trabajos han propuesto GANs para generar series temporales sintéticas.

Este trabajo tiene un doble objetivo. Primero, evaluar dos modelos de generación de series temporales basados en ANNs para crear datos de glucosa asociados con pacientes diagnosticados con DT1. Específicamente *Conditional Probabilistic Auto-Regressive model (CPAR)* y *DoppelGANger (DGAN)* son considerados. Segundo, analizar la efectividad de modelos basados en ANNs para la predicción de valores de glucosa en diferentes ventanas temporales. En concreto, se consideraron tres modelos, *Neural Basis Expansion Analysis Time Series Forecasting (N-BEATS)*, *Temporal Convolutional Network (TCN)* y *Neural Hierarchical interpolation for Time Series (N-HITS)*.

2. Materiales y métodos

En esta sección se describe la base de datos, y los modelos para la generación de series temporales sintéticas y predicción de valores de glucosa en ventanas temporales.

2.1. Base de datos

La base de datos que se ha utilizado para el desarrollo de este trabajo ha sido proporcionada por un paciente anónimo con DT1 y usando el dispositivo CGM Abbott FreeStyle Libre. Los valores fueron recopilados entre el 25/12/2019 y el 31/12/2021, con una frecuencia de muestreo regular de 5 minutos. A esta frecuencia de muestreo, en 24 horas un total de 288 registros de glucosa se recopilaron. La base de datos cuenta con un total de 191,156 registros, y los valores de glucosa se

encuentran en el intervalo [12-400] mg/dL.

2.2. Modelos de generación de series temporales sintéticas

Los modelos de generación de secuencias temporales considerados en este trabajo son CPAR y DGAN. CPAR es un modelo autoregresivo probabilístico y condicional diseñado para modelar datos multi-secuenciales [5], basado en ANNs que estima los parámetros para crear cada elemento de cada secuencia, y actualiza la red con las estimaciones [5]. Por otro lado, DGAN es un modelo tipo GAN [6] que genera muestras en bloque, reduciendo así el coste computacional y la probabilidad de que el modelo «aprenda» datos [6]. DGAN utiliza tres redes generadoras (RG) para crear datos [6]: (i) una RG para crear atributos estáticos; (ii) una RG para generar el mínimo y el máximo de cada serie temporal; y (iii) una red recurrente que usa los datos generados por las dos redes anteriores para generar mediciones. Un red discriminadora distingue entre valores reales y sintéticos.

Para evaluar la similitud estadística y temporal de los datos temporales sintéticos generados por CPAR y DGAN se han considerado las siguientes métricas: divergencia de Jensen-Shannon (JSD) [7], test de Kolmogorov-Smirnov [7], y *Dynamic Time Warping* (DTW) [7]. JSD mide la similitud entre la *probability density function* (PDF) de la serie temporal real y la PDF de la serie sintética. Valores cercanos a 0 en JSD indica que las PDFs comparadas son más similares. El test de Kolmogorov-Smirnov es una prueba no paramétrica que determina si dos PDFs son similares cuando se cumple la hipótesis nula \mathcal{H}_0 , es decir cuando $p_{value} < 0.05$. DTW es una medida de similitud que encuentra una alineación óptima entre dos series temporales y minimiza los efectos del desplazamiento en el tiempo al permitir la transformación elástica de las series temporales para detectar similitudes. Valores más bajos en DTW indican series temporales más similares [7].

2.3. Modelos neuronales para estimación de valores de glucosa en diferentes ventanas temporales

En este trabajo, se usan los modelos N-BEATS, TCN, y N-HITS para la predicción de valores de glucosa. N-BEATS es un modelo neuronal basado en el perceptrón multicapa con enlaces residuales, y utiliza una técnica llamada expansión de bases (del inglés, *basis expansion*) para descomponer la serie temporal original en componentes (más) básicos [8]. El modelo está formado por diferentes *stacks*, y cada uno de ellos realiza una predicción parcial. La estimación final es obtenida como una combinación lineal de las predicciones parciales realizada por los *stacks* [8].

TCN, una variante de una red neuronal convolucional, utiliza una combinación de ANNs profundas con capas residuales y se compone de capas convolucionales dilatadas con la misma longitud de entrada y de salida [9]. N-HITS es una extensión del modelo N-BEATS que mejora la precisión en las predicciones y reduce el coste compu-

tacional para el entrenamiento de la red [10]. Esto se logra mediante un muestreo de la serie temporal a diferentes tasas. De esta forma, el modelo aprende tanto a corto como a largo plazo en la serie temporal, y al generar las predicciones, combina las estimaciones realizadas a diferentes escalas de tiempo, proceso denominado interpolación jerárquica [10].

3. Resultados

En esta sección se presentan los resultados de la evaluación de los modelos generativos de series temporales sintéticas, y los resultados de los modelos usados para la predicción de valores de glucosa.

3.1. Evaluación de series temporales sintéticas

En este trabajo, se optó por generar 5 series temporales sintéticas utilizando los modelos CPAR y DGAN. Cada una de las series temporales representa los datos de glucosa de un paciente sintético. Este enfoque permite comparar el rendimiento de los modelos de generación de datos para diferentes pacientes sintéticos. En la Figura 1 se muestran los valores de JSD entre la PDF de los datos de glucosa reales y la PDF de los valores de glucosa sintéticos generados con CPAR y DGAN para 5 series temporales sintéticas. Tanto con DGAN como con CPAR, se obtuvo $p_{value} < 0.05$ en todos los casos, por lo tanto se acepta la hipótesis nula \mathcal{H}_0 , es decir, ambas distribuciones son estadísticamente similares.

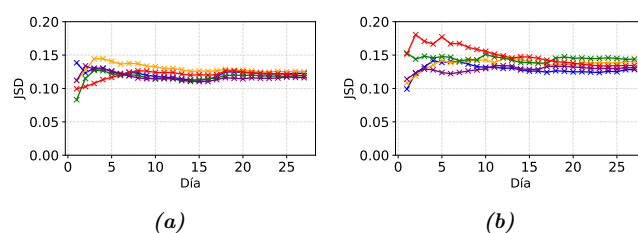


Figura 1: JSD entre los valores reales y sintéticos de glucosa generados con: (a) DGAN; y (b) CPAR para 5 pacientes sintéticos. JSD para cada paciente sintético es representado por un color distinto.

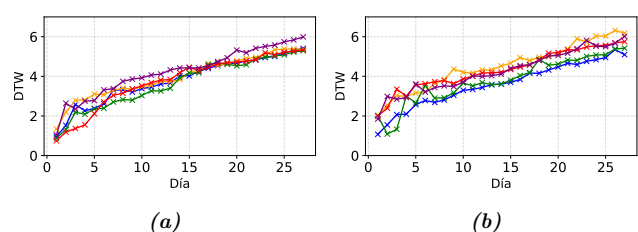


Figura 2: DTW entre la serie temporal real y varias series temporales sintéticas de glucosa generadas con (a) DGAN; y (b) CPAR para 5 pacientes sintéticos. DTW para cada paciente sintético se representa por un color distinto.

Para evaluar la similitud temporal entre series reales y sintéticas, se usa DTW. En la Figura 2 se representan

los valores de DTW obtenidos entre [1 – 28] días entre la serie temporal real y diferentes series temporales sintéticas. Los valores de DTW entre la serie real y cada una de las sintéticas generadas con DGAN, son similares a los valores de DTW obtenidos entre la serie real y cada una de las series sintéticas generadas con CPAR (véase Figura 2 (a) y Figura 2 (b), respectivamente).

3.2. Estimación de valores de glucosa

En esta sección se muestran los resultados obtenidos en la predicción de valores de glucosa usando una serie temporal real/sintética y en diferentes ventanas temporales (entre 1 y 8 horas). DGAN se seleccionó para generar series temporales sintéticas para su uso en la predicción de niveles de glucosa debido a que las secuencias generadas por este modelo exhibieron una mayor similitud con la serie temporal real tanto en términos estadísticos como temporales. Para evaluar los errores entre valores estimados y reales/sintéticos, tres figuras de mérito son consideradas: *Mean Square Error* (MSE), *Mean Absolute Error* (MAE) y *Mean Relative Absolute Error* (MRAE). En la Tabla 1 se muestran los errores obtenidos usando la serie real de glucosa. Como se observa, el modelo N-HITS muestra valores más bajos de MSE, MAE y MRAE (indicando menor error entre valores reales y estimados) que los modelos N-BEATS y TCN. En general, se observa que N-HITS presenta un mejor rendimiento en la predicción de valores de glucosa a corto (2 horas) y largo plazo (8 horas) en comparación con N-BEATS y TCN. Además, en la Tabla 1, se observa que los errores se van estabilizando a partir de una ventana de estimación superior a 4 horas, en particular, en N-HITS con valores de MAE y MRAE en 9 y 0.08, respectivamente.

En la Tabla 2, se muestran los errores calculados para cada uno de los modelos al considerar series temporales sintéticas generadas por el modelo DGAN. Se observa que N-BEATS demuestra un desempeño superior a los modelos TCN y N-HITS. Este hecho se fundamenta en los valores más reducidos de MSE, MAE y MRAE que presenta N-BEATS en relación con los otros modelos. En consecuencia, las estimaciones efectuadas por N-BEATS se asemejan de manera más precisa a los valores reales en comparación con las realizadas por TCN y N-HITS. Este resultado se manifiesta tanto en el corto plazo (2 horas) como en el largo plazo (8 horas), destacando el rendimiento de N-BEATS en la tarea de estimación de los niveles de glucosa en series temporales sintéticas. Es notable observar que en general los valores MSE, MAE, MRAE de la Tabla 2 son más bajos que aquellos de la Tabla 1. Los errores bajos en la estimación de valores de glucosa considerando una serie temporal sintética puede atribuirse principalmente a dos factores clave: la diferencia en complejidad y estructura de los datos, ya que las series temporales reales son más complejas, y la capacidad de los modelos de generación para generar datos, esto es, que la serie temporal sintética presentará una dispersión de los valores más baja.

	# registros	horas	MSE	MAE	MRAE
N-BEATS	12	1	124.9874	9.7183	0.0692
	24	2	84.2873	7.3130	0.0575
	36	3	356.4476	14.4105	0.1188
	48	4	505.4650	17.6894	0.1415
	60	5	430.7703	16.1206	0.1333
	72	6	402.0386	15.6217	0.1304
	84	7	366.8289	14.8233	0.1241
	96	8	332.4179	13.9128	0.1177
TCN	12	1	223.8272	12.0284	0.0848
	24	2	261.6353	13.3752	0.1099
	36	3	228.5306	12.3896	0.1027
	48	4	261.5450	12.7865	0.1037
	60	5	338.6953	14.4211	0.1248
	72	6	399.0167	16.1448	0.1433
	84	7	394.1630	16.1751	0.1431
	96	8	456.9607	17.5650	0.1582
N-HITS	12	1	36.5229	4.9912	0.0358
	24	2	185.4154	11.3488	0.0975
	36	3	208.3286	12.6056	0.1086
	48	4	167.7309	10.7605	0.0920
	60	5	148.3569	10.1516	0.0891
	72	6	135.4272	9.8014	0.0871
	84	7	119.6961	9.0580	0.0803
	96	8	120.7340	9.3044	0.0836

Tabla 1: MSE, MAE y MRAE obtenidos en diferentes ventanas temporales con series reales y N-BEATS, TCN y N-HITS.

	# registros	horas	MSE	MAE	MRAE
N-BEATS	12	1	122.5669	10.1645	0.0822
	24	2	121.1597	10.4836	0.1004
	36	3	95.1802	8.9203	0.0845
	48	4	76.7281	7.6896	0.0723
	60	5	65.4508	6.9769	0.0654
	72	6	71.5129	7.3118	0.0674
	84	7	107.6973	8.7441	0.0770
	96	8	129.2838	9.6986	0.0831
TCN	12	1	21.0726	3.8437	0.0313
	24	2	527.6457	17.9045	0.1911
	36	3	428.7735	16.9001	0.1732
	48	4	377.9128	16.4019	0.1638
	60	5	355.6582	16.3111	0.1600
	72	6	324.7853	15.6948	0.1518
	84	7	279.6863	13.8375	0.1331
	96	8	247.994498	12.6238	0.1204
N-HITS	12	1	155.4497	11.3777	0.0924
	24	2	193.6046	13.0341	0.1268
	36	3	163.3128	11.7534	0.1125
	48	4	124.2231	9.3359	0.0891
	60	5	133.2131	9.7672	0.0922
	72	6	160.2911	10.9086	0.1011
	84	7	150.6879	10.4639	0.0953
	96	8	150.5175	10.5298	0.0942

Tabla 2: MSE, MAE y MRAE obtenidos en diferentes ventanas temporales con series sintéticas generadas con DGAN y N-BEATS, TCN y N-HITS.

En la Figura 3 (a) se muestra la evolución temporal de

glucosa del conjunto de train, test y predicciones realizadas por el modelo N-HITS en una ventana temporal de 8 horas y considerando una serie temporal real. Se muestran las predicciones obtenidas en una ventana temporal DE larga duración (8 horas) para mostrar el potencial del método propuesto a largo plazo. A continuación, en la Figura 3 (b) se presenta el *Clarke Error Grid* (CEG) para una ventana temporal de 8 horas, usando de nuevo el modelo N-HITS y una serie temporal real. Se aprecia que todas las diferencias entre los valores reales de glucosa y estimados se encuentran en la zona A, es decir, es clínicamente aceptable. En la Figura 3 (c) se representa la evolución temporal de glucosa del conjunto de train, test y predicciones realizadas por el modelo N-HITS en una ventana temporal de 8 horas y considerando una serie temporal sintética generada con DGAN. De la misma manera, en la Figura 3 (d) se representa el CEG para una ventana temporal de 8 horas, empleando el modelo N-HITS y considerando una serie temporal sintética generada con DGAN. En este caso también se encuentran en la zona A todas las diferencias entre los valores reales de glucosa y los estimados por el modelo N-HITS.

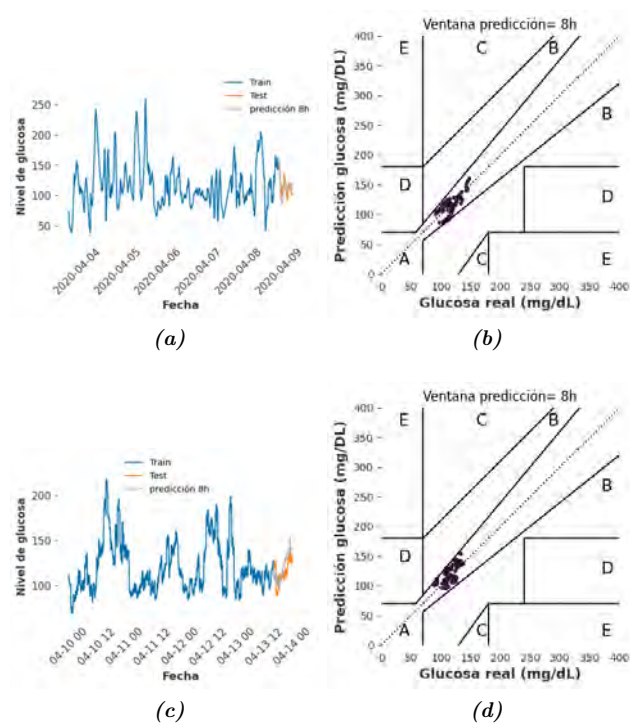


Figura 3: Evolución temporal de valores de glucosa (train, test) y predicciones obtenidos en una ventana temporal de 8 horas por N-HITS usando serie (a) real y (c) sintética generada con DGAN. CEG para predicciones de N-HITS (en 8 horas) usando serie (b) real y (d) sintética generada con DGAN.

4. Discusión y conclusiones

En este trabajo se evaluaron diferentes modelos basados en ANNs para la generación de series temporales sintéticas y para la predicción de valores de glucosa en diferentes ventanas temporales. Tras una evaluación mi-

nuciosa de los modelos de generación de series temporales, CPAR y DGAN, se observó que las series temporales sintéticas creadas con DGAN tienen mejor calidad que CPAR, en términos de similitud estadística (JSD, test de Kolmogorov-Smirnov) y temporal (DTW). Sin embargo, es conveniente señalar que DGAN no logra capturar y «aprender» valores de glucosa bajos en torno 40-60 mg/dl. Referente a modelos de predicción, tras una comparación exhaustiva entre los modelos N-BEATS, TCN y N-HITS en cuanto a su capacidad predictiva, se concluye que N-HITS es más robusto para predecir en ventanas de larga duración (8 horas). Además, alcanza valores más bajos en MSE, MAE y MRAE. El análisis comparativo de modelos de ANN para la predicción de valores de glucosa en diferentes ventanas temporales demuestra la robustez y capacidad predictiva de estos modelos, y pueden llegar a desempeñar un papel esencial en la generación de alarmas en situaciones de hipoglucemia e hiperglucemia. Su adopción e integración permitirían prevenir complicaciones relacionadas con los niveles de glucosa, mejorando el control de glucosa y mejorando sustancialmente la calidad de vida de los pacientes con diabetes.

Referencias

- [1] Aida Budreviciute and et al. Management and prevention strategies for non-communicable diseases (ncds) and their risk factors. *Frontiers in Public Health*, 8:788, 2020.
- [2] Majid Mobasseri et al. Prevalence and incidence of type 1 diabetes in the world: a systematic review and meta-analysis. *Health Promotion Perspectives*, 10(2):98, 2020.
- [3] David Slattery et al. Clinical use of continuous glucose monitoring in adults with type 1 diabetes. *Diabetes Technology & Therapeutics*, 19(S2):S-55, 2017.
- [4] Xin Yi et al. Generative adversarial network in medical imaging: A review. *Medical Image Analysis*, 58:101552, 2019.
- [5] Kevin Zhang et al. Sequential models in the synthetic data vault. *arXiv preprint arXiv:2207.14406*, 2022.
- [6] Zinan Lin, Alankar Jain, Chen Wang, Giulia Fanti, and Vyas Sekar. Generating high-fidelity, synthetic time series datasets with doppelganger. *arXiv preprint arXiv:1909.13403*, 2019.
- [7] Mikel Hernandez and et al. Synthetic data generation for tabular health records: A systematic review. *Neurocomputing*, 493:28–45, 2022.
- [8] Boris N Oreshkin and et al. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. In *International Conference on Learning Representations 2020*, Addis Ababa, Ethiopia, 2020.
- [9] Shaojie Bai et al. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*, 2018.
- [10] Cristian Challu and et al. Nhits: Neural hierarchical interpolation for time series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 6989–6997, 2023.

Computer-aided detection system for pulmonary embolism with integrated cardiac assessment based on embolic burden

I. Luque Del Toro¹, M. Masias¹ and G. Piella¹

¹ Department of Information and Communication Technologies, Pompeu Fabra University, Barcelona, Spain.

iluquedel@gmail.com, mireia.masias@upf.edu, gemma.piella@upf.edu.

Abstract

Pulmonary embolism (PE) is a cardiovascular disease resulting from occlusion(s) in the pulmonary arteries. Its definitive diagnosis relies mainly on imaging, being computerized tomography pulmonary angiogram the gold standard. Recently, there has been increasing interest in automatizing PE detection with the use of computer-aided detection systems, aiming to reduce workloads and enhance identification. Manual semiquantitative scores of embolic burden have also been proposed to assess PE severity and reinforce management. Yet, few attempts have been made to couple both. Here, we propose a deep learning-based system for PE detection, which exploits the visual explanations from the detector network to represent and quantize embolic burden. The resulting measurements of embolic burden are used to assess cardiac function, using a univariate logistic regression model. Particularly, we propose to predict right-to-left ventricle diameter (RV/LV) ratio ≥ 1 , a prognostic cardiac feature strongly associated with both embolic burden and ultimate clinical outcome. The detector network is based on a Squeeze-and-Excitation-ResNet50 and trained on a subset of the RSNA-STR Pulmonary Embolism CT dataset. For the PE detection task, we achieve an accuracy of 0.72, sensitivity of 0.73, and specificity of 0.82 on the test set, which is slightly below the performance of radiologists. As the cardiac assessment directly depends on the detector's performance, we are currently unable to successfully predict RV/LV ratio ≥ 1 . Nevertheless, we believe our system is theoretically feasible and could assist in both PE detection and severity assessment in the future.

1. Introduction

1.1. Pulmonary embolism

Pulmonary embolism (PE) is a treatable condition resulting from occlusion(s) in the pulmonary arteries. It ranks as the third most common type of cardiovascular disease, with an estimated incidence of 60 to 70 cases per 100,000 individuals.

The severity of PE depends on the location and size of the vascular occlusions, which can lead to various clinical outcomes; from natural resolution to adverse cardiac effects. Mortality in PE patients is mainly linked to subsequent cardiovascular events, particularly the failure of the right ventricle (RV). Untreated PE carries a significant mortality risk, of up to 30%, which can decrease to about 8% with timely detection and treatment.

The definitive diagnosis of PE is established through computerized tomography pulmonary angiography (CTPA). Additionally, the utilization of semiquantitative scores, such as the Q-score, has been proposed as a valuable diagnostic tool for the semiquantitation of em-

bolic burden in PE. Embolic burden measurements do not only aid in more accurately describing PE severity and assessing treatment options, but have also demonstrated correlations with short-term prognosis and a wide range of prognostic cardiac features.

Despite all this, detection and management of PE is challenging. Evidence suggests a potential overuse of CTPA and increased workloads in emergency departments, which could result in delayed diagnoses and human errors due to fatigue [1]. It is also suggested that triage may not be optimal and that current procedures are prompt to PE overdiagnosis [2]. Furthermore, embolic burden scores have limited applicability, as their semiquantitative nature makes them difficult to calculate and subject to low reproducibility and high inter-observer variability.

In light of this landscape, there is a growing interest in integrating computational solutions into the clinical workflow. These solutions could potentially reduce workloads, aid in patient characterization, and provide a second reader for direct comparison.

1.2. Computational solutions in the field of PE

The development of computer-aided detection (CAD) systems for PE has rapidly advanced, particularly after the release of RSPECT, the largest available dataset of weakly annotated CTPA studies [3]. Recent efforts [4, 5] have emphasized the importance of creating systems that mimic the thinking process of radiologists, considering the relationships between contiguous CTPA slices. These works used a convolutional neural network (CNN) for feature extraction followed either by a long short-term memory or a temporal CNN. Islam *et al.* [6] also exploited this idea by using consecutive slices, introduced in different image channels, and adding Squeeze-and-Excitation blocks to a CNN.

In spite of the advances, few attempts have been conducted to exploit PE-CAD systems, mainly intended for detection, to obtain and use embolic burden measurements. Both Liu *et al.* [7] and Huang *et al.* [8] proposed the total embolic volume (TEV), subtracted from CTPA with (semi)automatic tools, as an alternative measurement of embolic burden. Liu *et al.* [7] also showed TEVs to be significantly correlated with semiquantitative scores, such as Q-score ($p < 0.001$), as well as with cardiac prognostic features ($p \leq 0.004$), being the correlation with right-to-left ventricle diameter (RV/LV) ratio one of the strongest ($p < 0.0001$). Considering these results, the use of alternative embolic

burden measurements, obtained with computational tools, could be valid to aid in PE characterization and management.

1.3. Hypothesis and goals

We address the challenge of developing a system that integrates both PE detection and severity assessment. We hypothesize that the detection part of this system can be used to represent and quantize embolic burden, taking advantage of explainable tools, such as gradient-weighted class activation mapping (Grad-CAM)++, and following spatial weighing strategies. This differs from previous works ([7, 8]), which measured the embolic burden in terms of TEV, by proposing an embolic burden measurement that is more faithful to the theoretical principles of semiquantitative scores. We also hypothesize that these measurements can be used to predict the cardiac function (i.e., $RV/LV \text{ ratio} \geq 1$)*. A representation of the proposed PE-CAD system with integrated cardiac assessment is shown in *Figure 1*.

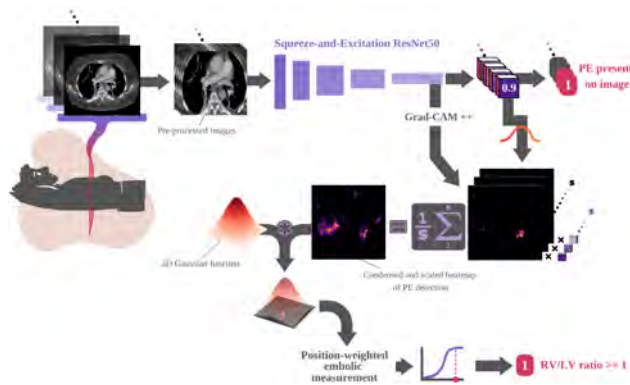


Figure 1. Pipeline of the proposed system.

2. Materials

The data used in this work was obtained from the RSPECT dataset [3]. This dataset comprises 12,195 CTPA studies and contains various annotations, including study-level information (e.g., $RV/LV \text{ ratio} \geq 1$, *Right sided PE*, *Central PE*...) and image-level details (e.g., *PE present on image* (1) or not (0) for each 2D slice in the axial plane). No demographic information was available.

As we focused on acute PE detection, we excluded studies with: (i) indeterminate or negative PE diagnosis, (ii) chronic, or acute and chronic PE, (iii) filling defects other than PE, and (iv) artifactual images. We created a smaller subset of CTPA studies that met the inclusion criteria ($n = 950$), thus ensuring a balanced representation of different clot locations, namely central, left-only, right-only, and double-sided. All steps were implemented using Python 3.6.9, coded in Kaggle Notebooks and Spyder 5.1.5, and ran in our institution's High Performance Computing cluster.

*The choice of RV/LV ratio as a cardiac prognostic variable was motivated by findings of a previously performed replication experiment. The experiment aimed at finding statistically significant relationships between cardiac features and embolic burden (Q-score), which, for the sake of simplicity, was not considered in the present article.

3. Methods

Our pipeline (see *Figure 1*) can be divided into two main components: a PE detector and a cardiac assessment module based on embolic burden. For the PE detector, we destined 920 CTPA studies (230 for each location group) to training and validation. This train-validation set was further balanced, after pre-processing, to account for the imbalance in the *PE present on image* label, thus ensuring an equal count for classes 0 and 1. The remaining 30 CTPA studies were destined to testing. For the cardiac assessment module, we used a total of 40 CTPA studies from the original subset; 28 (7 for each location group) as a training set, and 12 (3 for each location group) for evaluation.

3.1. PE detector

In the first part, the goal was to train and evaluate a CNN for automatic detection of PE using pre-processed CTPA studies. The underlying principle revolved around employing a binary 2D classifier to recognize slices that contained thrombi. We then extracted class activation maps of the trained network to pinpoint crucial areas for classification. In case of positive detection, and if the network was correctly trained, these areas should correspond to the clots' locations.

Images were pre-processed for harmonization as shown in *Figure 2*. This included: windowing, isometrizing, binarizing, subtracting 3D connected components, isolating the lungs, defining bounding boxes, focusing on the thoracic region, and 3-channeling the images.



Figure 2. Pre-processing's most relevant steps.

Several networks were explored for the detection task, including InceptionResNet V2, ResNet50, and DenseNet121. To account for contextual information, we integrated Squeeze-and-Excitation (SE) blocks into the architectures. This enabled the use of multi-channel inputs with sequential content (i.e., adjacent slices), while avoiding excessive computational costs by adhering to a single-step approach (i.e., a CNN). For the 2D classification task, the selected network was a SE-ResNet50, as illustrated in *Figure 3*, with SE blocks at the end of the Bottleneck blocks.

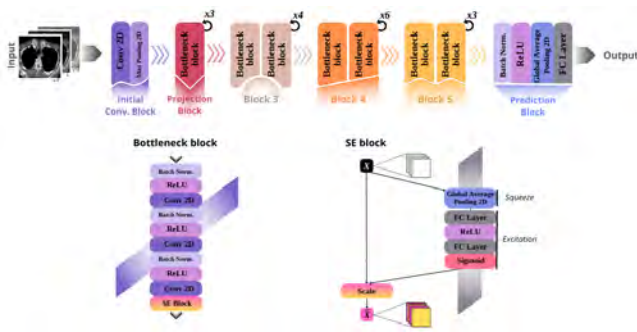


Figure 3. SE-ResNet50 architecture of our PE detector (top). Bottleneck and Squeeze-and-Excitation blocks are also included (bottom).

The SE-ResNet50 was trained from scratch (i.e., random weight initialization) using a batch size of 32. Binary cross-entropy was used as a loss function. The optimizer used was a gradient descent (with momentum), with a learning rate of 0.1, for a total of 10 epochs.

3.2. Cardiac assessment module

In the second part, we developed an independent module to derive a measurement of embolic burden, called position-weighted embolic measurement (PEM), from the PE detector’s visual explanations. We then used PEMs as inputs to predict a cardiac prognostic feature (RV/LV ratio).

First, Grad-CAM++ was used to obtain attention maps (i.e., heatmaps) from the *conv2d37* layer of the detector (i.e., a convolutional layer from Block 4, shortly before the last dimensionality reduction of feature maps). These maps served as visual explanations for the class 1 of the *PE present on image* label.

Individual heatmaps, of dimensions 32 x 32, were resized to 512 x 512 and weighed using the classifier’s prediction scores. Then, all heatmaps from each study were linearly combined, obtaining a single 2D heatmap, which was then normalized with an element-wise division by the number of slices. The weights used for the linear combination belonged to a 3D Gaussian function with mean equal to the center of mass. The 3D Gaussian function could be deconstructed as: (i) a 1D Gaussian function in the craniocaudal axis with standard deviation equal to 60, and (ii) a 2D Gaussian function in the axial plane with standard deviation equal to 15.

This was intended to further emphasize the theoretical basis of some semiquantitative scores, such as Q-score, which are known to give more importance to proximal occlusions in the final computation. This procedure resulted into a single PEM for each study.

The prediction of *RV/LV ratio* ≥ 1 from PEMs was assumed to be linearly separable. A univariate logistic regression model was considered. The model generated a probability ($p(x) = \frac{1}{1+e^{-(b_0+b_1 \cdot x)}}$) for each PEM (x), being b_0, b_1 the regression coefficients.

4. Results and Discussion

The proposed system achieves a precision of 0.8 and a recall of 0.92, with an accuracy of 0.87. When retrieving the network’s weights and evaluating on the test set, we get a precision of 0.5, a recall of 0.91, and an accuracy of 0.72. The false positive (FP) rate is 46, 34, and 31 FP/study for probability threshold values (T) of 0.5, 0.7, and 0.75, respectively. Noteworthy, the validation set, as well as the training set, included a balanced group of images for both the absence and the presence of PE. This was not ensured in the test set, as we tried to simulate a realistic input (i.e., full-size). Figure 4 shows the sensitivity vs. specificity plot of the test set for the *PE present on image* label at different probability thresholds.

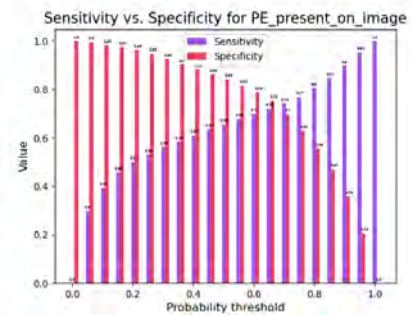


Figure 4. Sensitivity (purple) vs. specificity (red) of PE detection at different prediction threshold values.

As shown in Figure 4, when the threshold value is 0.7, the model can reach a sensitivity and specificity for PE detection of 0.73 and 0.82, respectively, on the test set. Sensitivity and specificity of radiologists performing the PE detection task range from 0.67 to 0.87 and 0.89 to 0.99, respectively. Therefore, while sensitivity is within radiologists’ range, our model’s specificity is slightly below the lower limit. On the other hand, when comparing the model performance with other state-of-the-art approaches, we find that it falls behind in terms of sensitivity and specificity, as some past approaches have been able to reach values above 0.8 for both [4, 5, 9].

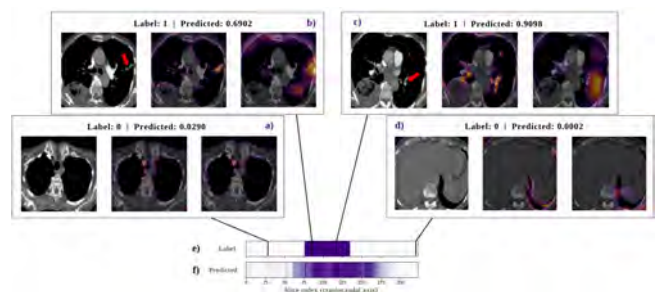


Figure 5. Examples of visual explanations. **a, d:** True negatives. **b, c:** True positives. For each sample, the pre-processed image (left) and overlaid attention maps from *conv2d37* (middle) and last convolutional (right) layers are collected. The arrows point to PE. Real (e) and predicted (f) labels are also indicated (white = 0, purple = 1).

Figure 5 shows examples of visual explanations for the predictions of PE presence. Our proposed detector effectively identifies PE, even if the data is weakly annotated and the model consists of a 2D classifier. Explanations from an intermediate layer (i.e., *conv2d37* layer), while well delineated, tend to highlight PE-alike regions, whereas explanations from the last convolutional layer, while more discriminative, tend to be spatially inaccurate.

Highly dependent on results of the previous part, our univariate logistic regression model for RV/LV ratio ≥ 1 label prediction did not perform better than random guess. The elapsed time to obtain a cardiac assessment for a single patient (i.e., including the *PE present on image* label prediction, heatmaps' obtention, PEM computation, and RV/LV ratio ≥ 1 label prediction) was 750.4 ± 172.2 sec, thus, approximately 12.5 min (computation based on a sample of 100 randomly selected pre-processed CTPAs).

The results of the cardiac assessment module indicate that our complete system would be unable to reliably predict RV dysfunction (in terms of RV/LV ratio) in patients undergoing CTPA. We believe such results are due to a) the high FP rate and b) the use of poorly post-processed visual explanations from an intermediate convolutional layer (*conv2d37*). Yet, we still consider the PEM to be a reasonable alternative to Q-score and TEV, as it exploits virtues from both, such as the emphasis on clot location, and the patient specificity due to derivation from CTPA, respectively.

There are several limitations in the present study. Concerning the PE detector, class balancing (according to the *PE present on image* label) is a major limitation, and a likely explanation for the high FP rates. In light of our results, we believe that future work should contemplate using full-sized imbalanced CTPAs for training and validation. Consequently, monitoring metrics such as recall and including a focal loss could help address the label disparity during training.

As for the cardiac assessment module, we consider layer selection a critical limitation, as visual explanations struggled to simultaneously achieve both high object discrimination and high spatial resolution. Future work should tackle the post-processing of visual explanations from intermediate convolutional layers (e.g., by conditioning heatmaps with automatic vessel segmentations). Further iterations should also focus on parameter tuning (e.g., the arbitrary standard deviation from Gaussian functions) and on comparing different embolic burden measurements.

Noteworthy, a central limitation of this phase remains the detector's performance, as it significantly influences the entire process. Therefore, improving the performance of the PE detector should be a priority before delving into the refinement of PEM computation and cardiac assessment.

5. Conclusions

We have presented a proof-of-concept system to address multiple challenges in the field of PE. In particular, our contribution can be summarized as follows:

1. Our proposed system can detect PE in CTPA, measure the embolic burden, and produce a cardiac assessment. Hence, PE detection and severity assessment are conducted in a single step.
2. Instead of using a two-step implementation for PE detection, our detection network incorporates sequence information thanks to multi-channel inputs and exploits it through SE blocks. Therefore, the resulting model is simpler and requires less time to train.
3. We produce visual explanations for the PE detection task thanks to Grad-CAM++.
4. Our cardiac assessment module is based on a new embolic burden measurement, called PEM, which highly emphasizes the role of clot location.

Although our system is still under improvement and further clinical validation is required, we believe that it can serve as a stepping stone to develop tools that can aid in PE characterization and management.

References

- [1] M. Perera *et al.* Underuse of risk assessment and overuse of computed tomography pulmonary angiography in patients with suspected pulmonary thromboembolism. *Internal medicine journal*, 47(10):1154–1160, 2017.
- [2] D. Corrigan *et al.* Pulmonary embolism: the diagnosis, risk-stratification, treatment and disposition of emergency department patients. *Clinical and experimental emergency medicine*, 3(3):117–125, 2016.
- [3] E. Colak *et al.* The RSNA Pulmonary Embolism CT Dataset. *Radiology*, 3, 2021.
- [4] X. Ma *et al.* A multitask deep learning approach for pulmonary embolism detection and identification. *Scientific Reports*, 12(13087), 2022.
- [5] H. Huhtanen *et al.* Automated detection of pulmonary embolism from CT-angiograms using deep learning. *BMC Med Imaging*, 22(43), 2022.
- [6] NU. Islam *et al.* Seeking an Optimal Approach for Computer-Aided Pulmonary Embolism Detection. *Machine Learning in Medical Imaging*, 12966:692–702, 2021.
- [7] W. Liu *et al.* Evaluation of acute pulmonary embolism and clot burden on CTPA with deep learning. *European radiology*, 30(6):3567–3575, 2020.
- [8] WM. Huang *et al.* Quantitative volumetric computed tomography embolic analysis, the Qanadli score, biomarkers, and clinical prognosis in patients with acute pulmonary embolism. *Sci Rep* 12, 7620, 2022.
- [9] M. Khan *et al.* IoMT-Enabled Computer-Aided Diagnosis of Pulmonary Embolism from Computed Tomography Scans Using Deep Learning. *Sensors*, 23(3), 2023.

Study of fibroblasts activation kinetics and identification of fibroblasts subpopulations in physiological and pathological conditions

L. Carreras-Vidal¹, Á. Martínez-Blanco¹, S. Noé¹, J. Alcaraz^{1,2,3}, R. Farré^{1,4,5}, N. Gavara^{1,2}

¹ Unit of Biophysics and Bioengineering, Universitat de Barcelona, Barcelona, Spain, lcarrevi29@alumnes.ub.edu, africa.martinez@ub.edu, sergio.noe@ub.edu, jalcaraz@ub.edu, rfarre@ub.edu, ngavara@ub.edu

² Institute for Bioengineering of Catalonia (IBEC), Barcelona Institute of Science and Technology, Barcelona, Spain

³ Thoracic Oncology Unit, Hospital Clínic Barcelona, Barcelona, Spain

⁴ CIBER de Enfermedades Respiratorias, Madrid Spain

⁵ Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain

Abstract

Fibroblasts undergo significant morphological and functional changes in response to specific environmental cues, adopting a novel phenotype when sustaining transformative activation in wound healing and cancerous processes. The aim of this project is to characterize cytoskeletal reorganization in both control fibroblasts' (CFs) and cancer-associated fibroblasts' (CAFs) activation. To achieve this, we propose a novel method based on the extraction of biophysical biomarkers from epifluorescence images of the cytoskeleton of individual fibroblasts, obtained from a patient with lung cancer and activated using TGF- β . These biophysical outputs were also used to identify CAFs subpopulations. While non-tumoral fibroblasts experience larger morphological changes characterized by an increase in area and the acquisition of a robust network of actin fibers, CAFs exhibited sustained larger areas throughout the process regardless of TGF- β administration, amongst other cytoskeletal transformations. The application of logistic regression has allowed for a classification between CFs and CAFs of 81% accuracy, highlighting the differences in the cytoskeleton of these cell types. Furthermore, the intragroup analysis provided by unsupervised clustering has enabled the identification of five clusters for non-activated CAFs, which converge at 72 hours post-activation into two clusters.

1. Introduction

Fibroblasts are diverse mesenchymal cells that, amongst other diverse functions, participate in tissue homeostasis and disease by producing complex extracellular matrix and creating signaling niches through biophysical and biochemical cues [1].

In wound healing processes, the early recruitment and activity of platelets and inflammatory cells promotes the migration of fibroblasts to the wound site. There, subsets of fibroblasts differentiate into myofibroblasts and deposit new extracellular matrix to fill the wound. [2] As a highly versatile cell, fibroblasts activation is not only functional but also morphological as quiescent fibroblasts lose their spindle-like shape and become more spread [3].

When the injury is perpetual, such as is the case in cancer, the status of fibroblasts activation is sustained due to chronic tissue repair response [3], and we talk about cancer-associated fibroblasts (or CAFs), which present a

similar phenotype to myofibroblasts [4]. However, we cannot talk about a unique type of CAFs as they have been defined to be extremely heterogeneous [5] and can even be divided into subpopulations [6].

While the role of fibroblasts in cancer is still being elucidated as both cancer-promoting and cancer-restraining, CAFs are considered to contribute to the structural, metabolic, and immune reprogramming of the tumoral microenvironment [3]. Therefore, understanding the morphological changes and characteristics of fibroblasts in cancer, as opposed to in physiological wound healing, is crucial in comprehending the mechanisms involved in disease's progression.

Over the past decade, significant advancements have been made in accurately and quantitatively characterizing the shape and morphology of cells and their cytoskeleton thanks to the development of image processing pipelines that can produce high-throughput, single-cell, and multiparametric outputs depicting the state of the cytoskeleton [7]. As a result, image-based quantification of cytoskeletal state has become a valuable approach for describing the biophysical condition of cells.

Current studies on CAF activation have not found unique biochemical biomarkers to identify fibroblasts [8]. Furthermore, no consensus has been reached on a univocal classification and definition of fibroblastic subpopulations [4]. Most of these studies on fibroblasts subpopulations, based on single-cell RNA sequencing and flux cytometry biomarkers, prove insufficient on subpopulations recognition based on functionality [6]. That is why instead of using conventional biomarkers, we propose biomedical engineering tools to explore changes in the cytoskeleton reorganization via novel biophysical biomarkers from immunofluorescence imaging.

2. Materials and methods

Our study is based on the extraction of biophysical biomarkers from epifluorescence images of the cytoskeleton of individual fibroblasts to characterize fibroblasts morphology during both pathological and physiological activation, as detailed in Figure 1.

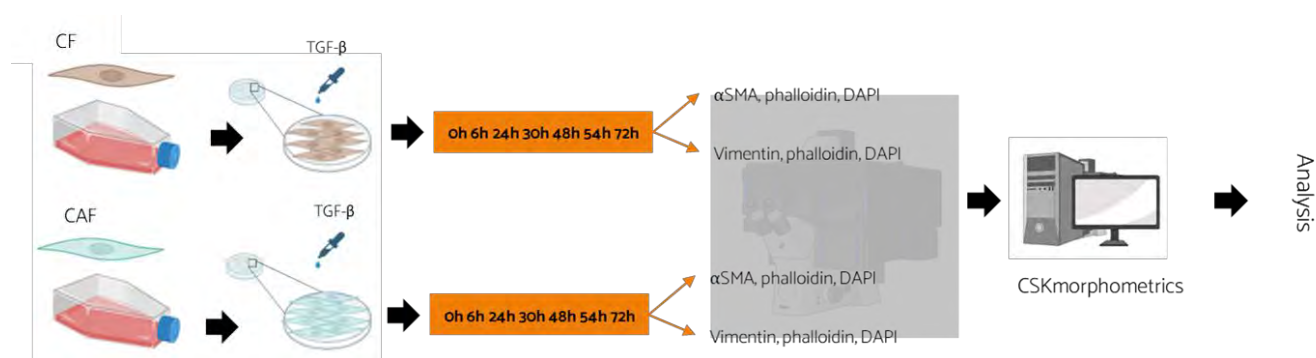


Figure 1. Tumoral (CAFs) and non-tumoral (CFs) fibroblasts from a single patient were acquired and cultured over glass coverslips. Fibroblasts were synthetically activated using TGF- β , and at different timepoints (6h, 24h, 30h, 48h, 54h, 72h) cells were fixed and subjected to immunostaining assays: one with α SMA to assess fibroblasts activation and the other with vimentin and phalloidin to visualize the cytoskeleton. CSKmorphometrics was used to extract features related to cell morphology, fiber organization, and the nucleus from epifluorescence microscopy images.

2.1. Cell acquisition and culture

Tumoral and non-tumoral fibroblasts were previously obtained from surgical lung adenocarcinoma patient #13 with the informed patient consent, using protocols approved by the Ethics Committee of the Hospital Clínic of Barcelona [9]. Cryotubes at passage #9 and #10, respectively, were acquired.

2.2. Experimental set up

40-mm diameter petri dishes were prepared, each holding three 12-mm glass coverslips. Of these, 22 dishes were designated for tumoral conditions: 11 for vimentin antibody, phalloidin, and DAPI staining to observe cell morphology, and 11 for α SMA antibody, phalloidin, and DAPI staining to gather fibroblast activation data. Another set of 22 dishes was readied for non-tumoral conditions. Each group includes the selected seven timepoints (0h, 6h, 24h, 30h, 48h, 54h, and 72h activation) and four additional controls.

2.3. Sterilization and Conditioning of Petri Dishes

Coverslips were sterilized with 70% ethanol for approximately 20 minutes, washed with PBS 1x and conditioned with cell media (DMEM, 10% FBS, 1% P/S).

2.4. Cell Seeding and Activation

A range of 1500-3000 cells per cm^2 were seeded on each of the petri dishes. After 24 hours, petri dishes were conditioned with activation media (DMEM, 1% ITS, 1% P/S). 2.5ng/ml of TGF- β were added per petri well to activate the cells. At the predetermined time points, cell activation was stopped by fixing the cells with paraformaldehyde (PFA) 4%.

2.5. Immune staining protocol

A 0.2% Triton X-100 solution was used to permeabilize the cell membranes for antibodies. Subsequently, a blocking buffer (10% FBS in PBS 1x) was applied to block unwanted antibody binding sites and reduce background.

Primary antibodies were prepared at concentrations of 1:1000 for anti-vimentin mouse monoclonal antibody and 1:500 for anti- α SMA rabbit monoclonal antibody. The appropriate secondary antibodies were selected based on the species of the primary antibodies. Both secondary

antibodies were used at a concentration of 1:500. After the secondary antibody incubation, conjugated phalloidin antibody (1:1000) was added to solution of PBS 1x and 1% bovine serum albumin (BSA). NucBlue™ Live ReadyProbes™ Reagent (1:1000) was used to stain the nucleus.

2.6. Microscopic analysis

Leica SP5 inverted microscope equipped with a CCD camera (C9100) with a Nikon CFI Plan Fluor 20XC MI objective with oil immersions was used for epifluorescence imaging. Composite images of TXRED, FITC, and UV-2A channels were acquired. Fifty pictures were obtained for each of the conditions. No cells were discarded unless they were clearly damaged due to external factors or too agglutinated to be analysed by the software.

2.7. Image analysis

CSKmorphometrics [10] is the custom-built pipeline used to extract features of the cells' cytoskeleton from immunofluorescence images. The features that are extracted include gross morphology, characteristics of the individual fibres, information about the organization of the fibres, and characteristics of the nucleus.

2.8. Data analysis

Throughout the manuscript, all values are reported as mean \pm SD. One-way ANOVA tests were used to assess the difference of the mean of 21 features obtained from actin fibers before activation and 24-, 48-, 72-hours after activation for control fibroblasts (CFs) and CAFs. Logistic Regression was employed to preliminarily assess transformative activation as a combination of the different obtained features. To assess fibroblast subpopulations, K-means clustering has been used, with the addition of R NbClust package to determine the optimal number of clusters.

3. Results

3.1. Activation of fibroblasts

To assess the effect of the synthetic activation of the fibroblasts using TGF- β , we have quantified the fluorescence of α SMA, a known biomarker of fibroblastic

activation, in both non-tumoral and tumoral fibroblasts before and after 72 hours of persistent activation stimulus.

Figure 2 reveals a significant disparity in α SMA between the pre-activated and activated cells, while no significant differences can be observed between CFs and CAFs, showing synthetic activation has been successful for both cell types.

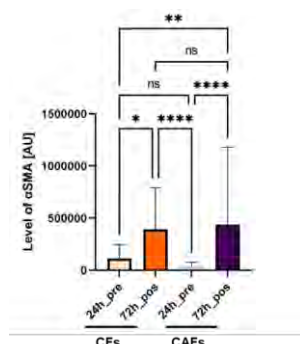


Figure 2. Level of α SMA before and after 72 hours of persistent activation for both CFs and CAFs.

3.2. Transformative activation on CFs

CFs exhibit larger morphological changes within the first 24 hours of the process of activation. Amongst the 21 extracted features, we highlight the significant increase in area ($p < 0.05$) between the preactivated stage ($2438 \pm 1876 \mu\text{m}^2$) and the activated stages ($9181 \pm 7661 \mu\text{m}^2$), that reaches already a plateau at 24 hours (Figure 3A). Simultaneously, there is a significant raise in F-actin fibers' fluorescence (Figure 4B) and an elongation of these fibers (Figure 4C, $21 \pm 9 \mu\text{m}$ in the preactivated stage and $34 \pm 14 \mu\text{m}$ at 72 hours post-activation stage). This is consistent with the transformation of fibroblasts to myofibroblasts, which acquire robust network of actin fibers.

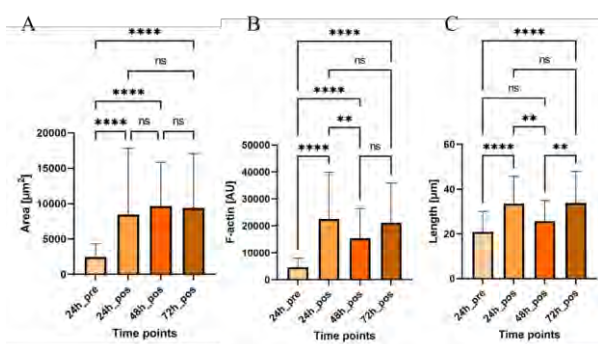


Figure 3. One-way ANOVA analysis of CFs' cytoskeleton extracted parameters at various activation stages. (A) shows the cell area, (B) shows the total fluorescence associated with the F-actin fibers inside the cells, and (C) shows their length.

3.3. Transformative activation on CAFs

CAF areas were sustained (Figure 4A) throughout the process of activation ($6067 \pm 5068 \mu\text{m}^2$). Out of the 21 extracted parameters, some of the most notable alterations in the cytoskeletal fibers were seen in terms of fiber length (Figure 4B) and preferent location of the fibers (Figure 4C), becoming gradually more concentrated due to activation (variance of fiber location of 0.165 ± 0.060

before activation and 0.138 ± 0.042 , $p < 0.05$), especially in the periphery of the cell. This may indicate an increased migratory capacity of activated CAFs.

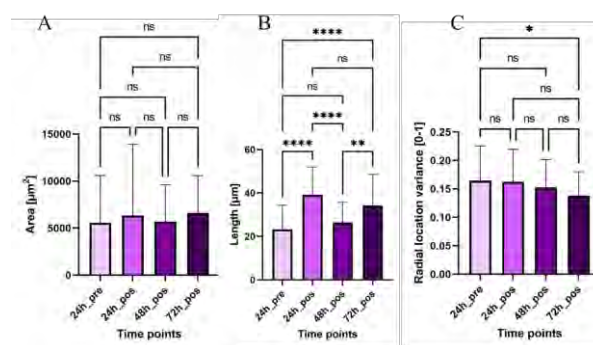


Figure 4. One-way ANOVA analysis of CAFs' cytoskeleton extracted parameters at various activation stages. (A) shows the cell area, (B) shows the length of the actin fibers, and (C) shows the preferent radial location of the actin fibers.

3.4. Application of machine learning

Logistic Regression provided a maximum classification score when distinguishing between non-activated CFs and CAFs (Table 1), meaning that -preliminarily- these two groups are different enough to be correctly classified.

Multiple-iteration average accuracy score	
Prediction of CFs activation stage	64%
Prediction of CAFs activation stage	58%
Prediction between non-activated CFs or CAFs	81%
Prediction between activated (72h) CFs or CAFs	68%

Table 1. Average accuracy score obtained from different train-test partition of Logistic Regression algorithm.

3.5. Clustering to identify CAFs subpopulations

K-Means clustering has given us insight on how CAF subpopulations evolve during activation. Figure 5A illustrates that non-activated CAFs have been optimally clustered into five different subgroups. Upon investigating the area of the different clusters, we see significant differences between the different groups. It is noteworthy to mention that the majority of cells are classified into clusters with smaller area (45 cells) while only two clusters present a larger area (22 cells).

After a 72-hour persistent activation, the majority of CAFs ($n=46$) fall into a subgroup of activated morphology ($7467 \pm 3624 \mu\text{m}^2$) and only a few cells ($n=8$) exhibited a non-activated morphology ($1741 \pm 599 \mu\text{m}^2$), as shown in Figure 5B.

4. Conclusions

This study has successfully implemented a single-cell approach using biophysical biomarkers from epifluorescence images to study the activation of CFs and CAFs. Cytoskeletal reorganization occurs during both physiological and pathological activation.

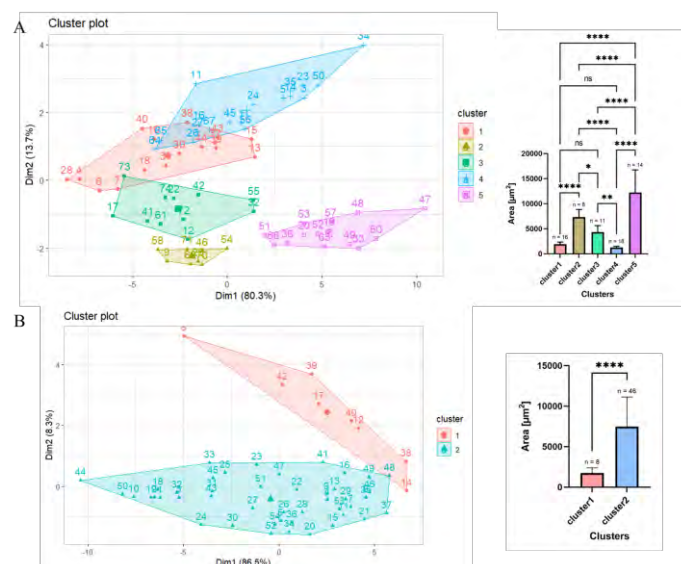


Figure 5. K-Means clustering for (A) 24h pre-activation tumoral fibroblasts and (B) 72h post-activation tumoral fibroblasts. The statistical differences on the area have also been displayed for (A) and (B).

While CFs transformed into myofibroblasts, forming a robust network of actin fibers within the first 24 hours of activation, CAFs already exhibited *myofibroblastic* morphologies prior to activation.

The extraction of numerous features from single cells allows for the creation of large datasets that can be analyzed using artificial intelligence tools. This approach holds the potential to identify CAFs in biopsy samples through the application of machine learning tools.

Additionally, the unsupervised clustering results indicate that the multiple subgroups observed before activation of CAFs converge into a single activated state. Notably, some of the pre-activated CAF subgroups already exhibited an activated morphology, suggesting that culturing on stiff glass coverslips artificially activates these fibroblasts. Therefore, the use of biomaterial scaffolds to better mimic the real cell environment, both in tumoral and non-tumoral conditions is necessary.

In future directions for this research, we aim to explore intermediate time points, particularly those occurring before the 24-hour milestone to further evaluate the transformative process of CF activation. Moreover, we intend to apply the proposed approach to study the process of CAF deactivation, a crucial aspect of normal wound healing that is conspicuously absent in cancer. This approach may also contribute to the screening and identification of drugs capable of deactivating CAFs, in the efforts of reinstating a healthy tumoral microenvironment.

Acknowledgements

This project has been carried out thanks to the resources and facilities available at the Unit of Biophysics of the Faculty of Medicine of University of Barcelona within the project PID2020-116808RB-I00 funded by the Spanish Ministry of Sciences, Innovation and Universities.

References

- [1] Plikus MV, Wang X, Sinha S, et al. Fibroblasts: Origins, definitions, and functions in health and disease. *Cell*, vol 184, sup 15, 2021, pp 3852-3872. (ISSN: 1097-4172)
- [2] Talbott HE, Mascharak S, Griffin M, Wan DC, Longaker MT. Wound healing, fibroblast heterogeneity, and fibrosis. *Cell Stem Cell*, vol 29, sup 8, 2022, pp 1161-1180 (ISSN: 1875-9777)
- [3] Kalluri, R. The biology and function of fibroblasts in cancer. *Nat Rev Cancer*, vol 16, 2016, pp 582–598 (ISSN: 1474-1768)
- [4] Xing F, Saidou J, Watabe K. Cancer associated fibroblasts (CAFs) in tumor microenvironment. *Front Biosci*, vol 15, sup 1, 2010, pp 166-179 (ISSN: 2768-6698)
- [5] Caligiuri G, Tuveson DA. Activated fibroblasts in cancer: Perspectives and challenges. *Cancer Cell*, vol 41, sup 1, 2023, pp 434-449 (ISSN: 1878-3686)
- [6] Lavie D, Ben-Shmuel A, Erez N, Scherz-Shouval R. Cancer-associated fibroblasts in the single-cell era. *Nat Cancer*, vol 3, sup 7, 2022, pp 793-807 (ISSN: 2662-1347)
- [7] Gavara N, Chadwick RS. Relationship between cell stiffness and stress fibre amount, assessed by simultaneous atomic force microscopy and live-cell fluorescence imaging. *Biomech Model Mechanobiol*, vol 15, sup 3, 2016, pp 511-523 (ISSN: 1617-7940)
- [8] Irvine AF, Waise S, Green EW, Stuart B, Thomas GJ. Characterising cancer-associated fibroblast heterogeneity in non-small cell lung cancer: a systematic review and meta-analysis. *Sci Rep*, vol 11, sup 1, 2021, pp 3727 (ISSN: 2045-2322)
- [9] Puig M, Lugo R, Gabasa M, et al. Matrix Stiffening and $\beta 1$ Integrin Drive Subtype-Specific Fibroblast Accumulation in Lung Cancer. *Mol Cancer Res*, vol 13, sup 1, 2015, pp 161-173 (ISSN: 1557-3125)
- [10] Zhang X, Flores LR, Keeling MC, Sliogeryte K, Gavara N. Ezrin phosphorylation at T567 modulates cell migration, mechanical properties, and cytoskeletal organization. *Int J Mol Sci*, vol 21, sup 2, 2020 (ISSN: 1422-0067)

Development of an electric stimulation system to analyze muscle damage in 3D Duchenne muscular dystrophy muscle culture

M. Ruiz Gutiérrez^{1,2}, A. Tejedera Villafranca², B. Giraldo Giraldo^{1,2,3}, J. Ramón Azcón^{2,4}, JM. Fernández Costa²,

¹ Universitat Politècnica de Catalunya (UPC), Dept. ESII – EEBE, Barcelona, España, martin.ruiz.gutierrez@estudiantat.upc.edu, Beatriz.Giraldo@upc.edu

² Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology (BIST), Barcelona, Spain, atejedera@ibecbarcelona.eu, jfernandez@ibecbarcelona.eu, jramon@ibecbarcelona.eu

³ CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

⁴ Institució catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Abstract

Duchenne muscular dystrophy (DMD) is a progressive muscle wasting disease that impairs the ability of myotubes to handle mechanical stress. Thus, muscle contraction leads to membrane damage in the form of sarcolemmal tears. Electrical pulse stimulation (EPS) can replace motor neuron activation in muscle tissues, allowing in vitro muscle contraction-related studies. This work aimed to provide a practical solution for the stimulation of 3D skeletal muscle tissues as an alternative to overcome the limitations of existing commercial systems. To this end, we developed an intuitive electric stimulation system tailored to the needs of the laboratory. Our system was divided into two parts designed from the ground up. The first part allowed the user to choose which samples to stimulate, while the second part delivered the stimulation signal via graphite electrodes. The system developed in this study was used to induce contraction in healthy and DMD affected samples under different stimulation regimes, and ELISA assays were used to quantify creatine kinase leakage as a muscle damage marker. Our system facilitated higher throughput experiments using multiple samples, proving its potential to serve as a valuable tool for future research in drug testing and DMD treatments.

1. Introduction

The musculoskeletal system is one of the major tissue/organ systems of the body. Skeletal muscle is the most abundant of the three muscle types and the one responsible for body movement. It is made up of thousands of fibres wrapped by sheaths made of connective tissue. The membrane which surrounds skeletal muscle fibre cells is known as the sarcolemma [1]. Muscle can be affected by several pathologies, with muscular dystrophies being one of the most common.

Duchenne muscular dystrophy (DMD) is an X-linked recessive disease caused by mutations in the *DMD* gene, responsible of the production of dystrophin [2]. This protein is the main component of the dystrophin-associated protein complex (DAPC). It ensures the integrity of muscle by establishing a link between the cell's cytoskeleton, the sarcolemma, and the extracellular matrix. The inability to synthesize dystrophin leads to irreparable muscle damage, caused by sarcolemmal tears. Patients lose muscular function as the disease progresses, leading to the need of wheelchairs and mechanical

ventilation. There is currently no cure for DMD and its treatments focus primarily on palliative care [3].

To study DMD, *in vitro* models can be used to explore tissue morphology, disease mechanisms, drug action and protein production. This model require cell culture, which can be done in 2D, also known as monolayer, or in 3D. 2D models cannot mimic the complexity of real tissues [4]. In recent years, the appearance of microphysiological systems known as Organs-on-a-Chip has allowed models to emulate *in vivo* tissue functionality with higher fidelity. These platforms can be used with 3D muscle cultures in contraction-related studies [5].

Muscle contractibility can be induced *in vitro* using electrical pulse stimulation (EPS), which effectively replaces neuron motor activation. EPS is conducted using a stimulation system which applies an electric field to the cultured tissues via electrodes [6]. There are some commercially available systems, but they do not always meet all research needs. Therefore, it is common for systems adapted to these needs to be developed.

In this work, we developed a novel electrical stimulation system for the functional study of 3D skeletal muscle tissues. Our system was designed to work with 24-well culture plates, implementing different stimulation modes. This improved on the throughput and versatility of the previous systems used in the laboratory. The system was validated studying the induction of sarcolemmal damage in DMD affected samples. After applying different stimulation regimes, Creatine Kinase (CK) was measured as a marker of muscle damage. Furthermore, myotube morphology and integrity was analysed after stimulation. Altogether, our system enabled more efficient experimentation with multiple samples.

2. Materials and Methods

2.1. Selector

The electric schematic of the selector was done using KiCad 6.0. It was divided in two main parts: an array of ten relays, with flyback diodes and controlled by the activation of optocouplers, and a dashboard consisting of two 7-segment displays and 24 white LEDs, each

representing a well of the culture plate. The device was powered at 12V. The stimulation signal was received through a BNC connector, while the branched signal was sent through a 10-way, 2 rows 5 columns IDC male connector. The selector was controlled by an Arduino Mega 2560 Rev3, with four pushbuttons to operate the system. The Arduino was programmed using Arduino IDE, and three stimulation modes were designed. The PCB was designed using KiCad 6.0, and the fabrication files were sent to Circuitos Impresos 2CI, S.L. to make the board. The components were manually welded.

The selector and the Arduino board were placed inside a 3D printed box. The parts of the box were designed using SOLIDWORKS and they were 3D printed using a Creality Ender 3 S1 Pro and Creality Ender PLA filament for the main body and TPU and PETG for its accessories.

2.2. Stimulator

The electric schematic of the stimulator and its PCB design were done using KiCad 6.0. A plated through-hole footprint was designed to hold the graphite electrodes in place with a separation of 10mm between their inner surfaces. The board was fabricated by Circuitos Impresos 2CI, S.L. The electrodes were cut manually from ultra high purity graphite plates of 100x100x2mm (Micro to nano) into smaller pieces of 25x5x2mm. They were glued to the board using Pattex Nural 20, due to the non-adherence of tin to graphite.

The support of the stimulator was designed and printed using SOLIDWORKS and a Creality Ender 3 S1 Pro. The filament used was PLA.

2.3. Electrical pulse stimulation experiments

Immortalized human muscle precursor cells from an unaffected control and a DMD patient were used for the EPS experiments. They were 2D cultured in 24-well culture plates coated with Matrigel. The myoblasts were left to differentiate into myotubes for seven days using Differentiation Medium and then were stimulated under different regimes of varying frequency and total stimulation time. The system was stimulated with a monophasic square wave of 10V_{pp} and 1ms wide pulses. The stimulation regimes were applied by rows, having three biological replicas of control and DMD muscle for each. After the EPS, the samples were left to rest for one hour, and then the supernatant was retrieved.

The stimulator was placed inside an XL S1 cell incubator maintaining physiological conditions (37 °C and 5% CO₂) during the EPS. Brightfield images and recordings of the contracting myotubes were taken inside using a Zeiss Axio Observer Z1/7.

2.4. CK ELISA

Creatine Kinase Enzyme-Linked ImmunoSorbent Assays were used to quantify the concentration of this enzyme in the supernatant as a measure of the damage caused by contraction to the muscle cultures. The assay was done using 96-well plates and substrates prepared following an established sandwich ELISA protocol [7]. Three

technical replicas of each sample were made. The supernatant of non-stimulated tissues was used as negative control, and cell lysates prepared using RIPA buffer were used as positive control. Some wells were filled with a known concentration of CK to make a calibration curve. The resulting data was processed using GraphPad Prism. The calibration curve was made using a sigmoidal 4 parametric logistic curve fitting, and it was used to interpolate the enzyme concentration in each well.

2.5. Immunostaining

An immunostaining assay was done on 6 samples after the EPS experiment. The antibodies used were sarcomeric α -actinin (anti-SAA), to mark the sarcomere; phalloidin, to mark the cytoskeleton (F-Actin filaments); and DAPI, to mark the nuclei. Fluorescence microscopy images were taken using Zeiss Axio Observer Z1/7.

3. Results and discussion

3.1. Experimental setup

A general diagram of the stimulation system when integrated in the used electrical pulse stimulation experimental setup is shown in figure 1. The stimulation signal is provided by a wave generator and amplified. Then, the selector branches this signal and sends it to the stimulation plate, making the selected tissues contract. An oscilloscope is used to check on the signal arriving to the tissues and a microscope to record their contractions. A computer establishes the stimulation regime by controlling the wave generator and stores the data recorded from the microscope.

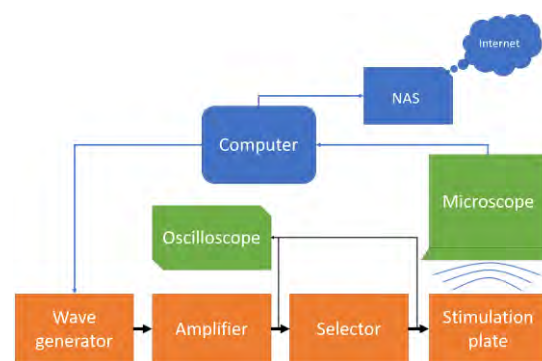


Figure 1. Block diagram of the experimental setup for electrical pulse stimulation experiments

The interaction between the selector and the stimulator is depicted in figure 2. The stimulation signal which arrives to the selector is branched and sent to the stimulator through a 10 way cable after passing through 10 relays. 4 ways send signal, while the other 6 are connected to the ground reference. This allows to establish a matrix organisation on the commercial 24-well plate, which consists of 4 rows and 6 columns. By connecting each row to one way of stimulation signal and each column to a ground reference, the operation of the relays permits the delivery of current to a single, a selection, or all wells at once.

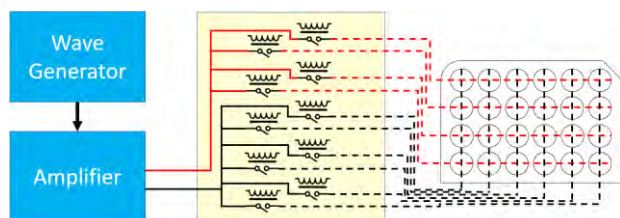


Figure 2. Connection of the two developed parts of the stimulation system

3.2. Selector

The Arduino was programmed to allow three stimulation modes. In all of them, the dashboard provided information to the user to better operate the device. The first mode, single mode, was used to stimulate a single sample at a time. The 7-segment displays showed the currently selected well according to the row letter and column number of the plate. The LED of the selected well was turned on once the relays were activated. The second mode, selection mode, allowed the stimulation of several wells at once. In this case, the user could choose between rows and columns, and add them to the selection. When a well would be stimulated due to the selection of its row and column at once, its representing LED turned on. Finally, the all samples stimulation mode activated all relays at once. In all cases, the displays blinked when the relays were activated, and the selection of different wells was locked until the active relays had been disconnected.

The box of the selector (figure 3) ensured electrical safety by enclosing all applied parts. The dashboard was covered by transparent 3D printed windows.

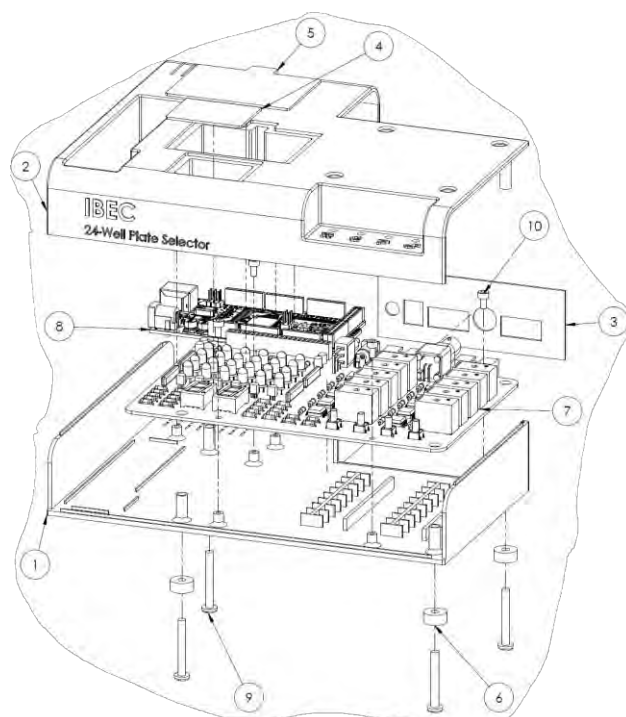


Figure 3. Exploded view of the selector. 1) Box base, 2) Box lid, 3) Back lid, 4) Displays window, 5) LEDs window, 6) Pads, 7) Selector PCB, 8) Arduino Mega 2560 Rev3, 9) M4x30 screws, 10) M3x6 screws

3.3. Stimulator

The PCB of the stimulator was designed so the electrodes were aligned with the wells of the commercial culture plate. It had holes between the electrode holders to allow microscopy imaging while conducting experiments.

Figure 4 shows the 3D designed and printed support used to seat the stimulator on top of the commercial culture plate (Thermofisher Nunc), trying to enclose and keep the cells under sterile conditions. The upper part has a notch to place the original lid, and the lower has a tongue to enclose the device. The pins of the connector were covered to restrict access to applied parts.

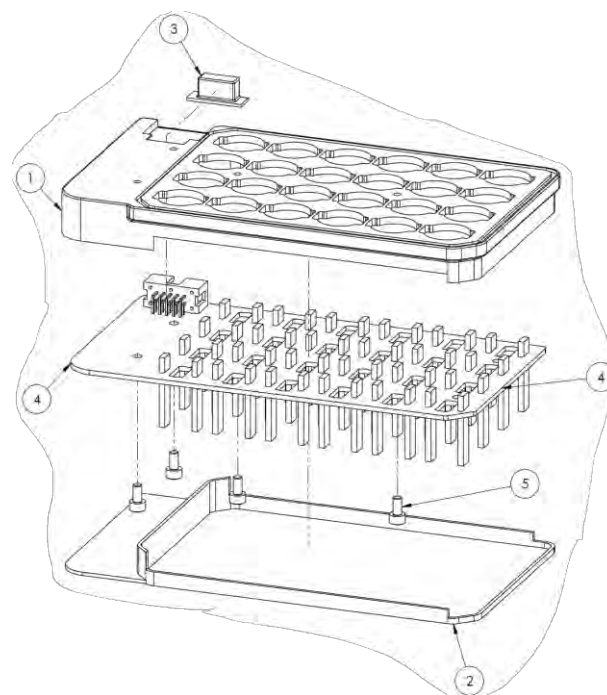


Figure 4. Exploded view of the stimulator. 1) Stimulator support, 2) Stimulator support base, 3) Stimulator connector cover, 4) Stimulator PCB assembly, 5) M3x6 screws

3.4. Evaluation of sarcolemmal damage

The validation of the system was made using 2D cultures due to them being easier to work with and their viability for a proof-of-concept study. We conducted EPS experiments using the newly developed system (figure 5) and quantified the sarcolemmal damage caused using ELISA assays (figure 6). As expected, the leakage of creatine kinase of DMD-affected cells was higher than that of healthy cells under almost all conditions. Longer stimulation times equalled more severe damage, but higher stimulation frequency did not. The results at 2 and 10Hz were very similar, while at 20Hz sarcolemmal damage was reduced. This could mean that twitch contractions cause more stress in the sarcolemma than tetanic ones.

Lysate samples had a high concentration of CK, but the difference between control and DMD was big. This could be attributed to the difference in cell density between healthy and DMD tissues (figure 7), or human errors during sample preparation, given the big disparity

between measurements. In most cases the variability of biological replicas was considerable, indicating that more replicas are needed.

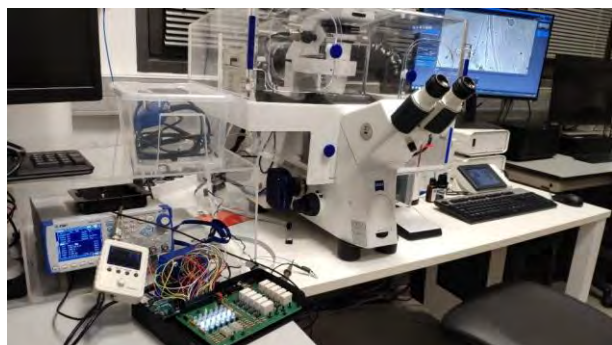


Figure 5. Real EPS experimental setup

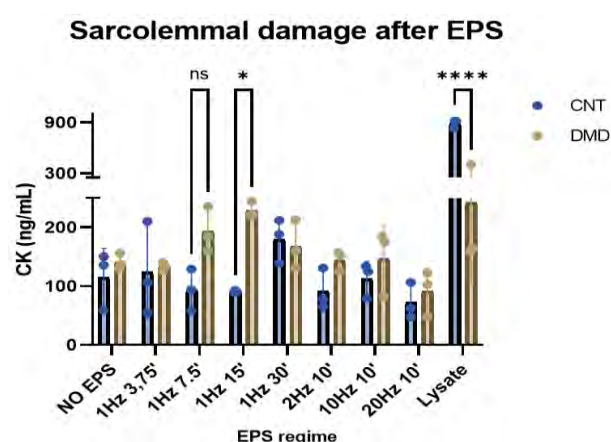


Figure 6. ELISA assay results

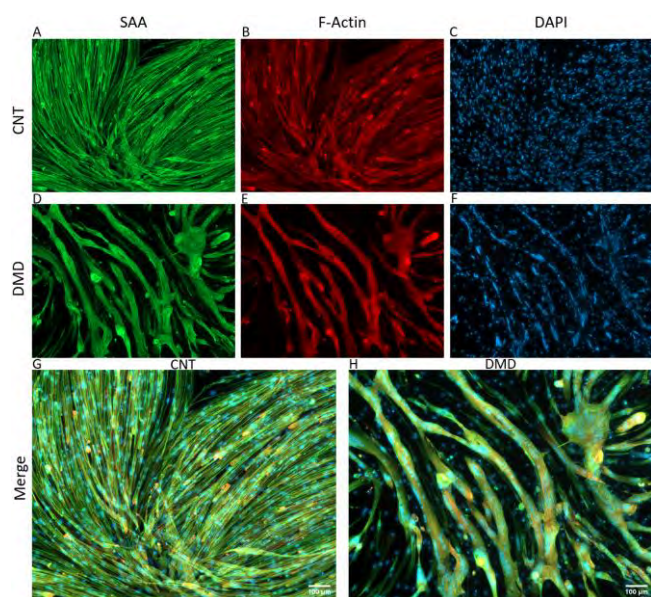


Figure 7. Control (A-C, G) and DMD-affected (E-F, H) 2D cultures immunostaining after EPS (10x). A, D) Sarcomeric α -actinin stained in green; B, E) F-Actin stained in red; C, F) Nuclei stained in blue; G, H) Image merge. Scale bars: 100 μ m.

3.5. Immunostaining

The fluorescence microscopy images clearly show differences in morphology between the two cell lines used. The healthy myoblasts fused into thinner myotubes, creating a dense net of muscle, whereas DMD affected cells formed thicker, more separated myotubes. The DMD culture appears more detached from the surface of the culture plate, further proving the increased susceptibility to contraction damaged caused by the disease.

4. Conclusions

The developed device accomplished faster electrical pulse stimulation of skeletal muscle cultures than the previous system. All designed and 3D printed parts worked as intended, ensuring electrical safety and sterile conditions. Research into sarcolemmal damage can help understand the main mechanisms of DMD and contribute to develop new treatments to fight the disease. The proposed system can also contribute to the study of other muscular dystrophies using 3D cultured tissues. In regard to Duchenne muscular dystrophy, it can be concluded that DMD-affected skeletal muscle is more susceptible to sarcolemmal damage under contraction than healthy muscle. Moreover, longer contraction periods result in bigger damage. Higher frequency contractions seem to produce less strain. However, these results need more experimentation to be confirmed.

Acknowledgements

The authors want to express our gratitude to all the members of the Biosensors for Bioengineering laboratory from IBEC. This work is supported in part by Duchenne Parent Project Spain.

References

- [1] H. D. Dave, M. Shook, and M. Varacallo, 'Anatomy, Skeletal Muscle', StatPearls, Aug. 2022, Accessed: May 21, 2023. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK537236/>
- [2] P. Soblechero-Martín, A. López-Martínez, L. de la Puente-Ovejero, A. Vallejo-Illarramendi, and V. Arechavala-Gomez, 'Utrophin modulator drugs as potential therapies for Duchenne and Becker muscular dystrophies', *Neuropathol Appl Neurobiol*, vol. 47, no. 6, pp. 711–723, 2021, doi: 10.1111/nan.12735.
- [3] E. M. Houang, Y. Y. Sham, F. S. Bates, and J. M. Metzger, 'Muscle membrane integrity in Duchenne muscular dystrophy: recent advances in copolymer-based muscle membrane stabilizers', *Skeletal Muscle* 2018 8:1, vol. 8, no. 1, pp. 1–19, Oct. 2018, doi: 10.1186/S13395-018-0177-7.
- [4] M. Kapalczyńska et al., '2D and 3D cell cultures – a comparison of different types of cancer cell cultures', *Arch Med Sci*, vol. 14, no. 4, p. 910, 2018, doi: 10.5114/AOMS.2016.63743.
- [5] C. Ma, Y. Peng, H. Li, and W. Chen, 'Organ-on-a-Chip: A New Paradigm for Drug Development', *Trends Pharmacol Sci*, vol. 42, no. 2, pp. 119–133, Feb. 2021, doi: 10.1016/J.TIPS.2020.11.009.
- [6] N. Nikolić, S. W. Görgens, G. H. Thoresen, V. Aas, J. Eckel, and K. Eckardt, 'Electrical pulse stimulation of cultured skeletal muscle cells as a model for in vitro exercise – possibilities and limitations', *Acta Physiologica*, vol. 220, no. 3, pp. 310–331, Jul. 2017, doi: 10.1111/APHA.12830.
- [7] Tabatabaei, M.S., Ahmed, M. (2022). Enzyme-Linked Immunosorbent Assay (ELISA). In: Christian, S.L. (eds) *Cancer Cell Biology. Methods in Molecular Biology*, vol 2508. Humana, New York, NY. https://doi.org/10.1007/978-1-0716-2376-3_10

Chamber design of a portable breathalyzer for disease diagnosis

D. Morales¹, T. Palacios², A. Díaz³

¹ Escuela Técnica Superior de Ingenieros de Telecomunicación, UPM, david.moralesl@alumnos.upm.es

² Microsystems Research Laboratories, Massachusetts Institute of Technology, tomasp@mit.edu

³ Escuela Técnica Superior de Ingenieros Industriales, UPM, andres.diaz@upm.es

Abstract

Breathalyzers allow, in a non-invasive way, to study the molecules present in the breath, enabling the diagnosis of some diseases due to the presence of low concentrations (ppb) of certain volatile organic compounds (VOCs). Developing these devices presents significant challenges, but the benefits of these techniques are tremendous, as they provide a noninvasive and cheaper diagnostic approach. The operation of these devices is simple; the patient only has to blow through a disposable mouthpiece into the measurement chamber, which houses the sensors that analyze the concentration of VOCs in individuals' breath.

Tomás Palacios group (Massachusetts Institute of Technology, Cambridge, United States) has built a graphene-based sensor array that can accurately measure the presence and concentration of different chemicals of interest. This bachelor's thesis develops a chamber design that allows this sensor to be used as a portable breathalyzer for disease diagnosis. This system could provide a non-invasive, cheap and rapid approach to disease diagnosis. Although significant research has been done in this field over the years, none has focused on the optimal chamber design of these devices, which must optimize contact between sensors and air samples and address issues such as moisture, air velocity control, recirculation and turbulence. This work studies the airflow properties in different chamber models and creates, with a careful component selection, a reusable in situ breath analyzer design with the help of fluid mechanics simulations and experiments with the analysis sensors.

1. Introduction

The field of medical diagnostics has undergone remarkable advancements in recent times, and the development of new technologies has enabled the early detection of various diseases with accuracy and ease. One such innovative solution are the breathalyzers, a cutting-edge technology that aims to diagnose various diseases by analyzing the volatile organic compounds (VOCs) in a patient's breath. This bachelor's thesis presents a comprehensive study on the design and development of a breathalyzer that, in the future, promises to bring a paradigm shift in the way many diseases are detected and diagnosed.

Breathing is an essential and natural process that occurs several times in a day, and the VOCs in our breath carry valuable information about our health. These VOCs are chemicals released by the body in normal conditions or in response to various diseases and health conditions, making our breath a rich source of information about our health [1]. However, traditional diagnostic methods are

often invasive and uncomfortable, leading to low patient compliance and limiting the potential of these methods.

The developed portable and reusable breathalyzer aims to address these limitations by providing a simple, non-invasive and painless solution to disease detection. The breathalyzer is designed to analyze the VOCs in a patient's breath and provide accurate diagnostic results within minutes. This device is portable and reusable, reducing the cost of diagnostic procedures and making it a cost-effective solution for patients and healthcare providers alike.

2. Design of the device

The goals of this project were to create a portable and reusable device that could perform prompt, safe and on-site analysis of breath samples to diagnose various diseases. To meet these goals, it was necessary to use disposable mouthpieces with one-way valves to minimize the patient's contact with the interior of the chamber. The chamber itself contains a sensor that analyzes the samples, as well as other components, and has an inlet and outlet for air circulation. It was essential to manage the sensor's performance across various scenarios due to the impact of moisture, turbulence and high air velocities, which caused significant signal noise. Additionally, preventing air recirculation within the chamber was necessary to prevent multiple readings of the same VOCs.

First, the various breath chamber models were 3D modeled to conduct fluid dynamics simulations and incorporate or modify elements to enhance their performance. Once the simulation results were good enough, the model was 3D printed and subjected to testing using the analysis sensor. This comprehensive testing evaluated the sensor's performance across diverse chamber models, various conditions, and the inclusion of elements such as filters and desiccants.

2.1. Fluid dynamics simulations

Hundreds fluid dynamics simulations were conducted to develop different generations of breathalyzers, with the objective of enhancing each model's features. The goal was to identify the optimal chamber geometry and size, component placement, as well as inlet and outlet positions to mitigate the effects mentioned in the previous paragraph, while maximizing contact between breath VOCs and the sensor surface.

The simulations were conducted using *COMSOL Multiphysics* software, with the complete device structure,

including the sensor and connector structures, positioned exactly as they were during the real tests. This approach, while time-consuming, ensured greater accuracy in the results.

Simulations, as well as previous studies, rapidly showed that the best chamber geometry for this application is cylindrical, since this shape helps reduce turbulence and air recirculation [2]. Furthermore, it was concluded that by preventing breath recirculation, and consequently, the repeated detection of the same VOCs, the optimal location to position the sensor was adjacent to the outlet. This placement allowed for the regulation of the velocity on the surface of the 1.7x1.7x0.8 cm measuring sensor through the outlet's size and geometry. The best outlet geometry is ellipsoidal to try to direct as many VOCs as possible to the surface of the sensor. For the same purpose, the sensor is placed on slightly inclined ramp and under a semi ellipsoidal piece connected to the outlet, that acts like a roof that direct breath to the sensor.

Through the simulations that were performed for this project using a fine mesh, the breath flow profiles, breath velocity field, turbulence variables, and the maximum and average Reynolds number were obtained, along with other useful data. The air velocity is always monitored to avoid generating noisy signals. Selecting the appropriate physics to simulate this system posed a challenge due to the intricate nature of fluid mechanics models and the range of options available in *COMSOL Multiphysics*. To measure turbulence with parameters other than the Reynolds number and air velocity on the sensor surface, the $k-\epsilon$ turbulent model was selected. This method has a faster convergence rate and a lower memory requirement than other turbulent models and is suitable for analyzing the turbulent kinetic energy and vorticity around complex geometries [3]. Other models, such as the $k-\omega$ turbulent model and laminar flow, were also considered. However, the latter only permits the calculation of the Reynolds number and none of the other useful parameters.

Figure 1 depicts the visual outcomes of the latest breathalyzer model simulations, while Table 1 displays some of the analyzed parameters and their evolution across the four breathalyzer generations, highlighting the significant improvements in the final model compared to its predecessors.

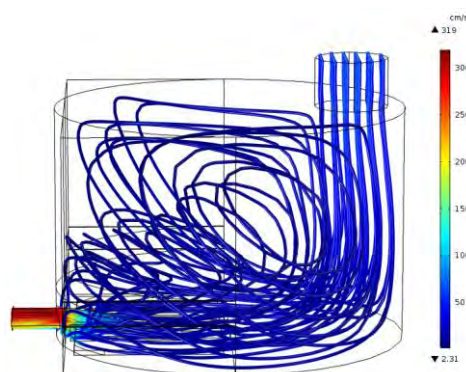


Figure 1. Airflow profile of the final breathalyzer generation.

Model	Inlet flow (mL/s)	Average Re	Max Re	Turbulent kinetic energy (cm ² /s ²)	Velocity on the sensor (cm/s)
First generation	25	17.57	29.86	34.89	1.89
Second generation	100	5.55	99.25	139.15	83.34
Third generation	100	5.78	83.32	121.94	113.12
Final generation	100	3.42	49.65	67.47	64.11

Table 1. Evolution of some of the analyzed parameters across the different models.

2.2. Materials

3D printing has the capability to address a multitude of fabrication obstacles through its highly automated process that solely necessitates a design file and can generate intricate features in all dimensions. Despite its convenience, the materials employed in 3D printing, such as ABS, are not ideal for exposure to VOCs, since the final prints contain residual uncured monomers other compounds resulting from these resins, which may react with VOCs [4].

While ABS may lead to issues, it is unnecessary to choose different materials or techniques to construct the breathalyzer. Instead, coating the surfaces that will come into contact with VOCs with parylene, a thermoplastic polymer, is sufficient [4]. Parylene is often applied to medical devices due to its chemical inertness and ability to act as a hydrophobic and impermeable barrier to small molecules, water, and gases [5]. By coating the interior walls of the chamber with this material, breath can be prevented from coming into contact with them and, thanks to its hydrophobic properties, breath condensation can be prevented, avoiding the loss of detection of highly hydrophilic VOCs like acetone and ethanol. Additionally, parylene can prevent the interaction of VOCs with ABS and create a better seal for the chamber, as this 3D printed material is quite porous.

Thermal deposition is the technique that allowed for coating the walls of the 3D-printed breathalyzer chamber. However, when assessing the roughness of the ABS walls to determine the required amount of parylene, it was observed that the mean roughness of horizontally and vertically printed surfaces was 16.88 μm and 58.14 μm , respectively. Given that thermal deposition of large quantities of parylene can become prohibitively expensive, especially considering that typical coating thicknesses typically range from 1 to 10 μm [6], alternative approaches were explored.

Given ABS's high reactivity to acetone, it was decided to use acetone vapor smoothing techniques to reduce wall roughness [7]. This process makes the polymer chains move smoothly against each other, which reduces porosity and fills gaps between adjacent layers,

which is advantageous for decreasing the outgassing effect in vacuum [7]. 3D printed ABS surfaces were exposed to acetone vapor in a controlled environment, which was very effective, reducing the mean roughness of horizontally and vertically printed surfaces to 2.65 μm and 5.60 μm , respectively.

Finally, after completing this process, the 3D-printed breathalyzer chamber was coated with a 10 μm layer of parylene using thermal deposition.

3. Final breathalyzer generation

The ultimate goal of the breathalyzer's final generation was to support clinical trials scheduled for the summer. Consequently, several additional components, in addition to the sensor and the chamber, needed to be incorporated, as depicted in Figure 2. The complete device, visible on the left side of the image, comprises two chambers. Patients effortlessly exhale through a disposable mouthpiece equipped with a one-way valve and a humidity reduction filter. The exhaled breath is then directed through a tube leading to the upper chamber. This compartment houses the sensor, its connectors, a heater, and a humidity sensor. The bottom part of the upper chamber, on which the sensor is placed on an inclined ramp, serves as the top section of the lower compartment, which includes a set of switches and LEDs to enable wireless device operation while it is in use. Additionally, this lower compartment houses a battery to supply power to the system and three stacked PCBs. These PCBs are responsible for controlling the sensor and managing the device's operation. They also incorporate a Bluetooth antenna for transmitting the collected data to a mobile and tablet app, which is connected to a database. Furthermore, this app employs artificial intelligence to perform diagnostic evaluations based on the sensor's detections.

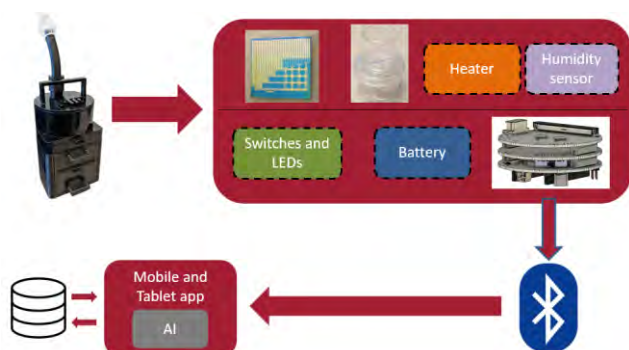


Figure 2. Final model workflow.

It is important to clarify that four people were involved in the system explained above. The development of the sensor and PCBs, data communication and storage, and the design of artificial intelligence were the contributions of my colleagues. However, the design of the chambers, the selection of the necessary components, and their placement were exclusively my work.

Each of the two chambers has a lid for opening the compartments to access the devices housed within them, should replacement be necessary. The heater in the top

chamber is designed to warm the interior, reducing the likelihood of condensation, potentially enhancing the sensor's performance, and helping to clean the surface of the sensor by removing any possible remaining VOCs from the previous patient. As previously noted, on the right side of the bottom chamber, a set of switches and LEDs is situated to operate the device. This enables healthcare professionals to power the entire device on and off, and it also indicates patients when to commence exhaling and analyze their breath samples.

4. Results and discussion

As mentioned, several experiments were conducted with each breathalyzer generation to continuously improve each model. Reducing moisture inside the chamber was crucial to enhance the sensor's performance. To address this, various approaches were explored, including the addition of dehumidifying filters and desiccants. However, for the sake of user convenience and to eliminate the need for replacing desiccant stones after each breath test, the decision was made to utilize disposable mouthpieces with integrated humidity reduction filters, replaced after each breath test. Also, it has been demonstrated that desiccants can absorb relevant VOCs used in diagnosis [8].

The typical signals measured by the sensor resemble those shown in Figure 3. This graphene-based sensor comprises 1024 independent devices, each functionalized with four different chemicals on its surface, which function as transistors and react to VOCs in various ways. The vertical axis represents I_{DS} , signifying the change in electrical current between the drain and source when individuals exhale into the chamber. In Figure 3, this change occurs at the 120-second mark, marked by a sudden increase in amplitude across most areas, indicating that VOCs are interacting with the sensor's chemicals. This event is succeeded by a plateau, which is the most crucial aspect for analysis, and a variable duration for the signal to return to the baseline level. The outcomes of various experiments, such as those shown in Figure 3, can sometimes be so similar that they cannot be reliably compared solely by visual inspection. This is why certain scripts were developed to measure the different parameters at each stage of the signals.

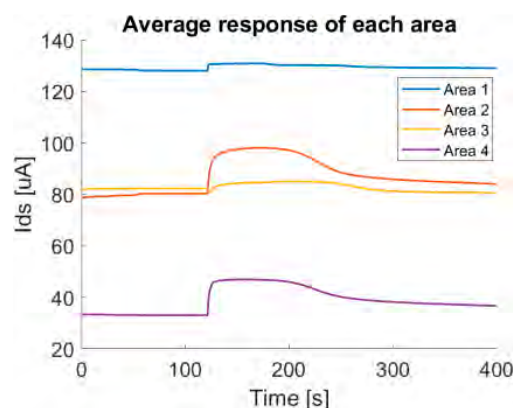


Figure 3. Typical measured signals corresponding to the latest breathalyzer model when no humidity reduction methods were used.

Although desiccants were not used in the final version of the device, their impact regarding humidity and their previously mentioned VOC absorption (which affects hydrophilic particles in particular) was tested [8]. The four areas functionalized with different chemicals react differently when in contact with breath. Figure 4 represents the average response of one of the most representative areas under four conditions: when no humidity reduction methods were used (bare chamber), when a filter was used, when a desiccant was used, and when both were used. It can be observed that when the bare chamber was used, both the amplitude and excitation time of the sensor were greater than in the other cases. When the filter was used, only the amplitude was reduced because water molecules were removed from the interior of the chamber, resulting in less I_{DS} . However, when a desiccant was used, both amplitude and excitation time were reduced, leading to fewer VOCs coming into contact with the sensor. In the last case, when both a filter and desiccant were used, humidity was significantly reduced along with the detected VOCs. Furthermore, this scenario gave rise to a very short and low signal that could not be analyzed by the artificial intelligence for diagnosis.

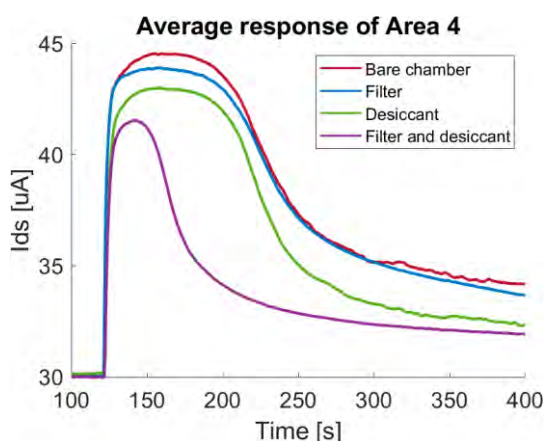


Figure 4. Average response to the breath in Area 4 of the analysis sensor.

5. Conclusions

This bachelor's thesis represents a comprehensive exploration into the design and creation of a portable and reusable breathalyzer, an innovation poised to redefine disease detection and diagnosis. By analyzing the volatile organic compounds (VOCs) present in our breath, this groundbreaking device holds the potential to provide invaluable insights into our health. VOCs, released by the body under various conditions, offer a wealth of information about our well-being. The portable and reusable nature of this device ensures swift and accurate diagnostic results, all while being simple, non-invasive, and painless for the user.

The primary goal of this device is to facilitate rapid and efficient analysis while minimizing patient interaction with the chamber's interior. This is achieved through the use of disposable mouthpieces. The design featured a sensor enclosed within a chamber, and the sensor's performance was influenced by factors such as high air velocities and moisture. These variables were successfully

controlled through chamber geometry adjustments and the incorporation of methods like dehumidifying filters at the inlet. To further enhance the breathalyzer's functionality, the inner walls of the final versions underwent acetone vapor smoothing and were coated with parylene. This coating reduced interaction between VOCs and the printing material while enhancing chamber sealing.

In conclusion, this bachelor's thesis has provided a comprehensive examination of the development of a portable and reusable breathalyzer. The research showcases the transformative potential of this innovative technology, offering a paradigm shift in disease detection and diagnosis. It offers a straightforward, non-invasive, and painless method for identifying various health conditions. The continued development of this device holds immense promise for the future of medical diagnostics and has the capacity to revolutionize the healthcare industry. Moreover, it is worth noting that this device has already seen successful application in clinical trials, underscoring its real-world potential and the positive impact it can have on healthcare practices.

References

- [1] Mika Shirasu and Kazushige Touhara. "The scent of disease: volatile organic compounds of the human body related to disease and disorder". *J. Biochem.* 150.3 (Sept. 2011), pp. 257–266.
- [2] Ramji Kalidoss, Snehalatha Umapathy, and Usha Rani Thirunavukkarasu. "A breathalyzer for the assessment of chronic kidney disease patients' breathprint: Breath flow dynamic simulation on the measurement chamber and experimental investigation". *Biomed. Signal Process. Control* 70.103060 (Sept. 2021), p. 103060.
- [3] Walter Frei. "Which turbulence model should I choose for my CFD application?" [En línea] Disponible en: <https://www.comsol.com/blogs/which-turbulence-model-should-choose-cfdapplication/>. Accedido: 2023-4-1. Julio 2017.
- [4] Brian J O'Grady et al. "Rapid prototyping of cell culture microdevices using parylenecoated 3D prints". Agosto 2021.
- [5] Monika Golda-Cepa et al. "Recent progress on parylene C polymer for biomedical applications: A review". *Prog. Org. Coat.* 140.105493 (Mar. 2020), p. 105493.
- [6] "How much does parylene cost?" [En línea] Disponible en: <https://www.paryleneconformalcoating.com>. Accedido: 2023-4-1.
- [7] Viviane Pestano, Mariana Pohlmann, and Fabio Pinto da Silva. "Effect of acetone vapor smoothing process on surface finish and geometric accuracy of fused deposition modeling ABS parts". *J. Mater. Sci. Chem. Eng.* 10.10 (2022), pp. 1–9.
- [8] Zhu, L., Shen, D., & Luo, K. H. (2020). A critical review on VOCs adsorption by different porous materials: Species, mechanisms and modification methods. *Journal of Hazardous Materials*, 122102. doi:10.1016/j.jhazmat.2020.122102

Análisis de electroencefalografía de amplitud integrada en neonatos: Evaluación del rendimiento de aCUP-E frente a los electrodos de gel líquido

A. Rodríguez Ballabriga¹, A. Fabregat-Sanjuan¹, A. Rigo Vidal², R. Pàmies-Vilà³, V. Rius Costa², S. Larrosa Capacés², V. Pascual Rubio²

¹ FUNCMAT, Departamento de Ingeniería Mecánica, Universitat Rovira i Virgili, Tarragona, España
{angel.rodriguez@estudiants.urv.cat, a.fabregat@urv.cat }

² NEUROEPIA, Institut d'Investigació Sanitària Pere Virgili. Departamento de Neurofisiología Clínica, Hospital Sant Joan de Reus, España.

³ BIOMECH, Departamento de Ingeniería Mecánica, Universitat Politècnica de Catalunya, España

Resumen

Los neonatólogos y neurofisiólogos clínicos coinciden en que no existen electrodos que permitan realizar registros de larga duración, limitando la capacidad de la electroencefalografía de amplitud integrada (aEEG) para diagnosticar daños cerebrales en las Unidades de Cuidados Intensivos Neonatales (UCIN). En este estudio se realiza un análisis comparativo entre electrodos comerciales de gel líquido y un electrodo nuevo y específico (aCUP-E) para evaluar su rendimiento en un ensayo clínico con 15 neonatos neurológicamente sanos. Los resultados muestran que las mejoras que presenta aCUP-E aumentan la estabilidad de la impedancia y la calidad de los registros, así como una clara disposición del personal de las UCIN a utilizar el nuevo electrodo. Como conclusión, esta investigación pone de manifiesto las limitaciones de los electrodos comerciales en aEEG y propone aCUP-E como una alternativa mejor.

1. Introducción

En las Unidades de Cuidados Intensivos Neonatales (UCIN) las urgencias neurológicas son una preocupación, tanto para los clínicos como para los padres, a un nivel en el que los cuidados neurocríticos del recién nacido requieren unos conocimientos específicos. Las patologías neurológicas representan el 20% de los ingresos actuales en las UCIN [1].

Aunque los avances en la monitorización neonatal han mejorado las tasas de supervivencia de los recién nacidos, la evaluación de la función cerebral en neonatos sigue siendo un área que requiere un mayor desarrollo. Como consecuencia de lo anterior, tanto los neonatólogos como los neurofisiólogos clínicos coinciden en la necesidad de mejorar las técnicas actuales de registro de la actividad cerebral, con el objetivo de mejorar la atención general de neonatos en estado crítico, permitir una intervención temprana y disminuir la probabilidad de retrasos en el desarrollo cerebral a largo plazo [2].

En este contexto, apareció la electroencefalografía de amplitud integrada (aEEG), que es una técnica usada en las UCIN alrededor del mundo. Sus ventajas principales radican en que es una técnica de monitoreo continuo y se usan menos electrodos que en la electroencefalografía (EEG), permitiendo el acceso craneal al personal médico y

un montaje más rápido. Sin embargo, todavía no se ha implantado completamente en la clínica como *gold standard*, ya que existe una controversia entre los profesionales clínicos sobre su interpretación, ya que los registros de aEEG habitualmente presentan muchos artefactos y pueden confundir al clínico en la toma de decisiones [3].

La aEEG fue propuesta en 1969 por D. Maynard y P.F. Prior [4] y su funcionamiento no ha cambiado mucho. La señal de aEEG se deriva del electroencefalograma y resulta en una representación comprimida (6 cm/h), filtrada y rectificadas mostrada en escala semilogarítmica para dar más importancia a la señal de menor voltaje. La señal muestra 2 bandas, el margen superior e inferior, relacionadas con la actividad de alto y bajo voltaje respectivamente. Será la evolución y los patrones que presenten estos límites lo que permitirá evaluar e identificar diversas patologías.

Al igual que en las otras técnicas de electroencefalografía, la impedancia juega un papel crucial. Este concepto está muy relacionado con el contacto piel-electrodo. En los neonatos, se ha supuesto que la impedancia es menor debido a una mayor hidratación de la piel inmadura. Sin embargo, la preparación del cuero cabelludo sigue siendo necesaria para obtener un buen registro [3]. En la aEEG, el montaje bipolar y el número reducido de electrodos provoca que la impedancia sea aún más importante para la interpretación y la toma de decisiones. Un mal contacto de uno de los electrodos provocará una mayor variabilidad de la impedancia, que puede producir picos en la señal capturada que pueden ser confundidos, por ejemplo, con una epilepsia.

En consecuencia, se puede afirmar que los electrodos son un factor clave para esta técnica, y aún quedan algunos retos por superar. Por ese motivo, el objetivo de este estudio es presentar y comparar en un ensayo clínico un nuevo electrodo, diseñado específicamente para el registro de aEEG en neonatos, con otro ya existente (considerado como el mejor para monitorización neonatal), evaluando si las mejoras presentes en el primero se traducen en una mejor captura de la actividad cerebral y una mejor usabilidad.

2. Materiales y métodos

El análisis se ha desarrollado dentro de un ensayo clínico. Para lograr el objetivo, se ha computado una interfaz gráfica para poder visualizar la aEEG, así como diversos parámetros para describir el rendimiento de los electrodos.

2.1. Electrodo

El Dr. Vicenç Pascual, neurofisiólogo clínico del *Hospital Universitari Sant Joan de Reus (HUSJR)*, reconoció las deficiencias y los problemas que presentan los electrodos para adultos cuando se aplican a neonatos. Así, junto con el Dr. Albert Fabregat, del *Departamento de Ingeniería Mecánica* de la *Universitat Rovira i Virgili*, desarrollaron un electrodo que cubriera las necesidades y se adaptara a la situación específica de los recién nacidos.

El resultado de esta innovación fue *Advanced Cup Electrode (aCUP-E)*, un nuevo electrodo diseñado específicamente para adaptarse a las necesidades del neuromonitoreo de larga duración de los neonatos en las UCIN: ajuste a la curvatura craneal de los recién nacidos, flexibilidad, adherencia firme, capacidad de recambiar el gel electroconductor y seguridad para su frágil piel.

Por otro lado, el electrodo Neuroline720™ de Ambu® ha sido seleccionado como el electrodo comercial en este estudio, ya que son considerados como los mejores para la neuromonitorización neonatal [5] y recomendados por los centros con más experiencia en aEEG. Son autoadhesivos y ya llevan introducido el gel electroconductor.

2.2. Plan de investigación clínica

Este trabajo se ha realizado en el marco de un ensayo clínico llevado a cabo en el Hospital Universitari Sant Joan de Reus por el Dr. Vicenç Pascual Rubio, y que dispone de la aprobación del comité de ética de investigación con medicamentos (CEIm) y de la agencia española de medicamentos y productos sanitarios (AEMPS).

La hipótesis inicial es que el nuevo electrodo aCUP-E ofrece al menos la misma seguridad y calidad de la señal bioeléctrica que los electrodos de gel líquido. Presenta una mejor adherencia, una impedancia más estable y permite una manipulación más cómoda del recién nacido.

2.3. Procedimiento

Se ha realizado el estudio a 15 neonatos (4 a término y 11 prematuros), mientras permanecían en la UCIN por causas no neurológicas. En las sesiones de registro se han utilizado simultáneamente los dos tipos diferentes de electrodos. Así, en un hemisferio se ha registrado con electrodos de gel líquido y el otro con el aCUP-E. Para el registro de aEEG se han colocado 5 electrodos en el cuero cabelludo siguiendo el sistema internacional 10-20: C3 y P3 (hemisferio izquierdo), C4 y P4 (hemisferio derecho) y Cz (referencia y tierra). Antes de iniciar el registro y colocar los electrodos, se ha preparado el cuero cabelludo del neonato para optimizar el registro de aEEG.

2.4. Adquisición de datos y procesado

Se ha utilizado el monitor *CFM Olympic Brainz (Natus Medical Incorporated)* para adquirir las señales aEEG. Del

monitor se han extraído los archivos de formato europeo de datos (EDF) que contienen la señal cruda de la EEG y las impedancias continuas. No existe ninguna opción para descargar los archivos de la señal aEEG en este dispositivo. El procesamiento de la señal se ha realizado principalmente utilizando MATLAB de *The MathWorks, Inc. (2022)*, pero los archivos se han convertido a CSV (*Comma Separated Values*) utilizando el paquete *MNE* de *Python*.

2.5. Obtención digital de la señal aEEG

Las empresas que comercializan los monitores de función cerebral consideran el procesamiento de la aEEG parte de su propiedad intelectual. Esto no permite la obtención directa del registro de aEEG ni conocer los detalles del procesado de la señal electroencefalográfica usado.

Así pues, para poder facilitar la visualización de la aEEG en el ordenador, procesar las señales y poder extraer características de los diferentes electrodos, se ha procesado la señal cruda de la EEG extraída del monitor, y se ha creado una interfaz gráfica para poder visualizar la aEEG de manera interactiva. Además, las variables se pueden guardar en el *workspace* para su análisis matemático.

El objetivo de este enfoque es imitar el monitor utilizado por los clínicos, en cualquier ordenador. Para ello, la interfaz debe tener, al menos, las mismas características que el dispositivo médico. El código para obtener la señal aEEG se ha adaptado de *Washington University Neonatal EEG Analysis Toolbox (WU-NEAT, 2020)*, y se ha basado principalmente en Zhang D. *et al.* [6] y Chen C. *et al.* [7]

El procesado incluye un filtro asimétrico que compensa por las distintas atenuaciones del cráneo a distintas frecuencias y elimina artefactos (el rango de 2 a 15 Hz es el de interés en aEEG), un rectificado de la señal (valor absoluto), detección de la envolvente (filtro Butterworth paso bajo), representación en escala semilogarítmica (eje y) y con una proporción definida de 6 cm/h (eje x).

Se ha ajustado el algoritmo (frecuencias de corte, orden de los filtros, proporciones de visualización...), para poder obtener una señal idéntica a la del monitor utilizado en la práctica clínica. La comparación entre las dos señales no se ha podido realizar estadísticamente, ya que la señal aEEG no se puede extraer del dispositivo médico. Así pues, se ha realizado una comparación visual entre las capturas de pantalla extraídas del monitor y la visualización de la interfaz de usuario (IU) creada, donde los médicos no han percibido diferencias para el diagnóstico.

También se han procesado los márgenes superior e inferior, así como la actividad cerebral no óptima (que probablemente no provenga del cerebro, contaminada) y se han representado en la IU, permitiendo ver más características de la señal que las que permite el monitor (Figura 1).

Para detectar la actividad cerebral óptima, se han propuesto dos criterios neurofisiológicos: Es poco probable que las ondas de EEG de más de 150 μV y de más de dos segundos de longitud de onda provengan de la actividad cerebral, y, que las amplitudes de pico a pico cercanas a 0 en aEEG se deben a otros factores (p.ej. puente salino).

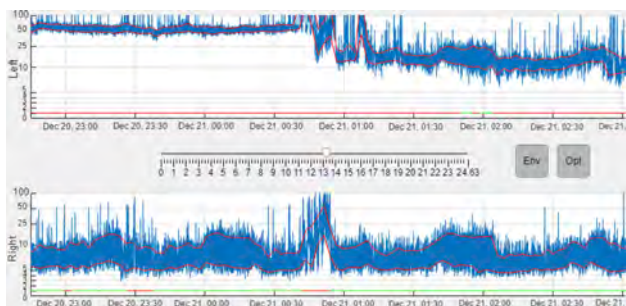


Figura 1: IU construida en MATLAB, simulando un monitor aEEG. La señal aEEG (azul), la envolvente (línea roja) y las partes donde se ha detectado actividad cerebral no cerebral (línea recta en la parte inferior de cada hemisferio). Verde: Actividad cerebral, Rojo: Artefactos detectados). La señal de la parte superior corresponde al hemisferio izquierdo (comerciales), mientras que la de la parte inferior corresponde al hemisferio derecho (aCUP-E). El botón *Env* activa los márgenes, *Opt* activa la línea óptima/no óptima y el *slider* permite navegar en el tiempo (horas).

2.6. Parámetros y análisis estadístico

Para comparar y caracterizar la diferencia entre los hemisferios de cada sujeto: estabilidad de la impedancia y calidad de la señal (*Signal to Noise Ratio*, SNR) (Ecuación 1), es necesario realizar una prueba estadística para evaluar si las variaciones son significativas. Se ha utilizado la prueba no paramétrica U de Mann-Whitney. La prueba se ha realizado unilateralmente, en la que la hipótesis alternativa es que aCUP-E funciona mejor (mayor o menor según el caso), con $\alpha=0.05$.

$$SNR = 10 \times \log_{10} \left(\frac{\text{PotenciaSeñal}_{2-15\text{Hz}} (\mu\text{V}^2)}{\text{PotenciaRuido}_{50\text{Hz}} (\mu\text{V}^2)} \right) \text{ (dB)} \quad (1)$$

Al evaluar la supervivencia de los electrodos (curvas de Kaplan-Meier), se ha utilizado una prueba Tarone-Ware ($p\text{-value} < \alpha$) para probar la hipótesis nula de que no hay diferencia entre los dos electrodos en observar un evento (cese de funcionamiento) en cualquier momento.

La susceptibilidad mide la captura de señal y ruido (SNR) dependiendo de los niveles de impedancia. Para evaluar las curvas, se ha realizado para cada tipo de electrodo un ajuste de curva que relacione la SNR (dB) y la impedancia en k Ω . Se han evaluado dos parámetros que describen la calidad del ajuste (R^2 *adjusted* y *Root Mean Square Error*, RMSE) de distintas regresiones.

Finalmente, se ha diseñado una encuesta, estructurada en 2 secciones principales para comprobar la facilidad de uso y la seguridad cutánea, además de recopilar algunos comentarios del personal de enfermería.

3. Resultados y discusión

De los 15 neonatos neurológicamente sanos que se han introducido en este estudio, se ha registrado un 20% más de actividad neuronal óptima con aCUP-E que con los electrodos comerciales. Esto es equivalente a 58 horas más de registro sin artefactos.

Centrándose únicamente en los valores de las impedancias a lo largo de estos registros, aCUP-E ha estado solo 9% del tiempo sobre 20 k Ω , mientras que los comerciales han estado el 26% del tiempo. Además, aCUP-E ha estado el

74% del tiempo por debajo de 10 k Ω (rango óptimo de impedancia), mientras que los comerciales el 67% del tiempo. Estas diferencias en porcentaje conllevan muchas horas de registro perdidas por parte de los comerciales, que se verían aumentadas si no se hubiese hecho un control exhaustivo sobre las impedancias, recambiando los electrodos comerciales siempre que no estuviesen registrando adecuadamente.

Estos ceses de funcionamiento de los electrodos se han detectado a posteriori mediante las señales de impedancia, permitiendo realizar una curva de supervivencia. Una vez detectados los eventos, estos han sido refinados de acuerdo con las observaciones hechas por el personal médico en el monitor, evitando considerar otros eventos como mal funcionamiento del electrodo. La Figura 2 muestra esta curva, donde se puede observar como a lo largo de las primeras 24 horas de registro, menos del 40% de los comerciales sobreviven mientras que más del 80% de los aCUP-E no deben ser recambiados.

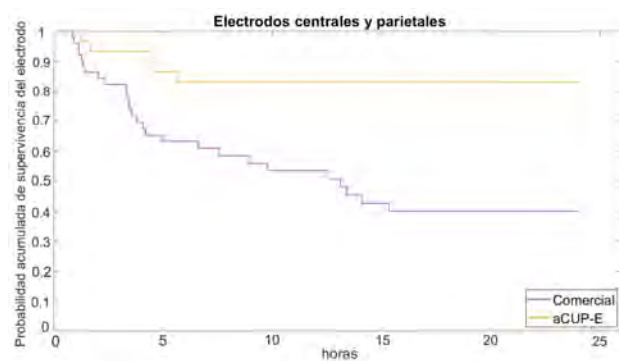


Figura 2: Análisis de supervivencia corregido (curva de Kaplan-Meier) en el que no se consideran las manipulaciones y los incumplimientos del protocolo como un cese del funcionamiento. (Tarone-Ware: $\chi^2=11,38$, $p\text{-value} = 7,42e-4$)

En el estudio realizado por Ters *et al.* [8] sobre la supervivencia, se realizó una comparación entre los electrodos de hidrogel y los electrodos de copa de oro. A las 24 horas, los electrodos de hidrogel habían sobrevivido de forma similar a los electrodos de gel líquido, mientras que los electrodos de copa de oro aún menos. Esto ayuda a reforzar la validez de nuestros resultados, y establece aCUP-E como el electrodo con mayor tiempo de vida.

Otro factor a tener en cuenta es la estabilidad de la impedancia. Tal y como se ha comentado anteriormente, esta puede producir simulaciones de epilepsia y provocar un errores de interpretación clínica. Para caracterizar este parámetro, se ha calculado una desviación estándar móvil para segmentos de 30 minutos, con un *overlap* del 50%.

Comparando la media de todos estos segmentos para cada registro, se ha obtenido un $p\text{-value}=0,049$, reforzando la idea de que aCUP-E se adhiere mejor al cuero cabelludo del neonato, permitiendo así menor variabilidad de la impedancia. El estudio realizado por Foreman *et al.* [3] muestra las dificultades de los electrodos de hidrogel para mantener una impedancia estable y la importancia de una buena preparación de la piel. Por lo tanto, aCUP-E podría ser una buena alternativa no invasiva para mejorar esta característica.

Otra cuestión de interés es preguntarse si ambos electrodos capturan la misma cantidad de señal para las distintas impedancias, y si hay alguna característica de aCUP-E que pueda afectar este comportamiento. Tal y como se puede observar en la Figura 3, ambos electrodos tienen una tendencia (coeficientes de regresión) similar. Se ha ajustado esta curva mediante una ecuación polinomial, una exponencial y una potencial, obteniendo siempre un coeficiente de ajuste (R^2 adjusted y RMSE) similar, con diferencias inferiores al 3%.

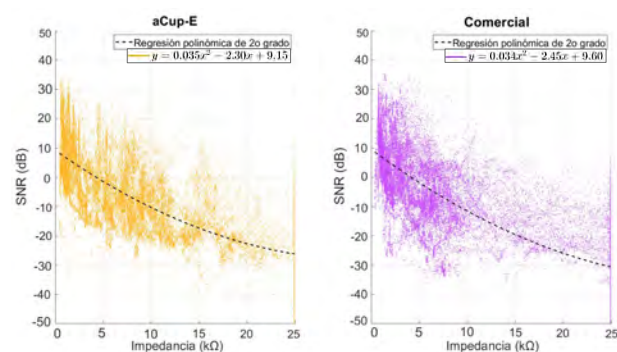


Figura 3: Relación señal/ruido (SNR) capturada para los electrodos aCUP-E (izquierda) y comercial (derecha) en dB a diferentes impedancias (kΩ). La línea negra discontinua representa la regresión (polinomio de 2º orden) de los datos y la leyenda muestra su ecuación (coeficientes de regresión).

La susceptibilidad al ruido es una característica de la placa del electrodo, donde el contacto es un factor crucial [9]. Por tanto, las similitudes encontradas entre electrodos eran de esperar, ya que el material del sensor no es un factor diferencial de aCUP-E. Las regresiones anteriores podrían ser una modelización de cuál es el comportamiento de un electrodo de Ag/AgCl al realizar aEEG.

Pero, tal y como hemos visto en los apartados anteriores, aCUP-E presenta mejores niveles de impedancia y menor variabilidad de esta. Por tanto, se puede concluir que esos factores también son motivo de tener una mayor calidad de la señal (SNR elevado y $p\text{-value} < 2,18e-4$ en los casos estudiados).

Por último, después de cada grabación se ha realizado una encuesta de opinión para medir la satisfacción del personal sanitario con el nuevo electrodo. Los resultados muestran que el personal de las UCIN considera que aCUP-E permite una mejor manipulación del recién nacido (94%), permitiendo el método canguro o piel con piel. Los electrodos comerciales son más fáciles de retirar, debido a su menor adherencia, mientras que ambos electrodos presentan las mismas condiciones de seguridad cutánea.

4. Conclusiones

En este estudio se ha desarrollado un código para crear la señal aEEG a partir de la señal electroencefalográfica cruda exportada desde el dispositivo médico. Este programa puede ser útil para clínicos e investigadores para obtener la señal aEEG y poder procesarla en un ordenador, con capacidades ampliadas.

Los resultados comparativos muestran que los electrodos aCUP-E presentan una mejor adherencia y una mayor

estabilidad de impedancia (40% menos variabilidad) que los utilizados actualmente en la práctica clínica. Esto se traduce en más del doble de vida útil del electrodo.

En consecuencia, las innovaciones presentes en aCUP-E, destacando la adhesión y la capacidad de recambio del gel electroconductor, permiten una mejor calidad de la señal registrada (20% más de registro óptimo). Este incremento permitiría mejorar el diagnóstico/tratamiento de los recién nacidos. Además, se ha visto una mejora en la usabilidad y una clara disposición del personal de las UCIN a usar el nuevo electrodo. Estos resultados son muy importantes para poner de manifiesto las deficiencias que suponen los electrodos utilizados actualmente en la práctica clínica, afirmando que aCUP-E es una alternativa viable.

Hay que tener en cuenta que estos resultados proceden de un ensayo clínico con neonatos, en el que las condiciones de cada recién nacido son diferentes y nunca ideales. Por lo tanto, las diferencias encontradas en las que aCUP-E rinde mejor, podrían ser realmente significativas para evaluar su introducción en la práctica clínica.

5. Agradecimientos

Este estudio ha sido financiado por “la Caixa” mediante el contrato LCF/TR/CI21/52650005.

6. Referencias

- [1] Benavente I, Lubián S, and Lechuga A. Estudio del electroencefalograma integrado por amplitud normal y patológico, y su relación con el pronóstico, en recién nacidos prematuros de muy bajo peso al nacimiento. *Tesis doctoral*, Universidad de Cádiz, Cadiz, 2015.
- [2] Glass H. *et al.* Neurocritical Care for Neonates. *Neurocrit Care*, vol 12, no. 3, 2010, pp 421–429 (ISSN: 1541-6933)
- [3] Foreman S *et al.* Electrode Challenges in Amplitude-Integrated Electroencephalography: Research Application of a Novel Noninvasive Measure of Brain Function in Preterm Infants. *Biological Research for Nursing*, vol 13, No. 3, 2011, pp. 251–259 (ISSN: 1099-8004)
- [4] Maynard D, Prior PF, and Scott DF. Device for continuous monitoring of cerebral activity in resuscitated patients. *British Medical Journal*, vol 4, No. 5682, 2019, pp. 545–546 (ISSN: 0007-1447)
- [5] Cordeiro M *et al.* Evaluation of the suitability and clinical applicability of different electrodes for aEEG/cEEG monitoring in the extremely premature infant. *An Pediatr*, vol.95, No. 6, 2021, pp 423–430 (ISSN: 1695-9531)
- [6] Zhang D and Ding H. Calculation of compact amplitude-integrated EEG tracing and upper and lower margins using raw EEG data. *Health N Hav*, vol 05, No. 05, 2013 pp. 885–891 (ISSN: 1949-4998)
- [7] Chen C *et al.* A digitized approach for amplitude-integrated electroencefalogram transformation towards a standardized procedure. *Biomed Signal Process Control*, vol 66, 2021 (ISSN: 1746-8108)
- [8] El Ters N *et al.* Long term electroencefalography in preterm neonates: Safety and quality of electrode types. *Clinical Neurophysiology*, vol 129, No. 7, 2018, pp 1366–1371 (ISSN: 1872-8952).
- [9] Lee M *et al.* Characterization of Ag/AgCl Dry Electrodes for Wearable Electrophysiological Sensing,” *Frontiers in Electronics*, vol 2, 2022 (ISSN: 2673-5857)

Analyses of H₂O₂ production in time during wound healing

C. Salamanca-González^{1,2}, I. Sorrentino¹, A.K Molina-Oviedo¹, E. Arévalo-Nuñez de Arenas¹, I. Medraño-Fernández^{1,2}

¹ Department of Bioengineering, University Carlos III de Madrid, 28911, Madrid, Spain

² Corresponding authors: 100428923@alumnos.uc3m.es ; imedrano@ing.uc3m.es

Summary

Reactive Oxygen Species (ROS) have been traditionally considered as toxic species due to their capacity to damage macromolecules. However, their deleterious effect is limited to high concentrations. At lower levels, further studies have shown that at least hydrogen peroxide (H₂O₂) acts as a second messenger and results crucial for basic biological functions. Wound healing is one of those: in several animal models, this molecule is emitted in cellular wound margins, leading to neutrophil chemoattraction, while inducing epithelial cell migration during skin regeneration. Nevertheless, there are no references to this phenomenon in human experimental set-ups. In this study, human keratinocyte cells (HaCaT), previously modified to express a H₂O₂-specific ratiometric fluorescent sensor (HyPer7) were monitored using live imaging after generating a mechanical-like injury. The time evolution ratio of the probe and the gap area were measured for both cells near the margin and for topologically secondary cells, establishing a relationship between cytosolic H₂O₂ fluxes and the process of refilling a wound. In addition, comparison of movement rates between HyPer7-expressing and non-expressing HaCaT cells demonstrated a profound dependency of H₂O₂ availability for the process to occur.

1. Introduction

1.1 Reactive Oxygen Species and H₂O₂

Reactive Oxygen Species (ROS) are intermediate molecules produced by the sequential reduction of molecular oxygen (O₂) to water (H₂O). They include: the superoxide anion (O₂^{•-}), hydrogen peroxide (H₂O₂) and hydroxyl radicals (•OH). These molecules cause a hormetic effect in the organism. They are toxic at higher concentrations but at lower doses, they are essential for biological processes, such as cell migration and proliferation.

Among all these species, H₂O₂ is the most stable molecule given its lack of unpaired electrons. Its concentration levels are controlled by several enzymatic mechanisms (SOD, PGX, CAT) that balance cellular redox homeostasis. Additionally, spatiotemporal control is possible given its compartmentalized production, as enclosed generation creates the concentration gradients necessary for signaling, permitting at the same time its timely release. This is true whether the source is the plasma membrane [1], the ER [2], or the mitochondria [3], all of them notorious H₂O₂-producing organelles that pour into the cytosol to reach their targets [4].

1.2. Keratinocytes and skin homeostasis

The skin contributes to general organismal homeostasis by forming a barrier of tightly attached keratinocytes

that acts as the first line of defense against external potentially damaging factors [5]. As such, the multilayered epithelium that composes the tissue is constantly under renewal. Daily cues, such as aging or sunlight exposure, cause perturbations in the skin, compromising redox balance. Thus, the paradigm of ROS has its maximum exponent in the skin, as pathological excessive levels promote damage, leading to cell apoptosis and death, whereas physiological levels allow for correct triggering of the pathways that successfully regenerate the tissue [6].

1.3. Wound healing and epithelial cell migration

Wound healing is a complex and very highly regulated mechanism [7, 8]. Skin wounds can be classified as epidermal or deep, depending on whether they affect only keratinocytes or also the dermis underneath. In both cases, regeneration of the epithelium is needed. To do so, progenitor cells situated in the interphase of the dermal and the epidermal layers proliferate and migrate towards the wound until tissue integrity is restored [9]. H₂O₂ has a primordial role as a second messenger during the process, as in several animal models [10, 11], epithelial cells emit this molecule resulting in the attraction of neutrophils to the wound area.

In addition, experimental work carried out in xenopus [12], zebrafish [13] and axolotl [14] also reports the need of H₂O₂ to achieve complete tissue regeneration. Although yet no experimental evidence has been achieved in humans, some reports have foreseen that these findings could also be extrapolated to our species [15, 16, 17].

3. Materials and Methods

3.1. Cells and culture conditions

This project was performed using an immortalized cell line (HaCaT) [18], cells stably expressing the H₂O₂-specific ratiometric HyPer7 probe [19] in the cytosol (*HaCaT Cyto*), already generated by the host lab.

Cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) commercially calcium depleted (Gibco; Thermo Fisher) and supplemented with 0.09mM CaCl₂, 2mM of GlutaMAX™-I (Gibco; Thermo Fisher), 100 U/ml of penicillin and 100 µg/ml of streptomycin (Pan Biotech). To avoid contamination-derived oxidative stress in the cell culture, also 2.5 µg/ml of the plasmocin™ prophylactic was added to the cultures. Cell density, a parameter promoting undesired cell differentiation, was tightly controlled at all times.

3.2. HaCaT migration and proliferation over time

The cells at the densities required in each experiment were seeded 48 hours before image acquisition in μ -dishes, 35 mm, low culture insert two-well plate (Ibidi). These plates have two chambers separated by a silicone dam. Removal of the dam generates a gap that mimics a mechanical wound.

3.3. Microscope settings and data processing

Live imaging was performed using a wide-field Leica Dmi8 microscope equipped with an ORCA-Flash4.0 LT3 digital CMOS camera (Hamamatsu) and a CO₂/O₂ manual mixer (Okolab). 10x confocal images were acquired every 15 minutes for 63.5 hours using consecutive excitation at two wavelengths, 488 nm and 405 nm produced by led lights at 40% intensity for 800 and 400 ms of exposition, respectively. Another set of images was taken using a phase contrast channel. The CO₂ was set to 5% and the temperature to 37°C inside the plate adaptor chamber. For the analysis of HyPer7 fluorescence profile and the wound area closure evolution over time, the open source imaging software Fiji was used. The experiments were performed at least three times to record biological variations. However, in parallel to the figures represented with error bars, additional figures have been inserted in this project showing the mean tendency line to allow better visualization.

3.4. Monitoring of cytosolic H₂O₂ levels: HyPer7 ratio

The probe used in this project presents a ratiometric nature. Its emission intensity using two different excitation wavelengths (488 and 405 nm) varies depending on the redox state of the probe and can be numerically determined by the computation of their mean intensity fluorescent ratio (488/405, HyPer7 ratio). To measure it, two channels were used in the image acquisition. The average of at least 10 cells at the front and the rear of the wound was computed and represented for each single experiment. Final graphs averaging different experiments performed during the same culture and supplementation conditions were produced using the Prism software.

3.5. Wound healing closure

The gap area was assessed by manually drawing a polygon enclosing the gap between the two fronts in the images corresponding to lapses of four hours using Fiji. The results were exported and averaged using the program Prism.

4. Results and discussion

4.1. Characterization of redox responses in HaCaT cells during mechanical injury

The aim of this experiment is to find a correlation between the migration and proliferation processes in 2D in HaCaT cells with the fluorescent probe and the ac-

tivation of intracellular redox influxes. As it can be observed in Figure 1, left panel, right after the removal of the dam, the HyPer7 ratio increases both at the front and rear cells, and then decreases during all the evolution of the time lapse recorded. There is a difference in the oxidation state of both populations from the initial oxidation peak until a second oxidation wave that occurs after 50 hours of monitoring. Remarkably, this second wave is recorded only for cells populating the rear populations.

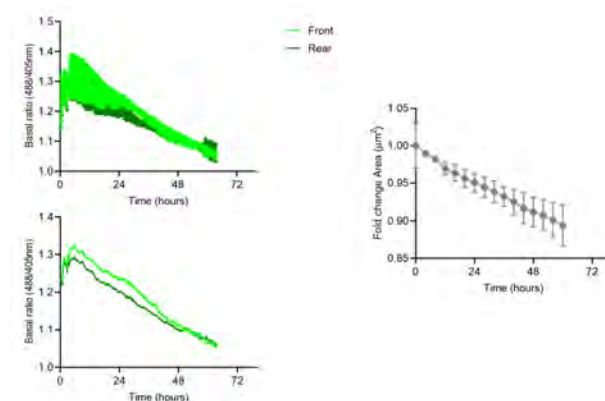


Figure 1. On the left panel, time evolution of HyPer7 basal ratio in HyPer Cyto cells, with and without SEM. On the right, the corresponding gap area reduction with SEM.

Regarding gap closure, there is a continuous decrease in the gap area, but the two fronts never met. (Figure 1, right panel). During the first 12 hours, and coincident with the maximum peak of HyPer7 ratio increase (Figure 1, left panel), cell closure is impressively homogeneous across the different experiments performed.

One possibility to be considered is that closure at the beginning of wounding is more dependent on proliferation than on migration. The data revealed by HyPer7 (Figure 1, left panel) would then be in accordance, as signaling downstream growth factor tyrosine kinase receptors depends on crosstalk with H₂O₂ generated by plasma membrane NADPH oxidases to be successful. Thus, an increase in the ratio of a cytosolic HyPer7 will be expected to increase during proliferation.

To test this possibility, the same experiments were performed after a 4-hour-long starving of the cells to synchronize cell response and achieve a boost in the effect. The former medium was reintroduced 30 minutes before starting the acquisition.

As it is shown in Figure 2 (left panel), intracellular basal ratio was higher under starving than under control conditions. Nevertheless, the initial rise in the redox state recorded previously (Figure 1, left panel) is no longer present, suggesting that the sudden reintroduction of growth factors induced a rapid redox wave of increased amplitude that was not registered due to the time employed on preparing the plates for microscope acquisition. Having different basal ratios among the cell populations from the beginning of the acquisition supports

this hypothesis.

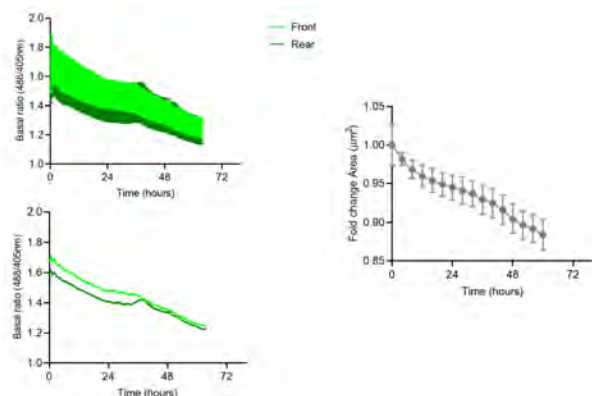


Figure 2. On the left panel, time evolution of HyPer7 basal ratio in HyPer Cyto cells after a 4-hour-long starvation treatment, with and without SEM. On the right, the corresponding gap area reduction with SEM.

A less likely possibility would be that the starving treatment is sensed as oxidative stress, thus increasing ROS levels. To control whether this is the case, a similar live imaging experiment was performed. This time, control and starved cells were plated in p35 plates at the same density to register their initial basal redox states in the absence of a gap to be filled. The results (Figure 3) reflected that the cytosolic HyPer7 ratio is similar for both culture conditions when the cells were stabilized by contact inhibition, thus confirming that a redox response is activated when cells perceive a wound margin, and that this response is dependent on growth factor activation.

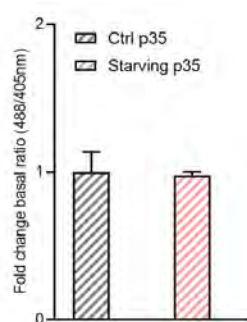


Figure 3. Fold change of the basal ratio (488/405 nm) of the HyPer7 probe in HaCaT Cyto seeded in p35 plates under control and 4-hour-long starving conditions, with SEM.

Additionally, in Figure 2, the second oxidation wave is more visible than in control conditions. At the same time, an acceleration in the gap area reduction (Figure 2, right) can be observed, which suggests that the second oxidation wave is mainly driven by a migratory signal.

4.2. Redox dependency of HaCaT migration

In all the experiments discussed so far, there is an incomplete wound closure even after 63.5 hours of continuous monitoring, which was not expected from a keratinocyte cell line. Based on the former results that indicate that migration is a redox process, and due to the working mechanism of the probe that consumes the signal as it reports its presence, a new migration experiment in cells that do not express the probe was performed to test whether closure was dampened by signal consumption. Thus, HaCaT wild type cells (*HaCaT empty*) were seeded in Ibidi plates to conduct an experiment in which H_2O_2 molecules are completely available for endogenous signaling routes. Cell culture conditions were the same as in control cells. Indeed, the absence of the probe restored the capacity of wound closure, so that the cells were able to almost close the gap by the end of the experiment (Figure 4). This result indicates that the increased sensitivity of the HyPer7 probe seriously compromises migration by eliminating fundamental signaling molecules that react with the biosensor instead of with their biological targets, a result of outmost importance in the redox field.

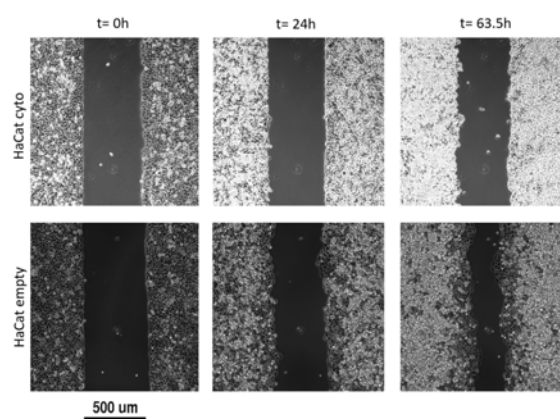


Figure 4. Snapshots of HaCaT cells at different time points ($t=0h$, $t=24h$ and $t=63.5h$) with the probe (top row) and wild type cells (bottom row). The wound area reduction is more evident in HaCaT cells without the probe.

Acknowledgements

This project is the result of previous work carried out by exceptional scientists in the redox field and has been developed in the UC3M facility core CleanRooms for Bioengineering. Financial support was provided by the Madrid Government (Comunidad de Madrid) under the Multiannual Agreement with UC3M in the line of "Research Funds for Beatriz Galindo Fellowships" (REDOXSKIN-CM-UC3M), and in the context of the V PRICIT (Regional Programme of Research and Technological Innovation", and from "Proyectos de I + D + I" (PID2020-114230 GA-I00) funded by MCIN/AEI/10.13039/501100011033/.

References

- [1] F. Magnani and A. Mattevi. Structure and mechanisms of ROS generation by NADPH oxidases. *Current Opinion in Structural Biology*, 59:91–97, 2019.
- [2] B. P. Tu, S. C. Ho-Schleyer, K. J. Travers, and J. S. Weissman. Biochemical basis of oxidative protein folding in the endoplasmic reticulum. *Science (New York, N. Y.)*, 290(5496):1571–1574, 2000.
- [3] Chang Li and Hai-Meng Zhou. The role of manganese superoxide dismutase in inflammation defense. *Enzyme Research*, 2011:e387176, 2011.
- [4] A. Glasauer and Navdeep S. Chandel. ROS. *Current Biology*, 23(3):R100–R102, 2013.
- [5] Lieveke Hellemans, Hugo Corstjens, Ann Neven, Lieve Declercq, and Daniel Maes. Antioxidant enzyme activity in human stratum corneum shows seasonal variation with an age-dependent recovery. *Journal of investigative dermatology*, 120(3):434–439, 2003.
- [6] Christopher Dunnill, Thomas Patton, James Brennan, John Barrett, Matthew Dryden, Jonathan Cooke, David Leaper, and Nikolaos T Georgopoulos. Reactive oxygen species (ros) and wound healing: the functional role of ros and emerging ros-modulating technologies for augmentation of the healing process. *International wound journal*, 14(1):89–96, 2017.
- [7] C. Qing. The molecular biology in wound healing & non-healing wound. *Chinese Journal of Traumatology*, 20(4):189–193, 2017.
- [8] Paulina Krzyszczyk, Rene Schloss, Andre Palmer, and François Berthiaume. The role of macrophages in acute and chronic wound healing and interventions to promote pro-wound healing phenotypes. *Frontiers in Physiology*, 9, 2018.
- [9] Alisha M. Mendonsa, Na Tae-Young, and Barry M. Gumbiner. E-cadherin in contact inhibition and cancer. 37(35):4769–4780, 2018. Num Pages: 4769-4780 Place: New York, United States Publisher: Nature Publishing Group.
- [10] Severina Moreira, Brian Stramer, Iwan Evans, Will Wood, and Paul Martin. Prioritization of competing damage and developmental signals by migrating macrophages in the drosophila embryo. *Current biology*, 20(5):464–470, 2010.
- [11] Philipp Niethammer, Clemens Grabher, A. Thomas Look, and Timothy J. Mitchison. A tissue-scale gradient of hydrogen peroxide mediates rapid wound detection in zebrafish. *Nature*, 459(7249):996–999. Number: 7249 Publisher: Nature Publishing Group.
- [12] Nick R Love, Yaoyao Chen, Shoko Ishibashi, Paraskevi Kritsiligkou, Robert Lea, Yvette Koh, Jennifer L Gallop, Karel Dorey, and Enrique Amaya. Amputation-induced reactive oxygen species are required for successful xenopus tadpole tail regeneration. *Nature cell biology*, 15(2):222–228, 2013.
- [13] Sa Kan Yoo, Christina M. Freisinger, Danny C. LeBert, and Anna Huttenlocher. Early redox, src family kinase, and calcium signaling integrates wound responses and tissue regeneration in zebrafish. *The Journal of general physiology*, 140(5):i7–i7, 2012.
- [14] Belfran Carbonell M, Juliana Zapata Cardona, and Jean Paul Delgado. Hydrogen peroxide is necessary during tail regeneration in juvenile axolotl. *Developmental dynamics*, 251(6):1054–1076, 2022.
- [15] Sandra Chang, Angela Linderholm, Lisa Franzi, Nicholas Kenyon, Helmut Grasberger, and Richart Harper. Dual oxidase regulates neutrophil recruitment in allergic airways. *Free radical biology medicine*, 65:38–46, 2013.
- [16] Alvin Eng Kiat Loo, Yee Ting Wong, Rongjian Ho, Martin Wasser, Tiehua Du, Wee Thong Ng, and Barry Halliwell. Effects of hydrogen peroxide on wound healing in mice in relation to oxidative damage. *PloS one*, 7(11):e49215–e49215, 2012.
- [17] Thomas S Lisse, Benjamin L King, and Sandra Rieger. Comparative transcriptomic profiling of hydrogen peroxide signaling networks in zebrafish and human keratinocytes: Implications toward conservation, migration and wound healing. *Scientific reports*, 6(1):20328–20328, 2016.
- [18] P. Boukamp, R. T. Petrussevska, D. Breitkreutz, J. Hornung, A. Markham, and N. E. Fusenig. Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line. 106(3):761–771.
- [19] Valeriy V. et al. Pak. Ultrasensitive genetically encoded indicator for hydrogen peroxide identifies roles for the oxidant in cell migration and mitochondrial function. *Cell metabolism*, 31(3):642–653.e6, 2020.

Nuevos modelos multicelulares para la predicción de la hepatotoxicidad idiosincrática

N. Pons Andrés^{1,2}, E. Villanueva Bádenas¹, G. Gallego Ferrer², L. Tolosa Pardo¹

¹ Unidad de Hepatología Experimental, Instituto de Investigación Sanitaria La Fe, 46026 Valencia, España

² Centro de Biomateriales e Ingeniería Tisular, Universidad Politécnica de Valencia, 46022 Valencia, España

Resumen

La hepatotoxicidad inducida por fármacos (DILI) se refiere al daño hepático resultante de la exposición del organismo a un medicamento o xenobiótico. Concretamente la de tipo idiosincrático (iDILI) es la principal causa de retirada de fármacos tanto en fases de desarrollo como ya comercializados.

Los estudios preclínicos con animales de experimentación presentan una baja predictividad, por ello, ha sido necesaria la búsqueda de nuevos modelos in vitro que permitan una mejor comprensión de los mecanismos que subyacen al iDILI para reducir su impacto. La mayoría de los modelos clásicos propuestos están basados en el cultivo de hepatocitos, pero no permiten modelar sus interacciones con las células no parenquimáticas como las células de Kupffer, que han demostrado estar implicadas en la hepatotoxicidad mediada por el sistema inmunitario.

En el presente trabajo se ha optimizado y caracterizado un modelo de co-cultivo de hepatocitos y macrófagos proinflamatorios, derivados de líneas celulares inmortalizadas (HepG2 y THP-1, respectivamente) para el estudio de la hepatotoxicidad idiosincrática. El modelo ha mostrado un adecuado mantenimiento de la viabilidad celular, expresión génica y funcionalidad hepática.

Asimismo, la idoneidad del co-cultivo como plataforma de detección de fármacos que inducen iDILI ha sido evaluada a través de su exposición a dos medicamentos asociados con este, el trovafloxacino y la troglitazona. Los resultados obtenidos han probado su mayor sensibilidad y activación de las rutas metabólicas que conducen al iDILI, frente a los sistemas de monocultivo utilizados tradicionalmente.

1. Introducción

El hígado es un órgano esencial, que realiza o controla una gran cantidad de funciones metabólicas. El 80% de las células presentes en el hígado son hepatocitos, que llevan a cabo las principales funciones hepáticas, mientras que el 20% restante son las células no parenquimáticas (CNP), entre las cuales destacan las células de Kupffer (KC), macrófagos residentes en el hígado que son activados a un estado proinflamatorio tras el daño hepático [1].

Una de las principales dificultades a las que se enfrentan tanto la salud pública como las agencias reguladoras y las empresas farmacéuticas es la predicción de las reacciones adversas a medicamentos durante el desarrollo de fármacos, y previamente a que éstos sean comercializados [2]. A pesar de que estas reacciones pueden afectar a diferentes órganos, el hígado es de los más recurrentes, recibiendo el nombre de daño hepático inducido por fármacos (DILI) y ha sido designado como la causa

principal de trasplantes e insuficiencia hepática aguda [3]. En particular, el DILI idiosincrático (iDILI) sucede únicamente en una minoría de pacientes, pero es impredecible y multifactorial, siendo así su complejo diagnóstico y predicción un obstáculo importante para el desarrollo farmacéutico [2].

Tradicionalmente, se han usado modelos animales para estudiar el DILI, pero debido a las diferencias en las rutas metabólicas entre especies, han mostrado ser poco predictivos. Debido a ello y a las crecientes iniciativas por reducir la experimentación animal, se han desarrollado diferentes sistemas de cultivo celular que permiten mimetizar las condiciones *in vivo* y servir de modelos para el estudio de la hepatotoxicidad provocada por los medicamentos en humanos [2,4]. Los modelos más simples son los monocultivos, siendo el de hepatocitos humanos primarios (PHH) el más estudiado [4]. Sin embargo, debido a su escasa disponibilidad, la desdiferenciación rápida que sufren y su corta supervivencia, ha sido indispensable la búsqueda de fuentes celulares alternativas [4,5]. En este sentido, se pueden obtener hepatocitos derivados de líneas celulares de hepatoma capaces de proliferar de manera indefinida bajo las condiciones adecuadas [5]. Aunque presentan una alta disponibilidad, fácil manejo y fenotipo estable durante largos periodos de tiempo; su genotipo suele ser inestable y expresan menos funciones metabólicas que los PHH [4,5]. No obstante, los monocultivos no mimetizan las interacciones dinámicas y complejas entre los hepatocitos y las CNP. Por ello, estudios recientes se han basado en el uso de co-cultivos con CNP, principalmente KC, para estudiar el papel de la inflamación. De hecho, se ha demostrado una mayor competencia metabólica y estabilidad de éstos frente a los monocultivos [6].

Debido a la ausencia de modelos apropiados capaces de predecir eficientemente el iDILI, surge la necesidad de desarrollar e investigar un nuevo sistema que reproduzca la fisiología del hígado y que pueda servir de fundamento para el desarrollo de futuros modelos más complejos. Por ello, el presente estudio ha tenido como objetivo principal el desarrollo y caracterización de un modelo de estudio del iDILI, haciendo uso del co-cultivo de células de la línea celular de hepatoma humano HepG2, que mimetizan los hepatocitos, en contacto directo con macrófagos diferenciados a partir de la línea celular monocítica THP-1 y activados al estado proinflamatorio (M1) para mimetizar las KC. Los objetivos concretos han sido: (1) optimizar y evaluar el mantenimiento del co-cultivo,

utilizando diferentes estímulos proinflamatorios, (2) evaluar la funcionalidad de los co-cultivos que mimeticen una situación fisiológica o inflamatoria, y (3) comprobar la idoneidad del modelo *in vitro* desarrollado para estudiar el iDILI.

2. Materiales y métodos

2.1. Materiales

En el proceso de realización de los cultivos celulares se utilizaron medios de cultivo y complementos adquiridos mayoritariamente de GIBCO (Gibco BRL, Paisley, UK) y Sigma Aldrich (Madrid, España). El suero bobino fetal (SBF) fue obtenido de Capricorn Scientific GmbH (Ebsdorfergrund, Alemania) y las citoquinas se adquirieron en Peprotech (Londres, UK).

2.2. Cultivo celular

Cultivo HepG2

La línea celular HepG2 (ATCC HB-8065) fue cultivada en medio de cultivo Ham's F12/Leibovitz L-15 (1:1 v/v), suplementado con 7% SBF, L-Glutamina (2,5 mM), Glucosa (5 mM), BSA (2 mM), Bicarbonato (12 mM), 10000 U penicilina/mL de medio y 10000 µg estreptomina/mL de medio. Las células fueron cultivadas a la densidad de 47000 células/cm² como monocultivo de control.

Cultivo THP-1

Los monocitos de la línea celular THP-1 (ATCC TIB-202) fueron cultivados con el medio comercial RPMI Glutamax suplementado con un 10% de SBF, 10000 U penicilina/mL y 10000 µg estreptomina/mL en medio.

Las células THP-1 se diferenciaron a macrófagos inactivos (M0) a través de su incubación con forbol-12-miristato-13-acetato (PMA, 20 ng/mL) durante 72 h, caracterizados por una morfología estrellada y su adherencia a la placa de cultivo. A continuación, fueron diferenciados a macrófagos proinflamatorios (M1) mediante su incubación durante 24 h con lipopolisacárido (LPS, 1 µg/mL) e interferón (IFN-γ, 20 ng/mL). Se estudió el mantenimiento del fenotipo bajo diferentes condiciones proinflamatorias durante 48h.

Co-cultivo de HepG2 y Macrófagos M1-THP-1

Los co-cultivos se efectuaron empleando células de tipo HepG2 y macrófagos M1 derivados de las células THP-1. Se utilizaron diferentes ratios de HepG2 y THP-1-M1 (1:10, 1:5 y 1:2 (M1:HepG2)) para mimetizar un estado fisiológico del hígado, un estado de transición intermedio y el estado de inflamación, respectivamente. Las células se sembraron a la densidad de 47000 células/cm² y el medio utilizado consistió en la mezcla 1:1 v/v de los medios de ambos tipos celulares, manteniendo el estímulo pro-inflamatorio en el medio de M1 a través de la adición de LPS y IFN-γ. Asimismo, se realizaron dos monocultivos de control de HepG2 y THP-1-M1 empleando este mismo medio.

2.3. Caracterización del modelo celular

Inmunofluorescencia

Las células fueron fijadas con paraformaldehído (PFA) al 4% durante 20 min y se lavaron 3 veces con PBS 1X. A continuación, las muestras fueron permeabilizadas con Tritón X-100 al 0,5% en PBS durante 20 min a temperatura ambiente (RT), y bloqueadas en PBS/3% BSA durante 1 h a RT. Seguidamente, se incubaron durante toda la noche con los anticuerpos primarios (anti-albúmina, CD68 y CD163 (1:200)). Al día siguiente, se lavaron 3 veces con tampón de lavado (PBS/1% BSA-0.1% Tween) y se incubaron con los anticuerpos secundarios durante 1h a RT. Finalmente, se lavaron una vez con PBS/1% BSA con Hoechst (1 µg/mL) para marcar los núcleos y dos veces con tampón de lavado. Las imágenes se adquirieron con el microscopio de fluorescencia Leica HMR (DMI8, Leica Microsystems, Barcelona, España).

Análisis transcriptómico

Se realizó la extracción del RNA de las muestras utilizando el Kit RNeasy Plus Mini (Qiagen, Madrid, España), y fue retrotranscrito empleando el protocolo elaborado por el laboratorio [7]. El cDNA diluido fue amplificado empleando el LightCycler DNA Master SYBR Green I (Roche Applied Science, Barcelona, España) y los primers correspondientes fueron amplificados haciendo uso del LightCycler 480 Instrument (Roche Applied Science, Barcelona, España). Los genes codificantes de la porfobilinógeno deaminasa (*PBGD*), la proteína *TBP*, la proteína de la subunidad subribosómica UL10 (*RPLP0*) y actina fueron utilizados para normalizar la expresión. El análisis de los resultados de la PCR se llevó a cabo por medio del método delta delta Ct ($\Delta\Delta Ct$).

Análisis de la funcionalidad de los co-cultivos

Se realizó la cuantificación de la albúmina producida por los co-cultivos tras 24 h de cultivo, empleando un ensayo comercial ELISA (Bethyl Laboratories, Montgomery, TX, USA). La capacidad ureogénica de los cultivos se midió por medio del kit de urea de *Quantichrom Urea Assay Kit* (BioAssays, Hayward, CA, USA).

2.4. Estudios de toxicidad

Las células HepG2 y los macrófagos M1 fueron sembradas solas o de manera combinada en placas de 96 pocillos (15000 células/pocillo). Pasadas 24 h, los co-cultivos o monocultivos se incubaron con un rango de concentraciones de los fármacos modelo seleccionados: el trovafloxacino (TVX) y la troglitazona (TGZ), asociados con el iDILI; y sus análogos estructurales no hepatotóxicos, el levofloxacino (LVX) y la rosiglitazona (RGZ). La viabilidad celular se determinó con MTT [6].

Además, se analizó la expresión génica de las muestras tratadas con los fármacos tal y como se indica en la *Sección 2.3*.

2.5. Análisis estadístico

Los datos obtenidos fueron estudiados e interpretados por medio del análisis ANOVA, considerando como cierta la hipótesis de homocedasticidad. El intervalo de confianza fue del 95% y se llevó a cabo utilizando el Software GraphPad Prism 8.0.2. Se realizó la prueba de Tukey como post-test, que permite identificar los subconjuntos de datos con diferencias estadísticamente significativas.

3. Resultados

3.1. Caracterización del proceso de diferenciación a macrófagos M1

Con el propósito de establecer las condiciones apropiadas para el mantenimiento de las células THP-1 diferenciadas a macrófagos proinflamatorios (M1) durante el tiempo de experimentación, se estudió la expresión de los marcadores génicos del fenotipo M1 tras la estimulación con la endotoxina LPS y la citoquina IFN- γ , así como los efectos de eliminar dichos estímulos pasadas 48 h desde su activación a M1.

La eliminación del LPS e IFN- γ resultó en una disminución significativa en la expresión de los marcadores proinflamatorios *TNF- α* y *CXCL10* (Fig.1), mientras que, si se mantenía dicho estímulo, estos mismos marcadores junto con el *IL-1 β* aumentaban su expresión. Asimismo, el gen *CD206*, marcador del fenotipo inactivo M0, aumentó significativamente su expresión cuando el cultivo no fue estimulado con LPS ni IFN- γ , y apenas se expresó en caso contrario. Se estudió también el efecto de la estimulación de HepG2 y THP-1-M1, ambas en monocultivo, sobre la viabilidad celular, sin encontrar diferencias estadísticamente significativas en ningún caso.

De esta manera, se decidió mantener el estímulo conjunto de LPS e IFN- γ para evitar la desdiferenciación del fenotipo M1 de los macrófagos.

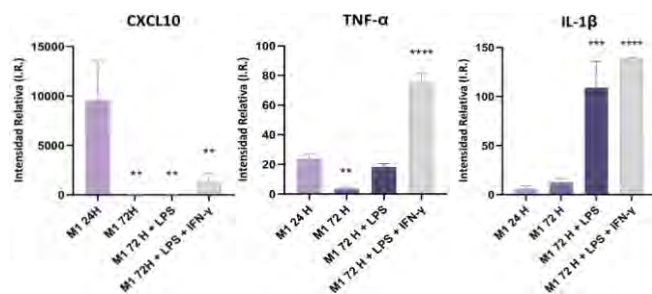


Figura 1. Expresión de los genes característicos *CXCL10*, *TNF- α* y *IL-1 β* en células THP-1-M1, pasadas 72 horas sin estimulación (M1 72 h), con estimulación de 1 μ g/mL de LPS (M1 72h + LPS) y con 1 μ g/mL de LPS y 20 ng/mL de IFN- γ (M1 72h + LPS + IFN- γ). Se representa la I.R. con respecto a su expresión en los macrófagos M0. Indicadas con * las diferencias estadísticamente significativas, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ y **** $p < 0.0001$ ($n=3$, Test ANOVA, post-test Tukey) respecto los macrófagos M1 tras 24 horas sin estímulos.

3.2. Co-cultivos HepG2 con células THP-1 diferenciadas a macrófagos M1

Los ensayos de viabilidad no mostraron diferencias estadísticamente significativas de los distintos ratios de co-cultivos tratados con LPS e IFN- γ con respecto a los no tratados. A través de la inmunofluorescencia se identificaron los marcadores propios de los macrófagos (CD68) y de los hepatocitos (albúmina) en todos los co-cultivos (Fig.2). Asimismo, la expresión génica tanto de marcadores hepáticos (*ALB*, *AFP*, *HNF-4*) como de macrófagos M1 (*CXCL10*, *IL-1 β* , *TNF- α*) fue favorable en todos ellos, aunque disminuyeron significativamente en relación con los monocultivos. Por otra parte, el citocromo *CYP3A5* y las enzimas *UGT1A1* y *GSTT1* incrementaron su expresión en los co-cultivos.

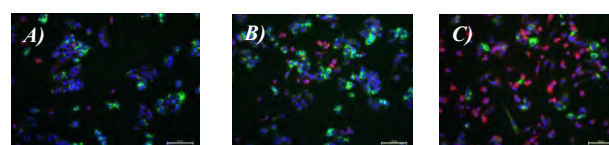


Figura 2. Imágenes de inmunofluorescencia de los co-cultivos (M1:HepG2): A) 1:10, B) 1:5, y C) 1:2. La albúmina se marca en color verde, y el marcador CD68 aparece en rojo y el color azul tiñe los núcleos celulares. La barra de escala: 100 μ m.

Adicionalmente, el estudio de la funcionalidad hepática en los co-cultivos mostró una disminución en la síntesis de albúmina, ya que esta función es específica de los hepatocitos; y un aumento de la producción de urea en los co-cultivos con mayor proporción de macrófagos M1.

3.3. Hepatotoxicidad inducida por fármacos

Para evaluar la idoneidad del modelo *in vitro* propuesto para el estudio del iDILI, los co-cultivos fueron tratados durante 24 h con los fármacos TVX y LVX, y sus análogos estructurales TGZ y RGZ, respectivamente.

El TVX y la TGZ provocaron muerte celular significativa a concentraciones de fármaco superiores a 50 μ M y 100 μ M, respectivamente (Fig.3). En ambos casos se observó una disminución de la viabilidad celular dependiente de la dosis de fármaco, mostrando así una posible actuación sinérgica del TVX y de la TGZ junto con los macrófagos proinflamatorios, ya que esta disminución fue estadísticamente más acusada en los co-cultivos, con mayor proporción de macrófagos proinflamatorios que en el monocultivo de HepG2. Este comportamiento no se produjo en los cultivos tratados con los fármacos LVX y RGZ, no asociados con el iDILI.

Finalmente, se analizó la expresión génica de los co-cultivos expuestos a los fármacos para explicar las diferentes sensibilidades en los co-cultivos para la detección de la toxicidad inducida por el TVX y la TGZ. El TVX aumentó significativamente la expresión del gen *ATF4*, del factor nuclear *NRF2* y de la enzima *GCLC* en todos los cultivos respecto a sus análogos no tratados, hecho que no sucedió con el LVX en ningún caso. Por su parte, la TGZ aumentó significativamente la expresión de los genes *CHOP*, *ATF6*, *NRF2*, *GCLC* y *GLCM* en todos los cultivos con respecto a los no tratados, pero

disminuyó la expresión de *CAT* y *SOD1*. La RGZ provocó una disminución en la expresión del gen *SOD1* y un aumento en *NRF2* y *GCLM*.

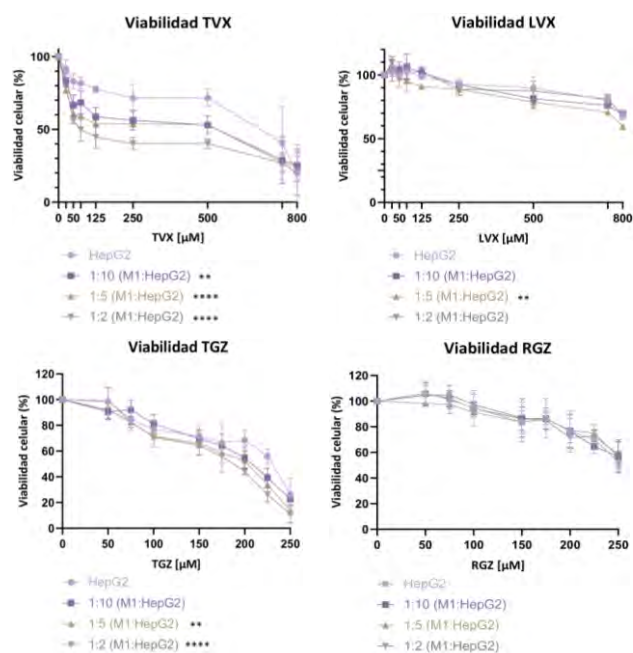


Figura 3. Toxicidad dependiente de la concentración de los fármacos testados en monocultivos de HepG2 y en co-cultivos de HepG2 y THP-1-M1. Se ha indicado con * las diferencias estadísticamente significativas entre la viabilidad celular en los co-cultivos con respecto al monocultivo de HepG2, con * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ y **** $p < 0.0001$ ($n = 3$, Test Two-Way ANOVA, post-test Tukey).

4. Discusión

La aparición de hepatotoxicidad idiosincrática inducida por ciertos medicamentos es un problema grave para el cual todavía no se ha hallado una solución efectiva. Por lo tanto, se necesitan nuevos modelos que permitan una mejor predicción y comprensión de los mecanismos subyacentes al iDILI para disminuir los riesgos en el proceso de desarrollo de fármacos.

En el presente trabajo hemos caracterizado completamente un sistema de co-cultivo directo de células HepG2, que poseen las principales características de los hepatocitos, y células monocíticas THP-1, cuya diferenciación a macrófagos proinflamatorios permite reemplazar a las KC primarias; para la detección de fármacos que producen iDILI inmunomediados.

Los resultados mostraron la obtención de co-cultivos bien establecidos, con funcionalidad hepática mantenida, y con expresión tanto de los genes y marcadores propios de los hepatocitos y los macrófagos, como de las principales enzimas presentes en el metabolismo hepatotóxico, proporcionando las interacciones celulares adecuadas para poder desarrollar el iDILI.

El estudio de la idoneidad del sistema para el estudio de la hepatotoxicidad inducida por fármacos *in vitro* mostró una posible actuación sinérgica del TVX y de la TGZ junto con los macrófagos M1 en los co-cultivos. Asimismo, el modelo permitió ver cambios en vías de

señalización implicadas en el iDILI, pero no explicó la mayor sensibilidad al fármaco de los co-cultivos con mayor proporción de macrófagos proinflamatorios. Para resolverlo, sería conveniente llevar a cabo otros estudios modificando parámetros como la concentración de fármaco a la cual se realiza el tratamiento de las células, o el tiempo de incubación. Finalmente, sería favorable el establecimiento de modelos de mayor complejidad, como esferoides, estructuras tridimensionales que mimetizan mejor el entorno hepático.

5. Conclusiones

El modelo *in vitro* propuesto para la evaluación de la hepatotoxicidad ha sido caracterizado y evaluado, mostrando su correcto mantenimiento, funcionalidad y la predicción de la hepatotoxicidad inducida por dos fármacos asociados con el iDILI, el TVX y la TGZ. La incorporación del componente inflamatorio proporciona una mayor sensibilidad en las evaluaciones de hepatotoxicidad *in vitro*, siendo la base para el desarrollo de sistemas más complejos para el estudio del iDILI.

Agradecimientos

A los comités organizadores del CASEIB por proporcionar las guías para la realización del presente documento.

Referencias

- [1] Godoy, P., ... & Hengstler, J. G. Recent advances in 2D and 3D in vitro systems using primary hepatocytes, alternative hepatocyte sources and non-parenchymal liver cells and their use in investigating mechanisms of hepatotoxicity, cell signaling and ADME. *Archives of toxicology*, vol 87, sup 8, 2013, pp 1315-1530.
- [2] Shaw, P. J., Ganey, P. E., & Roth, R. A. Idiosyncratic drug-induced liver injury and the role of inflammatory stress with an emphasis on an animal model of trovafloxacin hepatotoxicity. *Toxicological Sciences*, vol 118, sup 1, 2010, pp 7-18.
- [3] Ostapowicz, G., ... & US Acute Liver Failure Study Group*. Results of a prospective study of acute liver failure at 17 tertiary care centers in the United States. *Annals of internal medicine*, vol 137, sup 12, 2002, pp 947-954.
- [4] Guillouzo, A., & Guguen-Guillouzo, C. Evolving concepts in liver tissue modeling and implications for in vitro toxicology. *Expert opinion on drug metabolism & toxicology*, vol 4, sup 10, 2008, pp 1279-1294.
- [5] Mathieu Vinken, Vera Rogiers (eds.) Protocols in In Vitro Hepatocyte Research. Humana Press, 2015 (ISBN: 978-1-4939-2073-0).
- [6] Rose, K. A., Holman, N. S., Green, A. M., Andersen, M. E., & LeCluyse, E. L. Co-culture of hepatocytes and Kupffer cells as an in vitro model of inflammation and drug-induced hepatotoxicity. *Journal of pharmaceutical sciences*, vol 105, sup 2, 2016, pp 950-964.
- [7] Moya, M., ... & Jover, R. Foxal reduces lipid accumulation in human hepatocytes and is down-regulated in nonalcoholic fatty liver. *PLoS ONE*, vol 7, sup 1, 2012.

Estudio numérico para consolidación mediante fijación de fracturas verticales en cuello de fémur

E. Agulló Mertens, D. Bru Orts¹

¹ Dpto. de Ingeniería Civil, Universidad de Alicante, San Vicente del Raspeig, España, david.bru@ua.es

Resumen

Las fracturas del cuello femoral (FNF) son un problema creciente en ortopedia y traumatología, afectando a personas de todas las edades y géneros. El sistema de clasificación Pauwels divide estas fracturas en tres tipos según la verticalidad de ella, siendo la de tipo III (ángulo >50°) las más difíciles de tratar. La elección del tratamiento o sistema de fijación, sigue siendo un desafío debido a la falta de guías estandarizadas.

Este estudio se enfoca en comparar tres tipologías de sistemas de fijación interna, CS y XCS tipo 1 y tipo 2, a partir de 24 modelos numéricos, evaluando su efecto en la rigidez del sistema, así como de las tensiones tanto en el sistema óseo como en el propio sistema de fijación. El reducido número de investigaciones en esta línea de trabajo resalta la importancia de esta investigación para mejorar el conocimiento, el tratamiento de las FNF y la calidad de vida de los pacientes afectados.

1. Introducción

El fémur es el hueso más largo del cuerpo humano y desempeña un papel crucial en la función de la pierna. Las fracturas en el cuello femoral, especialmente en personas jóvenes debido a lesiones traumáticas, pueden ser complicadas y afectar su función y movilidad.

Las fracturas en el cuello femoral están sometidas a tensiones de compresión, tracción y torsión durante la locomoción y estas tensiones varían según la carga y la posición de la cadera [1]. La compresión adecuada es esencial para abordar estas fracturas de manera efectiva. En el caso de las fracturas intracapsulares, el flujo sanguíneo hacia la cabeza femoral puede interrumpirse, lo que puede llevar a complicaciones graves. La rápida reducción y fijación son clave para preservar la vascularización [2][3].

El proceso de consolidación ósea se divide en consolidación primaria cuando la deformación interfragmentaria es menor del 2% y en consolidación secundaria cuando la deformación interfragmentaria es menor del 10% [4], dependiendo de la inmovilización de la fractura. La elección del método de tratamiento influye en el tipo de consolidación que se logra. La posición y orientación de los tornillos utilizados en el tratamiento son cruciales y pueden variar según el caso.

Para las fracturas específicas, como las de Pauwels tipo III, se suelen utilizar tornillos canulado en la configuración de triángulo invertido (CS) debido a su menor invasividad. Sin embargo, esta configuración puede tener limitaciones en términos de fuerza y rigidez. Se han propuesto enfoques alternativos como tornillos fuera de eje (XCS) [6], [7] y la incorporación de alambre de cerclaje al CS [8], pero los resultados hasta ahora han sido inconsistentes.

En base a lo anterior, el objetivo de esta investigación se centra en analizar numéricamente el comportamiento mecánico de tres tipologías de fijación interna FNF (CS, XCS tipo 1 y XCS tipo 2) desde la perspectiva de las deformaciones en el plano de la fractura y las tensiones en el sistema de fijación y óseo, a fin de evaluar la efectividad de las mismas de cara a la consolidación de la fractura. Para ello se ha evaluado la influencia del diámetro de los tornillos, así como de la disposición de estos respecto a la fisura vertical (Pauwels III).

2. Métodos y materiales

En esta investigación, ha llevado a cabo la simulación numérica mediante el software de elementos finitos Ansys de una campaña experimental tomando como base de la metodología experimental los procedimientos seguidos por Kuan. et al. [10].

2.1. Casos de estudio

Se han llevado a cabo un total de 24 casos de estudio, agrupados en tres grandes grupos según el tipo de fijación: a) CS; b) XCS tipo 1 y c) XCS tipo 2, Figura 1. Para evaluar la eficacia de dichos sistemas, dentro de cada grupo, se ha variado la posición o separación entre tornillos, así como el tipo de material.

El primer grupo de estudio, basado en la fijación CS, se ha analizado utilizando solo tornillos canulados en una disposición de triángulo invertido, con tornillos de titanio y acero inoxidable, modelizándose un total de 8 casos. En concreto se han utilizado 3 tornillos canulados paralelos de 80 mm de longitud, 6.5 mm de diámetro, 16 mm de longitud de rosca e inclinación de 120°, según cuatro disposiciones diferentes dentro del plano de rotura, Figura 2. De esta forma en el grupo 1 de análisis se han evaluado 8 casos de estudio, siendo los modelos 1 a 4 de titanio y los modelos 5 a 8 de acero.

El segundo grupo, XCS tipo 1, se ha analizado utilizando dos tornillos canulados de iguales características a los usados en el grupo 1, y un tornillo cortical de 58 mm de longitud y 4.5 mm de diámetro, también en acero o titanio. De esta forma, para el grupo 2 se han modelizado también 8 casos de estudio, siendo los modelos 9 a 12 de titanio y los modelos 13 a 16 de acero inoxidable, Figura 3. Debe destacarse que, en este grupo, el tornillo cortical se ha dispuesto horizontalmente y mantiene su ubicación en todos los casos, a diferencia de los tornillos canulados.

El tercer grupo, XCS tipo 2, es similar al segundo, pero variando la posición del tornillo cortical, el cual está

dispuesto en una posición más distal. En la Figura 4 se resumen las cuatro posiciones utilizadas en este grupo, siendo los modelos 17 a 20 de titanio, y los modelos 21 a 24 de acero.

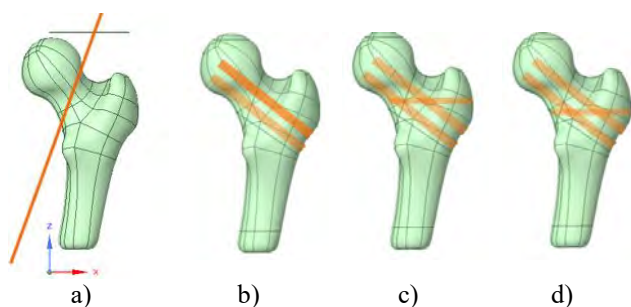


Figura 1. Fractura de Pauwels tipo III y Sistemas de fijación: a) Plano de fractura; b) CS; c) XCS tipo 1; d) XCS tipo 2.

Un detalle de las diferentes configuraciones de los sistemas de fijación se muestra en las Figuras 2, 3 y 4.

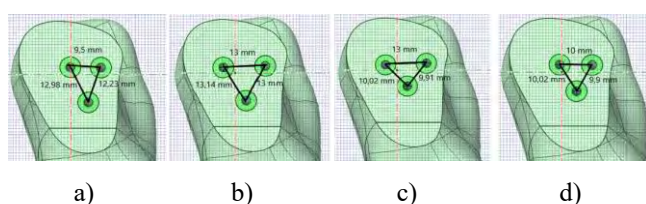


Figura 2. Sistema CS: a) Modelo 1 y 5; b) Modelo 2 y 6; c) Modelo 3 y 7; d) Modelo 4 y 8.

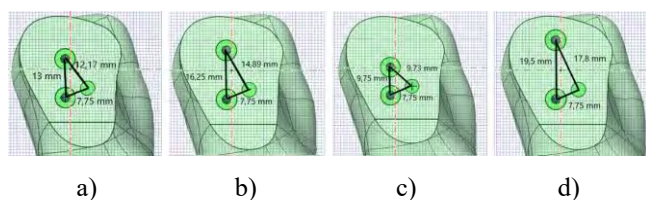


Figura 3. Sistema XCS tipo 1: a) Modelo 9 y 13; b) Modelo 10 y 14; c) Modelo 11 y 15; d) Modelo 12 y 16.

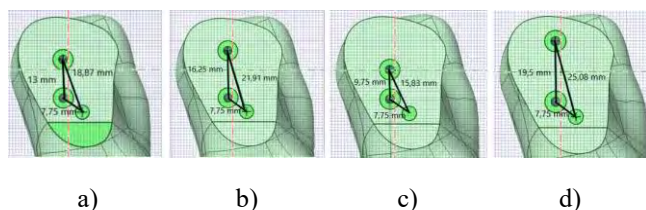


Figura 4. Sistema XCS tipo 2: a) Modelo 17 y 21; b) Modelo 18 y 22; c) Modelo 19 y 23; d) Modelo 20 y 24.

2.2. Materiales

Los materiales modelizados en esta investigación son: a) Hueso fémur Sawbone; b) Tornillos canulados; c) Tornillos corticales autorroscantes, siendo los dos tipos de tornillos evaluados tanto en titanio como en acero inoxidable. En concreto, el modelo de hueso fémur Sawbone de cuarta generación y tamaño medio, cuenta con un canal intramedular de 13 mm (Pacific Research Laboratories, Vashoon Island WA) y está constituido por espuma de poliuretano rígida con densidad 0,27 g/cm³ para la componente trabecular y resina epoxi reforzada con fibra de vidrio corta con densidad 1,64 g/cm³ [11].

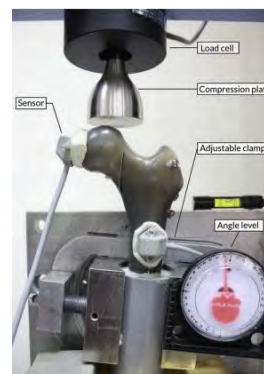


Figura 5. Diseño experimental [10]

2.3. Modelización numérica

La modelización numérica presentada en esta investigación toma como datos de partida la configuración del ensayo establecida por Kuan et al. [10], Figura 5. Dicha configuración se basa en un ensayo a compresión con un rango de aplicación de la carga entre 200 N y 6000 N a fin de evaluar el comportamiento a bajos y altos niveles de carga. A partir de dichos resultados, en esta investigación se han determinado las curvas carga-desplazamiento y se ha calculado la rigidez axial y la carga última de rotura en base a los datos facilitados por el fabricante.

Para la modelización numérica se han utilizado los programas Ansys Mechanical y SpaceClaim [9]. En concreto se han empleado elementos tetraédricos solid187 tanto para la modelización del fémur como de los tornillos canulados y corticales. Asimismo, se ha simplificado el análisis asumiendo comportamiento isotrópico y lineal [5]. Por otro lado, como condiciones de contorno se ha asumido un apoyo fijo para todos los puntos de la zona distal del fémur, así como una superficie de carga en la zona proximal, para reducir la concentración de tensiones, Figura 6.

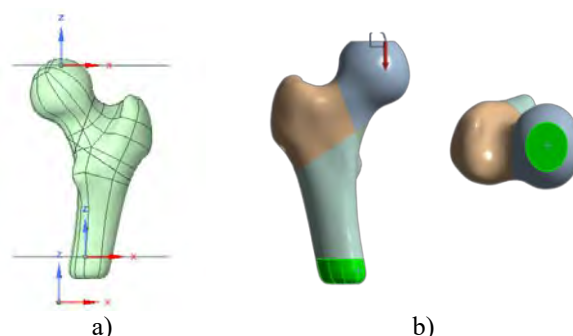


Figura 6. Condiciones de contorno: a) Planos de corte; b) Zona de apoyos y aplicación de la carga.

Para la preparación del modelo del fémur, se ha generado un plano de contacto asociado con la zona de fisura a 70°, Figura 1. Además, se ha diferenciado entre hueso cortical y hueso trabecular en la zona medial del fémur, manteniendo material trabecular en la cabeza del fémur de acuerdo con el fabricante e investigaciones previas [10].

Una vez preparado el modelo, se realizó la malla mediante elementos tetraédricos de tamaño 4,3 mm para el fémur y de 1 mm para los tornillos, Figura 7, y se fijaron los contactos entre superficies según la Tabla 1.

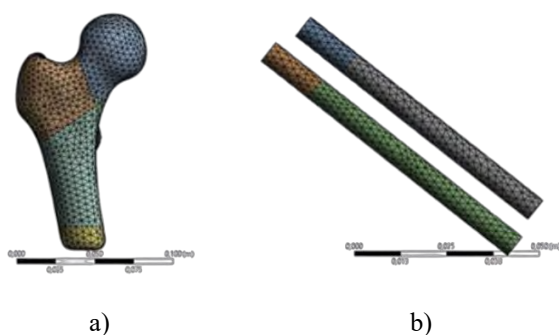


Figura 7. Modelo numérico: a) Fémur; b) Sistema de fijación

Elemento a unir	Tipo de contacto
Hueso - Rosca	Bonded
Hueso - Vástago	No separation
Hueso - Fractura	No separation

Tabla 1. Tipos de contacto

3. Resultados

Para el análisis comparativo de los diferentes modelos se han evaluado las siguientes variables: a) Rigidez axial en cabeza de fémur y en zona de fractura; b) Deformación; c) Tensión equivalente de Von Mises; y d) Carga de rotura. La primera se ha determinado como $k=F/L$, siendo F la fuerza aplicada en la cabeza del fémur, y L el desplazamiento vertical en sentido de la carga en la cabeza del fémur o en el plano de fractura. La segunda variable hace referencia a la deformación máxima registrada según la dirección de la carga en el modelo. La tercera variable está relacionada con la tensión equivalente máxima durante el ensayo tanto en tornillos como en zona de fémur. Finalmente, la cuarta variable hace referencia a la carga axial de rotura determinada en base a los valores facilitados por el fabricante.

En relación a la primera variable analizada, se puede observar en la Figura 8 que la rigidez axial en la cabeza y en la zona de fractura varían ligeramente entre los diferentes modelos de fijación. En concreto, los modelos CS de titanio (grupo 1, modelos 1, 2, 3 y 4) son los que menor rigidez axial tienen en comparación a los modelos XCS (grupo 2 y 3, modelos 9 a 24) que muestran una rigidez mayor, pudiendo proporcionar una mejor estabilidad y resistencia a la deformación.

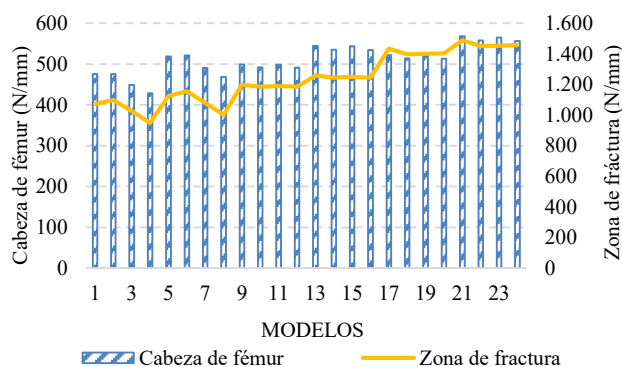


Figura 8. Variación de la rigidez axial

Respecto a los resultados asociados con la segunda variable, esta investigación pretende evaluar qué sistemas de fijación permitirían alcanzar la consolidación primaria y cuáles la secundaria, en base a los datos obtenidos a partir de investigaciones previas [4]. De esta forma, deformaciones inferiores al 2% en el plano de la fractura facilitan un proceso de consolidación primaria. En este sentido, la Figura 9 muestra como 23 de los 24 sistemas evaluados en esta investigación permitirán obtener una deformación inferior al 2%. En particular, la fijación relacionada con el modelo 4, Figura 2, muestra que la deformación interfragmentaria asociada con la fijación del tipo CS de titanio sería la única que podría provocar la consolidación secundaria en lugar de la primaria. Por otro lado, se debe destacar que los modelos asociados al sistema de fijación XCS tipo 2, Figura 4, tanto en su modalidad con titanio como con acero, son los que proporcionan mejores resultados desde la perspectiva de la consolidación primaria.

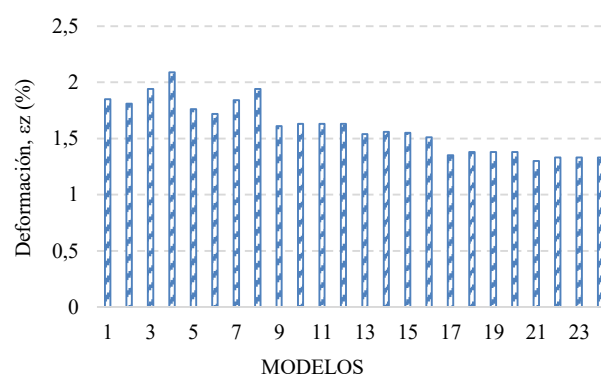


Figura 9. Variación de la deformación máxima

En relación a la tercera variable, la Figura 10 muestra los valores de la tensión de Von Mises en los puntos más solicitados tanto para los tornillos como para la zona del fémur. Los resultados indican que en las zonas de fijación los tornillos que presentan valores más elevados de tensión son los asociados con los sistemas XCS tipo 2, pudiendo llegar a generarse el fallo de los mismos en función del material elegido, con valores próximos a los 200 MPa. Por otro lado, se puede observar que los modelos relacionados con el sistema CS, muestran que, para el mismo nivel de carga, la tensión equivalente quedaría por debajo de los 150 MPa, asegurando un mayor nivel de seguridad. Además, en relación a las tensiones en el fémur, no hay tanta variabilidad de tensiones, estando todas entre los 30 y 40 MPa, destacando los modelos XCS tipo 1 de titanio (9, 10, 11 y 12) que son los que mayor tensión de von Mises alcanzan, y también destacan los modelos XCS tipo 2 de acero inoxidable (21, 22, 23 y 24) ya que estos están un poco por debajo de los 30 MPa, asegurando mayor nivel de seguridad en comparación con los otros modelos.

Finalmente, en relación a la última variable analizada, la Figura 11 muestra los resultados relacionados con los desplazamientos y la carga de rotura. En base a ellos, se puede deducir que los modelos de titanio presentan mayores desplazamientos frente a los de acero inoxidable, pero a la vez, permiten resistir ligeramente mayores cargas antes del fallo. En particular, se debe destacar el sistema de fijación

CS pues permiten resistir las mayores cargas a costa de permitir los mayores desplazamientos del sistema.

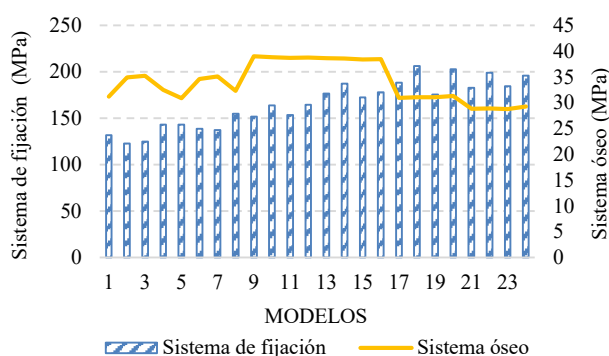


Figura 10. Variación de la tensión máxima de von Mises

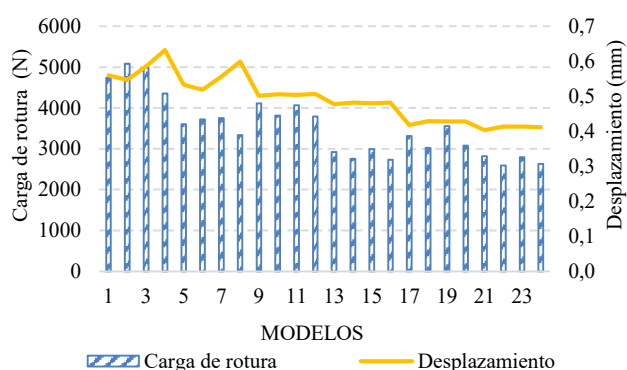


Figura 11. Variación del desplazamiento y la carga de rotura

4. Conclusiones

El presente estudio tenía como objetivo investigar y comparar diferentes sistemas de fijación interna utilizados en el tratamiento de fracturas del cuello femoral, centrándose en la fisura vertical. Se buscó evaluar la influencia de la disposición de los tornillos en la zona de fractura en términos de resistencia al desplazamiento y consolidación de la fractura.

Los resultados biomecánicos varían en los modelos debido a las diferencias en el diseño. Por ejemplo, los modelos 4 y 8, con una distancia menor entre todos los tornillos, muestran una menor rigidez en la fractura y carga de falla en comparación con el modelo 2, que tiene más distancia entre los tornillos. Esto sugiere que un mayor espacio entre los tornillos puede brindar una mejor estabilidad en la fijación CS, mientras que para las fijaciones XCS tipo 1 y tipo 2 es al revés, cuanto menos separación más estabilidad.

Considerando un equilibrio entre los resultados obtenidos, los modelos 2, 9, 11, 17 y 19 se destacan como opciones favorables en términos de rigidez, carga de falla, deformación y tensión, siendo todos los modelos de titanio.

Estos hallazgos respaldan la importancia de una cuidadosa planificación y selección de la configuración de los tornillos en el tratamiento de fracturas del cuello femoral. Una correcta disposición de los tornillos puede proporcionar una

mayor estabilidad y resistencia en la unión de la fractura, lo que puede contribuir a una mejor consolidación y a mejorar los resultados clínicos en los pacientes.

Agradecimientos

Queremos dar las gracias a la Sociedad Española de Ingeniería Biomédica y a la Universidad de Alicante por darnos la oportunidad de participar en los premios SEIB-FENIN, así como en el congreso anual de la Sociedad Española de Ingeniería Biomédica.

Referencias

- [1] M. Panteli, P. Rodham, and P. V. Giannoudis, “Biomechanical rationale for implant choices in femoral neck fracture fixation in the non-elderly,” *Injury*, vol. 46, no. 3, pp. 445–452, Mar. 2015, doi: 10.1016/j.injury.2014.12.031.
- [2] A. H. Schmidt and M. F. Swiontkowski, “Femoral Neck Fractures,” *Orthopedic Clinics of North America*, vol. 33, no. 1, pp. 97–111, Jan. 2002, doi: 10.1016/S0030-5898(03)00074-9.
- [3] “Search Results.” <https://surgeryreference.aofoundation.org/searchresults?q=ao%20classification%20bones&sort=relevancy> (accessed May 11, 2023).
- [4] S. Mehmood, U. Ansari, M. Najabat Ali, and N. Fatima Rana, “Internal fixation: An evolutionary appraisal of methods used for long bone fractures,” *International Journal of Biomedical and Advance Research*, vol. 5, no. 3, p. 142, Mar. 2014, doi: 10.7439/ijbar.v5i3.627.
- [5] J. Li, P. Yin, L. Zhang, H. Chen, and P. Tang, “Medial anatomical buttress plate in treating displaced femoral neck fracture a finite element analysis,” *Injury*, vol. 50, no. 11, pp. 1895–1900, Nov. 2019, doi: 10.1016/j.injury.2019.08.024.
- [6] M. Parker, K. Porter, D. Eastwood, M. Schembi Wismayer, and A. Bernard, “Intracapsular fractures of the neck of femur. Parallel or crossed garden screws?,” *J Bone Joint Surg Br*, vol. 73-B, no. 5, pp. 826–827, Sep. 1991, doi: 10.1302/0301-620X.73B5.1894675.
- [7] R. S. Garden, “STABILITY AND UNION IN SUBCAPITAL FRACTURES OF THE FEMUR,” *J Bone Joint Surg Br*, vol. 46-B, no. 4, pp. 630–647, Nov. 1964, doi: 10.1302/0301-620X.46B4.630.
- [8] M. J.-Q. Peng *et al.*, “Biomechanical analysis for five fixation techniques of Pauwels-III fracture by finite element modeling,” *Comput Methods Programs Biomed*, vol. 193, p. 105491, Sep. 2020, doi: 10.1016/j.cmpb.2020.105491.
- [9] G. Campoli, H. Weinans, and A. A. Zadpoor, “Computational load estimation of the femur,” *J Mech Behav Biomed Mater*, vol. 10, pp. 108–119, Jun. 2012, doi: 10.1016/j.jmbbm.2012.02.011.
- [10] F.-C. Kuan, K.-L. Hsu, C.-L. Lin, C.-K. Hong, M.-L. Yeh, and W.-R. Su, “Biomechanical properties of off-axis screw in Pauwels III femoral neck fracture fixation: Bicortical screw construct is superior to unicortical screw construct,” *Injury*, vol. 50, no. 11, pp. 1889–1894, Nov. 2019, doi: 10.1016/j.injury.2019.07.020.
- [11] GENERAL CATALOG”, Accessed: May 15, 2023. [Online]. Available: www.sawbones.com

Aplicación de Realidad Aumentada enfocada a la Salud Bucodental en edades tempranas

A. Aniorte García¹, D. Ojados González², G. A. Salcedo Eugenio²

¹Universidad Politécnica de Cartagena, Cartagena, España {alicia.aniorte}@edu.upct.es

²Servicio de Apoyo a la Investigación Tecnológica (SAIT), Universidad Politécnica de Cartagena, Cartagena, España {lola.ojados, gustavo.salcedo}@upct.es

Resumen

La realidad aumentada (RA) es una visión en tiempo real de un ambiente real físico que ha sido mejorado o aumentado añadiendo información virtual generada por computador; combina objetos reales con virtuales [1]. Esta puede ser una herramienta muy poderosa, ya que tiene diversas aplicaciones en campos de interés, como en el turismo, en juegos, en arquitectura e ingeniería, en diseño de interiores, y en el ámbito que reviste mayor interés para este trabajo: la educación [4]. La finalidad de este trabajo es el diseño de una aplicación educativa interactiva mediante el uso de realidad aumentada que pretende implementar de manera sencilla, utilizando el entorno de UnityHub y de Vuforia Engine, un juego que permita a los usuarios de edades tempranas aprender sobre la salud bucodental.

1. Introducción

La realidad extendida es el área de conocimiento que engloba las tecnologías de realidad virtual, realidad aumentada y la realidad mixta. Con el fin de comprender adecuadamente el contenido de este trabajo, es fundamental familiarizarse con estos conceptos:

- La realidad virtual (RV) es una experiencia inmersiva multisensorial, donde una persona puede explorar e interactuar con un ambiente virtual de 3 dimensiones generado por un ordenador.
- Se define la realidad aumentada (RA) como una visión en tiempo real de un ambiente real físico que ha sido mejorado o aumentado añadiendo información virtual generada por computador; combina objetos reales con virtuales. [1] Para el desarrollo de este trabajo se empleará la tecnología de realidad aumentada o RA.
- Finalmente, la realidad mixta (RM) es un ambiente donde la realidad física se combina con contenidos digitales de manera que permite la interacción tanto con los objetos reales como con los virtuales. [3]

Tras conocer la base de estos conceptos, cabe centrar este estudio en la tecnología de realidad aumentada, en la que se distinguen diferentes niveles de RA, dependiendo del tipo de parámetros y técnicas empleadas ([5] Aguirregoitia, 2016):

0. Hiperenlace al mundo físico: une el mundo físico al virtual. Códigos de barras o QR que enlazan un objeto real con información virtual.
1. Realidad aumentada basada en marcadores: permite reconocer marcadores físicos representados en 2D en tiempo real y enlazarlos a un contenido virtual.

2. RA sin marcadores: en este nivel se elimina la necesidad de marcadores y es posible superponer información basándose en el uso del sistema de posicionamiento GPS y la brújula.
3. Visión aumentada: entra ya en el mundo de los dispositivos *wearables* y comprende dispositivos de uso cotidiano y alejarse de la necesidad de un monitor.

El proceso de funcionamiento de la RA es el siguiente: mediante un dispositivo que disponga de una cámara al que previamente se le haya cargado una aplicación software de realidad aumentada generada y configurada con herramientas informáticas y motores gráficos habilitados a tal efecto, se enfoca a una imagen target (objetivo) que la aplicación software reconocerá, y en la pantalla del dispositivo aparecerá la imagen 3D o la información gráfica que se haya configurado.

La RA puede ser una herramienta muy poderosa. Algunas de las aplicaciones exitosas que más se observan en el día a día de la RA son en el turismo (museos), en juegos (por ejemplo, Pokemon Go), juegos serios (aplicación para PRL y riesgos en laboratorios Preven Go), en arquitectura e ingeniería, en diseño de interiores (IKEA), y en la educación. [4]

Pese a que la implementación de la realidad aumentada en la educación es muy reciente, ya se han comprobado los efectos positivos en el aprendizaje. [2]

En la educación, la RA puede ayudar a los estudiantes a comprender fácilmente la aplicación real del concepto que esté estudiando en ese momento, haciendo el aprendizaje más experiencial y, además, aumentando la motivación del estudiante. La imagen tridimensional generada con la aplicación de RA permite entender un concepto mejor que solo con la explicación. [4]

Las aplicaciones de realidad aumentada con fines educacionales están diseñadas para mejorar el aprendizaje de los usuarios a través de un enfoque más interactivo, que promueve una participación más alta. Convierte el aprendizaje en un “juego” divertido e inmersivo. [6]

Está indicado el uso de la realidad aumentada en edades tempranas (educación infantil) debido a las características de esta etapa evolutiva [2]. Los niños de edades tempranas tienen una capacidad de atención muy reducida; esto supone un gran desafío para los educadores, ya que resulta problemático para ellos el conseguir la atención prolongada de estos alumnos. [6]

El uso de aplicaciones de realidad aumentada puede ayudar a superar este problema, ya que hacer la educación entretenida es una manera excelente de promover la participación de los alumnos en una actividad que resulte en estos ganando conocimientos. [6]

Profundizando en el tema de la aplicación de RA que se pretende crear, su resultado puede ser un medio eficaz para que los niños pierdan el temor de asistir a una consulta o revisión odontológica.

2. Objetivos

El **objetivo principal** de este trabajo es el diseño de una aplicación educativa interactiva mediante el uso de realidad aumentada que pretende implementar de manera sencilla, utilizando el entorno de UnityHub y de Vuforia Engine, un juego que permita a los usuarios de edades tempranas aprender sobre la salud bucodental.

En cuando a los **objetivos secundarios**, se pueden destacar los siguientes:

- Ayudar a la comunidad de interés identificada a la que va dirigido a familiarizarse con el proceso de aprendizaje mediante contenidos desarrollados con nuevas tecnologías como las de aplicaciones de realidad aumentada.
- Ampliar los reducidos recursos existentes relacionados con las aplicaciones educativas y concretando en materia de salud, utilizando RA.
- Proponer una alternativa para la disposición de nuevas actividades divulgativas y educativas para campañas a favor de la salud.
- Ofrecer estos contenidos a las consejerías de educación como actividades formativas a desarrollar en las aulas de primaria.
- Poner a disposición de clínicas odontológicas y consultas médicas este tipo de aplicaciones para captar la atención de los más pequeños y motivarles a cuidar su salud bucodental.

3. Materiales

Para el desarrollo de este trabajo se han utilizado los siguientes recursos:

- UnityHub: software para el desarrollo del proyecto de realidad aumentada.
- Vuforia Engine: plataforma de desarrollo de aplicaciones de realidad aumentada.
- Dispositivo Android
- Modelos 3D procedentes de fuentes públicas, con posibilidad de creación de modelos propios [7-9]
- Procreate: software de diseño gráfico para preparación de la imagen objetivo.
- Equipos informáticos con capacidad mínima requerida para el buen funcionamiento de estos motores gráficos.

4. Metodología

4.1. Revisión bibliográfica y estado del arte

Se revisaron diferentes fuentes de información, analizando la información obtenida acerca del estado actual de conocimientos en el campo de la salud bucodental en el que se está trabajando, y recopilando diferentes aplicaciones de la Realidad Aumentada (RA), y las tendencias actuales en esta línea.

Uno de los estudios encontrados sobre el tema fue “Realidad Aumentada como Herramienta Lúdica para el Estudio de la Anatomía Humana. Caso de Estudio: Los Dientes” (Plata YL, 2015). [10]

4.2. Diseño de la propuesta

En esta fase se definieron las especificaciones de la aplicación que se quería implementar, usando RA, y el campo y la funcionalidad que presentaría.

Las especificaciones principales de este trabajo son:

- la aplicación trata el tema de la salud bucodental.
- los principales usuarios son personas de edades tempranas, alumnos de primaria principalmente.
- la aplicación está diseñada siguiendo el estilo de un juego, ya que esto puede generar un mayor atractivo para los niños, incrementando así el tiempo de concentración en la misma.
- la aplicación permitirá ofrecer diferentes contenidos de información, basándose en RA y con la posibilidad de interacción mediante botones virtuales.

4.3. Creación de la Imagen Target

Una vez clara la idea principal de la aplicación y la planificación para su desarrollo, se procedió al diseño de la imagen *target*. Una imagen *target* es una imagen que el software reconocerá como objetivo para desencadenar el comienzo del funcionamiento de la aplicación RA, y que servirá para el uso e interacción con la propia aplicación.

Para el diseño de esta imagen target, se utilizó la aplicación de dibujo y diseño gráfico del software Procreate que sirvió para representar al paciente potencial de un dentista con la boca abierta sin dientes (ya que éstos se añadirían posteriormente en forma de modelos 3D).

Se creó la imagen *target* en el proyecto RA con la imagen que previamente se diseñó en Procreate (Figura 1).



Figura 1. Imagen target de la aplicación de RA

A continuación, se creó una nueva licencia en Vuforia Engine para poder hacer uso de esta plataforma, que posteriormente se copió en el inspector del proyecto de Unity.

En la ventana de “Target Manager” de Vuforia, se creó una nueva base de datos, a la que se le añadió como Target la imagen diseñada, como se puede observar en la Figura 2.

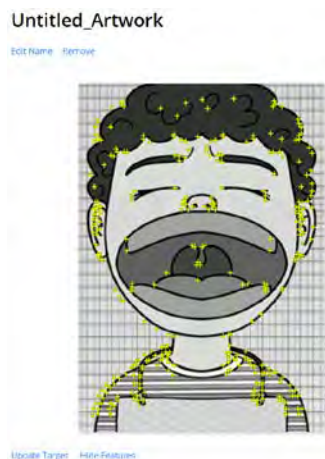


Figura 2. Imagen target en la página de Vuforia Engine

4.4. Importación de modelos 3D

Una vez creada la imagen *target* e incorporada en el proyecto de Unity, se procedió a importar los modelos 3D de los dientes. Estos modelos se redimensionaron y colocaron de manera que estuviesen bien alineados con la imagen.

Tras colocar los modelos 3D, se comprobó el funcionamiento de la aplicación, y se observó que se podía visualizar la raíz de los dientes, cosa que no estaba contemplada en la idea original, por lo que se diseñó una boca 3D simple para que apareciese a la vez que los dientes y tapase la raíz de estos (Figura 3), de manera que se ofreciera una imagen más natural del personaje.

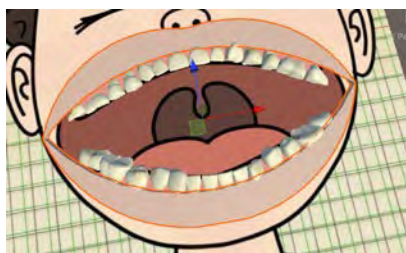


Figura 3. Diseño de la boca sobre las raíces de los dientes

4.5. Creación de botones virtuales

Una vez se colocaron correctamente los modelos 3D, se añadieron los 4 botones virtuales que agrupan las piezas dentales en: incisivos, caninos, premolares y molares.

El fin de la aplicación es aprender el nombre de los distintos conjuntos de dientes mientras se realizaba su limpieza.

Para conseguir este objetivo, se duplicaron todos los dientes ya colocados: al primer grupo de dientes se le dio

una textura de “diente limpio”, mientras que al segundo se le dio una textura de “diente sucio”.

El objetivo de los botones es pasar del modo sucio al modo limpio con su activación, mediante un sencillo código.

4.6. Generación de la apk

Finalmente, una vez comprobado el funcionamiento del proyecto, se exportó el paquete ejecutable de la aplicación (apk) a un dispositivo Android, en modo desarrollador, desde el software UnityHub.

5. Resultados

La aplicación resultante es una apk que al ejecutarse muestra la interface visible en la Figura 4.



Figura 4. Aplicación resultante: pantalla inicial

Esta apk puede activarse desde cualquier dispositivo con sistema operativo Android donde se instale y está conformada en forma de actividad de gamificación para el aprendizaje interactivo de cualquier usuario, aunque se dirija en concreto a un público de edades comprendidas entre 4 y 10 años como usuarios potenciales.

Además, la aplicación podría estar disponible para los usuarios de IOS de una forma simple si esta configuración se incluyese en Unity para este tipo de plataformas, si así se deseara.

6. Conclusiones

Todos los objetivos que se fijaron en el planteamiento de este trabajo han sido alcanzados; se consiguió desarrollar una aplicación que de manera sencilla enseña al usuario la localización de los 4 grupos principales de dientes y su proceso de higiene. Esto cubre el objetivo principal que dio pie a este estudio.

Además, gracias a este proyecto se ha podido utilizar de manera competente el entorno de UnityHub y de Vuforia Engine aplicado a un proceso que tiene que ver con el ámbito sanitario.

Existe una creciente adopción e impacto de la realidad extendida en la educación; según estudios estadísticos, el

uso de estas tecnologías en la educación puede mejorar el aprendizaje de los usuarios en un 76%. [11]

En esta línea, Edgar Dale, pedagogo estadounidense, desarrolló la pirámide de aprendizaje (o el Cono de la Experiencia) en 1946 (Figura 5). Esta pirámide representa y evalúa los métodos de aprendizaje a partir de una representación triangular, siendo los métodos que menos ayudan a la retención de la información el escuchar o leer, y los que más, el decir y hacer. Según la pirámide de aprendizaje, la enseñanza con RA ayudaría a retener un 75% de la información, ya que incorpora aprendizaje experimental. Estas tecnologías tienen el potencial de revolucionar el mundo de la enseñanza, haciendo que sea más divertida e interactiva. [12]



Figura 5. Pirámide de aprendizaje de Edgar Dale.

En base a estos estudios se puede concluir que la existencia de aplicaciones de RA como la desarrollada en este trabajo ayudan a una mejor y más eficaz retención del conocimiento en los temas incorporados en la información contenida en las mismas.

Agradecimientos

Cabe agradecer al comité organizador del CASEIB por la oportunidad de mostrar los resultados de trabajos realizados en el entorno académico universitario y de participar en el concurso propuesto a alumnos estudios enfocados al ámbito de la Ingeniería Biomédica. Así como al personal del FABLAB del SEDIC de la UPCT por su apoyo y atención.

Además, agradecer a mis profesores y tutores de este trabajo, Dolores Ojados y Gustavo Adolfo Salcedo, por su ayuda, motivación y su atención durante el desarrollo del mismo, así como a los servicios SEDIC y FABLAB, por el soporte proporcionado.

Referencias

- [1] Carmigniani, J., Furht, B. Augmented Reality: An Overview. In: Furht, B. (eds) Handbook of Augmented Reality. Springer, New York, NY, 2011. (ISBN: 978-1-4614-0064-6).
- [2] Página web EDUCACIÓN 3.0 – Lo que aporta la realidad aumentada a la educación (Jorge Calvo). <https://www.educaciontrespuntocero.com/tecnologia/realidad-aumentada-educacion/>
- [3] Página web TechTarget – DEFINITION: mixed reality (hybrid reality, extended reality) (Ivy Wigmore 2018). <https://www.techtarget.com/whatis/definition/mixed-reality>
- [4] Página web Value Universidad Europea – La realidad aumentada aplicada a la educación (Eva Roper, Gonzalo Mariscal). <https://value.universidadeuropea.com/la-realidad-aumentada-aplicada-a-la-educacion/>
- [5] Aguirregoitia A, Allende I, López JR, Artetxe E. Leihoa: Una ventana a la realidad aumentada en Educación Infantil. *2016 International Symposium on Computers in Education (SIIE)*
- [6] Página web Zoog - How Kids Improve Learning with Augmented Reality Apps (The Zoog Team 2022). <https://www.getzoog.com/augmented-reality-kids/>
- [7] Página web 3D Warehouse (Trimble Inc 2023). <https://3dwarehouse.sketchup.com/?hl=es>
- [8] Página web Turbosquid (Shutterstock 2023). <https://www.turbosquid.com/es/Search/3D-Models/free>
- [9] Página web Free3D (Free3D 2023). <https://free3d.com/es/>
- [10] Plata YL, García-Ojeda JC, Méndez F. Realidad Aumentada como Herramienta Lúdica para el Estudio de la Anatomía Humana. Caso de Estudio: Los Dientes. *2015*.
- [11] Página web Colegios Málaga - Estadísticas revelan el impacto de la realidad virtual en la educación (Colegios Málaga). <https://www.colegiosmalaga.com/estadisticas-de-la-realidad-virtual-en-la-educacion/>
- [12] Página web eLearning Adobe - Augmented Reality and Virtual Reality in School Education (Jitender 2020). <https://elearning.adobe.com/2020/01/augmented-reality-and-virtual-reality-in-school-education/>

Optimización de Servicios de Urgencias mediante Modelos de Inteligencia Artificial: Revisión Sistemática de la Predicción de Ingresos Hospitalarios desde Urgencias

L. Azkue Amondarain^{1,2}, J. Kerexeta Sarriegi^{2,3}, N. Larburu Rubio^{2,3}, U. Ayala Fernandez¹

¹ Departamento de Ingeniería Biomédica, Mondragon Unibertsitatea, 20500 Mondragón, España

² Vicomtech Foundation, Basque Research and Technology Alliance, (BRTA), 20009 Donostia, España

³ Biodonostia Health Research Institute, 20014 San Sebastián, España

Resumen

La creciente demanda en servicios de urgencias provoca congestiones en la gestión de pacientes y gran carga para los clínicos, lo cual dificulta un tratamiento óptimo al igual que los tiempos de espera llegan a ser excesivos. La inteligencia artificial (IA) se presenta como una solución innovadora para agilizar la atención de estos pacientes. Este artículo propone una revisión sistemática centrada en modelos de IA para anticipar ingresos hospitalarios desde urgencias lo cual mejora la eficiencia y la rapidez en la atención a los pacientes. Siguiendo la metodología PRISMA, se seleccionaron 14 estudios tras un análisis exhaustivo de 367 artículos. Los modelos, predominantemente basados en regresión logística y aumento de gradiente, mostraron AUC en el rango de 0.75–0.92. Se identificaron tres momentos en los que se han aplicado los modelos de IA con diferentes capacidades predictoras y variables disponibles: a la llegada del paciente, después del triaje y enfocado en pruebas de laboratorio. Cada vez que se recopilen más datos, la evaluación del riesgo de ingreso se actualizará mediante los tres modelos. El primer modelo se centrará en un cribado rápido de los pacientes, mientras que los otros modelos mejorarán la predicción y serán más precisos, especialmente para aquellos pacientes en estados más graves. Este enfoque dinámico garantiza una adaptación continua a medida que se disponga de nueva información, permitiendo una toma de decisiones más ágil y precisa en la gestión de urgencias médicas.

1. Introducción

En el contexto actual, la abrumadora demanda en servicios de urgencias ha desencadenado colapsos sistemáticos y una gestión mejorable a la hora de atender a los pacientes. Este escenario crítico se traduce en tiempos de espera excesivos para los pacientes, lo cual subraya la urgente necesidad de abordar y optimizar la atención médica de emergencia. La presión sobre los servicios de urgencias no solo resalta la importancia de mejorar la eficiencia, sino que también subraya la necesidad de una intervención innovadora y efectiva [1], [2].

En paralelo, el campo de la inteligencia artificial ha experimentado una notable evolución, especialmente en el ámbito sanitario. Los avances tecnológicos han permitido la aplicación de modelos predictivos y herramientas analíticas en tiempo real, transformando la forma en que abordamos los desafíos médicos. Esta revolución tecnológica ofrece oportunidades sin precedentes para optimizar la gestión de recursos en situaciones críticas,

como las emergencias médicas. La capacidad de procesar grandes volúmenes de datos de manera rápida y precisa brinda un potencial significativo para mejorar la toma de decisiones y, en última instancia, salvar vidas [3].

En este contexto, nuestro artículo busca integrar de manera práctica la necesidad apremiante en los servicios de urgencias mediante la aplicación de modelos de inteligencia artificial. La revisión sistemática [4] se enfocará en identificar características esenciales para el desarrollo de estos modelos, mientras que la implementación de múltiples modelos predictivos en distintos momentos de la urgencia ofrecerá una perspectiva detallada y escalonada. Nuestro propósito es mejorar la gestión de pacientes en situaciones críticas, proporcionando respuestas más efectivas y precisas a la creciente demanda en los servicios de urgencias. Este enfoque pragmático busca contribuir de manera significativa a la calidad de la atención médica de emergencia, aprovechando las capacidades de la inteligencia artificial de manera relevante y eficiente.

2. Métodos

Esta revisión se basa en la estructura PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [5]. Esta metodología proporciona un marco riguroso y transparente para la realización de revisiones sistemáticas.

2.1. Bases de Datos y Metodología PICO

Para la búsqueda de estudios, se implementa la metodología PICO [6] (Población de interés, Intervención, Comparación, Resultado) en las bases de datos de PubMed, Scopus, Google Scholar, entre otras.

- **Población de Interés:** Pacientes en el departamento de urgencias.
- **Intervención:** Modelos predictivos para ingresos hospitalarios desde el departamento de urgencias.
- **Comparación:** No aplicable.
- **Resultado:** Admisión del paciente.

Se utilizan términos y combinaciones de palabras clave específicas relacionadas con la población de interés, la intervención, y el resultado en las búsquedas de las bases de datos. Esta cuidadosa selección de términos y la

aplicación coherente de la metodología PICO aseguran una búsqueda exhaustiva y específica, centrada en la temática de modelos predictivos para ingresos hospitalarios desde el departamento de urgencias y sus efectos en la admisión de pacientes.

2.2. Criterios de Inclusión y Exclusión

En esta sección, mencionan los criterios que guían la selección y exclusión de estudios en la revisión sistemática. Estos criterios están diseñados para garantizar la relevancia y coherencia con los objetivos de nuestro estudio.

Se excluyen los estudios que cumplan alguno de estos criterios:

- Estudios enfocados exclusivamente en modelos predictivos para la readmisión hospitalaria.
- Investigaciones centradas en condiciones médicas específicas (e.g., asma o bronquiolitis) como población principal.
- Modelos que predicen eventos futuros o requieren datos posteriores al triaje.
- Estudios que solo proporcionan evaluación de factores de riesgo sin desarrollar un modelo predictivo.
- Modelos que no responden directamente a la pregunta de interés: prever la admisión desde el departamento de urgencias.
- Trabajos no disponibles en su totalidad.
- Estudios publicados en idiomas distintos al inglés o español.

3. Resultados

Esta revisión sistemática ha identificado inicialmente un total de 367 posibles artículos, reduciéndolos a 14 estudios al final del proceso (Ilustración 1). La cifra total de 367 artículos se ha obtenido de PubMed (n = 150), Scopus (n = 194), Google Scholar (n = 12) y otros sitios de búsqueda distintos a los anteriores (n = 11).

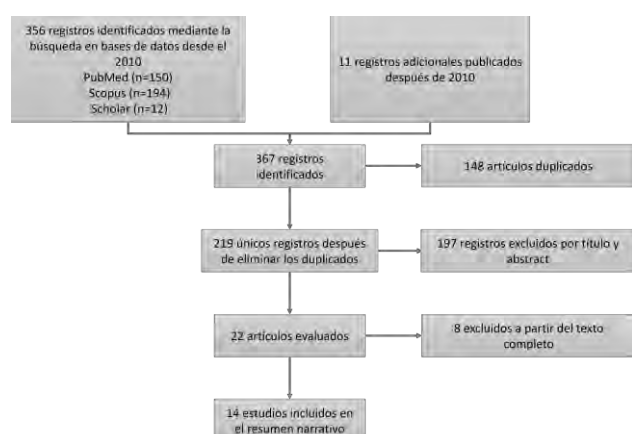


Ilustración 1. Diagrama de Flujo PRISMA de Modelos de Predicción de Ingresos Hospitalarios Incluidos del Servicio de Urgencias.

Como se observa en la Ilustración 1, tras la búsqueda automática, se han eliminado los artículos duplicados, dejando un total de 219 artículos. Posteriormente, al

analizar los títulos y resúmenes de cada uno de los artículos no duplicados, se descartan 197 estudios, y se seleccionan 22 estudios para un análisis más detenido.

Siguiendo los mismos criterios de exclusión, se analizan los textos completos de los 22 estudios restantes, resultando en la exclusión de seis de ellos y, finalmente, se han seleccionado 14 estudios relevantes. Estos 14 estudios han sido examinados uno por uno para extraer conclusiones y contribuir a los hallazgos de esta revisión.

3.1. Descripción de los estudios seleccionado

En el análisis final de esta revisión, se han identificado 14 estudios (Ilustración 1), distribuidos geográficamente de la siguiente manera: tres en España [7]–[9], tres en los Países Bajos [10]–[12], dos en Grecia [13], [14], dos en el Reino Unido [15], [16], dos en los Estados Unidos [17], [18], uno en Francia [19], y uno en Singapur [25]. En términos de la población estudiada, la mayoría de los artículos (11 de los 14) incluyeron a cualquier individuo que se presentara al departamento de urgencias. Sin embargo, tres estudios (n = 3) aplicaron restricciones a la población de estudio: uno excluyó a pacientes mayores de 70 años [11]; otro se centró únicamente en pacientes mayores de 18 años, dividiendo a los participantes en grupos según la edad (menores y mayores de 70 años) [12]; y el tercer estudio [7] incluyó a pacientes adultos que llegaron al departamento de urgencias en estado estable.

En relación con el tamaño de las muestras, este varió significativamente, desde 2476 [7] hasta 3,189,204 [8] visitas al departamento de urgencias. El porcentaje de ingresos en planta con respecto al total de instancias varió entre el 4.6% [7] y el 39% [10].

En términos de modelos, la regresión logística destacó como el algoritmo más empleado, utilizado en siete de los 14 estudios [7], [11], [12], [16], [17], [19], [20]. El segundo modelo más común fue el aumento de gradiente, presente en tres estudios [8], [15], [18]. Otros enfoques incluyeron Random Forest [14], Naïve Bayes Gaussiano [13], modelado de árboles de decisión con potencia de gradiente [10], y una red neuronal artificial [9]. Estos modelos, en su totalidad, se centran en predecir la probabilidad de ingreso en planta desde el departamento de urgencias, ofreciendo respuestas binarias o probabilísticas.

3.2. Variables utilizadas para crear el modelo predictivo

Todos los estudios utilizaron una categoría de datos de triaje. Un estudio en particular [15] destacó que la variable de triaje es uno de los factores más importantes en la predicción de ingresos hospitalarios. En 11 de los 14 artículos [3]–[5], [8]–[13], [15], [16], se utilizaron datos demográficos y la variable de edad fue la más utilizada, ya que se utiliza en todos los estudios. Cuatro de estos artículos identificaron la edad como un buen predictor de ingreso hospitalario [7], [15], [17], [20]. Las conclusiones también revelaron que los hallazgos clínicos y de laboratorio son la tercera categoría más utilizada, con nueve de los 14 artículos incorporando variables de esta categoría [7]–[15], [18]. Los datos de laboratorio se utilizaron de manera extensa, con cinco de los nueve

estudios haciéndolo [10]–[14]. Las variables de signos vitales se utilizaron en tres de los nueve estudios [10]–[12]. En seis estudios, se utilizaron variables relacionadas con el historial médico [8], [9], [12], [15], [16], [20], y uno de ellos determinó que admisiones previas en el último mes y en el último año son buenos predictores de ingreso hospitalario [15]. En un estudio, se incorporaron medicamentos [7]. Finalmente, nueve estudios utilizaron otras variables que no corresponden a estas categorías [5], [8]–[13], [15], [16]. De estas variables no clasificadas, el modo de llegada al departamento de urgencias es la más común y, según un artículo, es un buen predictor [17].

Es crucial subrayar que los modelos predictivos que se han examinado se basan en una consideración de datos en diversas etapas del proceso de atención médica de emergencia. Este análisis se realiza en tres momentos clave: antes del triaje, durante el propio proceso de triaje y después de su ejecución.

3.3. Modelos de inteligencia artificial

Los resultados de los 14 estudios analizados se presentaron utilizando la métrica del Área bajo la Curva (AUC, por sus siglas en inglés), que mide la separabilidad de las clases en un problema de clasificación. En la mayoría de los artículos, se calculó el intervalo de confianza del 95% (IC) para proporcionar un rango en el cual se estima que el modelo tiene un rendimiento con un 95% de probabilidad. Cuanto mayor sea el rango, mayor será la incertidumbre. Los estudios obtuvieron resultados de AUC en el rango 0.75–0.92. El estudio con el mejor resultado fue [18], realizado en los Estados Unidos, que obtuvo un AUC de 0.92 (IC del 95% 0.92–0.93). En segundo lugar se ubicó el artículo [8], realizado en España, con un AUC de 0.8938 (IC del 95% 0.8929–0.8948). En tercer lugar se encontró el estudio [16], también realizado en los Estados Unidos, con un AUC de 0.8774 (IC del 95% 0.8752–0.8796).

4. Discusión

Esta revisión sistemática abordó 14 estudios que desarrollaron modelos de aprendizaje automático para anticipar ingresos hospitalarios desde el departamento de urgencias (ED, por sus siglas en inglés). La diversidad en el número de variables utilizadas fue notable, ya que algunos estudios emplearon un número reducido y otros más de 20. La mayoría de los artículos analizados ($n = 11$) utilizaron datos de todos los pacientes en el ED para construir sus modelos, mientras que algunos establecieron restricciones en la población estudiada. Se observó que los modelos más comúnmente empleados para este tipo de predicción son la regresión logística y el aumento de gradiente, ya que 10 de los 14 artículos utilizaron alguno de estos dos, siendo la regresión logística la más prevalente, empleada en siete de los 14 estudios. En resumen, los resultados sugieren que estos modelos pueden ser efectivos en predecir con precisión la admisión del paciente en planta, con numerosos estudios informando sólidas relaciones entre sensibilidad y especificidad. El artículo que obtuvo los mejores resultados presentó un AUC de 0.92 [IC del 95% 0.92–0.93] [18], mientras que el de menor rendimiento registró un AUC de 0.766 (IC del 95% 0.759–0.781) [11].

Como resultado de este revisión sistemática proponemos como trabajo futuro la siguiente aproximación: Considerando que las variables utilizadas abarcan desde antes del triaje, durante y después del mismo, se proponen los modelos propuestos en la Ilustración 2, en el cual se desarrollarían tres modelos: (1) modelo a la llegada del paciente a urgencias, (2) modelo después de realizar el triaje y (3) un modelo focalizado en los pacientes que se sometan a pruebas de laboratorio. Este enfoque estratificado permitirá un cálculo de riesgo rápido sin necesidad de esperar los datos de triaje y laboratorio, e ir actualizando y haciendo más preciso el riesgo según se obtenga nueva información de paciente. Obteniendo así, una evaluación más precisa de las necesidades y condiciones de los pacientes, identificando de manera efectiva a aquellos que requieren una atención más urgente o se encuentran en un estado más grave en comparación con otros pacientes.

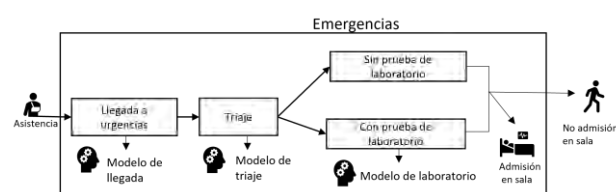


Ilustración 2. Propuesta Estratificada para Modelos de Predicción de Ingresos Hospitalarios desde Urgencias.

5. Conclusiones

Hasta la fecha, el algoritmo de regresión logística es el más ampliamente utilizado para predecir el ingreso en planta hospitalaria desde el departamento de urgencias. Estos modelos predictivos tienen el potencial de mejorar la calidad del servicio de urgencias, facilitando un manejo interno más eficiente. Al anticipar el número de ingresos, se posibilita una gestión más efectiva de los recursos, lo que convierte esta propuesta en una opción atractiva.

Agradecimientos

Los autores agradecen a INVIZA-Asunción Clínica y a STT Systems su apoyo en el proyecto INURGE. Este trabajo ha sido financiado por el proyecto de investigación INURGE (ZL-2022/00571) del programa HAZITEK del Gobierno Vasco de la agencia pública SPRI.

Referencias

- [1] F. Relinque-Medina, E. Pino-Moya, y J. Gómez-Salgado, «APROXIMACIÓN CUALITATIVA AL INCREMENTO DE LA DEMANDA ASISTENCIAL POR PROPIA INICIATIVA EN UN SERVICIO DE URGENCIAS HOSPITALARIAS», *Rev Esp Salud Pública*.
- [2] A. Pozuelo García, E. Torné Vilagrana, A. Guarga Rojas, M. G. Torras Boatella, M. Pasarin Rúa, y C. Borrell Thió, «Análisis de la demanda en los servicios de urgencias de Barcelona», *Atención Primaria*, vol. 32, n.º 7, pp. 423-424, 2003, doi: 10.1016/S0212-6567(03)70761-6.
- [3] M. Chen y M. Decary, «Artificial intelligence in healthcare: An essential guide for health leaders», *Healthc Manage Forum*, vol. 33, n.º 1, pp. 10-18, ene. 2020, doi: 10.1177/0840470419873123.
- [4] N. Larburu, L. Azkue, y J. Kerexeta, «Predicting Hospital Ward Admission from the Emergency Department: A Systematic Review», *Journal of Personalized Medicine*,

- vol. 13, n.º 5, Art. n.º 5, may 2023, doi: 10.3390/jpm13050849.
- [5] D. Moher, A. Liberati, J. Tetzlaff, y D. G. Altman, «Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement», *Ann Intern Med*, vol. 151, n.º 4, pp. 264-269, ago. 2009, doi: 10.7326/0003-4819-151-4-200908180-00135.
- [6] C. M. da C. Santos, C. A. de M. Pimenta, y M. R. C. Nobre, «The PICO strategy for the research question construction and evidence search», *Rev. Latino-Am. Enfermagem*, vol. 15, pp. 508-511, jun. 2007, doi: 10.1590/S0104-11692007000300023.
- [7] C. M. Elvira Martínez, C. Fernández, J. González del Castillo, J. J. González-Armengol, P. Villarroel, y F. J. Martín-Sánchez, «Prediction model for in-hospital admission in patients arriving in the Emergency Department», *Anales del Sistema Sanitario de Navarra*, vol. 35, n.º 2, pp. 207-217, ago. 2012, doi: 10.4321/S1137-66272012000200003.
- [8] J. Cusidó, J. Comalrena, H. Alavi, y L. Lluas, «Predicting Hospital Admissions to Reduce Crowding in the Emergency Departments», *Applied Sciences*, vol. 12, n.º 21, Art. n.º 21, ene. 2022, doi: 10.3390/app122110764.
- [9] A. Zlotnik, M. C. Alfaro, M. C. P. Pérez, A. Gallardo-Antolín, y J. M. M. Martínez, «Building a Decision Support System for Inpatient Admission Prediction With the Manchester Triage System and Administrative Check-in Variables», *CIN: Computers, Informatics, Nursing*, vol. 34, n.º 5, p. 224, may 2016, doi: 10.1097/CIN.0000000000000230.
- [10] A. De Hond *et al.*, «Machine learning for developing a prediction model of hospital admission of emergency department patients: Hype or hope?», *International Journal of Medical Informatics*, vol. 152, p. 104496, ago. 2021, doi: 10.1016/j.ijmedinf.2021.104496.
- [11] «Brink: Prediction admission in the older population... - Google Académico». Accedido: 10 de octubre de 2023. [En línea]. Disponible en: https://scholar.google.com/scholar_lookup?title=Prediction+admission+in+the+older+population+in+the+Emergency+Department:+The+CLEARED+tool&author=Brink,+A.&author=Almsa,+J.&author=Brink,+H.S.&author=De+Gelder,+J.&author=Lucke,+J.A.&author=Mooijaart,+S.P.&author=Zietse,+R.&author=Schuit,+S.C.E.&author=Lingsma,+H.F.&publication_year=2020&journal=Neth.+J.+Med.&volume=78&pages=357%E2%80%93367
- [12] J. A. Lucke *et al.*, «Early prediction of hospital admission for emergency department patients: a comparison between patients younger or older than 70 years», *Emerg Med J*, vol. 35, n.º 1, pp. 18-27, ene. 2018, doi: 10.1136/emermed-2016-205846.
- [13] G. Feretzakis *et al.*, «Predicting Hospital Admission for Emergency Department Patients: A Machine Learning Approach», en *Informatics and Technology in Clinical Care and Public Health*, IOS Press, 2022, pp. 297-300. doi: 10.3233/SHTI210918.
- [14] G. Feretzakis *et al.*, «Using Machine Learning Techniques to Predict Hospital Admission at the Emergency Department», *The Journal of Critical Care Medicine*, vol. 8, n.º 2, pp. 107-116, abr. 2022, doi: 10.2478/jccm-2022-0003.
- [15] «Using Data Mining to Predict Hospital Admissions From the Emergency Department | IEEE Journals & Magazine | IEEE Xplore». Accedido: 10 de octubre de 2023. [En línea]. Disponible en: <https://ieeexplore.ieee.org/document/8300528>
- [16] A. Cameron, K. Rodgers, A. Ireland, R. Jamdar, y G. A. McKay, «A simple tool to predict admission at the time of triage», *Emerg Med J*, vol. 32, n.º 3, pp. 174-179, mar. 2015, doi: 10.1136/emermed-2013-203200.
- [17] C. A. Parker, N. Liu, S. X. Wu, Y. Shen, S. S. W. Lam, y M. E. H. Ong, «Predicting hospital admission at the emergency department triage: A novel prediction model», *The American Journal of Emergency Medicine*, vol. 37, n.º 8, pp. 1498-1504, ago. 2019, doi: 10.1016/j.ajem.2018.10.060.
- [18] W. S. Hong, A. D. Haimovich, y R. A. Taylor, «Predicting hospital admission at emergency department triage using machine learning», *PLOS ONE*, vol. 13, n.º 7, p. e0201016, jul. 2018, doi: 10.1371/journal.pone.0201016.
- [19] G. Noel *et al.*, «Real-time estimation of inpatient beds required in emergency departments», *European Journal of Emergency Medicine*, vol. 26, n.º 6, p. 440, dic. 2019, doi: 10.1097/MEJ.0000000000000600.
- [20] Y. Sun, B. H. Heng, S. Y. Tay, y E. Seow, «Predicting Hospital Admissions at Emergency Department Triage Using Routine Administrative Data», *Academic Emergency Medicine*, vol. 18, n.º 8, pp. 844-850, 2011, doi: 10.1111/j.1553-2712.2011.01125.x.

Sesión Póster I

Miércoles 22 11:00 - 12:00
Claustro

Clasificación Supervisada de Pacientes con Síndrome de Brugada con Biomarcadores Derivados del ECG

Alba Isabel-Roquero¹, Pedro Gomis², Alvaro Leva³, Luis Tortosa², Flavio Palmieri^{2,4,*},
Elena Arbelo^{1,3,*}

¹ Institut d'Investigació August Pi y Sunyer (IDIBAPS), Barcelona, España isabel@recerca.clinic.cat

²ESAI Dept, EEBE, CREB, Universitat Politècnica de Catalunya, Barcelona, España pedro.gomis,
flavio.palmieri@upc.edu, tortosavalerol@gmail.com

³ Unidad de Arritmias, Hospital Clínic, Barcelona, España earbelo@clinic.cat,
alevalig7@alumnes.ub.edu

⁴ Fundació de Recerca Sant Joan de Déu, Esplugues-Barcelona, España

* Estos autores contribuyeron de la misma forma al desarrollo de este artículo

Abstract

El síndrome de Brugada (SB) es una rara canalopatía cardiovascular hereditaria asociada con el riesgo de fibrilación ventricular y muerte súbita (MS). Su estratificación de riesgo es un desafío y la mayoría de los pacientes son diagnosticados en la fase asintomática siendo el único tratamiento contra la MS la implantación de un desfibrilador eléctrico. En este estudio, presentamos herramientas de clasificación supervisada para la estratificación del riesgo del SB basadas en el análisis automático de información electrocardiográfica de alta resolución. Para ello, se utilizaron señales de ECGs de 24 horas y 12 derivaciones y variables clínicas de 64 sujetos con SB. Las señales de ECG se preprocesaron promediando la señal para reducir el ruido y obtener latidos individuales para el cálculo de 10 biomarcadores. Posteriormente, se probaron 4 algoritmos de clasificación supervisada diferentes basados en Random Forest (RF), Extreme Gradient Boosting (XGBoost), K-Nearest Neighbors (KNN) y Support Vector Machine (SVM). Los resultados del área bajo la curva fueron alrededor del 90%, sin embargo, los resultados de sensibilidad fueron alrededor del 50%. Los resultados no predicen de manera eficiente a los pacientes sintomáticos en riesgo de MS debido al reducido número de pacientes sintomáticos. Estudios posteriores con sujetos y variables adicionales pueden mejorar este pronóstico.

1. Introducción

El Síndrome de Brugada (SB) es una canalopatía hereditaria rara que muestra un patrón electrocardiográfico distintivo y está asociada con una alta incidencia de acci-

dentos arrítmicos súbitos e inesperados. Se cree que causa el 4-12% de todas las muertes súbitas (MS) y hasta el 20% entre los pacientes sin anomalías estructurales[1].

El SB se caracteriza por un patrón electrocardiográfico típico descrito como una elevación del segmento ST seguida de una pendiente en forma de aleta de tiburón (tipo 1) o en forma de silla de montar (tipo 2), los cuales pueden permanecer ocultos. La estratificación del riesgo y el manejo de los pacientes siguen siendo un desafío, ya que el único índice riesgo aceptado es la presencia de MS recuperada o un síncope arritmogénico y la gran mayoría de los pacientes son diagnosticados en la fase asintomática [1]. El único tratamiento disponible para prevenir la MS es la implantación de un desfibrilador cardíaco, que puede causar eventos adversos como descargas inapropiadas (hasta el 9%) o infecciones (20%) [2] y puede empeorar la salud cardíaca en los casos en los que el paciente no experimenta una MS.

A pesar del progreso en la comprensión de los mecanismos detrás de la manifestación de Brugada, actualmente no existen indicadores fiables para determinar con precisión el riesgo de MS en un individuo. Algunos investigadores han desarrollado algoritmos predictivos basados en el uso de señales de ECG de 10 segundos de visitas de control [3]. Sin embargo, estos pueden no ser suficientes para detectar el SB, ya que la mayoría de los pacientes no muestran el patrón electrocardiográfico en un período de tiempo tan reducido. El objetivo de este trabajo es el desarrollo de una herramienta de clasificación supervisada para la estratificación de pacientes con SB a partir del análisis de señales electrocardiográficas de alta resolución y otras variables clínicas.

Tabla 1. Conjuntos de datos obtenidos aplicando distintos algoritmos para encontrar el dataset optimizado para el entrenamiento de los modelos.

Datos	Descripción
Sin datos clínicos	
Dataset 0	Todos los biomarcadores del ECG
Dataset 1	Biomarcadores del ECG comunes entre los 3 tests
Dataset 3	5 mejores biomarcadores del ECG para cada derivación obtenido con correlación de Pearson
Con datos clínicos	
Dataset 2	Biomarcadores del ECG comunes entre los 3 tests y todos los datos clínicos
Dataset 4	5 mejores biomarcadores del ECG para cada derivación obtenidos con correlación de Pearson
Dataset 5	30 mejores variables obtenidas con Regresión Univariante
Dataset 6	30 mejores variables obtenidas con árboles de decisión

Tabla 2. Parámetros empleados para la optimización de parámetros (GridSearch) en cada uno de los modelos con el mejor dataset.

Modelo	VARIABLES
SVM	C= [0.01,0.1,10,100] , gamma=[0.1,0.001,1]
RF	n estimators= [100, 300, 500, 700], max depth = [None, 5, 10]
KNN	k =[1, raíz cuadrada del número de muestras]
XGBoost	eta = [0.01, 0.1, 0.3], max depth=[3, 6, 9], n estimators= [100, 300, 600]

2. Métodos

2.1. Materiales

El estudio incluyó señales de ECG del holters de 12 derivaciones de 24 horas de 64 sujetos con SB, contando con 54 pacientes asintomáticos y 10 pacientes sintomáticos. Además, se agregaron datos clínicos de los pacientes tales como la edad de diagnóstico, sexo, antecedentes documentados de MS en la familia, antecedentes documentados de fibrilación ventricular, patrón ECG de tipo 1 espontáneo, entre otros.

2.2. Preprocesado

En primer lugar, las señales se preprocesaron con MATLAB y se dividieron en segmentos de 3 minutos para realizar la delineación con un algoritmo llamado ECG-kit [4]. Este paso fue necesario para extraer los puntos fiduciales del ECG y obtener todos los latidos de cada segmento de 3 minutos. Seguidamente, las señales fueron filtradas, para lo cual se aplicó una técnica de promedio de señal, como se muestra en la Figura 1. Este proceso consistió en la obtención del latido promedio aplicando la media sobre todos los latidos de cada segmento. La señal promedio se comparó con cada latido individual del segmento de 3 minutos para obtener la similitud entre ellos. La similitud se evaluó realizando la correlación cruzada y se obtuvo un valor de desalineación entre los picos R de la señal promediada y cada latido. Se corrigieron las señales con valores de desalineación inferiores a la duración del

QRS y se rechazaron el resto.

Tras el filtrado, todos los latidos promediados de 1 segundo se delinearono para extraer los puntos fiduciales, incluyendo el pico, el inicio y el final de la onda P, el complejo QRS y la onda T. Luego, se calcularon 10 biomarcadores: desviación del ST en el punto J y en el punto J + 60 ms, pendiente del ST, intervalo PR, intervalo QT, intervalo QT corregido, potencia media del QRS (Pavg), valor absoluto del área del QRS (areaQRSabs) y variables de potenciales tardíos (LP) como la duración del QRS (QRSd) y la amplitud de la señal RMS en los últimos 40 ms (RMS40) del complejo QRS filtrado.

2.3. Clasificación supervisada

Después del preprocesado, se realizó un análisis estadístico de los biomarcadores, basado en la aplicación de la prueba t de Student, para estudiar las diferencias estadísticas entre pacientes asintomáticos y sintomáticos. Por otro lado, los biomarcadores obtenidos y los datos clínicos de cada paciente se utilizaron para entrenar varios modelos de clasificación con el fin de optimizar un modelo, en Python, para la predicción de pacientes sintomáticos y asintomáticos. Para este propósito, se probaron 4 modelos, incluyendo Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest (RF) y Extreme Gradient Boosting (XGBoost). Para entrenar los modelos, se probaron siete conjuntos de datos diferentes obtenidos mediante la combinación de biomarcadores derivados del ECG y datos clínicos aplicando los siguientes algoritmos: regresión univariante, árboles de decisión y correlación de

Tabla 3. Mejor subconjunto derivado de las 20 variables más significativas del test ANOVA.

Variabes	Descripción
Pavg aVL	Potencia promedio del complejo QRS en aVL
ST 0 I	Amplitud del punto J en I
AreaQRSabs aVL	Área absoluta del complejo QRS en aVL
ST 60 V3	Amplitud del punto J +60 ms en V3
ST slope V5	Pendiente del segmento ST en V5
ST 60 aVR	Amplitud del punto J +60 ms en aVR
ST 0 V5	Amplitud del punto J en V5
LP QRSd V3	Duración del QRS a partir de Potenciales Tardíos en V3
AreaQRSabs III	Área absoluta del complejo QRS en III
QTc V6	Duración corregida del QRS en V6
AreaQRSabs V2	Área absoluta del complejo QRS en V2
LP QRSd V1	Duración del QRS a partir de Potenciales Tardíos en V1
LP QRSd V4	Duración del QRS a partir de Potenciales Tardíos en V4
ST slope II	Pendiente del segmento ST en II
ST 60 V1	Amplitud del punto J +60 ms en V1

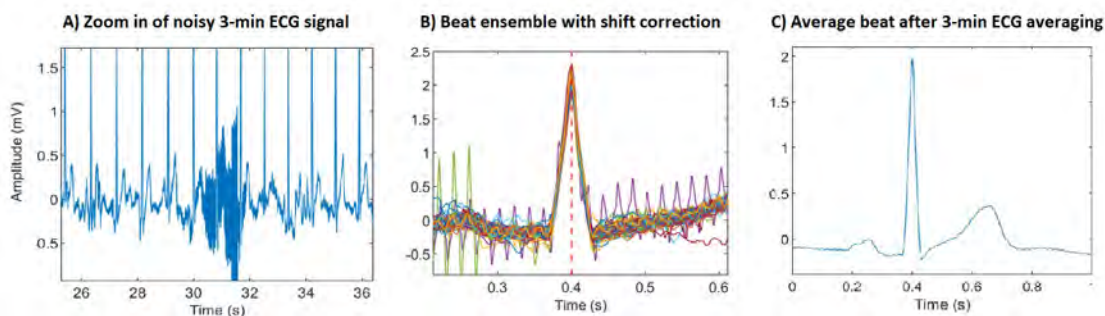


Figura 1. A) Intervalo de un segmento de ECG ruidoso de 3 minutos. B) Conjunto de latidos sin corrección de desalineación. C) Latido promedio obtenido.

Pearson. Los distintos conjuntos de datos se detallan en la Tabla 1.

Estos subconjuntos se probaron en cada uno de los 4 modelos. Para evaluar el rendimiento de cada modelo en los 7 subconjuntos, se aplicó una validación cruzada estratificada. Tras ello, el mejor conjunto de datos se dividió en entrenamiento (75%) y los modelos se optimizaron utilizando ajuste de hiperparámetros con Grid Search con los parámetros de la Tabla 2.

3. Resultados

Los resultados del preprocesado mostraron una reducción del ruido después de la aplicación del algoritmo de promediado. En el caso de segmentos ruidosos, el algoritmo de promediado mostró resultados mejorados, respecto a algoritmos sin corrección, al preservar mejor la amplitud de la señal original. Estos resultados se pueden ver en la Figura 1.

La evaluación de los mejores conjuntos de datos mostró

que el mejor conjunto de datos era el conjunto de datos 5, obtenido con el algoritmo de prueba ANOVA y las 20 variables más significativas, con mejores resultados en la clasificación. Este conjunto de datos incluyó las siguientes 4 variables clínicas, MS recuperada, antecedentes familiares de MS antes de los 55 años, la presencia de eventos previos de fibrilación auricular (FA) y la condición de probando (caso index en la familia). Por otro lado, el conjunto de datos incluyó 16 biomarcadores de ECG, y donde aproximadamente la mitad están relacionados con la elevación del ST. Las variables electrocardiográficas más significativas se muestran en la Tabla 3.

Los resultados del análisis mostraron que algunos biomarcadores del ECG, como la potencia promedio del complejo QRS o el área del QRS, fueron significativamente más altos en pacientes sintomáticos en algunas derivaciones precordiales como V1 ($p < 0,001$). Los resultados se pueden ver en la Tabla 4. Los mejores datos de clasificación se obtuvieron con RF y XGBoost. Las matrices de confusión mostraron resultados de sensibili-

Tabla 4. Análisis estadístico de los pacientes sintomáticos (n= 4089 latidos) y asintomáticos (n= 22215 latidos).

Variabes	Sintomáticos	Asintomáticos	P valor
Pavg V2 (mV^2)	0.143 ± 0.146	0.119 ± 0.126	p < 0.001
AreaQRSabs V2 ($\mu V \cdot s$)	32.78 ± 24.28	24.68 ± 13.36	p < 0.001
ST slope V1 ($\mu V/ms$)	-1.92 ± 1.89	-1.41 ± 1.81	p < 0.001
PR V1 (s)	0.205 ± 0.065	0.185 ± 0.048	p < 0.001
LP QRSd V1 (ms)	135.1 ± 45.2	115.2 ± 37.27	p < 0.001

Tabla 5. Resultados finales de rendimiento para la versión más optimizada de cada modelo.

Modelo	Conjunto de datos	Sensitividad	AUC
KNN	Dataset 5 (ANOVA)	31.60%	0.651
SVM	Dataset 1 (19 ECG-biomarkers)	27.22%	0.708
XGBoost	Subset 20 (ANOVA)	58.38%	0.888
RF	Subset 20 (ANOVA)	54.30%	0.907

dad (tasa de verdaderos positivos) del 46,6 % y del 47,07 % y resultados de especificidad (tasa de verdaderos negativos) del 99,85 % y del 46,6 %. Además, los resultados, que se muestran en la Tabla 5 y se obtuvieron después de la validación cruzada aplicada a Random Forest y XGBoost, mostraron un AUC de 0,907 y 0,888 y una sensibilidad del 54,30% y del 58,38 %, respectivamente.

4. Discusión y Conclusión

En este estudio, se ha evaluado la efectividad de modelos de clasificación supervisada entrenado con biomarcadores derivados del ECG y datos clínicos para diferenciar entre la condición sintomática y asintomática en el SB. En cuanto al preprocesamiento de los datos, se encontró que el algoritmo de promediado logra reducir considerablemente el ruido, se preservó mejor la amplitud de la señal original y se redujo la pérdida de información en el complejo QRS.

En lo que respecta a la clasificación, los mejores modelos fueron Random Forest y XGBoost, con un AUC de 0.907 y 0.888 respectivamente. A pesar de estos resultados positivos, los modelos presentaron sensibilidades del 53,3% y el 58,38%. Los modelos mostraron capacidades óptimas para detectar pacientes asintomáticos (especificidad > 80% en todos los casos), pero una identificación deficiente de pacientes sintomáticos (sensibilidad aproximadamente del 50%), lo cual se debe principalmente al desequilibrio entre las clases que no se pudo mejorar con la aplicación de modelos ponderados. A pesar de que el análisis estadístico mostró diferencias entre pacientes asintomáticos y sintomáticos en algunas variables, los modelos de clasificación no lograron encontrar estas diferencias.

En relación a los conjuntos de datos probados, el conjunto de datos óptimo fue el conjunto de Datos 5, específicamente en las 20 variables más significativas de este

subconjunto. Los 4 predictores clínicos refuerzan algunos de los hallazgos descritos en la literatura, como la presencia de eventos previos de fibrilación auricular y la MS recuperada como indicadores de riesgo. Además, la presencia de variables ST entre los predictores optimizados podría estar relacionada con el patrón del SB, caracterizado por una elevación descendente del segmento ST.

En conclusión, estos modelos, basados en el uso de biomarcadores extraídos de señales promediadas de holters de 12 derivaciones y 24 horas, tiene potencial de ser efectivos en la predicción de pacientes con riesgo de MS. De hecho, la consistencia clínica de las variables que ofrecen el mejor rendimiento alienta a evaluar estos modelos con conjuntos de datos más grandes y equilibrados. Se están llevando a cabo otros estudios, aumentando el número de casos con síntomas y preservando las características electrocardiográficas mediante el uso de datos de latido a latido.

Referenc

- [1] Brugada J, *et al.* Present status of brugada syndrome. *Journal of the American College of Cardiology* 8 2018;72:1046–1059. ISSN 07351097.
- [2] Hernandez-Ojeda J, *et al.* Patients with brugada syndrome and implanted cardioverter-defibrillators. *Journal of the American College of Cardiology* 10 2017;70:1991–2002. ISSN 07351097.
- [3] Liu CM, *et al.* A deep learning-enabled electrocardiogram model for the identification of a rare inherited arrhythmia: Brugada syndrome. *Canadian Journal of Cardiology* 2 2022; 38:152–159. ISSN 0828282X.
- [4] Demski AJ, *et al.* ecg-kit: a matlab toolbox for cardiovascular signal processing. *Journal of Open Research Software* 4 2016;4:8. ISSN 2049-9647.

Contacto: Flavio Palmieri, UPC CREB, Av. Diagonal, 647, 08028, Barcelona (España). flavio.palmieri@upc.edu

Autómata celular para simulación de fibrilación auricular: acercando el tiempo de cómputo a la práctica clínica

G. S. Romitti, A. Liberos, M. Termenón-Rivas, J. Barrios-Álvarez de Arcaya, P. Romero, D. Serra, I. García-Fernández, M. Lozano, R. Sebastian, M. Rodrigo

CoMMLab, Departament d'Informàtica, Departament d'Enginyeria Electrònica, Universitat de València, València, España
{giada.romitti, alejandro.liberos, miguel.rodrigo}@uv.es

Resumen

Las simulaciones biofísicas auriculares conllevan un alto coste computacional asociado al gran número de ecuaciones que deben resolverse en cada nodo en modelos con una densa discretización espacial, lo que dificulta su uso clínico. El Autómata Celular (AC), al tomar un número finito de estados, reduce el tiempo computacional y puede proporcionar aproximaciones interesantes para modelos de gemelos digitales. Este estudio explora el potencial de un AC que simula la electrofisiología auricular en condiciones sanas y patológicas, comparándolo con las simulaciones biofísicas.

El AC fue entrenado a partir de simulaciones biofísicas con diferentes ritmos de estimulación en láminas de tejido. Luego, el AC entrenado se utilizó para simular patrones de fibrilación auricular en una lámina 2D de tejido auricular y en una geometría auricular 3D completa, pudiendo así evaluar su rendimiento con respecto a las simulaciones biofísicas.

En el primer caso, la arritmia autosostenida bajo diferentes grados de remodelado electrofisiológico presentó un error promedio 6.25 ± 2.87 ms en la duración del ciclo. En la geometría tridimensional, el valor de los tiempos locales de activación difirió en 9.6 ± 7.6 ms para el ritmo sinusal y proporcionó patrones comparables para la fibrilación auricular crónica. Todos los resultados se acompañan de una reducción en los tiempos computacionales que se acercan a tiempo real (21 veces más rápido que las simulaciones biofísicas en equipos con GPU). Estos hallazgos sugieren que los modelos de AC tienen el potencial de proporcionar una herramienta eficiente para reproducir la electrofisiología auricular específica de cada paciente en tiempos clínicos.

1. Introducción

Entre las arritmias cardíacas, la fibrilación auricular (FA) es la más prevalente [1]. Desde principios del siglo XX, los mecanismos que subyacen a la iniciación y el mantenimiento de la FA han sido objeto de un debate continuo, ya que aún no se comprenden completamente. Por lo tanto, la elección de un tratamiento específico sigue siendo un problema clínico importante, lo que hace que los tratamientos farmacológicos y quirúrgicos sean sub-óptimos [2].

La tecnología de gemelos digitales tiene el potencial de perfeccionar la medicina personalizada y mejorar los resultados de los pacientes al crear una herramienta virtual que fusiona datos clínicos personales con modelos mecanicistas y estadísticos. En el contexto de la FA, pueden predecir el desarrollo de la enfermedad ayudando

a diseñar tratamientos personalizados, monitorear su efectividad y la respuesta del paciente [3].

Las simulaciones de gemelos digitales cardíacos se basan principalmente en modelos biofísicos [4], que reproducen el comportamiento electrofisiológico con gran precisión utilizando docenas de expresiones de canales iónicos que requieren alta resolución temporal, lo que resulta en un alto costo computacional. Por lo tanto, es necesario preguntarse si esta alta precisión es necesaria, especialmente si se piensa utilizar esos modelos en diagnósticos cotidianos. En esta dirección, se han desarrollado Autómatas Celulares (AC) para describir sistemas complejos mediante un conjunto discreto de estados [5]. Evidencias recientes [6] presentan simulaciones de condiciones sanas y patológicas en geometrías ventriculares reales utilizando AC con dos estados de excitación únicos dependientes de propiedades de restitución. Este enfoque ha logrado reducir los tiempos de computación.

El objetivo de este trabajo es validar un AC capaz de simular la actividad electrofisiológica auricular en tiempos asumibles en la práctica clínica. Para lograr esto, se realizaron primero simulaciones biofísicas para obtener las propiedades de restitución necesarias para el AC. Luego, se comparó el AC con simulaciones biofísicas en escenarios realistas sanos y en fibrilación en dominios 2D y geometría auricular 3D completa.

2. Materiales y Métodos

2.1. Modelo biofísico

Se realizaron simulaciones biofísicas utilizando el modelo auricular de Koivumäki [7]. Estas se resolvieron utilizando una combinación de los métodos directos de Euler y Rush-Larsen, y fueron implementadas en una plataforma de GPU [8].

Los dominios de las simulaciones son tres. El primero, utilizado para entrenar el modelo, es una lámina rectangular de tejido auricular ($0,3 \times 2 \times 0,025$ cm, 2106 células, malla hexaédrica, distancia entre nodos de 0,25 mm) activada desde la parte inferior frontal, ver Fig. 1a. Sobre esta geometría, se realizaron un total de 3000 protocolos de estimulación S1-S2 (S1 aleatorio $\in \{100, \dots, 1000\}$ ms repetido 15/16 veces antes de un S2 $\in \{100, \dots, 1000\}$ ms). Esto se replicó bajo diferentes porcentajes

de remodelado eléctrico ($\{0, 33, 66, 100, 135\}$ %), ver Fig. 1c. Un porcentaje de remodelado cero representa condiciones sanas, mientras que el 100% de remodelado representa el sustrato auricular de pacientes con fibrilación auricular crónica (cAF). Este último fue creado modificando el modelo de ritmo sinusal presentado en [7]: expresión de SERCA (-16%), relación PLB/SERC (+18%), relación SLN/SERCA (-40%), INCX máximo (+50%), sensibilidad de RyR a $[Ca^{2+}]SR$ (+100%), conductancia de I_{CaL} (-59%), conductancia de I_{to} (-44%), conductancia de I_{Kur} (-22%) y conductancia de I_{K1} (+100%). Cabe destacar que las modificaciones en I_{to} , I_{Kur} e I_{K1} se han alterado del modelo original de cAF [7]. La discretización temporal se estableció en 20 μs para todas las simulaciones.

A partir de los resultados de estas simulaciones, se obtuvieron los valores de la duración del potencial de acción (APD), la velocidad de conducción (CV) y el intervalo diastólico (DI), ver Fig. 1b. El APD se midió como la diferencia entre el tiempo de activación (instante de máxima dV/dt positiva) y el momento en el que se alcanza el 90% de la repolarización, el llamado APD_{90} . APD^0 y APD^{+1} denotan las últimas activaciones de S1 y S2, respectivamente. Los valores adquiridos se utilizaron para caracterizar las propiedades de restitución del modelo.

Con el objetivo de realizar una validación independiente del autómata celular, se realizaron simulaciones biofísicas en una lámina cuadrada de tejido auricular (7,0 x 7,0 x 0,03 cm, 109512 células, malla hexaédrica, distancia entre nodos de 0,30 mm, ver Fig. 3a-3b) y en un modelo 3D volumétrico de las aurículas humanas construido a partir de imágenes médicas [9] (284578 vértices, 1353783 tetraedros, distancia entre nodos de 673,38 μm). Este modelo 3D incorporó tejido anisótropo con dirección de fibras [9] con valores de velocidad de conducción transversal y longitudinal para las distintas regiones auriculares: aurícula derecha (AD), aurícula izquierda (AI), nodo sinoauricular (NSA), crista terminalis (CT), músculos pectíneos (MP), haz de Bachmann (HB), orejuela auricular derecha (OAD), orejuela auricular izquierda (OAI), vena cava inferior (VCI), vena cava superior (VCS), istmo cavotricuspidé (ICT), válvula mitral (VM), válvula tricúspide (VT), venas pulmonares (VP) y fosa oval (FO). En ambos dominios, se simuló reentradas funcionales a partir de la interacción entre frentes de onda consecutivos provenientes de diferentes regiones, que crearon una reentrada auto-sostenida [1], en los que se calcularon los valores de la duración del ciclo (CL). Para el caso del modelo 3D, se realizaron simulaciones de ritmo sinusal mediante la activación a 1Hz del nodo sinoauricular y se midió el tiempo de activación local (LAT).

2.2. Autómata celular

Se propuso un enfoque alternativo para modelar la electrofisiología del tejido cardíaco en un estudio anterior [6, 10]. Este enfoque utilizó un autómata celular (AC) extendido espacialmente, basado en eventos y

asincrónico, diseñado específicamente para analizar tejidos y geometrías cardíacas.

El AC considerado utiliza dos estados principales: 0 (inactivo o repolarizado) y 1 (activo o capaz de activar a los vecinos). Para ser activado, un nodo debe ser estimulado por sus vecinos. Esta estimulación sigue el método de *Fast Marching* [11], que alcanza a los vecinos con la onda plana que se propaga a la velocidad de conducción (CV). Posteriormente, un nodo finaliza su activación cuando ha transcurrido, desde que fue activado, un intervalo de tiempo igual a su APD asignado.

El APD, así como el valor de CV, deben calcularse para cada nueva activación. Para ello, se utilizan las propiedades de restitución extraídas de simulaciones biofísicas. Cabe destacar que estas propiedades se obtienen solo una vez para cada tipo de sustrato y luego serán válidas para cada geometría simulada con el AC.

El AC se implementó en Processing [12], que utiliza el lenguaje Java, y se puede utilizar en cualquier tipo de computadora sin necesidad de paralelización. Las simulaciones realizadas con el AC utilizan el mismo dominio y condiciones de estimulación que las realizadas con el modelo biofísico, con el fin de realizar un análisis por comparación.

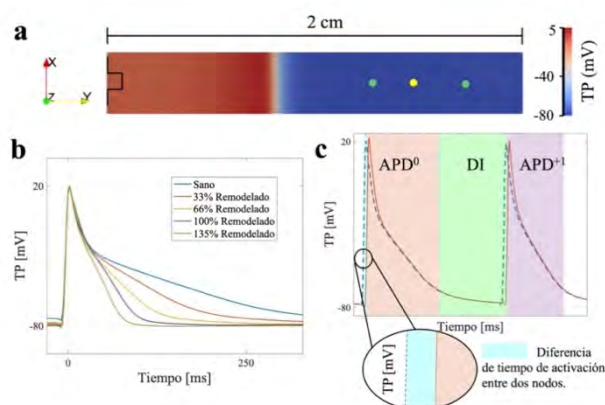


Figura 1. a) Modelo biofísico 2D utilizado para entrenamiento del AC. Los nodos amarillo y verde se utilizaron para medir la APD y el CV, respectivamente; b) Representación de curva de potencial transmembrana en diferentes grados de remodelado eléctrico; c) Biomarcadores extraídos: Duración del potencial de acción anterior y siguiente (APD^0 , rojo; APD^{+1} , morado), Intervalo diastólico (DI, verde), retraso en activación (azul).

3. Resultados

3.1. Caracterización del modelo biofísico

A partir de las simulaciones realizadas en la lámina de tejido auricular, se extrajeron los valores experimentales de APD, DI y CV. Estos deben resumirse en funciones numéricas que los relacionen entre sí.

Un análisis de la dependencia del APD^{+1} con las activaciones anteriores mostró que un ajuste con una sola curva basada en el APD^0 produjo un mayor error (caso de remodelado del 33%: 5.45 ± 3.48 ms) en comparación con considerar un ajuste de superficie que incluye el APD y el

DI anteriores, Fig. 2a (1.83 ± 1.47 ms, $p=0.02$, Fig. 2b). Esto no se observó en el caso del CV (Fig. 2c): ajustar una sola curva produjo un error igual al obtenido al considerar un ajuste de superficie. Por esta razón, se decidió considerar una superficie de restitución para los valores de APD y una curva de restitución para los valores de CV.

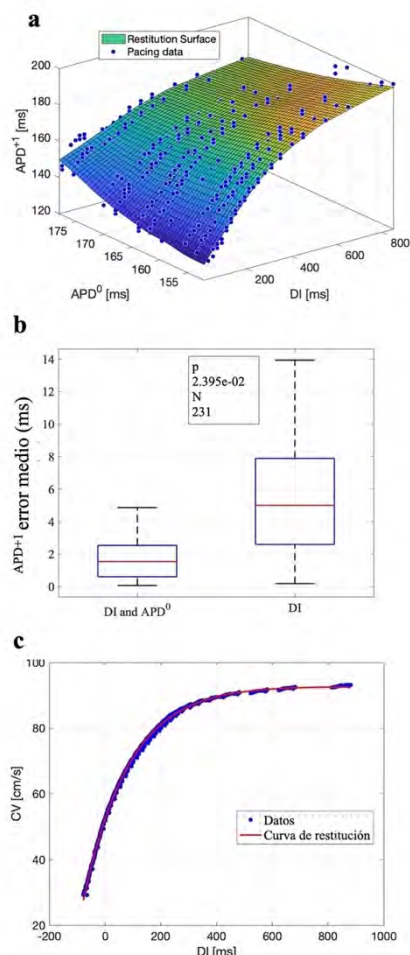


Figura 2. Caso remodelado 33%. a) Superficie de restitución del APD; b) Error medio en la estimación del APD; c) Curva de restitución de la CV.

3.2. Validación del autómata celular

Las simulaciones de ritmos complejos realizadas por el AC entrenado fueron comparadas con las simulaciones biofísicas correspondientes. Se consideraron tres casos para la comparación: arritmia autosostenida en la lámina cuadrada de tejido (Fig. 3a-3b), conducción del ritmo sinusal en la geometría 3D (Fig. 4a-4b) y arritmia autosostenida en la geometría 3D (Fig. 4c-4d).

Las simulaciones de arritmia autosostenida en la lámina de tejido se repitieron para valores de remodelado del {33, 66, 100, 135} %. En todos los casos, el estímulo S2 provocó una rotura del frente de onda que se estabilizó como una reentrada funcional. Al comparar los valores de CL entre la simulación biofísica y el AC, se observó que la diferencia promedio entre los valores obtenidos fue de 6.25 ± 2.87 ms (<5%), ver Fig. 3c, así como el mismo patrón de propagación del rotor, ver Fig. 3a-3b.

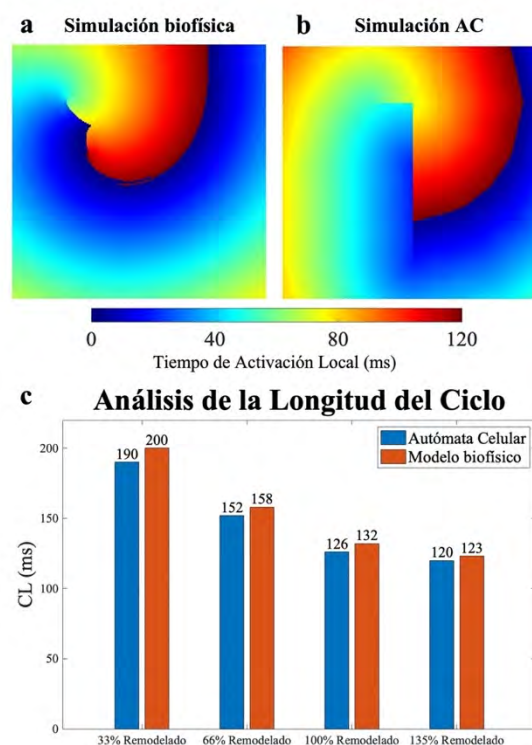


Figura 3. a) Reentrada funcional simulada con 100% de remodelado y modelo biofísico; b) Reentrada funcional simulada con 100% de remodelado y AC; c) análisis de las diferencias entre los CL del rotor obtenidos con AC y modelo biofísico a diferentes niveles de remodelado.

En cuanto a la geometría completa en 3D, las primeras simulaciones se realizaron en un modelo sano (0% de remodelado, anisotropía incluida como variación en los valores de difusión transversal y longitudinal, y con dirección de fibras), produciendo una activación sinusal. El tiempo de despolarización se observó prácticamente igual al del modelo biofísico correspondiente (110 ms, Fig. 4a-b). En promedio, la diferencia entre la simulación del AC y la simulación biofísica en los tiempos de activación local de cada nodo difirió en 9.6 ± 7.6 ms. Finalmente, también se reprodujo una reentrada autosostenida en la geometría completa en condiciones de fibrilación auricular crónica (100% de remodelado), reproduciendo una reentrada en forma de 8. En este caso, el AC reprodujo un comportamiento de reentrada similar, como se muestra en la Fig. 4c-4d, con ligeras diferencias en los patrones de propagación en la aurícula derecha.

Además, se encontró una fuerte evidencia de la disminución del tiempo de cálculo. Ambos simuladores se ejecutaron en un ordenador con un procesador Intel® Core™ i7-13700H, 14 núcleos, frecuencia de 3,70GHz a 5,00GHz turbo máximo, 24 MB Intel® Smart Cache, 16GB de RAM, ejecutando Linux versión 6.2.0-32-generic. El AC fue 21 veces más rápido que el solver biofísico equivalente: 17 frente a 360 segundos de tiempo de cálculo por segundo de tiempo simulado, respectivamente. Es importante enfatizar que este resultado se obtuvo utilizando un único procesador CPU para el AC y una GPU de 3840 núcleos para las simulaciones biofísicas.

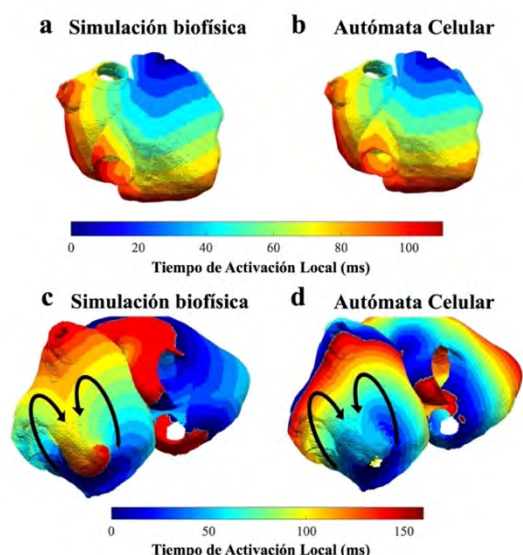


Figura 4. Simulación de un caso sano con a) modelo biofísico y b) AC. Reentrada funcional simulada con c) modelo biofísico y d) AC. Representación del tiempo de activación local.

4. Discusión y Conclusiones

Este estudio describe un autómata celular capaz de simular la actividad electrofisiológica auricular en condiciones patológicas en un tiempo computacional reducido. El AC fue entrenado utilizando el modelo atrial de Koivumäki en una pequeña porción rectangular de tejido auricular bajo condiciones de estimulación controlada. Luego, el AC fue validado frente a simulaciones biofísicas en un patrón de reentrada en 2D y finalmente en una geometría auricular tridimensional completa incluyendo dirección de fibras.

Al analizar la dependencia tanto de la velocidad de conducción (CV) como del APD^{+1} en el efecto de memoria a corto plazo de la simulación (APD^0), se encontró que el APD^{+1} variaba en función del APD^0 mientras que la CV permanecía invariable. Por lo tanto, se consideró el APD^{+1} como una función del intervalo diastólico (DI) y del APD^0 , mientras que la CV se consideró como una función únicamente del DI. Esto permitió crear propiedades de restitución adecuadas para el AC, lo que redujo el error en el valor estimado de APD^{+1} .

Se llevaron a cabo simulaciones de escenarios de reentrada para aproximarse a situaciones de fibrilación auricular, y los resultados se analizaron cualitativa y cuantitativamente para el tejido auricular cuadrado. El AC reprodujo el patrón de onda de activación obtenido con un modelo biofísico, tanto en el tejido auricular como en el dominio auricular completo. Además, las diferencias en cuanto a la duración del ciclo (CL) fueron consistentes con el nivel de remodelado y con un error aceptable en términos clínicos (< 5%). Aún mostrando el mismo fenómeno fibrilatorio (rotor o figura en ocho), se observaron diferencias locales en las simulaciones de la geometría auricular completa. Futuras investigaciones con múltiples geometrías y patrones de

estimulación deberán evaluar si estas diferencias suponen algún impedimento para el uso clínico de esta herramienta.

Los resultados del estudio indican que es posible entrenar un AC mediante la simulación de un modelo biofísico en un dominio reducido, y luego utilizar directamente el AC para simulaciones que impliquen un mayor número de nodos y comportamientos más complejos, logrando así simulaciones útiles con un tiempo de cálculo reducido.

Agradecimientos

Este trabajo ha sido financiado por la Generalitat Valenciana Subvención AICO/2021/318 (Consolidables 2021) y Subvención PID2020-114291RB-I00 financiada por MCIN/ 10.13039/501100011033 y por "ERDF A way of making Europe".

Referencias

- [1] Jalife J. Basic cardiac electrophysiology for the clinician. Wiley-Blackwell, 2009.
- [2] Krummen DE, Swarup V, Narayan SM. The role of rotors in atrial fibrillation. *Journal of thoracic disease*, vol 7(2), 142, 2015.
- [3] Corral-Acero J., Margara F., Marciniak M., Rodero C., Loncaric F., Feng Y., ... Lamata P. The 'Digital Twin' to enable the vision of precision cardiology. *European heart journal*, vol 41(48), 2020, pp 4556-4564.
- [4] Hodgkin AL, Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, vol 117(4), 500, 1952.
- [5] Von Neumann J, Burks AW. Theory of self-reproducing automata. *IEEE Transactions on Neural Networks*, vol 5(1), 1966, pp 3-14.
- [6] Serra D, Romero P, Garcia-Fernandez I, Lozano M, Liberos A, Rodrigo M, ... & Sebastian R. (2022). An automata-based cardiac electrophysiology simulator to assess arrhythmia inducibility. *Mathematics*, vol 10(8), 2022.
- [7] Koivumäki JT, Seemann G, Maleckar MM, & Tavi, P. (2014). In silico screening of the key cellular remodeling targets in chronic atrial fibrillation. *PLoS computational biology*, 10(5), e1003620.
- [8] García-Mollá VM, Liberos A, Vidal A, Guillem MS, Millet J, Gonzalez A, ... & Climent AM. Adaptive step ODE algorithms for the 3D simulation of electric heart activity with graphics processing units. *Computers in biology and medicine*, vol 44, 2014, pp 15-26.
- [9] Krueger MW, Schmidt V, Tobón C, Weber FM, Lorenz C, Keller DU, ... & Dössel O. Modeling atrial fiber orientation in patient-specific geometries: a semi-automatic rule-based approach. In *Functional Imaging and Modeling of the Heart: 6th International Conference, FIMH 2011*, New York City, NY, USA, May 25-27, 2011. Proceedings 6 (pp. 223-232). Springer Berlin Heidelberg.
- [10] Romitti GS, Romero P, Liberos A, Serra D, García-Fernandez I, Lozano M, ... & Rodrigo M. The Effects of Long-and Short-term Memory on Action Potential Duration for Atrial Cellular Automata. In *2022 Computing in Cardiology (CinC)* (Vol. 498, pp. 1-4). IEEE.
- [11] Ethian JA. Fast marching methods. *SAIM Review*, vol 41, 1999, pp 199-235.
- [12] Reas C, Fry B. Processing: a programming handbook for visual designers and artist. *Mit Press*, 2007.

Importancia del Pre-entrenamiento en el Desarrollo de una Red Neuronal Convolutiva Ligera para la Evaluación de la Calidad del ECG

A. Huerta Herraiz¹, A. Martínez Rodrigo¹, J.J. Rieta Ibáñez², R. Alcaraz Martínez¹

¹ Grupo de Investigación en Ingeniería Electrónica, Biomédica y de Telecomunicación, Universidad de Castilla-La Mancha, Cuenca, España, {alvaro.huerta, arturo.martinez, raul.alcaraz}@uclm.es

² BioMIT.org, Dep. de Ingeniería Electrónica, Universitat Politècnica de València, Valencia, España, jjrieta@upv.es

Resumen

La utilización de dispositivos portátiles para la adquisición de ECG de larga duración es cada vez mayor. Sin embargo, una señal de ECG tan extensa puede registrar ruido o interferencias, por lo que es crucial una evaluación precisa de su calidad. Recientemente se ha estandarizado el uso de redes neuronales convolucionales (RNCs) para este cometido, mostrando un rendimiento muy prometedor, sin embargo, requieren muchos recursos limitando así su uso en sistemas portátiles. En este trabajo se propone explorar la capacidad de las RNC ligeras, las cuales requieren muchos menos parámetros que las tradicionales RNCs pre-entrenadas. Así pues, en este trabajo se comparó el rendimiento de una RNC en dos circunstancias diferentes, entrenada desde cero y pre-entrenada sobre un conjunto de imágenes naturales con 70 clases diferentes. Además, como referencia, también se evaluó el rendimiento del conocido modelo GoogLeNet. Todas las RNCs se sometieron a un ajuste fino con un conjunto balanceado de 20.000 intervalos de ECG de 5 segundos de duración y se validaron externamente. Los resultados mostraron que GoogLeNet funcionó ligeramente mejor que la RNC ligera pre-entrenada. Sin embargo, los modelos ligeros fueron unas 12 veces más rápidos a la hora de clasificar cada intervalo de ECG. Estos resultados ponen de manifiesto la utilidad de pre-entrenar una RNC utilizando imágenes naturales, mejorando así la capacidad de generalización, manteniendo un rendimiento comparable al de redes mucho más profundas a la hora de evaluar la calidad del ECG, pero reduciendo notablemente el coste computacional.

1. Introducción

Tradicionalmente se utiliza el electrocardiograma (ECG) en reposo como la manera más común de monitorizar enfermedades cardiovasculares (ECVs) en entornos hospitalarios. Sin embargo, puede no ser suficiente si se pretende diagnosticar ECVs con una naturaleza intermitente y aleatoria, requiriendo una monitorización de manera continua [1]. Así pues, para realizar este cometido se propone la utilización de dispositivos vestibles o portátiles, los cuales permiten capturar la señal de ECG durante muy largos periodos de tiempo. Esto supone un avance muy importan-

te a la hora de detectar ECVs como la fibrilación auricular paroxística (FA), esto es de especial importancia ya que es la ECV más frecuente y con una gran prevalencia, causando una elevada morbilidad e incluso puede desencadenar en ictus isquémico [2].

A pesar de las enormes posibilidades que presentan este tipo de dispositivos, el hecho de monitorizar a los pacientes mientras realizan sus actividades cotidianas puede mermar la calidad general del ECG [3]. Por lo tanto, para evitar que estas perturbaciones puedan afectar al posterior diagnóstico de ECVs, es necesario un proceso previo de evaluación de la calidad de la señal de ECG. Existen varios tipos de ruidos que pueden estropear la grabación, ya sean inherentes al sistema de adquisición (línea de base, interferencias de red eléctrica, ruido muscular) o que provengan del exterior (artefactos de movimiento o pérdida de contacto con el electrodo entre otros) [4]. De ahí que junto al avance de la tecnología, sea necesario desarrollar sistemas capaces de evaluar la calidad de los distintos tramos de la señal de ECG.

El principal problema que presentan estos sistemas es la gran cantidad de datos registrados, haciendo inviable la visualización por parte de facultativos, ya que eso requiere de una gran cantidad de recursos humanos. Por este motivo, no son pocos los trabajos que se han centrado en evaluar de manera automática y desatendida la calidad del ECG de larga duración. En ese sentido, los sistemas basados en inteligencia artificial son la tendencia más utilizada actualmente, experimentando un gran avance en los últimos tiempos. Existen métodos previos basados en la extracción de puntos fiduciales o características morfológicas de los registros de ECG para su posterior análisis mediante técnicas de aprendizaje automático [5], sin embargo, los algoritmos recientes de aprendizaje profundo han superado esos enfoques. Se ha demostrado que estos modelos tienen una mayor capacidad de generalización, de ahí que presenten una mayor habilidad para distinguir entre segmentos de ECG de alta y baja calidad. Además, otra ventaja es que no es necesario un procesamiento de los datos pues se pueden tratar los datos en bruto [6].

En este sentido, las redes neuronales convolucionales (RNCs) se están estableciendo como la metodología más utilizada en diversas tareas como clasificación, regresión o detección. El problema es la enorme cantidad de datos necesarios para entrenar una RNC desde cero, por lo que para tratar este asunto se suele utilizar la técnica conocida como *transferencia de aprendizaje*, lo que significa que se puede aplicar el conocimiento previo de una red en una nueva tarea [7]. De esta manera, se han popularizado algunos modelos de RNCs, como GoogLeNet [8], compuesto por millones de parámetros y una estructura con ramas en paralelo. Por este motivo, el trabajo propuesto pretende investigar el rendimiento de las RNCs en función del pre-entrenamiento a la hora de evaluar la calidad de las señales ECG [9]. Además, se desarrolló una estructura RNC ligera con un número muy limitado de capas y parámetros, para poder evaluar su rendimiento en comparación con GoogLeNet como arquitectura de referencia.

2. Materiales y métodos

2.1. Bases de datos

El material utilizado para el desarrollo del presente trabajo queda dividido en dos grupos. En primer lugar, el conjunto de datos utilizado para entrenar desde cero la RNC ligera, el cual se basa en el conjunto original de ImageNet [10] (compuesto por 1.000 clases diferentes de objetos) pero limitado a 70 clases y 500 muestras de cada clase, alcanzando un total de 35.000 muestras. Debido a la limitación de recursos computacionales y de tiempo necesarios para entrenar una RNC desde cero, se seleccionó esta opción como la mejor solución de compromiso entre conjunto de datos y recursos disponibles.

Por otro lado, el segundo conjunto de datos corresponde a las grabaciones de ECG. Con el fin de garantizar la capacidad de generalización de la red y evitar el sobreentrenamiento, se utilizaron dos bases de datos de ECG con sistemas de adquisición muy diferentes, es decir, una base de datos propia (PDB) y otra base de datos pública perteneciente al PhysioNet/CinC Challenge 2017 (PC2017DB) [11]. Para la adquisición de la PDB se utilizó un dispositivo textil cuya frecuencia de muestreo es de 250 Hz con 12 bits de resolución. El proceso de etiquetado de los diferentes segmentos se llevó a cabo por dos expertos, diferenciando los intervalos de ECG de alta y baja calidad, marcando como ruidosos aquellos segmentos en los que todos los picos R no se podían diferenciar claramente. Este conjunto de datos se compone de 20.000 intervalos de ECG de 5 segundos de duración divididos de manera balanceada entre los dos grupos. Cabe destacar que dentro de la clase de alta calidad se incluyeron tanto episodios de ritmo sinusal normal (RSN) como de FA. Esta primera base de datos se utilizó para realizar el ajuste fino de los tres modelos de RNC bajo estudio.

Por último, de la base de datos PC2017DB se obtuvieron los segmentos de ECG utilizados para la realización de los test de cada uno de los experimentos llevados a cabo. A diferencia del dispositivo anterior, en este caso no

es vestible, sino un aparato portátil que registra la actividad eléctrica del corazón utilizando las yemas de los dedos cuando el paciente lo solicita. La frecuencia de muestreo es de 300 Hz con una resolución de 16 bits. El conjunto de datos se compone de cuatro clases diferentes, tales como, FA, RSN, otros ritmos (pertenecientes a alta calidad) y ruido. Este conjunto se compone de 48.607 intervalos de ECG de 5 segundos de longitud con las dos clases totalmente desbalanceadas, siendo 47.439 segmentos pertenecientes a la clase de alta calidad y 1.168 fragmentos pertenecientes a la clase de baja calidad.

2.2. Transformación del ECG

Como referencia se tomaron trabajos previos que utilizaban RNCs para la clasificación de la calidad de los diferentes segmentos de ECG de larga duración [6, 12]. Bajo este enfoque, las RNCs estaban diseñadas para aceptar imágenes como entrada, por lo que las estructuras propuestas y GoogLeNet tendrán entradas comunes. Por lo tanto, es necesario un paso previo que transforme cada uno de los segmentos de ECG en imágenes. Así pues, las grabaciones en bruto (sin preprocesar) de ECG se segmentaron en intervalos de 5 segundos de longitud y se convirtieron en imágenes utilizando la Transformada Wavelet Continua (TWC) aplicando a la transformación un mapa de color [13]. Se optó por esta transformación debido a que la TWC se ha mostrado muy útil en trabajos previos puesto que realiza las características de señales cuasiperiódicas no estacionarias, como lo es el ECG [13]. El desarrollo en detalle de la TWC utilizada queda definido en [6].

2.3. Arquitectura de GoogLeNet

Cada vez existen más modelos de RNCs pre-entrenadas y disponibles para su utilización de manera libre y gratuita, aprovechando el concepto de transferencia de aprendizaje, GoogLeNet es una de las más extendidas y utilizadas [8]. Se hizo famosa tras ganar una competición internacional muy importante (ImageNet Challenge 2014 [10]). Dentro de sus principales características presenta la capacidad de distinguir entre 1.000 clases de imágenes diferentes, además de añadir una novedad en su diseño, estando formada por varias ramas en paralelo, a diferencia de lo que se había hecho hasta el momento ya que solían ser diseños apilando exclusivamente una serie de capas en serie. Con este enfoque se consigue aumentar la profundidad reduciendo el coste computacional. La parte principal del diseño se conoce como *módulo Inception*, y es un bloque compuesto por núcleos de diferentes tamaños (1×1 , 3×3 , 5×5) colocados en paralelo. De esta manera, se puede extraer la información espacial perteneciente a diferentes niveles. En su estructura, GoogLeNet cuenta con nueve módulos Inception, además de dos capas convolucionales, cuatro capas de agrupación máxima, tres capas de agrupación media, cinco capas totalmente conectadas y tres capas con la función *Softmax*. Además, también incluye funciones de activación (ReLU) o regularizaciones dentro de su arquitectura. Finalmente, GoogLeNet se caracteriza por una profundidad de 22 capas con capacidad de aprendizaje y un total de 6,7 millones de parámetros.

2.4. Arquitectura propuesta

Como punto de partida es necesario especificar lo que se podría considerar como RNC ligera, para ello, se toma como referencia uno de los diseños disponible famoso por su ligereza y conocido como *SqueezeNet*, el cual contiene alrededor de 1,24 millones de parámetros [14]. Por lo tanto, el modelo propuesto en el presente trabajo se compone de cuatro capas con capacidad de aprendizaje, donde tres son convolucionales, y la última es una capa totalmente conectada. La primera capa convolucional tiene una dimensión espacial de 7×7 y 128 núcleos. La segunda capa convolucional reduce el tamaño de la primera y utiliza filtros de dimensión 4×4 y 128 núcleos. La última capa convolucional se compone de 384 núcleos de un tamaño de 3×3 y 256 núcleos. Después, la última capa con parámetros con capacidad de aprendizaje es la capa totalmente conectada, la cual consta de 70 neuronas, que se corresponden con el número de clases capaces de identificar. Además, a las capas convolucionales le siguen las funciones de activación ReLU y una capa de agrupación máxima para reducir las dimensiones espaciales del mapa de características. Tras la capa totalmente conectada, un clasificador Softmax se encarga de asignar las probabilidades de pertenecer a cada una de las clases. En la Tabla 1 se presenta un análisis detallado de la arquitectura al completo junto con las activaciones y parámetros con capacidad de aprendizaje. Finalmente, el modelo de red ligera propuesta suma un total de 1.221.574 parámetros con una profundidad de 4 capas.

Capa	Tipo	Activaciones	Parámetros
Entrada	Entrada de imagen	$227 \times 227 \times 3$	
Conv. 1	Convolución 2-D	$45 \times 45 \times 128$	18,944
ReLU 1	Función de activación	$45 \times 45 \times 128$	
AM 1	Agrupación máx. 2-D	$22 \times 22 \times 128$	
Conv. 2	Convolución 2-D	$10 \times 10 \times 128$	262,272
ReLU 2	Función de activación	$10 \times 10 \times 128$	
AM 2	Agrupación máx. 2-D	$9 \times 9 \times 128$	
Conv. 3	Convolución 2-D	$7 \times 7 \times 256$	295,168
ReLU 3	Función de activación	$7 \times 7 \times 256$	
AM 3	Agrupación máx. 2-D	$6 \times 6 \times 256$	
TC	Totalmente conectada	$1 \times 1 \times 70$	645,190
Softmax	Softmax	$1 \times 1 \times 70$	
Clasif.	Salida	$1 \times 1 \times 70$	

Tabla 1. Detalle de la arquitectura completa de la RNC ligera propuesta. Solamente cuatro capas tienen parámetros con capacidad de aprendizaje.

2.5. Desarrollo del experimento

Para realizar el presente experimento, se han considerado tres configuraciones diferentes: el modelo de RNC sin pre-entrenamiento; el mismo modelo de RNC entrenada con el conjunto de ImageNet 70; y GoogLeNet, entrenada con todo el conjunto de ImageNet. A cada uno de los tres modelos se le aplicó el mismo proceso de ajuste fino utilizando la base de datos PDB. Normalmente, para aplicar la técnica de transferencia de aprendizaje es necesario adaptar la arquitectura de las RNC para desempeñar la nueva tarea. Mientras que la red sin pre-entrenar se diseñó desde el inicio para distinguir solamente dos clases, la red ligera pre-entrenada y GoogLeNet necesitaron un pequeño ajuste en sus capas finales para diferenciar entre segmentos de

ECG de alta y baja calidad, pasando de 70 a 2 neuronas en el caso de la red ligera y de 1.000 a 2 neuronas en el caso de GoogLeNet. La configuración fue la misma en los tres experimentos realizados, repitiendo cada uno de los entrenamientos 10 veces, utilizando para ello un tamaño de lote de 32 imágenes y 10 épocas, obteniendo un promedio de los resultados utilizando PC2017DB para su evaluación.

3. Resultados

El conjunto de test presenta un desbalanceo muy notable entre el número de segmentos de ambas clases, siendo su gran mayoría (98 %) pertenecientes a la clase de alta calidad, mientras que sólo el 2 % restante pertenece a la clase de baja calidad. Por este motivo, además de las estadísticas típicas de sensibilidad (*Se*), especificidad (*Es*) y exactitud (*Ex*), se obtuvo la exactitud balanceada (*ExBa*). También se evaluó el tiempo de procesamiento de cada uno de los intervalos de ECG de 5 segundos. Mediante la utilización de una base de datos diferente a la de entrenamiento se consigue una validación robusta de los resultados, además, cada uno de las pruebas se repitió 10 veces y se promediaron los resultados. La Tabla 2 muestra dichos valores promedio obtenidos de cada experimento realizado junto con el tiempo de clasificación de los tres modelos basados en RNCs. Como se puede apreciar, GoogLeNet es la que mostró mejores métricas de rendimiento, siendo ligeramente superiores a las de la RNC ligera pre-entrenada. Por otro lado, se observó como los modelos ligeros fueron aproximadamente 12 veces más rápidos en clasificar cada fragmento de 5 segundos que GoogLeNet. Sin embargo, el modelo de RNC ligera con pesos vacíos funcionó sustancialmente peor que las dos anteriores. También se apreció que el desbalanceo entre los indicadores de *Se* y *Es* es del 5 % entre los modelos con pre-entrenamiento, mientras que esta diferencia se acentúa casi hasta el 16 % en el modelo sin pre-entrenamiento.

Modelo	Se	Es	Ex	ExBa	Tiempo(ms)
RNC ligera (Sin pesos)	0.776	0.934	0.780	0.855	0.65
RNC ligera (ImageNet 70)	0.836	0.884	0.836	0.860	0.64
GoogLeNet (ImageNet)	0.887	0.841	0.886	0.864	7.94

Tabla 2. Resultados de clasificación medios obtenidos por los tres modelos diferentes de RNCs junto con el tiempo de evaluación de cada intervalo de ECG de 5 segundos de duración.

4. Discusión

A la vista de los resultados, se puede afirmar que las redes pre-entrenadas utilizando un conjunto diverso de imágenes naturales mejoran su capacidad de generalización cuando se emplean para diferentes tareas de clasificación. En el caso de GoogLeNet a la hora de evaluar la calidad de la señal ECG, el rendimiento mejoró ligeramente con respecto a los resultados de la RNC ligera pre-entrenada. Por lo tanto, se precia una estrecha relación entre el conjunto de datos utilizado para el pre-entrenamiento y los resultados finales. Mientras que GoogLeNet se diseñó originalmente para distinguir entre 1.000 clases diferentes, el modelo de RNC ligera solamente se diseñó para discernir entre 70 clases. Por este motivo, GoogLeNet obtuvo mejores valores de *Se*, es-

to se debe a que la clase de alta calidad es más difícil de detectar que la clase de baja calidad. Por este motivo, la RNC ligera sin pre-entrenamiento tiene una menor capacidad de generalización, obteniendo mejores valores de E_s , ya que tiende a clasificar los segmentos de ECG como baja calidad debido a la aleatoriedad del ruido o artefactos que afectan a la señal de ECG, perdiendo capacidad de generalización. Además, otro hallazgo interesante observado es la capacidad de las RNCs ligeras, ya que pueden rendir a niveles similares al de redes más profundas. La arquitectura propuesta tiene tan sólo 4 capas, mientras que GoogLeNet tiene 22, del mismo modo, la estructura sugerida tiene 5 veces menos parámetros que GoogLeNet, lo que como era de esperar, significa un tiempo de procesamiento mucho menor (12 veces).

Aunque no es muy habitual encontrar en la literatura trabajos que hayan abordado problemas similares, los resultados obtenidos en estudios previos que utilizan arquitecturas de RNCs ligeras y evalúan su rendimiento con redes más complejas están en línea con la presente investigación. Por ejemplo, Garg y Singh compararon el rendimiento de un modelo de RNC ligera propia con arquitecturas de RNCs pre-entrenadas anteriores como son ResNet-50, Inception-V4, o MobileNet-V2 [15]. Consiguieron resultados de rendimiento similares entre las redes utilizadas, superando incluso a trabajos anteriores en el contexto de la clasificación del cáncer de mama, además de reducir el tiempo de ejecución al intervenir menos parámetros. Del mismo modo, Huang y Liao desarrollaron una estructura de RNC ligera basada en AlexNet y EfficientNetV2 destinada a detectar COVID-19 en imágenes de rayos X y TC [16]. Este algoritmo se probó en varios conjuntos de datos de imágenes, y obtuvieron valores de clasificación similares a los de trabajos anteriores utilizando una RNC con menos de 800.000 parámetros.

5. Conclusiones

Durante el presente trabajo se ha observado como el desarrollo y entrenamiento de RNCs ligeras desde cero para evaluar la calidad de los diferentes intervalos dentro de grabaciones de ECG de larga duración es una tarea muy compleja y que requiere mucho tiempo y recursos. Sin embargo, a la vista de los resultados se ha observado que es conveniente utilizar RNCs ligeras ya que presentan una buena relación entre velocidad y rendimiento en comparación con RNCs pre-entrenadas más complejas y profundas.

Agradecimientos

Este trabajo ha sido financiado por los proyectos de investigación competitiva PID2021-00X128525-IV0, PID2021-123804OB-I00 y TED2021-130935B-I00 pertenecientes al Gobierno de España 10.13039/501100011033 conjuntamente con el Fondo Europeo de Desarrollo Regional (EU), SBPLY/17/180501/000411 y SBPLY/21/180501/000186 de la Junta de Comunidades de Castilla-La Mancha, y AI-

CO/2021/286 de la Generalitat Valenciana, además de la compañía Daiichi Sankyo S.A.U.

Referencias

- [1] Sana F, Isselbacher EM, Singh JP, Heist EK, Pathik B, Armondas AA. Wearable devices for ambulatory cardiac monitoring: JACC state-of-the-art review. *Journal of the American College of Cardiology* 2020;75(13):1582–1592.
- [2] Lippi G, Sanchis-Gomar F, Cervellin G. Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge. *Int J Stroke* Feb 2021;16(2):217–221.
- [3] Smital L, Haider CR, Vitek M, et. al. Real-time quality assessment of long-term ECG signals recorded by wearables in free-living conditions. *IEEE Trans Biomed Eng* 2020; 67(10):2721–2734.
- [4] Nagai S, Anzai D, Wang J. Motion artefact removals for wearable ECG using stationary wavelet transform. *Health Technol Lett* 2017;4(4):138–141. ISSN 2053-3713.
- [5] Satija U, Ramkumar B, Manikandan MS. An automated ECG signal quality assessment method for unsupervised diagnostic systems. *Biocybern Biomed Eng* 2018;38:54–70.
- [6] Huerta Herraiz Á, Martínez-Rodrigo A, Bertomeu-González V, Quesada A, Rieta JJ, Alcaraz R. A deep learning approach for featureless robust quality assessment of intermittent atrial fibrillation recordings from portable and wearable devices. *Entropy* 2020;22(7):733.
- [7] Weiss K, Khoshgoftaar TM, Wang D. A survey of transfer learning. *Journal of Big data* 2016;3(1):1–40.
- [8] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015; 1–9.
- [9] Krishna ST, Kalluri HK. Deep learning and transfer learning approaches for image classification. *International Journal of Recent Technology and Engineering IJRTE* 2019; 7(5S4):427–432.
- [10] Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision IJCV* 2015;115(3):211–252.
- [11] Clifford GD, Liu C, Moody B, Lehman LWH, Silva I, Li Q, Johnson AE, Mark RG. AF classification from a short single lead ECG recording: The PhysioNet/Computing in Cardiology Challenge 2017. *Comput Cardiol* 2017;44.
- [12] Zhao Z, Liu C, Li Y, Li Y, Wang J, Lin BS, Li J. Noise rejection for wearable ECGs using modified frequency slice wavelet transform and convolutional neural networks. *IEEE Access* 2019;7:34060–34067.
- [13] Hong S, Zhou Y, Shang J, Xiao C, Sun J. Opportunities and challenges of deep learning methods for ECG data: A systematic review. *Comput Biol Med* 2020;122:103801.
- [14] Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and 0.5 mb model size. *arXiv preprint arXiv:160207360* 2016;.
- [15] Garg S, Singh P. Transfer learning based lightweight ensemble model for imbalanced breast cancer classification. *IEEEACM Transactions on Computational Biology and Bioinformatics* 2022;20(2):1529–1539.
- [16] Huang ML, Liao YC. A lightweight cnn-based network on COVID-19 detection using X-ray and CT images. *Computers in Biology and Medicine* 2022;146:105604.

Algoritmo de Sincronización del Electrocardiograma Para Análisis Retrospectivos Basados en Vectocardiograma

E. Ramírez Candela¹, R. T. Ors Quixal¹, S. Ruipérez-Campillo¹, R. Casado Arroyo², Francisco Castells¹, José Millet^{1,3}

¹ ITACA Institute, Universitat Politècnica de València, Valencia, Spain, elracan@upvnet.upv.es

² Electrophysiology Lab, Hôpital Erasme, Brussels, Belgium

³ Centro de Investigación Biomédica en Red Enfermedades Cardiovascular, Madrid, Spain.

Resumen

El Vectocardiograma (VCG) ofrece una representación tridimensional completa de la actividad eléctrica del corazón, lo que resulta de gran valor para el diagnóstico y tratamiento de afecciones cardiovasculares. El registro del VCG no se adquiere de manera común, pero puede reconstruirse a partir del electrocardiograma (ECG) a partir de transformaciones lineales. El formato de ECG clásico en papel o pdf registra doce derivaciones de manera intermitente cada 2.5 segundos y la derivación continua comúnmente la II. Esto supone un desafío para reconstruir el VCG, ya que los latidos en cada derivación ocurren en momentos temporales distintos.

El objetivo de esta investigación es proponer una metodología que permita sincronizar con precisión los latidos registrados para reconstruir el VCG y validarlo con un gran número de señales.

Los resultados demuestran una sincronización precisa, evidenciada por los valores de correlación de Pearson de $0,9840 \pm 0,0397$, un error absoluto medio (MAE) de $0,0133 \pm 0,0123$ mV y una raíz del error cuadrático medio (RMSE) de $0,0248 \pm 0,0214$ mV en la reconstrucción del VCG.

El método de sincronización propuesto permitirá la realización de estudios retrospectivos a partir del VCG en los que únicamente se dispone de registros convencionales de ECG.

1. Introducción

Aunque el registro de los ECG de manera digital es cada vez más común, existe una gran cantidad de señales registradas en papel. Estos registros en papel pueden poseer distintos formatos siendo el más común el uso de cuatro secciones de 2.5 segundos donde se registran tres derivaciones de manera simultánea y junto al registro continuo (10 segundos) de la derivación II [1].

El formato en papel o pdf presenta normalmente un problema a la hora de efectuar un análisis digital de la señal, de ahí que numerosos estudios se han dedicado a ofrecer soluciones precisas en el ámbito de la digitalización de este tipo de registros [2-4].

Por otro lado, el VCG es una técnica que ofrece información espacial acerca de la magnitud y dirección del vector cardiaco [5]. El VCG se obtiene a partir de las derivaciones de Frank [6], aunque actualmente en clínica el registro directo no es muy común. Sin embargo, esta representación tridimensional también es posible obtenerla a partir de un mapeo lineal a partir de las 12

derivaciones del ECG [7-8], siendo la Transformada Inversa de Dower la técnica más utilizada.

La información proporcionada por el VCG es de gran relevancia en la evaluación de ciertas patologías como el infarto de miocardio, isquemia e hipertrofia [5]. Sin embargo, el formato típico de registro asíncrono de cada una de las secciones impide la correcta reconstrucción del VCG. Es por ello que, previo a la transformación lineal, es necesaria una sincronización de los latidos. Hasta la fecha, únicamente Morales et al. [9] han proporcionado una solución para la sincronización previa a la reconstrucción del VCG. El método que proponen se basa en la alineación de cada derivación de 2.5 segundos basándose en correlación cruzada. Sin embargo, debido a que el intervalo RR en cada una de las derivaciones no es perfectamente homogéneo, los resultados que presentan no son completamente fiables.

Por tanto, existe una necesidad de desarrollar un método robusto de sincronización que permita la correcta reconstrucción del VCG y permita la realización de estudios retrospectivos haciendo uso de la representación tridimensional del vector cardiaco.

En este estudio se asume que el problema de digitalización ya ha sido satisfactoriamente abordado y se centra en la propuesta de un método que permita la sincronización precisa para la realización de estudios retrospectivos que requieran utilizar el VCG.

2. Materiales

Para el desarrollo y validación del método se ha utilizado de la base de datos pública PTB-XL ECG [10] disponible en Physionet [11]. Ésta contiene un gran número de registros, 21799 de 18869 pacientes, que permitirá la validación objetiva del método. Las señales tienen 12 derivaciones de 10 segundos y están disponibles a 100 y 500 Hz de frecuencia de muestreo. Un total de dos cardiólogos han anotado las señales utilizando cinco etiquetas: ECG normal, infarto de miocardio, cambios en ST/T, distorsión en la conducción e hipertrofia.

Del conjunto completo de señales, únicamente 125 señales tuvieron que ser descartadas por no mostraron al menos un latido de más de 0.6 segundos en alguna de las derivaciones.

3. Métodos

3.1. Fantoma

Con el fin de simular el formato de los registros de ECG en papel, se construyó un fantoma a partir de las señales completas (Figura 1) que actúa como entrada al método de sincronización. Dado que en este artículo únicamente se evaluará la sincronización, se usará directamente la señal digital asumiendo una digitalización idónea.

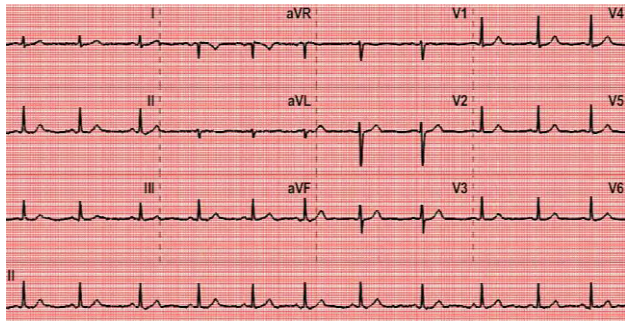


Figura 1. Ilustración del fantoma con el formato de un registro tradicional en papel.

El fantoma se genera a partir de la señal adquirida sincronamente durante 10 segundos. Éste se compone de cuatro secciones de 2.5 segundos: la primera contiene las derivaciones: I, II y III; la segunda de las aumentadas: aVR, aVL y aVF; la tercera de las tres primeras precordiales: V1, V2 y V3 y la cuarta de las tres precordiales restantes: V4, V5 y V6. Finalmente se representa en la última fila la derivación II en su versión completa de 10 segundos.

3.2. Referencia

Con el fin de evaluar el rendimiento del algoritmo, se ha seleccionado una sincronización y reconstrucción del

VCG de referencia con la que se compararán los resultados obtenidos.

Teniendo en cuenta que el VCG se construirá con latidos de derivaciones que pertenecen a instantes temporales distintos, cabe esperar que exista una variabilidad fisiológica entre ellos. Por tanto, la referencia se ha construido acorde a este hecho. Un promedio de todos los latidos del registro ofrecería un resultado que captura toda la variabilidad del registro; sin embargo, un promedio excesivo puede distorsionar la morfología del latido. Por ello se ha seleccionado el mínimo número de latidos que comprenda gran parte de la información de variabilidad del registro. Realizando la correlación entre el promedio de distinto número de latidos (N) y el promedio de todos los latidos se ha observado una tendencia no lineal (Figura 2). Situando un umbral en la correlación de 0.995 se obtiene que con 3 latidos ya se captura suficiente variabilidad.

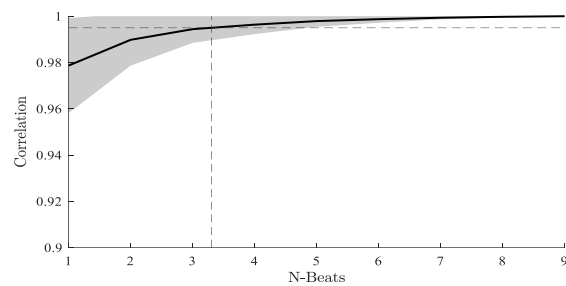


Figura 2. Correlación entre la media de N latidos y todos los latidos utilizando un umbral de selección en 0.995.

3.3. Algoritmo de Sincronización y Reconstrucción

El algoritmo de sincronización que se presenta (Figura 3), actúa a nivel de latido con el fin de evitar distorsiones en la alineación debido a la variabilidad del intervalo RR

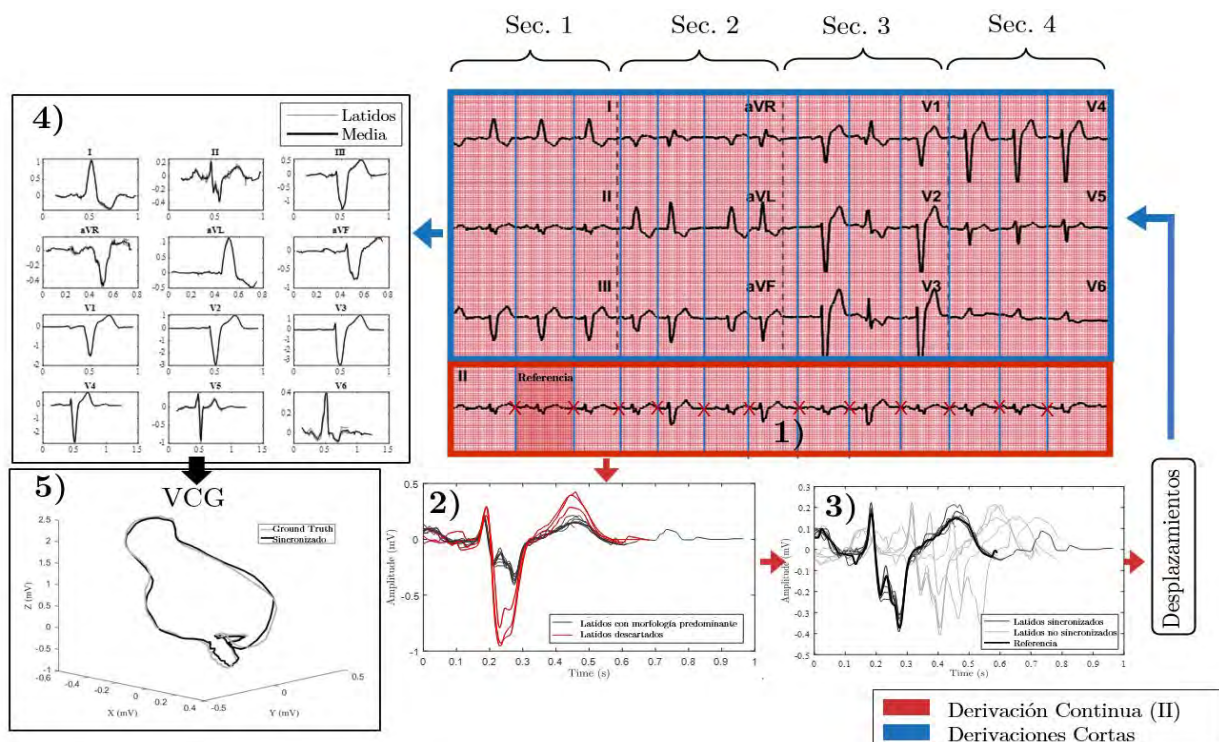


Figura 3. Ilustración del método de sincronización y reconstrucción propuesto.

presente en cada una de las secciones.

El método de sincronización diseñado se compone de las siguientes etapas:

- 1) Detección de latidos: Haciendo uso del software R-DECO [12], se han detectado los picos R de los latidos de la derivación continua (II). Se ha establecido un criterio para determinar los puntos de corte de los latidos. Éste se basa en la división del intervalo RR en 2/3 para el latido anterior y 1/3 para el posterior. Finalmente se selecciona un latido de referencia para la sincronización consistiendo en el primer latido completo de la derivación continua.
- 2) Descarte de latidos aberrantes: Para evitar sincronizaciones erróneas debido a la posible presencia de latidos aberrantes ya sean patológicos o debido a ruido, se han seleccionado los latidos con morfología predominante descartando aquellos cuya distancia Euclídea supere los 0.05 mV respecto del latido de referencia.
- 3) Alineamiento en la derivación continua: mediante el uso de la correlación cruzada se ha calculado el desplazamiento que se debe aplicar a cada latido para que esté perfectamente sincronizado.
- 4) Alineamiento en las 12 derivaciones: Una vez conocido los puntos de corte y el desplazamiento a realizar a cada latido, se aplica a cada segmento de latido mayor de 0.6 segundos del registro todas las derivaciones teniendo en este punto todos los latidos sincronizados.
- 5) Reconstrucción del VCG: a partir del promedio de latidos sincronizados de cada derivación, aplicando la Transformada Inversa de Dower [13] se reconstruye el VCG.

4. Resultados

4.1. Evaluación a nivel de ECG

En la Tabla 1 se recogen las métricas de sincronización tanto a nivel de morfología (Correlación de Pearson) como de magnitud (Error Cuadrático Medio, MAE; Raíz Cuadrada del Error Cuadrático Medio Normalizado, NRMSE). Se observa una correlación mayor de 90, un MAE menor de 2 mV y un NRMSE menor de 5% para todas las derivaciones siendo las derivaciones aumentadas (aVR, aVF y aVL) las que peor rendimiento arrojan.

Derivación	Correlación ($\times 10^{-2}$)	MAE (mV $\times 10^{-2}$)	NRMSE (mV $\times 10^{-2}$)
I	98.59±6.44	1.00±1.06	2.02±2.36
II	98.89±5.61	0.96±1.08	1.84±2.09
III	98.27±6.59	0.82±1.03	2.06±2.39
aVR	96.92±4.73	1.62±0.81	3.28±1.97
aVF	94.20±9.23	1.56±0.99	4.11±3.19
aVL	95.50±6.16	1.47±0.98	3.84±4.43
V1	97.35±5.91	1.78±1.40	2.98±2.33
V2	97.65±5.89	2.43±2.14	2.71±2.26
V3	97.34±6.01	2.56±2.57	2.74±2.21
V4	98.35±6.61	2.04±0.22	1.93±2.47
V5	98.47±5.75	1.85±0.16	1.97±2.79
V6	98.16±6.83	1.63±1.35	2.25±2.21

Tabla 1. Métricas para la evaluación de la sincronización a nivel de ECG ($\times 10^{-2}$): Correlación de Pearson, Error absoluto medio (MAE por sus siglas en inglés) y raíz cuadrada del error cuadrático medio normalizado (NRMSE por sus siglas en inglés).

Se ha realizado una comparación directa del método propuesto con el del estado del arte [9]. Para ello se ha replicado el modelo que se propone en [9] y se ha evaluado con la misma base de datos que el método que se expone en el presente estudio. En la Tabla 2 se observa un mayor rendimiento en el método propuesto en todas las métricas, en comparación con el método de Morales et al. [9].

Método	Correlación ($\times 10^{-2}$)	MAE (mV $\times 10^{-2}$)	NRMSE (mV $\times 10^{-2}$)
Morales	22.98±15.81	10.93±5.45	20.38±9.03
Propuesto	97.41±1.31	1.67±0.54	2.65±2.55

Tabla 2. Promedio de las métricas ($\times 10^{-2}$) para todas las derivaciones en la comparativa de los resultados presentados por el algoritmo de Morales et al.

4.2. Evaluación a nivel de VCG

Tras la reconstrucción del VCG a partir de la Transformada Inversa de Dower, se han extraído las métricas de correlación de Pearson ($\cdot 10^{-2}$), MAE ($\cdot 10^{-2}$ mV) y RMSE ($\cdot 10^{-2}$ mV) para todas las señales obteniéndose 98.15±4.26, 1.33±1.23 y 2.48±2.14 respectivamente.

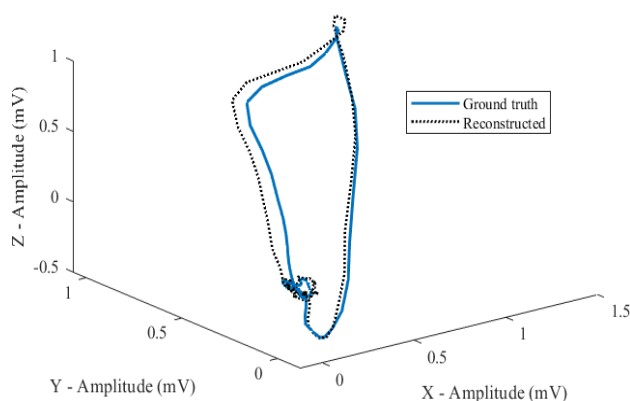


Figura 4. Representación tridimensional del VCG sincronizado y su correspondiente referencia con una correlación de Pearson de 98.61.

En la Figura 4, se muestra un ejemplo de reconstrucción del VCG cuyo valor de correlación es similar a la media de todas las señales.

5. Discusión

Los valores obtenidos en las métricas de correlación y error demuestran una estrecha similitud entre la señal y la referencia tanto a nivel de ECG como en su respectiva reconstrucción del VCG. Las pequeñas diferencias entre la sincronización y la referencia son debidas posiblemente a fuentes de error aleatorio pero que no presentan ninguna relevancia clínica.

Los resultados del método que se propone en este estudio superan en gran medida los obtenidos mediante el método del estado del arte [9] que actúa a nivel de derivación completa. La estrategia de abordar el problema a nivel de latido evita problemas asociados a la variabilidad en el intervalo RR o la heterogeneidad en el número de latidos entre secciones. Por tanto, esta comparativa ratifica a la adecuación de la sincronización a nivel de latido.

Este método de sincronización actúa como herramienta complementaria a los métodos de digitalización [2-4] pues permite al personal sanitario a emplear la información no solo de la señal ECG, sino que también del VCG proporcionando información complementaria para el diagnóstico y evaluación. Este algoritmo ha resultado ser el primero en la literatura que permite ofrecer la representación espacial del vector cardiaco con un alto nivel de fidelidad a partir de los registros en formato papel o pdf, estableciendo la posibilidad de realizar estudios retrospectivos a nivel de VCG que requieran señales en este tipo de formato.

6. Conclusión

El algoritmo de sincronización desarrollado ha resultado ser adecuado para la representación del vector cardiaco proporcionando resultados muy superiores al propuesto en la literatura y ha sido validado con una base de datos extensa. Esta reconstrucción fiable del VCG permite, por primera vez, la realización de estudios retrospectivos que requieran este tipo de formato.

Agradecimientos

Este trabajo ha sido parcialmente subvencionado por el PID2019-109547RB-I00 (Programa Nacional de Investigación, Ministerio de Ciencia e Innovación, Gobierno de España) y el PID2022-142514OB-I00 CIBERCV CB16/11/00486 (Instituto de Salud Carlos III).

Referencias

- [1] Goldberger JJ, Cain ME, Hohnloser SH, et al., American heart association/american college of cardiology foundation/heart rhythm society scientific statement on noninvasive risk stratification techniques for identifying patients at risk for sudden cardiac death. A scientific statement from the American heart association council on clinical cardiology committee on electrocardiography and arrhythmias and council on epidemiology and prevention, *J. Am. Coll. Cardiol.*, vol. 52, no. 14, pp. 1179–1199, Sep. 2008.
- [2] Millet-Roig J et al. A method capable of digitising every type of biosignal paper recordings as a tool for database improvement. *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Vol.20, Hong Kong, China, 1998, pp. 1238-1241 vol.3.
- [3] Gautam P, Sunkaria, RK and Sharma LD. Digitisation of paper-ECG using column wise-median approach. *Int. J. Comput. Appl. Technol.*, vol. 66, no. 2, 2022.
- [4] Baydoun M et al. High Precision Digitization of Paper-Based ECG Records: A Step Toward Machine Learning. *IEEE J Transl Eng Health Med*, vol. 7, 1900808, Nov. 2019.
- [5] Burger HC.. Heart and Vector: Physical Basis of Electrocardiography. *Gordon and Breach*; 1968.
- [6] Frank E, An accurate, clinically practical system for spatial vectorcardiography, *Circulation*, vol. 13, no. 5, pp. 737–749, 1956.
- [7] Edenbrandt L, Houston A, and Macfarlane PW. Vectorcardiograms synthesized from 12-lead ECGs: A new method applied in 1792 healthy children. *Pediatr. Cardiol.*, vol. 15, pp. 21–26, 1994.
- [8] Lingman M et al. Transient repolarization alterations dominate the initial phase of an acute anterior infarction—A vectorcardiography study, *J. Electrocardiol.*, vol. 47, pp. 478–485, 2014.
- [9] Morales E, Sevilla D, Pierluissi JH., Nazeran H. (2005). Digitization and Synchronization Method for Electrocardiogram Printouts. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society Conference*, 2, 1588-1591.
- [10] Wagner P, Strodthoff N, Boussejot R, Samek W, & Schaeffter, T. (2022). PTB-XL, a large publicly available electrocardiography dataset (version 1.0.3). *PhysioNet*, Jan, 2022.
- [11] Goldberger A, Amaral L, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* [Online], vol. 101, no. 23, pp. e215-e220, Jun, 2000.
- [12] Moeyersons J, Amoni M, et al. R-DECO: An open-source Matlab based graphical user interface for the detection and correction of R-peaks (version 1.0.0). *PhysioNet*.
- [13] Edenbrandt L, Pahlm O. Vectorcardiogram synthesized from a 12-lead ECG: Superiority of the inverse Dower matrix. *Journal of Electrocardiology*, vol. 21, no. 4, pp. 361-367, Apr. 1988. ISSN 0022-0736.

Clasificación EEG del estrés mental inducido por un *Serious Game* mediante *Deep Learning*

J.J. Gonzalez-Vazquez¹, L. Bernat¹, C. Romero¹, G.J. Garcia-Gomez¹, A. Ubeda¹

¹ Human Robotics Group, Universidad de Alicante
andres.ubeda@ua.es

Resumen

Este estudio propone un sistema de electroencefalografía (EEG) para la clasificación de diferentes niveles de estrés mental utilizando un Serious Game que consiste en esquivar obstáculos controlando un coche, donde en cada nivel de dificultad se aumenta el número de obstáculos presentes. Para ello, se ha medido el estrés tomando como referencia el nivel de dificultad jugado y se ha desarrollado un modelo de Deep Learning conformado por redes recurrentes (RNN), basado en unidades recurrentes cerradas (GRU). Los resultados muestran que el modelo es capaz de predecir los niveles de estrés con hasta una exactitud del 94 %, lo que sugiere que este sistema puede ser una técnica efectiva para predecir y clasificar el nivel de estrés mental.

1. Introducción

El estrés es de una de las condiciones más frecuentes en la sociedad contemporánea y su impacto en la salud y bienestar de las personas puede ser significativo. Cada individuo percibe y maneja el estrés en mayor o menor medida y un inadecuado manejo puede desencadenar estrés crónico, que está claramente relacionado con problemas de salud física y mental [1]. En los últimos años se han desarrollado modelos predictivos y de clasificación de estrés gracias a la inteligencia artificial (IA) y el aprendizaje automático (*machine learning*, ML) [2]. Una de las formas de evaluar este estado es a través de la medición de señales electroencefalográficas (EEG) en la superficie del cuero cabelludo, utilizando electrodos no invasivos [3]. Por ejemplo, se ha utilizado el agrupamiento *k-means* para clasificar los estados de estrés-relajación [4] y se ha evaluado cuantitativamente el estrés a través de la realidad virtual (VR) empleando técnicas de interpolación de datos y modelos individualizados de aprendizaje automático logrando una alta correlación (0.7-1) [5].

Recientemente, la irrupción del aprendizaje profundo (*Deep Learning*, DL) ha permitido mejorar la exactitud de estos sistemas. Un ejemplo es el estudio realizado por Pandey et al., que se centró en la detección de emociones, incluyendo el estrés a través de DL y clasificó las emociones en 2 grupos en función de la calma, tensión o tristeza de la emoción. Emplearon un modelo de red neuronal profunda (DNN) con una exactitud cercana al 60 % en los datos de prueba [6]. Finalmente, se destaca una revisión innovadora llevada a cabo por Xin Xu et al que trata la forma de clasificar y reducir el estrés utilizando DL combinado con *Transformers*, proporcionando resultados prometedores de hasta aproximadamente el 90 % de exactitud [7].

Sin embargo, la mayoría de estos estudios no utilizan

enfoques interactivos para medir el estrés mental, ya que muchos de ellos se basan en experiencias audiovisuales pasivas como método de monitorización del estrés. Por esta razón, este estudio propone un enfoque diferenciador e innovador basado en *Serious Games*. El objetivo es comprobar si este enfoque interactivo es efectivo para predecir y clasificar el estrés mental a partir de las señales EEG utilizando un modelo de redes recurrentes (RNN) con unidades recurrentes cerradas (GRU).

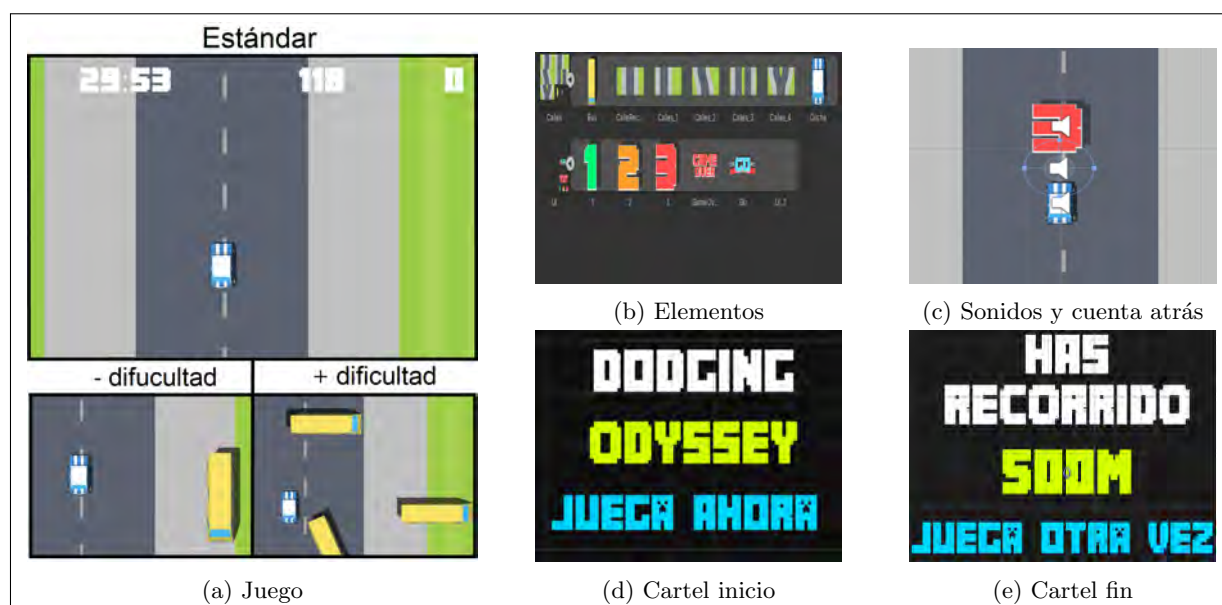
2. Materiales y métodos

2.1. Propuesta

El sistema consiste en un *Serious Game* basado en un juego de carreras enfocado en evitar obstáculos con un vehículo controlado mediante las teclas laterales del teclado del ordenador (ver Figura 1). El juego está dividido en 4 niveles de dificultad, diferenciados por el aumento del número de obstáculos a esquivar. El uso del juego como aplicación final se basa en la interactividad y atención que genera en los usuarios durante el juego. Se utiliza un dispositivo EEG de 8 canales como método de registro para capturar las señales cerebrales de un grupo de 19 usuarios, a partir del cual se obtiene un conjunto de datos de señales basado en el nivel jugado, la repetición, el canal y el usuario. El conjunto de datos pasa por una etapa de preprocesamiento para ingresar los datos óptimamente al modelo de clasificación. El objetivo es clasificar el estrés mental en 4 niveles de intensidad correlacionados con los 4 niveles de dificultad del *Serious Game*. El modelo de Aprendizaje Profundo consiste en redes neuronales recurrentes (RNNs) con una estructura de 7 capas basadas en unidades recurrentes de puertas (GRUs), donde su uso sobre las unidades de las redes neuronales convolucionales (CNNs) radica en la buena predisposición de los datos como serie temporal, aspecto óptimo como entrada de datos en las RNNs. El modelo se entrena con el conjunto de datos preprocesado para abordar el problema de clasificación multiclase del estrés mental en 4 niveles de intensidad.

2.2. Sistema EEG y comunicación

El sistema utilizado es el dispositivo EEG *Unicorn Hybrid Black*, el cual se encuentra conectado a un ordenador portátil que alberga la aplicación final del *Serious Game*, la cual ha sido programada en Unity. Asimismo, se ha desarrollado una API de registro específica para integrar el juego con el mencionado dispositivo de adquisición. *Unicorn Hybrid Black* es un dispositivo EEG que permite medir la actividad eléctrica cerebral en distintas áreas mediante el uso de 8 electrodos pasivos colocados siguiendo el Sistema Internacional 10/20 (Fz, C3, Cz,

Figura 1: *Serious Game*

C4, Pz, Po7, Oz, P08). Las señales EEG se muestrean a 24 bits y 250 Hz por canal, y se referencian a dos electrodos ubicados en los mastoides.

La integración del dispositivo EEG con el Serious Game se logra mediante la programación de una API que utiliza *sockets UDP* para permitir la comunicación entre ambas partes. El propósito principal de esta comunicación es enviar y recibir un elemento cuya lectura adecuada activará la función de grabación. Para ello, se ha implementado un cliente en el Serious Game, de manera que cuando el juego se inicia, se envía un mensaje al software de registro. La duración del registro está programada para ser exactamente de 90 segundos por nivel.

2.3. Protocolo experimental

Se han registrado las señales cerebrales de un grupo de 19 usuarios, compuesto por 15 hombres y 4 mujeres, con una edad promedio de 31 ± 14 años. Todos los participantes han firmado el consentimiento informado correspondiente, siguiendo el Protocolo del Comité de Ética (REF: UA-2023-02-08 2) de acuerdo con la Declaración de Helsinki. Se ha seguido un protocolo experimental en el que se establecen las posturas, descansos y tiempos a seguir. Durante la prueba, se registró a cada usuario por nivel y repetición, realizando 5 repeticiones para cada uno de los 4 niveles de dificultad del Serious Game, lo que resultó en 20 mediciones EEG de 90 segundos cada una para cada uno de los 19 participantes, generando así un registro EEG de 30 minutos por usuario.

2.4. Preprocesamiento de la señal EEG

Una vez capturadas las señales cerebrales de cada usuario, se procedió a establecer una ventana de preprocesamiento con el objetivo de eliminar el ruido, limpiar y transformar los datos de la señal EEG con el objetivo de para disponerlos de manera óptima para su poste-

rior clasificación. Para alcanzar este propósito, se han empleado una serie filtros de respuesta finita (*Finite Impulse Response, FIR*). Se ha aplicado un filtro paso bajo de 60 Hz y un filtro de paso alto de 4 Hz para eliminar cualquier tipo de artefacto motor y visual, incluyendo el parpadeo, la presión de teclas o movimientos corporales. Cabe recalcar, que la elección de 4 Hz es lo suficientemente elevada para continuar eliminando estos artefactos pese a que aumenten o disminuyan al cambiar el nivel de dificultad. Se ha implementado además un filtro *Notch* para eliminar la componente de 50 Hz originada por la interferencia de la red eléctrica. Adicionalmente, se aplica un filtro de promediado adaptativo para eliminar el ruido residual adaptándose a la señal de entrada y conservando las características relevantes de la misma.

Con el propósito de generar una cantidad significativa de datos para entrenar el modelo de DL, se ha procedido a dividir la señal EEG de los 8 canales en segmentos de 2 segundos con un solapamiento del 20%. La división en épocas permite generar datos uniformes con los que entrenar el modelo y el solapamiento permite aumentar las muestras de 400 a 500 datos por cada canal, en cada uno de los 8 canales muestreados en la repetición de cada nivel, lo que se traduce a decenas de miles de datos adicionales, donde su gran cantidad es necesaria para entrenar el modelo de DL. Posteriormente, los datos son etiquetados con un valor que varía de 0 a 3, correspondiendo al nivel de dificultad del nivel respectivo. La asignación específica de etiquetas es la siguiente: "0: Estrés bajo, 1: Estrés moderado, 2: Estrés intermedio, 3: Estrés alto". Por último, se han eliminado los segmentos con valores atípicos, es decir, aquellos que exceden el umbral absoluto de 75 mV, ya que sobrepasan el potencial de reposo de las neuronas [5]. Es relevante destacar que los modelos de aprendizaje profundo (DL) muestran un mejor rendimiento con datos normalizados en el rango de 0 a 1. En consecuencia, se

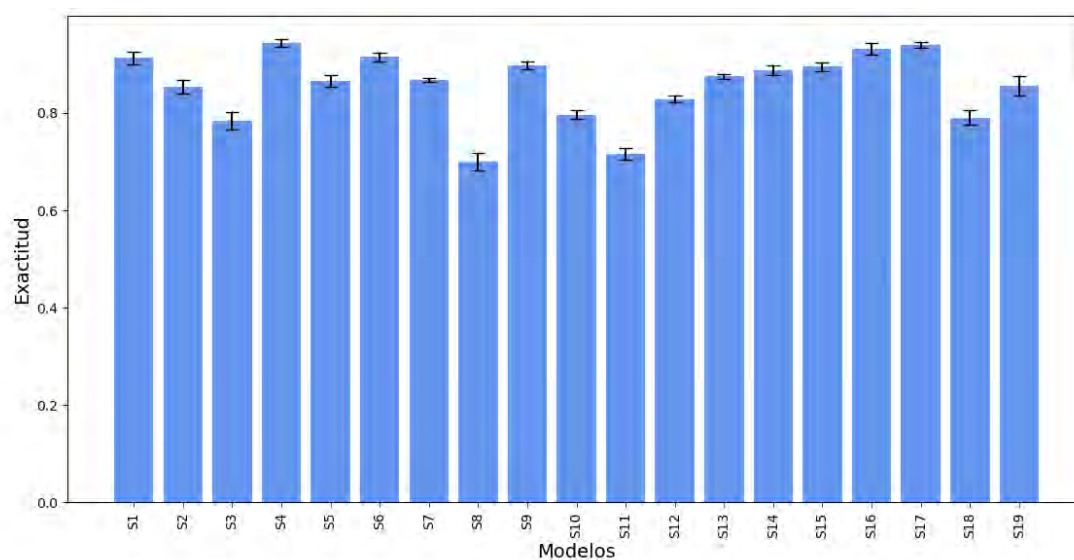


Figura 2: Exactitud y desviación estándar para cada modelo

ha llevado a cabo la normalización de los datos. Una vez que los datos han sido sometidos al preprocesamiento, el siguiente paso consiste en transformarlos de manera que sean adecuados para su entrada en el modelo. Para ello, los datos han sido transformados a una dimensión, lo que los hace aptos para modelos unidimensionales. Posteriormente, el conjunto de datos se convierte en un tensor y se divide en conjuntos de entrenamiento, prueba y validación. La distribución predeterminada es la siguiente: el 20 % de los datos totales se destina al conjunto de prueba (test), el 20 % del 80 % restante se utiliza para el conjunto de validación y el resto se asigna al conjunto de entrenamiento.

2.5. Clasificación multiclase del estrés mental

Para clasificar los distintos niveles de estrés mental se ha implementado un modelo de Red Neuronal Recurrente (RNN) basado en Unidades Recurrentes Cerradas (GRU). Para la construcción del modelo se ha implementado una capa de entrada junto a 4 capas GRU compuestas de 128 unidades a excepción de la última, con 256 unidades, seguidas de una capa de aplanamiento (flatten) y la capa de salida con la función de activación *softmax* esencial para el problema multiclase.

El aumento del número de unidades en la última capa, es debido al objetivo de capturar patrones más complejos y específicos en esta capa, aspecto que puede ayudar a detectar cambios concretos en las relaciones temporales de estrés. El hecho de implementar 4 capas es debido al objetivo de optimizar computacionalmente el entrenamiento del modelo, ya que el hecho de haber generado una gran cantidad de datos junto a la validación cruzada pueden ralentizar en gran medida el proceso de entrenamiento. Como hiperparámetros específicos se ha implementado un *learning rate* de 0.001 y un *batch size* de (32, 64 y 256) correspondiente a los modelos (individuales, intermedios y grandes).

En cuanto a la compilación del modelo, se ha imple-

mentado el optimizador *Adam* (Estimación del Momento Adaptativo). Este enfoque de optimización permite que la red se adapte a diferentes tasas de aprendizaje y ajuste el modelo al problema específico en cuestión. En relación con la función de pérdida, se ha optado por *Categorical Cross Entropy* debido a la naturaleza del problema, que implica la clasificación en múltiples clases. Como métrica de evaluación se ha empleado la exactitud (*accuracy*), que calcula la fracción de predicciones correctas en relación con el total de predicciones realizadas.

Por último, se han definido el número de épocas para el proceso iterativo, el tamaño del lote de datos previamente comentado y los datos de validación, manteniendo esos parámetros para todos los usuarios y obteniendo la exactitud para cada uno de ellos mediante validación cruzada.

3. Resultados y discusión

Los resultados obtenidos para cada usuario se muestran en la Figura 2. Se aprecia que los modelos S1, S4, S6, S9, S16 e S17 han alcanzado una exactitud superior al 90 %. Sin embargo, cuatro de ellos, S3, S7, S10 e S11, no logran una exactitud del 80 %. La variabilidad en estos resultados es relativamente baja y puede atribuirse a factores como el ruido durante las pruebas, los movimientos bruscos del usuario, una menor concentración en ciertos momentos durante el monitoreo o simplemente variaciones en la percepción del usuario sobre los niveles de estrés en comparación con la referencia establecida.

Se muestra también un ejemplo de gráfico de exactitud y pérdida para un usuario individual (Sujeto 17) (Figura 3). Al examinar el gráfico, se observa una disminución progresiva en la pérdida y un aumento en la exactitud a medida que aumenta el número de épocas. Esto sugiere que el modelo está extrayendo características y comenzando a clasificar el estrés correctamente. Es importante señalar que hay usuarios para los cuales la pérdida no

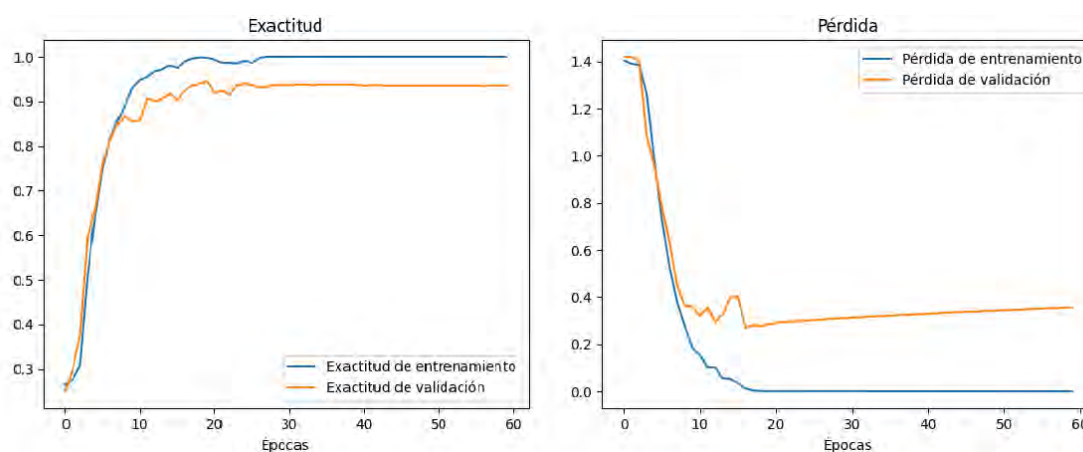


Figura 3: Métricas de evaluación del rendimiento del modelo

disminuye. Sin embargo, para un usuario individual, esto no tiene un gran impacto, ya que el modelo es capaz de clasificar correctamente el estrés en los datos de entrenamiento, los datos de validación y los datos de prueba no vistos.

Comparando con estudios anteriores, nuestros resultados muestran una exactitud superior al trabajo de Pandey et al. que apenas alcanza un 60% de acierto clasificando en 2 clases (valencia y excitación) y muestran que el DL es la técnica más efectiva en comparación con el ML en términos de desempeño durante la clasificación [6]. La revisión de Xu et al. se centra en el uso de Transformers con redes DL. Utilizan diferentes tipos de música para la relajación y tareas aritméticas para la inducción del estrés, clasificando 2 clases (estresado y no estresado) con 5 subniveles cada uno. Los resultados son comparables, con una exactitud ligeramente menor a nuestro estudio, alrededor del 62% y 92% como peor y mejor desempeño de cada modelo [7]. En el caso del estudio de Kalas et al, realizaron una investigación sobre detección de estrés utilizando agrupamiento *k-means* y un umbral para dividir el conjunto de datos en 2 clases [4]. No obstante, no emplearon una métrica de evaluación definida, por lo que es difícil comparar los resultados. Por último, el estudio reciente de Perez-Valero et al., muestra resultados muy altos de detección de cinco clases pero emplean una métrica de correlación (mayor que el 0.7) que está sujeta a limitaciones y que tampoco es comparable con nuestra propuesta [5]. En líneas futuras se propone adaptar cada modelo en tiempo real con un sesgo mayor de usuarios y aumentar la explicabilidad aplicando técnicas de interpretabilidad enfocadas a "Black-Box models".

4. Conclusiones

Se ha desarrollado un sistema basado en un Serious Game capaz de medir el estrés de una manera relativamente sencilla debido al control intuitivo que presenta el juego. Este enfoque representa un sistema distintivo en comparación con otros sistemas EEG utilizados para medir el estrés, los cuales involucran principalmente la aplicación de una experiencia visual y auditiva donde el

participante alterna entre estados de relajación y estrés. La diferencia radica en la interactividad del usuario con el juego, a diferencia de la pasividad de los sistemas tradicionales. Los resultados obtenidos muestran una exactitud del 94% superior a los estudios anteriores lo que permite prever un buen comportamiento en la futura generalización de los modelos y su evaluación en tiempo real para la implementación de sistemas reales en distintos ámbitos de aplicación.

Agradecimientos

Los autores quieren agradecer a todos los usuarios su participación en este estudio.

Referencias

- [1] Valdés, M., & De Flores, T. (1985). *Psicobiología del estrés*. Barcelona: Martínez Roca, 2.
- [2] Aggarwal, K., Mijwil, M. M., Al-Mistarehi, A. H., Alomari, S., Gök, M., Alaabdin, A. M. Z., & Abdulrhman, S. H. (2022). Has the future started? The current growth of artificial intelligence, machine learning, and deep learning. *Iraqi Journal for Computer Science and Mathematics*, 3(1), 115-123.
- [3] Palacios, L. (2002). Breve historia de la electroencefalografía. *Acta Neurol Colomb*, 18(2), 104-107.
- [4] Kalas, M. S., & Momin, B. F. (2016, March). Stress detection and reduction using EEG signals. In *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)* (pp. 471-475). IEEE.
- [5] Perez-Valero, E., Vaquero-Blasco, M. A., Lopez-Gordo, M. A., & Morillas, C. (2021). Quantitative assessment of stress through EEG during a virtual reality stress-relax session. *Frontiers in Computational Neuroscience*, 15, 684423.
- [6] Pandey, P., & Seeja, K. R. (2022). Subject independent emotion recognition from EEG using VMD and deep learning. *Journal of King Saud University-Computer and Information Sciences*, 34(5), 1730-1738.
- [7] Xu, X., Zhao, Y., Zhang, R., & Xu, T. (2022). Research on Stress Reduction Model Based on Transformer. *KSII Transactions on Internet and Information Systems (TIIS)*, 16(12), 3943-3959.

Diagramas de Recurrencia General para la Detección Automática de Apnea Obstruktiva del Sueño

D. Padovano¹, A. Martínez-Rodrigo¹, José M. Pastor¹, José J. Rieta², R. Alcaraz¹

¹ Grupo de Investigación en Ingeniería Electrónica, Biomédica y de Telecomunicación, Universidad de Castilla-La Mancha, España, danielle.padovano@uclm.es

² BioMIT.org, Departamento de Ingeniería Electrónica, Universitat Politècnica de Valencia, España

Resumen

La apnea obstructiva del sueño (AOS) es un síndrome respiratorio estrechamente relacionado con múltiples enfermedades cardiovasculares. La AOS se encuentra muy infradiagnosticada debido al coste prohibitivo de la polisomnografía, el estándar de oro para su detección. Ante este escenario, se han propuesto numerosas alternativas basadas en una derivación de ECG y la variabilidad del ritmo cardíaco (VRC), junto con diversas técnicas de aprendizaje automático (AA). Sin embargo, no existen muchos trabajos basados en el análisis de recurrencia de la VRC. Por ello, el presente estudio propone evaluar métodos tradicionales de AA y enfoques modernos de aprendizaje profundo (AP) para mejorar la detección de AOS por métodos alternativos a la PSG. En particular, se evaluará el diagrama de recurrencia general (DRG) de la VRC y se comparará su rendimiento de clasificación con su homónimo tradicional basado en el análisis de cuantificación de recurrencia (ACR). Asimismo, se emplearán bases de datos públicas del repositorio Physionet, con grabaciones de ECG divididas en intervalos de un minuto. Por último, a través de buenas prácticas de validación, se describirá y discutirá la forma en la que los modelos de AP con DRG superan a los modelos AA tradicionales basados en ACR en términos de exactitud, sensibilidad y especificidad.

1. Introducción

La apnea obstructiva del sueño (AOS) es un síndrome caracterizado por interrupciones repetitivas de la respiración durante el sueño [1]. Se estima que la prevalencia de este síndrome oscila entre el 9% y el 38% de la población general [2], aunque su estimación exacta resulta compleja debido a que se trata de una condición altamente infra-diagnosticada [3]. En los últimos años, se ha incrementado el número de estudios que relacionan la AOS con diversas enfermedades no comunicables [4], en especial modo, las enfermedades cardio-vasculares [5], revelando un riesgo tal que su detección temprana se ha convertido en un motivo de preocupación mayor a nivel socio-sanitario [6].

El estándar de oro para la detección de la AOS es la polisomnografía (PSG), una prueba que monitoriza las constantes vitales del paciente durante el sueño en un laboratorio especializado, cuyo coste suele estar limitado a zonas con grandes recursos o alto poder adquisitivo, dificultando así la detección temprana de esta condición [7]. Ante esta problemática, cientos de investigadores se han sumado a la exploración de nuevas vías de detección de AOS para paliar las limitaciones de la PSG [8]. En particular, los métodos basados en el análisis de la variabilidad del ritmo cardíaco (VRC), a partir de una derivación

de ECG, han encabezado el estado del arte [6]. La efectividad de la VRC radica en su capacidad para reflejar el comportamiento del sistema nervioso autónomo, encargado, entre otras funciones, de controlar la respiración durante el sueño [9].

En este contexto, existen diversos métodos alternativos a la PSG [6]. Muchos de ellos se enfocan en el dominio tiempo-frecuencia, utilizando herramientas como las transformadas de Fourier [10], periodogramas de Lomb-Scargle [11], o análisis por ondículas (wavelets) tanto a nivel discreto como continuo [12]. Sin embargo, existen muy pocos estudios basados en aspectos de recurrencia. Tales estudios se centran en caracterizar la VRC a través del análisis de cuantificación de recurrencia (ACR) [13, 14, 15], un enfoque no lineal que consiste en extraer características a partir un diagrama de recurrencia (DR) [16]. El DR presenta limitaciones en cuanto al umbral escogido para analizar el espacio de fases de la VRC, las cuales se pueden solventar mediante el uso de un diagrama de recurrencia general (DRG), que posibilita el análisis desde una perspectiva más holística [17]. Sin embargo, la investigación detrás de la sinergia entre el DRG y el AP en la detección de la AOS sigue siendo escasa, especialmente en el ámbito que ocupan los modelos basados en ECG de una sola derivación [6]. Por ende, esta brecha en la literatura resalta la necesidad de explorar los métodos basados en DRG y AP, ofreciendo una oportunidad única para impulsar nuevos avances en la detección de la AOS.

2. Materiales

Entre los materiales empleados para el presente estudio, se incluyen tres bases de datos públicamente accesibles del repositorio oficial de Physionet, concretamente: el *Computing in Cardiology 2000* (Apnea-ECG) [18], la base de datos polisomnográfica del MIT (MIT-BIH) [19], y la base de datos de la *University St. Vincent/University College de Dublín* (UCD-DB) [20].

La base de datos Apnea-ECG consta de 70 registros de ECG de 7 a 9 horas de duración, recopilados a partir de 30 sujetos varones y 5 mujeres, con edades comprendidas entre los 27 y 63 años. Esta base de datos incluye anotaciones minuto a minuto realizadas por expertos clínicos, categorizando cada minuto del registro de ECG como un episodio de apnea (etiquetado como A) o un episodio normal (etiquetado como N).

La base de datos MIT-BIH consta de 18 registros de PSG de 2 a 7 horas de duración, obtenidos de 16 su-

jetos varones con edades entre 32 y 56 años. Al igual que en la base de datos Apnea-ECG, expertos clínicos proporcionaron anotaciones sobre el ECG, pero estas se asignaron en periodos de 30 segundos.

Por último, la UCD-DB se compone de 25 registros completos de PSG, provenientes de 21 sujetos varones y 4 mujeres con edades entre 28 y 68 años. Las anotaciones de ECG de una sola derivación en tiempo real se realizaron siguiendo las reglas de *Rechtschaffen y Kales* [21], identificando varios eventos cardiorespiratorios además de episodios de apnea.

Finalmente, garantizar la uniformidad en los criterios de anotación, se adaptaron todas las bases de datos a la resolución temporal de la menos detallada, que es Apnea-ECG. Más concretamente, la base de datos MIT-BIH se etiquetó nuevamente al alinear las anotaciones con las etiquetas originales cada dos bloques de 30 segundos, mientras que la UCD-DB se etiquetó minuto a minuto. Este enfoque armonizó las anotaciones en todas las bases de datos.

3. Métodos

3.1. Procesado de la señal de ECG

Los registros de ECG se re-muestraron a 500 Hz para mejorar la detección de picos R. Después, para reducir el ruido, se aplicó un filtro Chebyshev de segundo orden con una frecuencia de corte de 100 Hz y se filtraron las interferencias de baja frecuencia con una frecuencia de corte de 0,5 Hz. Acto seguido, se empleó el algoritmo Pan-Tompkins para detectar los picos R [22], y se aplicó una ventana de corrección manual para corregir aquellos picos que pudieran haber sido omitidos o indebidamente colocados. Las grabaciones de ECG se segmentaron en intervalos de 1 minuto para el análisis de la VRC. Finalmente, se realizó una revisión manual de la señal para eliminar segmentos ruidosos, lo que resultó en más de 40000 segmentos de ECG de un minuto de duración.

3.2. Análisis de recurrencia

Un diagrama de recurrencia (DR) es una técnica de análisis de datos no lineal que visualiza las instancias en las que la trayectoria de un sistema dinámico vuelve a visitar el mismo área en el espacio de fases [17]. Un DR consiste en una matriz binaria con puntos blancos y negros basados en la distancia entre estados en la trayectoria del espacio de fases, en comparación con un umbral específico (ϵ). La serie temporal ($s(n) = s(1), s(2), \dots, s(N)$) se embebe en un espacio de m dimensiones utilizando el teorema del retardo temporal de Taken [23], lo que resulta en vectores de estado \vec{x}_i . Desde el punto de vista formal, el DR ($R(i, j)$) se puede definir como sigue [17]:

$$R(i, j) = \begin{cases} 1 & \text{si } \|\vec{x}_i - \vec{x}_j\| \leq \epsilon, \\ 0 & \text{en cualquier otro caso,} \end{cases} \quad (1)$$

El DR se representa visualmente con puntos negros para puntos recurrentes $\|\vec{x}_i - \vec{x}_j\| \leq \epsilon$ y puntos blancos en el caso contrario. A partir del DR ($\epsilon = 0,5$), se extrajeron características que constituyen el análisis de cuantificación de recurrencia (ACR) [24], incluyendo la tasa de recurrencia, el determinismo, la entropía de Shannon, la longitud promedio de las líneas diagonales y la divergencia [13].

Sin aplicar un umbral específico, la matriz de distancias (obtenida previamente a aplicar el umbral), también conocida como el diagrama de recurrencia general (DRG), ofrece una representación continua de la recurrencia en todos sus niveles. Finalmente, se emplearon los algoritmos de los falsos vecinos más cercanos [13] e información mutua promedio [25] para reconstruir el espacio de fases de la VRC con $m = 3$ y $\tau = 2$.

3.3. Modelos de clasificación

Por un lado, se utilizaron varios algoritmos tradicionales de AA, tales como la máquina de soporte de vectores (SVM), el algoritmo de los k-vecinos más cercanos (KNN) y árbol de decisión (AD). La SVM se configuró con un núcleo de función radial Gaussiana, la KNN incluyó el método de distancia Euclídea y el AD se conformó a partir del algoritmo convencional de CART. Todos los modelos utilizaron las mismas características extraídas del DR de la VRC.

Por otro lado, se implementó aprendizaje por transferencia, utilizando la red preentrenada de AlexNet, introduciendo representaciones del DRG como entrada (imágenes $227 \times 227 \times 3$ píxeles). Algunos de los parámetros de la red preentrenada fueron la tasa de aprendizaje, el número de épocas, tamaño de mini-lote y el optimizador ADAM [6]. Concretamente, los parámetros de entrenamiento consistieron en una tasa de aprendizaje de 0,001, 10 épocas y un tamaño de mini-lote de 128, utilizando el optimizador de estimación de momento adaptativo (ADAM).

3.4. Sistema de evaluación

Para evaluar los modelos generados, se utilizaron medidas de exactitud (Ac), sensibilidad (Se) y especificidad (Sp). La validación de tales modelos siguieron un mismo marco estándar de evaluación, entrenando los modelos en las bases de datos MIT-BIH y UCD-DB balanceadas (mismo número de observaciones por categoría) y evaluándolos sobre Apnea-ECG. Este enfoque permitió obtener resultados con un sesgo reducido y cumpliendo con las directrices TRIPOD [26].

4. Resultados

En la Tabla 1 se reflejan las métricas de rendimiento para los distintos modelos de clasificación. La SVM mostró una Ac del 67,13 %, con una Se del 22,95 % y una Sp del 86,6 %. El modelo KNN alcanzó una precisión del 62,29 %, con valores de Se y Sp del 37,21 % y

73,37%, respectivamente. El modelo AD obtuvo una Ac del 59,20%, con una Se del 43,23% y una especificidad del 66,27%. El modelo de DL, AlexNet, mostró la mayor Ac entre los modelos, registrando un 71,30%, junto con una Se del 58,95% y una Sp del 76,76%.

Modelo	Ac (%)	Se (%)	Sp (%)
SVM	67,13	22,95	86,6
KNN	62,29	37,21	73,37
AD	59,20	43,23	66,27
AlexNet	71,30	58,95	76,76

Tabla 1. Tabla de resultados

5. Discusión

Entre los modelos de AA tradicionales, la SVM demostró una Sp considerable pero a expensas de la Se. Por otro lado, KNN logró un equilibrio más adecuado entre Se y Sp, mientras que AD sobresalió en Sp pero no en Se. Por el contrario, AlexNet superó a todos los demás modelos en términos de Ac y mantuvo un buen equilibrio entre Se y Sp.

En base a estas observaciones, se puede comprobar que los modelos basados en AP muestran un alto rendimiento para la detección de AOS a través del análisis de la VRC utilizando el DRG. En consecuencia, la combinación de características de la VRC y el análisis de recurrencia ofrece una perspectiva distintiva, con AlexNet destacando como líder en términos de Ac y Se, lo que la posiciona como una herramienta valiosa para la detección temprana de la AOS.

En general, los clasificadores basados en AA tradicional pueden mostrar dificultades para lograr un balance adecuado entre valores de Se y Sp en comparación con los clasificadores basados en AP debido a su dependencia de la estrategia de extracción de características escogida. Más precisamente, mientras que los modelos de AA tradicional requieren de intervención humana para la extracción de características, los modelos de AP tienen la capacidad inherente de extraer características previamente no identificadas a partir de los datos de entrada [27], mejorando así la Ac en los resultados obtenidos al priorizar las características que más contribuyen al problema de clasificación.

Es importante destacar que estos resultados han sido sometidos a rigurosos procedimientos de validación de acuerdo con las directrices TRIPOD (*Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis*) [26]. A diferencia de los resultados relativamente optimistas obtenidos por medio de marcos de validación cruzada [28], los modelos presentados en este estudio fueron entrenados en dos bases de datos distintas y posteriormente probados en una tercera, completamente ajena a las de entrenamiento, garantizando una evaluación de rendimiento sólida y clínicamente relevante, aunque esto repercute en unos resultados numéricos menos llamativos. Sin embargo, aún es

preciso realizar ulteriores investigaciones y concentrar esfuerzos en los métodos de evaluación de clasificadores automáticos para aprovechar al máximo el potencial de estos en escenarios más realistas.

6. Conclusiones

El presente estudio resalta el potencial de las arquitecturas de aprendizaje profundo preentrenadas, como AlexNet, para detectar la apnea obstructiva del sueño (AOS) a través del análisis de la variabilidad del ritmo cardíaco con diagramas de recurrencia general. La rigurosidad seguida durante la validación y la alta sensibilidad del modelo generado proporcionan valor en la mejora de la detección de la AOS en intervalos de un minuto de registros de ECG, abordando así un aspecto importante de salud pública.

Agradecimientos

Esta investigación ha recibido apoyo financiero de subvenciones públicas PID2021-00X128525-IV0 y PID2021-123804OB-I00 del Gobierno de España en colaboración con el Fondo Europeo de Desarrollo Regional (UE), SBPLY/17/180501/000411 y SBPLY/21/180501/000186 de la Junta de Comunidades de Castilla-La Mancha, y AICO/2021/286 de la Generalitat Valenciana. Además, Daniele Padovano cuenta con una beca predoctoral 2022-PRED-20642, que está cofinanciada por el programa operativo del Fondo Social Europeo (FSE) 2014-2020 de Castilla-La Mancha.

Referencias

- [1] Jessica Vensel Rundo. Obstructive sleep apnea basics. *Cleveland Clinic Journal of Medicine*, 86(9 Suppl 1), September 2019.
- [2] Chamara V Senaratna, Jennifer L Perret, Caroline J Lodge, Adrian J Lowe, Brittany E Campbell, Melanie C Matheson, Garun S Hamilton, and Shyamali C Dharmage. Prevalence of obstructive sleep apnea in the general population: A systematic review. *Sleep Medicine Reviews*, 34:70–81, Aug 2017.
- [3] Yerem Yeghiazarians, Hani Jneid, Jeremy R. Tietjens, Susan Redline, Devin L. Brown, Nabil El-Sherif, Reena Mehra, Biykem Bozkurt, Chiadi Ericson Nduemele, Virend K. Somers, and null null. Obstructive Sleep Apnea and Cardiovascular Disease: A Scientific Statement From the American Heart Association. *Circulation*, 144(3):e56–e67, July 2021.
- [4] World Health Organization et al. Noncommunicable diseases country profiles 2018. *Publications on NCDs*, 2018.
- [5] Sigrid C. Veasey and Ilene M. Rosen. Obstructive Sleep Apnea in Adults. *New England Journal of Medicine*, 380(15):1442–1449, April 2019.
- [6] Mahsa Bahrami and Mohamad Forouzanfar. Sleep Apnea Detection from Single-Lead ECG: A Comprehensive

- Analysis of Machine Learning and Deep Learning Algorithms. *IEEE Transactions on Instrumentation and Measurement*, pages 1–1, 2022.
- [7] Amal M Osman, Sophie G Carter, Jayne C Carberry, and Danny J Eckert. Obstructive sleep apnea: current perspectives. *Nature and Science of Sleep*, 10:21–34, January 2018.
- [8] Anita Ramachandran and Anupama Karuppiyah. A Survey on Recent Advances in Machine Learning Based Sleep Apnea Detection Systems. *Healthcare*, 9(7):914, July 2021.
- [9] Subani Chandra, Anthony L Sica, Janice Wang, Viera Lakticova, and Harly E Greenberg. Respiratory effort-related arousals contribute to sympathetic modulation of heart rate variability. *Sleep Breathing*, 17(4):1193–200, Dec 2013.
- [10] Cheng-Yu Lin, Yi-Wen Wang, Febryan Setiawan, Nguyen Thi Hoang Trang, and Che-Wei Lin. Sleep apnea classification algorithm development using a machine-learning framework and bag-of-features derived from electrocardiogram spectrograms. *Journal of Clinical Medicine*, 11(1):192, 2021.
- [11] Zhao Dong, Xiang Li, and Wei Chen. Frequency network analysis of heart rate variability for obstructive apnea patient detection. *Ieee Journal of Biomedical and Health Informatics*, 22(6):1895–1905, November 2018.
- [12] Asghar Zarei and Babak Mohammadzadeh Asl. Automatic detection of obstructive sleep apnea using wavelet transform and entropy-based features from single-lead ECG signal. *IEEE Journal of Biomedical and Health Informatics*, 23(3):1011–1021, May 2019.
- [13] Hoa Dinh Nguyen, Brek A. Wilkins, Qi Cheng, and Bruce Allen Benjamin. An Online Sleep Apnea Detection Method Based on Recurrence Quantification Analysis. *IEEE Journal of Biomedical and Health Informatics*, 18(4):1285–1293, July 2014.
- [14] Trung Q. Le and Satish T. S. Bukkapatnam. Nonlinear Dynamics Forecasting of Obstructive Sleep Apnea Onsets. *PLOS ONE*, 11(11):e0164406, November 2016.
- [15] Sofia Martín-González, Juan L. Navarro-Mesa, Gabriel Julia-Serda, Jan F. Kraemer, Niels Wessel, and Antonio G. Ravelo-García. Heart rate variability feature selection in the presence of sleep apnea: An expert system for the characterization and detection of the disorder. *Computers in Biology and Medicine*, 91:47–58, December 2017.
- [16] Norbert Marwan, Niels Wessel, Udo Meyerfeldt, Alexander Schirdewan, and Jürgen Kurths. Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *American Physical Society*, 66(2):026702, August 2002.
- [17] Norbert Marwan, M. Carmen Romano, Marco Thiel, and Jürgen Kurths. Recurrence plots for the analysis of complex systems. *Physics Reports*, 438(5):237–329, January 2007.
- [18] Thomas Penzel, George B Moody, Roger G Mark, Ary L Goldberger, and J Hermann Peter. Apnea-ECG database, 2000.
- [19] Yuhei Ichimaru and George B Moody. MIT-BIH polysomnographic database, 1992.
- [20] Walter McNicholas, Liam Doherty, Silke Ryan, John Garvey, Patricia Boyle, and Eric Chua. St. Vincent’s University Hospital / University College Dublin Sleep Apnea Database, 2004.
- [21] Doris Moser, Peter Anderer, Georg Gruber, Silvia Parapatics, Erna Loretz, Marion Boeck, Gerhard Kloesch, Esther Heller, Andrea Schmidt, Heidi Danker-Hopfe, Bernd Saletu, Josef Zeitlhofer, and Georg Dorffner. Sleep Classification According to AASM and Rechtschaffen & Kales: Effects on Sleep Scoring Parameters. *Sleep*, 32(2):139–149, February 2009.
- [22] Jiapu Pan and Willis J. Tompkins. A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering*, BME-32(3):230–236, March 1985.
- [23] Lyle Noakes. The takens embedding theorem. *International Journal of Bifurcation and Chaos*, 01(04):867–872, December 1991.
- [24] Sofía Martín-González, Juan L. Navarro-Mesa, Gabriel Juliá-Serdá, G. Marcelo Ramírez-Ávila, and Antonio G. Ravelo-García. Improving the understanding of sleep apnea characterization using recurrence quantification analysis by defining overall acceptable values for the dimensionality of the system, the delay, and the distance threshold. *Public Library of Science*, 13(4):e0194462, April 2018.
- [25] Sebastian Wallot and Dan Mønster. Calculation of Average Mutual Information (AMI) and False-Nearest Neighbors (FNN) for the Estimation of Embedding Parameters of Multidimensional Time Series in Matlab. *Frontiers in Psychology*, 9, 2018.
- [26] G S Collins, J B Reitsma, D G Altman, and K G M Moons. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *British Journal of Surgery*, 102(3):148–158, 2015.
- [27] Jian-Wei Liu, Hui-Dan Zhao, Xiong-Lin Luo, and Jun Xu. Research Progress on Batch Normalization of Deep Learning and Its Related Algorithms. *Acta Automatica Sinica*, 46(6):1090–1120, 2021.
- [28] Gabriele B. Papini, Pedro Fonseca, Jenny Margarito, Merel M. van Gilst, Sebastiaan Overeem, Jan W.M. Bergmans, and Rik Vullings. On the generalizability of ECG-based obstructive sleep apnea monitoring: merits and limitations of the Apnea-ECG database. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6022–6025, July 2018.

Análisis integral de la vorticidad en la región anterior de las vías aéreas nasales

D. Sanz-Prieto^{1*}, M.A. Burgos², A. Pérez-Ramos³, F. Esteban-Ortega⁴, M. Bastir¹

¹ Departamento de Paleobiología, Museo Nacional de Ciencias Naturales – Centro Superior de Investigaciones Científicas, Madrid, España, daniel.sanz@mncn.csic.es

² Departamento de Ingeniería Térmica y de Fluidos, Universidad Politécnica de Cartagena, Murcia, España

³ Departamento de Ecología y Geología, Facultad de Ciencias, Universidad de Málaga, Málaga, España

⁴ Servicio de Otorrinolaringología, Hospital Universitario Virgen del Rocío, Sevilla, España

Resumen

Este estudio se centra en analizar la relación entre la morfología nasal y la dinámica del flujo de aire en la válvula nasal, con el objetivo de comprender su impacto en la función respiratoria. Se empleó una muestra de 40 tomografías computarizadas (TC) de individuos sanos de España y se aplicaron técnicas de dinámica de fluidos computacionales y morfometría geométrica 3D.

Los resultados revelaron una asociación significativa entre la morfología nasal y la vorticidad en la válvula nasal. Las narices alargadas, alas nasales estrechas y vestibulos nasales anchos (tendencia leptorrina) mostraron menor vorticidad, mientras que las narices chatas, alas nasales anchas y vestibulos nasales estrechos (tendencia platirrina) presentaron mayor vorticidad.

Estos hallazgos sugieren que la morfología nasal influye en la vorticidad del flujo en la válvula nasal, lo que podría afectar la función respiratoria en relación al acondicionamiento de aire. También se plantea la posibilidad de que estas diferencias morfológicas estén relacionadas con adaptaciones evolutivas en poblaciones humanas.

1. Introducción

Mlynski et al. [1] propusieron una regionalización organizada anteroposteriormente para las vías aéreas nasales (tejido blando) en tres unidades morfofuncionales distintas caracterizadas por diferentes características de flujo de aire:

- El tracto de entrada compuesto por el vestíbulo nasal, la válvula nasal y el cavum anterior.
- El tracto funcional, donde se localizan los meatos.
- El tracto de salida conformado por el cavum posterior, las coanas y la nasofaringe.

Si bien se ha considerado que el tracto funcional es el área principal responsable del acondicionamiento del aire [1], también se ha sugerido que el segmento nasal anterior desempeña un papel en este proceso hasta cierto punto [2].

Teniendo en cuenta que una de las principales funciones de las vías aéreas nasales es el acondicionamiento de aire, un reciente estudio empleando poblaciones humanas de distintas zonas geográficas halló que existen diferencias en el intercambio de calor y humedad en la región nasal anterior, pero que la eficiencia del proceso se igualaba en la región media y posterior [3].

Además, hipotetizaron que una de las posibles explicaciones funcionales para este suceso era una mayor vorticidad en la región nasal anterior, lo que aumentaría el contacto entre el aire y la mucosa nasal en las primeras secciones del tracto funcional.

El objetivo de este trabajo es analizar la vorticidad en la válvula nasal en una submuestra española de dicho estudio mediante dinámica de fluidos computacional y relacionar su valor con la configuración morfológica de la región anterior de las vías aéreas. Esto permitirá establecer relaciones estadísticas entre la morfología nasal y la funcionalidad respiratoria.

2. Materiales y métodos

Se analizaron retrospectivamente 40 tomografías computarizadas (TC) de una muestra de individuos sanos procedentes de España y balanceados por sexo. Los criterios de inclusión fueron la edad adulta, la disponibilidad de una TC facial que incluyera las vías aéreas nasales al completo, y que los pacientes que se sometieron a la adquisición de las TC por razones no relacionadas con enfermedades nasales.

El proceso de dinámica de fluidos computacional se realizó con el software Flowgy [4] y consistió en los siguientes pasos:

- Segmentación de las TC para la obtención de los modelos 3D de cada individuo, compuestos por la cara, las vías aéreas nasales y un área rectangular exterior simulando un ambiente atmosférico.
- Discretización de los modelos 3D en mallas computacionales compuestas por una serie de elementos finitos tetraédricos (entre 7 y 15 millones de celdas).
- Simulación del flujo de aire en inspiración a partir de unas condiciones de contorno individualizadas. Se calculó un flujo volumétrico aproximado de 15L/min asegurando un régimen laminar [5].
- Con el flujo resuelto, se realizó un plano de corte homólogo para todos los individuos a la altura de la válvula nasal. La vorticidad media en esta región para cada individuo fue extraída en este plano.

La morfometría geométrica 3D se utilizó para analizar la morfología de las vías aéreas nasales. Medimos 284 puntos

de referencia y utilizando el software ViewBox4 [6] en los modelos virtuales en 3D que se generaron previamente para los análisis de dinámica de fluidos computacional. Estos puntos de referencia recogieron las relaciones espaciales y la configuración estructurales de la región anatómica anterior de las vías aéreas nasales.

Se empleó la regresión de mínimos cuadrados parciales como una herramienta fundamental en nuestro enfoque metodológico integral. La regresión de mínimos cuadrados parciales es una técnica estadística que nos permite explorar y cuantificar la relación entre variables independientes multidimensionales y una variable dependiente. Este análisis se utilizó para analizar la covariación entre las coordenadas en tres dimensiones y la vorticidad en la válvula nasal, empleando los paquetes *Morpho* [7] y *geomorph* [8] del entorno y lenguaje de programación R.

3. Resultados

La regresión de mínimos cuadrados parciales mostró un patrón de covariación significativo (Tabla 1), es decir, para una determinada configuración nasal, existe un valor específico de vorticidad en la válvula nasal.

Valor singular	Coefficiente de correlación	p-valor
25.903	0.745	<0.01

Tabla 1. Resultados de la regresión de mínimos cuadrados parciales combinando los datos geomorfométricos y la variable funcional de la vorticidad. El p-valor asociado está basado en una prueba de permutación de 999 rondas.

Los individuos con una nariz más alargada, las alas nasales más estrechas y un vestíbulo nasal más ancho (tendencia leptorrina), muestran valores de vorticidad más bajos (Figura 1).

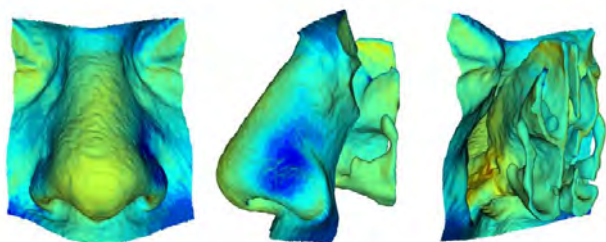


Figura 1. Modelo 3D de la forma media de la muestra para la región anterior de las vías aéreas nasales junto con un mapa de calor hacia los valores mínimos de vorticidad en la válvula nasal.

Por otro lado, los individuos con una nariz más chata, las alas nasales más anchas y un vestíbulo nasal más estrecho (tendencia platirrina) muestran valores de vorticidad más altos (Figura 2).

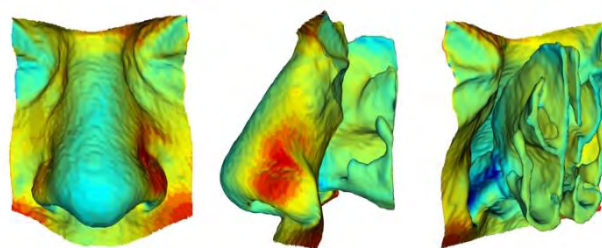


Figura 2. Modelo 3D de la forma media de la muestra para la región anterior de las vías aéreas nasales junto con un mapa de calor hacia los valores máximos de vorticidad en la válvula nasal.

4. Discusión

Estos hallazgos sugieren que la configuración nasal platirrina llevaría asociada una mayor vorticidad en la válvula nasal. Este hecho podría tener importantes implicaciones en el estudio de la función respiratoria y de la variabilidad morfológica nasal en poblaciones humanas.

Un reciente trabajo observó que las poblaciones de climas cálidos y húmedos (patrón platirrino) poseían un menor acondicionamiento del aire en la región nasal anterior, pero la eficiencia del proceso se igualaba hacia regiones posteriores [3]. Los presentes resultados apoyan la hipótesis de que un aumento de la vorticidad, asociado a una configuración platirrina, podría suponer un incremento del acondicionamiento de aire.

Sin embargo, para continuar testando esta asociación morfofuncional, sería necesario llevar a cabo análisis más amplios que incluyan las poblaciones de climas cálidos y húmedos. En futuros estudios, esperamos encontrar diferencias geográficas significativas para la vorticidad en la región nasal anterior, relacionadas con las diferencias morfológicas entre poblaciones de climas cálidos y húmedos frente a las de ambientes fríos y secos.

Además, esta investigación podría ser relevante en el contexto de la antropología y la evolución, ya que podría ayudarnos a comprender cómo ciertos rasgos nasales se han adaptado a lo largo del tiempo en respuesta a diferentes presiones selectivas.

5. Conclusiones

Nuestros hallazgos proporcionan una base para futuras investigaciones que podrían ayudar a esclarecer aún más la relación entre la morfología nasal y la función respiratoria en diferentes poblaciones humanas y en contextos evolutivos.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia e Innovación de España y la Unión Europea mediante los proyectos de investigación PID2019-105097RB-I00 a M.A.B. y PID2020-115854GB-I00 a M.B.

Referencias

- [1] Mlynski G, Grützenmacher S, Plontke S, Mlynski B, Lang C. Correlation of nasal morphology and respiratory function. *Rhinology*, vol 39, sup 4, 2001, pp 197-201.
- [2] Inthavong K, Wen J, Tu J, Tian Z. From CT scans to CFD modelling—fluid and heat transfer in a realistic human nasal cavity. *Engineering Applications of Computational Fluid Mechanics*, vol 3, sup 3, 2009, pp 321-335.
- [3] Sanz-Prieto D, Burgos MA, Pérez-Ramos A, Heuzé Y, Maréchal L, Evteev A, Toro-Ibacache V, Esteban-Ortega F, Bastir M. Beyond Skeletal Studies: A Computational Analysis of Nasal Airway Function in Climate Adaptation. *American Journal of Biological Anthropology*, under review.
- [4] Flowgy. The most affordable, fastest and efficient nasal surgery. <https://www.flowgy.com/>
- [5] Taylor DJ, Doorly DJ, Schroter RC. Inflow boundary profile prescription for numerical simulation of nasal airflow. *Journal of The Royal Society Interface*, vol 7, sup 44, 2010, pp 515-527.
- [6] Dhal Viewbox. 2D and 3D Cephalometric Software. <https://www.dhal.com/viewbox.htm>
- [7] Schlager S. Morpho and Rvcg – Shape Analysis in R. En Zheng G, Li S, Szekely G (eds.). *Statistical Shape and Deformation Analysis*, Academic Press, 2017, pp 217-256 (ISBN 9780128104934).
- [8] Baken E, Collyer M, Kaliontzopoulou A, Adams D. geomorph v4.0 and gmShiny: enhanced analytics and a new graphical interface for a comprehensive morphometric experience. *Methods in Ecology and Evolution*, vol 12, 2021, pp 2355-2363.

Análisis computacional del flujo espiratorio nasal: Investigación de las diferencias geográficas y de sexo en la resistencia y la asimetría

I. M. García^{1*}, D. Sanz Prieto², A. Pérez Ramos³, M. Bastir², F. Esteban Ortega⁴, L. Picazo Selva¹, M. Burgos Olmos¹

¹ Dpto. de Ingeniería Térmica y de Fluidos, Universidad Politécnica de Cartagena, Murcia, Spain

² Grupo de Paleoantropología, Departamento de Paleobiología, Museo Nacional de Ciencias Naturales – Centro Superior de Investigaciones Científicas, Madrid, España.

³ Dpto. de Ecología y Geología, Facultad de Ciencias, Universidad de Málaga, 29071, Málaga, Spain

⁴ Servicio Andaluz de Salud, Sevilla, Spain. Hospital Universitario Virgen del Rocío, Servicio de Otorrinolaringología.

Resumen

La mecánica de fluidos computacional (CFD) ha permitido simular flujos respiratorios y analizar parámetros como la velocidad y la presión. Estudios previos de nuestro grupo han utilizado CFD para analizar dos parámetros adimensionales en inspiración: el estimador R , que representa la resistencia nasal bilateral, y el estimador ϕ , que cuantifica la asimetría del flujo. Estos parámetros demostraron ser útiles para distinguir entre vías aéreas nasales normales y patológicas. Inicialmente se analizaron en una muestra española y recientemente en 95 tomografías de adultos sanos de tres poblaciones diferentes, sin encontrar diferencias significativas. Sin embargo, estos análisis se limitaron a la inspiración. El objetivo de este estudio fue evaluar los parámetros R y ϕ en espiración mediante CFD y compararlos con los de inspiración en la misma muestra poblacional. Para ello, se analizaron retrospectivamente 95 tomografías de adultos sanos balanceados por sexo y población. Se realizaron simulaciones CFD en espiración y se calcularon los parámetros R y ϕ , que luego se compararon con los de inspiración. Los resultados mostraron que los valores y distribuciones de R y ϕ fueron muy similares en espiración e inspiración, sin diferencias significativas ni geográficas ni sexuales. Este estudio complementa el análisis en inspiración y sugiere que R y ϕ podrían ser características prometedoras y robustas para asistir en el diagnóstico nasal utilizando CFD en diversas poblaciones.

1. Introducción

El flujo nasal presenta una amplia variabilidad intra e interindividual, al igual que la morfología nasal humana. Esto ha dificultado la identificación de patrones universales en el flujo nasal y su relación con la morfología, que permitan una descripción más precisa de las vías respiratorias nasales normales [1]–[5]. En este contexto, la mecánica de fluidos computacional (CFD) ha realizado importantes contribuciones en el estudio del flujo respiratorio humano, específicamente en el área biomédica [6].

La CFD permite simular y analizar distintos flujos y su interacción con las propiedades geométricas de superficies sólidas. Los resultados de estos análisis proporcionan información relevante sobre parámetros dimensionales del flujo como velocidad, temperatura o humedad. Además, algunos autores definieron dos parámetros adimensionales que caracterizan el flujo nasal: el estimador R representa la

resistencia nasal bilateral y el estimador ϕ simboliza una función de posibles asimetrías del flujo [7].

Estos parámetros adimensionales han demostrado distinguir entre vías aéreas nasales normales o patológicas [7]–[11]. Sin embargo, una limitación es que fueron definidos en una muestra española [7], por lo que se desconocían las variaciones en otros grupos.

Recientemente, nuestro grupo analizó posibles diferencias sexuales y geográficas en R y ϕ mediante CFD en inspiración [9], [10], [12]. No se encontraron diferencias significativas entre sexos ni poblaciones, excepto un leve dimorfismo sexual en ϕ en la muestra camboyana [12]. Estos resultados apoyaron el uso de R y ϕ para diagnóstico nasal computarizado en diversos grupos humanos.

Sin embargo, el análisis del flujo solo se realizó en la fase inspiratoria. Dado que la espiración presenta diferencias fisiológicas y de flujo respecto a la inspiración [13], [14], resulta relevante analizar si estos parámetros adimensionales también son útiles en la fase espiratoria del ciclo respiratorio.

El objetivo de este estudio es evaluar los parámetros R y ϕ en espiración mediante CFD en la misma muestra geográficamente diversa previamente analizada en inspiración. Los resultados se compararán con inspiración para determinar si existen diferencias. Esto permitirá establecer la utilidad de R y ϕ para el diagnóstico computarizado de la función nasal normal en espiración.

2. Material y Métodos

Se analizaron retrospectivamente 95 tomografías computarizadas (TC) de individuos adultos sanos de tres áreas geográficas: España (11 hombres y 10 mujeres), Rusia (20 hombres y 15 mujeres) y Camboya (19 hombres y 20 mujeres). La muestra y criterios de inclusión fueron los mismos que nuestro estudio previo [12].

El proceso de CFD se realizó con el software Flowgy [15]. Se segmentaron las TC y construyeron modelos 3D de las cavidades nasales y componentes faciales. Se generaron mallas de elementos finitos tetraédricos de 7-15 millones de celdas. Se simularon los flujos en espiración, estableciendo condiciones de contorno individualizadas.

Se calculó un flujo volumétrico aproximado de 15 L/min asegurando régimen laminar [13]. Se analizaron las soluciones CFD y se calcularon los parámetros R y ϕ .

Con el propósito de investigar la existencia de posibles diferencias significativas en los parámetros de dinámica de fluidos nasales durante la inspiración y espiración, se llevó a cabo un minucioso análisis estadístico tanto de la muestra total como estratificada por sexo y grupos poblacionales.

Las técnicas estadísticas empleadas incluyeron estadística descriptiva a través del cálculo de medias y desviaciones estándar, pruebas t de Student para muestras apareadas a fin de comparar promedios entre inspiración y espiración, coeficientes de correlación de Pearson para cuantificar la asociación lineal entre condiciones, y representaciones gráficas para la visualización de patrones.

Los valores de R y ϕ en espiración se compararon con los obtenidos previamente en inspiración [12]. Un valor de $p < 0,05$ se consideró estadísticamente significativo.

2.1. Detalles del Protocolo y Tratamiento de Datos

Para garantizar la transparencia y replicabilidad de nuestro estudio, es crucial detallar el protocolo seguido durante la recolección de datos y el tratamiento posterior de los mismos. Las tomografías computarizadas utilizadas en este

estudio se obtuvieron de hospitales colaboradores, siguiendo un protocolo ético aprobado por el comité de ética correspondiente.

En cuanto al tratamiento de los datos, las imágenes de tomografía computarizada se procesaron utilizando el software Flowgy [15], que permitió una segmentación precisa de las vías respiratorias. Posteriormente, se realizaron simulaciones de CFD utilizando OpenFoam [16]. Los datos se analizaron estadísticamente aplicando pruebas t-student para determinar la significancia de las diferencias observadas entre los grupos. Específicamente, la prueba t-student se utilizó para comparar las medias de los parámetros R y ϕ entre diferentes grupos, como por ejemplo entre hombres y mujeres o entre diferentes poblaciones geográficas. Todos los datos se almacenaron en un servidor seguro y se anonimizaron para garantizar la confidencialidad

3. Resultados

En las Tablas 1 y 2 se muestran los resultados estadísticos de R y ϕ obtenidos. Los resultados presentados en las tablas proporcionan un panorama completo de las similitudes en los coeficientes adimensionales de simetría y resistencia del flujo nasal durante la inspiración y espiración para las distintas poblaciones y ambos sexos.

Parámetro de Simetría							
Población	Sexo	Media Inspiración	Media Espiración	Desviación Típica Inspiración	Desviación Típica Espiración	Coefficiente Pearson	P
Camboyano	H	1.680	1.634	0.435	0.435	0.984	< 0.001
Camboyano	M	1.392	1.350	0.284	0.231	0.986	< 0.001
Ruso	H	1.530	1.486	0.561	0.526	0.995	< 0.001
Ruso	M	1.427	1.398	0.186	0.164	0.930	< 0.001
Español	H	1.238	1.195	0.091	0.110	0.918	< 0.001
Español	M	1.308	1.279	0.161	0.144	0.966	< 0.001
Todos	H y M	1.459	1.419	0.380	0.359	0.986	< 0.001

Tabla 1. Estadística descriptiva y correlacional del coeficiente adimensional de simetría del flujo nasal durante la inspiración y espiración para diferentes grupos poblacionales segregados por sexo.

Parámetro de Resistencia							
Población	Sexo	Media Inspiración	Media Espiración	Desviación Típica Inspiración	Desviación Típica Espiración	Coefficiente Pearson	P
Camboyano	H	9.288	8.415	6.426	5.753	0.998	< 0.001
Camboyano	M	7.702	7.436	4.460	4.286	0.847	< 0.001
Ruso	H	9.244	8.890	6.494	7.278	0.981	< 0.001
Ruso	M	7.507	6.997	4.028	4.094	0.967	< 0.001
Español	H	9.473	8.119	4.543	3.056	0.919	< 0.001
Español	M	6.068	5.344	3.116	2.420	0.978	< 0.001
Todos	H y M	8.326	7.700	39.205	32.947	0.951	< 0.001

Tabla 2. Estadística descriptiva y correlacional del coeficiente adimensional de resistencia del flujo nasal durante la inspiración y espiración para distintos grupos poblacionales separados por sexo.

Es importante aclarar el significado de los valores 'p' presentados en las tablas de esta sección. Estos valores se obtuvieron mediante pruebas t-student y representan el nivel de significancia en la comparación de las medias de los parámetros R y ϕ entre diferentes grupos (por ejemplo, entre hombres y mujeres o entre diferentes poblaciones geográficas). En este estudio, los valores 'p' se han calculado para cada comparación y se presentan en las tablas correspondientes para proporcionar un contexto completo de la significancia estadística de nuestras observaciones.

Se observa que los promedios de estos parámetros son prácticamente iguales entre las dos condiciones respiratorias, con diferencias menores al 10% en todos los casos. Asimismo, los rangos y dispersiones representados por las desviaciones estándar son equivalentes.

El análisis estadístico arroja correlaciones positivas muy fuertes ($r > 0.90$) entre inspiración y espiración para ambos coeficientes y todos los grupos, con significancia estadística ($p < 0.001$) que permite descartar la hipótesis nula de no asociación.

En conjunto, estos resultados indican que los parámetros adimensionales R y ϕ presentan valores y distribuciones muy similares en espiración comparado con inspiración. Además, en espiración tampoco se evidencian diferencias sexuales ni geográficas significativas.

4. Discusión

Nuestros hallazgos muestran que los parámetros ϕ y R presentan valores y distribuciones muy similares en espiración comparado con la fase inspiratoria previamente analizada [12].

Se evidencia una notable similitud en los valores medios del coeficiente de simetría entre la inspiración y espiración para todas las poblaciones y ambos sexos. Las pequeñas diferencias observadas no superan el 10%, indicando que este parámetro adimensional, que cuantifica la distribución del flujo entre fosas nasales, se conserva indistintamente de la fase respiratoria.

Esta equivalencia se refleja también en las desviaciones estándar, que representan la dispersión de los datos y resultan equivalentes entre inspiración y espiración dentro de cada grupo. Asimismo, los altísimos coeficientes de correlación positiva ($r > 0.9$) refuerzan desde la estadística inferencial la inexistencia de diferencias significativas en el coeficiente de simetría.

Resultados análogos se reproducen para el coeficiente adimensional de resistencia del flujo nasal. Y nuevamente, las poderosas correlaciones positivas ($r > 0.8$) avalan la similitud entre fases respiratorias.

La notable preservación de los coeficientes aerodinámicos adimensionales, a pesar de los contrastantes patrones estructurales de flujo durante la inspiración y espiración,

encuentra su génesis en la eficiente optimización fisiológica del sistema respiratorio humano. Dicho sistema procura satisfacer los requerimientos de intercambio gaseoso aprovechando la anatomía nasal disponible de la manera más efectiva, ya sea durante la inhalación o exhalación del ciclo respiratorio.

Esta versatilidad adaptativa se manifiesta en las diferencias de las estructuras de flujo, visualizadas por ejemplo en las líneas de corriente entre las distintas fases respiratorias, como se aprecia en la Figura 1 para sujetos representativos de las poblaciones analizadas.

5. Conclusiones

En este estudio se analizaron mediante CFD los parámetros adimensionales de resistencia (R) y asimetría (ϕ) del flujo nasal en espiración, en una muestra geográficamente diversa.

Los valores y distribuciones de R y ϕ fueron muy similares en espiración comparado con inspiración. No se observaron diferencias sexuales ni geográficas significativas en ninguno de los dos parámetros.

Estos hallazgos complementan nuestro estudio previo en inspiración, mostrando la utilidad potencial de R y ϕ para el diagnóstico computarizado de la función nasal normal en diversas poblaciones mediante CFD, tanto en inspiración como espiración.

Analizar los parámetros R y ϕ en una población con patologías nasales podría ofrecer nuevas perspectivas sobre la utilidad de estos parámetros en el diagnóstico y tratamiento de condiciones médicas específicas.

6. Agradecimientos

Yann Heuze (Universidad de Bordeaux) y Andrej Evteev (Lomonosov's Moscow State University, Anuchin's Research Institute and Museum of Anthropology) han proporcionado datos de CT de poblaciones rusas y camboyanos.

7. Financiación

Este trabajo ha sido patrocinado por el Ministerio de Ciencia e Innovación de España:

- La Subvención PID2019-105097RB-I00 a Manuel Burgos. está financiada por CIN/AEI/10.13039/501100011033 del Ministerio de Ciencia e Innovación de España y la Unión Europea.
- La subvención PID2020-115854GB-I00 a Markus Bastir está financiada por MCIN/AEI/10.13039/501100011033 del Ministerio de Ciencia e Innovación de España y la Unión Europea.

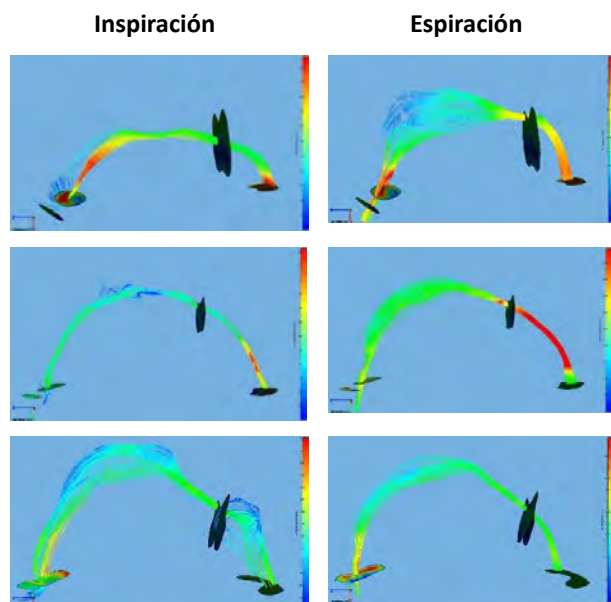


Figura 1. Líneas de corriente durante la inspiración (izquierda) y espiración (derecha) en sujetos representativos de las poblaciones camboyana (arriba), rusa (centro) y española (abajo). Las marcadas diferencias en la conformación de las trayectorias de flujo entre las fases respiratorias evidencian los contrastantes patrones estructurales, a pesar de la conservación cuantitativa de propiedades aerodinámicas demostrada en parámetros adimensionales.

Referencias

- [1] D. J. Doorly, D. J. Taylor, and R. C. Schroter, “Mechanics of airflow in the human nasal airways,” *Respir Physiol Neurobiol*, vol. 163, no. 1–3, pp. 100–110, Nov. 2008, doi: 10.1016/j.resp.2008.07.027.
- [2] R. A. Segal, G. M. Kepler, and J. S. Kimbell, “Effects of differences in nasal anatomy on airflow distribution: a comparison of four individuals at rest,” *Ann Biomed Eng*, vol. 36, no. 11, pp. 1870–1882, Nov. 2008, doi: 10.1007/s10439-008-9556-2.
- [3] T. R. Yokley, “Ecogeographic variation in human nasal passages,” *Am J Phys Anthropol*, vol. 138, no. 1, pp. 11–22, Jan. 2009, doi: 10.1002/ajpa.20893.
- [4] M. Bastir, P. Godoy, and A. Rosas, “Common features of sexual dimorphism in the cranial airways of different human populations,” *Am J Phys Anthropol*, vol. 146, no. 3, pp. 414–422, Nov. 2011, doi: 10.1002/ajpa.21596.
- [5] S. D. Maddux, L. N. Butaric, T. R. Yokley, and R. G. Franciscus, “Ecogeographic variation across morphofunctional units of the human nose,” *Am J Phys Anthropol*, vol. 162, no. 1, pp. 103–119, Jan. 2017, doi: 10.1002/ajpa.23100.
- [6] “Leong et al. - 2010 - A review of the implications of computational flui.pdf.”
- [7] E. Sanmiguel-Rojas, M. A. Burgos, C. Del Pino, M. A. Sevilla-García, and F. Esteban-Ortega, “Robust nondimensional estimators to assess the nasal airflow in health and disease,” *Int J Numer Method Biomed Eng*, vol. 34, no. 1, Jan. 2018, doi: 10.1002/cnm.2906.
- [8] M. A. Burgos, E. Sanmiguel-Rojas, N. Singh, and F. Esteban-Ortega, “DigBody®: A new 3D modeling tool for nasal virtual surgery,” *Comput Biol Med*, vol. 98, pp. 118–125, Jul. 2018, doi: 10.1016/j.compbiomed.2018.05.016.
- [9] M. A. Burgos, E. Sanmiguel-Rojas, R. Rodríguez, and F. Esteban-Ortega, “A CFD approach to understand nasoseptal perforations,” *Eur Arch Otorhinolaryngol*, vol. 275, no. 9, pp. 2265–2272, Sep. 2018, doi: 10.1007/s00405-018-5073-6.

- [10] M. A. Burgos *et al.*, “Linking Chronic Otitis Media and Nasal Obstruction: A CFD Approach,” *Laryngoscope*, vol. 132, no. 6, pp. 1224–1230, Jun. 2022, doi: 10.1002/lary.29882.
- [11] E. Sanmiguel-Rojas, M. A. Burgos, and F. Esteban-Ortega, “Nasal surgery handled by CFD tools,” *International Journal for Numerical Methods in Biomedical Engineering*, vol. 0, no. ja, p. e3126, Feb. 2018, doi: 10.1002/cnm.3126.
- [12] Daniel Sanz-Prieto, Markus Bastir, Alejandro Pérez-Ramos, Francisco Esteban-Ortega, Yann Heuzé, Laura Maréchal, Manuel A. Burgos, “Testing geographic and sexual differences in dimensionless parameters for the human nasal airflow.” Manuscript under review.
- [13] D. J. Taylor, D. J. Doorly, and R. C. Schroter, “Inflow boundary profile prescription for numerical simulation of nasal airflow,” *Journal of The Royal Society Interface*, vol. 7, no. 44, pp. 515–527, Mar. 2010, doi: 10.1098/rsif.2009.0306.
- [14] I. Hörschler, W. Schröder, and M. Meinke, “On the assumption of steadiness of nasal cavity flow,” *J Biomech*, vol. 43, no. 6, pp. 1081–1085, Apr. 2010, doi: 10.1016/j.jbiomech.2009.12.008.
- [15] “Flowgy The most affordable, fastest and efficient nasal surgery.” <https://www.flowgy.com/>, consultado en octubre, 2023.
- [16] “OpenFOAM - Official home of The Open Source Computational Fluid Dynamics (CFD) Toolbox.” <http://www.openfoam.com.> consultado en octubre, 2023.

Aplicación de técnicas de Inteligencia Artificial a datos de Diabetes de un repositorio OMOP *Common Data Model*

A. Licona-Beltrán¹, J. Tapia-Galisteo^{1,3}, M. Pedrera-Jiménez², P. Rubio-Mayo², F. J. Somolinos-Simon¹, G. García-Sáez^{1,3}, M. E. Hernando^{1,3}

¹ Grupo de Bioingeniería y Telemedicina, Universidad Politécnica de Madrid, Madrid, España, a.licona@alumnos.upm.es, {jose.tapia.galisteo, javier.ssimon, gema.garcia.saez, mariaelena.hernando}@upm.es

² Unidad de Datos, Hospital 12 de Octubre, Madrid, España, {miguel.pedrera, paula.rubio.externo}@salud.madrid.org

³ Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, CIBER, Madrid, España

Resumen

La Diabetes Mellitus (DM) es una de las enfermedades crónicas con mayor prevalencia. La DM es una enfermedad compleja y multifactorial, por lo que dentro de cada tipología hay pacientes con perfiles diferentes que responden de distinta manera a los tratamientos. OMOP es un modelo de datos estándar centrado en la información observacional en salud. Este trabajo utiliza técnicas de Inteligencia Artificial para estratificar a los pacientes con DM tipo 1 y tipo 2 de un repositorio OMOP de datos del mundo real del Hospital 12 de Octubre de Madrid. Se han extraído y pre-procesado 35.224 casos y se ha abordado el análisis de calidad del dato. Se han identificado cinco clústeres para DM tipo 1 y tres clústeres para DM tipo 2 en base a cuatro variables determinantes. La metodología propuesta en este trabajo es reutilizable en otros repositorios estandarizados con el modelo OMOP.

1. Introducción

La Diabetes mellitus (DM) es una patología crónica caracterizada por un desorden metabólico y está relacionada con factores genéticos, medioambientales y del estilo de vida. Existen varias tipologías de DM y las más prevalentes son la diabetes tipo 1 (DT1) y la diabetes tipo 2 (DT2), ambas objeto de este trabajo. La DT1 es una enfermedad autoinmune y representa el 10% de los casos de DM. Se diagnostica principalmente en niños y adultos jóvenes. Los pacientes necesitan la administración exógena de insulina para mantener sus niveles de glucosa en sangre dentro de los rangos de normalidad. La DT2 representa el 90% de los casos y suele estar asociada a la edad adulta, la obesidad y el sedentarismo. Las personas con DT2 presentan resistencia a la insulina y pueden ser asintomáticas, lo que influye en el número de casos no diagnosticados y la gravedad de las comorbilidades [1].

En 2021, el 10,5% de la población mundial tenía DM, y se espera que la prevalencia suba hasta el 12,2% en 2045. España es el cuarto país en el Top 5 europeo con mayor cantidad de población con DM, con 5,1 millones de pacientes en 2021. Los costes de la DM representaron el 11,5% del gasto mundial en salud en 2021. España es el noveno país en el mundo con el mayor gasto en salud relacionado con la DM, con un gasto anual de 15,5 millones de dólares [1].

La mayoría de las complicaciones de la DM son comunes en ambas tipologías. Las de mayor prevalencia son: enfermedades cardiovasculares, accidentes

cerebrovasculares, retinopatía, neuropatía y nefropatía [2]. Estas comorbilidades son prevenibles con un estilo de vida saludable y una gestión adecuada de la enfermedad.

Los Ensayos Controlados Aleatorizados (ECA) son considerados el estándar de referencia para obtener evidencia científica, aunque también tienen ciertas limitaciones relacionadas con los altos costes, con la representación limitada de la población, con un seguimiento temporal limitado y con cuestiones éticas [3].

En los últimos años ha surgido una corriente de investigación centrada en la reutilización de datos del mundo real (RWD, del inglés *Real-World Data*), denominado “uso secundario de la información en salud”. Los repositorios de RWD contienen información de la Historia Clínica Electrónica (HCE), incluyendo la gestión de la enfermedad, la práctica clínica y los datos de dispositivos externos [4].

Los repositorios de RWD permiten realizar análisis prospectivos para generar evidencia en salud, y se pueden utilizar para complementar a los ECA, mitigando sus principales limitaciones. El análisis con RWD permite la caracterización de una enfermedad a través de la práctica clínica y la consideración de variables que pueden ser pasadas por alto en los ECA [5].

En este contexto de análisis de datos de la HCE, es esencial mencionar el papel de los estándares de interoperabilidad, pues son fundamentales en el uso secundario de la información en salud, ya que permiten la combinación y reutilización de datos de diferentes sistemas. Uno de los estándares mayormente aceptados es el modelo de datos común del *Observational Medical Outcomes Partnership* (OMOP) [6], cuyo objetivo es resaltar el valor de la información observacional en salud y estructurar un modelo de datos que unifique la información de diferentes repositorios con diferentes estándares en un punto de acceso único para su uso secundario.

Al trabajar con RWD, es importante evaluar la calidad de los datos en seis dimensiones: consistencia, completitud, unicidad, estabilidad temporal, corrección y estabilidad multifuente. La evaluación de la calidad del dato permite determinar si los datos son precisos, únicos y consistentes con las reglas semánticas establecidas, y estudiar su variabilidad en función del tiempo y de su origen [7].

Las fuentes de RWD permiten explorar las enormes oportunidades que la aplicación de la Inteligencia Artificial (IA) puede aportar al cuidado de la DM y a la mejora de los tratamientos existentes [8]. Los sistemas de aprendizaje de máquina, como parte de la IA, son capaces de reconocer patrones y extraer conocimiento automáticamente de los datos. Estas técnicas incluyen los análisis de clústeres, clasificados como métodos no supervisados, que permiten la estratificación de grandes conjuntos de datos en base a características comunes [9].

En la literatura existen varios trabajos que aplican técnicas de *clustering* para la caracterización conjunta de pacientes con DT1 y DT2 [10,11] o de DT2 específicamente [12,13]. Según nuestro conocimiento, hasta la fecha no se ha realizado la caracterización en población española.

En este trabajo se presenta la clasificación en subgrupos de dos sub-cohortes de población española con DT1 y DT2, lo que permite analizar las diferencias en la evolución de la enfermedad, su gestión y el riesgo de comorbilidades.

2. Materiales y métodos

2.1. Fuentes de datos y herramientas

Este trabajo reutiliza datos de DM de la HCE y del repositorio OMOP del Hospital 12 de Octubre de Madrid. Para la identificación de los datos se usaron las terminologías estándar SNOMED, LOINC y CIE-9. El trabajo ha sido desarrollado con los lenguajes de programación SQL 16.0, R 4.2.1 y Python 3.10.12.

2.2. Identificación de la cohorte

Para la definición de la cohorte se establecieron dos criterios de inclusión:

- Pacientes con un diagnóstico único de DT1 o DT2, excluyendo a pacientes con un diagnóstico de DM genérica o con múltiples diagnósticos contradictorios.
- Pacientes mayores de 18 años.

En el repositorio OMOP se identificaron tres conceptos relacionados con el diagnóstico de DM, correspondientes con DT1, DT2 y DM genérica. A partir de estos tres conceptos, se determinaron en el repositorio OMOP los pacientes que cumplían los criterios de inclusión.

Adicionalmente, se extrajo información demográfica, de resultados de laboratorio y de comorbilidades. Para la identificación de estos conceptos de interés, se consideraron los mismos empleados en el proyecto HTx (*Next Generation Technology Assessment, grant agreement* No 825162). El Proyecto HTx es una iniciativa europea que se centra en la evaluación de tecnologías sanitarias de próxima generación que trabaja con dos grandes repositorios: el *Clinical Practice Research Datalink*, con información de la práctica clínica de pacientes con DT1 y DT2 del Reino Unido; y el *T1D Exchange Registry* con información longitudinal de pacientes con DT1 de Estados Unidos.

El repositorio OMOP carecía de tres conceptos esenciales identificados en HTx: el Índice de Masa Corporal (IMC), la Hemoglobina Glicosilada en sangre (HbA1c) y el colesterol en LDL (*Low-density lipoprotein*). Para

cumplimentar estos conceptos se realizaron extracciones de la HCE y se aplicaron las transformaciones de estructura y de unidades necesarias para cumplir el estándar OMOP del conjunto de datos.

2.3. Procesamiento de los datos

En el caso de registros repetidos de diagnóstico de DM o de comorbilidades, se consideró correcto el primero en el tiempo. Los registros de laboratorio fueron evaluados para detectar inconsistencias entre los datos extraídos del repositorio OMOP y los extraídos de la HCE, así como diferencias en unidades de medida y formatos de fecha.

Los valores de IMC requirieron un tratamiento exhaustivo, ya que se calcularon a partir de los registros de peso y altura de los pacientes. Para ello, los registros fueron unificados por paciente y fecha, posteriormente se gestionaron otras inconsistencias, como valores biológicamente imposibles, y se normalizaron las unidades previo al cálculo del IMC.

Después del procesamiento, se llevó a cabo un análisis descriptivo y de calidad para evaluar el conjunto de datos resultante. Con relación a la calidad, se evaluaron tres dimensiones: completitud, que indica como de completa es la información; unicidad, que evalúa si un registro es único dentro del conjunto; y la consistencia, que evalúa si la información es consistente dentro del conjunto y con las reglas de la semántica.

Finalmente se crearon dos conjuntos, uno para DT1 y otro para DT2, ambos con una estructura basada en episodios. Cada episodio representaba una visita anual con información sobre las variables de interés. Para la definición de cada episodio, se asoció el registro disponible más cercano al final del año de cada variable.

Se definieron cuatro conceptos como determinantes para la técnica de *clustering*:

- Para DT1: edad en el diagnóstico, duración de la diabetes, IMC y HbA1c.
- Para DT2: edad en el diagnóstico, IMC, HbA1c y la medida metabólica de resistencia a la insulina (*Metabolic score for Insulin Resistance, METS-IR*), calculada a partir de la fórmula de la Figura 1.

$$METS - IR = \ln \left(\frac{(2 * Glucosa + Triglicéridos) * IMC}{Colesterol\ en\ HDL} \right)$$

Figura 1. Fórmula METS-IR [14].

Los valores nulos y atípicos se descartaron en ambos conjuntos, definiéndose estos últimos como aquellos +/- 3 desviaciones estándar (SD) desde la media. Luego, ambos conjuntos se normalizaron con la función *StandardScaler*.

2.4. Aplicación de técnicas de *clustering*

Se usó la técnica *K-means* para la aplicación de *clustering*, tratándose de un algoritmo no supervisado que divide un conjunto de datos en K clústeres definidos por un centroide. Luego, cada elemento del conjunto se asigna a un clúster en función del centroide que tenga más cerca.

Esta técnica requiere indicar el número de clústeres a identificar como parámetro. Se usaron dos métodos para

determinar el número óptimo: el método del Codo, que representa el valor de la inercia de varias combinaciones de clústeres en una curva lineal, en la que se busca un cambio brusco o “codo”; y el método de la Silueta, que calcula el coeficiente de Silueta para cada número de clústeres usando la técnica *K-means*, y se busca el mayor coeficiente medio.

3. Resultados y discusión

Del repositorio OMOP con 965,374 pacientes, 35.224 (3,65%) tenían al menos un diagnóstico de DM. En cuanto a unicidad, el 6,28% de los pacientes tenían diagnósticos múltiples o repetidos del mismo tipo de DM, mientras que el 0,47% tenía diagnósticos de ambos tipos al mismo tiempo. Estos últimos pacientes se descartaron, ya que no cumplían con los criterios de inclusión de la cohorte.

Luego, la consistencia se evaluó para los registros de laboratorio e IMC, detectando 7,17% y 40,50% de los registros inconsistentes, respectivamente. Estos registros correspondían a valores biológicamente imposibles y fueron descartados, lo que no supuso un gran impacto, ya que la mayoría de los pacientes tenían al menos un registro consistente.

En cuanto a la completitud, se encontró que los registros de IMC suponían la mayor limitación, ya que solo el 47,60% de los pacientes tenían datos de esta variable.

Después del procesamiento de los datos, las cohortes se componían de 1.211 pacientes con DT1 y 34.008 pacientes con DT2. Sin embargo, para el análisis de clústeres solo se tuvieron en cuenta los pacientes con registros completos, siendo 524 con DT1 y 8.915 con DT2.

En el análisis descriptivo se encontraron diferencias en las tendencias de algunas variables. La mayoría de los pacientes con DT1 fueron diagnosticados antes de los 60 años (87,26%), mientras que gran parte de los pacientes de DT2 fueron diagnosticados después (57,77%). El 52,54% de los pacientes con DT1 fueron clasificados inicialmente dentro de las categorías de sobrepeso y obesidad, en contraste con el 82,01% de pacientes con DT2.

No se detectó una diferencia notable entre mujeres y hombres al ser evaluados de forma separada, excepto en la prevalencia de comorbilidades. Las mujeres tuvieron una mayor prevalencia de depresión e insuficiencia cardíaca, mientras que los hombres la tuvieron en hipertensión, infarto de miocardio y accidente cerebrovascular.

• Diabetes tipo 1

El conjunto final de DT1 consistió en 524 pacientes, con un periodo de observación de 12 años y una media de 6,48 años de seguimiento por paciente. Los resultados del método del Codo y de la Silueta para este conjunto no fueron concluyentes, dando un rango entre 3 y 5 clústeres. Por tanto, se decidió aplicar la técnica de *clustering K-means* con las tres configuraciones. La técnica fue aplicada al primer año de seguimiento, con 181 pacientes.

En la Tabla 1 pueden verse los resultados obtenidos con la configuración considerada más significativa, compuesta por 5 clústeres caracterizados por: 1) el menor valor medio

de HbA1c e IMC, en rangos de buen control glucémico y normopeso, respectivamente; 2) la mayor duración de la DM con IMC en rango de sobrepeso; 3) el mayor valor de HbA1c; 4) la mayor edad en el diagnóstico y el mayor IMC en rango de obesidad; 5) una combinación que no destaca en ninguna variable, pero con IMC en rango de normopeso y mal control glucémico.

Clúster (N)	Edad diagnóstica (años)	Duración diabetes (años)	HbA1c (%)	IMC (kg/m ²)
1 (44)	42,1 (±12,1)	2,6 (±2,8)	6,6 (±0,9)	23,7 (±3,6)
2 (11)	17,9 (±10,5)	42,4 (±8,4)	7,9 (±1,1)	28,1 (±4,6)
3 (25)	34,4 (±11,4)	8,3 (±6,2)	10,5 (±1,2)	24,4 (±5,2)
4 (52)	53,9 (±13,6)	7,5 (±6,3)	7,4 (±1,1)	32,6 (±4,7)
5 (49)	19,7 (±9,4)	13,7 (±5,8)	7,6 (±0,8)	24,4 (±3,5)

Tabla 1. Valores medios y desviaciones estándar de los 5 clústeres en cada variable para DT1.

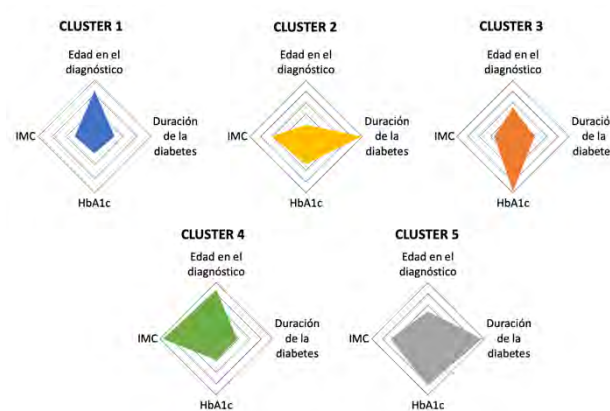


Figura 2. Gráficos de radar de cada clúster con las variables normalizadas para DT1.

Se analizaron las comorbilidades con una prevalencia mínima del 5% fueron: depresión, insuficiencia cardíaca e hipertensión. El clúster con mayor incidencia de todas ellas fue el caracterizado por la mayor edad en el diagnóstico y mayor IMC.

Finalmente, se evaluó la evolución temporal de estos grupos mediante un diagrama de *Sankey*. La mayoría de los pacientes permanecieron en el clúster inicial, pero destaca la transferencia de pacientes desde el grupo identificado con el mayor valor de HbA1c al resto de los grupos.

• Diabetes tipo 2

El conjunto de datos final incluyó 8.915 pacientes, con un periodo de seguimiento de 10 años y una media de 6,37 años de seguimiento por paciente. El resultado del método del Codo para este conjunto tampoco fue concluyente, con un rango de 3 a 5 clústeres. Sin embargo, con el método de la Silueta se determinó que el número óptimo era 3. Por ello, se aplicó la técnica de *clustering K-means* con este número y, en este caso, en el tercer año de seguimiento, ya que tenía un mayor número de pacientes (1.816).

En la Tabla 2 pueden verse los resultados obtenidos con la configuración de 3 clústeres, caracterizados por: 1) la

mayor edad en el diagnóstico; 2) el mayor valor medio de IMC y METS-IR; y 3) el mayor valor medio de HbA1c.

Clúster (N)	Edad diagnóstica (años)	METS-IR	HbA1c (%)	IMC (kg/m ²)
1 (778)	64,5 (±11,0)	40,7 (±6,4)	6,3 (±0,7)	26,5 (±3,5)
2 (560)	58,7 (±11,8)	60,6 (±9,6)	6,8 (±0,7)	36,7 (±4,0)
3 (478)	51,4 (±9,8)	46,1 (±7,3)	8,2 (±0,7)	28,3 (±3,4)

Tabla 2. Valores medios y desviaciones estándar de los 3 clústeres en cada variable para DT2.



Figura 3. Gráficos de radar de cada clúster con las variables normalizadas para DT2.

Las comorbilidades con una prevalencia mínima del 5% fueron: depresión, insuficiencia cardíaca, accidentes cerebrovasculares e hipertensión. El clúster con mayor incidencia fue el de mayor IMC y METS-IR.

Finalmente, la evolución temporal de los grupos se evaluó mediante un diagrama de Sankey. La mayoría de los pacientes permanecieron en el clúster inicial.

4. Conclusiones

En este trabajo se llevó a cabo una reutilización efectiva de los RWD recogidos del repositorio OMOP que, en conjunto con la aplicación de técnicas de *clustering*, demuestra ser un enfoque prometedor para futuros trabajos de investigación.

Con los resultados obtenidos se demuestra que las variables determinantes usadas en el análisis de clústeres permiten caracterizar y estratificar de manera más específica a pacientes con DT1 y DT2. En base a las características de los subconjuntos encontrados, y partir del conocimiento médico experto, podrían establecerse una serie de acciones específicas para cada subgrupo o asignar un nivel de riesgo diferente que permita un seguimiento más intensivo en los casos que lo requieran. Una mejor estratificación de los pacientes con diabetes puede ayudar a los profesionales clínicos a detectar aquellos casos que requieran un seguimiento más exhaustivo o un cambio en la gestión de la enfermedad, permitiendo evaluar también su evolución y progreso.

Para concluir, es esencial mencionar que la metodología propuesta en este trabajo es reutilizable, ya que la extracción de los datos y su procesamiento no depende de la patología. Esto nos permitirá reproducir los análisis en otras cohortes de DM de diferentes instituciones, nacionalidades, o incluso en otras patologías, así como obtener resultados más robustos cuando una cohorte más grande esté disponible.

Referencias

- [1] Página web de International Diabetes Federation. (IDF Diabetes Atlas, 10th edn. Brussels, Belgium: 2021). <https://www.diabetesatlas.org> (Consulta: Mayo 2023).
- [2] Aljehani, M.M., Mohamed, A.M., Ali, A.H., Alhaeti, H.R., Alribh, F.M., Alharbi, F.A., Alassiri, K.M., Nasser, A.M., Alghamdi, O.A., & Al-Hassan, F. (2019). Prevalence of Diabetic Complications in Relation to Risk Factors among Patients with Type 2 Diabetes at University Hospital in Riyadh, Saudi Arabia.
- [3] E. Hariton and J. J. Locascio. Randomised controlled trials - the gold standard for effectiveness research. *BJOG*, vol 125, no 13, pp 1716–1716, 2018. (DOI: 10.1111/1471-0528.15199).
- [4] Miguel Pedrera Jiménez. Propuesta de metodología para el registro, la gestión y la reutilización de datos de salud basada en modelos de arquitectura dual. Doctorado en Ingeniería Biomédica, Universidad Politécnica de Madrid, Madrid, 2022.
- [5] F. Liu and D. Panagiotakos. Real-world data: a brief review of the methods, applications, challenges, and opportunities. *BMC Med Res Methodol*, vol 22, no 1, pp 287, 2022. (DOI: 10.1186/s12874-022-01768-6).
- [6] Página web de Observational Health Data Sciences and Informatics. Standardized Data: The OMOP Common Data Model. <https://www.ohdsi.org> (Consulta: Mayo 2023).
- [7] David Moner. Calidad de datos de salud [Diapositivas]. *Veratech for Health*. www.veratech.es (Consulta: Mayo 2023).
- [8] T. Zhu, K. Li, P. Herrero, and P. Georgiou. Deep Learning for Diabetes: A Systematic Review. *IEEE J Biomed Health Inform*, vol 25, no 7, pp 2744–2757, 2021. (DOI: 10.1109/JBHI.2020.3040225).
- [9] C. M. Bishop. *Pattern Recognition and Machine Learning*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006. (ISBN-10: 0-387-31073-8).
- [10] Ahlqvist, Emma et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *The lancet. Diabetes & endocrinology*, vol 6,5, pp 361-369, 2018. (DOI:10.1016/S2213-8587(18)30051-2).
- [11] Zaharia, Oana P et al. Risk of diabetes-associated diseases in subgroups of patients with recent-onset diabetes: a 5-year follow-up study. *The lancet. Diabetes & endocrinology*, vol 7,9, pp 684-694, 2019. (DOI:10.1016/S2213-8587(19)30187-1).
- [12] Dennis, John M et al. Disease progression and treatment response in data-driven subgroups of type 2 diabetes compared with models based on simple clinical features: an analysis using clinical trial data. *The lancet. Diabetes & endocrinology*, vol 7,6, pp 442-451, 2019. (DOI:10.1016/S2213-8587(19)30087-7).
- [13] Bello-Chavolla, Omar Yaxmehen et al. Clinical characterization of data-driven diabetes subgroups in Mexicans using a reproducible machine learning approach. *BMJ open diabetes research & care*, vol 8,1, pp e001550, 2020. (DOI:10.1136/bmjdr-2020-001550).
- [14] Bello-Chavolla, Omar Yaxmehen et al. METS-IR, a novel score to evaluate insulin sensitivity, is predictive of visceral adiposity and incident type 2 diabetes. *European journal of endocrinology*, vol 178,5, pp 533-544, 2018. (DOI:10.1530/EJE-17-0883).

Control por bloques Adaptativo por Modelo de referencia *In Silico* para pacientes con Diabetes Tipo 1

Misael G. Jacinto Mata¹, Carlos E. Castañeda¹, Onofre Orozco-López¹, Agustín Rodríguez-Herrero^{2,3}

¹ Universidad de Guadalajara Centro Universitario de los Lagos, Lagos de Moreno, Jalisco México. misael.jacinto5670@alumnos.udg.mx, {carlo.e.castanedah, juan.onofre}@academicos.udg.mx.

² Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España; agustin.rodriguez@upm.es

³ Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, España.

Resumen

En este trabajo se presenta un controlador adaptativo por modelo de referencia basado en la transformación del modelo no lineal a la forma controlable por bloques (MRAC-NBC) y unilateral. El controlador está formado por dos lazos, uno donde se ajusta el seguimiento del modelo del paciente y el otro donde se propone la glucosa de seguimiento para el paciente virtual. Las ganancias del NBC son ajustadas de forma heurística mediante la comparación de los estados del modelo de referencia con respecto a la estimación de los estados del paciente. En el lazo del modelo de referencia la dinámica se obtiene mediante el modelo de Bergman con un conjunto medio de parámetros, mientras que el lazo de paciente utiliza el mismo modelo pero parametrizado de forma diferente. El MRAC-NBC respecto del NBC es capaz de conducir la glucemia del paciente a una glucosa objetivo independientemente de la variabilidad interpaciente y de las ingestas, aumentando el Target In Range en más del 18 %.

1. Introducción

La diabetes tipo 1 (DMT1) es una enfermedad crónica caracterizada por altos niveles de glucemia que ocasionan complicaciones a medio y largo plazo causado por la falta de producción de insulina. El control de la normoglucemia se puede realizar con insulina exógena vía subcutánea mediante bomba de infusión (CSII, por sus siglas en inglés, *continuous subcutaneous insulin infusion*). Una terapia con CSII está formada por: 1) bolos de insulina para absorber los carbohidratos (CHOs); y 2) insulina basal para conseguir normoglucemia entre ingestas y en periodos de ayuno [1].

Las personas con DMT1 necesitan evaluar los efectos de la terapia continuamente mediante frecuentes medidas de glucosa, típicamente con la monitorización continua de glucosa (CGM, por sus siglas en inglés, *continuous glucose monitoring*) y en función de ellas tomar continuamente decisiones sobre cambios en CSII, este procedimiento es manual y se conoce como control en lazo abierto. Para evitar la continua toma de decisiones CGM se conecta mediante un algoritmo de control automático a CSII, esto se conoce como control en lazo cerrado (LC) y el sistema que lo implementa se llama páncreas artificial (PA) [2]. Un PA tiene una serie de retos a resolver para mantener normoglucemia: control unidireccional, control repetitivo, control de la glucemia estimada por CGM, retardos, calibraciones de los sensores, perturbaciones persistentes, variabilidad inter- e intra-paciente,

personalización de los algoritmos de control y problemas tecnológicos [3, 4]. Algoritmos de control típicos en PA son control reactivo, inteligencia artificial, controles predictivos y control no lineal basado en modos deslizantes, entre los más relevantes todos ellos citados en [1]. Algunos de estos algoritmos requieren de modelos matemáticos de la dinámica insulina-glucosa para su diseño o para ser personalizados a los pacientes. Los modelos contienen parámetros y ecuaciones no lineales para describir el comportamiento de la dinámica glucosa-insulina, cada conjunto de parámetros define lo que se denominará “paciente virtual”. El modelo de Bergman se utiliza para definir la dinámica glucosa-insulina de una persona sana [1]:

$$\begin{aligned} \dot{G} &= -p_1 [G - G_b] - XG + d(t) \\ \dot{X} &= -p_2 X + p_3 [I - I_b] \\ \dot{I} &= -\eta [I - I_b] + \gamma [G - h]^+ \tau + \frac{S_1}{t_{max,I} V_I} \\ \dot{S}_1 &= -\frac{S_1}{t_{max,I}} + \frac{S_2}{t_{max,I}} \\ \dot{S}_2 &= -\frac{S_2}{t_{max,I}} + u(t) \end{aligned} \quad , \quad (1)$$

donde G es la glucemia (mg/dl); $d(t)$ es la ingesta de CHOs (g) (la perturbación al sistema); $u(t)$ es la ley de control que representa la infusión subcutánea de insulina (IU); X representa la acción de insulina sobre la eliminación de glucosa (min^{-1}). Con γ y p_1 iguales a 0 el modelo se ajusta a una persona con DMT1.

Los controladores adaptativos por modelo de referencia no han sido referenciados sistemáticamente antes del año 2019 [5] aunque ha habido intentos simulados de su aplicación en la modalidad de seguimiento [6].

En este trabajo exploramos un controlador adaptativo por modelo de referencia aplicado a un controlador no lineal por bloques (MRAC-NBC) basado en la transformación del modelo no lineal a la forma controlable por bloques como controlador primario. Dicho controlador ha sido testeado sobre pacientes virtuales que usa información de las ingestas y de variaciones de los parámetros del modelo. Los resultados muestran robustez ante perturbaciones externas, variaciones de paciente y dinámicas no modeladas. La robustez del controlador ya ha sido demostrada en trabajos aplicados en otros ámbitos [7, 8]. Las principales aportaciones del trabajo son: 1) dos lazos de control, uno ajusta glucosa del paciente y el otro las ganancias del NBC; 2) los lazos disponen de sendos modelos de Bergman parametrizados de diferente forma y 3) sistema de control híbrido.

2. Materiales y métodos

En esta sección se describe el procedimiento para diseño del controlador no lineal por bloques (NBC, por sus siglas en inglés de *Non linear Block Controller*) adaptativo por modelo de referencia (MRAC, *model reference adaptive controller*), el cual es aplicado a pacientes virtuales con diabetes tipo 1. Al controlador adaptativo se le referirá como MRAC-NBC.

2.1. Diseño del modelo de referencia

El modelo de referencia se diseña usando como controlador primario, uno basado en NBC [7], una técnica de linealización por realimentación de estado que puede ser aplicada a modelos lineales y no lineales. Con NBC se transforma el sistema dinámico (1) en uno equivalente, donde las nuevas variables de estado representan la dinámica del error entendido como las desviaciones respecto de los valores deseados. Debido al espacio limitado de este documento, esta transformación de variables no se explicita en este trabajo pero se puede consultar con detalles en [7]. Entonces, reasignando las variables de estado como $x_1 = G$, $x_2 = X$, $x_3 = I$, $x_4 = S_1$ y $x_5 = S_2$ y transformando el sistema (1) a la forma controlable por bloques, queda de la siguiente forma general:

$$\begin{aligned} \dot{e}_i &= k_i e_i + b_i e_{i+1} \\ \dot{e}_r &= \bar{f}_r(e, \mathbf{x}_{id}, \dot{\mathbf{x}}_{id}, k_i, \theta_i) + \bar{b}_{r-1}(e) \mathbf{u}(t) + \bar{\varphi}_r(\theta) \end{aligned} \quad (2)$$

donde $e_i = \mathbf{x}_{id} - \mathbf{x}_i$ con $i = 1, 2, \dots, r - 1$ que representa el i -ésimo vector de error de seguimiento, con \mathbf{x}_{id} como la señal deseada para cada i -ésimo vector de estado, con $\bar{\varphi}_r(\theta) = \theta_7 r + \theta_8 d(t) + \theta_9$, y la entrada $\mathbf{u}(t)$ es un bloque escalar. Los elementos k_i son componentes de la matriz diagonal \mathbf{K} la cual es Hurwitz y garantiza la convergencia del sistema en LC [8]. Para encontrar la ley de control de la forma controlable por bloques que provienen del sistema (2), es necesario despejar $u(t)$ no sin antes igualar \dot{e}_5 a $v(t)$:

$$\begin{aligned} u_m(t) &= [\theta_1(k_1)e_{m1} + \theta_2(k_1, k_2)e_{m2} \\ &+ \theta_3(k_1, k_2, k_3)e_{m3} + \theta_4(k_1, k_2, k_3, k_4)e_{m4} \\ &+ \theta_5(k_1, k_2, k_3, k_4)e_{m5} - v + \theta_7 r + \theta_8 d(t) + \theta_9] \theta_6^{-1} \end{aligned} \quad (3)$$

Si $v(t) = k_5 e_{m5} + k_r r$ y se sustituye la señal de control $u_m(t)$ (3) en la forma general controlable por bloques (2), el sistema en LC que representa al modelo de referencia queda de la siguiente forma:

$$\dot{e}_m = \mathbf{A}_m e_m + \mathbf{B}_m r \quad (4)$$

desarrollando el modelo de referencia (4) se tiene:

$$\begin{bmatrix} \dot{e}_{m1} \\ \dot{e}_{m2} \\ \dot{e}_{m3} \\ \dot{e}_{m4} \\ \dot{e}_{m5} \end{bmatrix} = \begin{bmatrix} k_1 & b_1 & 0 & 0 & 0 \\ 0 & k_2 & b_2 & 0 & 0 \\ 0 & 0 & k_3 & b_3 & 0 \\ 0 & 0 & 0 & k_4 & b_4 \\ 0 & 0 & 0 & 0 & k_5 \end{bmatrix} \begin{bmatrix} e_{m1} \\ e_{m2} \\ e_{m3} \\ e_{m4} \\ e_{m5} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ k_r \end{bmatrix} r. \quad (5)$$

Para el caso del controlador de lazo de seguimiento o del modelo de referencia, las variables de estado del modelo de paciente para el análisis son $\mathbf{x}'(t) =$

$[G' \ X' \ I' \ S'_1 \ S'_2]^\top$ y para el controlador para el lazo de paciente son $\mathbf{x}(t) = [G \ X \ I \ S_1 \ S_2]^\top$ y donde la ley de control para éste último es $u(t)$ teniendo como parámetros adicionales el valor de $k_5(t)$ y $k_r(t)$ que se van ajustando a cada iteración.

2.2. Mecanismo de Adaptación

Para este trabajo se utiliza la ley de adaptación con dos parámetros utilizando el procedimiento de diseño del método de estabilidad de Lyapunov [9], el cual consiste de la siguientes etapas: 1) encontrar la ecuación de error en las variables de estado de la forma controlable a bloques de la forma $e_s(t) = e - e_m$, que representa el error de estado entre los vectores de estado del lazo de paciente e y el lazo de seguimiento e_m ; 2) proponer una función candidata de Lyapunov como una función del error entre las variables de estado de la forma controlable a bloques ($e_s(t)$) y del error entre los parámetros del proceso y del modelo $\phi = [(A_{lc}(t) - A_m) \ (B_{lc}(t) - B_m)]$. En su forma más simple, esta función toma la forma de $V(e_s, \phi) = \frac{1}{2} [\gamma e_s^T P e_s + \phi^T \phi]$ donde P es una matriz definida positiva la cual es solución de la ecuación Lyapunov presentada en [9], y que depende de A_m ; 3) calcular la derivada de la función de Lyapunov que generalmente toma la forma de $\dot{V} = -\frac{\gamma}{2} e_s^T Q e_s + \delta$, donde δ representa algunos términos extra incluyendo ϕ . El primer término garantiza que la derivada es definida negativa por lo que haciendo el resto igual a cero se obtiene una posible solución, y siendo Q la matriz identidad definida positiva; y 4) obtener la ley de adaptación obtenida al hacer los términos extra igual a cero, por consiguiente la ley de adaptación es:

$$\dot{k}_5(t) = \gamma_A (p_{51} e_{s1} + p_{52} e_{s2} + p_{53} e_{s3} + p_{54} e_{s4} + p_{55} e_{s5}) e_5 \quad (6)$$

$$\dot{k}_r(t) = \gamma_B (p_{51} e_{s1} + p_{52} e_{s2} + p_{53} e_{s3} + p_{54} e_{s4} + p_{55} e_{s5}) r \quad (7)$$

donde $\dot{k}_5(t)$ es obtenida de $\dot{A}_{lc} = -\gamma_A P e_s e^\top$ (la matriz de estado en LC), y $\dot{k}_r(t)$ de $\dot{B}_{lc} = -\gamma_B P e_s r$ (la matriz de entrada en LC); r es valor deseado y γ es la tasa de velocidad de adaptación para cada parámetro. Debido al espacio limitado de este documento el procedimiento para obtener el mecanismo de adaptación (6)-(7) no se desarrolla con detalle en este trabajo, para mayor información se debe consultar [9]. La Figura 1 describe el diagrama de bloques del MRAC-NBC, está conformado por dos grandes macrobloques: 1) lazo de seguimiento (macrobloque naranja de la Figura 1) que define el modelo de referencia, la finalidad del macrobloque es situar la glucemia del modelo de paciente representado por el modelo no lineal de Bergman (1) ($\mathbf{x}'(t)$), en niveles de normoglucemia; y 2) lazo de paciente (macrobloque azul de la Figura 1) que define el control glucémico del paciente, también está diseñado por la forma no lineal controlable por bloques, utiliza el vector de estado $\mathbf{x}(t)$ y es alimentado además por las variables adaptativas $k_5(t)$ y $k_r(t)$.

Las variables adaptativas $k_5(t)$ y $k_r(t)$ son calculadas por un algoritmo de adaptación (6)-(7) que corregirá los errores de seguimiento entre las variables de estado del modelo del paciente del lazo de seguimiento (modelo de referencia) y el paciente virtual del lazo de paciente modificando las ganancias en línea del controlador.

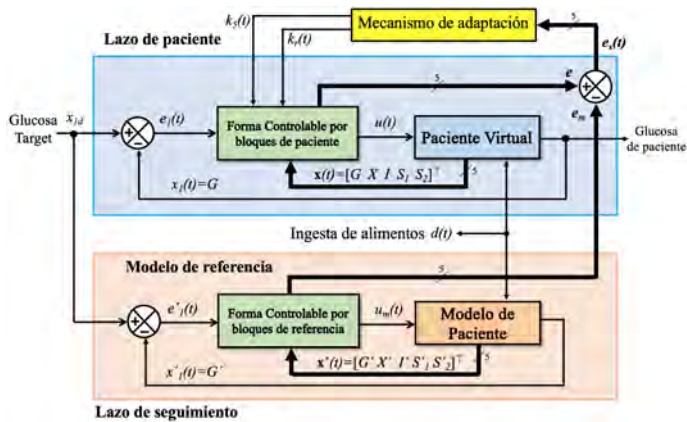


Figura 1. Diagrama de bloques del MRAC-NBC, se detallan los macrobloques de lazo de paciente (azul) y lazo de seguimiento (naranja). Además en la parte superior se muestra el bloque del mecanismo de adaptación (amarillo).

2.3. Diseño de experimentos

El algoritmo de control es implementado en Matlab® (version R2019b, Mathworks, USA) y Simulink® (version 10.0, Mathworks, USA). El método de integración es Runge-Kutta con paso fijo de un segundo.

El lazo de seguimiento se basa en un modelo de paciente modelado con el sistema (1) parametrizado con el paciente medio mostrado en la Tabla 1. Sin embargo, para el lazo de paciente se utiliza como paciente virtual al mismo modelo pero parametrizado con el resto de parámetros relativos a los pacientes virtuales 1, 2 y 3. El rendimiento del MRAC-NBC se compara con el rendimiento del control con ganancias constantes en LC (NBC) sin adaptación, cuando se pretende conseguir en el paciente una glucemia de 90 mg/dl. Para medir el rendimiento se usa la división de la glucosa en las bandas establecidas por el *International Consensus on Time in Range* [10]: TAR *Target Above Range* para glucemias mayores que 180 mg/dl, TIR *Target In Range* para glucemias comprendidas entre 70 y 180 mg/dl y TBR *Target Below Range* para glucemias menores de 70 mg/dl.

Parámetro	Media	Paciente 1	Paciente 2	Paciente 3
p_1	0	0	0	0
p_2	0.0112	0.0123	0.0072	0.0142
p_3	4.48×10^{-8}	8.2×10^{-8}	3.6×10^{-8}	1.656×10^{-6}
γ	0	0	0	0
η	0.2646	0.2659	0.2465	0.2814
$t_{max,I}$		55		

Tabla 1. Parámetros del modelo de Bergman.

3. Resultados y discusión

En esta sección se presenta el rendimiento del MRAC-NBC híbrido y se compara con el rendimiento del NBC

híbrido, en ambos casos se hace uso de la prealimentación informando al algoritmo de algún aspecto de las ingestas. Para el caso del MRAC-NBC se parametrizaron con los pacientes 1, 2 y 3 de la Tabla 1 y ajustando las ganancias $k_5(t)$ y $k_r(t)$ al cual se le aplicó el mecanismo de adaptación (6)-(7) mencionado en la subsección 2.2.

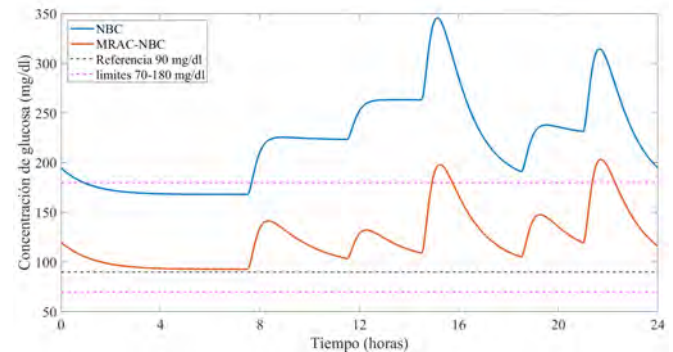


Figura 2. Comparación del desempeño de los controladores NBC vs. MRAC-NBC para el seguimiento de concentración de glucosa $x_1 = G$ para el paciente virtual 1.

En las Figuras 2, 3 y 4 se muestra la comparación del rendimiento de los controladores NBC vs. MRAC-NBC para el seguimiento de concentración de glucosa $x_1 = G$ a una señal de referencia $x_{1d} = 90$ mg/dl para los pacientes virtuales 1, 2 y 3 respectivamente de la Tabla 1. En todas las figuras, en color azul se ve como regula el controlador NBC los niveles de glucosa a lo largo de un día y en color naranja se muestra como mejora la respuesta del MRAC-NBC. Como se aprecia claramente el Paciente 1 (Figura 2) y el Paciente 3 (Figura 4) presentan altas concentraciones de glucosa, incluso en periodo nocturno donde no existen ingestas. Después de las ingestas se aprecian concentraciones de hiperglucemia severa siendo más exagerado el comportamiento del Paciente 1. Asimismo es posible apreciar como el mecanismo de adaptación mejora el rendimiento del controlador al mantener las concentraciones de glucosa lo más cercanas posibles al valor de referencia, ver la tabla 3.

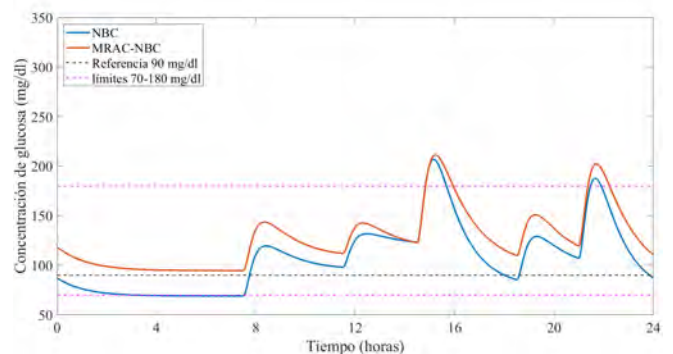


Figura 3. Comparación del desempeño de los controladores NBC vs. MRAC-NBC para el seguimiento de concentración de glucosa $x_1 = G$ para el paciente virtual 2.

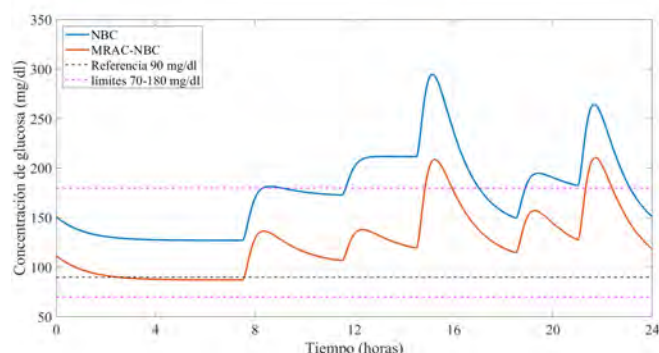


Figura 4. Comparación del desempeño de los controladores NBC vs. MRAC-NBC para el seguimiento de concentración de glucosa $x_1 = G$ para el paciente virtual 3.

	Banda	Paciente 1	Paciente 2	Paciente 3
NBC	TAR	71.8	5.3	43
	TIR	28.2	76.4	57
	TBR	0	18.3	0
MRAC	TAR	7.2	8.2	9
	TIR	92.8	91.8	91
	TBR	0	0	0

Tabla 2. Evaluación de los porcentajes de tiempo por rangos de glucosa.

La comparación del comportamiento de los controladores MRAC-NBC y NBC es cuantificado en la Tabla 2, haciendo constar que para los Pacientes 1 y 3, el controlador MRAC-NBC presenta valores menores que 9% en TAR, valores superiores que 91% en TIR y 0 en TBR, mientras que el controlador NBC presenta valores superiores a 43% en TAR, mayores que 28% en TIR y 0 en TBR.

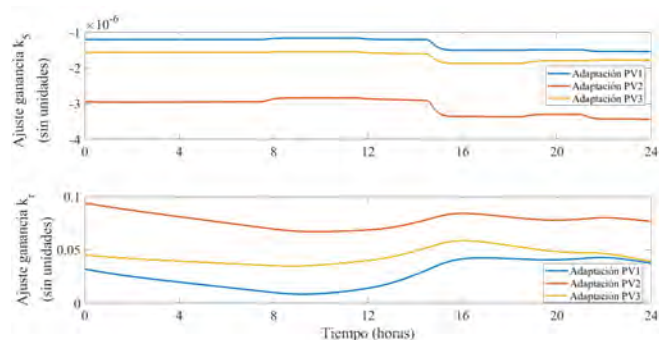


Figura 5. Variación de las ganancias del MRAC-NBC en 24 horas, superior: k_5 ; inferior: k_r .

En la Figura 5 se muestra como el algoritmo de adaptación ajusta las ganancias $k_5(t)$ y $k_r(t)$ a lo largo del día para compensar en el algoritmo de control las diferencias dinámicas entre el modelo de paciente y el lazo de seguimiento.

4. Conclusiones

Cuando se utiliza el mismo modelo para realizar el modelo de referencia y para el lazo de paciente parametrizado de diferente forma, se obtienen grandes beneficios eliminando la presencia de la glucosa en TBR, aumentando considerablemente en TIR y disminuyendo en TAR.

Parece ser que el controlador NBC es poco sensible al target puesto que para ninguno de los tres pacientes virtuales no se alcanza en periodo nocturno, hecho que se consigue con el controlador con seguimiento por tanto el MRAC-NBC es más sensible al modelo de referencia que al propio target.

Dado que el NBC se inicializa con las mismas ganancias para los tres pacientes por lo que no existe un ajuste óptimo para alguno de los casos por ello el TAR es grande. Sin embargo el MRAC-NBC ajusta la ganancia en tiempo real para aumentar el TIR.

El NBC adaptativo es capaz de conducir la glucemia del paciente a una glucosa objetivo independientemente de la variabilidad interpaciente y de las ingestas, aumentando el TIR en más del 18%.

Referencias

- [1] Onofre Orozco-López, Carlos E. Castañeda, Gemma García-Sáez, M Elena Hernando, and Agustín Rodríguez-Herrero. Personalized hybrid artificial pancreas using unidirectional sliding-modes control algorithm. *Biocybernetics and Biomedical Engineering*, 42(4):1218–1235, 2022.
- [2] Ariane Quintal, Virginie Messier, Rémi Rabasa-Lhoret, and Eric Racine. A critical review and analysis of ethical issues associated with the artificial pancreas. *Diabetes & metabolism*, 45(1):1–10, 2019.
- [3] Keren Zhou and Diana Isaacs. Closed-loop artificial pancreas therapy for type 1 diabetes. *Current Cardiology Reports*, 24(9):1159–1167, 2022.
- [4] Munachiso Nwoko and Roman Hovorka. The artificial pancreas and type 1 diabetes. *The Journal of Clinical Endocrinology & Metabolism*, 108(7):1614–1623, 2023.
- [5] Griselda Quiroz. The evolution of control algorithms in artificial pancreas: A historical perspective. *Annual Reviews in Control*, 48:222–232, 2019.
- [6] A Rodríguez-Herrero, C. Pérez-Gandía, EJ. Gómez, and ME. Hernando. A simulation study of an adaptive inverse controller for closed-loop control in type 1 diabetes. 2010.
- [7] Onofre A. Morfin, Riemann Ruiz-Cruz, Jesus I. Hernández, Carlos E. Castañeda, Reymundo Ramírez-Betancour, and Fredy A. Valenzuela-Murillo. Real-time sensorless robust velocity controller applied to a dc-motor for emulating a wind turbine. *Energies*, 14(4):868, 2021.
- [8] Juan Cristobal Alcaraz Tapia, Carlos E. Castañeda, Héctor Vargas Rodríguez, and P. Esquivel. Design of a port-hamiltonian control for an alt-azimuth liquid-mirror telescope. *Mathematics*, 11(16):3443, 2023.
- [9] Karl J. Astrom and Bjorn Wittenmark. *Adaptive Control*. DOVER PUBLICATIONS, INC, 2008.
- [10] Tadej Battelino, Thomas Danne, Richard M Bergenstal, Stephanie A Amiel, Roy Beck, Torben Biester, Emanuele Bosi, Bruce A Buckingham, William T Cefalu, Kelly L Close, et al. Clinical targets for continuous glucose monitoring data interpretation: recommendations from the international consensus on time in range. *Diabetes care*, 42(8):1593–1603, 2019.

Colimación Automática para Sistemas de Radiología mediante Aprendizaje Profundo

C. F. Del Cerro^{1,2}, J.M. Ortega³, M. Desco^{1,2,4,5}, M. Abella^{1,2,4}

¹ Depto. Bioingeniería, Universidad Carlos III de Madrid, Madrid, España

² Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España

³ SEDECAL S.A., Madrid, España

⁴ Centro Nacional Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, España

⁵ Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid, España

Resumen

En la adquisición de una radiografía, el haz de rayos X es colimado dependiendo de la posición radiológica ajustando la zona irradiada. Esta colimación es seleccionada por un técnico de rayos manualmente antes de la adquisición. Una colimación incorrecta puede dar lugar a una degradación del contraste en la imagen o un aumento de la dosis de radiación que recibe el paciente. Este trabajo presenta un método para realizar la colimación del haz de rayos X de forma automática basado en aprendizaje profundo. La detección automática de la región de colimación se realiza a partir de una fotografía de la paciente tomada antes de la adquisición de rayos X. Los resultados sugieren la viabilidad del método para su uso en sistemas reales con el fin de facilitar el flujo de trabajo de adquisición reduciendo la aparición de errores de exposición y evitando dosis de radiación innecesarias administradas a los pacientes.

1. Motivación

La calidad de una radiografía, es decir, el contraste y la relación señal/ruido, depende de la selección óptima de los parámetros de adquisición: el voltaje de la fuente (kVp), la corriente y el tiempo de exposición (mAs), la distancia entre la fuente y el detector y la región de colimación. La selección incorrecta de estos parámetros, que dependen de la posición radiológica, puede dar lugar a imágenes no óptimas para el diagnóstico, que pueden incluso requerir una repetición [1]. Estas repeticiones tienen consecuencias tanto en términos de sobreexposición del paciente como en el mal uso de los recursos hospitalarios, personales y de equipos. Los valores de kVp, mAs y distancia entre la fuente y el detector son ajustados por el técnico de rayos desde la consola de control del equipo en base a recomendaciones del fabricante [2]. Para determinar la región de colimación, el técnico debe entrar a la sala para ajustar el colimador manualmente con la ayuda de una luz que muestra el área que será irradiada antes de realizar la radiografía. La automatización de la selección del área de colimación reduciría los tiempos de adquisición y aseguraría una colimación óptima que resultaría en imágenes mejor calidad de imagen y una menor dosis de radiación ionizante recibida por el paciente.

No existen trabajos previos que hayan abordado la detección automática de la región de colimación de rayos X previa a la adquisición. En [3] los autores proponen un método para delinear el área de colimación en imágenes

radiográficas basado en la detección de bordes. Otros trabajos han abordado un problema relacionado, como es la detección automática de la posición radiológica, mediante redes neuronales convolucionales (CNN) [4-6], también sobre radiografías.

Un enfoque alternativo para la estimación de la región de colimación antes de la adquisición de rayos X, consiste en utilizar una fotografía del paciente. En la literatura, se han propuesto diversos métodos para la detección de partes del cuerpo humano en fotografías. En [7] se presenta un nuevo método para el reconocimiento de partes del cuerpo humano utilizando una CNN para la predicción de regiones. El principal problema de este método es que es relativamente complejo y requiere una gran cantidad de recursos computacionales, lo que podría dificultar su implementación en sistemas reales. También existen métodos para la detección de la posición humana que requieren anotaciones manuales largas y complejas [8].

Este trabajo presenta un método basado en aprendizaje profundo para determinar el área de colimación en una adquisición de rayos X a partir de una fotografía del paciente posicionado en el equipo.

2. Materiales y Métodos

El método propuesto permite detectar automáticamente el área de colimación a partir de una fotografía del paciente, y la muestra superpuesta en la fotografía para su validación por el técnico de rayos. Para minimizar errores en casos complicados, si se detecta una segunda posible área de colimación con una confianza mayor del 40%, el método muestra las dos áreas, para permitir la elección por parte del técnico. A continuación, se describen la creación de la base de datos y el modelo de aprendizaje profundo.

2.1. Base de datos

Para este estudio, se seleccionaron 66 posiciones radiográficas utilizadas habitualmente en la práctica clínica en colaboración con técnicos de rayos del Hospital Universitario Gregorio Marañón (Madrid, España). Se adquirieron 75 voluntarios en dos salas de rayos X: 50 voluntarios en una en el departamento de investigación de SEDECAL en Algete (Madrid) y 25 voluntarios en el laboratorio de radiología de la Universidad Carlos III de Madrid (UC3M). Aunque la adquisición no involucraba

radiación, todos los participantes firmaron un consentimiento legal que reconoce la protección de sus datos según lo establecido en la normativa nacional e internacional (Reglamento General de Protección de Datos, Reglamento UE 2016/679).

La configuración de adquisición de datos en ambas instalaciones fue la siguiente: una mesa de rayos X, un detector de pared y una fuente de rayos X montada en el techo equipada con una cámara *YI 4K Action* fijada a la carcasa del colimador mediante un monopie (Figura 1, arriba). La adquisición fue llevada a cabo por técnicos de rayos del Hospital Universitario Gregorio Marañón que seleccionaron la colimación óptima para cada caso. Se adquirieron dos tipos de fotografías para cada voluntario: una con la luz del colimador apagada y otra con luz del colimador encendida (Figura 1, abajo). Para minimizar la influencia de los elementos del fondo, las imágenes se recortaron a una región cuadrada con una anchura de 90 cm, correspondiente a la anchura de la mesa de rayos X (Figura 1, abajo, cuadrado rojo). A partir de las imágenes con luz del colimador se extrajeron manualmente las coordenadas de los píxeles de las esquinas del área iluminada, correspondientes con la región de colimación, que sirvieron de referencia para la evaluación del modelo.



Figura 1. Disposición de la instalación de rayos X de la UC3M y configuración de la cámara. Posición AP del tórax en con y sin luz de colimador (abajo).

El tamaño de píxel (TP) para calcular el área recortada se calculó mediante la siguiente relación:

$$TP = \frac{TR}{TM} = \frac{A_s \times DFO}{DLS \times TM} = m \times DFO = m \times (DFD - d) \quad (1)$$

donde TM es el tamaño de matriz de la imagen en píxeles, A_s es la altura del sensor de la cámara, DFO es la distancia entre la fuente y el paciente, DLS es la distancia entre la lente y el sensor de la cámara, m es el parámetro de pendiente de la cámara, DFD es la distancia entre la fuente y el detector, y d es la distancia entre el paciente y el detector (Figura 2). El tamaño de píxel, TP , calculado en nuestro caso, fue de 0,0513 cm, 0,0798 cm y 0,0969 cm para una DFD de 100 cm, 150 cm y 180 cm respectivamente.

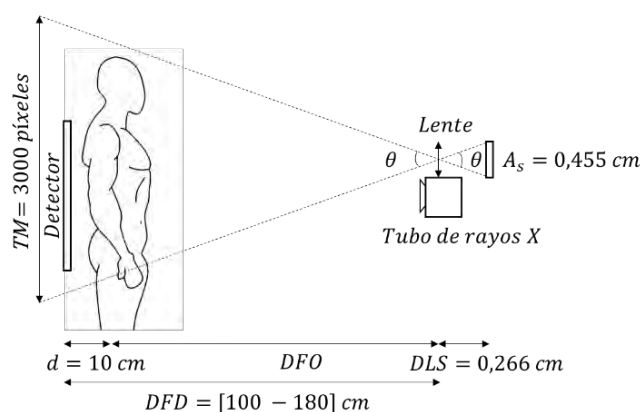


Figura 2. Geometría del sistema de adquisición completo.

2.2. Modelo

La entrada del modelo es la fotografía del paciente con la luz del colimador apagada y salida del modelo son las coordenadas del área de colimación y la confianza de la predicción.

La arquitectura seleccionada fue la versión 5 del modelo *You only look once (YOLOv5)* [9], normalmente aplicado a la segmentación y detección de objetos. Este modelo fue desarrollado en Python utilizando librerías de aprendizaje profundo: *pytorch* [10], *fastai* [12] e *icevision* [13]. Para compensar el limitado tamaño de la base de datos, se aplicó la estrategia de aprendizaje por transferencia, conocida como *fine-tuning*, utilizando una versión del modelo *YOLOv5* preentrenado con la base de datos *ImageNet-1K* [14]. Para adaptar las imágenes de nuestra base de datos a características del modelo preentrenado, fueron reescaladas a 224×224 píxeles y se normalizaron usando la media y la desviación estándar de *ImageNet-1K*.

La función de coste se calculó como una combinación de tres componentes individuales según la propuesto en [15]: un componente para estimar el error en la clasificación de la posición radiológica, a partir de la entropía cruzada binaria, otro componente para calcular errores en la detección del colimador, basado también en la entropía cruzada binaria, y un tercer componente para calcular la localización del colimador mediante la intersección sobre la unión. Para minimizar la función de coste se seleccionó el optimizador Adam [16].

El entrenamiento se hizo en dos etapas: en la primera se entrenó la parte del predictor durante 1 época y en la segunda se entrenó el modelo completo durante 50 épocas. La tasa de aprendizaje óptima fue calculada con el test de Leslie N. Smith [17], dando como resultado 10^{-2} para la primera etapa y 10^{-3} para la segunda etapa. Esta tasa de aprendizaje se dividió por 10 cuando el error de validación no bajó más de 0.1 durante 3 épocas seguidas.

Para llevar a cabo el desarrollo se utilizó un equipo con Windows® con un procesador AMD Ryzen 7 5800X de 8 núcleos, con 32 GB de memoria RAM, y una unidad de procesamiento gráfico (GPU) NVIDIA RTX 3090.

2.3. Metodología de evaluación

Debido al limitado tamaño de la base de datos, se utilizó validación cruzada (VC) con tres conjuntos de validación aleatorios para evitar sesgos en los resultados. En todos los experimentos, la base de datos se dividió en 60 voluntarios para el entrenamiento y 15 para la validación.

Para evaluar el rendimiento del modelo en cada conjunto de validación, se utilizó la media del promedio de precisión (*mPP*). La precisión se refiere a la proporción de predicciones correctas realizadas en relación con el total de predicciones hechas, con valor entre 0 y 1. La precisión se basa en el concepto de intersección sobre unión (*IoU*), que mide la similitud entre dos regiones de interés, siendo 1 cuando las dos regiones coinciden perfectamente y 0 cuando no coinciden en absoluto. Para cada conjunto de validación, se calculó la *mPP* para dos valores de *IoU*: 0,5 y 0,75.

3. Resultados

La Tabla 1 presenta la media (μ) y la desviación estándar (σ) de las métricas de evaluación para cada conjunto de validación durante la validación cruzada.

VC	<i>IoU</i> = 0.5	<i>IoU</i> = 0.75
1	0.914	0.673
2	0.933	0.643
3	0.917	0.705
$\mu \pm \sigma$	0.921 \pm 0.008	0.674 \pm 0.025

Tabla 1. Valores de *mPP* obtenidos por cada conjunto de validación.

No se aprecian cambios estadísticamente significativos en ninguno de los tres conjuntos de validación, lo que respalda la robustez del método propuesto. En promedio, las áreas de colimación predichas coinciden con las de referencia en al menos la mitad de sus píxeles en el 92% de los casos. Sin embargo, esta coincidencia se reduce al 67% cuando se requiere que el área de colimación predicha coincida en al menos tres cuartas partes de sus píxeles con el área de referencia (Figura 3A y 3B, con una coincidencia de 92,1 % y 93,6% de los píxeles, respectivamente). Sin embargo, una inspección visual detallada de los resultados muestra que la métrica *mPP* no representa bien su validez clínica, como se puede ver en la Figura 3C y 3D, que a pesar de tener una coincidencia de píxeles del 60,4 % y 39,3% de píxeles, respectivamente, resultan en una colimación válida para el diagnóstico.

En la Figura 3E se muestra a un ejemplo de una predicción incorrecta del área de colimación que el algoritmo devuelve con un nivel de confianza bajo, quizá debida a los objetos que se ven sobre la mesa. Finalmente, la Figura 3F muestra uno de los casos en los que el modelo predijo dos áreas de colimación con una confianza mayor del 40%, que representan el 7% del conjunto de validación.

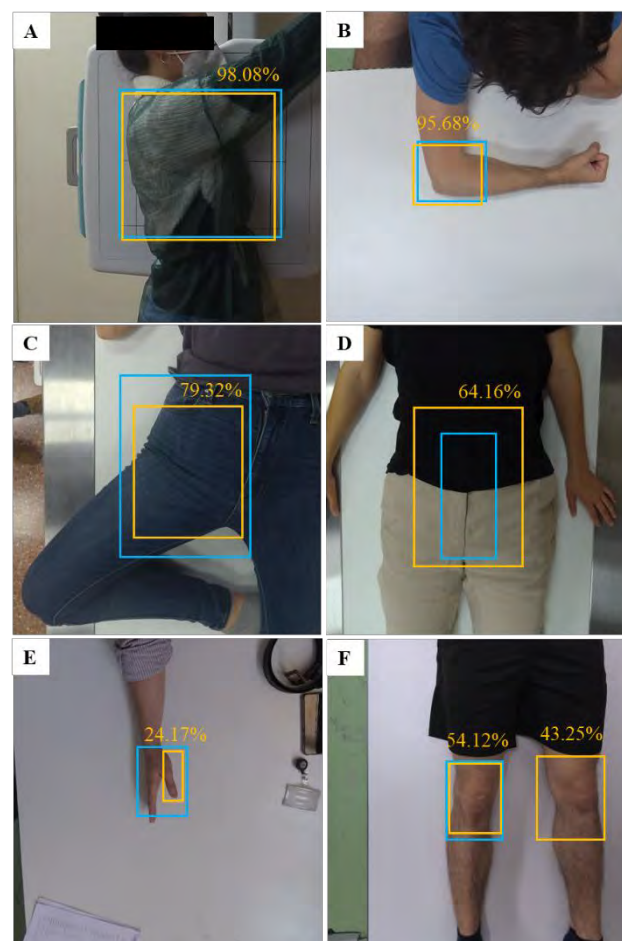


Figura 3. Imágenes de la base de datos con las áreas de colimación de referencia (azul) y predichas por el modelo con su nivel de confianza (amarillo) para seis casos del conjunto de validación.

4. Discusión

Este trabajo presenta un método para la colimación automática en sistemas de radiología basado en una fotografía del paciente, obtenida mediante una cámara fijada al colimador. La fotografía se preprocesa y se introduce en un modelo de aprendizaje profundo que devuelve las coordenadas del área de colimación y la confianza en la predicción en tiempo real. A partir de esta información se muestra el área de colimación sobre la fotografía para su validación por parte del técnico de rayos.

La evaluación preliminar mostró que el método propuesto es capaz de predecir áreas de colimación que coinciden con las de referencia en al menos la mitad de sus píxeles en el 92% de los casos. A pesar de que esto puede parecer insuficiente, la inspección visual mostró que la métrica utilizada no refleja la validez clínica del área de colimación detectada. El trabajo futuro incluye la evaluación visual de los resultados por parte de técnicos de rayos.

El método propuesto puede ser fácilmente implementado en escenarios reales, ya que sólo requiere una cámara, un programa ejecutable instalado en el PC de control del sistema y un nuevo flujo de trabajo con una interfaz de usuario para que el técnico de rayos confirme la región del colimador sugerida. Dado que la radiografía es la modalidad de imagen médica más utilizada en clínica [18],

esta automatización de la adquisición podría tener un impacto significativo en el sistema sanitario en términos de seguridad del paciente y optimización del flujo de trabajo del departamento de radiología.

Agradecimientos

Este trabajo ha sido apoyado por el Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación, proyecto PDC2021-121656-I00 (MULTIRAD), financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea 'NextGenerationEU'/PRTR. También por la Comunidad de Madrid, proyecto S2017/BMD-3867 RENIM-CM, cofinanciado por el Fondo Estructural y de Inversión Europeo. Y por el Instituto de Salud Carlos III a través de los proyectos PMPTA22/00121 y PMPTA22/00118, cofinanciados por la UE 'NextGenerationEU'/PRTR. El CNIC cuenta con el apoyo del Instituto de Salud Carlos III, el Ministerio de Ciencia e Innovación y la Fundación Pro CNIC.

Referencias

- [1] D. H. Foos, W. J. Schnert, B. Reiner, E. L. Siegel, A. Segal, and D. L. Waldman, "Digital radiography reject analysis: data collection methodology, results, and recommendations from an in-depth investigation at two hospitals," *Journal of digital imaging*, vol. 22, pp. 89-98, 2009.
- [2] H. Menzel, H. Schibilla, and D. Teunen, "European guidelines on quality criteria for computed tomography," *Luxembourg: European Commission*, vol. 16262, 2000.
- [3] Y. Yu, J. Wang, and S. T. Acton, "X-ray collimator shutter detection by active-rods," in *2016 IEEE International Conference on Image Processing (ICIP)*, 2016: IEEE, pp. 2350-2354.
- [4] R. Hosch, L. Kroll, F. Nensa, and S. Koitka, "Differentiation between anteroposterior and posteroanterior chest X-ray view position with convolutional neural networks," in *RöFo-Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*, 2021, vol. 193, no. 02: Georg Thieme Verlag KG, pp. 168-176.
- [5] T. J. Saun, "Automated classification of radiographic positioning of hand X-rays using a deep neural network," *Plastic Surgery*, vol. 29, no. 2, pp. 75-80, 2021.
- [6] X. Fang, L. Harris, W. Zhou, and D. Huo, "Generalized Radiographic View Identification with Deep Learning," *Journal of Digital Imaging*, vol. 34, no. 1, pp. 66-74, 2021.
- [7] L. Yang, Q. Song, Z. Wang, M. Hu, and C. Liu, "Hier R-CNN: Instance-level human parts detection and a new benchmark," *IEEE Transactions on Image Processing*, vol. 30, pp. 39-54, 2020.
- [8] C. Zheng *et al.*, "Deep learning-based human pose estimation: A survey," *arXiv preprint arXiv:2012.13392*, 2020.
- [9] Ultralytics. "YOLOv5: Ultrafast Object Detection." <https://ultralytics.com/yolov5> (accessed August, 2023).
- [10] A. Paszke *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *Advances in neural information processing systems*, vol. 32, 2019.
- [11] A. Paszke *et al.*, "Automatic differentiation in pytorch," 2017.
- [12] J. Howard and S. Gugger, "Fastai: A layered API for deep learning," *Information*, vol. 11, no. 2, p. 108, 2020.
- [13] Airtic. "IceVision: An Agnostic Computer Vision Framework." <https://github.com/airctic/icevision> (accessed August, 2023).
- [14] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, 2009: Ieee, pp. 248-255.
- [15] Ultralytics. "YOLOv5: Architecture Description." https://docs.ultralytics.com/yolov5/tutorials/architecture_description/#3-training-strategies (accessed August, 2023).
- [16] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [17] L. N. Smith, "Cyclical learning rates for training neural networks," in *2017 IEEE winter conference on applications of computer vision (WACV)*, 2017: IEEE, pp. 464-472.
- [18] N. England and N. Improvement, "Diagnostic imaging dataset statistical release," in "London: Department of Health," 2023.

Hacia una protonterapia adaptativa: Evaluando el potencial de la nnUNet para la segmentación de órganos en riesgo

B. Rodriguez-Gonzalez¹, B. Rodriguez-Vila¹, N. Malpica¹, A. Torrado-Carvajal¹

¹ Medical Image Analysis and Biometry Lab, Universidad Rey Juan Carlos, Madrid, España, {blanca.rodriguez, borja.rodriguez.vila, norberto.malpica, angel.torrado}@urjc.es

Resumen

La segmentación de órganos en riesgo es un proceso esencial en el tratamiento de protonterapia, especialmente en pacientes con cáncer de cabeza y cuello, donde la presencia de estructuras en riesgo es elevada. Debido a la necesidad de replanificación del tratamiento, así como a la gran carga de trabajo asociada a una delineación manual de las estructuras, es frecuente el uso de algoritmos automáticos de procesamiento de imagen para esta labor. En este marco se encuentra el trabajo propuesto, donde se prueba el potencial de la nnU-Net para la segmentación de 45 estructuras cerebrales. En este sentido, se entrenan cuatro modelos (un modelo bidimensional y tres tridimensionales a diferentes resoluciones) sobre 65 sujetos. Tras la comparación con métodos del estado del arte para la segmentación de estructuras cerebrales, destaca el potencial de la nnU-Net como herramienta para la segmentación de órganos en riesgo en la clínica, no solo por sus resultados, donde se obtiene un Coeficiente de Dice de 0.873 empleando arquitecturas tridimensionales; si no por su reducido tiempo de procesamiento (3:46min) y la robustez arrojada en la fase de test, donde se realiza una evaluación de incertidumbre con el objetivo de determinar la variabilidad de respuesta del modelo ante variaciones sistemáticas en la entrada.

1. Introducción

La terapia con protones, propuesta por Robert R. Wilson en 1946, se ha convertido en una herramienta fundamental en la lucha contra numerosos tipos de cáncer, siendo muy frecuente su uso para los tumores cerebrales y de cabeza y cuello [1]. La protonterapia ofrece mejoras significativas en comparación con las terapias de radiación convencionales debido a sus características físicas, las cuales permiten una mayor precisión al maximizar el poder terapéutico sobre el tumor y reducir el daño en los **órganos en riesgo** (*organs at risk*; OARs) que lo rodean [2].

Los protones poseen una distribución de dosis característica, la cual se concentra en un área específica conocida como pico de Bragg y donde reside el potencial terapéutico de las partículas. La profundidad del pico varía según la energía inicial de los protones, lo que permite ajustar la mayor parte de la dosis en el área tumoral, reduciendo la dosis en los tejidos circundantes. El ajuste de este pico a las características anatómicas del paciente se realiza en una etapa inicial de planificación, la cual se revisa a lo largo de la terapia con el fin de recoger posibles cambios y obtener los mejores resultados posibles [3].

La planificación y simulación del tratamiento de protonterapia comienza con la segmentación del tumor y los OARs que podrían verse afectados, lo que permite calcular la dosis que se depositará en estas estructuras

desde diferentes ángulos de incidencia y determinar la estrategia óptima de irradiación [2]. La segmentación es un paso esencial en este proceso. Por lo general se utiliza la tomografía computarizada (TC) para la planificación. Sin embargo, también es frecuente emplear imágenes complementarias como la resonancia magnética (RM), la cual proporciona una mayor precisión en la delineación de las estructuras de tejido blando.

Tradicionalmente, la segmentación de estructuras se realiza de manera manual por un profesional; sin embargo, debido a la variabilidad de este proceso, el uso de métodos automáticos se ha vuelto una práctica común. En este sentido, la inteligencia artificial ha revolucionado el campo. Desde 2019, se observa un creciente interés en el uso de métodos de **aprendizaje profundo** (*Deep Learning*; DL) para la segmentación de OARs, especialmente en el contexto de tumores de cabeza y cuello, región de elevada complejidad debido a la presencia de numerosas estructuras en riesgo. Destaca el uso de redes neuronales convolucionales basadas en U-Nets tridimensionales sobre TC [4, 5] logrando segmentar entre 8 y 35 estructuras.

Gran parte de los estudios relacionados con la segmentación de OARs se engloba dentro de la radioterapia, sin encontrar ninguna investigación específica para protonterapia, donde una delimitación precisa de las estructuras es crucial. En este marco, se plantea la evaluación de la *no-new UNet* (**nnUNet**) [6], adaptada para esta problemática. Concretamente, se realiza una segmentación de **45 estructuras cerebrales**, de las cuales muchas son evaluadas en la planificación de protonterapia. Además de métricas estándar de rendimiento, se introduce un proceso de **cuantificación de incertidumbre** en la evaluación del método propuesto. En una primera instancia y como prueba de concepto el estudio se realiza sobre imágenes de RM.

2. La base de datos

Para el diseño del conjunto de datos se emplean bases de datos de libre acceso. Se han incluido estudios de RM ponderados en T1, adquiridos sobre sujetos sanos en equipos de 3T con vóxeles isométricos de 1mm.

Atendiendo a los criterios de inclusión mencionados, se genera una base de datos conformada por **89 volúmenes** de RM a partir de tres bases de datos públicas contenidas en [OpenNeuro](#) (ds004516 (1) [7], ds002322 (2) [8] y ds004513 (3) [9]). Se incluye una descripción demográfica de las bases de datos empleadas en la **Tabla 1**.

	Sujetos N	Edad μ ($\pm\sigma$)	Mujeres n (%)	Equipo de RM
1	36	20.27 (1.63)	14 (53.84)	General Electrics
2	43	28.28 (9.64)	24 (55.81)	Siemens Prisma
3	20	34.45 (9.71)	10 (50.00)	Siemens Biograph
	89	27.33 (9.63)	48 (58.93)	

Tabla 1. Descripción demográfica de las bases de datos empleadas y de la base de datos total resultante.

3. Metodología

Con el fin de realizar una segmentación precisa de los OARs, se propone un marco metodológico compuesto por una etapa de preprocesamiento, seguida de las etapas de entrenamiento y evaluación (Figura 1).

3.1. Preprocesado

Los algoritmos automáticos de segmentación se basan, generalmente, en los niveles de intensidad de la imagen; es por ello por lo que se requiere una etapa previa de mejora de la imagen. En primer lugar, se elimina la inhomogeneidad de campo mediante el algoritmo N4ITK implementado en 3DSlicer [10, 11]. Posteriormente, se estandariza el tamaño de los volúmenes a 256x256x256, asegurando mantener todas las estructuras de interés.

El siguiente paso consiste en generar las máscaras multietiqueta de los OARs. Debido al interés por delinear estructuras cerebrales, se emplea FreeSurfer [12], la herramienta por excelencia de segmentación cerebral en numerosos estudios de investigación; sin embargo, debido al elevado tiempo de procesado, esta herramienta no es aplicable en la práctica clínica diaria.

Se procesan los 89 volúmenes de RM con FreeSurfer v7.3.2 en un servidor Intel Core i9 que opera en Ubuntu 22.04.2 LTS, obteniendo las máscaras contenedoras de 45 estructuras cerebrales. Tras un proceso de revisión manual, se reduce el número de sujetos a 74 debido a errores de

segmentación, de los cuales se emplean 65 sujetos para la etapa de entrenamiento y validación, reservando, por tanto, 9 sujetos para la evaluación seleccionados proporcionalmente a la distribución de los equipos de RM empleados en el presente estudio.

3.2. El modelo: nnUNet

La nnUNet es un método de segmentación basado en DL capaz de adaptar automáticamente su estructura a tareas nuevas. De forma simple, el modelo extrae características de la base de datos las cuales utiliza para establecer distintos parámetros del modelo, incluyendo la tasa de aprendizaje o el número de capas de la red. La nnUNet, permite, por tanto, diseñar modelos bidimensionales y tridimensionales a diferentes resoluciones [6].

Para la segmentación de estructuras cerebrales, se entrenan UNets 2D (*nnUNet 2D*) sobre el plano coronal, así como todas las arquitecturas 3D disponibles. El marco de la nnUNet permite el diseño de arquitecturas 3D a alta resolución (*nnUNet 3D full-resolution*), las cuales trabajan con parches que preservan el tamaño de píxel original, y arquitecturas 3D de baja resolución (*nnUNet 3D low-resolution*), trabajando con una imagen muestreada a un mayor tamaño de píxel. En todos los casos se emplea un esquema de validación cruzada de 5 particiones, con 1000 épocas por partición para los modelos 2D y 350 épocas por partición para las 3D en alta y baja resolución. Los resultados de este último modelo se refinan mediante el uso de una U-Net en alta resolución (estructura en cascada; *nnUNet cascade*) entrenada durante 50 épocas adicionales. Esta estructura toma la imagen original junto con las segmentaciones realizadas, comparando su información con el fin de generar segmentaciones de mayor calidad. Todos los modelos fueron entrenados con un optimizador SGD con una tasa inicial de aprendizaje del 0.01 y una combinación del **Coefficiente de Dice** (DC) y la entropía cruzada como función de pérdida. Los modelos se construyen en un servidor local equipado con una NVIDIA GeForce RTX 3080 de 10GB.

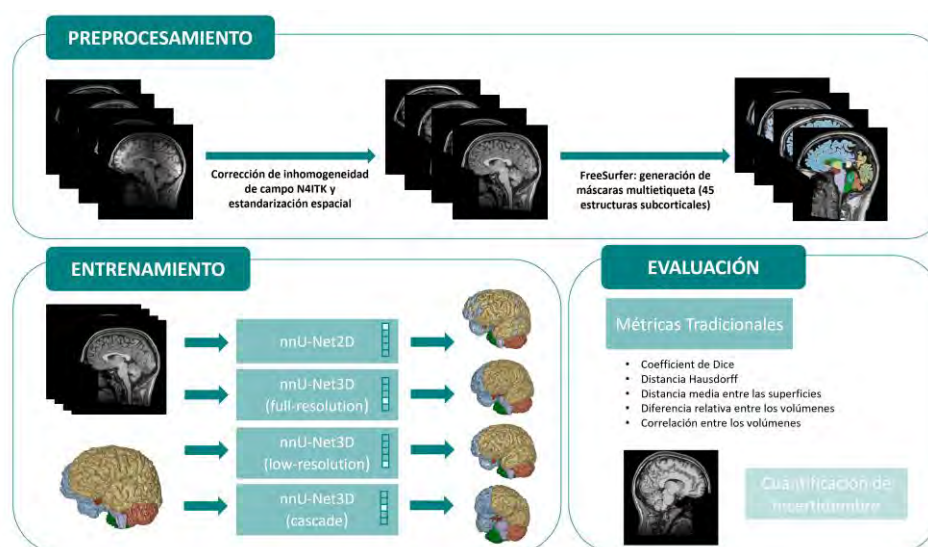


Figura 1. Marco metodológico seguido. Tras una etapa de preprocesamiento en la que se corrigen y estandarizan los volúmenes, se entrenan 4 modelos de DL basados en la nnUNet. Posteriormente, se evalúan los resultados basándose en métricas clásicas y se conduce una evaluación de incertidumbre.

	nnUNet 2D	nnUNet 3D full-resolution	nnUNet 3D low-resolution	nnUNet Cascade	FastSurfer
<i>DC</i>	0.816	0.873	0.842	0.834	0.833
<i>HD</i>	11.665	5.412	6.001	5.671	4.960
<i>ASD</i>	1.049	0.336	0.339	0.363	0.343
<i>RVD</i>	-0.007	0.015	-0.061	0.005	0.011
<i>VC</i>	0.934	0.983	0.958	0.940	0.921
<i>Varianza</i>	0.00045	0.00018	0.00017	0.00019	0.002
<i>Entropía</i>	34078.89	14575.49	13260.85	14163.99	16668.74
<i>Tiempo de predicción</i>	1:35min	3:46min	3:50min	3:44min	55min

Tabla 2. Resultados cuantitativos.

3.3. La evaluación

Se realiza una evaluación tradicional sobre el conjunto de test. En ella, se evalúan métricas específicas, incluyendo el **DC**, la **distancia Hausdorff** (DH), la **distancia media entre superficies** (*Average Surface distance*; ASD), la **diferencia relativa entre los volúmenes** (*relative volumen difference*; RVD) y la **correlación entre los volúmenes** (*volumen correlation*; VC).

Estas métricas aportan una evaluación de la calidad de las segmentaciones, lo que permitirá la elección cuantitativa del mejor modelo. Sin embargo, estas métricas no llevan a una comprensión completa de los mismos, ya que su rendimiento y fiabilidad no están completamente caracterizados. Debido a la importancia de estas características en la clínica y la necesidad de desarrollar modelos explicables, se somete a los modelos a un proceso de **cuantificación de incertidumbre**.

La cuantificación de incertidumbre en problemas de segmentación permite estimar la confianza de las predicciones arrojadas por un modelo, ofreciendo, a su vez, una herramienta para interpretar los resultados. En este sentido, se apuesta por una estrategia de **aumento de tiempo en la evaluación**, en el que se estima la incertidumbre de un modelo evaluando la salida de este frente a variantes de una misma imagen.

El conjunto de test se ve sometido, por tanto, a un proceso de aumento de datos aleatorio, donde se aplican diferentes transformaciones a los volúmenes. En concreto, se mide la robustez espacial mediante la rotación de las imágenes de entrada, generando dos volúmenes rotados en el eje X y dos rotados en el eje Y. En ambos casos, el ángulo de rotación (r), medido en radianes, se selecciona de manera aleatoria a partir de una distribución uniforme ($r \sim U(-0.3, 0.3)$). Además, se mide la robustez del modelo frente a cambios de intensidad empleando dos volúmenes distorsionados con ruido gaussiano (centrado en cero con desviación estándar 10).

Los seis volúmenes alterados se introducen en cada uno de los modelos propuestos. Sobre sus segmentaciones (transformadas de vuelta cuando es necesario), se calcula la **varianza y la entropía** como estimadores cuantitativos de incertidumbre. Además, se evalúa la incertidumbre de

forma cualitativa, mediante la generación de mapas de error entre la segmentación promedio y la segmentación original, así como con la imagen de *ground truth* (GT).

4. Resultados y Discusión

La **Tabla 2** recoge los resultados obtenidos, incluyendo el valor medio de las métricas descritas, así como los tiempos de predicción medios de cada uno de los modelos sobre una única imagen medidos en CPU. Estas métricas de rendimiento se comparan con **FastSurfer**, un algoritmo capaz de producir segmentaciones parecidas a FreeSurfer en un tiempo menor de predicción [13].

Los modelos propuestos ofrecen resultados similares a aquellos arrojados por FastSurfer, en especial aquellos entrenados empleando datos 3D. Es esencial, sin embargo, considerar las particularidades del modelo en 2D. En general, estos modelos presentan peor rendimiento que sus homólogos en 3D, especialmente notorio en las métricas que indican distancia. Al evaluar las predicciones arrojadas por estos modelos notamos que, en ciertos cortes, la red confunde sistemáticamente las estructuras que se encuentran en la derecha e izquierda. Esto se debe, probablemente, a la naturaleza intrínseca de las imágenes por RM: dado que la red solo ha visto información bidimensional, ésta no posee suficiente información geométrica para discernir entre el hemisferio izquierdo y derecho, arrojando, por tanto, segmentaciones de peor calidad, lo que resulta en peores métricas cuantitativas y modelos menos robustos.

Este problema se solventa mediante el uso de estructuras de DL en 3D, las cuales, a pesar de trabajar sobre parches de dimensión pequeña en lugar de sobre el volumen completo, poseen más información global, y, por tanto, son capaces de aprender características distintivas entre izquierda y derecha. Además, estos modelos prueban ser más robustos en el proceso de cuantificación de incertidumbre, ya que producen predicciones similares a la original cuando se modifica ligeramente la entrada (menor varianza y entropía).

Siendo la *nnUNet 3D full-resolution* la arquitectura con los mejores resultados cuantitativos, se decide medir su incertidumbre de manera cualitativa a través del cómputo de mapas de error (**Figura 2**).

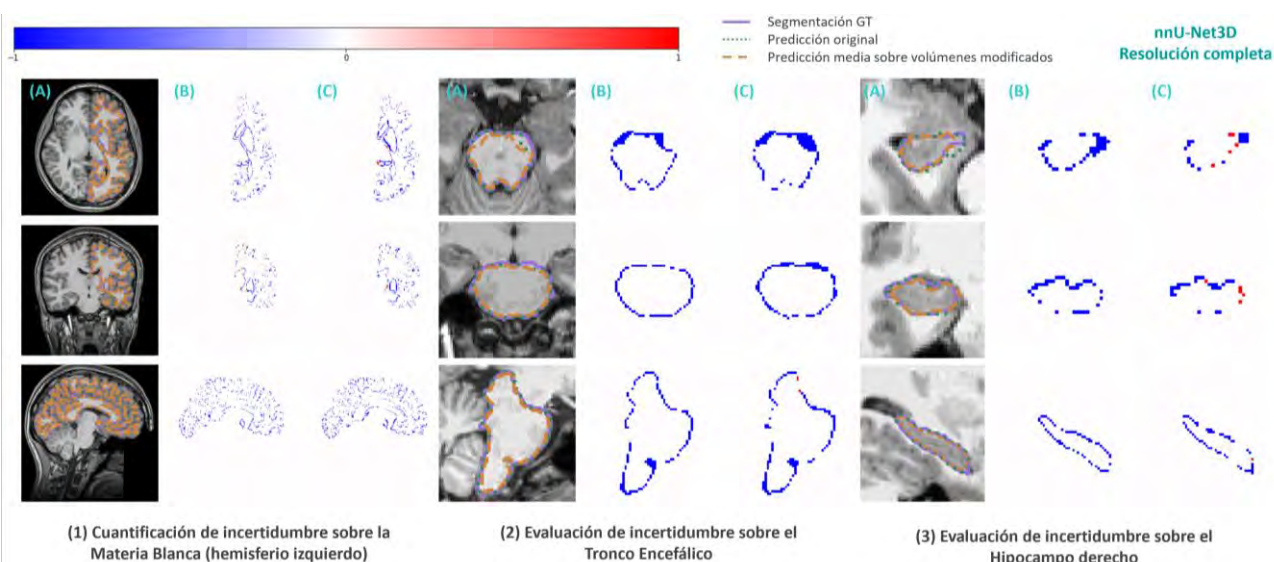


Figura 2. Cuantificación de incertidumbre sobre el modelo nnUNet 3D full resolution medido a partir de mapas de error sobre tres estructuras. La columna (A) muestra los contornos predichos sobre el volumen original (verde) y los volúmenes modificados (naranja) en comparación con el GT (morado). La columna (B) recoge la diferencia de la predicción media sobre los volúmenes modificados y la predicción original. La columna (C) muestra la diferencia entre la predicción media sobre los volúmenes modificados y el GT.

Tal y como se observa en la **Figura 2**, este modelo es capaz de producir segmentaciones relativamente similares a la original sobre estructuras de diferente tamaño cuando se modifica ligeramente la entrada, imitando incluso los errores cometidos en la predicción original frente a *ground truth*. Estos resultados, junto con los obtenidos en la evaluación cuantitativa avalan a la nnUNet 3D trabajando en resolución completa como un modelo robusto para la incorporación en la clínica para la segmentación de OARs sobre RM. Sin embargo, es imperativo continuar con la investigación, especialmente enfocada en la segmentación de OARs sobre TC, lo cual ampliará significativamente la aplicabilidad clínica, dado que este tipo de información es con la que frecuentemente se opera en el flujo de trabajo diario en entornos clínicos. Por lo tanto, adaptar y validar métodos en dicha modalidad es esencial para una integración efectiva en la rutina clínica.

5. Conclusiones

La segmentación de OARs es un paso esencial en la planificación de protonterapia. Debido al éxito del marco de la nnUNet, se decide probar la efectividad de ésta para esta tarea. La nnUNet demuestra ser una herramienta eficaz para la segmentación de 45 estructuras cerebrales tras someterla a una evaluación tradicional, en especial aquellas estructuras entrenadas con parches 3D a resolución completa. Además, se muestra su estabilidad y robustez mediante técnicas de cuantificación de incertidumbre, lo que permite proponer de forma certera el uso de este método para su incorporación en la clínica.

Referencias

- [1] Bortfeld, T., Paganetti, H., & Kooy, H. (2005). MO-A-T-6B-01: proton beam radiotherapy—the state of the art. *Medical Physics*, 32(6Part13), 2048-2049.
- [2] Paganetti, H., Blakely, E., Carabe-Fernandez, A., Carlson, D. J., Das, I. J., Dong, L., ... & Willers, H. (2019). Report of the AAPM TG-256 on the relative biological effectiveness of proton beams in radiation therapy. *Medical physics*, 46(3), e53-e78.
- [3] Mohan, R., & Grosshans, D. (2017). Proton therapy—present and future. *Advanced drug delivery reviews*, 109, 26-44.
- [4] Oktay, O., Nanavati, J., Schwaighofer, A., Carter, D., Bristow, M., Tanno, R., ... & Nori, A. (2020). Evaluation of deep learning to augment image-guided radiotherapy for head and neck and prostate cancers. *JAMA network open*, 3(11), e2027426-e2027426.
- [5] Lucido, J. J., DeWees, T. A., Leavitt, T. R., Anand, A., Beltran, C. J., Brooke, M. D., ... & Patel, S. H. (2023). Validation of clinical acceptability of deep-learning-based automated segmentation of organs-at-risk for head-and-neck radiotherapy treatment planning. *Frontiers in Oncology*, 13, 1137803.
- [6] Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 18(2), 203-211.
- [7] Sava-Segal, C., Richards, C., Leung, M., & Finn, E. S. (2023). Individual differences in neural event segmentation of continuous experiences. *Cerebral Cortex*, bhad106.
- [8] Bhattasali, S., Brennan, J., Luh, W. M., Franzluebbers, B., & Hale, J. (2020). The Alice Datasets: fMRI & EEG observations of natural language comprehension. In *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 120-125).
- [9] Castrillon, G., Epp, S., Bose, A., Fraticelli, L., Hechler, A., Belenya, R., ... & Riedl, V. (2023). An energy costly architecture of neuromodulators for human brain evolution and cognition. *bioRxiv*, 2023-04.
- [10] Kikinis, R., Pieper, S. D., & Vosburgh, K. G. (2013). 3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support. In *Intraoperative imaging and image-guided therapy* (pp. 277-289). New York, NY: Springer New York.
- [11] Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: improved N3 bias correction. *IEEE transactions on medical imaging*, 29(6), 1310-1320.
- [12] Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., ... & Dale, A. M. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3), 341-355.
- [13] Henschel, L., Conjeti, S., Estrada, S., Diers, K., Fischl, B., & Reuter, M. (2020). FastSurfer - A fast and accurate deep learning-based neuroimaging pipeline. *NeuroImage*, 219, 117012.

Enhancing MRI Reconstruction Efficiency Through Multi-GPU Parallelization

E. López-Ales¹, R.M. Menchón-Lara¹, M. Martín-Fernandez¹, F. Simmross-Wattenberg¹, C. Alberola-López¹

¹ Laboratorio de Procesado de Imagen, Universidad de Valladolid, Valladolid, Spain. elopezales@gmail.com, rmenchon@lpi.tel.uva.es, marcma@uva.es, fedsim@uva.es, caralb@tel.uva.es.

Abstract

Dynamic cardiac MRI (cMRI) is essential for diagnosing cardiovascular diseases, demanding high resolution and image quality. However, achieving superior quality increases data volume and reconstruction time. To tackle this, we propose a solution using parallel imaging and Compressed Sensing (CS) with high-capacity computing devices (e.g., GPUs) for accelerated reconstruction of undersampled data. GPU memory limitations, especially in 3D cMRI, present challenges. Our scalable approach splits the reconstruction problem and employs multiple GPUs (or multiple multi-core CPUs) to perform multiple optimizations in parallel using the well-known NESTA algorithm, while preserving smoothness between adjacent frames in the temporal dimension. Preliminary results on 5D cMRI reconstruction show that our parallel proposal achieves equivalent reconstruction quality in less time, enabling larger data processing and cost reduction with smaller, more affordable GPUs, as opposed to a single, highly expensive GPU. Moreover, the adoption of the OpenCLIPER framework eliminates proprietary GPU technologies. Exploration of adaptability to other sampling schemes opens new possibilities in this field.

1. Introduction

Dynamic cardiac magnetic resonance imaging (cMRI), also known as cardiac CINE, assumes a pivotal role in the evidence-based diagnosis and treatment of cardiovascular diseases, affording clinicians the ability to meticulously assess cardiac function and discern pathological alterations with a remarkable degree of precision [1]. Ensuring the attainment of dependable diagnostic outcomes via cMRI hinges upon the meticulous fulfillment of stringent spatio-temporal resolution and image quality requirements, which, in turn, require obtaining complex representations of cardiac dynamics.

In the pursuit of higher image quality, the substantial increase in acquired data volume represents a significant challenge. Due to the increased time and computing resources required, the practice of using undersampled data in reconstruction techniques has become commonplace. By taking advantage of these data, it is possible to mitigate resource demands while maintaining quality requirements. Although these techniques, while indisputably efficacious, have the potential to introduce image artifacts, which is a significant impediment to diagnostic procedures. Consequently, it is of paramount importance to address these challenges with the primary goal of streamlining clinical workflows and optimizing the allocation of medical imaging resources. Consequently, a compelling need arises to develop advance iterative algorithms, designed to enhance data retrieval and regularization, thereby elucidating the complexities inherent in undersampled cMRI reconstruction.

Our approach aims to exploit the potential of a scalable number of GPUs or multi-core CPUs. Our solution, paired with a customized version of NESTA algorithm that use parallel local optimizations and the seamless GPU management facilitated by OpenCLIPER, enables parallel 3D reconstruction even with high-dimensional data volumes. By judiciously distributing the computational workload across multiple GPUs, our goal is to enhance the speed and efficiency of cMRI reconstruction while effectively addressing the challenges posed by GPU memory limitations. This approach aims to significantly streamline cMRI workflows, thereby promoting cardiovascular research.

The remainder of this paper is structured as follows: Section 2 briefly surveys the state of the art, Section 3 discusses the data employed, Section 4 explains the methods used, along with details concerning the multi-GPU process and the operators used, and Section 5 presents the results and a brief discussion thereof. Finally, Section 6 includes the main conclusions.

2. Background

Presently, the landscape of MRI reconstruction with GPU utilization is extensively surveyed by Wang et al. [2], revealing a noticeable upsurge in publications at the intersection of GPUs, MRI, and reconstruction following the introduction of NVIDIA's CUDA technology in 2007. However, once the fundamental algorithms for conventional MRI reconstruction have been comprehensively explored and implemented, the rate of publication in this domain appears to plateau.

Within this realm of publications, the domain of CS stands out prominently. The popularity of CS has spurred numerous investigations harnessing GPU acceleration for MRI reconstruction. For instance, Nam et al. have proposed a GPU-accelerated parallel implementation of an iterative CS reconstruction algorithm tailored for radial data acquisitions [3]. On a similar note, Sabbagh et al. have conducted a comprehensive comparative analysis of various parallel platforms, focusing on accelerating nonlinear CS reconstruction in cMRI using iterative optimization algorithms [4]. Meanwhile, Xue et al. have introduced Gadgetron as a framework for leveraging distributed computing in reconstruction, demonstrating the feasibility of achieving nonlinear image reconstruction with clinically acceptable latency [5]. Nonetheless, it is essential to highlight a prevailing practice in the field where parallel implementations tend to focus predominantly on optimizing the speedup achievable through GPU utilization in the reconstruction process, often relying on NVIDIA's CUDA technology. However,

it is noteworthy that there has been a limited exploration of leveraging multiple GPUs to accelerate the resolution of optimization problems. This underutilization of multiple GPUs for parallelization remains a notable gap in current research and implementation practices.

There appears to be a broad consensus within the MRI research community to acknowledge and harness the capabilities of GPUs [6]. In summary, the incorporation of GPUs into MRI reconstruction workflows has yielded substantial acceleration benefits, especially within the realm of CS. Nonetheless, it is imperative to exercise prudence regarding the inherent limitation of GPU memory capacity. This limitation often constitutes a bottleneck, constraining the volume of data that can be efficiently reconstructed. This constraint becomes even more pronounced in the context of 5D cMRI reconstruction (3 spatial dimensions + 2 temporal dimensions, i.e., cardiac phase and respiratory phase), where extensive datasets are commonly encountered.

In this context, we bring to the forefront the NESTA algorithm [7], a pivotal tool in the realm of signal an image reconstruction which operates iteratively, particularly tailored to combat the intricacies posed by undersampled data. This highlights NESTA’s role in achieving precise reconstructions of large-scale signals or images characterized by extensive dynamic ranges.

3. Materials

In this work, we use synthetic data generated with the numerical phantom XCAT [8], that provides detailed internal anatomy and realistic cardio-respiratory deformation models. A bSSFP acquisition was simulated in free-breathing (FB) with the following relevant parameters: TR/TE=3.0/1.5ms, flip angle=60°, field of view (FOV) of (168 mm)³ with spatial resolution of 2 mm³. Multi-coil data were simulated with a total of 16 coils. The fully sampled data were retrospectively undersampled using 3D variable density spiral-like Cartesian sampling scheme (VD-CASPR) [9] to avoid gridding operations during reconstructions. Finally, data were binned in 20 cardiac phases and 4 respiratory phases, corresponding to 80 reconstructed frames with an acceleration factor (AF) of 4.

The performance evaluation of our research was conducted on a dedicated hardware system consisting of 4 Nvidia RTX A5000 GPUs, each equipped with 24GB of VRAM. The system features 2 AMD EPYC 7513 CPUs (32 cores/64 threads each) @ 2,6 GHz. Additionally, the system is equipped with 1TB of RAM.

4. Methods

In the context of the implementation of algorithms mentioned in Section 2, especially on GPUs, the choice of a programming platform is crucial. Most literature implementations prefer CUDA due to its well-established ecosystem and performance optimizations, but it’s limited to NVIDIA hardware. OpenCL, an open standard, offers universality but may sacrifice, at times, performance. Hence, there’s notable need for an integrated solution to reduce reliance on proprietary technologies.

Within this framework, the OpenCLIPER tool [10] emerges as an exceptional development resource. It offers precise control over GPU utilization intricacies. Leveraging OpenCLIPER enhances algorithm execution speed by enabling load distribution across multiples devices. Incorporating the OpenCLIPER framework into our proposed solution offers several advantages. Additionally, the framework simplifies transparent GPU resource management, mitigating the potential bottlenecks and optimizing overall system performance.

To address the gap in the underutilization of multiple GPUs for parallelization and the imperative need for significant reductions in computational time, this study introduces the parallelization of optimization techniques within the context of cMRI reconstruction. Instead of relying on a global optimization strategy, we advocate the utilization of localized optimizations, while preserving smoothness between adjacent frames in the temporal dimension. This approach yields substantial reductions in processing time, bringing the outcomes remarkably close to those achieved through global optimization.

In pursuit of this objective, we have adapted the NESTA algorithm and integrated it with the temporal Total Variation (tTV) transform, serving as the primary separation technique. NESTA, being an iterative algorithm, can be computationally intensive, particularly when confronted with extensive datasets, which also pose a problem in terms of memory space. This computational challenge is further compounded when dealing with 3D free-breathing acquisitions (5D cMRI), because of the extra temporal dimension, i.e., the respiratory phase. In response, our approach shifts the focus towards resolving this computational bottleneck through parallel optimization, as opposed to a global approach.

RECONSTRUCTION WORKFLOW USING OpenCLIPER

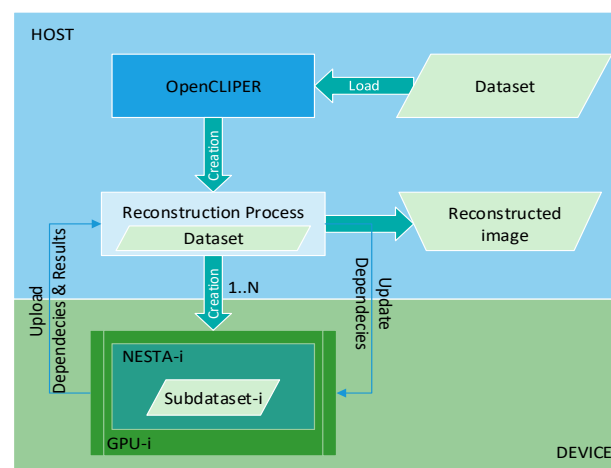


Figure 1: Reconstruction workflow using OpenCLIPER. Parallelization is performed using multiple GPUs: each one runs a NESTA process.

To overcome this computational challenge, we have leveraged the potential of the OpenCLIPER framework [10]. Figure 1 illustrates how this framework functions as the backbone of the process, facilitating an efficient inter-

face between the algorithm and the underlying hardware. OpenCLIPER is renowned for its support for a wide range of computing devices, especially GPUs, which are recognized for their ability to efficiently handle parallel processing tasks.

A pivotal facet of this approach pertains to the methodology adopted for data processing. Rather than processing data in large blocks, constrained by the memory limitations of the utilized device, we meticulously partition the data into many subsets as there are available devices on the execution platform, typically GPUs. Given that the cardiac dimension encompasses a greater number of phases in comparison to the respiratory dimension, data partitioning is conducted along the former dimension. In this manner, the requisite dependencies for each GPU are significantly reduced as compared to partitioning along the respiratory dimension. Moreover, dataset partitioning considers the memory capacity of each GPU ensuring that each GPU receives, in accordance with its capabilities, a subset of frames from a portion of the cardiac dimension alongside the entire associated respiratory dimension. Although this strategy is poised to unlock the full potential of parallel processing while circumvent memory limitations, it is not devoid of challenges. A highlighted challenge involves the emergence of temporal dependencies in both the cardiac and respiratory dimensions.

Temporal dependencies manifest due to two primary factors: the cyclical nature of the data in its temporal dimensions (cardiac and respiratory phases) and the inclusion of tTV within the regularization term. These dependencies pose challenges, disrupting the seamless processing of data and introducing inconsistencies in results, such as discernible discontinuities between reconstructed frame sequences.

To address the issue dependencies, we've established a dynamic runtime communication channel among processing devices. This ensures that each device gains access to its temporal dependencies across both dynamic dimensions. However, owing to the iterative nature of the algorithm's convergence, synchronization is indispensable, given that information related to temporal dependencies evolves with each iteration. Achieving this periodic synchronization among devices during runtime necessitates the application of mutexes and condition variables. These elements grant access to essential resources and ensures the correct sequence of access to dependencies. Moreover, this concurrent programming approach is meticulously crafted to minimize idle periods and empower devices to work and transmit their outcomes without unwarranted delays. We have strategically incorporated synchronization mechanisms at the conclusion of each iterative cycle and precisely at the point when they are required to access dependencies. This deliberate implementation ensures that our computational devices meticulously arrange their local results as each iteration culminates, ensuring they stand prepared to promptly address their dependencies when the need arises. Consequently, after each individual iteration of the algorithm, dependencies

become readily available to the involved devices.

Once the dynamic communication channel between devices and synchronization protocols are firmly in place, it becomes paramount to comprehend how temporal variation operators adapt to account for temporal dependencies in each dimension. To elucidate this adaption, it's worth noting that the equation governing the 5D reconstruction problem aligns with Eq. (1):

$$\hat{\mathbf{m}} = \underset{\mathbf{m}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{E}\mathbf{m} - \mathbf{y}\|_{\ell_2}^2 + \lambda \|\Phi\mathbf{m}\|_{\ell_1} \quad (1)$$

where \mathbf{y} denotes the acquired k-space data, $\hat{\mathbf{m}}$ is the image to be reconstructed, \mathbf{E} is the encoding operator (including coil sensitivity map, Fourier Transform and sampling matrix), λ the regularization parameter, and Φ stands for a sparsifying transform (in our case, the aforementioned tTV). This optimization problem is solved iteratively using the NESTA algorithm. At each iteration, the gradient of the function to be minimised is evaluated according to the following equation:

$$\nabla f(\mathbf{m}_k) = \mathbf{E}^H(\mathbf{E}\mathbf{m}_k - \mathbf{y}) + \lambda \cdot \underbrace{\Phi^H f'_{\mu}(\Phi\mathbf{m}_k)}_{\Lambda} \quad (2)$$

where k -index denotes the iteration number, f_{μ} stands for the Huber function used as an approximation of the ℓ_1 -norm in Eq.(1), and $\{\mathbf{E}^H, \Phi^H\}$ denotes the corresponding adjoint operators. Specifically, our work is on the parallelization of the second term of the gradient calculation (Λ).

Our proposed approach for parallelizing this problem entails the application of sparsifying transform to data subsets within each device. Consequently, as the manner of applying the operator to the matrix hinges on the operational direction, each result is derived by manipulating the data in concert with its dependencies. This is represented in Eq. (3):

$$\Lambda_{i,j} = f_{\mu}(m_{i,j} - m_{i-1,j}) - f_{\mu}(m_{i+1,j} - m_{i,j}) + f_{\mu}(m_{i,j} - m_{i,j-1}) - f_{\mu}(m_{i,j+1} - m_{i,j}) \quad (3)$$

were i and j are the indices related to cardiac and respiratory phases. This seamlessly integrated mechanism facilitates parallel processing of data subsets without encountering any impediments.

5. Results and discussion

In the evaluation of reconstruction quality, we have employed the Structural Similarity Index (SSIM) as a reference metric. Figure 2 illustrates that no significant differences are discernible between our parallel version with local optimizations (with 2, 3 and 4 GPUs) and the sequential global optimization (CPU). This finding supports the notion that our approach represents a substantial advancement in terms of efficiency.

The implementation of the parallel version of the NESTA algorithm has enabled us to attain reconstructions of a quality comparable to those achieved by the sequential algorithm. Furthermore, it is pertinent to underscore the efficiency of our approach, as we have succeeded in significantly curtailing the time required for reconstruction.

One of the principal challenges in this context pertains to the ability to process substantial volumes of data without succumbing to hardware limitations, particularly in terms of GPU memory capacity, which may pose an insurmountable obstacle. Our methodology surmounts this hurdle by leveraging the computational workload across multiple GPUs to improve computational performance, with results virtually indistinguishable from those generated using the sequential reconstruction process.

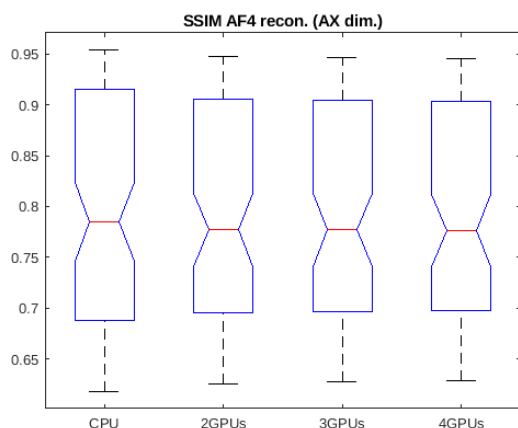


Figure 2: Distributions of the SSIM.

Conversely, the results presented in Table 1 provide compelling evidence that execution times required for reconstruction in a multi-GPU environment are considerably shorter than those necessitated for performing the same task on a CPU without parallelization (since a single GPU is insufficient due to memory constraints). These findings, coupled with the quality measures, strongly support the claim that the use of multiple GPUs in optimization problems results in at least the expected level of efficiency enhancement.

5D cMRI (AF4)	
CPU	1879 s
Multi-GPU (2)	32,9 s
Multi-GPU (3)	17,4 s
Multi-GPU (4)	15,4 s

Table 1: Reconstruction times of the 5D cMRI dataset (3D spatial with cardiac and respiratory dimensions) for 80 reconstructed frames.

Despite the promising results obtained, our research now enters a phase of rigorous validation using clinical data. These validations will not only bolster the credibility of our approach but also highlight areas amenable to enhancement. Furthermore, we actively exploring the adaptability of our parallel algorithm to different sampling schemes, thereby investigating its versatility.

6. Conclusions

The findings demonstrate that the parallelized local optimization version with the NESTA algorithm achieves reconstructions of equivalent quality with the iterative approach employing global optimization, which in turn substantially reduces reconstruction time. As expected, this research corroborates the increase in parallelization efficiency provided using of multiple GPUs in optimization tasks. While the validation of these findings with clinical data remains outstanding, and the exploration of adaptability to various sampling schemes is ongoing, the preliminary results augur promising prospects for the development of efficient, high-quality solutions in the realm of cMRI reconstruction research.

Acknowledgements

This work is partially supported by MINECO under grants TEC2017-82408-R, PRE2018-086922, and by the Agencia Estatal de Investigación under grants PID2020-115339RB-I00 and TED2021-130090B-I00.

References

- [1] R. Turner et al. Challenges of imaging structure and function with MRI. *IEEE Engineering in Medicine and Biology Magazine*, 19(5):42–54, 2000.
- [2] H. Wang et al. A survey of GPU-based acceleration techniques in MRI reconstructions. *Quantitative imaging in medicine and surgery*, 8(2):196, 2018.
- [3] S. Nam et al. Compressed sensing reconstruction for whole-heart imaging with 3D radial trajectories: a graphics processing unit implementation. *Magnetic resonance in medicine*, 69(1):91–102, 2013.
- [4] M. Sabbagh et al. Cardiac MRI compressed sensing image reconstruction with a graphics processing unit. In *2016 10th International Symposium on Medical Information and Communication Technology (ISMICT)*, pages 1–5. IEEE, 2016.
- [5] H. Xue et al. Distributed MRI reconstruction using gadgetron-based cloud computing. *Magnetic resonance in medicine*, 73(3):1015–1025, 2015.
- [6] S. S. Stone et al. How GPUs can improve the quality of magnetic resonance imaging. In *The First Workshop on General Purpose Processing on Graphics Processing Units*, pages 39–55, 2007.
- [7] S. Becker et al. NESTA: A fast and accurate first-order method for sparse recovery. *SIAM Journal on Imaging Sciences*, 4(1):1–39, jan 2011.
- [8] W. P. Segars et al. 4D XCAT phantom for multimodality imaging research. *Medical Physics*, 37(9):4902–4915, 2010.
- [9] C. Prieto et al. Highly efficient respiratory motion compensated free-breathing coronary MRA using golden-step cartesian acquisition. *Journal of Magnetic Resonance Imaging*, 41(3):738–746, 2015.
- [10] F. Simmross-Wattenberg et al. OpenCLIPER: an OpenCL-based C++ framework for overhead-reduced medical image processing and reconstruction on heterogeneous devices. *IEEE journal of biomedical and health informatics*, 23(4):1702–1709, 2018.

Modelo de predicción de la respuesta al tratamiento de quimio-radioterapia en pacientes con cáncer de pulmón de células no pequeñas localmente avanzado irresecable mediante la aplicación de radiómica en imágenes de TC

G. Rozalen¹, A.P. Seiffert¹, E.J. Gómez^{1,2}, J.J. Martín-Pinacho³, P. Sánchez-González^{1,2}

¹ Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España;
g.rozalen@alumnos.upm.es, {ap.seiffert, enriquejavier.gomez, p.sanchez}@upm.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, España

³ Servicio de Radiodiagnóstico, Hospital Universitario Ramón y Cajal, Madrid, España;
jesusjavier.martin@salud.madrid.org

Resumen

El cáncer de pulmón, con 2,2 millones de nuevos casos y 1,8 millones de muertes en 2020, es la segunda neoplasia más común y la principal causa de muerte por cáncer a nivel global. Se clasifica en cáncer de pulmón de células no pequeñas (CPCNP) (85% de los casos) y cáncer de pulmón de células pequeñas (CPCP) (15% de los casos). Actualmente, el método diagnóstico más empleado para evaluar el tumor y los efectos secundarios del tratamiento son las imágenes de tomografía computarizada (TC). En los últimos años, la radiómica ha emergido como una herramienta valiosa para el análisis cuantitativo de imágenes médicas, mejorando el diagnóstico y personalizando el tratamiento. Este estudio tiene como objetivo desarrollar un modelo predictivo de respuesta al tratamiento radical de quimio-radioterapia en pacientes con CPCNP basado en el análisis cuantitativo de imágenes de TC. Se exploran dos enfoques de selección de características (correlación más análisis estadístico univariante y LASSO) y se crean siete modelos predictivos, obteniendo mejores resultados con LASSO, alcanzando AUCs de 0,88 (Regresión Logística) y 0,90 (Perceptrón Multicapa). Estos resultados demuestran el potencial de la radiómica como herramienta para la predicción de la respuesta al tratamiento.

1. Introducción

El cáncer de pulmón es una enfermedad caracterizada por el crecimiento anormal, maligno y descontrolado de las células en el tejido pulmonar [1]. En el año 2020, este cáncer tuvo una estimación de 2,2 millones de nuevos casos y causó 1,8 millones de muertes, posicionándose como la segunda neoplasia más diagnosticada y la principal causa de fallecimiento relacionada con el cáncer a nivel global [2]. Es innegable que el tabaquismo sigue siendo el factor de riesgo más importante y modificable en relación con el cáncer de pulmón, siendo atribuido a aproximadamente el 80-90% de los casos [3]. Este tipo de cáncer se divide en dos categorías histológicas principales: el cáncer de pulmón de células pequeñas (CPCP) que representa el 15% de los diagnósticos totales; y el cáncer de pulmón de células no pequeñas (CPCNP) que abarca el 85% restante de los casos [4]. A su vez, el CPCNP se subdivide en tres subtipos histológicos: adenocarcinoma (ADC) (40%), carcinoma de

células escamosas (CCE) (25-30%) y carcinoma de células grandes (10-15%) [5].

El diagnóstico del cáncer de pulmón es un proceso crucial y complejo que involucra la recopilación de información de diversas fuentes, como la historia clínica, síntomas, exámenes físicos y pruebas de imagen, entre ellas la tomografía computarizada (TC) [6]. Una vez localizado y clasificado, el cáncer de pulmón se estudia utilizando el sistema TNM en cuatro estadios: inicial (I-II), localmente avanzado (III) y metastásico (IV) [7]. El tratamiento se basa en el tipo de cáncer, tamaño, ubicación, estadio, metástasis y la salud general del paciente, y puede incluir cirugía, radioterapia, quimioterapia, inmunoterapia y terapia molecular dirigida, a menudo combinados de diversas terapias [1].

Actualmente en la práctica clínica no existen herramientas para identificar los pacientes respondedores a la quimioterapia, por lo que todos han de recibir un tratamiento que no está exento de toxicidad (30% de efectos adversos graves) y el seguimiento es uniforme, independientemente de la probabilidad de respuesta. Poder predecir la respuesta tumoral personalizada puede cambiar la pauta de tratamiento o el tipo de tratamiento, con una mejor calidad de vida en consecuencia para el paciente oncológico. Por ello, el objetivo principal de este trabajo es desarrollar un modelo predictivo de respuesta al tratamiento de pacientes con CPCNP utilizando análisis cuantitativo de imágenes TC. Para ello, se evalúan diversas técnicas de *machine learning* incluyendo métodos de selección de características y basándose en variables clínicas y características radiómicas extraídas de imágenes TC sobre el tumor segmentado.

2. Materiales y Métodos

2.1. Materiales

Para realizar este estudio, la cohorte inicial disponible incluye todos los pacientes del Servicio de Oncología Radio-terápica del Hospital Universitario Ramón y Cajal que entre los años 2015 y 2022 han sido diagnosticados con cáncer de pulmón, sin especificar si se trata de células pequeñas o células grandes, y han recibido al menos tratamiento con radioterapia. Los criterios de inclusión establecidos se

fijan en: (1) Pacientes mayores de 18 años, (2) con diagnóstico histológico de CPNCP irreseccable de estadios IIIA, IIIB o IIIC, (3) tratamiento completo con quimioterapia y radioterapia con intención radical, (4) al menos dos ciclos de quimioterapia basada en platino, (5) dosis de radioterapia mínima de 60 Gy e (6) independientemente de si completan el tratamiento con Durvalumab durante un año. Y los criterios de exclusión son: (1) TC de tórax con grosor de corte de más de 3 mm, (2) sin contraste intravenoso, (3) realizado más de 8 semanas después de haber finalizado tratamiento concurrente, (4) sin imágenes en el repositorio PACS y (5) pacientes que no completan el tratamiento radical por cualquier motivo.

Para la obtención de imágenes, se siguen protocolos clínicos estándar en el Hospital Universitario Ramón y Cajal, Madrid. Se utiliza un contraste yodado a razón de 1 ml por kilogramo de peso corporal, con velocidad de infusión de 5 ml por segundo y retraso de 40 segundos. Se obtienen imágenes de TC de tres escáneres: dos de 64 cortes (*Philips Brilliance* y *Siemens Sensation*) y uno de 128 cortes (*Philips Incisive*). La segmentación del tumor se realiza con *Synapse 3D* (v. 7.2) (<https://synapse-emea.fujifilm.com/synapse-3d.html>). Además, se emplea *Python* (v. 3.11.0) (<https://es.python.org/>) para implementar el código necesario para el análisis estadístico y los modelos de clasificación. Se emplean bibliotecas como *Pandas* (<https://pandas.pydata.org/>) y *Numpy* (<https://numpy.org/>) para la gestión de las bases de datos y operaciones, *Pyradiomics* (<https://pyradiomics.readthedocs.io/en/latest/>) para la extracción de características cuantitativas, y *scikit-learn* (<https://scikit-learn.org/stable/>) para los modelos de aprendizaje automático, entre otros.

2.2. Métodos

La metodología de este estudio se divide en cuatro etapas: (1) segmentación del tumor, (2) extracción de características, (3) selección de características y (4) construcción de los modelos de clasificación.

Segmentación del tumor. Para limitar la región de interés (ROI), en este caso el tumor pulmonar primario, se localiza el tumor tomando como referencia las descripciones del estudio de la imagen. A continuación, se establece el diámetro máximo del tumor, creando una ROI en 3D que engloba la masa. Este resultado se puede editar o bien manual en 2D o 3D o semiautomáticamente basada en contornos guiados por el usuario. Todas las segmentaciones se validan por un radiólogo torácico experto.

Extracción de características. Previo a la extracción de características, se utiliza el interpolador *sitkBSpline*, un espaciado de 1 píxel en las tres dimensiones de la imagen y un nivel de intensidades *binWidth* igual a 25, tras el remuestreo siguiendo las indicaciones de la *Image Biomarker Standardisation Initiative* (IBSI) [9]. Se aplican filtros no espaciales como *Wavelet*, *Laplacian of Gaussian* (LoG) con 5 escalas de filtrado, *Gradient* y *Local Binary Pattern in 3D* para resaltar propiedades específicas de las imágenes. Se extraen 1652 características radiómicas de cada lesión primaria, que se dividen en características de primer orden (morfológicas e intensidad) y características textura-

les de segundo orden u orden superior (basadas en las matrices *gray-level co-occurrence matrix* (GLCM), *gray-level run length matrix* (GLRLM), *gray-level size zone matrix* (GLSZM), *gray-level dependence matrix* (GLDM) y *neighborhood gray-level difference matrix* (NGLDM)). Además, se utilizan variables clínicas recopiladas retrospectivamente. Estas variables incluyen: Edad, Sexo, Tabaquismo, Índice de paquetes de cigarrillos anuales (IPA), *Eastern Cooperative Oncology Group* (ECOG), Histología, Estadio, Tipo de Tratamiento y Número de ciclos de quimioterapia.

Selección de características. Para seleccionar las características más relevantes se emplean dos enfoques: (1) Estudio de la correlación de las características radiómicas seguido de un análisis estadístico univariante y (2) *Least Absolute Shrinkage and Selection Operator* (LASSO). Antes de cada selección, todas las características son estandarizadas con la función de *sklearn StandardScaler*. Respecto a la selección (1), tras la correlación, las variables binarias o categóricas se evalúan con la prueba chi-cuadrado. En el caso de las variables continuas se estudia la normalidad mediante la prueba *Shapiro-Wilk*, y dependiendo del resultado se emplea la prueba T o prueba U de *Mann Whitney*. Para todas las pruebas se considera un *p*-valor < 0,05 estadísticamente significativo. Respecto a la selección (2), LASSO es una técnica de selección y regularización en aprendizaje automático. Utiliza el parámetro α para controlar la intensidad de la penalización, lo que reduce algunos coeficientes hacia cero (regularización L1), contribuyendo a la retención de características relevantes, con $\alpha=0.1$ en este análisis.

Construcción de los modelos de clasificación. Las características seleccionadas se emplean para crear los modelos de clasificación utilizando los métodos *Support Vector Machine* (SVM), Vecinos más próximos (KNN, *K-nearest neighbors*), *Gaussian Naive Bayes* (GNB), *Random Forest* (RF), *Gradient Boosting* (GB), *Logistic Regression* (LR) y *Multi-Layer Perceptron Classifier* (MLP). Para ello, los datos de entrada se distribuyen mediante validación simple con una ponderación de 70-30 en entrenamiento y prueba. Con los datos de entrenamiento se realiza una validación cruzada para la optimización de hiperparámetros con la función de *sklearn GridSearchCV* con 5 pliegues. Con los datos de prueba y los modelos con los mejores parámetros, se evalúa el rendimiento mediante parámetros como la exactitud (Acc), la sensibilidad (S), la precisión (P), la especificidad (E) y el área bajo la curva (AUC). En la *Figura 1* se resume la secuencia descrita.

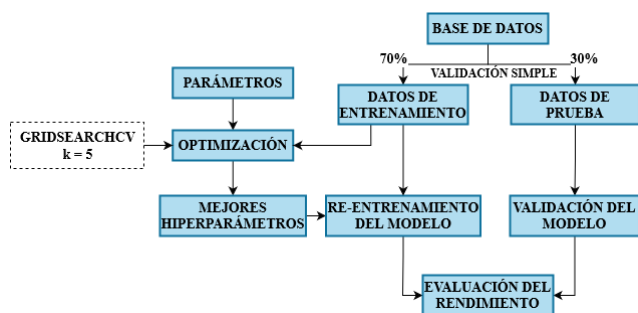


Figura 1. Flujo de trabajo seguido para cada modelo predictivo.

3. Resultados y Discusión

3.1. Pacientes

A partir de los criterios de inclusión y exclusión especificados en el apartado 2.1 de Materiales, de 870 pacientes iniciales se seleccionan 75. Las características demográficas de los pacientes a estudio se muestran en la *Tabla 1*. Para llevar a cabo el estudio se divide en pacientes respondedores o no respondedores según los criterios RECIST 1.1 [10].

Características	
Edad	
Medias ± Desviación estándar	77 ± 8,90
Sexo	
Mujer	20 (26,67%)
Varón	55 (73,33%)
Fumador	
No fumador	2 (2,67%)
Activo	28 (37,33%)
Exfumador	45 (60,00%)
Estadio TNM	
Estadio IIIA	35 (46,67%)
Estadio IIIB	38 (50,67%)
Estadio IIIC	2 (2,67%)
Respuesta al tratamiento	
Sí	47 (62,67%)
No	28 (37,33%)
Histología	
ADC	30 (40,00%)
CCE	39 (52,00%)
Otros	6 (8,00%)

Tabla 1. Características demográficas de los pacientes.

3.2. Segmentación

La segmentación del tumor con *Synapse 3D* se ilustra en la *Figura 2*. En la captura A se muestra la interfaz principal de la aplicación y en la B cómo se traza la línea que define los límites del nódulo y genera automáticamente la ROI. En la captura C se observa cómo se puede editar la segmentación de forma semiautomática utilizando contornos guiados por el usuario, y en la D se presenta la ROI en 3D al finalizar el proceso.



Figura 2. Capturas del programa *Synapse 3D*

3.3. Selección de características

Se parte de 1652 características radiómicas junto con 9 variables clínicas para la respuesta al tratamiento. Con la selección (1), se toman 7 características radiómicas y 1 clínicas (“Número de ciclos de quimioterapia”). Con la selección (2) se obtienen 5 características radiómicas y 2 clínicas (“ECOG”, “Número de ciclos de quimioterapia”). En ambas selecciones se observa que la única característica común es el “Número de ciclos de quimioterapia”. La relación que se extrae es que, a más ciclos, hay menos respuesta de los pacientes, posiblemente debido a tumores resistentes o al impacto en el organismo. A menos ciclos, hay más respuesta debido a la sensibilidad del tumor y la eficacia del tratamiento.

3.4. Modelos de aprendizaje automático

A continuación, en la *Tabla 2* y *Tabla 3* se muestra el rendimiento de los resultados obtenidos al evaluar los siete modelos de *machine learning* mencionados en la subsección de Métodos con los datos de prueba. Se muestran los resultados tanto para la selección (1) como para la selección (2).

	Acc	S	P	E	AUC
SVM	0,78	0,94	0,81	0,20	0,70
KNN	0,74	0,83	0,83	0,40	0,67
GNB	0,52	0,44	0,89	0,80	0,74
RF	0,70	0,78	0,82	0,40	0,69
GB	0,61	0,72	0,76	0,20	0,66
LR	0,70	0,78	0,82	0,40	0,61
MLP	0,52	0,50	0,82	0,60	0,69

Tabla 2. Resultados de predicción de la respuesta al tratamiento con la selección (1).

	Acc	S	P	E	AUC
SVM	0,78	0,83	0,88	0,60	0,77
KNN	0,70	0,78	0,82	0,40	0,80
GNB	0,74	0,89	0,80	0,20	0,66
RF	0,87	0,94	0,89	0,60	0,70
GB	0,87	0,94	0,89	0,60	0,73
LR	0,87	0,89	0,94	0,80	0,88
MLP	0,87	0,94	0,89	0,60	0,90

Tabla 3. Resultados de la predicción de la respuesta al tratamiento con la selección (2).

Respecto a la selección (1), se pueden observar en la *Tabla 2*, que a pesar de que GNB y SVM arrojen los AUCs más altos, el resto de las métricas no están muy equilibradas, lo que indica que no está generalizando correctamente. Esto podría deberse a que estos modelos, con las características seleccionadas se ven más afectados debido al desequilibrio de clases. Igualmente, en cuestión de balance, los modelos KNN y RF presentan mejores valores en relación con las métricas, con el AUC de 0,67 y 0,69, respectivamente. Y respecto a la selección (2), se puede observar en la *Tabla 3*, los resultados son mejores a la selección (1) con métricas más balanceadas, obteniendo modelos con mejores AUCs, como el LR y el MLP con AUCs de 0,88 y 0,90, respectivamente. Otro aspecto importante a comentar es que, no coincide los algoritmos que mejores hallazgos derivan.

Esta discrepancia puede deberse a las diferencias en el enfoque y las características utilizadas en cada una de ellas. En general, la capacidad de predecir la respuesta al tratamiento entre pacientes que responden y no responden al tratamiento es consistente con los hallazgos reportados en estudios previos [8].

La principal limitación del presente estudio es el relativamente bajo número de pacientes (75), lo que puede conllevar a un desequilibrio entre las clases objetivo afectando a la capacidad de generalización y a un sobreajuste. Otras limitaciones que considerar son: la posible heterogeneidad de las muestras con diferentes estadios y enfoques de tratamiento; y la configuración de los parámetros de adquisición y los algoritmos de reconstrucción de los equipos de TC.

Finalmente, algunas líneas futuras para mejorar los resultados son la estandarización del proceso de adquisición, reconstrucción y segmentación de imágenes de TC; utilización de otras técnicas como el entrenamiento de redes neuronales para automatizar la identificación y segmentación de tumores pulmonares o realización de estudios de mayor dimensionalidad y multicéntricos, entre otros.

4. Conclusiones

El cáncer de pulmón es el segundo cáncer más diagnosticado y la principal causa de muerte relacionada con el cáncer en todo el mundo. A pesar de avances recientes, el manejo de esta enfermedad en etapas avanzadas sigue siendo desafiante. La radiómica y el aprendizaje automático se han convertido en herramientas muy útiles al mejorar la evaluación radiológica, la predicción de respuestas al tratamiento y la identificación de pacientes con mejores perspectivas de supervivencia. Esto se puede observar en los resultados obtenidos en este estudio, los cuales avalan el uso de modelos predictivos basados en características cuantitativas de imágenes de TC para predecir la respuesta al tratamiento de quimio-radioterapia en pacientes con CPCNP. Concretamente, la selección de LASSO muestra un rendimiento ligeramente superior en comparación con la correlación y el análisis estadístico univariante. Y, además, se destacan características comunes entre diferentes selecciones, lo cual indica una mayor relevancia en la tarea objetivo, como por ejemplo el “Número de ciclos de quimioterapia”. En conclusión, el presente estudio plantea una oportunidad para desarrollar aplicaciones de modelos predictivos basadas en imágenes médicas, que ayuden a modo de "biopsia virtual" evitando pruebas diagnósticas invasivas, en la determinación de enfoques de tratamiento más personalizados al evaluar la posible respuesta antes de su aplicación.

Agradecimientos

Este trabajo se ha realizado dentro del marco de la comunidad EELISA Health in the City.

Referencias

- [1] S. Lareau, C. Slatore, and R. Smyth, “Lung Cancer,” *Am J Respir Crit Care Med*, vol. 204, no. 12, pp. P21–P22, Dec. 2021, doi: 10.1164/RCCM.20411P21.
- [2] H. Sung *et al.*, “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” *CA Cancer J Clin*, vol. 71, no. 3, pp. 209–249, May 2021, doi: 10.3322/CAAC.21660.
- [3] P. M. de Groot, C. C. Wu, B. W. Carter, and R. F. Munden, “The epidemiology of lung cancer,” *Transl Lung Cancer Res*, vol. 7, no. 3, p. 220, Jun. 2018, doi: 10.21037/TLCR.2018.05.06.
- [4] A. A. Thai, B. J. Solomon, L. V. Sequist, J. F. Gainor, and R. S. Heist, “Lung cancer,” *The Lancet*, vol. 398, no. 10299, pp. 535–554, Aug. 2021, doi: 10.1016/S0140-6736(21)00312-3.
- [5] M. B. Schabath and M. L. Cote, “Cancer Progress and Priorities: Lung Cancer,” *Cancer Epidemiology, Biomarkers & Prevention*, vol. 28, no. 10, pp. 1563–1579, Oct. 2019, doi: 10.1158/1055-9965.EPI-19-0221.
- [6] F. Nasim, B. F. Sabath, and G. A. Eapen, “Lung Cancer,” *Med Clin North Am*, vol. 103, no. 3, pp. 463–473, May 2019, doi: 10.1016/J.MCNA.2018.12.006.
- [7] F. C. Detterbeck *et al.*, “The International Association for the Study of Lung Cancer Staging Project: Methods and Guiding Principles for the Development of the Ninth Edition TNM Classification,” *Journal of Thoracic Oncology*, vol. 17, no. 6, pp. 806–815, Jun. 2022, doi: 10.1016/j.jtho.2022.02.008.
- [8] B. Chen, L. Yang, R. Zhang, W. Luo, and W. Li, “Radiomics: an overview in lung cancer management—a narrative review,” *Ann Transl Med*, vol. 8, no. 18, pp. 1191–1191, Sep. 2020, doi: 10.21037/ATM-20-4589.
- [9] A. Zwanenburg *et al.*, “The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping,” *Radiology*, vol. 295, no. 2, pp. 328–338, May 2020, doi: 10.1148/RADIOL.2020191145/SUPPL_FILE/R191145S_UPPF2.JPG.
- [10] L. H. Schwartz *et al.*, “RECIST 1.1-Update and clarification: From the RECIST committee,” *Eur J Cancer*, vol. 62, pp. 132–137, Jul. 2016, doi: 10.1016/J.EJCA.2016.03.081.

Estimación de la proporción de hidrops endolinfático en pacientes con la enfermedad de Ménière

G.R. Delicado-Correa ¹, C. Fuster-Barceló ^{1,2}, V. M. Suárez-Vega ³, N. Pérez-Fernández ³, G. Bastarrika ³, A. Muñoz-Barrutia ^{1,2}

¹ Departamento de Bioingeniería, Universidad Carlos III de Madrid, Leganés, España; 100428677@alumnos.uc3m.es; cafuster@pa.uc3m.es; mamunozb@ing.uc3m.es

² Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España

³ Clínica Universidad de Navarra, Pamplona and Madrid, España; {vvega, nperezfer, gbastarrika}@uva.es

Resumen

El hidrops endolinfático (EH) ha sido identificado como una manifestación patológica de la enfermedad de Ménière (MD) desde hace más de 85 años, aunque su confirmación era solo posible mediante análisis histológicos *post-mortem*. En los últimos años, se han logrado avances significativos en las técnicas de adquisición de imagen del oído interno utilizando resonancia magnética (MRI) lo que ha permitido la visualización de la EH *in vivo* en pacientes con sospecha de MD. En este estudio, se han utilizado las secuencias MRI 3D-T2-SPACE-MRC y 3D-REAL-IR de 90 pacientes facilitadas por la Clínica Universidad de Navarra. Sobre ellas se ha establecido una estrategia para estimar de manera automática la proporción de EH mediante la segmentación automática del vestíbulo y la endolinfa contenida. Asimismo, se han evaluado diferentes modelos de predicción para optimizar el rendimiento y se ha encontrado que este enfoque sienta las bases para estimar automáticamente la proporción de EH en pacientes con MD. Estos hallazgos abren una vía para futuros estudios de perfeccionamiento y ofrecen una herramienta prometedora para el diagnóstico y seguimiento de la MD.

1. Introducción

La enfermedad de Ménière (MD, de sus siglas en inglés) es un trastorno del oído interno caracterizado por episodios recurrentes de vértigo, hipoacusia, acúfenos y presión en el oído [1]. A pesar de su relevancia clínica, la causa subyacente de esta enfermedad aún se desconoce [2].

Debido a la subjetividad de los síntomas, el diagnóstico de la MD presenta ciertos desafíos [3]. La manifestación estructural asociada con la MD se caracteriza por la distensión del compartimento endolinfático dentro del oído interno. El aumento de endolinfa provoca hinchazón y presión, que conduce a una distorsión del equilibrio y de la información del sonido, que viaja al cerebro [4]. Esta patología fue descrita por primera vez en 1938 y se denomina hidrops endolinfático (EH, de sus siglas en inglés). Sin embargo, durante años, solo era posible confirmar esta observación patológica a través de análisis histológicos *post mortem*.

La fisiopatología precisa que subyace al EH sigue siendo incierta; sin embargo, las teorías predominantes proponen un desequilibrio entre la secreción y reabsorción de endolinfa en la cóclea [5]. Dado la falta de biomarcadores alternativos *in vivo*, la resonancia magnética (MRI) ha despertado un gran interés para el diagnóstico del EH, especialmente a través del uso de secuencias post-gadolinio de adquisición retardada [6]. En las últimas dos décadas, ha surgido una cantidad

creciente de evidencia que destaca la utilidad diagnóstica de la resonancia magnética para la detección de EH y su correlación con los síntomas audiovestibulares [7]. Como resultado de estos avances, ahora es posible visualizar el EH *in vivo* en pacientes sospechosos de MD [2], aunque la estimación de la proporción en la práctica clínica se realiza de forma manual. Para fines anatómicos, la secuencia 3D Sampling Perfection with Application-optimized Contrasts using different flip angle Evolutions Magnetic Resonance Cisternograph (3D-SPACE-MRC) produce una imagen ponderada en T2 de muy alta resolución con un excelente contraste entre estructuras que facilita la distinción de diversos tejidos y espacios llenos de líquido. Esto permite una visualización detallada de las estructuras y la anatomía del oído interno [8].

Park *et al.* en [9] presentaron una estimación automática de la proporción de EH en MRIs adquiridas con la secuencia 3D FLuid Attenuated Inversion Recovery (3D-FLAIR), una secuencia con recuperación de la inversión que neutraliza los fluidos. A diferencia de 3D-FLAIR, la secuencia 3D Inversion-Recovery with REAL reconstruction (3D-REAL-IR) permite diferenciar la visualización de la perilinfia, la endolinfa y el hueso [10].

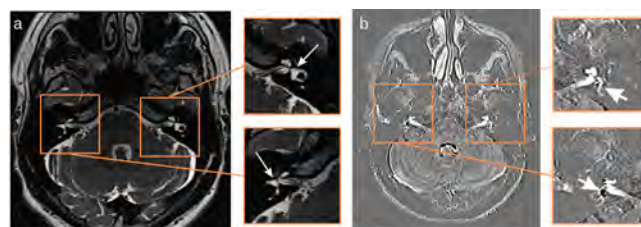


Figura 1. Secuencias MRC (a) y 3D-REAL-IR (b) con subimágenes del oído derecho e izquierdo. La imagen REAL-IR ha sido mejorada mediante un proceso de estiramiento de contraste, donde se ha ajustado la escala de intensidades para resaltar los detalles. Flechas largas: vestíbulo. Flechas cortas: endolinfa vestibular.

El objetivo de este estudio ha sido desarrollar una metodología para estimar cuantitativamente la proporción de EH en imágenes MRI (3D-SPACE-MRC y 3D-REAL-IR). Esto implica el uso de técnicas de pre-procesamiento de imágenes y una red neuronal pre-entrenada, la U-Net, para segmentar ambas imágenes. Luego, se calcula el ratio de EH como el cociente entre el área del espacio endolinfático vestibular (segmentado en la imagen 3D-REAL-IR) y el área total del vestíbulo (segmentado en la imagen 3D-SPACE-MRC). Ambos tipos de imagen, así como las áreas de interés, se muestran en la Figura 1. Esta contribución avanza hacia la

automatización del cálculo del EH, lo que podría mejorar el diagnóstico y el seguimiento de la enfermedad.

2. Materiales y Métodos

2.1. Pacientes

La cohorte de este estudio incluye 90 individuos con EH, de los cuales 83 tienen MD unilateral definitiva. Estos pacientes fueron seleccionados del archivo clínico de la Clínica Universidad de Navarra (CUN). Las edades varían entre los 28 y 75 años, con un promedio de 55 años. La composición de la cohorte fue equitativa en cuanto al género, con un 51,11 % hombres y un 48,89 % mujeres.

2.2. Adquisición de las imágenes

Los estudios MRI se realizaron en la CUN utilizando dos imanes Siemens de 3 Tesla con una bobina dedicada para cabeza/cuello: un Siemens Magnetom Vida con una bobina de 20 canales y un Siemens Magnetom Skyra con una bobina de 32 canales. Las imágenes se obtuvieron cuatro horas después de administrar una dosis única de gadolinio intravenoso (Gadovist; Bayer-Schering Pharma, Berlín, Alemania; 1.0 mmol/mL a una dosis de 0.1 mmol/kg).

Los parámetros de adquisición para las dos secuencias utilizadas en este estudio (3D-SPACE-MRC y 3D-REAL-IR) se muestran en la Tabla 1

Parámetro	MRC	3D REAL IR
Grosor de sección (mm)	0.5	0.8
TR (Tiempo de Repetición, ms)	1400	16000
TE (Tiempo de Eco, ms)	155	551
TI (Tiempo de Inversión, ms)	-	2700
Ángulo de inclinación (°)	120	125
Bandwidth (Hz/píxel)	289	434
Tamaño del voxel (mm)	0.5x0.5x0.5	0.5x0.5x0.8
Número de excitaciones	2	2
Tamaño (n. píxeles)	320x320	324x384
Número de secciones	56	112
Tiempo (min)	5	11

Tabla 1: Parámetros de Adquisición MRI para las secuencias MRC y 3D REAL IR.

2.3. Anotación de las imágenes

Las imágenes del estudio provienen de archivos DICOM proporcionados por la CUN y se enfocan en el vestíbulo para las imágenes MRC y en la endolinfa contenida en él para las imágenes REAL-IR. Se han seleccionado cuidadosamente los cortes principales que contienen estas áreas de interés máximas para cada oído, junto con cortes anterior y posterior inmediatos, siguiendo criterios previamente establecidos. El número de cortes seleccionados varía de 3 a 6 por paciente y secuencia, dependiendo de la ubicación de las áreas de interés. Las anotaciones han sido realizadas con el plugin Labkit¹ de ImageJ, siguiendo las pautas de neuroradiólogos de la CUN. El conjunto de datos anotados incluye 438 imágenes MRC y 409 imágenes 3D-REAL-IR, junto con sus máscaras binarias utilizadas para entrenar el modelo.

¹<https://imagej.net/plugins/labkit/>

²<https://www.image-net.org/>

2.4. Aumento de datos

Con el propósito de mejorar la capacidad de generalización de nuestra red, se ha aumentado el conjunto de datos de entrenamiento utilizando transformaciones tradicionales, incluyendo volteo horizontal, rotación (máximo 5°), desplazamiento horizontal y vertical (máximo 10%), y cambio de intensidad/brillo (rango 0.5-1.5). Estas transformaciones han sido aplicadas aleatoriamente tanto a las imágenes como a sus máscaras correspondientes, con excepción del cambio de intensidad/brillo que solo ha sido aplicado a las imágenes debido a la naturaleza binaria de las máscaras.

El aumento de datos solo ha sido aplicado en el conjunto de entrenamiento para garantizar que el modelo es evaluado en datos realistas, reflejando así la variabilidad y las características reales de las imágenes que encontradas durante la implementación.

2.5. Preprocesado

2.5.1. Recorte

Se ha realizado un recorte semiautomático en las imágenes de resonancia para extraer las regiones de interés relevantes, eliminando áreas innecesarias y enfocando el modelo en las estructuras clave para la segmentación (ver Figura 1). Este proceso también ha ayudado a compensar el desbalance de clases (0, 1). El recorte se ha aplicado tras calcular manualmente los centros de las regiones de interés correspondientes a las orejas derecha e izquierda en ambas secuencias, lo que ha permitido la extracción de sub-imágenes de 96x96 píxeles dentro de las imágenes originales.

2.5.2. Normalización

Para estandarizar los valores de intensidad de las imágenes, reduciendo el impacto de las variaciones y mejorando la compatibilidad entre ellas, se ha utilizado el escalado min-max.

2.5.3. Mejora de imágenes 3D-REAL-IR

La mejora de la calidad de las imágenes 3D-REAL-IR ha sido un aspecto central de interés en esta investigación debido a la inherente baja calidad de estas imágenes. Para abordar este problema, hemos aplicado un ajuste de contraste adaptativo basado en la media y la desviación estándar. Además, dado que la región de interés se presenta como negra sobre un fondo gris, hemos invertido la intensidad de los píxeles para resaltar la región de interés en blanco, al igual que las imágenes de secuencia MRC. Estas modificaciones se han realizado con el objetivo de mejorar la visualización y aumentar el rendimiento de la red con las imágenes 3D-REAL-IR.

2.6. Entrenamiento

En este estudio, se ha empleado la arquitectura U-Net en combinación con el encoder ResNet34, pre-entrenado en ImageNet² para la segmentación de imágenes. El mismo modelo se ha utilizado para las secuencias MRC y 3D-REAL-IR. Se han entrenado diferentes configuraciones utilizando funciones de costo Tversky (con α y β como hiperparámetros que controlan el equilibrio entre los falsos positivos y los falsos negativos) y Jaccard, junto con el optimizador Adam y una tasa de aprendizaje de 0.0001. La métrica utilizada para

la evaluación ha sido la Intersección sobre Unión (IoU, por sus siglas en inglés). El entrenamiento se ha llevado a cabo en una GPU NVIDIA A100 y tomó aproximadamente 7 minutos, utilizando un conjunto de datos (tras el aumento) de 12,200 imágenes MRC y 11,360 imágenes 3D-REAL-IR, con un total de 15 epochs.

2.7. Evaluación

Además de la IoU utilizada durante el entrenamiento, se ha evaluado el rendimiento de los modelos en el conjunto de prueba mediante otros parámetros clave como la exhaustividad (*recall* en inglés), la precisión (*precision* en inglés) y el *F-1 score* (el balance entre precisión y exhaustividad). Estas métricas proporcionaron una visión integral del desempeño de los modelos y facilitaron su comparación.

En la Figura 2, se presenta el proceso de pre-procesado, entrenamiento y evaluación del modelo.

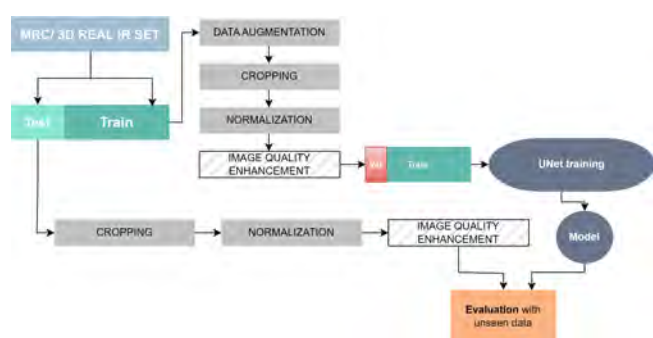


Figura 2. Proceso de entrenamiento y evaluación del modelo, que incluye separación de datos, aumento de datos, recorte en sub-imágenes, normalización, mejora de calidad (solo para 3D-REAL-IR), entrenamiento y evaluación con datos de prueba no vistos previamente.

2.8. Estimación de la proporción del hi-drops endolinfático

Una vez que se han obtenido modelos precisos para la segmentación automática del vestíbulo utilizando imágenes de MRC y la segmentación automática de endolinfa vestibular utilizando imágenes 3D-REAL-IR, el cálculo es directo. Se multiplican los píxeles positivos en la máscara predicha por el tamaño de los píxeles, obteniendo así la superficie del vestíbulo (V) y la superficie de la endolinfa vestibular (EV) y se estima EH como el cociente de dichas áreas:

$$EH = \frac{\text{Superficie}_{EV}}{\text{Superficie}_V}$$

3. Resultados y discusión

En la Tabla 2 se muestran los resultados de la evaluación del rendimiento de la red en tres casos diferentes, utilizando las imágenes 3D REAL IR, tanto en su estado original como con las dos mejoras aplicadas. Se evidencia que el ajuste de contraste adaptativo ha tenido un impacto considerable en el rendimiento, y que este se ha visto potenciado aún más con la posterior inversión de las intensidades de píxeles.

Técnica	Precision	Recall	F1 score	IoU score
<i>Original</i>	1,00	0,76	0,86	0,16±0,20
<i>Ajustado</i>	0,92	0,94	0,93	0,21±0,19
<i>Invertido</i>	0,92	0,94	0,93	0,33±0,26

Tabla 2: Comparación resultados de las predicciones de la secuencia 3D REAL IR con diferentes técnicas de mejora de la calidad. *Ajustado*: ajuste contraste adaptativo. *Invertido*: posterior inversión de las intensidades. Función: Tversky ($\alpha=0,8$; $\beta=1,6$).

Tras comprobar la efectividad de las mejoras, se han analizado los resultados para tres tipos de configuraciones de función de coste en imágenes REAL-IR y MRC, presentados en la Tabla 3. En el caso de las imágenes MRC, la función Tversky, con valores de $\alpha = 0,8$ y $\beta = 1,6$, muestra el mejor desempeño en términos de Índice de Jaccard (IoU) y puntuación F1 (*F1-score*). De media, un 68 % del área predicha coincide con el área de referencia, lo que indica una segmentación precisa y cercana a las áreas reales.

En la misma tabla, se presentan los resultados de la segmentación de la endolinfa vestibular en la secuencia 3D-REAL-IR con mejoras aplicadas. En este caso, la función Tversky con $\alpha = 0,5$ y $\beta = 0,8$, ha resultado ser la mejor opción, logrando un equilibrio óptimo entre precisión y exhaustividad, y obteniendo el mayor índice IoU (33 %).

Función	Precision	Recall	F1 score	IoU score
Tversky (0,5;0,8)	0,97	0,99	0,98	0,65±0,23
Tversky (0,8;1,6)	0,98	0,99	0,98	0,68±0,24
Jaccard	0,98	0,98	0,98	0,63±0,23
Tversky (0,5;0,8)	0,94	0,94	0,94	0,33±0,25
Tversky (0,8;1,6)	0,92	0,93	0,93	0,33±0,26
Jaccard	0,98	0,91	0,94	0,27±0,23

Tabla 3: Resultados de las predicciones de las secuencias MRC (arriba) y 3D REAL IR invertida (abajo).

Los resultados de la evaluación revelan una amplia gama de valores en el índice IoU, subrayando la variabilidad en la calidad de las predicciones. Es importante destacar que las áreas a segmentar son muy pequeñas, especialmente en las imágenes 3D-REAL-IR (generalmente menos de 10 píxeles), por lo que una diferencia de un solo píxel se traduce en una variación que supera el 10 %. Además de por el menor tamaño, el rendimiento de la secuencia 3D-REAL-IR es inferior debido a la limitada calidad de estas imágenes y a la mayor variabilidad en forma e irregularidad del área, en comparación con las imágenes MRC.

Para una mejor apreciación del desempeño del modelo, resulta esencial llevar a cabo una evaluación visual de las segmentaciones predichas (ver ejemplos en la Figura 3). En el caso de las imágenes MRC, a menudo se observa que las segmentaciones están mejor delineadas que la verdad de referencia (GT, por sus siglas en inglés), a pesar de las penalizaciones en los valores IoU. Por otro lado, en las imágenes 3D-REAL-IR, es común que las máscaras predichas no abarquen por completo el área objetivo, lo que sugiere posibles mejoras me-

diante pos-procesamiento, como por ejemplo, aplicando un algoritmo de crecimiento de regiones basado en intensidades similares.

En ambos casos, se han identificado pequeñas discrepancias y errores en el conjunto de GT. Aproximadamente, 10 sub-ímagenes no tenían segmentación y cuando el modelo ha predicho la máscara correctamente, se ha penalizado con un IoU de 0, lo que ha afectado significativamente la media de los resultados.

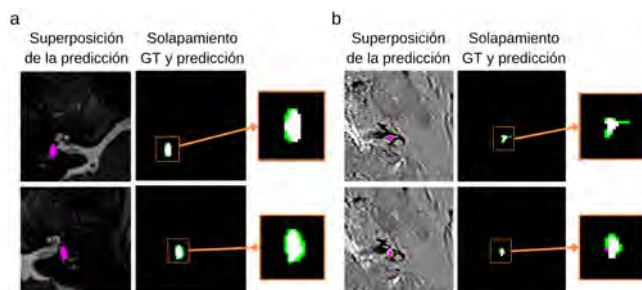


Figura 3. Comparación de los resultados MRC (a) y REAL-IR invertida (b). La primera columna muestra la predicción (magenta) sobre la imagen a segmentar, y la segunda columna muestra el solapamiento entre GT (Verdad de Referencia, verde) y la predicción, con el área en común representada en blanco.

Por último, se han comparado las áreas obtenidas manualmente para el cálculo del EH ratios con las automáticas como prueba de concepto. Los resultados se presentan en la Tabla 4, y como era de esperar, se observan diferencias debido a las discrepancias en la superficie endolinfática.

Caso	Superficie Vestibular (mm ²)		Superficie Endolinfática (mm ²)		EH ratio	
	Man.	Auto	Man.	Auto	Man.	Auto
	1	14,75	15,00	10,04	6,24	0,68
2	13,25	13,50	5,15	2,71	0,39	0,20
3	14,75	14,00	5,15	5,43	0,35	0,39
4	14,50	14,75	8,41	8,95	0,58	0,61
5	15,00	14,00	3,80	3,53	0,25	0,25

Tabla 4: Mediciones manuales (Man.) y automáticas (Auto) de superficie vestibular, superficie endolinfática y ratio EH en varios casos.

4. Conclusiones

En resumen, los modelos de segmentación entrenados con la función Tversky muestran un rendimiento preciso. La aplicación de técnicas de mejora de imágenes 3D-REAL-IR no solo mejora significativamente el rendimiento de los modelos, sino que también aporta un valor clínico crucial al mejorar la visualización de estas imágenes. Aunque los modelos analizados en este artículo, especialmente el destinado a la segmentación del vestíbulo en MRC, generan segmentaciones confiables, reconocemos la necesidad de introducir técnicas de pos-procesamiento para mejorar aún más la precisión en la estimación automática del EH. En conjunto, estos resultados sientan una base prometedora para futuras innovaciones

en el diagnóstico de enfermedades vestibulares como MD.

Agradecimientos

Este trabajo ha estado financiado parcialmente por el Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación, bajo el contrato PID2019-109820RB-I00, MCIN/AEI/10.13039/501100011033, y por el Instituto de Salud Carlos III mediante el proyecto "19/", cofinanciados por el European Regional Development Fund (ERDF, 'Á way of making Europe').

Referencias

- [1] J. A. Lopez-Escamez, J. Carey, W.-H. Chung, J. A. Goebel, M. Magnusson, M. Mandalà, D. E. Newman-Toker, M. Strupp, M. Suzuki, F. Trabalzini, and A. Bisdorff. Diagnostic criteria for menière’s disease. *Journal of vestibular research*, 25(1):1–7, 2015.
- [2] A. Jasińska, M. Lachowska, E. Wnuk, K. Pierchała, O. Rowiński, and K. Niemczyk. Correlation between magnetic resonance imaging classification of endolymphatic hydrops and clinical manifestations and audiovestibular test results in patients with definite ménière’s disease. *Auris Nasus Larynx*, 49(1):34–45, 2022.
- [3] W. Zhang, L. Hui, B. Zhang, L. Ren, J. Zhu, F. Wang, and S. Li. The correlation between endolymphatic hydrops and clinical features of meniere disease. *The Laryngoscope*, 131(1):E144–E150, 2021.
- [4] MD John C Li. Ménière disease (idiopathic endolymphatic hydrops). <https://emedicine.medscape.com/article/1159069-overview>, Oct 2021.
- [5] Y.-Jr Jeng and Y.-H. Young. Comparison of inner ear deficits in meniere’s variants and their significance. *Acta Oto-Laryngologica*, 141(7):684–688, 2021.
- [6] S. E. J. Connor and I. Pai. Endolymphatic hydrops magnetic resonance imaging in ménière’s disease. *Clinical Radiology*, 76(1):76–e1, 2021.
- [7] S. Connor, M. T. Grzeda, B. Jamshidi, S. Ourselin, J. V. Hajnal, and I. Pai. Delayed post gadolinium mri descriptors for meniere’s disease: a systematic review and meta-analysis. *European Radiology*, pages 1–23, 2023.
- [8] J. Cc Benson, M. L. Carlson, and J. I. Lane. Mri of the internal auditory canal, labyrinth, and middle ear: how we do it. *Radiology*, 297(2):252–265, 2020.
- [9] C. J. Park, Y. S. Cho, M. J. Chung, Y.-K. Kim, H.-J. Kim, K. Kim, J.-W. Ko, W.-H. Chung, and B. H. Cho. A fully automated analytic system for measuring endolymphatic hydrops ratios in patients with meniere disease via magnetic resonance imaging: deep learning model development study. *Journal of Medical Internet Research*, 23(9):e29678, 2021.
- [10] M. Zhao, H. Jiang, S. Zhang, Z. Liu, Q. Li, X. Wen, J. Hu, M. Zeng, X. Xie, and J. Lin. Comparison of an optimized 3D-REAL IR and a 3D-FLAIR with a constant flip angle in the evaluation of endolymphatic hydrops. *European Journal of Radiology*, 158:110614, 2023.

Estudio comparativo de algoritmos de segmentación de imágenes médicas para la generación de modelos para la planificación quirúrgica

T. Corpas del Moral^{1,*}, A.P. Seiffert^{1,*}, E.J. Gómez^{1,2}, J. Pedraja³, P. Sánchez-González^{1,2}

¹ Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España; mt.corpas@alumnos.upm.es, {ap.seiffert, enriquejavier.gomez, p.sanchez}@upm.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, España

³ Hospital Virtual Valdecilla, Santander, España; jpedraja@hvvaldecilla.es

* T. Corpas del Moral y A.P. Seiffert son primeros autores compartidos.

Resumen

La planificación quirúrgica es un proceso esencial en la cirugía hepática, tratamiento principal en las primeras fases del cáncer de hepatobiliar. De especial importancia es la relación que guarda el tumor con la anatomía vascular hepática. Para ello, se adquieren imágenes de tomografía computerizada (TC) o resonancia magnética antes de la cirugía. Se segmentan los elementos esenciales y se generan modelos tridimensionales que, posteriormente, el cirujano puede visualizar mediante gafas de realidad virtual, aumentada o mixta. En este trabajo, se lleva a cabo un estudio comparativo para evaluar distintos métodos para la generación de modelos de la vasculatura hepática a partir de imágenes TC preoperatorias. En concreto, se comparan 4 algoritmos de segmentación y 3 grados de suavizado del modelo final. Para llevar a cabo la comparación, se extraen métricas cuantitativas y se comparan estadísticamente con los modelos generados previamente empleados para la planificación quirúrgica. Los resultados alcanzan valores de correlación intraclass de hasta 0,79, comparando los volúmenes de las mallas generadas. El estudio resalta la necesidad de la combinación de algoritmos automáticos y un ajuste o adición manual para obtener modelos precisos y usables para la planificación quirúrgica.

1. Introduction

El cáncer de hepatobiliar es una enfermedad por la se forman células cancerosas en los tejidos del hígado, órgano situado detrás de las costillas en la parte superior del abdomen. Es la tercera causa principal de muertes relacionadas con el cáncer en todo el mundo con unas 830 180 defunciones en 2020 [1]. Se estima que 905 677 personas fueron diagnosticadas con cáncer de hígado en 2020 [1].

La elección del tipo de tratamiento depende del estadio del tumor. Cada paciente es cuidadosamente evaluado y el tratamiento seleccionado está basado en sus características individuales como edad, calidad de vida, condición del hígado y salud general del paciente [2]. Existen diferentes tratamientos, siendo los más comunes la cirugía, el trasplante de hígado, la ablación, la terapia intraarterial y las terapias sistémicas. La cirugía y el trasplante son las principales opciones curativas en las primeras fases del cáncer, mientras que la ablación y la terapia intraarterial se aplican en etapas intermedias. Por último, las terapias sistémicas son el principal tratamiento para los pacientes en etapa avanzada [3].

Para llevar a cabo una cirugía o un trasplante seguros, con la menor morbimortalidad, es necesario realizar una evaluación preoperatoria adecuada y un estudio con imágenes de alta precisión [4]. El cirujano debe conocer a la perfección la anatomía del hígado, morfológica y funcional. Para realizar una buena planificación quirúrgica, es aconsejable hacer una tomografía computerizada (TC) y una resonancia magnética (RM) 4-6 semanas antes de la cirugía [5]. Gracias a la adquisición de imágenes estructurales 3D, es posible identificar y segmentar elementos esenciales en la cirugía hepática como la anatomía vascular [4]. Posteriormente, estas segmentaciones una visualización 3D en gafas de realidad virtual, aumentada y mixta. Esto permite a los cirujanos interactuar con entidades tangibles y añadir información relevante en tiempo real como instrucciones, constantes vitales o detalle de la anatomía que no pueda ver en imágenes 2D. Se ha demostrado que el uso de modelos 3D y la RV permiten planificar tareas más intuitivas con una interacción más rápida [6].

El objetivo de este trabajo llevar a cabo un estudio comparativo de algoritmos de segmentación de imágenes médicas para la generación de modelos para la planificación quirúrgica en pacientes con tumores hepáticos. Para ello se basa en métricas cuantitativas de las mallas 3D generadas a partir de segmentaciones con 7 algoritmos y como *ground truth* aquellas mallas empleadas para la planificación quirúrgica.

2. Material y métodos

2.1. Materiales

Para este trabajo se han obtenido imágenes radiológicas de 12 pacientes anonimizados proporcionados por el Hospital Virtual Valdecilla, todos de fase portal. A todos estos pacientes, con carcinoma hepatocelular en distintos estadios, se les sometió a una cirugía. La segmentación de los vasos sanguíneos hepáticos y la generación de los modelos se realiza mediante Dicom to Print (D2P) (<https://oqton.com/d2p/>), un software aprobado por la FDA para la planificación quirúrgica preoperatoria. Para extraer las métricas cuantitativas de los modelos generados se emplea MeshLab (<https://www.meshlab.net/>). Finalmente, el análisis estadístico se lleva a cabo con IBM

SPSS Statistics (<https://www.ibm.com/es-es/products/spss-statistics>). Además de las imágenes, se dispone de las segmentaciones de cada paciente realizadas para su planificación quirúrgica, consideradas el *ground truth*.

2.2. Métodos

En primer lugar, se realiza siete segmentaciones con distintos algoritmos (semi)automáticos de los vasos sanguíneos de las imágenes TC. A continuación, se extraen distintas métricas cuantitativas tanto de los nuevos modelos como del *ground truth*. Por último, se realiza un análisis estadístico comparando las métricas entre ambos modelos.

Segmentación de los vasos sanguíneos. Esta tarea se realiza en la aplicación D2P, la cual tiene integrada diferentes algoritmos de segmentación. Incorpora tanto algoritmos automáticos como semiautomáticos para distintas anatomías. En este trabajo se utilizan dos algoritmos automáticos (A1 y A2) y otros dos semiautomáticos (S1 y S2) para los vasos sanguíneos. También incorpora métodos de segmentación más manuales como la interpolación (I), el cual requiere la segmentación de los vasos en varios cortes y el algoritmo rellena los puntos intermedios. Además, se realiza la mezcla de los dos algoritmos automáticos (MA) y de los dos semiautomáticos (MS).

Una vez se obtiene todas las mallas con los distintos algoritmos de segmentación, se les aplica ninguna, 3 o 7 iteraciones del algoritmo de suavizado de D2P. De esta manera se genera un total de 21 mallas en cada fase por cada paciente e imagen TC disponible.

Extracción de las métricas. Para realizar la comparación cuantitativa con el *ground truth* se extrae la superficie, el volumen y el número de triángulos de las mallas generadas usando los distintos algoritmos de segmentación y tipos de suavizado. Dado que las mallas del *ground truth* están divididas por vasos, en arteria, porta y cava, es necesario unirlos previamente a la extracción de las métricas cuantitativas. Una vez conectados todos los vasos se extrae las mismas métricas.

Análisis estadístico. Una vez obtenidos los datos del análisis cuantitativo, se comparan cuantitativamente las métricas entre las mallas generadas y su respectivo *ground truth*. Para ello, se calcula el Índice de Correlación Intraclass (ICC, por sus siglas en inglés) (3, k) el cual se denomina “*Two-Way Mixed*” de acuerdo absoluto (1).

Para evaluar si existen diferencias significativas entre las mallas y el *ground truth*. Se realiza la prueba de rangos con signo de Wilcoxon para muestras relacionadas y se considera un p -valor $< 0,05$ estadísticamente significativo.

3. Resultados

3.1. Segmentación de los vasos sanguíneos

Se ejemplifican los resultados de la segmentación para el paciente 6, una mujer de 56 años con un carcinoma hepatocelular a la que se le realizó un estudio previo a un trasplante. La segmentación de los vasos sanguíneos se realiza en el software D2P, donde se visualiza el TC del paciente y la representación 3D de lo segmentado.

En la Figura 1 se observa las dos segmentaciones automáticas y las dos segmentaciones semiautomáticas. En ellas se puede apreciar que las segmentaciones en las imágenes B y D tienen una menor extensión anatómica que las segmentaciones en la imagen A y C. Además, se observa las dos mezclas de los algoritmos automáticos (E) y los semiautomáticos (F). Se puede apreciar que en la imagen E se segmenta más de lo debido, incluyendo parte del riñón. Por último, en la Figura 2, mostrando la segmentación I, se puede observar que la segmentación realizada es incompleta, ya que solo incluye los vasos grandes.

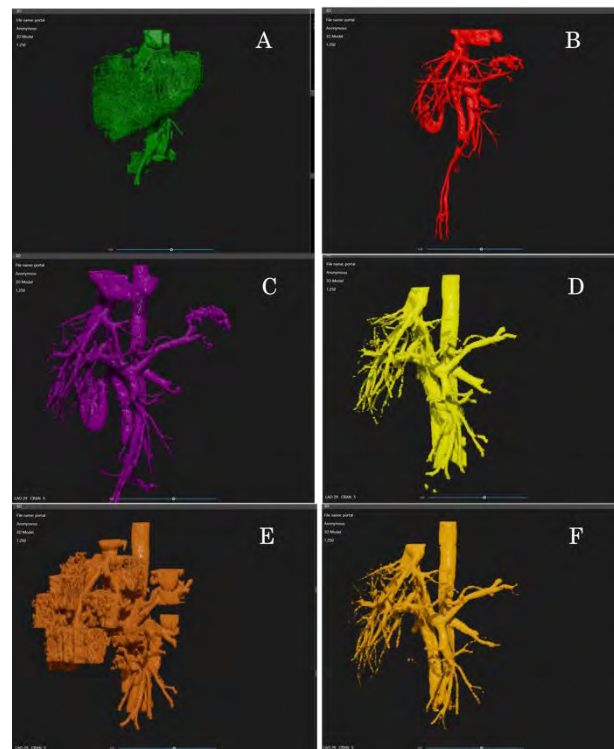


Figura 1. Segmentación A1 (A), A2 (B), S1 (C), S2 (D), MA (E), MS (F) del paciente 6.



Figura 2. Segmentación I del paciente 6

Tras aplicar los siete algoritmos de segmentación en cada paciente, se aplica a cada tipo de segmentación ninguna, tres o siete iteraciones de suavizado. A continuación, se muestra los distintos suavizados que se aplican a las mallas:

3.2. Análisis estadístico

En la primera tabla, la Tabla 1, se obtiene el ICC y el p -valor de la comparación de los volúmenes con el *ground truth*. En ella se puede apreciar que los mejores resultados se obtienen en interpolación, con un valor mayor a

0,75 en todos los suavizados, por lo que se considera una buena fiabilidad. El mayor valor (0,79), se obtiene en la malla con ningún suavizado. En la Tabla 2 se muestran los valores de ICC con sus intervalos de confianza del 95% y el p-valor de las comparaciones del número de triángulos con el *ground truth*. En esta se observa que la

segmentación con mejores resultados es la MS, con un valor de ICC de 0,563, en todos los suavizados. Por último, se encuentra la Tabla 3 donde aparecen los valores de ICC y el p-valor de la comparación entre las superficies de las mallas con el *ground truth*.

	S0		S3		S7	
	ICC (IC 95%)	p-valor	ICC (IC 95%)	p-valor	ICC (IC 95%)	p-valor
A1	0,005 [-0,102-0,257]	0,002	0,005 [-0,103-0,258]	0,002	0,005 [-0,104-0,26]	0,002
A2	0,025 [-0,278-0,466]	0,028	0,025 [-0,288-0,472]	0,034	0,027 [-0,306-0,485]	0,034
S1	0,005 [-0,209-0,384]	0,003	0,004 [-0,214-0,387]	0,003	0,005 [-0,217-0,391]	0,003
S2	0,068 [-0,496-0,597]	0,433	0,065 [-0,475-0,588]	0,347	0,061 [-0,432-0,573]	0,239
MA	-0,076 [-0,363-0,38]	0,06	-0,078 [-0,371-0,382]	0,071	-0,082 [-0,385-0,386]	0,071
MS	0,581 [0,087-0,854]	0,182	0,594 [0,103-0,86]	0,272	0,627 [0,14-0,874]	0,308
I	0,790 [0,409; 0,935]	0,182	0,783 [0,385-0,933]	0,117	0,772 [0,342-0,93]	0,117

Tabla 1. ICC y p-valor entre los volúmenes de las mallas de la fase portal y el ground truth. Aparece en negrita los valores con buena fiabilidad (ICC>0,75).

	S0		S3		S7	
	ICC (95%)	p-valor	ICC (95%)	p-valor	ICC (95%)	p-valor
A1	0,011 [-0,153-0,338]	0,002	0,011 [-0,153-0,338]	0,002	0,011 [-0,153-0,338]	0,002
A2	0,061 [-0,211-0,473]	0,012	0,061 [-0,211-0,473]	0,012	0,061 [-0,211-0,473]	0,012
S1	0,037 [-0,261-0,471]	0,006	0,037 [-0,261-0,471]	0,006	0,037 [-0,261-0,471]	0,006
S2	0,257 [-0,406-0,718]	0,583	0,257 [-0,406-0,718]	0,583	0,257 [-0,406-0,718]	0,583
MA	-0,067 [-0,317-0,358]	0,028	-0,067 [-0,317-0,358]	0,028	-0,067 [-0,317-0,358]	0,028
MS	0,563 [0,055-0,847]	0,06	0,563 [0,055-0,847]	0,06	0,563 [0,056-0,847]	0,06
I	0,229 [-0,098-0,643]	0,003	0,229 [-0,098-0,643]	0,003	0,229 [-0,098-0,643]	0,003

Tabla 2. ICC y p-valor entre los triángulos de las mallas de la fase portal ground truth.

	S0		S3		S7	
	ICC (95%)	p-valor	ICC (95%)	p-valor	ICC (95%)	p-valor
A1	-0,001 [-0,17-0,333]	0,002	-0,003 [-0,166-0,326]	0,002	-0,003 [-0,201-0,362]	0,002
A2	0,023 [-0,389-0,519]	0,005	0,044 [-0,241-0,328]	0,008	-0,052 [-0,266-0,338]	0,015
S1	-0,061 [-0,31-0,392]	0,003	-0,072 [-0,337-0,365]	0,004	-0,073 [-0,375-0,392]	0,006
S2	-0,015 [-0,562-0,542]	0,182	-0,025 [-0,625-0,551]	0,308	-0,052 [-0,673-0,541]	0,53
MA	0,015 [-0,374-0,504]	0,008	-0,093 [-0,31-0,315]	0,015	-0,103 [-0,332-0,319]	0,023
MS	0,267 [-0,146-0,675]	0,019	0,298 [-0,148-0,701]	0,023	0,332 [-0,152-0,729]	0,06
I	0,190 [-0,074-0,596]	0,002	0,179 [-0,072-0,579]	0,002	0,170 [-0,07-0,565]	0,002

Tabla 3. ICC y p-valor entre las superficies de las mallas de la fase portal y el ground truth.

4. Discusión y Conclusión

A continuación, se discuten los resultados obtenidos de las 7 segmentaciones aplicadas a las TCs de los 12 pacientes. A primera vista, se puede apreciar que la segmentación A1 y S1 se descartarían, como se puede ver en Figura 1 y Figura 2. Son algoritmos de segmentación que, en solitario, no aportan mucha información. Se observa, que los algoritmos de A1 y S1 abarcan una proporción mucho mayor que A2 y S2.

A la hora de realizar la segmentación por mezcla de los dos algoritmos automáticos la mayoría de las veces se utiliza A1 para la segmentación de vasos grandes y A2 para los más pequeños. Sin embargo, en 9 TCs no es posible aplicar el A1 pues segmenta más de lo debido y, por lo tanto, en esos casos, la segmentación MA coincide con A2. Por otro lado, en la segmentación MS se utiliza S1 para los vasos grandes y S2 para los más finos. Aunque, como ocurre en MA, hay varias TCs a los que no se puede aplicar S1 por la misma razón que A1 no se puede aplicar en MA. En este caso, ocurre en solo 3 TCs de los 22 segmentados.

En todas las comparaciones hay que tener en cuenta que la obtención del *ground truth* es diferente a la de este proyecto pues se le aplican más ajustes manuales. Además, el *ground truth* está compuesto por vasos de TCs de distintas fases, mientras que, en este caso, se compara una segmentación completa de una misma fase. Otra modificación que es importante destacar es la reducción de triángulos a la que se ve sometida el *ground truth*. Esto es necesario debido a que son usadas para la planificación quirúrgica en el hospital donde la red es limitada y por ello el número de triángulos que soporta las gafas de RV es menor. En la comparación de volúmenes los mejores resultados de ICC se obtienen para I y luego para la MS como se puede ver en la Tabla 1, por lo que son las segmentaciones que más se parecen al *ground truth*. Esto se debe a que la interpolación es un modelo de segmentación semi manual y por ello los vasos segmentados coinciden con los del *ground truth*. El problema de la interpolación es que no es posible segmentar vasos pequeños con este algoritmo de segmentación, además de que se tarda mucho tiempo (una media de 14:30 minutos). Aunque el resto no tiene altos resultados de ICC, ninguno tiene diferencias significativas ($p > 0,05$) excepto A1 y S1 que, como ya se comenta anteriormente, se puede percibir a simple vista. Además, se puede apreciar que al incrementar el suavizado, el ICC mejora excepto en S2 e I. Esto se puede deber a que como al *ground truth* se le reduce tanto el número de triángulos el volumen se reduce y por ello a menor volumen, en una misma segmentación, mayor ICC. En el caso de los triángulos, en la mayoría de las segmentaciones al incrementar el suavizado, no se modifican el número de triángulos y por ello el ICC y p-valor es el mismo para una misma segmentación. Al ser un software de pago en la documentación no aparece como funciona el algoritmo de suavizado. Una posibilidad es que se suavice los vértices promediando los ángulos entre las caras. Esta opción es más probable, pues explicaría la reducción del volumen y de la superficie al aplicar el algoritmo de suavizado. En la comparación de número de triángulos (Tabla 2) se observa que

MS tiene los mejores resultados de ICC, aunque no lo suficientemente bueno para decir que hay una fiabilidad excelente, sino una fiabilidad moderada. El resto tienen un valor muy bajo, incluso la I que obtiene el mejor ICC en la comparación de volúmenes. Esto se explica por lo comentado anteriormente, al poder segmentar solo los vasos grandes tiene una cantidad de triángulos muy pequeña. Como en el caso del volumen, A1 y S1 presentan diferencias significativas y a ellos se les une la interpolación. Por otro lado, en la comparación de la superficie se obtienen valores de ICC mucho peores que en triángulos y volúmenes. Los mejores resultados se dan en la MS. En este caso I obtiene unos valores muy bajos de ICC por el mismo motivo que en el caso de los 38 triángulos. En el resto de los casos se han obtenido unos valores muy bajos. Como en el caso de los triángulos, las segmentaciones A1, S1 e I presentan diferencias significativas.

Este proyecto, presenta varias limitaciones. Solo se puede realizar en 12 pacientes porque era el número que el hospital proporciona. Si el número fuese mayor podría extrapolarse los resultados. Además, no se posee las fases de todos los pacientes y por ello, solo se centra en la fase portal.

En conclusión, según la evaluación cuantitativa, el mejor resultado es la mezcla de los algoritmos semiautomáticos. Además, supone un tiempo menor que cualquiera de los semiautomáticos. Sin embargo, es importante tener en cuenta que, a pesar de usar el mismo programa para la segmentación, las mallas de *ground truth* son modificadas para obtener una mejor representación 3D.

Agradecimientos

Este trabajo se ha realizado dentro del marco de la comunidad EELISA Health in the City.

Referencias

- [1] Sung H, *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 2021, p caac.21660, (ISSN: 0007-9235).
- [2] Bruix J, *et al.* Evidence-Based Diagnosis, Staging, and Treatment of Patients With Hepatocellular Carcinoma. *Gastroenterology*, vol 150, no 4, 2016, pp 835–853, (ISSN: 0016-5085).
- [3] Vogel A, *et al.* Hepatocellular carcinoma. *Lancet*, vol 400, no 10360, 2022, pp 1345–1362, (ISSN: 1474-547X).
- [4] Lizana JR, Castro CM. Nuevos avances en cirugía hepática: Evaluación y planificación preoperatoria (Parte 1). *Revista de cirugía*, vol 74, no 3, 2022, pp 309–317, (ISSN: 2452-4549).
- [5] Robles Campos R, Paricio Pascual P. Cirugía hepática. Arán Ediciones, S.L, 2018, (ISBN: 978-84-17554-12-5).
- [6] Pelanis E, *et al.* Use of mixed reality for improved spatial understanding of liver anatomy. *Minimally invasive therapy & allied technologies*, vol 29, no 3, 2020, pp 154–160, (ISSN: 1365-2931).

Desarrollo de nuevos biomarcadores de imagen para la detección de cáncer de mama.

C. Ortiz Abellán¹, E. Aguado Sarrió¹, J.M. Prats Montalbán^{1,2}, J. Camps-Herrero³, A. Ferrer Riquelme^{1,2}

¹ Kenko Imalytics SL, Valencia, España, {eric,carmen.ortiz}@kenkoimalytics.es

² Grupo de Ingeniería Estadística Multivariante (GIEM), Departamento de Estadística e IO Aplicadas y Calidad, Universitat Politècnica de València, Valencia, España, {joprmon,aferrer}@eio.upv.es

³ Área de Mama, Grupo Ribera Salud, España, jcamps@riberasalud.es

Resumen

El cáncer de mama es uno de los de mayor incidencia en el mundo. Los esfuerzos se orientan hacia una detección más efectiva y temprana. En muchos hospitales, la identificación se realiza cualitativamente por radiólogos, por lo que se necesitan parámetros objetivos (cuantitativos) calculados a partir de biomarcadores de imagen (BIs), capaces de localizar y cuantificar la presencia y grado de desarrollo de determinados procesos biológicos, mejorar el diagnóstico y dar información sobre cómo se desarrollan los tumores. Una forma de abordar el desarrollo de nuevos BIs es utilizar modelos estadísticos multivariantes de variables latentes, que ayudan a comprender las relaciones entre las variables y los fenómenos fisiológicos subyacentes. De esta manera, es posible obtener nuevos BIs fáciles de interpretar para los radiólogos, independientemente de su formación y experiencia, además de identificar píxeles con valores no fiables.

1. Motivación

El cáncer de mama es uno de los tipos de cáncer con mayor incidencia. En 2020 se diagnosticaron 2.261.419 casos de cáncer de mama en todo el mundo [1]. Debido a su prevalencia en las mujeres, los esfuerzos se orientan hacia la búsqueda de una detección más efectiva y temprana. Para la identificación del cáncer, en la mayoría de los servicios de radiología, el estudio de las imágenes de resonancia magnética se realiza de forma cualitativa, a partir del criterio del experto. Por ello, se necesita apoyar esta metodología con parámetros objetivos que informen sobre las características de los tejidos, i.e. biomarcadores de imagen (BIs). Los BIs son mapas de localización y cuantificación relacionados con la presencia y desarrollo de determinados procesos biológicos (v.g., tumores) que proporcionan, para cada píxel, información cuantitativa y objetiva que es interpretada por los facultativos, mejorando el diagnóstico. Estos BIs proceden de secuencias de perfusión (DCE-MRI) y de difusión (DW-MRI).

En la difusión, si bien en determinados tejidos (v.g., la próstata) las restricciones pueden considerarse similares para todas las direcciones y, en consecuencia, puede suponerse isotropía, en el caso de otros órganos (v.g., la mama o el cerebro) hay direcciones preferenciales, por lo que se trata de un caso de difusión anisotrópica donde no es adecuado caracterizar la difusión mediante los modelos

mono o bi exponenciales desarrollados para tejidos isotrópicos. En su lugar, se debe utilizar el tensor de difusión (*Diffusion Tensor Imaging*, DTI) [2-4].

Sin embargo, el uso del DTI da lugar a BIs cuya interpretación clínica no resulta sencilla y pueden confundir al facultativo en su diagnóstico. Además, dichos BIs se basan en el coeficiente de difusión aparente, ADC (principal parámetro de los modelos exponenciales) en cada una de las direcciones principales de difusión en cada uno de los píxeles, y se calculan píxel a píxel, desaprovechando la estructura de correlación interna entre los mismos.

Dentro de los órganos isotrópicos, el modelo más utilizado es el monoexponencial [5], si bien existen otros modelos de tipo biexponencial [6-7], capaces de distinguir los dos fenómenos puros asociados a la difusión: la difusión pura y la pseudo-perfusión. Sin embargo y hasta la fecha, en tejidos anisotrópicos, tan sólo se ha modelado el comportamiento monoexponencial.

El presente trabajo presenta una metodología para el desarrollo de BIs de fácil y directa interpretación clínica, a partir del cálculo del DTI y la posterior modelización de las caídas monoexponenciales por medio de la aplicación del Análisis Multivariante de Imágenes (MIA) usando modelos estadísticos multivariantes de variables latentes [8]. Estos modelos aprovechan la estructura de correlación interna entre los píxeles, obteniendo BIs más robustos, y permitiendo el cálculo de nuevos biomarcadores.

Dada la naturaleza de las secuencias de imágenes analizadas, en este trabajo se aplican modelos de resolución multivariante de curvas (MCR) [9-10]. Mediante MCR, se obtienen nuevos biomarcadores de imagen fáciles de interpretar, independientemente de la formación y experiencia de los facultativos. Además, se proporcionan imágenes de error, localizando aquellos píxeles de valores no fiables.

2. Materiales y métodos

2.1. Base de datos

La base de datos consta de 24 casos de cáncer de mama confirmados histológicamente. Para cada uno de estos

casos se tiene un total de 81 cortes, con una resolución de 336x336 píxeles. Las secuencias DTI se adquirieron en 16 direcciones diferentes con $b=800 \text{ s/mm}^2$, más una imagen adicional adquirida con $b=0$. Por tanto, cada uno de los cortes tiene asignado un total de 17 imágenes, una para el valor base $b=0$ y otras 16 imágenes, una para cada dirección de gradiente. Es decir, 1377 imágenes por caso.

Un radiólogo experto identificó las regiones de interés. Todos los pacientes dieron su consentimiento para usar las imágenes, las cuales fueron anonimizadas previamente.

2.2. Modelos de difusión isotrópicos

Como se ha comentado, dependiendo del entorno en el que se produce la difusión, se pueden distinguir dos tipos: difusión isotrópica, cuando las restricciones del movimiento de las partículas son iguales para todas las direcciones; y difusión anisotrópica, cuando existen direcciones preferentes para la difusión (v.g., la mama).

En la difusión isotrópica, el modelo más utilizado es el modelo monoexponencial, el cual tiene un parámetro asociado, Coeficiente de Difusión Aparente (ADC) que integra el efecto global tanto de los desplazamientos lentos como de los rápidos, a partir de la ecuación 1:

$$\frac{s}{s_0} = e^{-b(ADC)} \quad \text{ec. 1}$$

donde s_0 es el valor inicial de la señal s cuando b es cero.

Existe un modelo pseudo biexponencial denominado movimiento incoherente intravoxel (IVIM), capaz de separar la difusión lenta de la difusión rápida (ec. 2):

$$\frac{s}{s_0} = (1 - f)e^{-bD} + fe^{-b(D+D^*)} \quad \text{ec. 2}$$

donde D es el coeficiente de difusión lenta, D^* el coeficiente de difusión rápida y f la fracción vascular. Es decir, la señal normalizada s/s_0 se modela como una media ponderada de la difusión lenta y la difusión rápida.

2.3. Tensor de difusión: DTI

Para el caso anisotrópico, existen direcciones preferentes para la difusión del agua. Por lo tanto, no es adecuado caracterizar la difusión por medio del ADC. Por ello, se define una matriz \mathbf{D} , tensor de difusión, de dimensiones 3x3, la cual modela la difusión aparente en cada dirección del espacio en un píxel en concreto [14]:

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{pmatrix} \quad \text{ec. 3}$$

Generalizando el modelo exponencial al caso anisotrópico, el modelo que relaciona la intensidad medida con el tensor de difusión en función de la dirección de gradientes es:

$$I = I_0 e^{-\mathbf{B}:\mathbf{D}} \quad \text{ec. 4}$$

donde:

I : intensidad en un píxel al aplicar el gradiente de difusión.

I_0 : intensidad en dicho píxel sin la aplicación del gradiente.

\mathbf{B} : matriz que define el gradiente aplicado y su dirección.

\mathbf{D} : tensor de difusión en dicho píxel.

La matriz \mathbf{B} puede descomponerse en el parámetro b del modelo isotrópico y una matriz \mathbf{G} la cual contiene información de la dirección del gradiente:

$$\mathbf{B} = b\mathbf{G} \quad \text{ec. 5}$$

La matriz \mathbf{G} se obtiene del vector director del gradiente, \mathbf{g}_M , formado por tres componentes (g_x, g_y, g_z) (Figura 1).

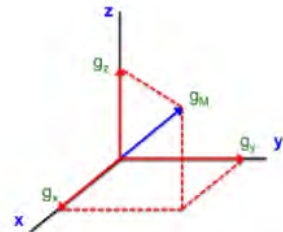


Figura 1. Vector gradiente \mathbf{g}_M y sus componentes

Teniendo en cuenta el vector anterior se forma la matriz \mathbf{G} :

$$\mathbf{G} = \mathbf{g}_M \mathbf{g}_M^T = \begin{pmatrix} g_x^2 & g_x g_y & g_x g_z \\ g_x g_y & g_y^2 & g_y g_z \\ g_x g_z & g_y g_z & g_z^2 \end{pmatrix} \quad \text{ec. 6}$$

El producto $\mathbf{G}:\mathbf{D}$ se conoce como producto tensorial y se calcula a través de la expresión siguiente:

$$\mathbf{G}:\mathbf{D} = g_x^2 D_{xx} + g_y^2 D_{yy} + g_z^2 D_{zz} + 2g_x g_y D_{xy} + 2g_x g_z D_{xz} + 2g_y g_z D_{yz} \quad \text{ec. 7}$$

Este modelo puede ser utilizado para calcular el tensor de difusión a partir de las imágenes y de los gradientes usados en su adquisición. Mientras que para caso del modelo isotrópico sólo se requerían dos valores, el base y uno con un parámetro b diferente a cero; para el modelo anisotrópico se requiere un mínimo de siete valores, el base y seis asociados a diferentes direcciones del gradiente.

El cálculo del tensor de difusión se ha realizado en este trabajo por medio de modelos de regresión por mínimos cuadrados parciales (PLS) [11], con el fin de aprovechar la estructura de correlación interna y mejorar la estimación del DTI calculado tradicionalmente mediante mínimos cuadrados ordinarios, OLS [4].

El tensor de difusión \mathbf{D} se puede ver gráficamente como un elipsoide (Figura 2). En general, los ejes del elipsoide no tienen por qué estar orientados en las direcciones del espacio.

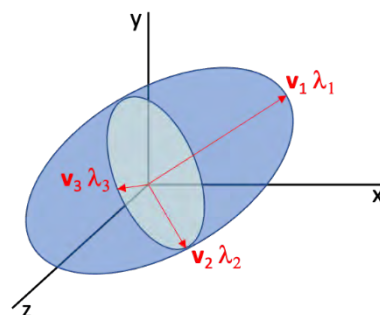


Figura 2. Elipsoide representando el tensor de difusión \mathbf{D}

Si las direcciones de los ejes del elipsoide estuvieran alineadas con los ejes del espacio, el tensor de difusión sería una matriz diagonal (ec 8):

$$D' = \begin{pmatrix} D'_{xx} & 0 & 0 \\ 0 & D'_{yy} & 0 \\ 0 & 0 & D'_{zz} \end{pmatrix} \quad \text{ec. 8}$$

Estos valores del tensor rotado corresponden con los valores propios del tensor de difusión original D , por lo que, para calcularlos, se usa la descomposición en valores y vectores propios [4]:

$$D\mathbf{v}_i = \lambda_i\mathbf{v}_i \quad \text{ec. 9}$$

$$D = V \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} V^T \quad \text{ec. 10}$$

De esta forma, se obtienen los valores propios λ_1 , λ_2 y λ_3 , correspondientes a la magnitud de la difusión en los tres ejes del elipsoide, además de sus vectores propios, los cuales muestran la dirección asociada al valor propio correspondiente. A partir de los valores propios se pueden calcular diferentes biomarcadores.

Cuando un píxel está “sano”, el ducto no está taponado; es decir, hay facilidad de movimiento en la dirección del ducto, por lo que λ_1 va a ser más grande que los otros dos. Sin embargo, si el ducto está taponado, es decir, hay celularización (indicio de proceso tumoral), el valor de λ_1 en la dirección del ducto va a ser mucho menor y similar a λ_2 y λ_3 .

2.4. Modelos de resolución multivariante de curvas

El modelo de resolución multivariante de curvas (MCR), y en concreto el método basado en mínimos cuadrados alternados (MCR-ALS) [12] es un método iterativo que realiza una descomposición de una matriz X por medio de una optimización de mínimos cuadrados alternados ALS:

$$X = CS^T + E \quad \text{ec. 11}$$

donde C es la matriz de contribuciones y S una matriz de componentes o fenómenos puros subyacentes. De esta manera, se asume que una observación (fila) de X acaba siendo el resultado de una media ponderada de los fenómenos puros subyacentes presentes en las columnas de S , en función de la contribución de cada uno de ellos almacenada en la fila correspondiente de C . La matriz E es la matriz de residuos, a partir de la cual se puede obtener el vector de suma de cuadrados residual de cada fila (\mathbf{rss}).

El hecho de no imponer ortogonalidad entre los componentes hace que el número de soluciones sea infinito, por lo que, para llegar a soluciones “reales” desde un punto de vista fisiológico, han de introducirse restricciones.

Para que una imagen pueda ser analizada por MCR es necesario desplegar cada imagen de dimensiones ($n_1 \times n_2 \times J$), donde ($n_1 \times n_2$) es el tamaño de la imagen y J son los diferentes valores b , en una matriz de dos dimensiones donde se tienen ($n_1 \times n_2$) filas y J columnas. Además, para analizar todo el volumen del órgano, se despliega cada imagen de cada corte y se apilan una debajo de otra. Este

tipo de despliegue, conservando las variables J en las columnas, asume que los fenómenos subyacentes son los mismos en todo el volumen del órgano.

Al reconstruir la matriz C nuevamente a la dimensión espacial original, se obtienen nuevas imágenes, denominadas mapas de distribución, que ubican aquellos píxeles relacionados con cada uno de los comportamientos fisiológicos correspondientes proporcionados por el modelo. Al reconstruir el vector \mathbf{rss} se obtiene la imagen de la suma de cuadrados de los residuos del modelo, que localiza aquellos píxeles para los que aparecen ciertos comportamientos que no han podido ser bien modelados.

2.5. Metodología

A partir del cálculo del DTI en cada uno de los píxeles de las imágenes de la secuencia, es posible obtener, para cada píxel, los valores propios λ_1 , λ_2 y λ_3 , además de sus vectores propios \mathbf{v}_1 , \mathbf{v}_2 , y \mathbf{v}_3 escalados a norma uno.

Si en un píxel se aplica un gradiente en la primera dirección principal \mathbf{v}_1 , que se asume la del píxel en dicha localización, la señal que se obtendría sería, a partir de las ecuaciones 3 a 10:

$$I = I_0 e^{-bG':D'} = I_0 e^{-bg'v_1^T \lambda_1} = I_0 e^{-b\lambda_1} \quad \text{ec. 12}$$

Generalizando, se puede decir que a partir de ec. 12 se puede reconstruir la curva de difusión para cada píxel en cada dirección principal y obtener así una estimación de la curva exponencial decreciente de difusión que se habría obtenido si se hubiese aplicado en ese píxel y en esa dirección principal un gradiente con diversos valores de b .

Si esto se realiza para cada píxel, utilizando en cada uno el valor de su primer valor propio, se obtendría una matriz de tantas filas como píxeles y tantas columnas como valores de b se usen para reconstruir la señal de decaimiento exponencial, consiguiendo así una estructura parecida al caso isotrópico. De esta manera, la estructura matricial desarrollada sí es apta para aplicar MCR y estimar un modelo **biexponencial**, ya que el almacenamiento de todos los píxeles en una única matriz permite aprovechar la estructura de correlación interna entre píxeles, obteniendo así nuevos biomarcadores derivados de MCR para tejidos anisotrópicos con una aproximación biexponencial.

Una vez comprobado que los comportamientos puros obtenidos tienen sentido (i.e. una caída exponencial lenta y una caída exponencial rápida), se representan los mapas de distribución (nuevos BIs) en los que se localizan los píxeles más relacionados con cada uno de los comportamientos fisiológicos: caída lenta (celularización) y caída rápida (vascularización). También se representa la suma de cuadrados residual (\mathbf{rss}) para localizar aquellos píxeles donde el modelo no se ajusta correctamente y, por lo tanto, los biomarcadores obtenidos no son fiables.

3. Resultados

Los resultados obtenidos se presentan, a modo de ejemplo, en un caso de estudio. Tras aplicar la metodología planteada, en la Figura 3 se aprecia una caída lenta que se

corresponde con celularización y una caída rápida que se corresponde con vascularización, mientras que la Figura 4 muestra el mapa de distribución correspondiente a la caída lenta de uno de los cortes. Antes de visualizar las imágenes cabe destacar que la zona de interés para este estudio es la mama, es decir, sólo la mitad superior de las imágenes que se van a mostrar.

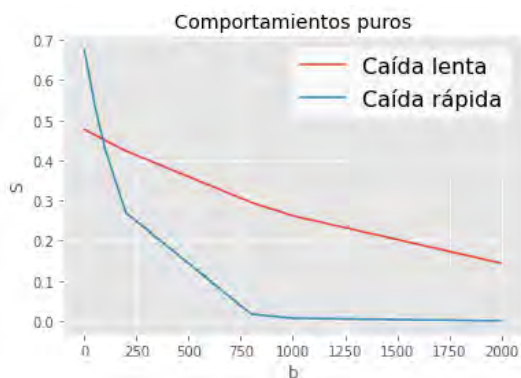


Figura 3. Curvas de los comportamientos puros

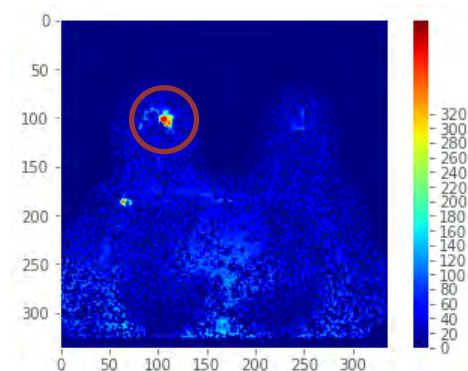


Figura 4. Mapa de distribución de la caída lenta

La región marcada tiene una contribución mayor del comportamiento de la caída lenta, lo cual indica que los ductos tienen mayor restricción al movimiento de las moléculas de agua. Es decir, están más obstruidos, lo cual indica indicio de tumor. Para confirmar que lo que se aprecia en la Figura 4 es tumor, se pasa a observar los residuos del modelo en ese mismo corte (Figura 5). El residuo es muy bajo en la zona de la lesión, lo cual indica que el modelo es correcto, y por tanto la imagen de la Figura 4 se puede considerar fiable.

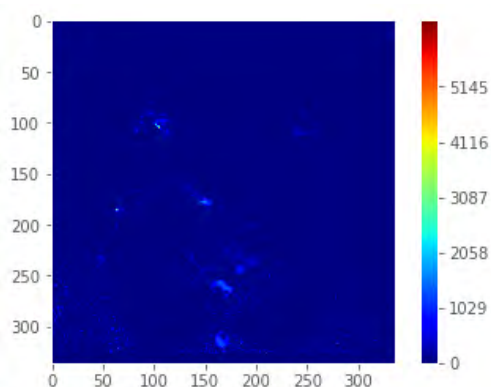


Figura 5. Mapa de distribución de los residuos del modelo

4. Conclusiones

Se ha propuesto un nuevo método para calcular biomarcadores de cáncer de mama a partir de imágenes de resonancia magnética de difusión basados en el tensor de difusión mediante MCR. Los biomarcadores desarrollados permiten una detección clara y de fácil interpretación clínica, con lo que cualquier facultativo, independientemente de su experiencia, podría identificar dónde se encuentra la lesión a través de los biomarcadores de imagen. Al utilizar MCR, se señalan directamente los píxeles relacionados con cada comportamiento fisiológico, facilitando así el diagnóstico clínico.

Agradecimientos

Esta investigación ha sido financiada por el Gobierno de España (Ministerio de Ciencia e Innovación) en el marco del proyecto PID2020-119262RB-I00.

Referencias

- [1] «Cancer.Net» Enero 2022: <https://www.cancer.net/es/tipos-de-c%C3%A1ncer/c%C3%A1ncer-de-mama/estad%C3%ADsticas>. [Último acceso: Septiembre 2022].
- [2] P. B. Kingsley. Introduction to Diffusion Tensor Imaging Mathematics: Part I. Tensors, Rotations, and Eigenvectors. Wiley InterScience, Vol. 28A(2), pp. 101-122, 2006.
- [3] P. B. Kingsley. Introduction to Diffusion Tensor Imaging Mathematics: Part II. Anisotropy, Diffusion-Weighting Factors, and Gradient Encoding Schemes. Wiley InterScience, Vol. 28A(2), pp. 123-154, 2006.
- [4] Kingsley, P.B.. Introduction to difusión tensor imaging mathematics: Part III. Tensor calculation, noise, simulations, and optimization. Wiley InterScience, Vol. 28A(2), 155-179, 2006.
- [5] D. Le Bihan, Molecular diffusion nuclear magnetic resonance imaging, Magn. Reson. Q. 7 (1991) 1–30.
- [6] D. Le Bihan, et al., MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders, Radiology. 161 (1986) 401–407.
- [7] E. Aguado-Sarrió, et al. (2015) Prostate Diffusion Weighted-Magnetic Resonance Image analysis using Multivariate Curve Resolution Methods. Chemometrics and Intelligent Laboratory Systems, 140, 43-48
- [8] J.M. Prats-Montalbán, A. Ferrer, A. de Juan. Multivariate image analysis: a review with applications, Chemom. Intell. Lab. Syst. 107 (2011) 1–23.
- [9] R. Tauler, A.K. Smilde, B.R. Kowalski, Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution, J. Chemometrics. 9 (1995) 31–58.
- [10] R. Tauler, Multivariate curve resolution applied to second order data, Chemom. Intell. Lab. Syst. 30 (1995) 133–146.
- [11] E. Aguado-Sarrió et al. Partial Least Squares - Diffusion Tensor Imaging (PLS-DTI): A novel approach for biomarker imaging in breast cancer. Chemom. Intell. Lab. Syst. 235 (2023) 104777
- [12] A. de Juan, R. Tauler. Chemometrics applied to unravel multicomponent processes and mixtures: Revisiting latest trends in multivariate resolution. Anal. Chim. Acta 500 (2003) 195-210.

Synchronization in Retrospective Respiratory Gating using Deep Learning

Fernando Bayo¹, Claire J. Brooks¹, David Pérez-Benito¹, Patricio López-Expósito¹, Juan José Vaquero^{1,2}

¹ Departamento de Bioingeniería, Universidad Carlos III de Madrid, Madrid, Spain

² Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain

Abstract

Breathing motion introduces artifacts during CT acquisition, what affect the quality and subsequent reconstruction of the images. This study aims to reduce artefact in CT images using deep learning techniques. Specifically, we propose the implementation of an autoencoder based on convolutional neural networks. Once the model was trained, we employed a morphing technique to generate new images with reduced respiratory motion. By analyzing the respiratory signal, we classified the different images into phases and selected those most suitable for correction. Subsequently, we applied the described method, obtaining a more homogeneous data set. The results demonstrate a significant reduction in motion when comparing intensity changes within the regions most affected by motion. Thus, we validated the efficacy of the proposed approach to mitigate breathing-induced artifacts. The application of artificial intelligence (AI) in this field represents a significant advance. This study provides promising initial results and opens up new possibilities for research and development. By complementing existing techniques, AI offers enhanced motion reduction capabilities, thus improving the quality of CT images. The potential for future advances in this direction are substantial, promising further reduction in respiratory motion artifacts and improvement to the overall quality of CT images.

1. Introduction

Tuberculosis remains a threatening global health problem, causing 1.6 million deaths in 2021 [1]. Mouse models aid in the understanding of tuberculosis due to their similarity to human infection and easy manipulation, but respiratory motion in CT images poses difficulties in the form of artifacts that hinder accurate visualization of lung lesions due to diaphragm motion and overlapping structures. Techniques like breath-holding and retrospective respiratory gating have been devised to mitigate respiratory motion artifacts in CT scans. While the latter involves selecting images from a particular phase of the respiratory cycle, the variability in a subject's breathing pattern can lead to missing images for specific phases. To overcome this limitation, our project introduces a novel AI-based solution that ensures the availability of images for each projection.

2. Materials and Methods

2.1. Algorithm

The approach proposed was based on the use of AI techniques to address the problem of motion introduced by respiration in CT images. The main objective was to replace images affected by excessive motion with images artificially corrected to remove the former. The first

step consisted of extracting the respiratory signal from each image to process and identify the level of motion introduced. With this information, images with moderate motion were selected to be corrected by generating new images that resemble the target data set with a lower level of motion. Those with excessive motion were discarded. To perform the correction, an image generating model based on deep learning techniques has been used. This model allowed the creation of a new image by combining two input images, influenced by both inputs in different degrees as established. By applying this morphing process using AI, it was expected to obtain corrected images that were more similar to the desired data set, which improved the quality and usefulness of CT images by reducing motion artifacts caused by breathing. The pipeline is illustrated in Figure 1.



Figure 1: Process pipeline design.

2.2. Data set

Data acquisition was performed in cine mode, using the "step and shoot" technique, in which the X-ray tube was rotated around two subjects in 720 steps covering a 360° rotation. As a result, data sets consisted of 720 projections, each containing a total of 16 frames with a size of 570x518 pixels. While first data set was used to train the model, the second was used for test tasks.

2.3. Respiratory Signal

Since we were acquiring repeated axial CT images during a certain period of time for each position, it was possible to visualize the movement derived from the animal's breathing in the acquired images. However, if we simultaneously displayed all the frames contained in a single projection, it was practically impossible to perceive such motion with the naked eye. This is the case represented in Figure 2, in which the 16 frames contained in projection 0 are shown.

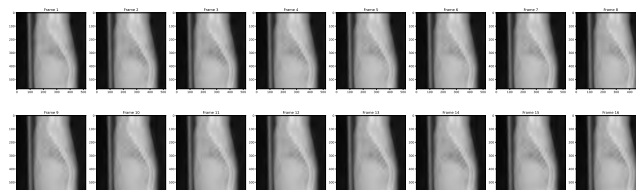


Figure 2: All 16 frames from projection '0'.

To quantify the movement we computed the image resulting from averaging all the frames for the same projection, and we subtracted to each individually frame the average image [4].

The respiratory phase corresponding to each frame, whether inspiration or expiration, was easy to distinguish by looking at the signal intensity of the aforementioned area. This was the case of Figure 4 in which the average image of all frames and the subtraction between both was represented, resulting in another image with a darker region at the height of the diaphragm in which negative intensity values predominate. This indicated that the image corresponded to the inspiration phase. If the opposite case was observed, in which the area was highlighted in white, we would have been looking at an image related to the expiration phase.

Once the identification of respiratory phases has been defined, we proceed with the classification of all the images contained in our data set in three phases that we will call: expiration, inspiration and intermediate.

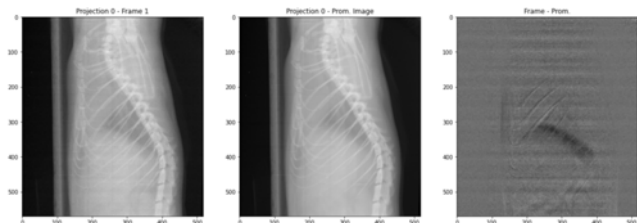


Figure 4: Identification of expiration frame.

To do this, we selected a region of interest (ROI) in the image resulting from the difference between the frame and the average image. We performed a summation of the intensities of the pixels contained in this strategically chosen ROI, which kept the highlighted part related to the movement of the diaphragm. Figure 5 shows the sum of intensity for the 16 frames of the projection we were working with. In it, it was possible to see how the lowest intensity values correspond to those frames that have a black highlighted area in the Figure 4 and the highest values correspond to the images with white highlighted areas.

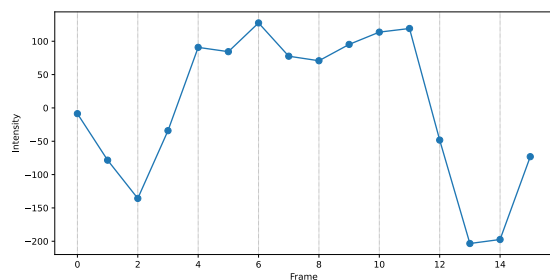


Figure 5: Respiratory signal relative to projection 0.

To classify all available images and determine which ones should be corrected to reduce the motion caused

by respiration, we used the Multi-Otsu (MO) method to divide our data set. The MO is a thresholding algorithm that separates the pixels of an input image into several different classes, each obtained according to the intensity of the gray levels within the image by exhaustively searching for the threshold that minimizes the intra-class variance [5]. Applying this method, we managed to divide the data set into the three phases previously defined and *outliers*. In Figure 6, the green samples correspond to the expiration phase, the yellow ones to the inspiration phase and the red ones to the class known as 'intermediate'. Likewise, the samples that are both above threshold 3 and below threshold 4 are classified as *outliers*, those images of both expiration and inspiration with considerable movement that will introduce a lot of noise if used. In this case, the total count of images classified in phases resulted in 1088 expiration, 9032 intermediate and 1400 inspiration images. Finally, we selected the expiration samples that were above threshold 4, as well as the inspiration samples that were below threshold 3, since these were the samples in which we corrected the movement by morphing to the sample of the same projection with the closest intensity.

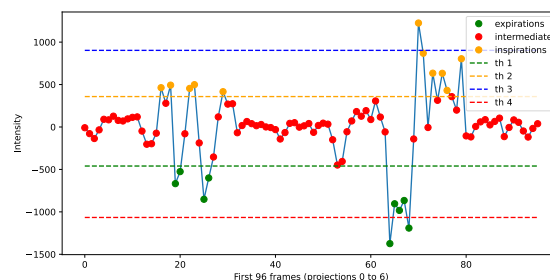


Figure 6: Classified Respiratory signal.

2.4. Model

The use of generative image models has opened up new possibilities in the field of visual content generation. These models use machine learning algorithms to learn the features and patterns present in an image set and generate new images that resemble those in the original ones. In this way, synthetic images are generated that can be used to improve the understanding of features of interest or generate additional training data sets for machine learning algorithms.

The model chosen in this case has been an autoencoder, which is a specific type of a neural network mainly designed to encode the input into a compressed and meaningful representation, and then decode it back such that the reconstructed input is similar as possible to the original one. We proposed an autoencoder architecture based on an encoder (Figure 7) which progressively reduce the dimensions and extract the essential features from the input images and a decoder (Figure 8) that aims to reconstruct the original images from the latent representation. The details of each convolutional neural network (CNN) layer used in the architecture were: The architecture of the autoencoder, comprising encoder

and decoder components, facilitated the compression of input images into a lower-dimensional latent representation and the subsequent reconstruction of images. By training the autoencoder to minimize the Mean Squared Error (MSE) loss between the reconstructed images and the original inputs, the model learned to capture significant features and generate precise reconstructions.

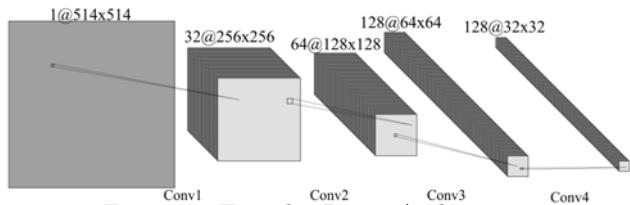


Figure 7: Encoder Layer Architecture

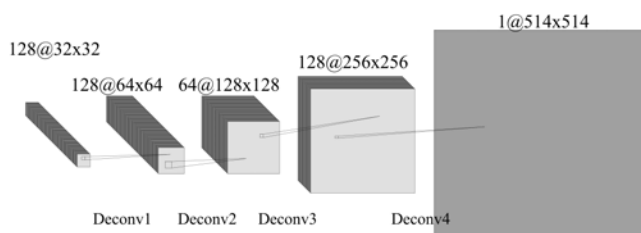


Figure 8: Decoder Layer Architecture

2.5. Morphing Technique

Morphing is a technique that had been used to generate new images lying in the feature space between two original images [6, 7, 8]. The process involved blending the latent representations of two images using a weight factor, which determined the contribution of each image to the morphed result. The weight factor ranged from 0.0 to 1.0 in increments of 0.1, indicating blending proportions. The blending was achieved by multiplying the latent representation and feature maps of the first image by 'w' and the second image by '(1-w)', with 'w' representing the weight factor. Iterating through each weight value, a morphed image was generated by passing the blended latent representation and feature maps through the decoder of the extended autoencoder [9, 10]. This morphing technique allowed for the creation of a series of images smoothly transitioning between the two originals, offering a continuous variation in the feature space, as shown in Figure 9. This enabled the exploration and generation of new artificial images with diverse characteristics.

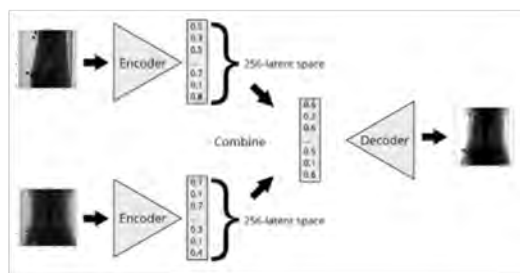


Figure 9: Fusion of the latent representation for both input images.

4. Results

The resulting reconstruction of an image is shown in Figure 10 as an example. This will be smaller due to the adjustment of dimensions that we made on the original image at the input of the autoencoder.

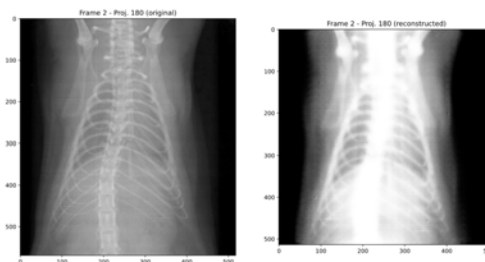


Figure 10: An example of image reconstruction using the model.

We applied the morphing technique to the dataset in order to reduce motion. Figure 6 gave an idea of how many images needed to be corrected. These images corresponded to the expiration and inspiration moments that were in the threshold ranges 1-4 and 2-3, respectively. To correct these images, we selected the frame of the same projection that has the most similar intensity in the range of *intermediate* images. In this way, we transformed the expiration and inspiration images into *intermediate* images. The evaluation method consisted of intensity profiles, drawing a line over the region of the images where we expected to find the larger motion and collecting the intensity values along that line, as shown in Figure 11.

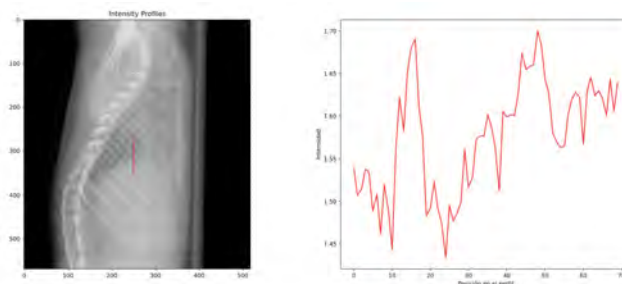


Figure 11: Method of evaluation of diaphragmatic movement on CT images.

We repeated this method for all the frames of the same projection, both for the original and the reconstructed projection. We evaluated the motion reduction after applying a Gaussian filter to remove the noise and thus better see the difference. Figure 12 showed the effect of the movement correction; the difference in intensity between the eight images depicted in the figure reflected the existing motion, as the traces do not match. In order to compare the motion between data sets (original - reconstructed) in different projections, we chose four projections at rotation degrees 1, 105, 220 and 288. The table below showed, for each angle, the number and type of images collected in the modified database: original,

zeros and reconstructed. Therefore, we did not always have the same number of images in the same projection for both data sets. Finally, the right side of the table showed the motion quantification obtained from the profiles drawn in the projections relative to the views in both cases. The results in the table showed a tendency towards motion reduction in the modified data set. The projection at 220 degrees uses 7 original images and 8 transformed images, and it was the one that presented a greater reduction of the movement compared to the use of the original data set, with a difference of 0.183 between the intensity ranges.

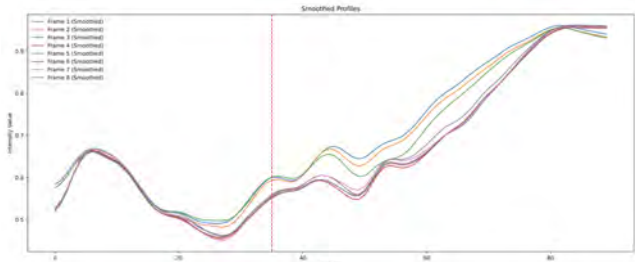


Figure 12: Motion measurement.

View	Proj.	Total Images Used			Intensity range	
		original	zeros	reconst.	original	reconst.
Right	2	11	2	3	0.082	0.017
Frontal	210	13	0	3	0.036	0.013
Left	440	3	6	7	0.161	0.121
Back	576	7	0	8	0.284	0.101

TABLE I

5. Conclusions

The fundamental objective of this work has been to address the challenge of reducing the artifact derived from the breathing of the animal under study.

Our findings conclusively support that by applying AI techniques, in particular by implementing autoencoders and neural networks, it is possible to effectively reduce artifacts and improve the quality of CT images as we use more images to average the same projection. This innovative approach has allowed us to overcome the limitation given by the use of a very small number of images associated with traditional methods.

The improvement of results using AI as a tool in this context represents a promising advance. By leveraging machine learning algorithms, we have been able to model respiratory motion more accurately and dynamically. We are excited about future research and developments in this area, with the goal of further improving medical care and clinical diagnosis. As identified limitations are addressed and more advanced approaches are explored, we are confident that even greater advances in reducing respiratory motion and optimizing CT imaging will be achieved for the benefit of patients and healthcare professionals.

Acknowledgements

This work has been supported by the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No. 853989. The JU receives support

from the European Union’s Horizon 2020 research and innovation programme and EFPIA and Global Alliance for TB Drug Development non-profit organisation, Bill & Melinda Gates Foundation and University of Dundee.

References

- [1] Organización Mundial de la Salud. (2022). Informe mundial sobre la tuberculosis 2021. Licencia: CC BY-NC-SA 3.0 IGO
- [2] PAN, Tinsu, et al. 4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT. *Medical physics*, 2004, vol. 31, no 2, p. 333-340.
- [3] NEHMEH, S. A., et al. Quantitation of respiratory motion during 4D-PET/CT acquisition. *Medical physics*, 2004, vol. 31, no 6, p. 1333-1338.
- [4] CHAVARRÍAS, Cristina, et al. Extraction of the respiratory signal from small-animal CT projections for a retrospective gating method. *Physics in Medicine Biology*, 2008, vol. 53, no 17, p. 4683.
- [5] LIAO, Ping-Sung, et al. A fast algorithm for multilevel thresholding. *J. Inf. Sci. Eng.*, 2001, vol. 17, no 5, p. 713-727.
- [6] LAURENT, Remy, et al. A morphing technique applied to lung motions in radiotherapy: preliminary results. *Acta Polytechnica*, 2010, vol. 50, no 6.
- [7] JI, Dinghuang, et al. Deep view morphing. En *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017. p. 2155-2163..
- [8] JIN, Shi, et al. Learning to dodge a bullet: Conccyclic view morphing via deep learning. En *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018. p. 218-233.
- [9] GONG, Maoguo, et al. A coupling translation network for change detection in heterogeneous images. *International Journal of Remote Sensing*, 2019, vol. 40, no 9, p. 3647-3672.
- [10] DIAZ-PINTO, Andres, et al. Predicting myocardial infarction through retinal scans and minimal personal information. *Nature Machine Intelligence*, 2022, vol. 4, no 1, p. 55-61.

Propuesta de modelo basado en redes neuronales para la clasificación de regiones del esqueleto a partir de imágenes TC

E. Milara¹, P. Sarandeses^{2,3}, A. Gómez-Grande^{2,3}, E.J. Gómez^{1,4}, P. Sánchez-González^{1,4}

¹ Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España; {eva.milara.hernando, enriquejavier.gomez, p.sanchez}@upm.es

² Servicio de Medicina Nuclear, Hospital Universitario 12 de Octubre, Madrid, España; {mariadel Pilar.sarandeses, adolfo.gomez}@salud.madrid.org

³ Facultad de Medicina, Universidad Complutense de Madrid, Madrid, España

⁴ Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, España

Resumen

Existen múltiples contextos clínicos en los que la segmentación y diferenciación de las regiones del esqueleto pueden resultar de gran utilidad para observar la diferente afectación de los huesos en función de su localización. Actualmente, existen algunos estudios que han desarrollado algoritmos basados en árboles de segmentación jerárquica con este fin. Sin embargo, no existen estudios que apliquen técnicas de Machine Learning para lograr esta clasificación. Por este motivo, en el presente estudio se propone un modelo 2D basado en redes neuronales para la clasificación del esqueleto en 8 regiones anatómicas a partir de imágenes de Tomografía Computarizada. Para ello, se ha realizado una previa clasificación manual de 64 volúmenes de esqueleto en estas etiquetas, utilizando 54 para entrenamiento y 10 para test. Posteriormente, se ha evaluado la base de datos con 32 modelos 2D de diferente complejidad, variando el número de capas ocultas (N) y el número de neuronas en estas capas (U). Se observa como un aumento de la complejidad no siempre se corresponde con una mejora del rendimiento, a pesar de que si se requiere cierta complejidad para alcanzar resultados satisfactorios. Por otro lado, se observa como un modelo con un mayor N requiere de un mayor valor de U para conseguir una mejor clasificación. Finalmente, se ha concluido que el modelo de mejor rendimiento es aquel formado por 4 capas ocultas y 200 neuronas por capa, alcanzando un índice Dice general de 0,863, y Dice mayor a 0,7 para todas las regiones.

1. Introducción

La segmentación del esqueleto a partir de imágenes anatómicas resulta de gran utilidad en multitud de contextos clínicos. Un claro ejemplo reside en el uso de esta segmentación para la detección de los cambios morfológicos que sufre el hueso cuando se ve afectado por metástasis, especialmente en pacientes de cáncer de mama o próstata [1][2]. Además, una buena segmentación también podría resultar beneficiosa para la evaluación de fracturas o debilidad ósea en pacientes tratados con nuevas terapias oncológicas (como terapia con radioligandos) o pacientes con osteoporosis [3]–[5]. Estos cambios afectan de forma distinta al hueso dependiendo de la región en la que se encuentre. Por este motivo, resulta de gran interés poder distinguir las diferentes regiones que componen el esqueleto completo.

Actualmente, existen diversos estudios que proponen modelos para segmentar individualmente cada vértebra [6], [7]. Sin embargo, existen solo algunos estudios que desarrollan una metodología dirigida a la diferenciación de las regiones del esqueleto. Fu et al. [8] implementa un algoritmo de diferenciación de los huesos individuales a través de un atlas del esqueleto completo con un índice Dice de 0,9 probado en 5 pacientes. Por otro lado, Bieth et al. [9] propone una clasificación similar a Fu et al. con menor rendimiento con un Dice de 0,774 a 0,856, pero probada en 3 conjuntos de pacientes con un total de 70 imágenes diferentes. Ambos estudios basan su metodología en árboles de segmentación jerárquica.

En cuanto al uso de técnicas de *Machine* o *Deep Learning* para conseguir una clasificación de las diferentes regiones del esqueleto, el estudio de Pradhan et al. [10] consigue diferenciar las radiografías en función de a que región del cuerpo pertenecen. A pesar de su alto rendimiento, esta aplicación no permite diferenciar varias regiones en una misma imagen, limitándose a clasificar las imágenes de forma general.

El objetivo del presente trabajo es desarrollar un modelo para imágenes 2D de Tomografía Computarizada (TC) basado en redes neuronales para la clasificación de la máscara del esqueleto en 8 regiones anatómicas distintas: fémur derecho, fémur izquierdo, pelvis, columna vertebral, costillas, esternón y clavículas, húmero y hombro derechos y húmero y hombro izquierdos.

2. Metodología

En este estudio se utiliza un total de 64 imágenes de TC, recogidas en el Servicio de Medicina Nuclear del Hospital Universitario 12 de Octubre de Madrid, de pacientes de mieloma múltiple y de pacientes de linfoma. A partir de estas imágenes se realiza la segmentación de la máscara del esqueleto con la herramienta desarrollada en [11], generando un conjunto de máscaras recortadas desde el fémur hasta el comienzo del cráneo.

Posteriormente, para el desarrollo de los modelos basados se utiliza el entorno de lenguaje de desarrollo Python 3.9 (Python Software Foundation, EEUU).

2.1. Clasificación por regiones

Una vez se tiene la segmentación del esqueleto, se realiza una clasificación manual de las diferentes regiones evaluando cada corte 2D de todas las imágenes en una herramienta auxiliar desarrollada en MATLAB 2021a y mostrada en la *Figura 1*. Las regiones se numeran de la siguiente manera: (0) fondo, (1) fémur derecho, (2) fémur izquierdo, (3) pelvis, (4) columna vertebral, (5) costillas, (6) esternón y clavículas, (7) húmero y hombro izquierdo, y (8) húmero y hombro derecho.



Figura 1. Interfaz de aplicación para clasificación manual de regiones del esqueleto.

Una vez se tienen todos los cortes de todas las imágenes clasificados, se separan 54 pacientes para el grupo de entrenamiento y 10 para el grupo de test. El grupo de entrenamiento se ha aleatorizado tanto en número de corte como en ID del paciente para evitar el sobreajuste del modelo a regiones o pacientes concretos.

2.2. Arquitectura de modelos basados en redes neuronales

Para la clasificación de las diferentes regiones, se han generado múltiples modelos 2D. En primer lugar, se ha realizado un etiquetado de los elementos no nulos presentes en un corte del volumen del esqueleto como se muestra en la *Figura 2*. Una vez diferenciados estos elementos, se ha utilizado la imagen binaria en la que los píxeles a 1 son aquellos pertenecientes a uno de los elementos no nulos (por ejemplo, el número 1) y el resto de los píxeles se encuentran a 0, como entrada 2D del modelo desarrollado. Este proceso se repite con todas las etiquetas pertenecientes a los elementos no nulos de un mismo corte y, una vez acabado el proceso, se repite con el siguiente corte. El fondo de la imagen, es decir, la ausencia de hueso también aparece como un elemento no nulo a clasificar.

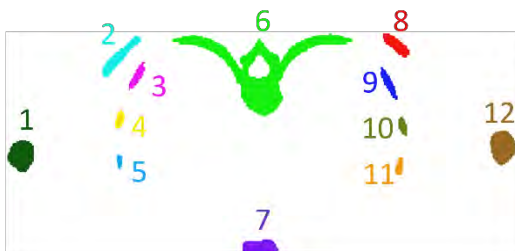


Figura 2. Etiquetado de elementos no nulos presentes en un corte.

Con el objetivo de explorar la aportación de una menor o mayor complejidad de los modelos a una mejor o peor clasificación, se ha probado con dos arquitecturas de red distintas, ambas siguiendo el esquema mostrado en la *Figura 3*.

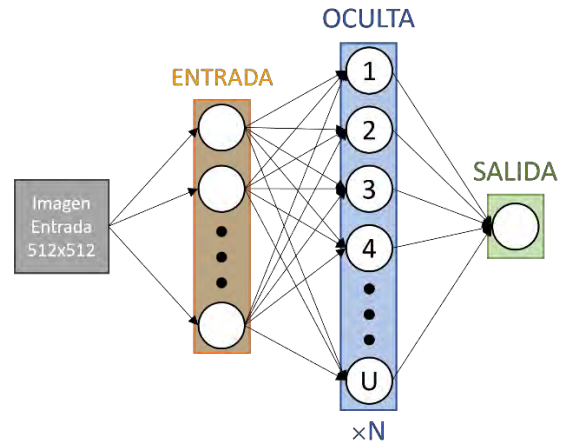


Figura 3. Arquitectura general de los modelos de clasificación desarrollados. *N*: número de capas densas ocultas, *U*: número de unidades o neuronas que forman la capa densa oculta.

Los modelos están formados por un número distinto de capas ocultas (distinto *N*) con activación de tipo ReLU (*Rectified Linear Unit*) entre la capa de entrada de tipo *Flatten* y la capa de salida de tipo densa con activación de tipo *softmax*:

- **Modelo 1:** formado por una capa densa entre la de entrada y salida ($N=1$).
- **Modelo 2:** formado por dos capas densas entre la de entrada y salida ($N=2$).
- **Modelo 4:** formado por cuatro capas densas entre la de entrada y la de salida ($N=4$).
- **Modelo 8:** formado por cuatro capas densas entre la de entrada y la de salida ($N=8$).

Además, se ha variado el valor del parámetro unidades o número de neuronas por capa representado como *U* (5, 10, 20, 50, 75, 100, 150 y 200) de las capas ocultas de ambos modelos para variar la dimensionalidad del espacio de salida, es decir, para generar modelos de mayor y menor complejidad.

2.3. Evaluación de los modelos

Para la evaluación de la correcta clasificación de los elementos no nulos que forman la imagen, los píxeles se clasifican para cada una de las etiquetas, es decir, si hay 8 etiquetas, se evalúan 8 veces. Por ejemplo, para el caso de la etiqueta número 1 (fémur derecho), los píxeles se diferencian en:

- **Verdaderos positivos (VP):** píxeles con la etiqueta real y predicha igual a 1.
- **Verdaderos negativos (VN):** píxeles con la etiqueta real y predicha ambas distintas de 1.
- **Falsos positivos (FP):** píxeles con la etiqueta real distinta de 1, y la predicha igual a 1.
- **Falsos negativos (FN):** píxeles con la etiqueta real igual a 1, y la predicha distinta de 1.

Para evaluar el rendimiento de los modelos desarrollados, se utiliza la métrica del índice Dice para cada etiqueta, calculada mediante la ecuación (1):

$$Dice = \frac{2 \times VP}{2 \times VP + FN + FP} \quad (1)$$

Finalmente, se calcula el valor de Dice promedio de todas las etiquetas correspondientes a regiones del esqueleto, es decir, excluyendo únicamente la etiqueta 0 correspondiente al fondo, así como el promedio incluyendo la etiqueta correspondiente al fondo.

3. Resultados y discusión

3.1. Variación de número de neuronas

En primer lugar, se evalúa el efecto de la variación del valor de U en las diferentes arquitecturas de red, es decir, para distinto valor de N. En la *Tabla 1* se muestra como, por lo general, un aumento del valor de U genera una mejor clasificación (mayor valor del índice Dice) para el Modelo 4 y 8, mientras que para los Modelos 1 y 2 alcanza el mejor rendimiento para U=100. Se puede discernir entonces, que un mayor valor de U no implica un mejor rendimiento del modelo. Finalmente, se puede determinar que del conjunto de modelos pertenecientes a Modelo 1 y 2, el modelo de mejor rendimiento es aquel con U=100, mientras que de los grupos de Modelo 4 y 8 es aquel de U=200.

También cabe destacar que los modelos del grupo Modelo 1, no alcanzan un rendimiento mejor a los modelos de los grupos restantes para ningún valor de U. Por este motivo, se determina que al menos se requiere un mínimo de complejidad para obtener una clasificación óptima de las regiones. Por otro lado, el Modelo 8 solo alcanza un valor mayor de Dice para U=75, por lo que, de nuevo, una mayor complejidad no implica un mejor rendimiento.

U	Mod. 1	Mod. 2	Mod. 4	Mod. 8
5	0,585	0,751	0,693	NC
10	0,814	0,782	0,827	0,751
20	0,832	0,840	0,836	0,816
50	0,847	0,850	0,813	0,833
75	0,840	0,843	0,841	0,849
100	0,849	0,855	0,857	0,830
150	0,839	0,852	0,846	0,844
200	0,844	0,847	0,863	0,857

Tabla 1. Índice Dice promedio de los modelos variando el valor de U. Marcado en negrita el valor máximo para cada valor de U (filas). Mod: modelo. NC: No converge.

3.2. Variación de capas ocultas

En cuanto a la variación del parámetro N, en la *Tabla 2* se muestra el resultado de los valores Dice para cada etiqueta para el valor de U de mejor rendimiento en cada uno de los modelos, U=100 para Modelo 1 y 2 y U=200 para Modelo 4 y 8. Se puede observar que para la mayoría de las etiquetas (4 de 8) el Modelo 4 obtiene los mejores resultados, comparado con el Modelo 2 (3 de 8), el Modelo 1 (0 de 8) y el Modelo 8 (2 de 8), siendo estas cuatro restantes no muy superiores a las obtenidas con el Modelo 4. Cabe destacar que, a pesar de que tanto para el Modelo

4 como para el Modelo 8 no hay ninguna etiqueta que obtenga un Dice menor de 0,7, el Dice promedio del Modelo 4 es superior al del resto de modelos, tanto si se incluye el fondo como si no.

	Modelo			
	1	2	4	8
(0) Fondo	1,000	1,000	1,000	1,000
(1) Fémur der.	0,884	0,906	0,916	0,891
(2) Fémur izq.	0,852	0,844	0,894	0,868
(3) Pelvis	0,912	0,918	0,932	0,932
(4) Columna vertebral	0,965	0,970	0,977	0,972
(5) Costillas	0,854	0,864	0,861	0,861
(6) Esternón+clavículas	0,702	0,675	0,704	0,715
(7) Húmero+hombro izq.	0,826	0,855	0,823	0,821
(8) Húmero+hombro der.	0,800	0,811	0,794	0,798
Promedio sin fondo	0,849	0,855	0,863	0,857
Promedio con fondo	0,866	0,871	0,878	0,873

Tabla 2. Valores Dice del rendimiento de los modelos en función de las etiquetas. der: derecho(s), izq: izquierdo(s). Marcado en negrita el valor máximo para cada etiqueta o promedio (por fila).

3.3. Selección y evaluación del modelo de mejor rendimiento

Teniendo en cuenta que es el modelo que mejores resultados obtiene tanto de promedio como para cada etiqueta, se considera que el modelo óptimo para la clasificación de regiones es aquel con U=200 y N=4 (Modelo 4). Para este modelo, se obtiene la matriz de confusión mostrada en la *Figura 4*, con el objetivo de evaluar cuales son las regiones que se clasifican de forma errónea de forma más habitual. Se puede observar, que las confusiones más habituales son las de la etiqueta 6 (esternón y clavículas) con las etiquetas 3 (pelvis) y 5 (costillas), así como la de la etiqueta 8 (húmero y hombro derecho) con la 1 (fémur derecho), y la 7 (húmero y hombro izquierdo) con la 5 y la 2 (fémur izquierdo).

Matriz de confusión (%píxeles)

	0	1	2	3	4	5	6	7	8
0	100	0	0	0	0	0	0	0	0
1	0	91.7	0	0	0	0	0	0	10.3
2	0	0	95.5	1.1	0	1	0.4	9.2	0
3	0	2.6	0	93.1	1.7	2.5	17.3	2.9	2.8
4	0	0	0	1.6	98.2	0.7	3.8	0.1	0
5	0	1.1	1.9	0.6	0.1	83.6	6.7	8.7	3.3
6	0	1.8	2.1	1.7	0	2.6	71.7	1.9	1.4
7	0	0	0.5	0.4	0	4.1	0	77.3	0
8	0	2.8	0	1.5	0	5.6	0.1	0	82.2

Etiqueta Predicha

Etiqueta Real

Figura 4. Matriz de confusión del modelo con mejor rendimiento. *Los colores han sido escalados, tonos más oscuros en los valores mayores a 0 y menores de 20, para mejorar la visualización y diferenciación de los porcentajes de clasificación errónea.

Aquellas clasificaciones incorrectas que confunden húmero y fémur son comprensibles debido a su forma y localización similares. Por otro lado, la clasificación de la etiqueta 6 como 3, puede producirse por la zona más alta del esternón, donde se produce su unión a las clavículas, ya que supone un área de mayor tamaño, similar al de la pelvis.

A pesar de que el modelo propuesto obtiene peor rendimiento que el de Fu et al. [8] se pretende obtener un modelo con menor sobreajuste al haber sido entrenado y evaluado con un número mayor de pacientes. Por otro lado, este modelo es comparable al de Bieth et al. [9] ya que obtiene un índice Dice similar, aunque ligeramente superior. Sin embargo, el modelo propuesto se ha probado única y exclusivamente imágenes de pacientes en riesgo de padecer alguna afectación medular, mientras que el de Bieth et al. se prueba tanto en dos contextos clínicos distintos, como en sujetos sanos.

En el contexto de la aplicación de técnicas de *Deep Learning*, se propone como posible continuación del estudio probar con modelos de mayor complejidad, es decir, con modelos con un mayor número de capas ocultas o mayor valor de U para N=4 o N=8, e incluso modelos con diferentes tipos de capas, no solo densas con activación de tipo ReLu.

4. Conclusiones

Actualmente, la evaluación ósea en enfermedades que implican cambios morfológicos interesa no únicamente a nivel global del esqueleto completo, sino también a nivel individual de cada hueso o conjunto de ellos para conocer el patrón de afectación. En el presente trabajo se desarrolla un modelo de clasificación automática de las diferentes regiones del esqueleto basado en redes neuronales. Se prueba con modelos de diferente complejidad, seleccionando el de mejor rendimiento para su posterior uso y aplicación. Para este modelo, se obtiene un valor de Dice general de 0,863 excluyendo el fondo y 0,878 incluyendo el fondo.

Agradecimientos

Este trabajo está parcialmente financiado por una beca predoctoral (ayuda del Programa Propio de I+D+i 2020) de la Universidad Politécnica de Madrid. Este trabajo se ha desarrollado dentro del marco de la comunidad *EELISA Health in the City*.

Referencias

- [1] M. Fujino, K. Suzuki, M. Nishio, N. Nishiyama, and Y. Osaka, “Strategy of radiation therapy for bone metastases and MSCC in breast cancer patients,” *Breast Cancer*, vol. 18, no. 4, pp. 238–243, Oct. 2011, doi: 10.1007/s12282-011-0288-z.
- [2] T. Bäuerle and W. Semmler, “Imaging response to systemic therapy for bone metastases,” *Eur Radiol*, vol. 19, no. 10, pp. 2495–2507, Oct. 2009, doi: 10.1007/s00330-009-1443-1.
- [3] B. Grubmüller et al., “Response assessment using 68Ga-PSMA ligand PET in patients undergoing 177Lu-PSMA radioligand therapy for metastatic castration-resistant prostate cancer,” *Eur J Nucl Med Mol Imaging*, vol. 46, no. 5, pp. 1063–1072, May 2019, doi: 10.1007/s00259-018-4236-4.
- [4] M. M. Heck et al., “Systemic Radioligand Therapy with [177] Lu Labeled Prostate Specific Membrane Antigen Ligand for Imaging and Therapy in Patients with Metastatic Castration Resistant Prostate Cancer,” *Journal of Urology*, vol. 196, no. 2, pp. 382–391, Aug. 2016, doi: 10.1016/j.juro.2016.02.2969.
- [5] H. K. Genant, K. Engelke, and S. Prevrhal, “Advanced CT bone imaging in osteoporosis,” *Rheumatology*, vol. 47, no. Supplement 4, pp. iv9–iv16, Jul. 2008, doi: 10.1093/rheumatology/ken180.
- [6] N. Lessmann, B. van Ginneken, P. A. de Jong, and I. Išgum, “Iterative fully convolutional neural networks for automatic vertebra segmentation and identification,” *Med Image Anal*, vol. 53, pp. 142–155, Apr. 2019, doi: 10.1016/j.media.2019.02.005.
- [7] S. Ruiz-España et al., “Automatic segmentation of the spine by means of a probabilistic atlas with a special focus on ribs suppression,” *Med Phys*, vol. 44, no. 9, pp. 4695–4707, Sep. 2017, doi: 10.1002/mp.12431.
- [8] Y. Fu, S. Liu, H. H. Li, and D. Yang, “Automatic and hierarchical segmentation of the human skeleton in CT images,” *Phys Med Biol*, vol. 62, no. 7, pp. 2812–2833, Apr. 2017, doi: 10.1088/1361-6560/aa6055.
- [9] M. Bieth et al., “Segmentation of Skeleton and Organs in Whole-Body CT Images via Iterative Trilateration,” *IEEE Trans Med Imaging*, vol. 36, no. 11, pp. 2276–2286, Nov. 2017, doi: 10.1109/TMI.2017.2720261.
- [10] N. Pradhan, V. Singh Dhaka, and H. Chaudhary, “Classification of Human Bones Using Deep Convolutional Neural Network,” *IOP Conf Ser Mater Sci Eng*, vol. 594, no. 1, p. 012024, Aug. 2019, doi: 10.1088/1757-899X/594/1/012024.
- [11] E. Milara et al., “Bone marrow segmentation and radiomics analysis of [18F]FDG PET/CT images for measurable residual disease assessment in multiple myeloma,” *Comput Methods Programs Biomed*, vol. 225, p. 107083, 2022, doi: 10.1016/J.CMPB.2022.107083.

Herramienta para la detección de los márgenes de resección en cirugía conservadora de cáncer de mama

M. Calvo de Mora¹, E. Milara¹, E.J. Gómez^{1,2}, M. Chiva de Agustín³, P. Sánchez-González^{1,2}

¹ Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España; m.calvodemora@alumnos.upm.es, {eva.milara.hernando, enriquejavier.gomez, p.sanchez}@upm.es

² Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, España

³ Sección de Radiología, Hospital Universitario Ramón y Cajal, Madrid, España; miguel.chiva@salud.madrid.org

Resumen

El cáncer de mama constituye un problema de salud pública, siendo el tipo de cáncer más frecuente en las mujeres en todo el mundo. Si se diagnostica en una etapa temprana, siendo pequeño y localizado, se realiza una cirugía conservadora como tratamiento. Previo a la cirugía, se puede realizar la implantación de una semilla prequirúrgica mediante aguja guiada por imagen con el fin de ayudar al cirujano a identificar la ubicación exacta de la lesión. Durante dicha intervención quirúrgica, se extirpa el tumor junto con un margen de tejido normal circundante y se realiza una evaluación radiológica intraoperatoria para valoración de los márgenes quirúrgicos de la pieza extirpada y determinar si es necesaria una resección adicional. El procedimiento actual demanda una participación activa del radiólogo, lo que supone una interrupción de su labor asistencial diaria. El presente trabajo tiene por objetivo diseñar y desarrollar una herramienta de apoyo a la decisión, que permita garantizar de forma automática la completa extracción de la lesión, el marcador y la semilla prequirúrgica, así como la distancia a los bordes quirúrgicos.

1. Introducción

La mama es un órgano que se extienden desde la segunda hasta la sexta costilla, medialmente hasta el esternón y lateralmente hasta la línea media axilar. Anatómicamente, la mama está formada principalmente por tejido adiposo (grasa) y tejido mamario (glándula mamaria) [1]. El cáncer de mama es la neoplasia maligna más frecuentemente diagnosticada en mujeres en todo el mundo, especialmente en los países occidentales [2]. Se estima que cada año se diagnostican 2,3 millones de nuevos casos de cáncer de mama en todo el mundo. Además, también es la principal causa de muerte por cáncer en las mujeres de todo el mundo. A nivel mundial fue responsable de 684.996 muertes en 2020 [3].

Actualmente, las técnicas utilizadas para el diagnóstico del cáncer de mama incluyen autoexploración, exploración clínica, pruebas de imagen, punción de lesiones y biopsias [4]. Entre las técnicas de imagen más utilizadas, destacan la mamografía y la ecografía, siendo la primera más utilizada en pacientes mayores de 50 años, sintomáticas o de alto riesgo [5], [6].

En tumores pequeños o aquellos tumores grandes que tras tratamiento con quimioterapia reducen su tamaño, se opta por una cirugía conservadora. Para ello, previamente se realiza la implantación de una semilla y un marcador prequirúrgicos mediante aguja guiada por imagen

radiológica con el fin de ayudar al cirujano a identificar la ubicación exacta de la lesión para su extirpación en quirófano. Durante la cirugía conservadora, se extirpa el tumor junto con un margen de tejido normal circundante. Además, se realiza una evaluación intraoperatoria de los márgenes quirúrgicos de la pieza extirpada para determinar si es necesaria una resección adicional. Esta evaluación se realiza mediante una mamografía y/o ecografía a la pieza extraída en quirófano, la cual se compara con la mamografía prequirúrgica para asegurar la correcta extracción de la lesión y la presencia de bordes libres de tumor [7]. El margen quirúrgico se define en la práctica clínica como la distancia entre las células tumorales y el borde de la pieza extirpada, siendo: positivo cuando la distancia es menor o igual a 2 mm; negativo pero próximo cuando se encuentra entre 2 y 5 mm; y negativo cuando es mayor de 5 mm [8]. Este control radiológico del espécimen quirúrgico supone una interrupción de la labor asistencial del experto, siendo un proceso tedioso. Además, la medida de los márgenes de la lesión a los bordes de la pieza se realiza de forma manual, estando sujeta a la cierta subjetividad.

El objetivo del presente estudio es desarrollar una herramienta capaz de segmentar y garantizar de manera precisa y automática la completa extirpación de la lesión, el marcador y la semilla prequirúrgicos. Además, esta herramienta pretende estimar la distancia a los bordes quirúrgicos, mediante la localización de dichos elementos en la imagen intraoperatoria comparada con la mamografía prequirúrgica en casos de cirugía conservadora de cáncer de mama.

2. Metodología

Se realiza un estudio observacional analítico a partir de una cohorte de 43 pacientes sometidas a cirugía conservadora de cáncer de mama a las que se les implantó exclusivamente dos marcadores metálicos: uno de tipo *coil*, con forma helicoidal, y otro de tipo semilla, con forma cilíndrica y alargada. Por cada paciente, se utilizan dos imágenes bidimensionales: una mamografía prequirúrgica y una imagen mamográfica de la pieza quirúrgica. Todas las imágenes provienen del mamógrafo *Siemens MAMMOMAT Revelation* y son recogidas en el Servicio de Radiodiagnóstico del Hospital Universitario Ramón y Cajal.

Para la implementación completa de la herramienta, tanto su interfaz como su algoritmo de segmentación de la lesión

y medida de márgenes quirúrgicos se utiliza el entorno de lenguaje de desarrollo Python 3.9 (Python Software Foundation, EEUU).

2.1. Algoritmo para estimación de márgenes quirúrgicos

2.1.1. Cálculo de la transformación de la imagen

Con el objetivo de conseguir el mismo posicionamiento de la imagen quirúrgica de la pieza y la imagen mamográfica prequirúrgica, se localiza el *coil* en ambas imágenes y se calcula las transformaciones necesarias para desplazar y reorientar la imagen de la pieza.

Para la localización del *coil*, se aplica una función de umbralización utilizando como umbral el máximo de la imagen menos un valor de 150 niveles de intensidad de gris, ya que al ser el marcador un elemento metálico tiene el valor de máxima intensidad de la imagen. De esta forma, se mantienen en la imagen exclusivamente los marcadores (semilla y *coil*), así como algunas calcificaciones mamarias. Posteriormente, se localiza la semilla (mayor área) y su respectivo centro, y luego el elemento cuyo centro sea más cercano (medido en distancia euclidiana) al de la semilla, localizando así el *coil*. Para ello, se presupone que ambos elementos son los más cercanos, ya que ambos se implantan en la lesión.

Una vez se tiene localizado el *coil* en ambas imágenes (mamografía y pieza quirúrgica) y sabiendo que las dimensiones y orientación de ambas imágenes difieren, se calcula el redimensionamiento de la imagen de la pieza para que el *coil* tenga un tamaño y posición aproximados en ambas imágenes. Para ello, como muestra la *Figura 1*, se calcula la relación de tamaño del *coil* en ambas imágenes (*Rsize*), el desplazamiento en las direcciones horizontal y vertical (*dX* y *dY*, respectivamente) y el ángulo que forman el *coil* en ambas imágenes, es decir, el ángulo de rotación necesario para conseguir la misma orientación (*Rot*).

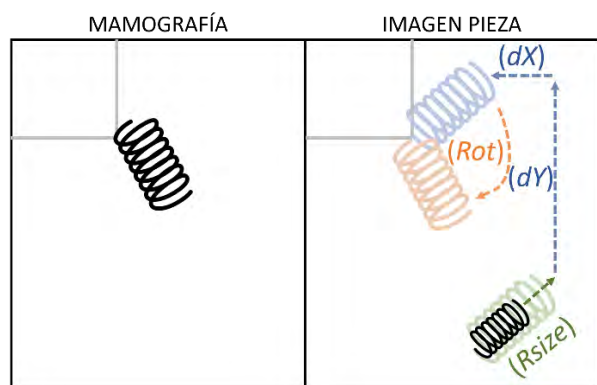


Figura 1. Esquema de las transformaciones necesarias para posicionar el coil de la imagen de la pieza lo más similar posible al coil de la mamografía. Negro: coils originales; verde: adaptación de tamaño; azul: desplazamiento horizontal y vertical; naranja: rotación.

2.1.2. Detección de bordes de la pieza y la lesión

Partiendo de la imagen original de la imagen quirúrgica de la pieza, se realiza una umbralización para conseguir la segmentación de la pieza completa. Además, también se

realiza una segmentación de la lesión mediante un algoritmo de crecimiento de regiones, cuya semilla se define como los píxeles más cercanos al *coil*. Concretamente, se aplican tres crecimientos de regiones variando el umbral de nivel de intensidad de gris (de más a menos estricto). El algoritmo incluye los nuevos píxeles si la media de ellos es mayor que la media de la semilla inicial menos el umbral y menor que la media de la semilla inicial más el umbral. El criterio de parada se alcanza en la iteración que no se añaden píxeles. Una vez se obtienen las áreas de los tres procesos, el algoritmo se queda con aquel resultado de menor área que cumpla las siguientes condiciones: tener un área menor al 70% del de la pieza original y mayor que el de la semilla inicial.

Finalmente, tanto para la imagen resultante de la segmentación de la pieza como la de la lesión, se aplica el algoritmo Canny para la localización de sus bordes. Una vez localizados los bordes de la pieza quirúrgica y la lesión, se aplica sobre la imagen de bordes las transformaciones geométricas calculadas anteriormente que permiten reposicionar la imagen de la pieza quirúrgica de la misma manera que la de la mamografía prequirúrgica. De esta forma, se consigue visualizar los bordes obtenidos de la imagen de la pieza (una vez transformados) sobre la imagen de la mamografía prequirúrgica.

2.1.3. Cálculo de márgenes quirúrgicos

Una vez obtenidos los bordes de la pieza y la lesión, y antes de la transformación calculada en el apartado 2.1.1. , se realiza la medida de los márgenes quirúrgicos. En primer lugar, se identifica si la mama estudiada representa la mama derecha o la izquierda mediante un análisis del perfil de intensidad de la imagen original de la mamografía con el objetivo de proporcionar las referencias adecuadas.

Posteriormente, se estima el punto central de la lesión a partir del punto medio en las direcciones horizontal y vertical. A partir de este punto, se trazan dos líneas, una horizontal y otra vertical, y se calculan los puntos de intersección de ambas líneas con el borde detectado para la lesión (X_{L1} , X_{L2} , Y_{L1} e Y_{L2}) y de la pieza (X_{P1} , X_{P2} , Y_{P1} e Y_{P2}), tal y como se muestra en la *Figura 2*.

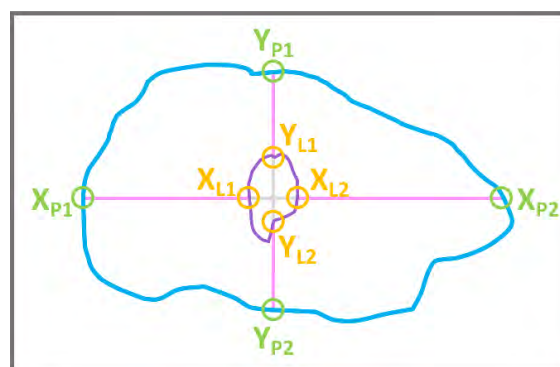


Figura 2. Esquema de obtención de los márgenes quirúrgicos. Azul: borde de la pieza; morado: borde de la lesión; rosa: márgenes quirúrgicos; gris: ejes que definen el punto medio de la lesión; naranja y verde: puntos de intersección de los ejes horizontal (*X*) y vertical (*Y*) con el borde de la lesión y pieza respectivamente.

Obtenidos los puntos de intersección, se realiza una resta para conocer el número de píxeles que hay entre el borde de la lesión y de la pieza en las cuatro direcciones y se multiplica por el tamaño del píxel en mm (*PixelSpacing*), siguiendo las ecuaciones (1), (2), (3) y (4):

$$\text{Margen}_{\text{posterior}} = (Y_{L1} - Y_{P1}) \times \text{PixelSpacing} \quad (1)$$

$$\text{Margen}_{\text{anterior}} = (Y_{P2} - Y_{L2}) \times \text{PixelSpacing} \quad (2)$$

$$\text{Margen}_{\text{exterior}} = (X_{L1} - X_{P1}) \times \text{PixelSpacing} \quad (3)$$

$$\text{Margen}_{\text{interior}} = (X_{P2} - X_{L2}) \times \text{PixelSpacing} \quad (4)$$

2.2. Diseño de la herramienta

Todo el algoritmo explicado con anterioridad se ejecuta a través del uso de una herramienta llamada *RX Margin Tester*. Al iniciar la herramienta, aparece durante unos segundos una ventana con su logo. Una vez inicializada, el usuario puede seleccionar las imágenes que desea procesar a través del explorador de archivos. Posteriormente, el usuario puede visualizar las imágenes importadas en pantalla y seleccionar el botón *Siguiente* en caso de que esté seguro de que las imágenes han sido seleccionadas correctamente. Mientras se realiza el procesamiento de las imágenes, se muestra una pantalla de espera al usuario, a través de la cual se indica un mensaje del procesamiento que se está realizando. Si no se localiza correctamente la lesión, aparece un mensaje de información en el que se indica que los márgenes se miden con respecto al borde del *coil*.

Finalmente, una vez completado todo el procesamiento, se muestra una ventana con los resultados visuales y cuantitativos obtenidos. Por un lado, se muestran los bordes de la pieza quirúrgica y la lesión superpuestos en la mamografía y en la imagen de la pieza quirúrgica. Por otro lado, se muestran los valores numéricos de los márgenes quirúrgicos calculados.

3. Resultados y discusión

Para ejemplificar el resultado de la metodología aplicada, se muestran las imágenes resultantes de cada paso de la aplicación del algoritmo en una misma paciente.

3.1. Cálculo de la transformación de la imagen

En la *Figura 3*, se muestra el resultado de localizar y superponer el *coil* de ambas imágenes, mamografía prequirúrgica y radiografía quirúrgica de la pieza, tras aplicar la función de transformación en la segunda imagen. Como se puede observar, la gran mayoría de píxeles coinciden para ambas imágenes.

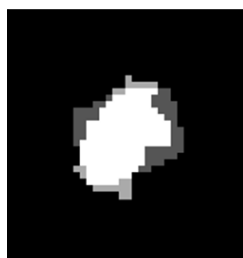


Figura 3. Superposición de las imágenes binarias de los coils; gris oscuro: coil de mamografía prequirúrgica; gris

claro: coil de la radiografía quirúrgica de la pieza tras las transformaciones para corregir su posicionamiento; blanco: píxeles coincidentes de ambos coils. La imagen ha sido ampliada al borde los píxeles no nulos más 10 píxeles para mejorar su visualización e interpretación.

Aunque la umbralización da resultados satisfactorios para la mayoría de los casos, es incapaz de eliminar el contacto directo entre ambos marcadores, situación que sucede en 7 casos. En ciertos casos, existe la presencia de una microcalcificación muy cercana a la semilla, incluso más cerca que el *coil*, provocando que sea identificada por el algoritmo como dicho *coil* y se logre un mal redimensionamiento, desplazamiento y superposición con la mamografía original. Es necesario considerar que dichos casos son muy escasos y la mayoría de las veces, ambos marcadores se encuentran muy próximos y la presencia de microcalcificaciones no afecta a la localización del *coil*.

3.2. Detección de bordes de la pieza y la lesión

En la *Figura 4* se muestra el resultado del paso correspondiente a la detección de los bordes de la pieza y la lesión a partir de la radiografía quirúrgica de la pieza, en azul y morado respectivamente. Además, se muestra en rosa la representación de los márgenes quirúrgicos medidos y cuantitativamente medidos en el siguiente paso.

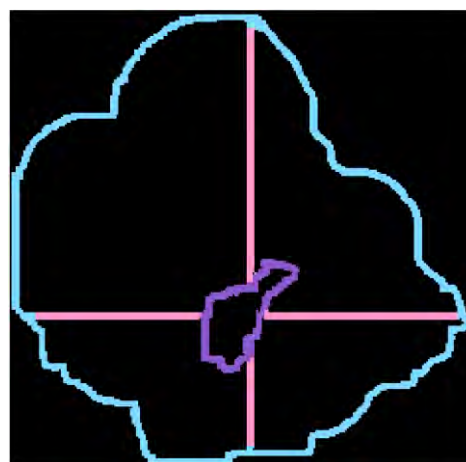


Figura 4. Detección de bordes. Azul: bordes de la pieza; morado: bordes de la lesión; rosa: márgenes quirúrgicos. La imagen ha sido ampliada al límite del borde de la pieza para mejorar su visualización e interpretación.

El algoritmo tampoco muestra un funcionamiento adecuado en el caso de pacientes con alta densidad mamaria (8 casos), ya que no consigue una detección precisa del área de la pieza, impidiendo así la obtención correcta de sus bordes. En la *Figura 5*, se muestra el ejemplo en el que la segmentación de la mama genera una imagen de muchas áreas pequeñas dispersas en lugar de un área grande que recoja todos los puntos de la pieza debido a su alta densidad mamaria y, en consecuencia, la gran diferencia en los niveles de intensidad de gris con el cambio de densidad.



Figura 5. Ejemplo de caso en el que la segmentación de la pieza no es adecuada debido a su alta densidad mamaria.

3.3. Implementación de la herramienta

En la *Figura 6* se muestra la pantalla resultante tras la aplicación de la metodología completa en el caso de la paciente usada para el ejemplo. Se puede observar como se muestra en rojo los bordes de la lesión y de la pieza sobre ambas imágenes, mamografía prequirúrgica en la izquierda y radiografía quirúrgica de la pieza en la izquierda. Además, se muestra una columna con los márgenes quirúrgicos medidos entre ambos bordes. Para ambos casos, un acercamiento de la imagen ha sido aplicado.

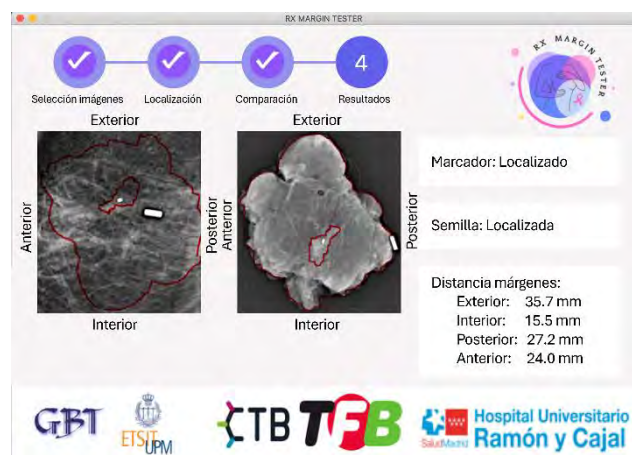


Figura 6. Interfaz gráfica de la herramienta desarrollada con la pantalla resultante al procesar un caso.

El profesional implicado en el desarrollo de la herramienta probó y valoró positivamente la sencillez, la estética, la usabilidad y la visualización de resultados. Además, teniendo en cuenta el proceso completo del procesamiento usado en la herramienta, resalta su rapidez, a pesar de ser menos veloz que los propios especialistas. Por ello, se considera un instrumento útil para el radiólogo profesional para evitar la interrupción asistencial diaria que supone la evaluación manual de los márgenes, proporcionando una herramienta totalmente automatizada.

4. Conclusiones

En el presente trabajo se propone una herramienta de segmentación de lesión en pacientes de cáncer de mama tratadas mediante cirugía conservadora, y de cálculo de los correspondientes márgenes quirúrgicos para garantizar de manera precisa y automática la completa extirpación de la lesión, el marcador y la semilla prequirúrgica. La herramienta se considera un sistema de ayuda a la decisión clínica, automatizando un proceso desarrollado

actualmente de forma manual por los radiólogos profesionales.

Agradecimientos

Este trabajo está parcialmente financiado por una beca predoctoral (ayuda del Programa Propio de I+D+i 2020) de la Universidad Politécnica de Madrid. Este trabajo se ha desarrollado dentro del marco de la comunidad *EELISA Health in the City*.

Referencias

- [1] H. Ellis and V. Mahadevan, “Anatomy and physiology of the breast,” *Surgery (Oxford)*, vol. 31, no. 1, pp. 11–14, Jan. 2013, doi: 10.1016/j.mpsur.2012.10.018.
- [2] J. Zhang, H. Zhou, S. Zheng, J. Liu, and Y. Han, “Automated breast ultrasound system (ABUS) for breast cancer screening and diagnosis: The preliminary results,” *Ultrasonics*, vol. 70, pp. 200–204, Sep. 2016, doi: 10.1016/j.ultras.2016.05.012.
- [3] S. Łukasiewicz, M. Czezelewski, A. Forma, J. Baj, R. Sitarz, and A. Stanisławek, “Breast Cancer—Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies—An Updated Review,” *Cancers (Basel)*, vol. 13, no. 17, p. 4287, Aug. 2021, doi: 10.3390/cancers13174287.
- [4] E. Leon-Rodriguez, C. Molina-Calzada, M. M. Rivera-Franco, and A. Campos-Castro, “Breast self-exam and patient interval associate with advanced breast cancer and treatment delay in Mexican women,” *Clinical and Translational Oncology*, vol. 19, no. 10, pp. 1276–1282, Oct. 2017, doi: 10.1007/s12094-017-1666-6.
- [5] R. Sardiñas Ponce, “Autoexamen de mama: un importante instrumento de prevención del cáncer de mama en atención primaria de salud,” 2009.
- [6] E. Pastor-Ausín, J. A. López-Ruiz, M. Ruiz-Delgado, Z. Fernández-Temprano, I. Aguirre-Olóriz, and J. J. Echevarría-Uraga, “Cáncer de Mama: el papel del Radiólogo,” *Gaceta Médica de Bilbao*, vol. 104, no. 3, pp. 120–126, Jan. 2007, doi: 10.1016/S0304-4858(07)74588-2.
- [7] M. Martín, A. Herrero, and I. Echavarría, “El cáncer de mama,” *Arbor*, vol. 191, p. 773, 2015.
- [8] I. T. Rubio and V. Marco, “La importancia de los márgenes quirúrgicos en la cirugía conservadora en el cáncer de mama,” *Cir Esp*, vol. 79, no. 1, pp. 3–9, Jan. 2006, doi: 10.1016/S0009-739X(06)70807-7.

Beyond DTI: Insights in Heart Infarcted Tissue

Justino R. Rodríguez Galván¹, Susana Merino Caviedes¹, David Filgueiras Rama², Javier Sánchez González³, Antonio Tristán Vega¹, and Carlos Alberola López¹

¹Laboratorio de Procesado de Imagen (LPI), Universidad de Valladolid.

²Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid.

³Philips Ibérica, Madrid.

Abstract

Diffusion magnetic resonance imaging (dMRI) is an advanced clinical imaging technique based upon characterizing the water flow within body tissues. Although models of higher order (HO) than diffusion tensor imaging (DTI) have been successfully applied to brain dMRI, the peculiarities of the heart hinder its application on this muscle. However, these models have been proven as a powerful way to get insights in the microstructure of organs. In this work we have applied both DTI and a HO model to two ex-vivo pig heart acquisitions for measuring their differences in a metric that we call “Non-Gaussianity” and then check how these differences are gathered on infarcted tissue. Further researching may be needed to both confirm this correlation and make these models feasible to clinical practice.

1 Introduction

Magnetic resonance imaging (MRI) is a clinical imaging modality based on the measurement of the alignment of spins in the nuclei of body tissues after the application of magnetic fields. MRI is considered the gold standard for cardiac imaging because, moreover than giving images of the best quality and resolution, it is, on the one hand, a non-ionizing procedure and, on the other hand, the chemical contrast used when needed is gadolinium, which is innocuous for the great majority of patients. A fundamental in MRI is the way images are acquired. Instead of being taken in the image domain, they are acquired on the so-called k -space, which is a translation of spatial coordinates into frequencies through the Fourier transform. Cardiac modality of MRI (CMR) is based on the acquisition of a sequence of all cardiac phases called “cine”. These cine acquisitions are extremely difficult to achieve, since an acquisition of the heart at every stage is needed, making the patient to hold their breath several times in order to stabilize the torso movement for the readouts. Notice that complexity arises from the dependence on synchronization with beating rate and the uncomfortableness of the patient, who must repeat this exhausting process plenty of times.

For alleviating these exams, several techniques has been developed, some of them consist in the application of a mathematical model for reconstruction, e.g. SENSE, GRAPPA, k - t sampling, compressed sensing [1–4]. Others are based on the application of deep learning [5] or making 3D breath free acquisitions, the so-called 5D MRI [6].

Diffusion MRI (dMRI) is a fashion of MRI focused on characterize the Brownian motion of water within body tissues. For describing dMRI, q -space should be presented beforehand. As k -space, it is relatd to image domain through Fourier transform an it is also where sampling is made by spreading direction-encoding

gradients over it. At every place where a gradient is placed, a diffusion weighted image (DWI) [7] is retrieved. The distribution of those gradients along the q -space is layered. Beginning at the center, where at least one is placed, each of them become part of concentric spherical surfaces, the so-called shells, and thus, the insight provided comes by two ways: the position of the gradient and the distance to the center, determined by the radius, the b -value, which is expressed in s/mm^2 . Those non-diffusion images taken at the center receive the name of baselines. From here on out, we can also differentiate two types of acquisitions depending on the number of shells used; if they use one, they are monoshell acquisitions, if more than one are used, they are multishell acquisitions.

Using DWIs as basis, models are computed. The simplest model is the 3D gaussian isotropic, which is fully described by a single parameter: the apparent diffusion coefficient (ADC) [7]. Diffusion tensor imaging (DTI) would take a step forward in complexity by using a tensor made of a 3×3 semidefinite matrix, growing the number of degrees of freedom (DoF) to 6 [7]. Until this point, the presented models only need one shell to be defined—monoshell samplings—but for higher order (HO) models than DTI, such as DSI (diffusion spectrum imaging), a multishell approach must be followed.

Next step in the process is the computation of the relevant functions. Essentially, they reduce to two: the ensemble average propagator (EAP) and the attenuation $E(q)$. Since the former is closely related to the chosen representation model, only a definition of its Fourier counterpart, $E(q)$ will be displayed.

$$E(q) = \frac{S(q)}{S(0)}, \quad (1)$$

where $S(q)$ is the signal measured at position q and, $E(q)$, or attenuation, measures the degradation of signal as we move away from the center.

Despite its simplicity, DTI may get some insight needing only 6 DWI acquisitions at the cost of losing resolution when fibers cross. For getting the full information, HO models shall be chosen, at the cost of much longer examinations.

Notice that for fully understanding the microstructure of an organ, knowing the behavior of fibers at intersect points is essential. Neurological dMRI has successfully applied HO models, being capable of characterizing structural differences for distinct diseases as Alzheimer or migraine [8, 9]. Although, the application of dMRI in the cardiac counterpart is adding a layer of complexity at the top of an extremely entangled structure, having to discern the diffusion motion from the bulk motion, which is

much bigger.

Furthermore, there are plenty of difficulties to overcome, for instance, the T_2 transversal relaxation rate is approximately half the brain, 40 ms, having a much shorter echo time (T_E). Another factor to take into account is the fact that regular sequences as PGSE (pulse gradient spin echo) are not suitable for in-vivo studies and most sequences must be programmed ad-hoc [10]. All these handicaps, combined with the fact of much longer acquisition times for HO models, have arisen reluctance in the community about the usefulness of going beyond DTI. Nonetheless, the reward of a successful application would mean a much deeper insight on the organization of the microstructural chambers of the myocardium, in charge of conducting electricity and motion, the cardiomyocytes, which are supposed to align in laminar structures called sheetlets. Another major achievement is the fact that with these kind of HO models, gadolinium contrast could become unnecessary.

Advances in dMRI in the heart are, almost at its entirety, made using DTI as their representation model. Since the publication of [11], where moreover than an explanation of every step in the process (acquisition sequences, relevant metrics, applications), a shade of concern is put over the utility of HO models. Thus, the efforts in diffusion research are mainly concentrated on continuing the existent lines on DTI instead of opening new ones in HO models, thereby this work should be taken as a proof-of-concept.

In this work, a first approach was made for determining how apart are DTI and HO models in the myocardium, applying both representation models and then, taking advantage of the gaussian nature of DTI, computing the Non-Gaussianity (NG) of the difference between models, as metric of differentiation. Results suggest that higher values of this metric are placed on infarcted tissue.

The rest of the article is organized as follows. Section 2 describes the materials that have been used. Section 3 is focused on the proposed methodology. After those explanations, section 4 is used for both showing the obtained results at subsection 4.1 and their discussion at subsection 4.2. All in all, section 5 will of the limitations of these experiments.

2 Materials

Two dense pig heart ex-vivo acquisition with the sampling scheme shown on Table 1 were used. The acquisitions were made on a 3T Philips Achieva Scanner at the Centro Nacional de Investigaciones Cardiovasculares (CNIC). The dimensions of the images taken for both volumes are 128×128 , having 130 slices for the first volume and 123 for the second. For both volumes, spatial resolution is $1.09156 \times 1.09156 \times 1.1$ mm per voxel. Echo time was set to 0.113725 seconds. For the sake of completeness, both T_1 and T_2 acquisitions were made.

b-value	Number of directions
300	8
600	16
900	32
1200	32
2000	64
3600	64

Table 1. Sampling scheme for each volume that has been used. Right column values determine the number of gradients placed at the radius shown on left column.

3 Methods

The purpose of this work is to determine how apart are DTI and HO models. In this case, there is only one HO model put into

test: DSI.

Our means to check this difference lies on the definition of DTI. DTI models the fibers in q -space with 6 DoF which correspond to the unique elements of the correlation matrix of a trivariate gaussian distribution. As DTI is a sort of gaussian distribution, hence the way for measuring the distance between approaches is determined by the amount of non-gaussianity.

As the relevant area of the heart for this work is the myocardium, a mask for each volume was made manually, then, the procedure followed the scheme in fig 1.

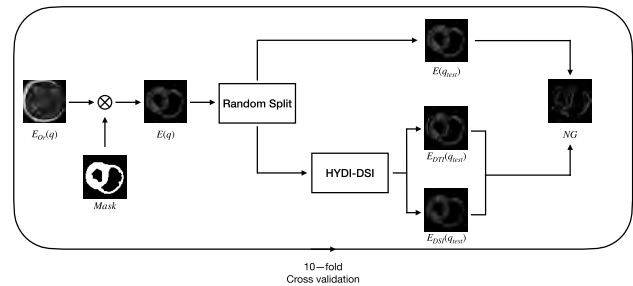


Figure 1. Pipeline for methods applied. First of all, mask is applied to each volume for only focusing on the myocardium area. After this masking procedure, a 10-fold cross validation scheme is put into action. In either of these folds, 3 directions per shell are randomly selected and reserved for an ulterior comparison. With the rest of the directions, both DTI and DSI models are computed by means of the voxel-wise general cross validation HYDI-DSI procedure [12], from which the EAP is obtained and thereby, synthetic DWIs at the same position as the reserved gradients are computed.

Considering the limited amount of data at our disposition, a bootstrap method was implemented. To do so, a 10-fold cross validation strategy was designed. For each iteration, 3 directions per shell are randomly selected by creating a hull sampling scheme of 6 shells and 3 samples per shell using the method described in [13], applying a random rotation and then selecting the DWIs made with the nearest direction-encoding gradients to the ones created.

The selected DWIs will be considered from now on as a test set. The meaning of this consideration is, as in a Deep Learning approach, that they will not intervene in the computation of the representation models, though they will be predicted by them. DTI and DSI models EAP will be computed by means of a voxel-wise general cross validation HYDI-DSI procedure [12], using the remainder directions and then making synthesized DWIs at the positions of the test set gradients.

Using both real and synthetic DWIs, Mean Square Errors for each DTI and DSI model are computed in order to compute the NG metric, defined as

$$NG = \frac{\overline{(E(q_{test}) - E_{DTI}(q_{test}))^2) - \overline{(E(q_{test}) - E_{DSI}(q_{test}))^2}}{\overline{E(q_{test})^2}}, \quad (2)$$

where both $E_{DTI}(q_{test})$ and $E_{DSI}(q_{test})$ represent the synthetic attenuation at the positions of the test set gradients and $E(q_{test})$ the real counterpart. It is assumed that by subtracting both MSEs, noise contribution can be neglected. NG index will be grouped then by shell, taking each 3 directions per shell for each iteration and putting altogether and then computing an average map per shell.

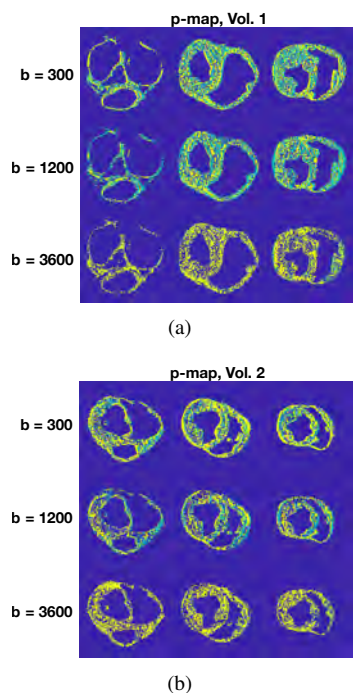


Figure 2. Wilcoxon test results for three different slices across b -values. Yellow parts show the areas where DSI error is smaller than DTI error, green parts show the vice versa areas. Blue values represent the area where the null hypothesis is not rejected for $p < 0.05$. (a) depicts the results on volume 1, (b) depicts the results on volume 2.

4 Results and discussion

4.1 Results

Obtained results are shown on the subsequent figures. Figure 2 depicts the results for a Wilcoxon test for a slice. Figures 3 and 4 display the average NG maps for certain b -values and a 3D render for the 3600 shell average NG map for each volume.

4.2 Discussion

All presented results are in the form of maps because of the great difference between the subtraction of DTI minus DSI errors from one voxel to another due to the difference between areas or boundary errors. To hurdle this set of circumstances, a Wilcoxon test was performed over the difference of MSEs. The results shown on figure 2 suggest that DSI and DTI has significant differences between them, having DSI a smaller error than DTI, and as we move further from the center, increasing the b -value, DSI gets better results.

Figure 3 details how NG increases as b -value gets bigger. Figure 4 shows a 3D render for NG computed with the data of the shell of 3600 s/mm^2 . Two main points can be taken from these figures: first, NG has continuity among slices and, second, higher values are aligned with squeezed part of the myocardium wall, sign of infarcted tissue, which may lead to new insight in this area of research by means of HO methods.

5 Conclusion

The study of HO models in cardiac tissue seems to be relevant and offers new applicable insight for the study of different pathologies. However, as the acquisition conditions are not suitable to be used on clinical practice, the application of this methodology is still uncertain for the moment.

6 Acknowledgements

The authors acknowledge grants PID2020-115339RB-I00, TED2021-130090B-I00 and PID2021-124407NB-I00 from the Ministerio de Ciencia e Innovación of Spain. The support from ESAOTE Ltd. is also acknowledged.

References

- [1] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, y P. Boesiger, "SENSE: sensitivity encoding for fast mri," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 42, no. 5, pp. 952–962, 1999.
- [2] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, y A. Haase, "Generalized autocalibrating partially parallel acquisitions (grappa)," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 47, no. 6, pp. 1202–1210, 2002.
- [3] M. Lustig, J. M. Santos, D. L. Donoho, y J. M. Pauly, "kt sparse: High frame rate dynamic mri exploiting spatio-temporal sparsity," en *Proceedings of the 13th annual meeting of ISMRM, Seattle*, vol. 2420. Citeseer, 2006.
- [4] M. Lustig, D. L. Donoho, J. M. Santos, y J. M. Pauly, "Compressed sensing mri," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 72–82, 2008.
- [5] E. Martín-González, T. Sevilla, A. Revilla-Orodea, P. Casaseca-de-la Higuera, y C. Alberola-López, "Group-wise non-rigid registration with deep learning: an affordable solution applied to 2d cardiac cine mri reconstruction," *Entropy*, vol. 22, no. 6, p. 687, 2020.
- [6] L. Feng, S. Coppo, D. Piccini, J. Yerly, R. P. Lim, P. G. Masci, M. Stuber, D. K. Sodickson, y R. Otazo, "5d whole-heart sparse mri," *Magnetic resonance in medicine*, vol. 79, no. 2, pp. 826–838, 2018.
- [7] P. Hagmann, L. Jonasson, P. Maeder, J.-P. Thiran, V. J. Wedeen, y R. Meuli, "Understanding diffusion mr imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond," *Radiographics*, vol. 26, no. suppl_1, pp. S205–S223, 2006.
- [8] J. R. Harrison, S. Bhatia, Z. X. Tan, A. Mirza-Davies, H. Benkert, C. M. Tax, y D. K. Jones, "Imaging alzheimer's genetic risk using diffusion mri: A systematic review," *NeuroImage: Clinical*, vol. 27, p. 102359, 2020.
- [9] Á. Planchuelo-Gómez, D. García-Azorín, Á. L. Guerrero, S. Aja-Fernández, M. Rodríguez, y R. de Luis-García, "White matter changes in chronic and episodic migraine: a diffusion tensor imaging study," *The journal of headache and pain*, vol. 21, pp. 1–15, 2020.
- [10] S. NIELLES-Vallespin, A. Scott, P. Ferreira, Z. Khalique, D. Pennell, y D. Firmin, "Cardiac diffusion: technique and practical applications," *Journal of Magnetic Resonance Imaging*, vol. 52, no. 2, pp. 348–368, 2020.
- [11] C. Mekkaoui, T. G. Reese, M. P. Jackowski, H. Bhat, y D. E. Sosnovik, "Diffusion MRI in the heart," *NMR in Biomedicine*, vol. 30, no. 3, p. e3426, 2017.
- [12] A. Tristán-Vega, T. Pieciak, G. París, J. R. Rodríguez-Galván, y S. Aja-Fernández, "Hydi-dsi revisited: Constrained non-parametric eap imaging without q-space re-gridding," *Medical Image Analysis*, vol. 84, p. 102728, 2023.
- [13] J. R. Rodríguez-Galván, G. París, A. Tristán-Vega, y C. Alberola-López, "Diffusion sampling schemes: A generalized methodology with nongeometric criteria," *Magnetic Resonance in Medicine*, vol. 89, no. 6, pp. 2270–2280, 2023.

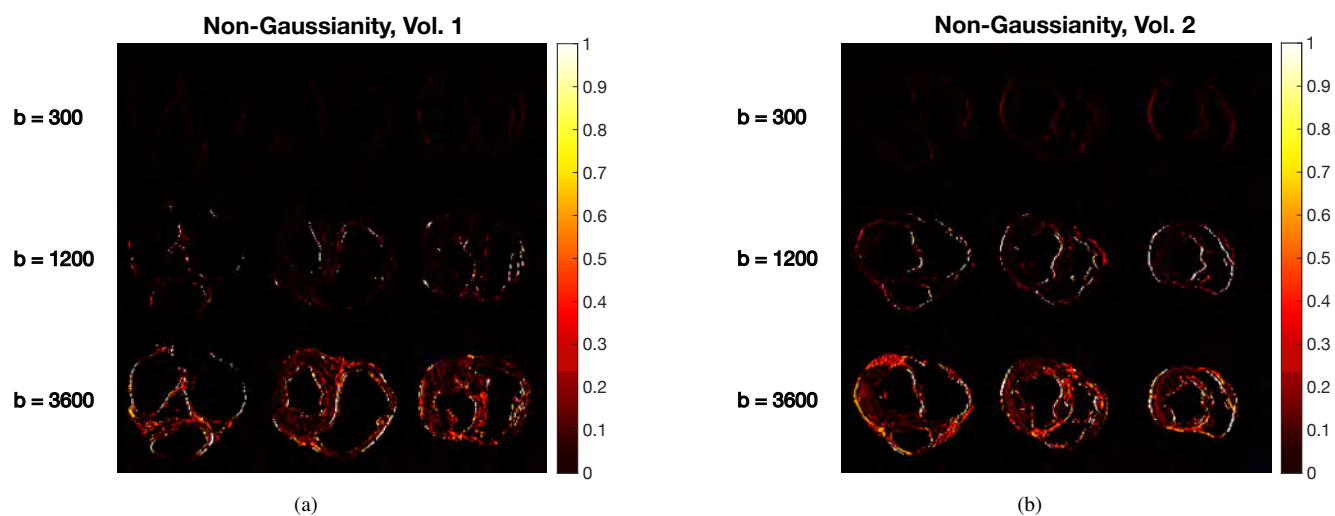


Figure 3. NG maps for three different slices across b -values. (a) depicts the behavior of this metric on volume 1, (b) depicts the behavior of this metric on volume 2.

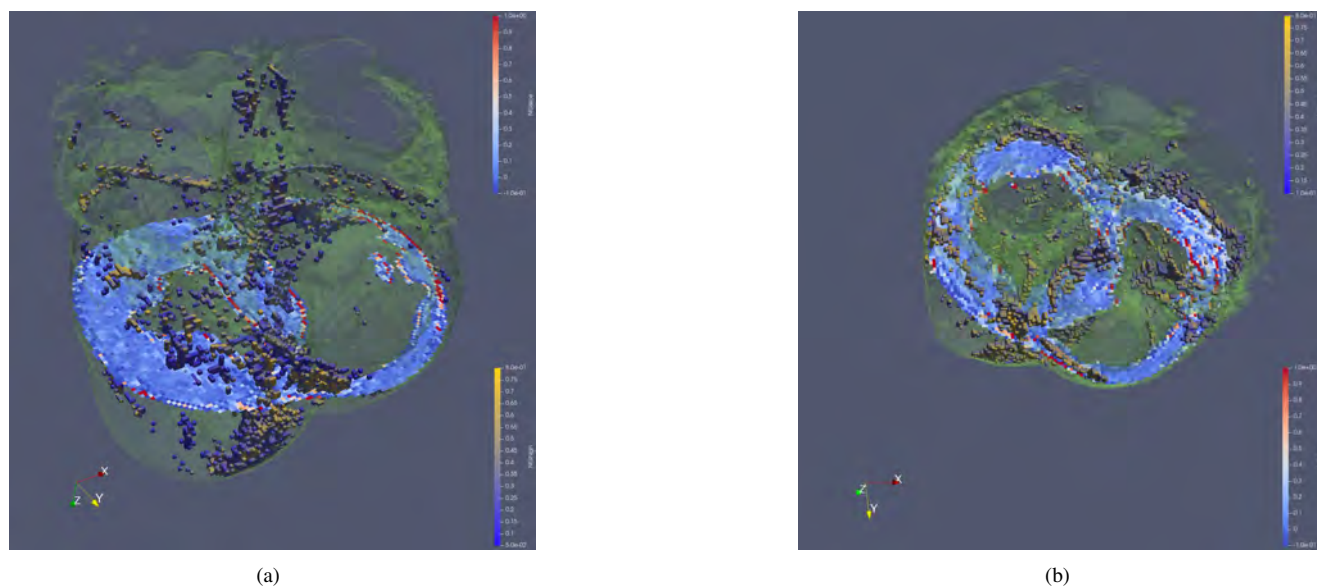


Figure 4. 3D renders for NG maps for each volume. In green is shown the render of the mask, light blue is used for give sight of a slice. Cubes throughout the volume show the amount of NG, from blue to yellow; different colors on the slice show the amount of NG, from blue to red. (a) is the 3D render of vol. 1, (b) is the 3D render of vol. 2.

Uso de cuestionarios digitales en el aula y realización de seminarios en la unidad de cuidados intensivos como innovaciones docentes en Ingeniería Biomédica

*Virginia Reverte^{1,2}, *José Francisco Garrido-Peñalver³, María Teresa Llinás^{1,2}, José Manuel Allegue³, Joaquín Roca-González⁴, María Trinidad Herrero^{2,5}

¹ Departamento de Fisiología, Universidad de Murcia, Murcia, España, vrr2@um.es, mayte@um.es

² Instituto Murciano de Investigación Biomédica (IMIB) Pascual Parrilla, Murcia, España

³ Servicio Medicina Intensiva, Hospital General Universitario Santa Lucía, Cartagena, España, josegarrido.uci@gmail.com, josem.allegue@carm.es

⁴ Departamento de Automática, Ingeniería Eléctrica y Tecnología electrónica, Universidad Politécnica de Cartagena, Cartagena, España, jroca.gonzalez@upct.es

⁵ Clinical & Experimental Neuroscience (NICE), Universidad de Murcia, Murcia, España, mtherrero@um.es

*Misma contribución, primeros autores

Resumen

La Ingeniería Biomédica proporciona una herramienta esencial para la formación de nuevos profesionales capaces de diseñar y gestionar la próxima generación de dispositivos médicos. La interacción de los futuros ingenieros e ingenieras y los profesionales de la medicina es importante para lograr un correcto aprendizaje de los alumnos y asegurar la adecuada atención a los pacientes. Los problemas más frecuentes detectados por los profesores y profesoras del Grado han sido la baja asistencia y la falta de interés de los alumnos y alumnas por no encontrar un enfoque práctico a los conocimientos adquiridos. Con el objetivo de mejorar esta situación, se introdujeron dos nuevas herramientas docentes en el módulo de Fisiología del Grado de Ingeniería Biomédica: la resolución de cuestionarios digitales en las clases y la realización de seminarios docentes en la unidad de cuidados intensivos del Hospital General Universitario Santa Lucía. La realización de los cuestionarios mejoró significativamente los resultados académicos de aquellos alumnos y alumnas que completaron como mínimo el 50% de los mismos, así como la asistencia a clase, alcanzándose un 73% del total de alumnos matriculados. La eficacia del seminario se evaluó mediante encuestas anónimas y opiniones de texto libre. Todos los/as estudiantes encuestados se mostraron satisfechos (6%) o muy satisfechos (94%) con esta actividad. En resumen, el uso de estas nuevas metodologías aumentó las calificaciones, la asistencia a clase y la motivación de los alumnos y alumnas, mejorando la adquisición de conocimientos, el desarrollo de habilidades comunicativas y su capacidad crítica.

1. Introducción

La llegada a la medicina de modernas técnicas como la robótica o la inteligencia artificial requiere de nuevos profesionales capaces de diseñar y gestionar la próxima generación de dispositivos médicos. El grado en Ingeniería Biomédica (GIB) recientemente implantado en la Escuela Técnica Superior de Ingenieros Industriales (ETSII) de la Universidad Politécnica de Cartagena (UPCT), proporciona la formación idónea a través de una sólida base en ingeniería y electrónica industrial que se complementa con contenidos específicos en biología,

anatomía y fisiología humana y procesamiento de bioseñales [1] entre otras materias.

Actualmente la enseñanza universitaria apuesta por un formato híbrido que combina las clases presenciales con las tecnologías de la información y la comunicación (TICs) ya que la presencialidad enriquece el proceso de aprendizaje del alumno y le ayuda a socializar [2]. Los principales problemas a los que nos enfrentamos los profesores y profesoras son la baja asistencia a clase y la falta de atención e interés por parte del alumnado. En este trabajo proponemos una herramienta innovación docente con la que, mediante el uso de las TICs, se amplifique la acción de los/as docentes y el aprendizaje de los/as estudiantes. Se utilizaron cuestionarios instantáneos de respuesta múltiple (IM-CQ) como herramienta que permite realizar evaluaciones en entornos digitales y conocer los resultados al instante.

Por otro lado, entre lo que más preocupa a los/as estudiantes, desmotivándolos hasta el punto de abandonar sus estudios, es no ver la aplicación práctica de los conocimientos que están adquiriendo [3]. En este sentido, los ingenieros e ingenieras biomédicos (IB) son los encargados de diseñar y desarrollar los dispositivos utilizados en el ámbito médico. Por tanto, es necesaria una estrecha relación entre los/as IB y los/as profesionales sanitarios para optimizar las innovaciones tecnológicas que permitan una mejor atención al paciente. Con el fin de establecer un flujo de conocimientos entre médicos e IB ideamos la segunda metodología de innovación docente: la realización de un Seminario Docente (TW) en la unidad de cuidados intensivos (ICU) del Hospital General Universitario Santa Lucía sobre los sistemas tecnológicos utilizados en dicho servicio tutorizado por un médico especialista en medicina intensiva.

Con la introducción de estas herramientas docentes pretendemos aumentar la presencia de alumnos y alumnas en las aulas, su grado de implicación y motivación, así como mostrarles la aplicación en pacientes reales de los

conocimientos adquiridos durante su formación académica. Finalmente, todo ello va a aumentar la capacidad de los IB para conseguir futuros avances tecnológicos que mejoren el bienestar de las personas.

2. Material y métodos

Las herramientas de innovación docente descritas en este estudio se introdujeron en un grupo de 33 estudiantes universitarios de 2º curso matriculados en la asignatura de Anatomía y Fisiología Humanas del segundo curso de GIB. La participación fue voluntaria. Además, los alumnos que asistieron a la ICU-TW firmaron un formulario de confidencialidad para mantener la privacidad de los pacientes. Las gráficas y los análisis estadísticos se realizaron con GraphPad Prism 6. Las diferencias entre los grupos se examinaron mediante un t-test no pareado de doble vía. La significación estadística se definió como $P < 0.05$.

2.1. Cuestionarios instantáneos de respuesta múltiple (IM-CQ)

Para la realización de los cuestionarios, se eligió la herramienta digital Socrative [4]. Se trata de una herramienta gratuita que permite realizar evaluaciones en entornos digitales, que se puede utilizar desde la web o descargarla en un ordenador, tableta o smartphone y que facilita enormemente la labor docente en el ámbito del proceso de evaluación, ya que permite generar cuestionarios de forma sencilla y atractiva para los alumnos. Además, la inmediatez a la hora de recibir los resultados de las pruebas supone una importante reducción de tiempo en el trabajo de corrección por parte del docente. Para empezar a utilizar Socrative, el profesor tiene que crear un “Room” para la clase, pudiendo incluir hasta 50 alumnos. Después, los alumnos tendrán que acceder a ella haciendo clic en “Student login”: sólo necesitarán el nombre de la “Room” que el profesor habrá creado previamente. Una vez registrados, podrán participar en todas las actividades propuestas por el profesor hasta que cierren la sesión. Entre las diferentes actividades disponibles en la aplicación seleccionamos “Quiz” y el tipo de respuestas “multiple choice”. Utilizando esta herramienta digital, en los últimos 15 minutos de cada una de las 11 sesiones teóricas, se realizaron los cuestionarios de preguntas tipo test directamente relacionadas con el contenido de la sesión teórica impartida. Estos cuestionarios, constaban de entre 10- y 17 preguntas cada uno, se cumplimentaron exclusivamente en clase y suponían el 10% de la nota final de la asignatura. Para evaluar la influencia del IM-CQ en la asistencia y aprendizaje de los alumnos y alumnas y en sus calificaciones, se calculó el porcentaje de asistencia y se compararon las puntuaciones obtenidas en el examen de Fisiología de los/as estudiantes que completaron más de la mitad de los cuestionarios ($>50\%$) con las de los/as estudiantes que completaron menos de la mitad de los cuestionarios ($<50\%$). Los datos se expresaron como media \pm SE.

2.2. Seminario docente en la unidad de cuidados intensivos (ICU-TW)

Esta metodología docente se llevó a cabo en la ICU del Hospital General Universitario Santa Lucía (Cartagena, Murcia, España). Se eligió esta unidad como óptima para el perfil de los/as IB porque requiere de una tecnología muy avanzada y eficaz debido a la gravedad del estado de los/as pacientes allí ingresados [5].

Los médicos del servicio diseñaron un nuevo TW en el que, divididos en grupos de seis, los/as estudiantes conocieron el funcionamiento de diferentes dispositivos y procesos: monitorización de las constantes básicas del paciente, monitorización hemodinámica invasiva/no invasiva, seguimiento de la analgesia y sedación de pacientes críticos, asistencia respiratoria invasiva/no invasiva y terapias eléctricas como marcapasos y desfibriladores. Se centraron en sus características tecnológicas, su aplicación en los pacientes ingresados y en las principales limitaciones técnicas que ofrecían actualmente.

Tras finalizar la sesión, se pidió a los estudiantes que, mediante una encuesta de calidad, utilizando una escala Likert de cinco puntos, valoraran de manera anónima su aprendizaje y comprensión clínica durante la ICU-TW y que calificaran su satisfacción con el seminario. La encuesta también incluía comentarios de los alumnos y alumnas en formato de texto libre que nos permitió obtener información más detallada sobre su experiencia en esta actividad para introducir posibles mejoras en futuras sesiones. Los datos se expresaron como frecuencias (%). La asistencia y participación en este seminario constituyó un 10% de la nota final de Fisiología.

3. Resultados

3.1. Cuestionarios instantáneos de respuesta múltiple (IM-CQ)

De los 33 alumnos y alumnas matriculados en el curso, 25 cumplimentaron más del 50% y 8 de ellos menos del 50%. Como esperábamos, la puntuación en el examen de Fisiología de los alumnos que completaron al menos la mitad de los IM-CQ fue significativamente superior a la de aquellos que sólo completaron menos de la mitad ($<50\%$) (8.5 ± 0.2 vs. 6.5 ± 0.3 , respectivamente) (Figura 1).

Dado que estos cuestionarios sólo podían cumplimentarse en clase tras la correspondiente sesión teórica, la presencia media en el aula fue de 24 alumnos lo que representó una tasa de asistencia del 73%. La nota obtenida por los alumnos en los distintos cuestionarios fue sobresaliente, con una puntuación media de 9.3 ± 0.1 , siendo 10 la máxima calificación. No hubo diferencias entre los resultados obtenidos a lo largo del curso, con una puntuación de 9.2 ± 0.3 en el primer cuestionario y de 9.1 ± 0.2 en el último.

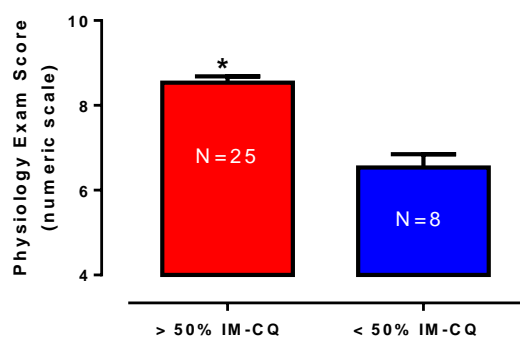


Figura 1. Puntuación del examen de Fisiología en los estudiantes que completaron más (N=25) o menos (N = 8) del 50% de los cuestionarios. Los datos son la media ± SE. * P<0.001 frente a los estudiantes que completaron menos del 50% del IM-CQ.

3.2. Seminario docente en la unidad de cuidados intensivos (ICU-TW)

De los/as 23 estudiantes que participaron en esta actividad de innovación docente, 16 de ellos (7 no respondieron) proporcionaron comentarios sobre la capacidad del ICU-TW para abordar la comprensión clínica y la función de los diferentes dispositivos utilizados en la unidad de críticos (Figura 2A). El 100% de los alumnos se mostró muy de acuerdo con los conocimientos proporcionados sobre los dispositivos. Además, el 81% estaba muy de acuerdo o de acuerdo (62% y 19%, respectivamente), y el 19% se mostró indiferente con la duración del ICU-TW. Por último, el 94% se mostró muy de acuerdo y el 6% de acuerdo con que este seminario tuviera lugar en la unidad de cuidados intensivos porque permitía comprender mejor cómo funcionan los sistemas tecnológicos en los pacientes críticos que las clases teóricas por sí solas. La sección final de la encuesta estaba diseñada para determinar la satisfacción de los/as estudiantes con el taller de enseñanza (Figura 2B). El 100% de los alumnos y alumnas se mostraron muy satisfechos con las habilidades del ponente y la calidad de las presentaciones. Además, se mostraron muy satisfechos con el diseño del seminario (30 min de introducción teórica fuera de la unidad de tratamiento + 150 min de demostración práctica), ya que el 100% de ellos no cambiaría nada al respecto. Por último, el 100% de los/as estudiantes se mostraron muy satisfechos o satisfechos con la utilidad del ICU-TW para su titulación (el 94% se mostró muy satisfecho y el 6% satisfecho).

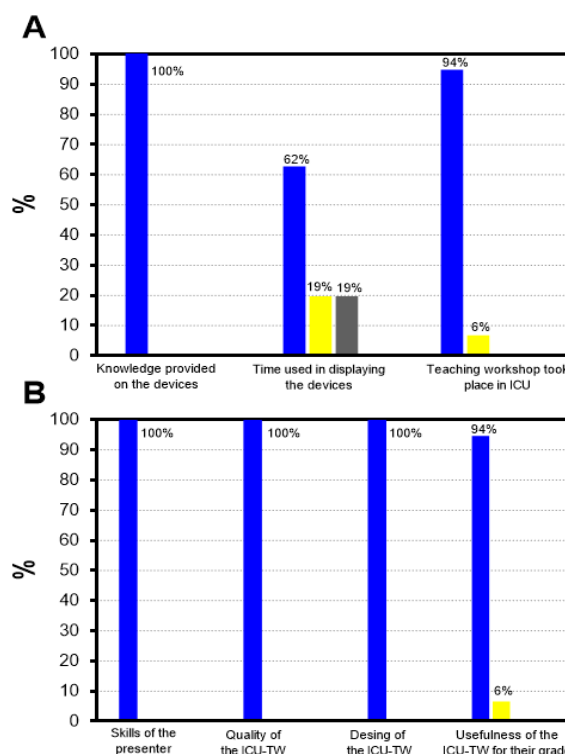


Figura 2. Feedback para recoger las opiniones de los/as estudiantes sobre **A)** Capacidad de la ICU-TW para abordar la comprensión clínica y el funcionamiento de los diferentes dispositivos, utilizando una escala Likert de 5 puntos: totalmente de acuerdo (azul), muy de acuerdo (amarillo), indiferente (gris), en desacuerdo (verde), totalmente en desacuerdo (rojo) (N=16) y **B)** Grado de satisfacción de los/as estudiantes con diferentes aspectos de la ICU-TW utilizando una escala Likert de 5 puntos: totalmente satisfecho (azul), muy satisfecho (amarillo), indiferente (gris), insatisfecho (verde), totalmente insatisfecho (rojo) (N= 16). Los datos se expresaron como frecuencias (%).

Además, se recogieron los comentarios individuales de los alumnos y alumnas en formato de texto libre. Los datos cualitativos proporcionaron una visión más amplia de las sesiones por parte los alumnos. Recibimos 16 respuestas escritas, la mayoría breves pero muy positivas como se muestra en la Tabla 1.

Tabla 1. Opiniones y comentarios en formato de texto libre.
1. La forma en que el médico explicó el funcionamiento de los diferentes aparatos y su profesionalidad.
2. En primer lugar, el médico explicó el funcionamiento de los aparatos y después pudimos ver cómo se utilizan en la vida diaria de un paciente ingresado en medicina intensiva.
3. Ver los diferentes dispositivos funcionando en un paciente real
4. La explicación de cómo funciona un marcapasos.
5. Ver los instrumentos con los que yo voy a trabajar en un futuro.
6. Poder ver en primera persona los diferentes dispositivos y cómo funcionan en un paciente real.
7. Ver los diferentes dispositivos usados en la ICU y descubrir algunos que no conocía.
8. Ver la aplicación práctica de los conocimientos que estoy adquiriendo en el GIB.

9. Me gustó mucho ver un dispositivo siendo utilizado por un estudiante.
10. La explicación teórica sobre los diferentes dispositivos y cómo funcionan en un paciente real.
11. Cómo se puede monitorizar el dolor de una persona a través de la forma de la onda marcada en un dispositivo.
12. Ver individualmente cómo es el equipamiento utilizado en un box de la ICU y cómo se utiliza en un paciente.
13. Las explicaciones del doctor y ver los diferentes dispositivos funcionando en un paciente real.
14. Ver la aplicación de los diferentes equipos médicos en un paciente real y conocer las condiciones y gastos que conlleva un ingreso en una ICU.
15. Ver los diferentes aparatos con los que trabajaremos en el futuro como ingenieros biomédicos.
16. Ver cómo funcionan los diferentes dispositivos, que espero mejorar en el futuro.

Tabla 1. Opiniones y comentarios de lo/as estudiantes sobre lo que más les gustó del ICU-TW recogidos mediante encuestas de satisfacción a través de formato de texto libre.

4. Discusión

La primera parte de este estudio demostró la utilidad de las TICs en la educación universitaria para mejorar los niveles de atención, participación y compromiso del alumnado [6]. Esta idea se vio respaldada por el hecho de que los alumnos y alumnas que completaron más de la mitad de los IM-CQ propuestos obtuvieron calificaciones significativamente más altas en el examen del módulo de Fisiología de la asignatura con respecto a los que no lo hicieron. La asistencia a clase mejoró en comparación con la media de asistencia universitaria (57%), alcanzando el 73% del aforo total [2]. Asistir regularmente a clase permite a los/as estudiantes mejorar su comunicación, aprender a trabajar en equipo y tener una buena convivencia. Este hecho es importante porque al relacionar la asistencia a clase con la tasa de rendimiento y la tasa de eficiencia, se comprobó que, en las titulaciones con menor asistencia a clase, ambas tasas también eran más bajas [7].

La ICU-TW se ideó para implementar la formación de los/as estudiantes de IB con conocimientos clínicos más específicos y establecer un flujo de conocimientos entre médicos/as y futuros ingenieros/as. Los/as estudiantes consideraron que la ICU-TW proporcionaba un entorno de aprendizaje más activo que mejoró significativamente su comprensión fisiológica, y que les aportó conocimientos muy útiles para su desarrollo profesional. De este modo, los alumnos y alumnas pueden ver la aplicación real de los conocimientos que adquirirán durante la titulación [8].

En resumen, afrontar los retos de un sistema sanitario en rápida evolución requiere una formación y unos conocimientos que van más allá de la educación universitaria tradicional. El uso de IM-CQ aumenta la tasa de asistencia a clase y la puntuación obtenida en la asignatura. Además, una mayor inclusión de los/as IB, desde las primeras etapas de su formación, en las filas de la asistencia sanitaria a través de este ICU-TW refuerza la confianza en los conocimientos adquiridos y en sus habilidades. Ambas capacidades son cruciales para establecer el canal de comunicación médico-ingeniero necesario para resolver las limitaciones técnicas actuales y lograr juntos una medicina mejor.

Agradecimientos

Por último, agradecer a los anteriores comités organizadores del CASEIB por su amabilidad al permitirnos usar sus guías de estilo como referencia para la realización de este documento.

5. Referencias

- [1] Página web de la Escuela Técnica Superior de Ingeniería Industrial (ETSII): Grado en Ingeniería Biomédica (5211). <https://estudios.upct.es/grado/5211/plan-de-estudios>. (Consultada: septiembre 2023)
- [2] Vicerrectorado de Calidad, Planificación e Innovación. Universidad de Oviedo. La importancia de asistir a clase en la universidad. 2004.
- [3] Cheng, E. C. K. (2021). Knowledge transfer strategies and practices for higher education institutions. *VINE Journal of Information and Knowledge Management Systems*, vol 51, sup 2, 2021, pp 288-301. DOI: 10.1108/VJKMS-11-2019-0184.
- [4] Balta N, Tzafilkou K. Using Socratic software for instant formative feedback in physics courses. *Education and Information Technologies*, vol 24, sup 1, 2018, pp: 307-23. DOI:10.1007/s10639-018-9773-8
- [5] Halpern NA. Innovative Designs for the Smart ICU. *Chest*, vol 145, sup 3, 2014, pp 646-58. doi: 10.1378/chest.13-0004.
- [6] Jolly A. Role of ICT in Quality Enhancement of Higher Education. *International Journal of Engineering Research and Technology (IJERT)*, vol 08, Issue 10, 2019. DOI: 10.17577/IJERTV8IS100072.
- [7] Martins P, Walker I. University Classes and Student Achievement. *Workshop on Education and Training*. Mannheim. Proceedings, 2005.
- [8] Lizzio A, Wilson K. Self-managed learning groups in higher education: students' perceptions of process and outcomes. *Br J Educ Psychol*, vol 75, 2005, pp 373–90.

Predictores de adherencia a una app de salud en el proyecto europeo PHARAON

A. Martínez Lax¹, L. Carrillo López², M.A. Conesa Carrillo³, C. Martínez Galera⁴, M. Jódar Reverte⁵, G. Sánchez Nanclares⁶, S. Pérez Martos⁷

^{1, 2, 3, 4, 5} Especialistas en Enfermería Familiar y Comunitaria, Proyecto Pharaon. Servicio Murciano de Salud, Murcia, España, enfermerosproyectopharaon@carm.es

⁶ Especialista en Medicina Familiar y Comunitaria. Responsable de Innovación del Servicio Murciano de Salud, Murcia, España, gorka.sanchez@carm.es

⁷ Ingeniero de telecomunicaciones. Universidad politécnica de Cartagena, Murcia, España, salvador.perez@upct.es

Resumen

En el marco de trabajo del proyecto europeo Pharaon se está llevando a cabo un piloto de telemonitorización para 150 pacientes adultos con insuficiencia cardiaca crónica de la Región de Murcia. Consiste en el despliegue de un conjunto de tecnologías integradas en una aplicación móvil (app) para monitorización clínica. El objetivo es determinar los predictores de adherencia en el uso de la app en este grupo de pacientes. Para ello, se ha establecido un indicador de adherencia. Más de la mitad de los usuarios cumplen los requisitos mínimos de usabilidad. Aunque los resultados no tienen significación estadística cabe destacar la influencia del sexo, el nivel de autonomía del usuario y el grado de autocuidado sobre la adherencia. Resulta imprescindible considerar el perfil del usuario para la implementación de las tecnologías en salud en el sistema sanitario.

1. Introducción y objetivo

El proyecto PHARAON, financiado por la Comisión Europea en el marco del programa Horizonte 2020 con el acuerdo de subvención nº 857188, tiene el objetivo de mejorar la vida de nuestros mayores en aspectos sociales o sanitarios a través de las nuevas tecnologías, haciendo realidad un modelo sanitario diferente mediante la creación de un conjunto de plataformas, dispositivos y herramientas que incluye el internet de las cosas (IoT), inteligencia artificial (IA), robótica, computación en la nube, dispositivos portátiles inteligentes, big data e inteligencia analítica, ofreciendo soluciones innovadoras centradas en el usuario.

Existen 6 ecosistemas piloto en España (Murcia y Andalucía), Portugal, Países Bajos, Eslovenia e Italia que desarrollan 6 plataformas diferentes. Estas herramientas serán agrupadas en una sola plataforma y los resultados de la difusión extenderán el conocimiento generado para promover el desarrollo de nuevas soluciones similares a Pharaon. La validación de las 6 plataformas se realiza en dos etapas: prevalidación (fase previa ya superada por todos) y pilotos a gran escala (LSP).

El piloto murciano está centrado en usuarios con insuficiencia cardiaca crónica (ICC), una de las enfermedades con mayor prevalencia en la población

mayor y que más limitación causa, requiriendo de monitoreo continuo para mejorar su seguimiento y el pronóstico de la enfermedad [1].

Actualmente, el control de las enfermedades crónicas supone un gran desafío para los servicios sanitarios, que cuentan con recursos limitados, debido al aumento de su prevalencia por el envejecimiento de la población. Entre los avances tecnológicos, las aplicaciones para teléfonos inteligentes (app), surgen como posible solución a este reto, prestando una asistencia sanitaria inteligente, a bajo coste, accesible y que mejora la calidad de vida de las personas [2]. Sin embargo, el acceso de los mayores a los teléfonos inteligentes es reducido y existe un alto porcentaje de usuarios con escasas habilidades tecnológicas, lo que dificulta el buen uso de este tipo de aplicaciones [3].

Con el fin de mejorar nuestra comprensión sobre el uso de la tecnología en el seguimiento de la salud que realizan los usuarios que integran el proyecto Pharaon, el objetivo de este trabajo es determinar los predictores de adherencia en el uso de una app en salud en pacientes con ICC.

2. Material y método

Tras los primeros 6 meses de intervención permanecen activos 126 usuarios con ICC, mayores de 55 años, que estarán en seguimiento hasta completar 12 meses. Éstos se clasifican en 3 categorías según las tecnologías que prueban (“Silver, Gold y Premium”), se categorizan en estos grupos en función de su grado de autonomía. Todos los participantes disponen de unas tecnologías “básicas”, siendo éstas una aplicación instalada en su teléfono móvil (app), un peso, un tensiómetro y una smart-band, que transmiten datos por bluetooth a la app requiriendo del usuario unos sencillos pasos. La categoría “Silver” sólo dispone de las tecnologías básicas, mientras que los “Gold” tiene añadidos también sensores de presencia y movimiento en el domicilio. Los “Premium”, disfrutan de las tecnologías de los anteriores grupos y también disponen de sensores de consumo eléctrico en diferentes dispositivos el hogar.

La app, además, incluye unos cuestionarios validados y de elaboración propia que nos ayudan a conocer el estado sociosanitario de los usuarios [4] y los cambios producidos durante la participación en el proyecto, que deben rellenar al comienzo del estudio, a los 6 meses y a los 12. Dichos cuestionarios son: Clasificación funcional de la New York Heart Association (NYHA) [5,6], Conocimiento sobre ICC, Psicosocial, Barber (anciano de riesgo) [7], Barthel (nivel de autonomía) [8], Dieta, nivel de autocuidado, Morinsky motivación y conocimiento (sobre el tratamiento) [9], Goldberg- (ansiedad/depresión) [10] y Calidad de vida de Minnesota [11]. Con ellos detectamos necesidades y desviaciones del estado de salud de los usuarios. Todos los participantes y sus cuidadores han recibido formación específica sobre el manejo de la app y las diferentes tecnologías.

Todos los datos registrados en la app y recogidos por los diferentes dispositivos son transmitidos mediante la app a una solución informática a la que tienen acceso un call center formado por enfermeros especialistas en enfermería familiar y comunitaria, a través del cual pueden observar en tiempo real las mediciones y datos registrados y actuar en caso necesario de manera remota. Esto permite intervenir pronto en el caso de descompensaciones de la enfermedad y conocer el estilo de vida los pacientes, adelantándonos a posibles complicaciones y actuando de forma proactiva ayudándolos a obtener un estilo de vida más saludable de una forma personalizada y asistida. Además, este equipo interviene animando a los usuarios a participar cada vez que se detecta una disminución de la adherencia.

Para determinar la adherencia al proyecto se han establecido indicadores de uso de la app, teniendo en cuenta las tareas habituales que son compromiso del paciente y/o cuidador, como son la frecuencia de registro semanal de las cifras de peso, tensión arterial (TA) y frecuencia cardíaca (FC). No hemos tenido en cuenta los cuestionarios puesto que solamente se registran 3 veces durante el proyecto y por tanto no son buenos determinantes de la adherencia, tampoco se han tenido en cuenta el número de accesos a la app puesto que no nos aportan información relevante si no han dejado registrada una medición objetiva. Se estableció como umbral mínimo el registro de al menos un valor de cada indicador a la semana y como umbral ideal una frecuencia de registro de al menos 3 valores de peso (más importante clínicamente en los pacientes con ICC) y 2 valores de TA y FC a la semana. Una frecuencia de registro inferior al mínimo supone la no cumplimentación y la necesidad de incidir sobre dichos pacientes, intentando recaptar su atención (tabla 1).

Indicador	Alerta	Mínimo	Ideal
Peso	0	1	3
TA	0	1	2
FC	0	1	2

Tabla 1. Indicadores y frecuencias de cumplimentación semanales

Según los resultados de adherencia clasificamos a los pacientes en tres categorías:

- “Cumplidores”: Aquellos pacientes que cumplen los umbrales mínimos de registro semanales (1 registro de peso, TA y FC). Cuando además alcanzan los umbrales ideales (3 registros de peso y 2 de TA y FC) se consideran “cumplidores ideales”.
- “Ausentes”: Pacientes que no cumplimentan el umbral mínimo de usabilidad de la app entre 1 y 3 semanas consecutivas después de recibir la formación, sin motivo justificativo. Estos pacientes exigen un recordatorio o ayuda para regresar al uso.
- “Resistentes”: Aquellos pacientes que no cumplimentan el umbral mínimo de usabilidad de la app durante 4 semanas consecutivas a pesar de los intentos de recaptación.

Para el análisis estadístico se revisaron los datos sociodemográficos y los indicadores de adherencia, con el software estadístico SPSS v27. Las variables cuantitativas se describieron mediante media y desviación típica y las variables cualitativas mediante frecuencia absoluta y relativa. Se empleó las pruebas chi cuadrado para establecer una relación entre variables cualitativas, con un nivel de confianza del 95% ($p < 0,05$) y para las variables cuantitativas se utilizó la regresión lineal.

3. Resultados y discusión

Tras 4 meses de monitorización clasificamos a los participantes según su adherencia: 51 usuarios (40,48%) son “cumplidores ideales”, 25 (19,84%) “cumplidores”, 26 (20,63%) son “ausentes” y 24 (19,05%) “resistentes”. Es decir, más de la mitad de los usuarios (60,36%) tiene adherencia a la aplicación, y de ellos la mayoría lo hace de manera ideal (figura 1).

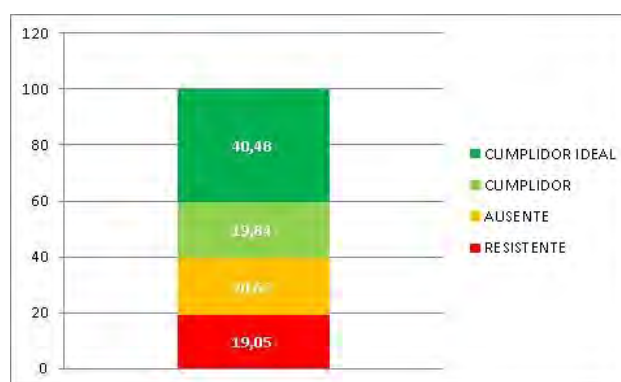


Figura 1. Porcentaje de usuarios según las categorías de cumplimentación

Centrándonos en el detalle del análisis semanal por registros (en lugar de por usuarios) los indicadores se comenzaron a contabilizar de manera semanal (todos los lunes) desde el 07 de mayo, periodo en el que estaban todos los pacientes formados y preparados para utilizar la app y las diferentes tecnologías. Desde esa fecha hasta la última semana contabilizada, del 14 al 21 de agosto, los

indicadores fueron variando continuamente, especialmente durante los 2 primeros meses, es decir, tras el inicio de los registros por parte de los pacientes, siendo un periodo de adaptación que se necesitó un refuerzo constante por parte del call-center. También se observa un descenso en los registros en aquellos periodos en que la aplicación ha tenido “caídas” por problemas técnicos (3 de mayo, 17 junio (varios días) y 27 de julio) y durante el periodo vacacional, siendo festivo en la región el día 9 de junio (figura 2). Los datos durante estas “caídas” se han tenido en cuenta en el análisis estadístico. En estas ocasiones se ha requerido una intervención motivacional del call center para restaurar los índices de adherencia previos a las caídas.

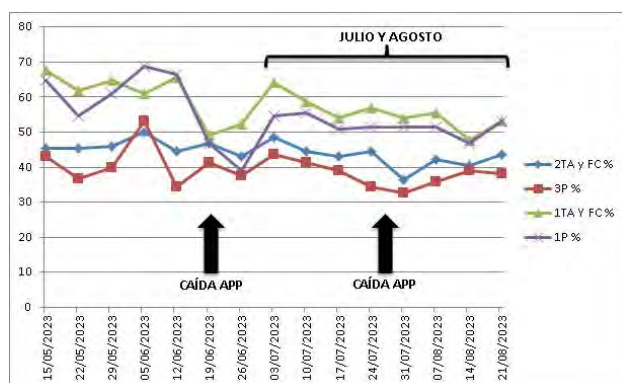


Figura 2. Cambios porcentuales en los indicadores semanales de registro de umbrales mínimos y máximos de peso, TA y FC.

El indicador que más variabilidad presenta es registro del peso (tabla 2), esto es debido a que es la medición que puede costar más trabajo realizar a los usuarios, por problemas de movilidad que les dificultan subirse y bajarse de la báscula de manera frecuente o por desplazamientos fuera del hogar en los que no pueden llevarse el dispositivo, que es pesado y ocupa bastante espacio, siendo menos manejable que el tensiómetro.

Medidas registradas en app	Media
1 Peso	54,47%
3 Pesos	39,33%
1 TA y FC	57,75%
2 TA y FC	44,29%

Tabla 2. Media del porcentaje de registros semanales en la app de las cifras de peso, TA y FC

Tras analizar los datos sociodemográficos de los usuarios hemos obtenido que 75 (59,5%) son hombres y 51 (40,5%) mujeres. La edad media es de 71,52 años. Respecto al estado civil el 8,1% son solteros/as, el 63,7% casados, 23,4% viudos, 4,8% separados. En cuanto al nivel de estudios el 7,9% no tienen estudios, el 53,5% tienen estudios básicos, el 21,9% estudios intermedios y 16,7% estudios superiores.

Según la categoría del usuario dentro del proyecto, de acuerdo a la tecnología de la que disponen 68 (54%) son “Silver”, 34 (27%) “Gold y 24 (19%) “Premium”.

Se ha realizado un análisis entre el indicador de adherencia y la respuesta a los diferentes cuestionarios, con el fin de establecer perfiles que nos determinen el nivel de adherencia de cada usuario.

Respecto al cuestionario NYHA el 54,8% de los pacientes presentan un nivel 1. El 62,9% presenta un conocimiento bajo de su enfermedad, un 68,5% puede mejorar sus conocimientos sobre la alimentación. Por otro lado, el 26,4% precisan valoración psicosocial profesional. El 66,3% de la muestra son ancianos de riesgo, sin embargo, el 76,9% son independientes. El 54,1% presenta un autocuidado excelente frente al 39,8% mejorable. En cuanto a la calidad de vida el 43,0% manifiestan tener una buena calidad, el 39% mejorable, 16% mala y solamente el 2% muy mala.

Se realizó el análisis estadístico entre los datos sociodemográficos y las respuestas de los cuestionarios con los resultados de adherencia. No se encontró relación estadísticamente significativa entre estas variables. Tampoco entre la adherencia y la categorización de los usuarios (silver, gold y premium).

Hay que tener en cuenta que aunque no se hayan establecido relaciones estadísticamente significativas sí se han detectado tendencias que pueden ser útiles para detectar patrones de adherencia y adelantarnos a la hora de conocer qué tipo de usuarios responden mejor a este tipo de tecnologías. Las tendencias encontradas relacionan la adherencia a la app con el sexo, así como los cuestionarios Barthel y autocuidado, observándose mayor interacción en la app por parte de los hombres frente a las mujeres (p=0,108), en aquellos participantes independientes para las actividades básicas de la vida diaria (ABVD) (p=0,182) y en los que presentan mayor nivel de autocuidado (p=0,197).

Los resultados con respecto a la menor adherencia a la app por parte de las mujeres concuerdan con los resultados aportados por diversos autores [12,13] que determinan que las mujeres históricamente se han dedicado a las tareas del hogar y al cuidado de otros por lo que tienen menos tiempo para cuidar de su propia salud, objetivándose en la alta prevalencia de trastornos crónicos de salud y en el abandono de su autocuidado. Por otro lado, a través del contacto con los usuarios, hemos tenido constancia de cómo en algunos casos de hombres cumplidores la app es manejada realmente por la cuidadora principal del usuario, normalmente su esposa o hija.

Como era de esperar, las personas que tienen mayor puntuación en el índice de Barthel, y por tanto mayor independencia para realizar las ABVD, tienen una mayor adherencia a la app, pues esta situación es esencial para el autocuidado. En este sentido, también se ha encontrado relación entre la adherencia a la app y el autocuidado. También, probablemente estas personas posean mayores habilidades para el uso de las tecnologías,

independientemente de la edad, donde no hemos encontrado relación.

4. Conclusiones

Más de la mitad de los usuarios utilizan la app de manera frecuente, cumpliendo los requisitos mínimos de usabilidad a los que se comprometieron. Incluso la mayoría de estos lo realizan de forma ideal, llegando a hacerlo de manera diaria. Hemos de tener en cuenta que los problemas en la aplicación, así como los días festivos influyen de forma negativa en el cumplimiento de los participantes, exigiendo un rescate motivador del usuario por parte del call center. Destacar la importancia de que exista un equipo humano que de soporte a la app de salud. También influye el tipo de dispositivo a usar teniendo en cuenta su complejidad, manejabilidad y grado de autonomía del paciente.

Aunque no existen relaciones significativas, teniendo en cuenta las tendencias se ha observado que el usuario que mayor adherencia presenta se corresponde con un perfil masculino, independiente y con un autocuidado excelente. Los indicadores además de determinar la adherencia también valoran la implicación del paciente en su salud.

Si bien no conviene generalizar, los resultados de este trabajo evidencian que es necesario considerar la perspectiva de género y el nivel de autonomía del posible usuario. Se evidencia que se trata de un buen modelo de prevención secundaria. Todos estos factores deben ser tenidos en cuenta para el desarrollo de aplicaciones en salud desde el sistema sanitario. De esto modo se deben adaptar las app empleadas a las características de los usuarios con el objetivo de mejorar la adherencia y usabilidad.

Referencias

- [1] McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M et al. Guía ESC 2021 sobre el diagnóstico y tratamiento de la insuficiencia cardíaca aguda y crónica. *Rev Esp Cardiol.* 2022;75(6):523.e1-523.e114 DOI: 10.1016/j.recesp.2021.11.027
- [2] Shah Nazir, Yasir Ali, Naeem Ullah, Iván García-Magariño. "Internet of Things for Healthcare Using Effects of Mobile Computing: A Systematic Literature Review", *Wireless Communications and Mobile Computing.* 2019. DOI:10.1155/2019/5931315
- [3] González de León B, León Salas B, Del Pino-Sedeño T, Rodríguez-Álvarez C, Bejarano-Quisoboni D, Trujillo-Martín MM. Aplicaciones móviles para mejorar la adherencia a la medicación: revisión y análisis de calidad [Mobile applications to improve drug adherence: Review and quality analysis]. *Aten Primaria.* 2021;53(9):102095. Spanish. DOI: 10.1016/j.aprim.2021.102095.
- [4] López Alonso, Sergio R., & Morales Asencio, José Miguel. ¿Para qué se administran las escalas, cuestionarios, tests e índices? *Index de Enfermería,* 2005: 14(48-49), 7-8.
- [5] Jill A. Bennett, Barbara Riegel, Vera Bittner, Joyce Nichols. Validity and reliability of the NYHA classes for measuring research outcomes in patients with cardiac. *2002(4):262-70.* <https://doi.org/10.1067/mhl.2002.124554>.
- [6] Martín, J., Carvajal, A., & Arantzamendi, M. Instrumentos para valorar al paciente con insuficiencia cardíaca avanzada: una revisión de la literatura. *Anales del Sistema Sanitario de Navarra,* 2015: 38(3), 439-452. <https://dx.doi.org/10.4321/S1137-66272015000300008>
- [7] Barber, T. X., & Wilson, S. C. (1978 -1979). The Barber Suggestibility Scale and the Creative Imagination Scale Experimental and clinical applications. *American Journal of Clinical Hypnosis,* 21(2-3), 84-108. <https://doi.org/https://doi.org/10.1080/00029157.1978.10403966>
- [8] Gompertz P, Pound P, Ebrahim S. A postal version of the Barthel Index. *Clinical Rehabilitation.* 1994;8(3):233-239. doi:10.1177/026921559400800308
- [9] Jiménez-Chala EA, Durantez-Fernández C, Martín-Conty JL, Mohedano-Moriano A, Martín-Rodríguez F, Polonio-López B. Use of Mobile Applications to Increase Therapeutic Adherence in Adults: A Systematic Review. *J Med Syst.* 2022 Oct 22;46(12):87. doi: 10.1007/s10916-022-01876-2. PMID: 36272023.
- [10] Carbonell, Marta Martin; DIAZ, Raquel Perez; MARIN, Antonio Riquelme. Diagnostic usefulness of anxiety and depression scale goldberg (EAD-G) in Cuban adults/Valor diagnóstico de la escala de ansiedad y depresión de Goldberg (EAD-G) en adultos cubano. *Universitas Psychologica,* 2016, vol. 15, no 1, p. 177-192.
- [11] Palomino, G., Pareja, M. J., Pareja, M. C., Nevado, N., Mendoza, F., Dávila, F. A., & Jaramillo, C. Utilidad de la Escala de Minnesota en el seguimiento de los pacientes con insuficiencia cardíaca crónica. *CardiCore,* 2018: 53(2), 55-60.
- [12] Marcos-Marcos J. Desigualdades de género en el deterioro de la salud como resultado del cuidado informal en España. *Gac Sanit.* 2011;25:100-7
- [13] Marcos-Marcos J, et al. Gender identity in informal care: impact on health in Spanish caregivers. *Qual Heal Res.* 2013;23:1506-2

Aplicación para la gestión personalizada de pacientes con diabetes

F. Pastor García¹, A. De Ramón Fernández¹, D. Ruiz Fernández¹, D. Marcos Jorquera¹

¹ Departamento de Tecnología Informática y Computación, Universidad de Alicante, San Vicente del Raspeig (Alicante), España, {aderamon@dtic.ua.es, druiz@dtic.ua.es, dmarcos@dtic.ua.es}

Resumen

El término "diabetes mellitus" se emplea para describir un grupo de trastornos metabólicos caracterizados por la hiperglucemia debido a la deficiencia en la secreción de insulina, acción deteriorada de la insulina o ambas condiciones. Además, se erige como una de las enfermedades no contagiosas más comunes en el mundo. Los individuos con diabetes requieren un monitoreo constante de sus niveles de glucosa y otros factores que afectan la enfermedad, lo que genera una carga importante en los centros de salud que se centran en tareas de control y seguimiento. Con el avance de las tecnologías, han surgido aplicaciones móviles orientadas a mejorar la vida de las personas afectadas por esta enfermedad. Este estudio presenta una aplicación diseñada para la gestión personalizada e integral de la diabetes, abordando las limitaciones que tienen otras aplicaciones existentes en el mercado. La aplicación captura, monitorea y analiza todos los factores relevantes que contribuyen a la enfermedad, al tiempo que evalúa el cumplimiento del paciente de los objetivos establecidos. Esta contribución aspira a mejorar la calidad de vida y las capacidades de autocuidado de las personas que padecen diabetes, ofreciendo un enfoque global de su condición.

1. Introducción

Bajo el término “diabetes mellitus” se engloban un grupo de enfermedades metabólicas cuyo nexos común es la hiperglucemia debido a un déficit de la secreción de insulina, defectos de la acción de esta o ambos trastornos. Es, además, una de las enfermedades no contagiosas más frecuentes del mundo. En España afecta a casi 6 millones de personas. Los datos recogidos por las encuestas nacionales y europeas de salud reflejan una tasa de 7,51 de cada 100 personas, mostrando mayor prevalencia en personas de más de 65 años. En España la población diabética se ha incrementado en un 3,71% desde 1991[1].

La carga de esta dolencia supone el 8,2% del total del presupuesto sanitario, lo que implica 5.809 millones de euros al año de los que 2.143 son, precisamente, debidos a las dificultades añadidas [2]. Por su parte, los pacientes reclaman mejoras en el proceso de atención como acceso a sistemas de monitorización y programas formativos, entre otros.

Los pacientes de diabetes necesitan de una monitorización constante de sus niveles de glucosa y del resto de factores que pueden incidir en la enfermedad. Esto hace que, a menudo, la carga asistencial de los centros de salud para

tratar a este tipo de pacientes sea muy elevada solo en cuanto a tareas de control y seguimiento se refiere.

2. Estado del arte

Con el auge de las tecnologías, en los últimos años han ido surgiendo aplicaciones móviles orientadas a facilitar la vida a las personas que sufren esta enfermedad. Una de las más completas es Diabetes:M [3], que permite añadir registros de comida, de medicación, de glucosa en sangre, y de deporte, así como visualizar tendencias y gráficas de la glucosa en sangre. También permite realizar consultas a la base de datos alimenticia Edamam [4], y la conexión con glucómetros Bluetooth aunque estas funcionalidades no son gratuitas.

SocialDiabetes [5], es otra de las aplicaciones más reconocidas en este ámbito. Al igual que en el caso anterior, permite controlar la toma de medicación, medidas de glucosa, las ingestas de alimentos y la de actividad física, además de controlar la HbA1c (Hemoglobina Glicosilada), el peso, la tensión arterial, y las cetonas. Sin embargo, sólo permite realizar análisis y gráficas de los niveles de glucosa en sangre. También permite sincronizar los registros de actividad con los almacenados en el dispositivo móvil y la conexión con glucómetros Bluetooth.

Otro ejemplo es la aplicación Diabetescontrole [6]. Se trata de una aplicación muy intuitiva y fácil de usar, pero permite registrar un número inferior de factores, almacenando sólo los valores de glucosa en sangre para su posterior análisis.

Estas aplicaciones son sólo algunos de ejemplos aplicaciones orientadas a la gestión de la diabetes, y que, sin embargo, presentan algunas limitaciones que las convierten en soluciones parciales. Muchas de estas aplicaciones ofrecen ciertas funcionalidades sólo en su modalidad de suscripción o previo pago. Una de ellas es el acceso a la información nutricional y en concreto al consumo de azúcares e hidratos de carbono que cada paciente realizar por comida, información vital para el control de la glucosa. Otra limitación de estas aplicaciones es que no suelen ofrecer la posibilidad de configurar valores de control o límite para detectar valores críticos. Además, estas aplicaciones analizan cada factor de manera individual y no de manera integral y no realizan una valoración global de la adherencia que está mostrando al cumplimiento de sus objetivos. Es por ello,

que en este trabajo se presenta una aplicación orientada a la gestión de la diabetes de manera personalizada e integral y que supere las debilidades comentadas.

3. Requerimientos y arquitectura del sistema.

Para el diseño de la aplicación se establecieron una serie de requerimientos funcionales y no funcionales que debían de ser cumplimentados. Los requerimientos funcionales fueron los siguientes:

- REQ FUN 01. Acceso al sistema controlado por usuario y contraseña.
- REQ FUN 02. Registro de un nuevo usuario.
- REQ FUN 03. Modificación de los datos personales del usuario.
- REQ FUN 04. Identificación de módulos por patrón de colores: cada módulo de la aplicación debe estar identificado con un color para que puede ser fácilmente identificable por el usuario.
- REQ FUN 05. Visualizar los registros de cada módulo en formato texto.
- REQ FUN 06. Filtrar de los registros por fecha.
- REQ FUN 07. Añadir registros nuevos.
- REQ FUN 08. Visualizar los registros gráficamente mediante gráficas diarias y/o semanales.
- REQ FUN 09. Eliminar registros.
- REQ FUN 10. Modificar la medicación asignada.
- REQ FUN 10. Visualizar una puntuación semanal global del grado de cumplimiento de los objetivos.
- REQ FUN 11 Visualizar una puntuación semanal individual para cada módulo.
- REQ FUN 12 Emisión de notificaciones para recordar la toma de medicación.
- REQ FUN 13 Asignar valores límites para cada factor a controlar.

Por su parte, los requerimientos no funcionales que debía cumplir la aplicación fueron los siguientes:

- REQ NO_FUN 01. La aplicación tiene que ser accesible y usable: la información debe mostrarse de forma clara y sencilla y utilizar fuentes, tamaños y colores que faciliten el uso y la lectura de datos para personas con problemas de visión.
- REQ NO_FUN 02. La aplicación debe almacenar los datos recogidos de forma permanente.
- REQ NO_FUN 03. El acceso a la aplicación debe implementar medidas de seguridad.
- REQ NO_FUN 04. La aplicación debe ser multiplataforma.
- REQ NO_FUN 05. La aplicación debe permitir integraciones de otras aplicaciones externas.

Para desarrollar un prototipo de aplicación que cumpla con los requerimientos funcionales y no funcionales descritos, se ha diseñado un sistema distribuido de arquitectura cliente-servidor basado en la orientación a servicios (SOA) y adoptando el estilo RESTful, asegurando así la reusabilidad, escalabilidad, bajo acoplamiento, interoperabilidad y seguridad necesarios [7].

Para el desarrollo de los servicios en estilo RESTful se ha utilizado Node.js como motor de ejecución y Express como framework que permite simplificar la creación de APIs RESTful de manera eficiente. Para garantizar la persistencia de los datos se ha incorporado una base de datos MySQL.

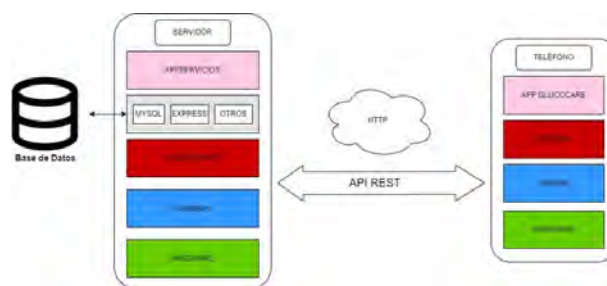


Figura 1. Arquitectura del sistema propuesto.

4. Diseño y funcionamiento de la aplicación móvil Glucocare

La aplicación Glucocare está pensada para una gestión integral de la diabetes, teniendo en cuenta todos aquellos factores que según la literatura tienen un impacto en la salud del paciente. Inicialmente la aplicación muestra la opción de registrarse como nuevo usuario o de iniciar sesión mediante un sistema de usuario y contraseña que ha sido implementado utilizando Json Web Tokens como mecanismo de seguridad. Una vez dentro de la aplicación, ésta se encuentra estructurada en módulos o pantallas que permiten gestionar la información de cada una de las variables que se controlan:

Módulo de Registros de actividad física

En esta pantalla se puede consultar los registros de actividad del usuario de la última semana y que hayan sido recogidos a través de una pulsera inteligente y sincronizados mediante la aplicación Google Fit. Los registros de actividad física son representados mediante un gráfico de barras de color, siguiendo el enfoque planteado en [8]: si en ese día el usuario ha realizado 20 minutos o más de actividad, se representará en color verde; si ha realizado entre 15 y 20 minutos se asigna el color naranja y si la actividad física realizada ha sido inferior a 15 min se asigna color. También se pueden añadir registros de actividad de forma manual, incluyendo distancia recorrida, tiempo de actividad, calorías consumidas y frecuencia cardíaca.

Módulo de Registros de glucosa en sangre

Este módulo se centra en el seguimiento de las medidas de glucosa en sangre. El usuario puede visualizar sus registros de glucosa y una gráfica con el valor medio de cada día (ver figura 2 sup.izq. y 2 sup.der.). Las barras del gráfico serán de un color u otro siguiendo el siguiente sistema de reglas [9]: si el valor de glucosa se encuentra entre los valores máximos y mínimos preestablecidos, se representará de color verde; si se encuentra entre 70 mg/dL y su valor mínimo u entre su valor máximo y 180 mg/dL se mostrará de color naranja; y si es inferior a 70 mg/dL o superior a 180 mg/dL se representará de color rojo.

Módulo de Registros de comida

Este módulo permite al usuario llevar un registro actualizado de las comidas realizadas: qué alimentos ha ingerido, a qué hora, kilocalorías consumidas en cada comida y su valor nutricional a partir de la medición de hidratos de carbono y azúcares (ver figura 2 inf.izq.). Cuando se añade un alimento y su cantidad (por ejemplo, 100 gramos de pechuga de pollo o 1 manzana), se realiza una consulta a la base de datos nutricional de Edamam que gracias a su API devuelve los valores nutricionales. También permite filtrar, consultar y eliminar los registros de comidas registrados en el sistema. La cantidad de kilocalorías, hidratos de carbono y azúcares también se representa mediante un gráfico de barras de colores, que sigue la siguiente codificación: En primer lugar se establece un valor límite superior, y un valor intermedio. Se representará de color verde si el valor no supera el límite intermedio, en color naranja si se encuentra entre el límite intermedio y el límite superior y en color rojo si se supera el límite superior. Estos límites son parametrizables en la propia aplicación, según las indicaciones del especialista.

Módulo de Registros de medicación

Este módulo permite gestionar la medicación asignada. Entre sus funcionalidades permite que el usuario consulte la medicación que tiene asignada (número de tomas, dosis, forma de administración, observaciones...), el registro de cuántas tomas ha realizado de cada medicamento, de cuántas le quedan por tomar, añadir una nueva medicación o eliminar una ya asignada (ver figura 2 inf. der.). También permite programar notificaciones para avisar al usuario de que debe realizar una toma.

Puntuación individual de cada módulo y puntuación media global.

En función de los resultados obtenidos y del cumplimiento de los niveles aconsejados por el especialista para cada módulo se puntuará el rendimiento del usuario en cada uno de ellos mediante una puntuación de estrellas, otorgando un máximo de 3 estrellas cuando

el usuario ha conseguido mantenerse de forma constante en los valores óptimos aconsejados y 0 estrellas cuando estos han sido incumplidos de forma constante (ver Figura 3 izq.).



Figura 2. Interfaces de la app Glucocare. Sup.izq) Registros de glucosa; Sup. der.) Gráfico de los registros de glucosa; Inf. izq.) Kilocalorías consumidas e Información nutricional del alimento/s ingerido/s; Inf. der) Registro de medicación

Para cada módulo, a partir de sus valores medios diarios, se establecerá un valor medio semanal. Al pulsar sobre cada botón aparece una ventana emergente que muestra un resumen de los datos obtenidos durante la semana y cuál era el objetivo a cumplir (ver Figura 3 der.).

La puntuación semanal e individual de cada módulo servirá para obtener una puntuación media semanal global del usuario, que se mostrará en su módulo

correspondiente. Dentro de este módulo, también se valorará el índice de masa corporal (IMC) del usuario según [10], de tal forma que si el IMC se encuentra entre 18.5-24.9 el peso se indica que el peso es normal, si está por debajo de 18.5 se indica bajo peso, y si es mayor que 25 o que 30 se indica sobrepeso u obesidad, respectivamente.



Figura 3. Izq.) Puntuación semanal individual de cada módulo; Der.) Información sobre el cumplimiento de los objetivos por parte del usuario.

5. Conclusiones

Las cifras alarmantes de prevalencia y el impacto económico y asistencial de la diabetes han impulsado la búsqueda de soluciones tecnológicas eficaces que alivien la carga tanto para los pacientes como para los sistemas de salud. La aplicación que se presenta en este trabajo aborda la gestión personalizada de la diabetes desde un punto de vista integral, basándose en la recolección, monitorización y análisis de todos los factores que influyen en la enfermedad así como valorando el grado de cumplimiento que tiene el paciente con cada objetivo propuesto y cumple con los requisitos funcionales y no funcionales establecidos durante la fase de concepción de la misma. Esta contribución hacia la mejora de la calidad de vida y el autocuidado de las personas que padecen esta enfermedad.

Referencias

[1] La diabetes en España y en el mundo, en datos y gráficos. (7 de septiembre de 2023). <https://www.epdata.es/datos/diabetes-espana-datos-graficos/472>

[2] Crespo, C., Brosa, M., Soria-Juan, A., Lopez-Alba, A., López-Martínez, N., & Soria, B. (2013). Costes directos de la diabetes mellitus y de sus complicaciones en España (Estudio SECCAID: Spain estimated cost Ciberdem-Cabimer in Diabetes). *Avances en diabetología*, 29(6), 182-189.

[3] Diabetes:M - Your Diabetes Management App - Keep diabetes under control. (7 de septiembre de 2023). Diabetes:M - Your Diabetes Management App. <https://diabetes-m.com/>

[4] Edamam - Food Database API, Nutrition API and recipe API. (7 de septiembre de 2023). <https://www.edamam.com/>

[5] Gestión Integral para el control de la diabetes - SocialDiabetes. (7 de septiembre de 2023). <https://www.socialdiabetes.com/es>

[6] Diabetes controle - glucemia - aplicaciones en Google Play. (7 de septiembre de 2023). <https://play.google.com/store/apps/details?id=com.bloodglucosetracker.diabetesreading&hl=es&gl=US>

[7] Ebert, J. (2013). SOA with REST: principles, patterns & constraints for building enterprise solutions with REST by Thomas Erl, Benjamin Carlyle, Cesare Pautasso, Raj Balasubramanian. *ACM SIGSOFT Software Engineering Notes*, 38(3), 32-33.

[8] ¡Haga actividad física! | Viviendo con diabetes | Diabetes | CDC. (7 de septiembre de 2023). <https://www.cdc.gov/diabetes/spanish/living/beactive.html>

[9] Monitoreo del nivel de azúcar en la sangre. (7 de septiembre de 2023). <https://www.cdc.gov/diabetes/spanish/living/bloodglucosemonitoring.html>

[10] World Health Organization: WHO. (7 de septiembre de 2023). Obesidad y sobrepeso. [www.who.int. https://www.who.int/es/news-room/fact-sheets/detail/obesity-and-overweight](https://www.who.int/es/news-room/fact-sheets/detail/obesity-and-overweight)

Diferenciación por métodos estadísticos convencionales y *machine learning* entre tejido pulmonar sano y patológico de medidas de impedancia eléctrica

G. Company Se¹, R. Bragós Bardia¹, V. Pajares Ruiz², A. Rafecas Codern², J. Rosell Ferrer¹, P. J. Riu Costa¹, L. Nescolarde Selva^{1*}

¹ Department of Electronic Engineering, Universitat Politècnica de Catalunya, Barcelona, Spain,
{georgina.compnay, ramon.bragos,pere.riu, javier.rosell, lexa.nescolarde}@upc.edu

² Department of Respiratory Medicine, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain,
{vpajares, arafecas}@santpau.cat

Resumen

El uso de la espectroscopia de impedancia eléctrica para la diferenciación del tejido pulmonar es una oportunidad para mejorar el diagnóstico clínico. El objetivo es distinguir entre diferentes estados del tejido pulmonar mediante el estudio de las diferencias entre los parámetros del espectro de impedancia entre dos frecuencias separadas (15 kHz y 307 kHz) en la región de la dispersión beta. Además, también se aplican algoritmos de machine learning para la clasificación automática de patologías pulmonares. Se han encontrado diferencias significativas entre aquellos tejidos que experimentan un aumento en la densidad (neoplasia y fibrosis) y aquellos tejidos que sufren una destrucción del tejido (enfisema). Además, los algoritmos automáticos son capaces de clasificar con gran exactitud ($\geq 86\%$) muestras de tejido neoplásico. Son necesarios más estudios y mayor validación de los algoritmos para diferenciar aquellos estados del tejido que son más similares entre sí.

1. Introducción

El correcto diagnóstico de las enfermedades pulmonares, así como la correcta toma de muestras de tejido, es esencial para un buen pronóstico del paciente. La adquisición de muestras de tejido en lesiones centrales en aquellos pacientes con sospecha de neoplasia es relativamente sencilla. Sin embargo, en lesiones periféricas, ese diagnóstico sigue siendo un reto.

Con el fin de complementar los actuales métodos de diagnóstico de las enfermedades pulmonares, las medidas de espectroscopia de impedancia eléctrica (EIS) podrían ofrecer la posibilidad de diferenciar entre tejido sano y tejido patológico, así como entre distintos tipos de tejido patológico, para ayudar y facilitar la elección de una localización específica para la adquisición de material biológico. Además, podría ofrecer la selección de la zona a biopsiar en tiempo real.

La bioimpedancia (Z) se define como la oposición que el tejido biológico ofrece a un determinado flujo de corriente eléctrica administrado. Cuando dicho flujo de corriente es de corriente alterna, la bioimpedancia deviene dependiente de la frecuencia. A su vez, cuando se usa un amplio espectro de frecuencias para medir la bioimpedancia, se aplica la técnica de medida de espectroscopia de impedancia eléctrica. La Z está compuesta por un elemento

resistivo (R), que define la oposición que ofrece el fluido presente tanto en el medio extracelular como en el medio intracelular, y un elemento de reactancia capacitiva (X_c), que define el comportamiento capacitivo de las membranas celulares. Derivado de estos dos componentes, se puede expresar la Z por su módulo ($|Z|$) definido como $\sqrt{R^2 + X_c^2}$ y por su fase (PA) definida como $\text{atan}\left(\frac{X_c}{R}\right)$. Este ángulo de fase aparece debido a que la X_c produce un retraso entre la corriente y el voltaje [1]–[3].

Debido a que las membranas celulares tienen un comportamiento capacitivo, se produce una relajación, denominada dispersión beta entre las decenas de kHz y las decenas de MHz. Este fenómeno conlleva a una disminución de la permitividad (ϵ) así como a un incremento de la conductividad. Además, dependiendo de las propiedades del tejido, se producen diferencias en la relajación [4]–[6].

El uso de inteligencia artificial para el análisis de datos clínicos ha cobrado importancia en los últimos años. En términos generales hay numerosos estudios usando machine learning para la clasificación entre pacientes patológicos y grupos control [7], [8] así como para la clasificación entre respondedores y no-respondedores ante un tratamiento específico [9]. También podemos encontrar estudios para clasificación de riesgo ante una patología [10] así como también para la clasificación en subgrupos de una enfermedad en concreto [11]. Por otro lado, la aplicación de algoritmos de clasificación de *machine learning* presenta una oportunidad para predecir resultados, así como desarrollar nuevos métodos de diagnóstico y mejorar el pronóstico [12]. La aplicación de dichos algoritmos facilitaría el muestreo en tiempo real de tejido pulmonar mediante el uso de una biopsia electrónica.

Hasta donde alcanza el conocimiento de los autores, por el momento solo se ha estudiado la diferenciación entre tejidos pulmonares mediante EIS usando valores medios a una frecuencia determinada para cada uno de los parámetros de bioimpedancia y usando únicamente métodos estadísticos clásicos [13].

El objetivo de este estudio es la diferenciación entre distintos estados del tejido pulmonar (neoplasia, fibrosis, neumonía, sano y enfisema) mediante el estudio del comportamiento de los distintos parámetros de la bioimpedancia ($|Z|$, PA, R y X_c) en la dispersión beta obtenidos mediante medidas mínimamente invasivas de EIS adquiridas mediante broncoscopia. Además, también se aplicarán algoritmos supervisados de *machine learning* para la clasificación de muestras de impedancia entre los distintos tejidos recogidos.

2. Materiales y métodos

Participantes

Se llevaron a cabo medidas de impedancia en un total de 102 pacientes (Edad: 66 ± 14 años; Peso: 74.5 ± 17.2 kg; IMC: 26.8 ± 4.3 kgm⁻²) con broncoscopia prescrita entre el mes de noviembre del 2021 y el mes de agosto del 2022 en el Hospital de la Santa Creu i Sant Pau de Barcelona. En total, se recogieron 30 muestras de tejido pulmonar sano, 29 muestras de tejido neoplásico, 23 muestras de tejido enfisematoso, 12 muestras de fibrosis y 22 muestras de neumonía.

Se obtuvo la aprobación ética del “Hospital de la Santa Creu i Sant Pau” (CEIC-73/2020) según principios de la Declaración de Helsinki para experimentos con seres humanos. Todos los pacientes acreditaron consentimiento informado firmado.

Mediciones de EIS

Para la adquisición de las medidas de impedancia mediante EIS se empleó el método de 3 electrodos a través de la inyección de una corriente multisenso comprendida entre 1 kHz y 1 MHz de frecuencias entre un catéter tetrapolar distal y un electrodo superficial. El voltaje inducido derivado de la aplicación de corriente se mide entre el electrodo distal y un segundo electrodo superficial. La impedancia se mide durante un periodo de 12 segundos a una frecuencia de muestreo de 60 espectros por segundo. Se puede obtener una información más detallada sobre el método de medida en Company-Se et al [14].

Protocolo de medida

Las medidas de impedancia mínimamente invasivas se obtuvieron mediante el proceso de broncoscopia, mediante la inserción de un catéter a través de uno de los puertos de trabajo del broncoscopio. Previo a la broncoscopia se realizó una evaluación radiológica mediante tomografía computarizada (TC) de tórax o en combinación con la tomografía por emisión de positrones (PET/TC). Durante el proceso de adquisición de las medidas de impedancia, los pacientes recibieron anestesia en las vías aéreas superiores y les fue proporcionada sedación intravenosa. La exploración endoscópica se realizó de acuerdo con las guías de práctica clínica.

Análisis de datos

A diferencia del estudio llevado a cabo por Company-Se et al [13], donde se emplearon valores absolutos de los parámetros de impedancia para realizar la diferenciación entre los diferentes estados del tejido, en este estudio se emplea la diferencia en los valores de los parámetros de

impedancia ($|Z|$, PA, R y X_c) entre bajas (15 kHz) y altas (307 kHz) frecuencias, con el fin de evaluar los cambios en los valores de impedancia derivado de la dispersión beta.

Se empleó el test de Shapiro-Wilk para evaluar la distribución de normalidad de las variables (diferencias entre bajas y altas frecuencias en $|Z|$, PA, R y X_c). Se realizó el test de varianza ANOVA con prueba de post-hoc Tamhane t2 para la determinación de diferencias estadísticamente significativas entre los diferentes estados del tejido (neoplasia, fibrosis, neumonía, sano y enfisema).

Para la visualización de los resultados se emplean los gráficos Boxplot para cada uno de los parámetros de la impedancia. La significación estadística entre pares de tejido se muestra con una línea horizontal y asterisco.

Para el empleo de algoritmos de clasificación (“Árbol de decisión” y “Naïve Bayes”) se usan como variables de entrada el módulo y la fase obtenidos en todas las frecuencias comprendidas entre los 15 kHz y los 307 kHz, aprovechando todo el espectro de medida válido. Se realiza una partición de los datos de 75% para entrenar y 25% para validar. Para el caso del algoritmo “Árbol de decisión” se realiza el ajuste del hiperparámetro de profundidad del árbol. Para el caso del “Naïve Bayes” se optimiza la ratio de aprendizaje “learning rate” así como el número de estimadores empleados. En ambos casos se emplea la técnica “Grid Search” para la optimización de los parámetros mencionados.

Para la visualización de los resultados de clasificación de ambos algoritmos se emplea la matriz de confusión donde se indica el número de casos clasificados correctamente para cada uno de los tipos de tejido estudiados.

El análisis estadístico ha sido desarrollado con el Software IBM®SPSS® versión 24 (IBM Corp, Armonk, NY, Estados Unidos). Los algoritmos de clasificación han sido entrenados y aplicados mediante la aplicación Spyder para programación en lenguaje Python.

3. Resultados

Diferenciación estadística de medidas mínimamente invasivas de impedancia eléctrica entre tejido pulmonar sano y patológico

De la *Figura 1* a la *Figura 4* se muestran los Boxplot para cada parámetro de la impedancia ($|Z|$, PA, R y X_c) y para cada tipo de tejido pulmonar (neoplasia, fibrosis, neumonía, sano y enfisema) resultante de la diferencia de los valores de impedancia a 15 kHz y 307 kHz. Además, también se muestra la significancia estadística obtenida, si aplica, entre pares de tejido en cada uno de los parámetros de impedancia, visualizada mediante líneas horizontales con asteriscos azules.

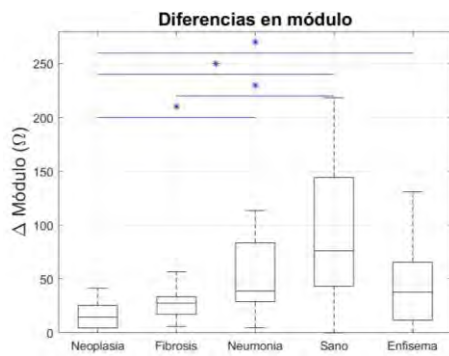


Figura 1. Boxplot de las diferencias entre bajas y altas frecuencias en $|Z|$ para cada tipo de tejido y significancia estadística.

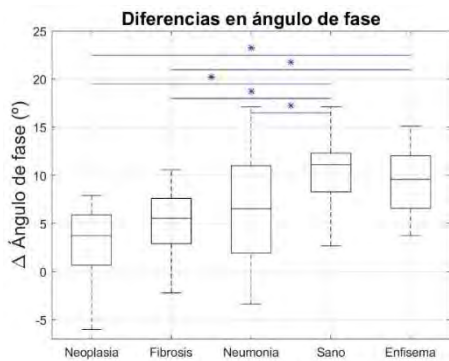


Figura 2. Boxplot de las diferencias entre bajas y altas frecuencias en PA para cada tipo de tejido y significancia estadística.

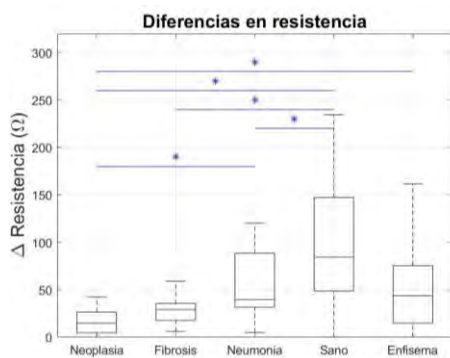


Figura 3. Boxplot de las diferencias entre bajas y altas frecuencias en R para cada tipo de tejido y significancia estadística.

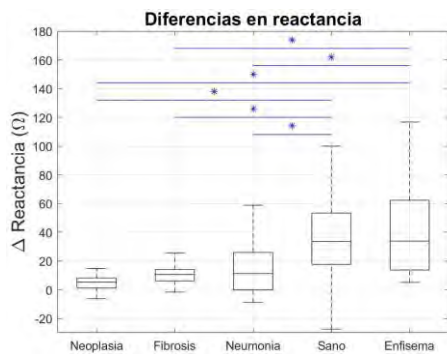


Figura 4. Boxplot de las diferencias entre bajas y altas frecuencias en X_c para cada tipo de tejido y significancia estadística.

Aplicación de algoritmos supervisados de machine learning para clasificación de tejido pulmonar sano y patológico

Las **Figura 5** y **Figura 6** muestran la matriz de confusión obtenidas en la validación de los algoritmos de árbol de decisión y Naïve Bayes respectivamente. El eje vertical muestra la categoría correcta de cada muestra de impedancia mientras que el eje horizontal muestra la categoría predicha por el algoritmo.

Categoría correcta	Enfisema	3	0	2	0	1
	Fibrosis	0	2	1	0	0
	Sano	0	0	6	0	2
	Neoplasia	0	0	0	6	1
	Neumonía	0	1	1	1	2
		Enfisema	Fibrosis	Sano	Neoplasia	Neumonía
Categoría predicha						

Figura 5. Matriz de confusión resultante de la validación del algoritmo de árbol de decisión.

Categoría correcta	Enfisema	2	0	4	0	0
	Fibrosis	0	1	0	1	1
	Sano	2	0	6	0	0
	Neoplasia	0	0	0	7	0
	Neumonía	0	1	1	1	2
		Enfisema	Fibrosis	Sano	Neoplasia	Neumonía
Categoría predicha						

Figura 6. Matriz de confusión resultante de la validación del algoritmo de Naïve Bayes.

4. Discusión

La dispersión beta, que se produce entre las decenas de kHz y las decenas de MHz, se produce debida a la polarización de las membranas celulares [4]. Esto hace que la diferencia entre los valores a bajas y altas frecuencias sea diferente a cero, además de ser diferente en función del estado del tejido y las alteraciones celulares derivadas de la patología, tal y como podemos observar en las **Figura 1** a **Figura 4**. En estas figuras se muestra como las diferencias entre los valores a baja y alta frecuencia son menores en el caso de neoplasia y fibrosis, comparadas con las diferencias en tejido sano y enfisema. Por tanto, dicha dispersión beta influiría más sobre estos dos últimos tejidos que sobre los primeros.

La dispersión beta produce diferencias en los valores medios de impedancia entre bajas y altas frecuencias produciendo diferencias entre neoplasia y neumonía ($|Z|$, R), tejido sano ($|Z|$, PA, R y Xc) y enfisema ($|Z|$, PA, R y Xc); fibrosis y tejido sano ($|Z|$, PA, R y Xc) y enfisema (PA y Xc); neumonía y tejido sano (PA, R y Xc) y enfisema (Xc). No se obtienen diferencias estadísticamente significativas entre fibrosis y neoplasia, fibrosis y neumonía y entre tejido sano y enfisema.

El contenido de aire presente en los pulmones en proporción al tejido es mayor en el tejido pulmonar sano y el enfisema en comparación con las neoplasias, la fibrosis y la neumonía. En el enfisema se produce una destrucción de las paredes alveolares. La neoplasia se caracteriza por un crecimiento celular y un aumento de la vascularización y la fibrosis se caracteriza por un aumento del tejido, pero sin un aumento en la vascularización del mismo. La similitud en ambas patologías en cuanto al incremento de tejido y, a su vez, de concentración celular, hace que no sea posible distinguir entre ambas patologías. A pesar de las diferencias clínicas entre neumonía y fibrosis, existen varias fases patológicas que podrían ocultar sus diferencias.

Por otro lado, la aplicación de algoritmos supervisados de clasificación de *machine learning* podría, tras un correcto entrenamiento y mayor validación de los algoritmos, proporcionar un sistema en tiempo real de biopsia electrónica y, de este modo, complementar los actuales métodos de diagnóstico.

Los resultados que se visualizan tanto en la **Figura 5** como en la **Figura 6** muestran una gran capacidad de clasificación de las muestras de neoplasia con un 86% de sensibilidad (6 de 7 muestras) para el árbol de decisión y un 100% para el algoritmo de Naïve Bayes. A diferencia de los resultados mostrados en las **Figura 1** a **Figura 4** donde no sería posible distinguir entre un tejido neoplásico de uno fibrótico.

5. Conclusión

El estudio de las diferencias en los parámetros de impedancia en frecuencias separadas en la región de dispersión beta debido a cambios en el estado del tejido pulmonar puede usarse para diferenciar entre diferentes patologías pulmonares. La diferencia en impedancia debida a la dispersión beta es mayor entre aquellas patologías que conducen a un aumento de tejido (neoplasia y fibrosis) y aquellas patologías que conducen a un mayor contenido de aire en los pulmones (enfisema). El uso de mediciones de bioimpedancia mínimamente invasivas para diferenciar patrones pulmonares tiene como objetivo ampliar las herramientas de diagnóstico disponibles actualmente en broncoscopia. Por ello, los algoritmos de *machine learning* pueden llegar a proporcionar un sistema de biopsia electrónica en tiempo real puesto que son capaces de identificar muestras correspondientes a una neoplasia con gran exactitud. Sin embargo, son necesarios más estudios para diferenciar las enfermedades pulmonares que son más similares entre sí.

Agradecimientos

Agradecer a los pacientes sin los cuales este estudio no habría sido posible. Además, agradecer a Marta Navarro Colom, Laura Romero Roca y Margarita Castro Jiménez de la Unidad de Neumología Intervencionista del Servicio de Neumología del Hospital de la Santa Creu i Sant Pau por su gran apoyo.

Este trabajo fue apoyado por el Ministerio español de Ciencia e Innovación (PID2021-128602OB-C21).

Referencias

- [1] H. C. Lukaski, 'Biological indexes considered in the derivation of the bioelectrical impedance analysis', *Am. J. Clin. Nutr.*, vol. 64, no. 3, pp. 397S-404S, Sep. 1996, doi: 10.1093/ajcn/64.3.397S.
- [2] H. C. Lukaski, N. Vega Diaz, A. Talluri, and L. Nescolarde, 'Classification of Hydration in Clinical Conditions: Indirect and Direct Approaches Using Bioimpedance', *Nutrients*, vol. 11, no. 4, p. 809, Apr. 2019, doi: 10.3390/nu11040809.
- [3] S. Khalil, M. Mohhtar, and F. Ibrahim, 'The Theory and Fundamentals of Bioimpedance Analysis in Clinical Status Monitoring and Diagnosis of Diseases', *Sensors*, vol. 14, no. 6, pp. 10895–10928, Jun. 2014, doi: 10.3390/s140610895.
- [4] H. P. Schwan, 'Electrical properties of tissues and cell suspensions: mechanisms and models', in *Proceedings of 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Baltimore, MD, USA, 1994, pp. A70–A71. doi: 10.1109/IEMBS.1994.412155.
- [5] G. Martinsen, S. Grimnes, and H. P. Schwan, 'INTERFACE PHENOMENA AND DIELECTRIC PROPERTIES OF BIOLOGICAL TISSUE', 2002.
- [6] H. P. Schwan and S. Takashima, 'Electrical Conduction and Dielectric Behavior in Biological Systems', in *digital Encyclopedia of Applied Physics*, Wiley-VCH Verlag GmbH & Co. KGaA, Ed. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, 2003, p. eap106. doi: 10.1002/3527600434.eap106.
- [7] K. R. V et al., 'Machine learning detects EEG microstate alterations in patients living with temporal lobe epilepsy', *Seizure*, vol. 61, pp. 8–13, Oct. 2018, doi: 10.1016/j.seizure.2018.07.007.
- [8] G. S. Chilla, L. Y. Yeow, Q. H. Chew, K. Sim, and K. N. B. Prakash, 'Machine learning classification of schizophrenia patients and healthy controls using diverse neuroanatomical markers and Ensemble methods', *Sci. Rep.*, vol. 12, no. 1, p. 2755, Feb. 2022, doi: 10.1038/s41598-022-06651-4.
- [9] M. Cikes et al., 'Machine learning-based phenogrouping in heart failure to identify responders to cardiac resynchronization therapy', *Eur. J. Heart Fail.*, vol. 21, pp. 74–85, 2019, doi: 10.1002/ejhf.1333.
- [10] L. Papp et al., 'Supervised machine learning enables non-invasive lesion characterization in primary prostate cancer with [68Ga]Ga-PSMA-11 PET/MRI', *Eur. J. Nucl. Med. Mol. Imaging*, vol. 48, no. 6, pp. 1795–1805, Jun. 2021, doi: 10.1007/s00259-020-05140-y.
- [11] S. Rastegar et al., 'Radiomics for classification of bone mineral loss: A machine learning study', *Diagn. Interv. Imaging*, vol. 101, no. 9, pp. 599–610, Sep. 2020, doi: 10.1016/j.diii.2020.01.008.
- [12] M. May, 'Eight ways machine learning is assisting medicine', *Nat. Med.*, vol. 27, no. 1, pp. 2–3, Jan. 2021, doi: 10.1038/s41591-020-01197-2.
- [13] G. Company-Se et al., "Differentiation using minimally-invasive bioimpedance measurements of healthy and pathological lung tissue through bronchoscopy," *Frontiers in Medicine*, vol. 10, Apr. 2023, doi: 10.3389/fmed.2023.1108237.
- [14] G. Company-Se et al., 'Effect of Calibration for Tissue Differentiation Between Healthy and Neoplasm Lung Using Minimally Invasive Electrical Impedance Spectroscopy', *IEEE Access*, vol. 10, pp. 103150–103163, 2022, doi: 10.1109/ACCESS.2022.3209809.

Señales Biomédicas (II)

Jueves 23 8:30 - 10:00

Salón de Grados

New Approach for the Assessment of Sarcopenia Using Core Shape Modelling

I. Junquera Godoy¹, J.L. Martínez-de-Juan¹, J. Gomis-Tena¹, J. Saiz¹, G. Prats-Boluda¹

¹ Centro de Investigación e Innovación en Bioingeniería, Universitat Politècnica de València, 46022 Valencia, Spain

{ijungod¹, jlmartinez¹, jgomiste¹, jsaiz¹, gprats¹} @ci2b.upv.es

Summary

Sarcopenia, the age-related decline in muscle mass and strength, poses diagnostic challenges due to limited, costly, and specialized evaluation methods. This study suggests using surface electromyography (sEMG) as a diagnostic tool. This research aims to evaluate the performance of the core shape modelling, a novel biomarker obtained from the sEMG signal for sarcopenia assessment. The study involved nine sarcopenic and eleven healthy individuals performing consecutive thirty-second sit-to-stand movements while a bipolar sEMG signal was recorded from the rectus femoris and biceps femoris muscles. Traditional sEMG parameters (root mean square and mean frequency) and core shape modelling features were analyzed. The statistical comparison revealed no significant differences in conventional parameters between groups for either muscle. However, the core shape modelling approach displayed significant distinctions, particularly in left shape distance (rectus femoris) and right shape distance (biceps femoris) between sarcopenic and control subjects. These findings suggest that the proposed biomarker holds promise in improving sarcopenia diagnosis.

1. Introduction

Rosenberg introduced the term "sarcopenia" to define this muscle mass decline [1]. Although sarcopenia is typically observed in older adults, it can also impact younger individuals with inflammatory ailments [2]. Sarcopenia encompasses not only muscle mass loss but also a decrease in muscle strength and function. This decline in physical function has far-reaching repercussions, contributing to postural instability, modified thermoregulation, compromised bone health, disrupted glucose equilibrium, and decreased basal energy production [3].

Sarcopenia's influence on individuals is noteworthy as both muscle mass and strength steadily decline. By the age of 50, a significant number of individuals experience a 10% reduction in muscle mass, with the potential for this decline to reach 70% by the time they are 70 years old [3]. Effectively addressing sarcopenia poses challenges, mainly due to the complexities involved in evaluating its three core components: muscle strength, muscle mass, and physical performance like walking speed. Conventional non-invasive gold standard techniques such as Computed Tomography (CT), Dual-energy X-ray Absorptiometry (DXA), and Magnetic Resonance Imaging (MRI) [4] are infrequently used in clinical practice due to their lack of portability, high equipment costs, and the need for specialized personnel. Surface electromyography (sEMG) is a non-invasive measurement technique that shows promise for monitoring muscle fatigue and assessing muscle function. EMG signals

represent the electrical activity generated by the muscle's motor units during contractions. The sEMG method is safer and less invasive than the intramuscular technique, utilizing non-invasive surface electrodes [5].

While numerous investigations have explored EMG signal usage in various medical contexts, limited attention has been dedicated to employing sEMG for monitoring and assessing sarcopenia. Much of the existing research tends to focus on scrutinizing muscle age disparities rather than sarcopenia per se [6,7]. Consequently, there exists a vital necessity to explore the potential of sEMG as a valuable diagnostic tool for sarcopenia. Some studies have delved into the temporal and spectral characteristics of sEMG signals during specific muscle activities, unearthing potential sarcopenia markers [8]. However, the evaluation was limited to time (root mean square) and frequency (instantaneous median frequency) characteristics of the sEMG signals.

Recent research has identified potential alterations in the probability density function (PDF) of surface electromyographic (sEMG) data in contexts, such as fatigue and muscle force increase. Boudaoud et al [9] proposed a new method to analyze PDF shape modifications using the Core Shape Modelling (CSM). The functional statistics employed in this study utilize kernel density estimation and shape distances of probability density functions (PDFs) to assess modifications in shape. This method was proven to be more robust than other shape parameters, such as kurtosis or skewness especially when dealing with small sample sizes. In a prior study [7], this technique was employed to assess muscle fatigue across distinct age groups, revealing notable distinctions between older and younger participants.

The main purpose of this study is to evaluate the performance of the Core Shape Modelling as a novel sEMG biomarker for the diagnosis of sarcopenia, in comparison with traditional biomarkers.

2. Materials and methods

2.1. Participants

Twenty participants, 10 male and 10 female volunteered for the experiment where nine out of twenty participants were diagnosed with sarcopenia. Groups present homogeneous characteristics, (age: 72.7 ± 7.6 years; weight: 71.1 ± 12.4 kg; height: 1.59 ± 0.1 m). Candidates needed to fulfill specific criteria: being 60 years of age or older and having no substantial underlying health conditions that might influence the testing procedure. Before engaging in the experiments, all

participants provided documented informed consent. The study was conducted at the Dawako clinic.

2.2. Experimental protocol and data acquisition

The protocol consisted of one simple task. Participants were directed to execute consecutive sit-to-stand movements using only their legs, abstaining from using their hands, for a 30-second interval. They were encouraged to complete as many repetitions as possible within the given time frame.

A pair of adhesive electrodes were placed in each muscle following SENIAM recommendations for the preparation of the skin, muscle fiber direction, and positioning [10]. The sEMG signals were recorded using the BTS-FREEMG device (BTS-FREEMG1000; BTS Bioengineering, Milan, Italy). The device was set to a sampling rate of 1000 Hz per channel. Myoelectric signals were recorded simultaneously from the rectus femoris (RF) and biceps femoris (BF) muscles of the dominant leg. The motion artifacts and high-frequency noises were removed using a zero-phase, Butterworth bandpass filter (10–500 Hz). Afterward, assisted segmentation was applied to obtain muscle activations from the signals [11]. Following the segmentation of the signal, sEMG features were extracted in each activation throughout the exercise. Then, the average value is calculated for each parameter of each subject.

2.3. Temporal analysis

The amplitude of sEMG signals is determined by various factors, including the number of active motor units, their discharge rates, and the characteristics of the intracellular action potentials (IAPs) [12,13]. The parameter used in this study to quantify the intensity of the sEMG is the root mean squared value (RMS).

2.4. Spectral analysis

Spectral analysis of sEMG signals provides insights into muscle fatigue, identifying displacement of the spectrum to low frequencies [14]. It also aids in studying motor control and coordination by analyzing frequency characteristics during different tasks. The parameter computed to analyze these characteristics is the mean frequency of the power spectral density (MNF) [15].

2.5. Shape analysis

Three distances are obtained from the CSM method, using equations 1-3. The Center Shape Distance (CSD), which evaluates PDF peakedness, the Left Shape Distance (LSD), and the Right Shape Distance (RSD), which both evaluate PDF asymmetry [8].

$$CSD_{(p,q)} = \sqrt{\int_{0.4}^{0.6} (\alpha P^{-1}(y) + \beta - G^{-1}(y))^2 dy} \quad (1)$$

$$LSD_{(p,q)} = \sqrt{\int_0^{0.25} (\alpha P^{-1}(y) + \beta - G^{-1}(y))^2 dy} \quad (2)$$

$$RSD_{(p,q)} = \sqrt{\int_{0.75}^1 (\alpha P^{-1}(y) + \beta - G^{-1}(y))^2 dy} \quad (3)$$

Where P and G are the PDF of the sEMG signal and a Gaussian PDF, respectively. In the continuous domain this PDFs are defined as p and q. The α and β values are approximated using a constrained linear regression between $P^{-1}(x)$ and $G^{-1}(x)$.

2.6. Statistical analysis

The Wilcoxon Rank Sum test, a nonparametric statistical test, is used to compare both groups and evaluate the performance of each parameter as a possible biomarker for the detection of sarcopenia. The level of significance was set to 0.05.

3. Results

The following results illustrate the performance of the sEMG features for the comparison between sarcopenic and control subjects. The box and whisker plots depict the distribution of parameter values for each group (sarcopenic and control) and each muscle (RF and BF), providing insight into the variability and differences between the two groups.

Fig. 1 illustrates the results obtained from the RMS parameter when comparing sarcopenic and control groups. Although the results are not statistically significant, it can be observed that the sarcopenic group in BF presents sEMG signals with higher amplitude than the control group.

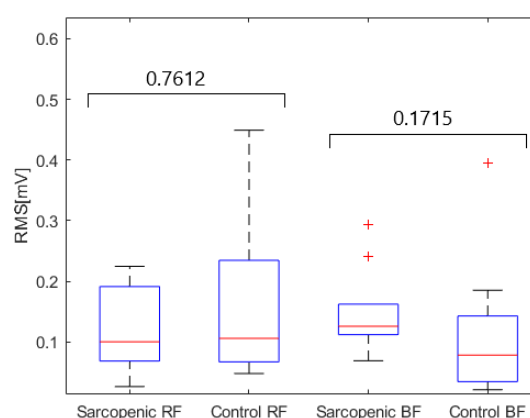


Fig. 1: Box & Whisker plots and p-values of the RMS for both muscles (RF and BF) and groups (sarcopenic and control). P-values, being (*) statistically significant.

Regarding the results obtained from the spectral analysis, Fig. 2 represents the results from MNF. Although the results for RF are very similar for both groups, BF does present differences. For this muscle, the obtained p-value is close to the limit, but still not statistically relevant. Evaluating this parameter, sarcopenic participants presented lower values in BF compared to controls.

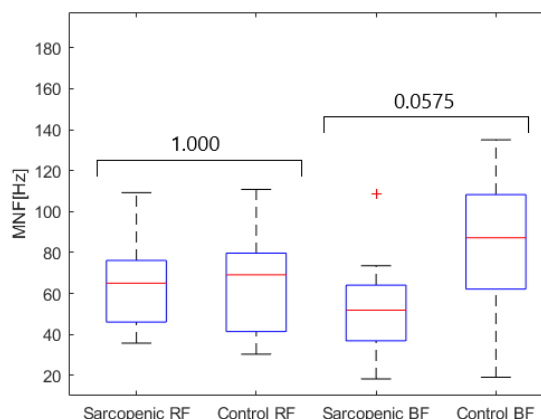


Fig. 2: Box & Whisker plots and p-values of the MNF for both muscles (RF and BF) and groups (sarcopenic and control). P-values, being (*) statistically significant.

Fig. 3 illustrates the PDFs of the sEMG signals from the sarcopenic and control groups, juxtaposed with a Gaussian PDF, for the RF muscle. Notably, there is a more pronounced deviation from the Gaussian distribution (black) in the sarcopenic group (red), particularly on the left side of the curve, when compared to the controls (green). This disparity in deviation is subsequently mirrored in the calculated distances, specifically the LSD within the RF muscle. Importantly, this difference is further substantiated by the presence of statistically significant distinctions between the groups in the computed LSD values. The same evaluation can be done with the BF muscle.

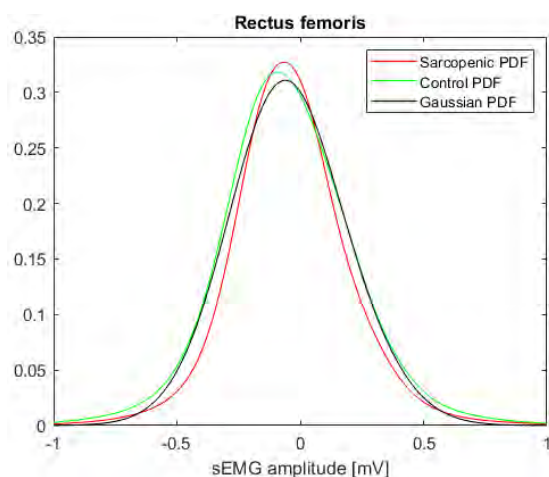


Fig. 3. Average PDFs from sarcopenic (red) and control (green) sEMG signals from RF muscle, juxtaposed with gaussian PDF (black).

Concerning the shape differences between the sEMG PDF from Gaussian PDF, the first distance of evaluation is the CSD in Fig. 4. This distance does not present differences between groups for either muscle.

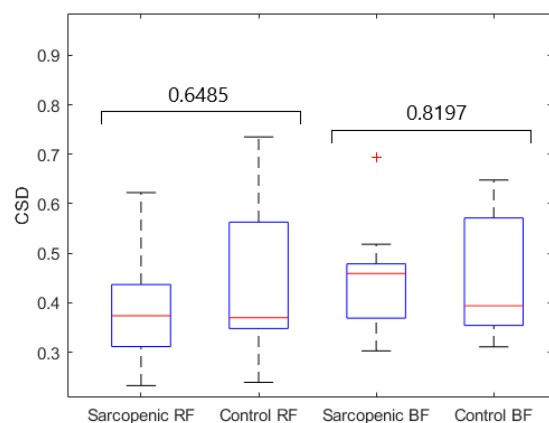


Fig. 4: Box & Whisker plots and p-values of the CSD for both muscles (RF and BF) and groups (sarcopenic and control). P-values, being (*) statistically significant.

Examining LSD results in Fig. 5, the distance is somewhat higher in the sarcopenic group, though not significantly different in the BF muscle. In the RF muscle, the values are also higher in the sarcopenic group, and in this case, the differences are statistically significant.

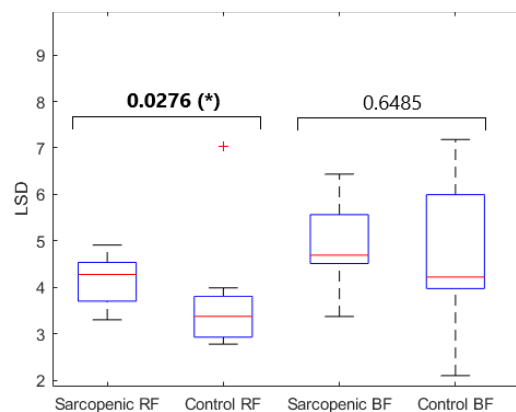


Fig. 5. Box & Whisker plots and p-values of the LSD for both muscles (RF and BF) and groups (sarcopenic and control). P-values, being (*) statistically significant.

Finally, evaluating the results of the last parameter computed in this study, in Fig. 6 it is also noticeable the higher values of the RSD distance for the sarcopenic group in both muscles, emphasizing the significant difference in BF muscle.

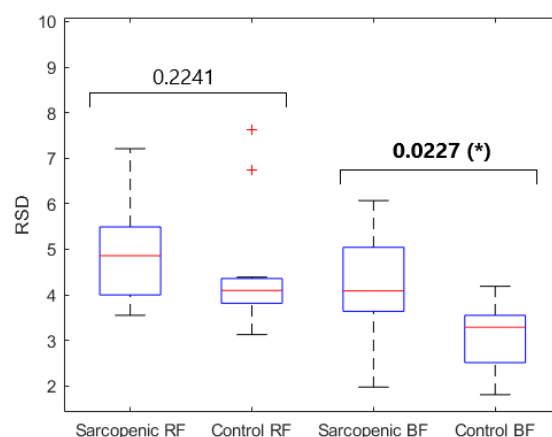


Fig. 6. Box & Whisker plots and p-values of the RSD for both muscles (RF and BF) and groups (sarcopenic and control). P-values, being (*) statistically significant.

4. Discussion

This study aimed to evaluate and compare the effectiveness of a newly introduced sEMG parameter alongside traditional measures, with the potential to serve as a detection biomarker for sarcopenia. None of the traditional parameters evaluated in this study (RMS and MNF) presented statistically significant differences between groups.

However, although not significant, some differences can be found in the BF muscle between groups, for these traditional parameters. On the other hand, no observable differences can be found in RF muscle. This disparity of results between muscles could be due to the very act of performing the exercise, or the difference in muscle size, among other possibilities. Sarcopenic patients show higher RMS values compared to the control group. This might be caused by the need for higher muscle recruitment by the sarcopenia group to do the same task as the control group [15].

Regarding the results obtained from BF muscle, in the spectral analysis, it can be observable lower values for sarcopenic participants in MNF. These results can be

attributed to several causes. One potential explanation is that the sarcopenic group may have a higher proportion of slow-twitch muscle fibers, which tend to generate lower-frequency electrical signals during muscle contractions. Additionally, it could imply reduced motor unit recruitment and increased muscle fatigue in sarcopenic individuals, leading to lower frequency sEMG signals [16].

Conversely, the parameters derived from CSM exhibited robust performance, better than the traditional parameters. While the CSD did not uncover any distinctions, the parameters assessing signal asymmetry, such as RSD and LSD demonstrated a consistent trend in both muscles. These trends manifested as elevated computed distance values within the sarcopenic group, indicating that the PDF of this group deviated more significantly from the Gaussian distribution than that of the control group. Notably, these findings reached statistical significance for RSD in the BF muscle and LSD in the RF muscle. The main limitation of this study was the reduced number of participants. Further investigation needs to be done with a more extensive database to corroborate these findings.

5. Conclusions

This study aimed to assess novel sEMG parameters alongside traditional measures to identify their potential as a sarcopenia biomarker. Specifically, the sarcopenic group presented higher RMS values and lower MNF values compared to the control group, suggesting heightened motor recruitment and increased muscle fatigue but without statistically significant differences between groups. By contrast, RSD and LSD distances derived from CSM exhibited statistically significant differences between the groups. These parameters, assessing signal asymmetry, showed higher values for sarcopenic participants compared to control ones. Additionally, though traditional RMS and MNF parameters lacked statistical significance, they showcased noteworthy contrasts within the BF muscle.

Acknowledgments

This study was supported by the Agència Valenciana de la Innovació project ‘iSARC-GENETICS, Plataforma Inteligente para la Detección Temprana, Ayuda al Diagnóstico y Seguimiento de la Sarcopenia’. (INNEST/2021/365).

References

- [1] Rosenberg IH. Sarcopenia: Origins and clinical relevance. *J Nutr.* 1997;127(5 Suppl):990S-991S. doi: 10.1093/jn/127.5.990S
- [2] Evans WJ, Campbell WW. Sarcopenia and age-related changes in body composition and functional capacity. *J Nutr.* 1993;123(2 Suppl):465-468. doi: 10.1093/jn/123.suppl_2.465
- [3] A. Cruz-Jentoft, J. Baeyens, J. Bauer, Y. Boirie, T. Cederholm, F. Landiet al., "Sarcopenia: European consensus on definition and diagnosis", *Age and Ageing*, vol. 39, no. 4, p. 412-423, 2010. <https://doi.org/10.1093/ageing/afq034>
- [4] Morley JE. Sarcopenia: diagnosis and treatment. *J Nutr Health Aging.* 2008 Aug-Sep;12(7):452-6. doi: 10.1007/BF02982705. PMID: 18615226.
- [5] Merletti R, Parker PJ. *Electromyography: physiology, engineering, and noninvasive applications.* Wiley-IEEE Press; 2004.
- [6] L. Imrani, S. Boudaoud, J. Laforêt, and K. Kinugawa, "Ageing effect evaluation on HD-SEMG signals using CCA approach," *IRBM*, vol. 42, no. 4, pp. 294-299, 2021. doi:10.1016/j.irbm.2021.05.002
- [7] Bharath Krishnan a b et al. (2022) Age-sensitive high density surface electromyogram indices for detecting muscle fatigue using core shape modelling, *Biomedical Signal Processing and Control.* Available at: <https://www.sciencedirect.com/science/article/pii/S1746809422009004> (Accessed: 22 August 2023).
- [8] Habenicht, R., Ebenbichler, G., Bonato, P., Kollmitzer, J., Ziegelbecker, S., Unterlerchner, L., ... & Kienbacher, T. (2020). age-specific differences in the time-frequency representation of surface electromyographic data recorded during a submaximal cyclic back extension exercise: a promising biomarker to detect early signs of sarcopenia. *J NeuroEngineering Rehabil.* 1(17). <https://doi.org/10.1186/s12984-020-0645-2>
- [9] S. Boudaoud, H. Rix, M. A. Harrach and F. Marin, "Robust functional statistics applied to probability density function shape screening of sEMG data," 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA, 2014, pp. 2213-2216, doi: 10.1109/EMBC.2014.6944058
- [10] Stegeman, Dick & Hermens, Hermie. (2007). Standards for surface electromyography: The European project Surface EMG for non-invasive assessment of muscles (SENIAM). 1.
- [11] Iranzo, S., Piedrabuena, A., García Torres, F., Martínez de Juan, J. L., Prats Boluda, G., Sanchis, M., & Belda Lois, J. M. (2022). Assessment of a passive lumbar exoskeleton in material manual handling tasks under laboratory conditions. <https://doi.org/10.20944/preprints202204.0024.v1>
- [12] Moritani T, Muro M, Nagata A. Intramuscular and surface electromyogram changes during muscle fatigue. *J Appl Physiol* (1985). 1986 Apr;60(4):1179-85. doi: 10.1152/jappl.1986.60.4.1179. PMID: 3700300.
- [13] Dimitrova NA, Dimitrov GV. Amplitude-related characteristics of motor unit and Mwave potentials during fatigue. A simulation study using literature data on intracellular potential changes found in vitro. *J Electromyogr Kinesiol* 2002;12(5):339-49.
- [14] B. Gerdle, B. Larsson, and S. Karlsson, "Criterion validation of surface EMG variables as fatigue indicators using peak torque," *Journal of Electromyography and Kinesiology*, vol. 10, no. 4, pp. 225-232, 2000. doi:10.1016/s1050-6411(00)00011-0
- [15] A. Zaman, T. Sharmin, M. Ali Khan, and M. Ferdjallah, "Muscle Fatigue Analysis in young adults at different MVC levels using EMG metrics," *Proceedings 2007 IEEE SoutheastCon*, 2007. doi:10.1109/secon.2007.342930
- [16] A. Phinyomark, S. Thongpanja, H. Hu, P. Phukpattaranont, and C. Limsakul, "The usefulness of mean and median frequencies in electromyography analysis," *Computational Intelligence in Electromyography Analysis - A Perspective on Current Applications and Future Challenges*, 2012. doi:10.5772/50639

Comparación de métodos manuales para el estudio del jitter neurofisiológico

C. Valle Morgado¹, A. Malanda Triguero¹, O. Garnés-Camarena², D. W. Stashuk³

¹ Departamento Ingeniería Eléctrica, Electrónica y de Comunicación, Universidad Pública de Navarra, Pamplona, España, valle.108450@e.unavarra.es, malanda@unavarra.es

² Servicio de Neurofisiología Clínica, Hospital de la Fundación Jiménez Díaz, Madrid, España, oscar.garnes@quironsalud.es

³ Department of Systems Design Engineering, University of Waterloo, Ontario, Canada, stashuk@uwaterloo.ca

Resumen

La diferencia temporal con la que se activan las fibras musculares de una motoneurona puede variar en sucesivas contracciones. Esta variabilidad se conoce como jitter neuromuscular y resulta aumentada en diversas patologías neurógenas y de la unidad motora. En la actualidad los equipos comerciales que se utilizan para evaluar el jitter (Keypoint y Synergy) implementan dos métodos distintos para realizar la medida: el método de los picos y el método de los flancos. Sin embargo, no existen evidencias de que las medidas obtenidas con ambos métodos ofrezcan resultados comparables. Para realizar este estudio se ha utilizado un banco de señales con 184 registros. Todos los registros han sido alineados por alguno de los picos utilizando el método de los picos y posteriormente el de los flancos. Una vez alineado, sobre los demás picos presentes en el registro, se han tomado medidas de jitter utilizando ambos métodos (picos y flancos), de manera que por cada par de picos válidos presentes se han obtenido cuatro medidas de jitter distintas. Para comparar si existían diferencias estadísticamente significativas entre las medidas registradas para cada una de las cuatro configuraciones posibles se ha utilizado el test de Wilcoxon. El resultado que se obtiene es que únicamente no existen diferencias significativas para las medias obtenidas con la configuración pico-flanco y flanco-pico. Para el resto de configuraciones existen diferencias estadísticamente significativas, lo cual demanda nuevos estudios para determinar qué métodos son los más precisos desde el punto de vista clínico.

1. Introducción

Una unidad motora (UM) está compuesta por una neurona motora y las fibras musculares que su axón inerva. Cuando la neurona es estimulada un impulso nervioso o potencial de acción se transmite a través de su axón. Cuando este potencial se transmite a cada fibra muscular, si su nivel de excitación está por encima del umbral de activación de la fibra, se produce la despolarización de la misma. Esta despolarización es registrada con electrodos en el medio extracelular generando un potencial de fibra muscular (PFM) por cada una de las fibras que compone la UM. La suma de todos los PFM es lo que se conoce como potencial de acción de unidad motora (PAUM). En las contracciones musculares la unidad motora es activada sucesivamente registrándose PAUMs consecutivos que es lo que se conoce como tren de potenciales de acción de la unidad motora (TPAUM).

En condiciones clínicas normales todas las fibras que componen una unidad motora son activadas de forma más

o menos sincrónica. Sin embargo, ante ciertas patologías, existe una pequeña variabilidad entre los instantes de activación de las distintas fibras. Esto se debe principalmente al tiempo de transmisión del potencial hacia la fibra a través de la unión neuromuscular. Debido a esto aparece un intervalo entre los potenciales de cada par de fibras de la unidad motora (IPI). A la medida de la variabilidad de los IPI en los potenciales sucesivos de un TPAUM se le conoce técnicamente como “jitter” [1]. Un incremento de esta variabilidad tiene origen fisiológico y puede deberse a enfermedades de la unión neuromuscular tales como la miastenia gravis, el síndrome de miastenia congénita, el síndrome miasténico de Lambert-Eaton, etc. [1] [3]. Tradicionalmente los estudios de jitter se han realizado por la técnica de electromiografía de fibra simple (EMGFS) que consiste en registrar un PAUM que contenga únicamente dos PFM (Figura 1) conocidos como potenciales de fibra simple (PFS).

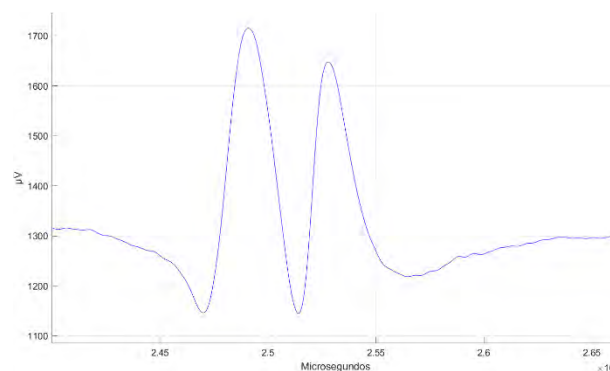


Figura 1. PAUM compuesto de dos PFM.

Estos potenciales tienen que estar lo suficientemente separados como para que no exista solapamiento y se pueda medir el intervalo que hay entre ellos (IPI).

Si se producen repetidas contracciones y se registran los sucesivos PAUMs se obtiene un TPAUM en el que se puede medir el jitter. Éste se calcula como la media de las diferencias de IPIs consecutivos (MCD) [1] (Figura 2.A).

$$MCD = \frac{|IPI_1 - IPI_2| + |IPI_2 - IPI_3| + \dots + |IPI_{N-1} - IPI_N|}{N - 1}$$

(1)

Donde IPI_i es el intervalo entre dos PFM y N es el número de PAUMs que forman del tren.

En nuestro estudio, como todos los potenciales de un TPAUM se alinean por alguno de los picos o flancos, para computar el MCD, en lugar de tomar la medida de los IPIs (Figura 2.A), basta con obtener los instantes temporales (t_i , $i=1\dots N$) en los que aparece un segundo pico (Figura 2.B) o flanco (Figura 2.C). El MCD se obtiene de forma equivalente como:

$$MCD = \frac{|t_1 - t_2| + |t_2 - t_3| + |t_3 - t_4| + \dots + |t_{N-1} - t_N|}{N - 1} \quad (2)$$

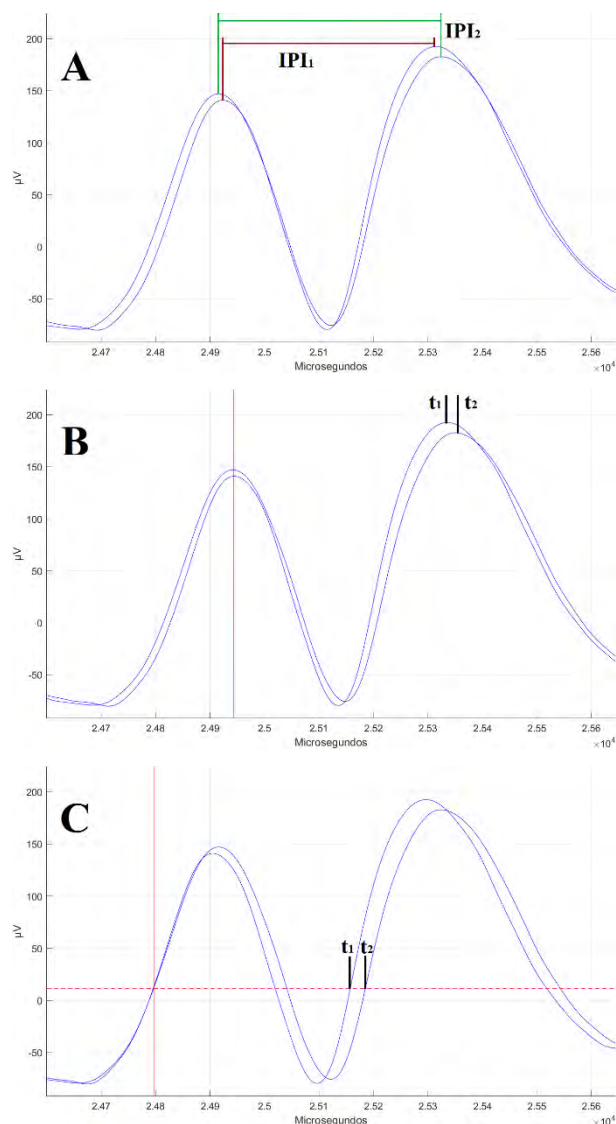


Figura 2. (A) IPIs de un TPAUM con dos potenciales. (B) TPAUM alineado por el primer pico. Instantes temporales para calcular el MCD asociados al segundo pico. (C) TPAUM alineado por el primer flanco. Instantes temporales para calcular el MCD asociados al segundo flanco.

Sanders [3] estableció cuantitativamente las condiciones para poder considerar potenciales válidos para la medida de jitter: una amplitud pico-pico (V_{pp}) entre 200 μV y 20 mV, un tiempo de subida inferior a 300 μs , una separación entre picos de al menos 150 μs y la ausencia de artefactos

conocidos como “shoulders”, referidos al cabalgamiento de dos picos consecutivos.

En la actualidad los equipos comerciales más utilizados para realizar los estudios de jitter emplean el método de los picos o el de los flancos para obtener las medidas. Sin embargo, no existen estudios que comparen los resultados obtenidos con ambos métodos. Por este motivo, en este trabajo se comparan las medidas que se obtienen con los métodos de picos y flancos para así determinar si existen diferencias estadísticamente significativas en los valores de jitter obtenidos con ambos métodos.

2. Material y métodos

2.1. Registros

Para la elaboración del estudio se utilizó un banco de señales de EMG con 184 registros tomados de 17 pacientes con patologías de la unión neuromuscular o neuropatías y recogidos en el Hospital Universitario de la Fundación Jiménez Díaz de Madrid. Tanto las extracciones como el resto del estudio fueron aprobados por el Comité Ético de este Hospital. Para el registro de los potenciales se utilizaron agujas concéntricas faciales con un área de registro de 0.019 mm² y un equipo Keypoint v.3.22. Las señales en bruto fueron filtradas con un filtro paso banda con frecuencias de corte de 500 ó 1000 Hz y 20 kHz, muestreadas a 48 kHz, exportadas en ficheros anonimizados y finalmente procesadas con el software DQEMG [2] para su descomposición en trenes de PAUMs, extrayéndose un total de 184 TPAUMs. Estas señales fueron sobremuestreadas hasta 192 kHz para aumentar su resolución y permitir una medición precisa del jitter. Del total de los 184 TPAUMs extraídos se descartaron 51: 30 de ellos por no presentar un par de picos válidos para realizar la medida que cumplen los criterios establecidos en [3], 11 por no haber una separación entre picos de al menos 150 μs , 5 por presentar únicamente potenciales con un solo pico y 5 por no tener al menos un 50 % de potenciales válidos para realizar la medida. Estos criterios constituyen estándares muy exigentes, convenidos en la investigación, pero que no siempre se llegan a cumplir en la práctica clínica, debido a las limitaciones del tiempo de exploración del paciente.

Para reproducir manualmente las medidas de jitter en registros de EMG tal como se realiza con los equipos profesionales, se utilizó una herramienta desarrollada por los autores y basada en Matlab®, que implementa los métodos de medida de picos y flancos [5].

2.2. Selección de potenciales

Para cada tren se realiza una evaluación manual previa que permite descartar aquellos potenciales que resultan desaconejables para el estudio: potenciales con presencia de bloqueos, que no sigan la forma de onda del conjunto de potenciales del tren o que no cumplan con los criterios mencionados de V_{pp} , tiempo de subida de flanco, etc. En nuestro caso, cuando el número de potenciales descartados superó la mitad del total se descartó el tren.

2.3. Alineamiento

Los potenciales de los TPAUMs son alineados antes de realizar las medidas. Este alineamiento se realiza sobre alguno de los picos que cumpla con los criterios establecidos. Dicho pico se selecciona dentro de los márgenes de una ventana de longitud variable y establecida por el operador. Una vez seleccionado se aplica alguno de los métodos (picos o flancos) para alinear los potenciales.

En el método de los picos cada PAUM del tren es desplazado hacia su derecha o izquierda de tal manera que el pico seleccionado para el alineamiento ocurra en idéntico instante temporal en todos los PAUMs del tren (Figura 2.B). En el método de los flancos el operador establece un umbral de amplitud (recta horizontal punteada en la Figura 2.C) y cada PAUM del tren es desplazado hacia su derecha o izquierda de tal manera que el flanco de subida del pico seleccionado de este PAUM corte el umbral de amplitud en el mismo instante temporal para todos los PAUMs del tren (Figura 2.C).

2.4. Medida del jitter

Una vez alineado el TPAUM (con el método de picos o flancos) se eligen los picos para realizar la medida cumpliendo los criterios establecidos por Sanders.

Una vez seleccionado el pico donde realizar la medida se utiliza el método de picos para obtener una medida de jitter. Con este método se obtienen las referencias temporales (t_1, t_2, \dots, t_N) del valor máximo de cada potencial para ese pico concreto (Figura 2.B). Con esos instantes temporales se obtiene el jitter utilizando la definición (2) de MCD. A continuación, sobre el mismo pico, se aplica el método de flancos, para lo cual el operario define de forma manual un umbral (Figura 2.C) de acuerdo con su experiencia clínica. Con este método se obtienen las referencias temporales (t_1, t_2, \dots, t_N) del instante en que el umbral corta a cada potencial. Con estas referencias temporales se obtiene otra media de jitter.

En nuestro estudio este procedimiento se repitió para cada uno de los picos seleccionados para ser medidos. Además, se volvió a alinear el TPAUM utilizando el método de los flancos y se repitió nuevamente todo el proceso: se seleccionaron uno a uno los mismos picos que en el caso anterior y se aplicaron los métodos de picos y flancos para obtener nuevamente dos medidas de jitter. De esta manera para cada pico seleccionado para el análisis se obtuvieron cuatro medidas de jitter distintas y que son el resultado de las cuatro configuraciones alineamiento-medida posibles (pico-pico, pico-flanco, flanco-pico y flanco-flanco). Todo este proceso se repitió para los 184 TPAUMs disponibles.

2.5. Métodos estadísticos

Para comprobar la distribución estadística que seguían los datos de los valores de jitter obtenidos con cada una de las cuatro configuraciones posibles y de las diferencias existentes entre todas las configuraciones se utilizó el test de Anderson-Darling ($\alpha=0,05$). Para determinar si existían diferencias estadísticamente significativas entre todas las configuraciones de alineamiento-medida posibles se utilizó el test de Wilcoxon ($\alpha=0,05$).

3. Resultados

De los 133 TPAUMs que se seleccionaron para el estudio se obtuvieron un total de 157 pares de picos válidos para la medida (algunos de los trenes presentaban más de un pico sobre el que poder medir). Para cada uno de estos pares de picos se obtuvieron cuatro valores de MCD, uno por cada configuración alineamiento-medida (Figura 3 y Tabla 1).

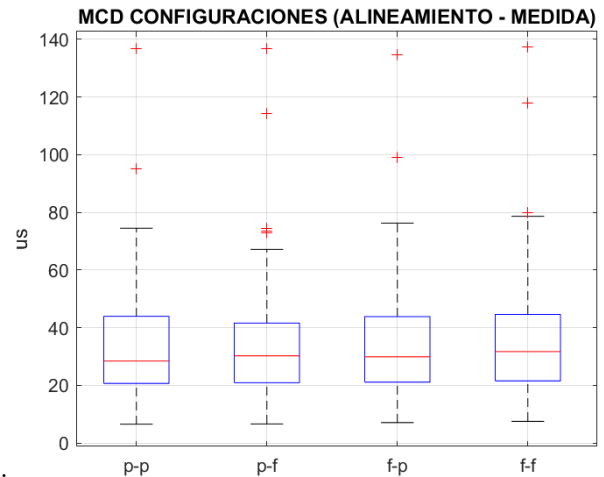


Figura 3. Boxplots de los valores de MCD obtenidos para cada configuración alineamiento-medida: pico-pico, pico-flanco, flanco-pico y flanco-flanco.

Alineamiento	Medida	Mediana (μs)	Percentiles [25 th, 75 th] (μs)	N.º Outliers
Pico	Pico	28,4	[20,7, 43,9]	2 (1,5 %)
Pico	Flanco	30,2	[20,9, 41,6]	5 (3,8 %)
Flanco	Pico	29,9	[21,1, 43,8]	2 (1,5 %)
Flanco	Flanco	31,7	[21,5 44,6]	3 (2,3 %)

Tabla 1. Estadísticas de los valores de MCD obtenidos para cada configuración alineamiento-medida.

Los resultados muestran que los rangos, las medianas y las distribuciones estadísticas de los valores de MCD de cada configuración son similares. Para evaluar las posibles desigualdades entre las distintas configuraciones se obtuvieron las diferencias entre ellas (Figura 4 y Tabla 2).

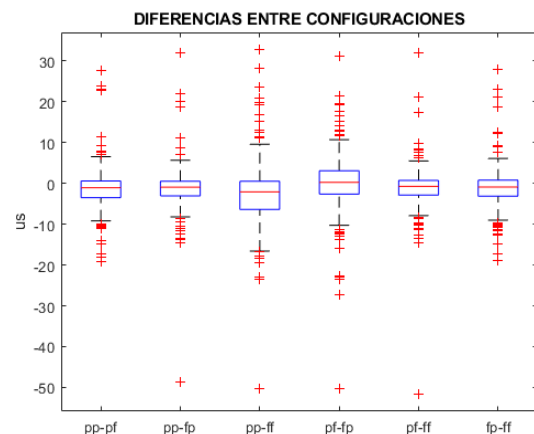


Figura 4. Boxplots de las diferencias entre las medidas de MCD obtenidas con cada configuración alineamiento-medida.

Configuración	Mediana (µs)	Percentiles [25 th, 75 th] (µs)	N.º Outliers
pp-pf	-1,07	[-3,39 0,91]	12 (7,8 %)
pp-fp	-0,97	[-3,29 0,67]	11 (7,1 %)
pp-ff	-2,21	[-3,43 3,07]	12 (7,8 %)
pf-fp	0,17	[-6,47 0,49]	9 (5,8 %)
pf-ff	-0,83	[-3,17 0,51]	16 (10,4 %)
fp-ff	-0,94	[-3,73 0,66]	13 (8,4 %)

Tabla 2. Estadísticas de los valores de las diferencias de MCD obtenidos entre las distintas configuraciones alineamiento-medida.

Las diferencias de las medidas que se obtuvieron entre las cuatro configuraciones posibles dejan ver que no existen diferencias importantes entre los valores obtenidos con cada una de ellas. Las medianas siempre están muy próximas a cero y comprendidas, en términos absolutos, entre 0,17 y 2,21 µs lo que solo representaba un error de 0,34 y 4,42 % con respecto a una situación de jitter patológico, el cual se considera a partir de los 40 µs. Se observa también que el rango de los valores para las diferencias de las configuraciones pp-pf, pp-fp, pf-ff y fp-ff es muy similares entre sí y aproximadamente la mitad del valor del rango de las diferencias de las configuraciones pp-ff y pf-fp y que está entre 6,96 y 6,5 µs respectivamente.

Como los datos registrados para las distintas configuraciones no seguían una distribución normal según el test de Anderson-Darling con un nivel de significación de 0.05 (p-valor = 0) y además se trataba de datos emparejados se utilizó el test de Wilcoxon ($\alpha = 0,05$) para determinar si las diferencias de los valores de jitter obtenidos para cada una de las cuatro configuraciones posibles de alineamiento-medida eran estadísticamente significativas (Tabla 3).

Configuración	Mediana (µs)	Intervalo confianza (µs)	p-valor
pp-pf	-1,22	[-1,77, -0,75]	1·10 ⁻⁶
pp-fp	-1,05	[-1,5, -0,61]	5·10 ⁻⁶
pp-ff	-2,24	[-3,01, -1,43]	2·10 ⁻⁷
pf-fp	0,28	[-0,47, 0,98]	0,47
pf-ff	-0,86	[-1,32, -0,43]	14·10 ⁻⁵
fp-ff	-1,09	[-1,65, -0,57]	12·10 ⁻⁵

Tabla 3. Resultados estadísticos del test de Wilcoxon aplicado a las diferencias entre configuraciones.

Con un nivel de significación de 0.05 se obtuvo que no existen diferencias estadísticas significativas entre las configuraciones pico-flanco y flanco-pico (p-valor = 0,47). Para las demás configuraciones existen diferencias estadísticas significativas siendo la diferencia obtenida entre las configuraciones pico-pico y flanco-flanco para la que se obtiene el intervalo de confianza más amplio [-3.01, -1.43] µs. Para este caso se obtiene también la mayor estimación de la mediana de las diferencias, -2.24 µs.

4. Conclusiones y líneas futuras

1. Se ha realizado un estudio para determinar si existen diferencias estadísticamente significativas entre los métodos que implementan los equipos médicos actuales para realizar los estudios de jitter: el método de picos y el de flancos. Para llevarlo a cabo se ha reproducido el proceso de alineamiento y el de medida de los TPAUM. Los métodos objetos del estudio han sido aplicados tanto en el proceso de alineamiento como en la medida de manera que para cada uno de los picos presentes en los TPAUM y válidos para realizar la medida se han obtenido cuatro medidas distintas. Las medidas obtenidas con cada una de las cuatro configuraciones alineamiento-medida posibles han sido comparadas estadísticamente.
2. De esa comparación se concluye que no existen diferencias significativas entre las medidas obtenidas con las configuraciones pico-flanco y flanco-pico. También se concluye que existen diferencias estadísticamente significativas entre el resto de configuraciones: pico-pico y pico-flanco, pico-pico y flanco-pico, pico-pico y flanco-flanco, pico-flanco y flanco-flanco y flanco-pico y flanco-flanco.
3. En particular las configuraciones que tradicionalmente implementan los equipos profesionales utilizados en neurofisiología clínica para medir el jitter (pico-pico y flanco-flanco) ofrecen medidas distintas en algunos µs.
4. El estudio para conocer qué método ofrece un resultado más fiable se debería realizar sobre señales simuladas, ya que es la única forma de conocer a priori el valor de jitter presente en el tren.
5. Para conocer con más profundidad como varían las medidas obtenidas con ambos métodos sería conveniente realizar estudios sobre músculos concretos. Esto permitiría conocer de qué manera afecta el tipo de músculo estudiado a la variabilidad de las medidas. Con el mismo fin también se podrían realizar estudios con patologías concretas.

Referencias

- [1] J.V. Trontelj, E. Stålberg. Single fiber and macro-Electromyography, in Clinical evaluation for diagnostics tests for neuromuscular disorders. T. E. Bertorini, Ed. Amsterdam, The Netherlands: Butterworth Henemann, 2002, pp. 417-447.
- [2] T.J. Doherty, D.W. Stashuk. Decomposition-based quantitative electromyography: methods and initial normative data in five muscles. Muscle Nerve. 2003 Aug;28(2):204-11.
- [3] D.B. Sanders, J.A. Kouyoumdjian, E. Stålberg. Single fiber electromyography and measuring jitter with concentric needle electrodes. Muscle Nerve. 2022 Jun 13.
- [4] A. Malanda, D. W. Stashuk, J. Navallas, J. Rodríguez-Falces, I. Rodríguez-Carreño, C. Valle, O. Garnés-Camarena. Automatic jitter measurement in needle-detected motor unit potential trains. Computers in Biology and Medicine. 149:105973. 2022.
- [5] C. Valle, A. Malanda, O. Garnés-Camarena, D. W. Stashuk. Interfaz para el estudio del jitter neuromuscular, CASEIB 2021.

Análisis del proceso de llenado de la señal sEMG a medida que aumenta gradualmente la fuerza en el cuádriceps

J. Rodríguez Falces¹, A. Malanda Trigueros¹, Cristina Mariscal², J. Navallas Irujo¹

¹ Dpto. Ing. Eléctrica y Electrónica, Universidad Pública de Navarra, Pamplona, España. javier.rodriguez@unavarra.es

² Departamento de Neurofisiología Clínica, Hospital Universitario de Navarra, Pamplona, Navarra, 31008 España.

Resumen

Objetivos: No existe una comprensión completa del modo en que la señal EMG de superficie se llena progresivamente de potenciales de unidad motora (MUP) a medida que aumenta la fuerza. Intentamos investigar este proceso de llenado de sEMG.

Métodos: Se registraron señales EMG superficiales del cuádriceps de sujetos sanos a medida que la fuerza aumentaba gradualmente de 0 a 40% MVC. El proceso de llenado sEMG se analizó midiendo el factor de llenado EMG (calculado a partir de los momentos no centrales de la señal sEMG rectificadas).

Resultados: (1) Al aumentar gradualmente la fuerza, aparecieron uno o dos saltos bruscos prominentes en la amplitud del sEMG entre el 0 y el 10% de la fuerza MVC en los vastos lateral y medial.

(2) Los saltos de amplitud se originaban cuando aparecían en la señal de sEMG unos pocos MUP de gran amplitud, que destacaban claramente de la actividad de sEMG anterior.

(3) Cada vez que se producía un salto brusco en la amplitud del sEMG, se iniciaba una nueva fase de llenado del sEMG.

Conclusiones: El proceso de llenado del sEMG tuvo una o dos etapas en los músculos vastos, estando el sEMG casi completamente lleno a fuerzas muy bajas (2-12% MVC).

Importancia: El factor de llenado es una herramienta prometedora útil para analizar el proceso de llenado EMG.

1. Introducción

La señal del electromiograma de superficie (sEMG) es la suma algebraica de trenes de potenciales de unidad motora (MUPs) generados por unidades motoras activas. A medida que la fuerza de una contracción aumenta gradualmente, aumenta el número de unidades motoras activas, se reclutan progresivamente unidades más grandes y las unidades ya reclutadas aumentan su frecuencia de disparo. Por lo tanto, la amplitud de la señal sEMG aumenta con el incremento de la fuerza [1]. Sin embargo, quedan por resolver varias cuestiones relacionadas con este aumento: (1) ¿El aumento de la amplitud del sEMG se produce de forma gradual y suave o se produce a saltos? (2) ¿La actividad sEMG a niveles de fuerza bajos se compone de unos pocos MUPs de gran amplitud (actividad "pulsátil"), o está formada por muchos MUPs de pequeña amplitud que se solapan y apenas sobresalen del ruido (actividad "continua")? (3) ¿Cuál es el nivel de fuerza en el que la señal de sEMG se llena por completo?

A pesar de los recientes avances en las técnicas de sEMG, aún no se comprende completamente la forma en que la

señal de sEMG se rellena progresivamente de MUPs a medida que aumenta el nivel de activación muscular. El conocimiento del proceso de llenado de la señal de sEMG puede proporcionar información útil para detectar la pérdida y la reinervación de unidades motoras (MU) [2], y también es valioso para el separar movimientos [3].

El proceso por el cual la señal sEMG se llena de MUPs a medida que aumenta la fuerza de contracción se refleja en cambios en la función de densidad de probabilidad (PDF) de las amplitudes sEMG (Fig. 1). En este sentido, se ha observado que, a bajas fuerzas de contracción, la PDF de sEMG se aproxima a una Laplaciana debido al escaso solapamiento entre MUPs de diferentes unidades motoras, mientras que, a medida que aumenta la fuerza, el solapamiento también aumenta y la PDF converge a una Gaussiana [4], [5].

En un estudio reciente de nuestro grupo presentamos un nuevo índice, el factor de llenado EMG, para cuantificar el grado de llenado de una señal EMG [6]. El factor de llenado informa sobre la forma de la distribución de la PDF sEMG, y también sirve como una medida de la superposición entre MUPs (Fig. 1). En nuestros experimentos preliminares comprobamos que el factor de llenado aumenta su valor a medida que el sEMG se llena progresivamente de MUPs (y aumenta el solapamiento de MUPs), y la forma de la PDF evoluciona hacia una gaussiana [6]. Para ilustrar este proceso, en la Fig. 1 se muestran tres señales sEMG registradas en el *vasto lateral* a tres niveles de contracción isométrica distintos (1, 3 y 9% de la fuerza de contracción máxima) y por tanto a 3 grados llenado diferentes.

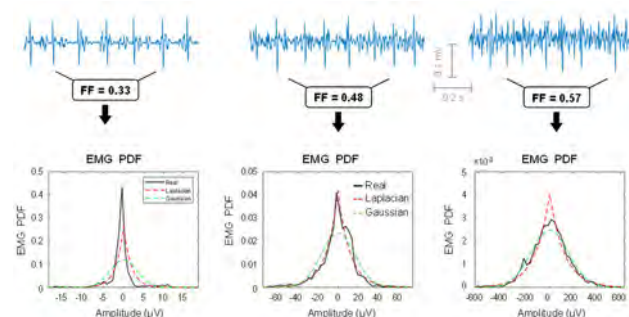


Fig. 1 Primera fila - Tres señales sEMG (1s) representativas registradas en el vasto lateral a niveles de contracción distintos (1, 3 y 9% de la fuerza máxima). El factor de llenado (FF, por sus siglas en inglés) se indica en el recuadro. Segunda fila - PDF correspondientes a las sEMG de la primera fila (líneas continuas) junto con las distribuciones de referencia laplaciana (líneas discontinuas) y gaussiana (líneas de puntos).

El objetivo del presente estudio fue investigar cómo la señal sEMG se llena de MUPs durante el incremento gradual de la fuerza muscular en los músculos cuádriceps. La nueva herramienta de análisis, el factor de llenado, tiene el potencial de caracterizar el proceso de llenado EMG, así como la capacidad de detectar posibles saltos bruscos en la amplitud sEMG. El análisis del factor de llenado también proporcionará una estimación del nivel de fuerza en el que la actividad sEMG se convierte en un patrón de interferencia completo. Dado que algunos procesos neurogénicos pueden alterar el proceso de llenado EMG [2], el análisis del factor de llenado puede tener potencial para detectar escenarios de pérdida de unidad motora. La pérdida de unidades motoras es característica de algunas enfermedades degenerativas de la motoneurona, como la esclerosis lateral amiotrófica, así como en otras patologías como la miopatía crónica y la atrofia muscular espinal, y, por último, es un rasgo típico del envejecimiento.

2. Material y métodos

2.1. Participantes

Veinte nueve participantes varones de entre 20 y 27 años (media \pm DE: 23 \pm 2 años) participaron en el estudio. Los experimentos fueron aprobados por el Comité de Ética de la Universidad Pública de Navarra, España (PI-010/21).

2.2. Montaje experimental y registro de la fuerza

Los participantes se sentaron en una silla a medida con un ángulo de rodilla de 90° y un ángulo tronco-muslo de 100°. La fuerza del cuádriceps se registró durante las contracciones isométricas voluntarias utilizando un extensómetro (STS, SWJ, China, sensibilidad 2 mV/V y 0,0017 V/N, rango lineal: 0-2452 N) que se fijó a la silla y se sujetó firmemente al tobillo con un molde hecho a medida. La señal de fuerza (extensión de la rodilla) se muestreó a 1000 Hz utilizando un sistema de conversión analógico-digital (MP150; BIOPAC, Goleta, EE.UU.).

2.3. Registros electromiográficos

Se registró sEMG del *vasto lateral*. La EMG se registró utilizando electrodos superficiales autoadhesivos (Ag/AgCl, Kendall 100), de forma circular (diámetro de registro 10 mm), colocados en configuración bipolar (distancia entre electrodos, 20 mm). El par de electrodos se colocó en el vientre del *vasto lateral* (zona de inervación). El electrodo de "tierra" se colocó sobre el tendón rotuliano. Antes de la colocación, la piel se exfolió y limpió para reducir la impedancia en la interfaz electrodo-piel. La sEMG se amplificaron (ancho de banda: 10 a 1000 Hz) y digitalizaron (frecuencia de muestreo: 5 kHz) utilizando un sistema de conversión analógico-digital (MP150; BIOPAC, Goleta, CA).

2.4. Protocolo experimental

En primer lugar, se calculó la fuerza voluntaria máxima (MVC) y para ello se realizaron 3 MVCs de extensión de rodilla. El valor más alto de MVC se utilizó como referencia para la determinación del nivel de fuerza correspondiente al 40% de MVC.

A continuación, los sujetos realizaron una contracción isométrica en "rampa" de 60 segundos de duración, durante la cual la fuerza se incrementó linealmente desde 0 hasta el 40% de la fuerza MVC. La trayectoria de fuerza deseada se mostraba en una pantalla de ordenador frente al participante, junto con la fuerza ejercida por el participante.

La contracción en rampa de 60 s se repitió tres veces en el mismo día, con un intervalo de descanso de 10 min entre contracciones. Esta sesión se repitió a los dos días.

2.5. Factor de llenado EMG

El factor de llenado se calculó a partir de los dos primeros momentos no centrales de la señal sEMG rectificadas, que se estiman de la siguiente manera:

$$m_1 = \frac{1}{N} \sum_{n=0}^{N-1} |x[n]| \quad (1)$$

$$m_2 = \frac{1}{N} \sum_{n=0}^{N-1} |x[n]|^2 \quad (2)$$

donde $x[n]$ es la señal sEMG muestreada y N es el número de muestras en cada segmento. Nótese que m_1 representa la media sEMG rectificadas, mientras que m_2 es el cuadrado de la media cuadrática sEMG (RMS).

A continuación, el factor de llenado (FF) EMG se calcula como la relación entre m_1 y m_2 de la siguiente manera [6]:

$$FF = \frac{m_1^2}{m_2} \quad (3)$$

2.6. Análisis del proceso de llenado EMG

El proceso de llenado del sEMG se analizó calculando el factor de llenado de segmentos sucesivos no solapados de la señal de sEMG. La duración de los segmentos fue de 0.7s, similar a la utilizada por Navallas et al. (2023) [6]. A continuación, los valores del factor de llenado se plotearon frente a la media rectificadas del sEMG (es decir, el m_1). En el presente estudio se registró la fuerza simultáneamente con la señal de sEMG para tener una indicación del nivel de esfuerzo al que se llena el sEMG (véase más adelante).

Uno de los objetivos de la presente investigación era determinar si el aumento de la amplitud del sEMG se produce de forma gradual o escalonada. Para ello, buscamos posibles saltos bruscos en la amplitud de la señal de sEMG. Para identificar la presencia de un salto de amplitud, deben satisfacerse dos criterios:

(1) Un criterio de amplitud. Se calcularon las amplitudes peak-to-peak máximas de 5 segmentos EMG secuenciales de 200 ms de duración, antes y después del salto. La media de las 5 amplitudes máximas después del salto debe superar en más del 100% la media de amplitudes antes del salto.

(2) Un criterio de factor de llenado. El valor del factor de llenado de un segmento sEMG de 0.7 s inmediatamente después del salto debe ser inferior a 0.45.

Otro objetivo del estudio era proporcionar una estimación del nivel de fuerza en el que la señal sEMG se rellena por completo. Esto corresponderá al momento en que el factor de llenado alcanzó el valor umbral 0.55, algo por debajo del valor correspondiente a la gaussianidad, 0.63 [6].

Aparte del factor de llenado, se midieron las siguientes 3 variables: (1) el número de saltos en la amplitud sEMG durante una contracción en rampa de 60 s; (2) el porcentaje medio de fuerza MVC en el que se produjo el salto; y (3) el porcentaje medio de fuerza MVC en el que el factor de llenado alcanzó 0.55.

3. Resultados

3.1. Ejemplo representativo del sEMG filling process

La figura 2 muestra un ejemplo representativo de una señal sEMG (plot b) registrada en un participante a medida que la fuerza aumentaba gradualmente de 0 a 40% (plot a). Los valores del factor de llenado extraídos de la señal sEMG se muestran en el gráfico c. Antes de iniciar la contracción, se registró ruido puro (gaussiano), que se sabe que tiene un valor de factor de llenado de 0.63. Se puede observar que el proceso de llenado del sEMG se produjo en 2 etapas secuenciales. La primera etapa comenzó cuando, a medida que el participante aumentaba lentamente la fuerza, aparecieron unos pocos MUP de gran amplitud, que sobresalían claramente del ruido, en 1.0s, (en plot b, véase la flecha blanca y el recuadro con FF=0.32). La aparición de estos nuevos MUPs de "gran amplitud" provocó una caída abrupta del factor de relleno hasta ~ 0.32 (véase la flecha blanca en el gráfico c). Después, el participante continuó aumentando la fuerza gradualmente, y otros MUPs entraron en juego, contribuyendo a llenar el EMG durante el intervalo de tiempo [1.0s, 12s]. Durante este primer periodo de llenado del EMG, el factor de llenado aumentó de ~ 0.34 a ~ 0.58 (gráfico c).

La segunda fase de llenado comenzó cuando, a medida que el participante seguía aumentando la fuerza, aparecieron a los 12s unos cuantos MUPs con una amplitud notablemente mayor que la actividad sEMG anterior (en plot b, véase la flecha gris y el recuadro con FF=0.36). La aparición de estos MUP de gran amplitud provocó una disminución del factor de llenado a ~ 0.36 (véase la flecha gris en el gráfico c). A partir de entonces, a medida que la fuerza seguía aumentando, entraron en juego otras MUPs, que contribuyeron a una segunda fase de llenado (plot b) y al aumento del factor de llenado de ~ 0.37 a ~ 0.63 (gráfico c). Durante este segundo periodo de llenado EMG, el factor de llenado alcanzó el valor de 0.55 a 11.3% MVC, y la meseta a aproximadamente 20% MVC. Las fuerzas correspondientes al primer y segundo saltos bruscos en la amplitud sEMG fueron del 0.6% y el 8% de la fuerza MVC, respectivamente (plot a).

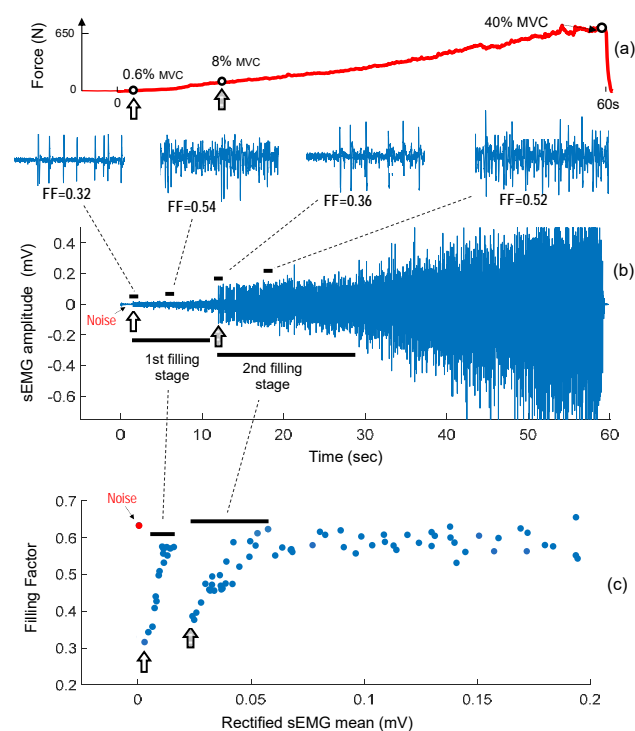


Fig. 2 Ejemplo representativo del proceso de llenado del sEMG en el vasto lateral de un sujeto (b) al aumentar gradualmente la fuerza de 0 a 40% MVC (a). Los valores del factor de llenado (FF) calculados a partir del sEMG se muestran en (c). Obsérvese que se produjeron dos saltos bruscos en la amplitud del sEMG a 1s y 12s (flechas blanca y gris en el gráfico b, respectivamente). La aparición de estos saltos en amplitud provocó que el factor de llenado disminuyera bruscamente (flechas blanca y gris en el gráfico c, respectivamente).

3.2. Resultados de grupo

La actividad sEMG inicial fue "pulsátil" (sEMG formada por unos pocos MUPs de gran amplitud, destacando claramente del ruido subyacente) en un gran porcentaje de los músculos analizados (82%, 24 de 29). En todos los casos (100%), se observó al menos un salto brusco prominente en la amplitud del sEMG. Se observó un salto de amplitud en el 46%, y dos saltos en el 54%.

La fuerza media a la que se produjeron los saltos en la amplitud sEMG fue del 2.8% MVC, rango [0.2%, 10% MVC]. Para muchos sujetos, se produjo un salto en la amplitud del sEMG justo al inicio de la contracción, es decir, el sEMG inicial fue pulsátil. Los demás valores comprendidos entre el 1.5 y el 10% MVC corresponden principalmente al segundo salto de amplitud de los sujetos en los que se observaron dos saltos. La fuerza media a la que el factor de llenado alcanzó el nivel de 0.55 fue del 5.7% MVC, rango [1.5%, 12.4% MVC].

4. Discusión

El factor de llenado es especialmente útil para analizar el proceso de llenado de la señal sEMG. En este estudio observamos que el factor de llenado aumentaba a medida que se añadían MUP adicionales a la señal sEMG, pero este aumento no era ilimitado. De hecho, la curva de aumento podía dividirse en dos regiones separadas: una primera

región (fuerzas bajas) en la que la curva del factor de llenado aumentaba rápidamente con el reclutamiento adicional, y una segunda región en la que la curva se aplanaba a medida que aumentaba el solapamiento entre MUPs, y el factor de llenado se aproximaba asintóticamente al valor de 0,63 (valor de la distribución gaussiana). Además, el factor de llenado es útil para detectar saltos en la amplitud del sEMG: a saber, una caída abrupta en el factor de llenado indica inequívocamente la aparición de unos pocos picos de gran amplitud que sobresalen de la actividad sEMG anterior.

Es importante destacar que el factor de llenado aporta información específica sobre el grado de llenado del sEMG que no puede obtenerse a partir de parámetros más clásicos como el RMS, más centrado en las características de amplitud del sEMG. Por ejemplo, en los ejemplos de las Fig. 2 puede observarse que, a medida que se aumentaba gradualmente la fuerza, el factor de llenado alcanzaba la meseta en torno al 20% MVC, mientras que la amplitud del sEMG seguía aumentando hasta el final de la contracción (40% MVC). Esto significa que la señal sEMG se llenó completamente al 20% MVC, a pesar del hecho de que el reclutamiento de unidades motoras adicionales continuó para fuerzas mayores.

Además, nuestro análisis del factor de llenado indica que, en los músculos vastos, se alcanzó un alto grado de llenado EMG (factor de llenado = 0.55) a fuerzas de contracción muy bajas, concretamente a ~5.7% MVC, lo que significa que el proceso de llenado sEMG se produjo muy rápidamente en estos músculos

Se proponen 2 teorías para explicar los saltos bruscos en amplitud de la señal sEMG: (1) que unas pocas unidades motoras de tamaño se recluten al 1.5-10% MVC; (2) que una unidad motora significativamente más cercana al electrodo, en comparación con las reclutadas previamente, fue llamada al 1.5-10% MVC. La primera teoría es coherente con el principio de Henneman [7], mientras que la segunda se basa en la idea de que, a medida que aumenta la fuerza, se activan más unidades motoras y, por lo tanto, es más probable que una UM recién reclutada se encuentre a una distancia menor del electrodo. En este punto, debe tenerse en cuenta el fenómeno de "regionalización" en el vasto lateral (es decir, que las unidades motoras pequeñas se localizan en la profundidad del músculo, mientras que las grandes se sitúan en regiones más superficiales) [8]. De hecho, una posible explicación sería que, para fuerzas de contracción muy bajas, el reclutamiento de unidades motoras comienza en las regiones más profundas del músculo, donde se localizan las unidades motoras de tipo I, y, a medida que aumenta la fuerza, el reclutamiento avanza hacia regiones más superficiales, donde las unidades de tipo II son mayoría.

Por último, es esperable que, en pacientes con pérdida de unidades motoras, las curvas de llenado sEMG estén alteren en comparación con las curvas de llenado de individuos sanos. Si esto ocurriera, el análisis del factor de llenado sería de gran utilidad para la detección de enfermedades y procesos neurogénicos.

5. Conclusiones y líneas futuras

Se ha descubierto que, a medida que la fuerza aumentaba lentamente en el vasto lateral, la amplitud del sEMG no aumentaba de forma gradual y suave, sino que aparecían uno o dos saltos bruscos y prominentes en la amplitud entre el 0 y el 10% de la fuerza MVC.

Los saltos de amplitud se originaban cuando aparecían en la señal de sEMG algunos MUP de gran amplitud, que destacaban claramente de la actividad de sEMG anterior o del ruido.

Cada vez que se producía un salto brusco en la amplitud del sEMG, se iniciaba una nueva etapa de llenado del sEMG, en la que se reclutaban sucesivamente MUP adicionales, que contribuían a llenar el sEMG en esta etapa. El proceso de llenado de la señal sEMG fue rápido, siendo casi completo al 2-12% MVC.

El factor de llenado, un índice basado en la distribución PDF de la señal sEMG, disminuyó significativamente cada vez que se produjo un salto abrupto en la amplitud sEMG, y este índice aumentó progresivamente a medida que se incorporan sucesivamente MUPs adicionales a la señal sEMG

Del presente estudio se desprende que el factor de llenado es una herramienta útil para analizar el proceso de llenado del sEMG.

Referencias

- [1] Disselhorst-Klug, C.; Schmitz-Rode, T.; Rau, G. Surface electromyography and muscle force: Limits in sEMG-force relationship and new approaches for applications. *Clin. Biomech.* 2009, 24, 225–235.
- [2] Nandedkar SD, Barkhaus PE, Stålberg EV. Form factor analysis of the surface electromyographic interference pattern. *Muscle Nerve.* 2020;62(2):233-238.
- [3] Nazarpour K, Sharafat A, Firoozabadi S (2007) Application of higher order statistics to surface electromyogram signal classification. *IEEE Trans Biomed Eng* 54(10):1762–1769
- [4] Nazarpour K, Al-Timemy AH, Bugmann G, Jackson A. A note on the probability distribution function of the surface electromyogram signal. *Brain Res Bull.* 2013; 90:88-91.
- [5] Ayachi FS, Boudaoud S, Marque C. Evaluation of muscle force classification using shape analysis of the sEMG probability density function: a simulation study. *Med Biol Eng Comput.* 2014;52(8):673-84.
- [6] Navallas J, Eciolaza A, Mariscal C, Malanda A, and Rodriguez-Falces J. EMG probability density function: a new way to look at EMG signal filling from single motor unit potential to full interference pattern. *IEEE Trans Neural Syst Rehabil Eng.* 2023 Feb 1;PP. doi: 10.1109/TNSRE.2023.3241354. Online ahead of print.
- [7] Henneman E, Somjen G, carpenter D. Functional significance of cell size in spinal motoneurons. *J Neurophysiol.* 1965; 28:560-80.
- [8] Knight CA, Kamen G. Superficial motor units are larger than deeper motor units in human vastus lateralis muscle. *Muscle Nerve.* 2005; 31(4):475-80.

Coeficiente de coactivación muscular como indicador del control motor: caso de estudio para flexo-extensión de codo en adultos

P. Martín Sierra^{1*}, C. Sánchez López de Pablo¹, E. Urendes Jiménez¹, R. Raya¹

¹ Departamento de Tecnologías de la Información, Escuela Politécnica Superior, San Pablo CEU, Madrid, España,
*pablo.martinsierra@ceu.es

Resumen

El índice de coactivación es una métrica que muestra el porcentaje de activación simultánea de un músculo agonista con su antagonista en un movimiento muscular concreto. En el caso de la flexo-extensión del codo, varios estudios previos resaltan la importancia de comparar el bíceps braquial con el tríceps braquial largo. Se escoge este movimiento por la sencillez de realización y medición, así como la falta de fuerza o control fino que requiere. El coeficiente de coactivación ha sido relacionado en la literatura con una mayor espasticidad muscular. En este contexto, este proyecto trata de analizar la viabilidad del uso de esta métrica como apoyo en la clasificación del estado del control motor en pacientes con parálisis cerebral. Para este propósito, y como primer paso en la consecución del objetivo anterior se ha desarrollado un algoritmo que la calcula, y posteriormente se ha validado en 10 adultos sanos, de edad comprendida entre 20 y 50 años, y se han comparado los resultados obtenidos con la bibliografía al respecto. Los resultados indican que existe un mayor nivel de coactivación durante la flexión del movimiento, mientras que en la extensión es más reducido. Además, el aumento de la frecuencia del movimiento implica un aumento de la coactivación en todo el movimiento.

1. Introducción

En pacientes con Parálisis Cerebral (PC) el control motor se ve comprometido por diversas condiciones como la debilidad muscular, la falta de control fino, la espasticidad y una elevada coactivación. Con el fin de estudiar la evolución del control motor de estos pacientes, se ha decidido analizar uno de estos factores influyentes, la espasticidad, relacionada con la rigidez muscular. Esta ha sido relacionada directamente con la coactivación muscular [1], al incrementar la coactivación aumenta la rigidez. Para el cálculo del Coeficiente de Coactivación (CC), se han propuesto diversas métricas basadas en las mediciones de electromiografías de los músculos, como la de Falconer y Winter [2] o la de Kellis y Arabatzi [3]. Para este experimento, se ha escogido el CC según Falconer y Winter [2] ya que este índice es útil para la comparación entre los sujetos medidos tras aplicar una normalización a la señal de electromiografía con respecto al pico máximo registrado de cada músculo en cada sesión de medición en vez de normalizar en función de la Máxima Contracción Voluntaria (MCV). Se escoge esta normalización debido a que no es aconsejable normalizar con respecto a MCV en pacientes con PC, puesto que resulta difícil conseguir de manera correcta este valor. Así pues, a pesar de trabajar con adultos sanos en esta primera aproximación, los algoritmos podrán ser comparables con los obtenidos

después en pacientes con PC. El coeficiente de Kellis y Arabatzi [3] no está pensado para este tipo de normalización, por lo que aplicamos el de Falconer y Winter [2].

En este contexto, en este trabajo, se ha generado un algoritmo que permite calcular el CC según esa métrica, y se han validado los resultados obtenidos en un caso particular: la coactivación entre el bíceps braquial y el tríceps braquial largo en movimientos de flexo-extensión de codo, a diferentes frecuencias de movimiento en adultos sanos. Se implementan estas frecuencias del movimiento ya que gracias a un ritmo externo de movimiento (con la ayuda de un metrónomo) se consigue comparar correctamente la actividad muscular y el rango de movimiento (ROM por sus siglas en inglés) realizado por cada sujeto y, además, esto permite una reducción del desplazamiento del tronco durante el movimiento, reduciendo compensaciones de músculos ajenos a la propia FE [4].

Las conclusiones obtenidas en este estudio preliminar serán utilizadas para poder aplicar esta métrica, junto a otras, que permitan un estudio y una cuantificación objetiva de la evolución motora de pacientes de PC durante su proceso de rehabilitación. Por tanto, los objetivos de este estudio son: la obtención una herramienta de cálculo del CC de una forma sencilla y la comprobación de que los resultados obtenidos con dicha herramienta se corresponden con los mostrados en la bibliografía.

2. Material y métodos

2.1. Electromiógrafo mDurance™

La herramienta utilizada para la obtención de las señales de electromiografía ha sido mDurance™. Este sistema inalámbrico registra la actividad de 2 músculos de forma simultánea. Además, incorpora un sensor inercial 3D que calcula el ROM a la vez que se mide la actividad muscular. El dispositivo presenta conectividad Bluetooth y una aplicación propia para su utilización. Esta aplicación permite ver la actividad muscular y el ROM en tiempo real, con lo que resulta sencillo detectar errores en la etapa de medición.

2.2. Segmentación del movimiento de flexo-extensión en sus fases de aceleración y deceleración

El movimiento se divide en la flexión y la extensión del propio codo y además estos movimientos se subdividen en sus fases de aceleración y deceleración, siguiendo el

ejemplo de A. Sarcher en su estudio de la extensión de codo en niños con CP [5]. De esta manera es posible dividir el movimiento y estudiar la coactivación dentro de esas divisiones.

La separación se consigue analizando el punto de máxima velocidad en las fases de flexión y extensión. Este punto es obtenido utilizando el ROM para calcular la velocidad en cada movimiento, y escogiendo el instante con el máximo valor de velocidad.

2.3. Descripción del protocolo

El experimento fue llevado a cabo con 10 adultos (5 mujeres y 5 hombres con una edad de entre 20 a 50 años) que no presentaban ningún tipo de problema motor ni cognitivo. Se registró el movimiento de Flexo-Extensión (FE) del codo del brazo dominante y no dominante. Este movimiento fue escogido por varios factores: la sencillez a la hora de realizarlo, la falta de fuerza o control fino que requiere y que sólo tiene un grado de libertad, lo que reduce el ruido de la actividad de otras articulaciones como la muñeca [6].

A los participantes se les indicó que se sentaran en una silla con la espalda recta y con las rodillas flexionadas en un ángulo de 90° con respecto al suelo. La medición de cada uno de los brazos se realizó de forma independiente, colocando 5 electrodos en cada uno de ellos. Estos electrodos fueron colocados en pares en el bíceps braquial y en el tríceps braquial largo, siguiendo las recomendaciones de SENIAM [7]. El electrodo de referencia fue colocado en el codo. El sistema mDurance™ se ajustó al antebrazo del paciente mediante una correa. Su posición era relevante para la obtención correcta del ROM, por lo que estaba colocado en el centro del antebrazo en el plano del movimiento de FE.

Se registró la FE con el brazo en posición neutra de pronosupinación, para replicar el estudio previo por Sarcher [5]. Los participantes recibieron instrucciones para realizar la FE cubriendo el ROM máximo sin forzar el movimiento partiendo y terminando el ciclo en la posición de máxima flexión. Las articulaciones de los hombros y las muñecas debían permanecer lo más estables posibles, para no afectar al movimiento del codo. Se realizaron 3 repeticiones a 3 frecuencias de movimiento distintas, marcadas con la ayuda de un metrónomo incluido en la aplicación de control del dispositivo mDurance™. Estas frecuencias marcaban la duración del ciclo completo de FE, de modo que los sujetos debían conseguir acompañar la máxima flexión y la máxima extensión con cada uno de los tonos del metrónomo. Se establecieron las tres mismas frecuencias que en el citado estudio de Sarcher, siendo 42, 60 y 78 revoluciones por minuto (RPM) con una pausa entre los cambios de frecuencia para modificar el valor del metrónomo y guardar la medición en un archivo diferente.

El ciclo comenzaba tras unos segundos desde el inicio del metrónomo, con el fin que el sujeto interiorizara la cadencia del movimiento y pudiera acompañar mejor las siguientes repeticiones. Para ayudar aún más al sujeto, se imitaba la FE en frente del sujeto a la velocidad y ritmo adecuados.

El sujeto realizaba 5 ciclos completos de FE para cada una de las pruebas, con lo que se registraron 6 extensiones y 5 flexiones. Se descarta la primera extensión para tener el mismo número de extensiones que de flexiones. Entre cada una de las pruebas, el sujeto descansaba 2 minutos para evitar la fatiga muscular.

2.4. Procesamiento

El sistema mDurance™ proporciona los datos de EMG procesados, con una frecuencia de muestreo de 1024 Hz y tras aplicar un filtrado paso-banda con un filtro Butterworth de cuarto orden, estableciendo una frecuencia de corte inferior de 20 Hz y una de corte superior de 450 Hz. La señal es además suavizada con una ventana de 0.025s de media cuadrática (RMS por sus siglas en inglés). El resultado es una envolvente de la señal de EMG con una frecuencia de muestreo de 4 Hz. Sin embargo, la frecuencia de la envolvente del ROM es de 20 Hz, por lo que, posteriormente, se llevó a cabo un procedimiento de sobremuestreo posterior sobre la envolvente del EMG para poder ajustar bien los puntos de máxima extensión y máxima flexión.

El algoritmo del cálculo del CC se desarrolló en Python 3, utilizando Jupyter Notebook 6.4.5 como interfaz. Los datos procesados de EMG y ROM presentaban la señal entera de los 5 ciclos de FE, por lo que para calcular el CC de cada una de las flexiones y extensiones, así como de sus fases de aceleración y deceleración, se debía segmentar la señal. Para esta segmentación de la envolvente del EMG se comenzó obteniendo los valores máximos y mínimos de la señal de ROM. Esta está medida de tal forma que 0° se corresponden a la posición del brazo en 90°, los valores máximos se corresponden a las máximas flexiones y los valores mínimos a las máximas extensiones, como puede verse en la Figura 1.

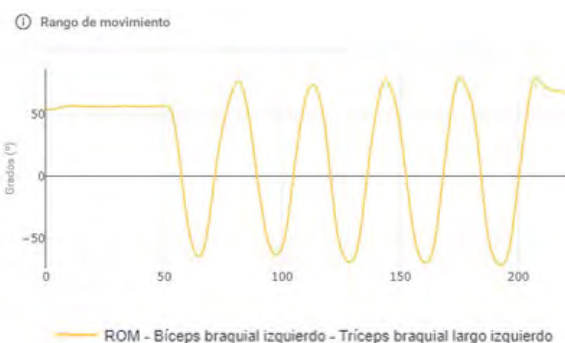


Figura 1. Medida del Rango de Movimiento de mDurance™

Se calcula la velocidad angular como la derivada discreta con respecto al tiempo del ROM, obteniendo sus valores máximo y mínimo que se corresponden con la máxima velocidad en la flexión y los de máxima velocidad en la extensión, respectivamente. Con estos puntos se segmenta la señal entera de ROM en 5 divisiones de flexión y 5 de extensión, que eran a su vez separados en los tramos de aceleración y deceleración, para un total de 20 segmentos distintos. Además, se calcula la línea base de la envolvente del EMG es ajustada a una línea base de actividad muscular que es obtenida a partir de los momentos iniciales de

reposo previo a empezar la prueba, durante los cuales el EMG tiene la ventana de actividad mínima.

Se normaliza la envolvente del EMG al pico de contracción máximo de cada músculo registrado durante cada ejercicio en lugar del MCV. Es precisamente por este tipo de normalización que se aplica la fórmula para el cálculo del CC de Falconer y Winter [3], que puede verse en la ecuación (1):

$$CC = 2 \cdot \frac{\int(Ag, Antag)}{\int Ag + \int Antag} \cdot 100 \quad (1)$$

donde *Ag* es el área bajo la envolvente del músculo agonista, *Antag* el área bajo la envolvente del músculo antagonista y $(Ag, Antag)$ el área común de ambos músculos.

2.5. Análisis de datos

Se calcula los CC medios y la desviación estándar correspondiente de todos los pacientes, para cada una de las frecuencias del ciclo de FE propuestas, separando los resultados para cada una de las 4 fases del ciclo consideradas, la aceleración de la flexión y la extensión y la deceleración de las mismas.

3. Resultados

En la Figura 2 se muestran las señales de EMG del bíceps y el tríceps combinadas y separadas. Puede verse las distintas áreas de activación marcadas, tanto las individuales como las superpuestas. Con el fin de visualizar si el cálculo del CC en los segmentos era correcto, se probó a calcular el CC con la envolvente del EMG sin segmentar.

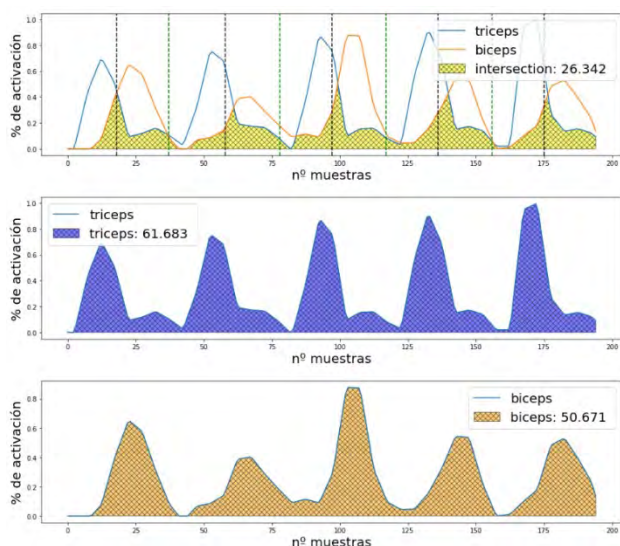


Figura 2. Visualización de las envolventes de EMG de un sujeto con las áreas de activación marcadas. Las líneas discontinuas negras indican los puntos de máxima extensión y las verdes indican los puntos de máxima flexión

Se obtienen diagramas de caja para cada uno de los segmentos empleando la Ecuación 1, obteniendo así los resultados para la aceleración de la flexión (CC_AccFlex), la deceleración de la flexión (CC_DeccFlex), la aceleración de la extensión (CC_AccExt) y la deceleración de la extensión (CC_DeccExt). Esto se realiza para cada repetición y posteriormente se obtiene la media y desviación estándar de cada sujeto. Un ejemplo de estos diagramas se puede ver en la Figura 3.

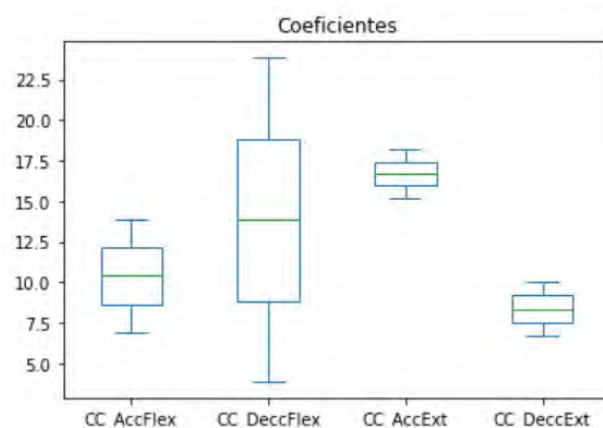


Figura 3. Ejemplo de diagrama de caja de los coeficientes de coactivación para uno de los sujetos

Se calcularon también los valores medios y las desviaciones estándar para cada uno de estos 4 valores. Esto se repite para cada una de las 3 frecuencias, para cada brazo y para cada uno de los sujetos. Los CC resultantes de todos los participantes se pueden ver divididos en segmentos del ciclo y en la frecuencia del movimiento en la Tabla 1.

Frec.	42 rpm	60 rpm	78 rpm
CC_Acc. Flex	15,92± 4,37%	20,29± 5,60%	21,53± 4,08%
CC_Decc. Flex	26,94±7,56%	28,89±9,19%	20,93±5,89%
CC_Acc. Ext	8,64±3,21%	8,64±3,21%	9,32±4,02%
CC_Decc. Ext	6,35± 2,90%	6,88±4,03%	11,70±3,71%

Tabla 1. Coeficientes de Coactivación del conjunto de sujetos, en función de la frecuencia y de los segmentos del ciclo

4. Discusión y conclusiones

Se puede observar en la Tabla 1 que el porcentaje de coactivación muscular en la fase de flexión es significativamente más elevado que aquellos obtenidos en la fase de extensión. Esto podría ser debido a lo observado durante las fases de medición, pues los sujetos solían contraer más el brazo al flexionar, pudiendo explicar así un aumento en la coactivación del músculo antagonista (el tríceps braquial largo) su parte final del ciclo. Cabe pensar que esto se debe a la instrucción de mantener el ritmo marcado por el metrónomo y tratar de cuadrar el punto de máxima flexión y extensión con el pitido de este.

Al tener que ejercer una fuerza mayor para superar a la fuerza gravitatoria, que se opone al movimiento de flexión, sobre todo comparado con la fuerza necesaria para la extensión, el aumento del CC cabría dentro de los valores esperados. Siguiendo la misma lógica, podemos explicar el aumento del CC directamente proporcional al aumento de la frecuencia del metrónomo. Al aumentar el tiempo que deben seguir los participantes, el movimiento es más rápido, siendo así más común tener una contracción muscular más elevada en el brazo para poder alcanzar la máxima flexión en el tiempo marcado.

Comparando los resultados obtenidos con la literatura precia del estudio de A. Sarcher [5,6], se puede comprobar que los valores de CC en extensión son inferiores, lo cual implica una menor coactivación muscular. Debido a que la muestra del experimento eran adultos, en vez de niños, es posible que los resultados inferiores puedan ser explicados por un mayor control motor por parte de los adultos, lo cual resulta en una reducción en coactivaciones innecesarias en la articulación. Observando los resultados obtenidos en este caso de estudio, y aquellos obtenidos en la bibliografía, se podría esperar que el CC de los pacientes de PC fuese significativamente más elevado que los sujetos sanos.

Este estudio proporciona los datos sobre el CC en la flexión y en la extensión en sanos, estudiando el movimiento completo en lugar de sólo la fase de la extensión. Esto resulta provechoso puesto que, en un próximo estudio, se quiere calcular el CC en pacientes con PC y analizar otras métricas que puedan resultar útiles para estudiar el progreso de rehabilitación de este tipo de pacientes, cuantificando de forma objetiva la mejora motora que pueda conseguir.

1. Agradecimientos

Este experimento forma parte se del proyecto Neurometrics financiado por MCIN/ AEI/ 10.13039/501100011033/ FEDER, UE, con número de beca: PID2021-127096OB-100.

Referencias

- [1] Rinaldi M, Ranavolo A, Conforto S, Martino G, Draicchio F, Conte C, Varrecchia T, Bini F, Casali C, Pierelli F, Serrao M. Increased lower limb muscle coactivation reduces gait performance and increases metabolic cost in patients with hereditary spastic paraparesis. *Clin Biomech (Bristol, Avon)*. 2017 Oct;48:63-72. doi: 10.1016/j.clinbiomech.2017.07.013. Epub 2017 Jul 29. PMID: 28779695.
- [2] Kellis E, Arabatzi F, Papadopoulos C. Muscle co-activation around the knee in drop jumping using the co-contraction index. *J Electromyogr Kinesiol*. 2003 Jun;13(3):229-38. doi: 10.1016/s1050-6411(03)00020-8. PMID: 12706603.
- [3] Falconer K, Winter DA. Quantitative assessment of co-contraction at the ankle joint in walking. *Electromyogr Clin Neurophysiol*. 1985 Mar-Apr;25(2-3):135-49. PMID: 3987606.
- [4] van Roon D, Steenbergen B, Meulenbroek RG. Trunk use and co-contraction in cerebral palsy as regulatory mechanisms for accuracy control. *Neuropsychologia*. 2005;43(4):497-508. doi: 10.1016/j.neuropsychologia.2004.07.014. Epub 2004 Oct 3. PMID: 15716140.
- [5] Sarcher A, Raison M, Ballaz L, Lemay M, Leboeuf F, Trudel K, Mathieu PA. Impact of muscle activation on ranges of motion during active elbow movement in children with spastic hemiplegic cerebral palsy. *Clin Biomech (Bristol, Avon)*. 2015 Jan;30(1):86-94. doi: 10.1016/j.clinbiomech.2014.10.009. Epub 2014 Nov 5. PMID: 25467763.
- [6] Sarcher A, Raison M, Leboeuf F, Perrouin-Verbe B, Brochard S, Gross R. Pathological and physiological muscle co-activation during active elbow extension in children with unilateral cerebral palsy. *Clin Neurophysiol*. 2017 Jan;128(1):4-13. doi: 10.1016/j.clinph.2016.10.086. Epub 2016 Oct 29. PMID: 27866118.
- [7] Página web del proyecto SENIAM (Surface ElectroMyoGraphy for the Non-Invasive Assessment of Muscles). http://seniam.org/sensor_location.htm (Consultada: Septiembre 2023).

Monitorización de la depresión mediante el análisis de la circadianidad del ritmo cardíaco proporcionado por un dispositivo *wearable*

S. Pérez¹, S. Kontaxis^{1,2}, E. García^{2,3}, S. Siddi⁴, N. Cummins⁵, S. Vairavan⁶, F. Matcham^{5,7}, J.M. Haro⁴, M. Hotopf^{5,8}, F. Lamers^{9,10}, B. Penninx^{9,10}, R. Dobson⁵, V. Narayan⁶, R. Bailón^{1,2}, A. Martín-Yebra^{1,2}, the RADAR-CNS consortium

¹ Grupo BSICoS, I3A, IIS Aragón, Universidad de Zaragoza, Zaragoza, España, rbailon@unizar.es ² Centro de Investigación Biomédica en Red – BBN (CIBER-BBN), Zaragoza, España ³ Microelectrónica y Sistemas Electrónicos, Universidad Autónoma de Barcelona, CIBER-BBN, Barcelona, Spain ⁴ Parc Sanitari Sant Joan de Déu, Fundació Sant Joan de Déu, CIBERSAM, Universitat de Barcelona, Barcelona, Spain ⁵ King's College London, Institute of Psychiatry, Psychology and Neuroscience, London, UK ⁶ Research and Development Information Technology, Janssen Research & Development, LLC, Titusville, NJ, USA ⁷ School of Psychology, University of Sussex, Falmer, UK ⁸ South London and Maudsley NHS Foundation Trust, London, UK ⁹ Department of Psychiatry, Amsterdam UMC, Vrije Universiteit, Amsterdam, the Netherlands ¹⁰ Amsterdam Public Health Research Institute, Amsterdam, the Netherlands

Resumen

En este estudio se ha aplicado el método de ajuste Cosinor, por mínimos cuadrados a una función senoidal, a los datos de frecuencia cardíaca (FC) de 203 pacientes con depresión, registrados de manera continua durante un transcurso de 18 meses por un dispositivo *wearable*, en condiciones de vida cotidiana. El objetivo es evaluar si la posible pérdida del ritmo circadiano, modulador de la frecuencia cardíaca, está asociada a una depresión más severa. Estos datos coexisten con resultados de pruebas médicas para la evaluación de la sintomatología de la depresión, como el *Patient Health Questionnaire* (PHQ-8) [1] y el *Inventory of Depressive Symptomatology* (IDS) [2], que permiten determinar la presencia y gravedad del trastorno.

El estudio U de Mann-Whitney sobre el ajuste Cosinor de la frecuencia cardíaca, sincronizado a los registros de PHQ-8 e IDS basales de cada paciente, ha permitido encontrar diferencias significativas según la gravedad del trastorno: la amplitud derivada del ajuste Cosinor (es decir, la oscilación de la FC a lo largo del día) es significativamente menor en aquellos pacientes con depresión severa. Este resultado se cumple en todas las ventanas temporales de datos sobre las que se ha realizado el ajuste Cosinor (1 día, 1 semana y 2 semanas), así como para los ajustes sincronizados con PHQ-8 e IDS. Esto supone una pérdida en la circadianidad cuando la depresión es severa.

1. Introducción

El trastorno depresivo mayor fue considerado la tercera mayor carga de enfermedad mundial en 2008 por la Organización Mundial de la Salud, y se espera que sea la primera en 2030 [3]. Al estar relacionado con diversas alteraciones fisiológicas (cambios en el estado de ánimo, problemas de sueño...), muchas de ellas relacionadas con el sistema nervioso autónomo, la monitorización de la frecuencia cardíaca, obtenible de forma continua, robusta y poco costosa mediante dispositivos *wearable*, puede tener un papel importante. En particular, la monitorización continua y objetiva de la enfermedad, a través de parámetros fisiológicos como la FC, así como la realización de tests con respuestas subjetivas, basa-

das en la percepción del individuo (por ejemplo, *Patient Health Questionnaire*, PHQ-8 e *Inventory of Depressive Symptomatology*, IDS), pueden ayudar a los equipos terapéuticos a conocer síntomas no detectables en un entorno extrahospitalario, y desarrollar un tratamiento adaptado a las necesidades de cada paciente.

Todos los seres vivos presentan, con el transcurso del tiempo, funciones oscilatorias influenciadas por cambios lumínicos o estacionales de la Tierra. Se distingue entre estados de vigilia y sueño. El ritmo circadiano, proveniente de las palabras latinas “*circa*” (alrededor) y “*diem*” (día), es uno de los ritmos biológicos que rigen nuestro organismo, existiendo una periodicidad biológica de, aproximadamente, 24 horas. Las variaciones circadianas son conocidas desde hace tiempo, debido a su influencia en los sistemas fisiológicos, entre ellos, el cardiovascular. Su relevancia clínica lo ha convertido en objeto de estudio en los últimos años [4]. Uno de los aspectos estudiados es su influencia en el ritmo cardíaco. Una forma de evaluar dicha circadianidad del ritmo cardíaco es el uso del método Cosinor. Por medio de mínimos cuadrados, este método identifica y evalúa la función matemática coseno que mejor se ajusta a los datos.

La hipótesis de partida es que la depresión, especialmente la severa, causa alteraciones en el sistema nervioso y el ciclo de sueño, que se manifestarán en alteraciones de la circadianidad del ritmo cardíaco, y que podrán cuantificarse mediante el ajuste Cosinor.

El objetivo de este estudio es, por tanto, encontrar posibles diferencias significativas en los parámetros del ajuste Cosinor (MESOR, amplitud y acrofase), producidas según la severidad de la depresión.

2. Materiales

Se han utilizado los datos de FC de 203 pacientes, pertenecientes al proyecto europeo RADAR-MDD [5], con trastorno depresivo mayor. Han hecho uso de tecnologías de medición remotas y, a la par, se ha realizado un seguimiento de sus resultados clínicos.

El proyecto RADAR ha creado una plataforma IoT de

código abierto, RADAR-base, donde se han recolectado los datos a tiempo real de los dispositivos *wearable*, a través de Bluetooth y una aplicación en el *smartphone* del paciente. Con tal objetivo, todos ellos han utilizado una pulsera Fitbit Charge 2/3 durante su vida cotidiana, que ha recogido datos de manera continua de su frecuencia cardíaca. Estos dispositivos poseen un sensor óptico, que parpadea con luces LED, así como fotodiodos, que detectan la luz reflejada, en función de los cambios de volumen sanguíneo en los capilares de la muñeca. De esta forma, se obtiene la señal fotopletolegráfica (PPG), de cuyo post-procesado se deriva la frecuencia cardíaca.

De forma paralela, han hecho uso de una aplicación móvil, con la que han respondido los cuestionarios PHQ-8 e IDS en el teléfono móvil cada 2 semanas y 3 meses, respectivamente, que se utilizarán como referencia para evaluar la severidad de la depresión. Están compuestos por una serie de preguntas evaluadas en una escala de ‘0’ a ‘3’ y su puntuación se suma, obteniendo un resultado final. Se consideran un PHQ-8 ≥ 10 y un IDS ≥ 26 indicativos de depresión severa. La duración de la toma de datos es de 18 meses.

Los datos de los pacientes han sido recopilados por el Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (España), el King’s College de Londres (Reino Unido) y el Amsterdam University Medical Centre (Países Bajos).

Hay un total de 203 pacientes, procedentes de Reino Unido (112), España (51) y Países Bajos (40). 114 son mujeres y 59 son hombres, con una edad media de $53,24 \pm 15,15$ años (mínimo 24, máximo 84 años).

El software utilizado para el tratamiento de los datos y estudio estadístico ha sido MATLAB.

3. Métodos

3.1. El método Cosinor

El método Cosinor es un modelo de regresión ampliamente utilizado en el análisis de series temporales de datos [6]. Este método utiliza mínimos cuadrados para ajustar dichas series a funciones sinusoidales. Es aplicable a datos no equidistantes, ventaja significativa en este análisis. Se aplica este método con el objetivo de estudiar potenciales alteraciones en los ritmos biológicos/circadianos de la frecuencia cardíaca, que pueden estar asociadas a la depresión.

El modelo de una componente se define:

$$Y(t) = M + A \cos(2\pi t/\tau + \phi) + e(t) \quad (1)$$

donde M es el MESOR (*Midline Estimating Statistic of Rhythm*, una media de la frecuencia cardíaca ajustada al ritmo), A es la amplitud de la senoide (la mitad de la extensión de la variación rítmica de la frecuencia cardíaca en un ciclo), ϕ es la acrofase (intervalo temporal donde se esperan los valores más altos de frecuencia cardíaca), τ es el periodo (duración de un ciclo) y $e(t)$ es el término de error del ajuste (Figura 1).

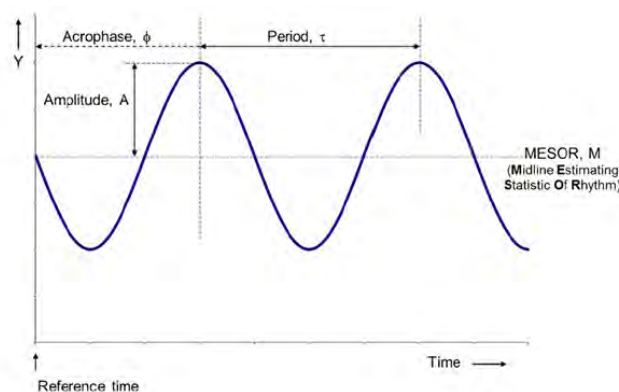


Figura 1. Parámetros de la función Cosinor [6].

La oscilación circadiana se corresponde siempre con un periodo de 24 horas, no obstante, existe la posibilidad de que tomando series de datos más largas se consiga un ajuste más preciso. En esos casos, la función sinusoidal tendrá tantas oscilaciones como ciclos de 24 horas haya en la ventana temporal.

Debido al gran volumen de datos y la carga computacional que supone la aplicación del método Cosinor, así como el aspecto ruidoso de los datos crudos, se opta por trabajar con la frecuencia cardíaca media por horas, señal mucho más suavizada.

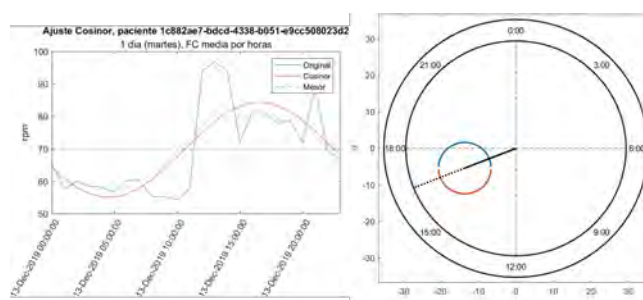


Figura 2. Aplicación del ajuste Cosinor a los datos de FC media por horas, en una ventana temporal de 1 día.

La Figura 2 muestra la aplicación gráfica del método, sobre la FC media por horas de 1 día. A la izquierda aparece el ajuste por mínimos cuadrados a un coseno. A la derecha se muestra el diagrama de reloj, donde la hora corresponde con la acrofase, la longitud de su línea es la amplitud y la circunferencia que la acompaña es el intervalo de confianza del 95%. Cuando la circunferencia no recoge el centro del reloj (ejemplo de la izquierda), el test de amplitud-cero (“no existe ritmo circadiano”) se rechaza, y la hipótesis alternativa es que existe una oscilación significativa de los datos durante ese periodo de tiempo. Si la circunferencia recogiera el centro del reloj (ejemplo de la derecha), la hipótesis nula “no hay variación significativa de la amplitud” se aceptaría.

3.2. Definición de umbrales de pérdida de datos

Una circunstancia habitual en las medidas tomadas por dispositivos *wearable* es la pérdida de datos por una mala colocación de la pulsera, movimientos bruscos, periodos de carga de la batería... Partiendo de la frecuencia media por horas de un paciente de ejemplo, se seleccionan ventanas temporales de 1 día, 1 semana y 2 semanas

donde no se aprecie falta de datos. Sobre esas ventanas se eliminan datos, en ráfagas de distinta longitud y posiciones aleatorias, para estudiar la robustez del método frente a situaciones adversas.

3.3. Cálculo de los parámetros Cosinor sincronizados a los valores de PHQ-8 e IDS

Con el objetivo de relacionar los parámetros Cosinor con los resultados de los test PHQ-8 e IDS, se realiza un ajuste Cosinor sincronizado a dichos indicadores de la depresión. Esto quiere decir que, tomando como referencia la fecha de registro de cada uno de los indicadores, se aplica el método Cosinor a los datos en ventanas de 1 y 2 semanas que le preceden. Además, para comprobar posibles diferencias, se realiza el ajuste Cosinor también en ventanas de 1 día, durante 1 y 2 semanas previas. De esta forma, cada valor de PHQ-8 e IDS tendrá asociado un valor o una serie de parámetros obtenidos del análisis Cosinor (de los que podrá hacerse un promediado) que describan la variación circadiana de la FC.

Se añade, además, un cribado adicional: para que un paciente sea incluido en el análisis, se exige que tenga, al menos, 12 valores de PHQ-8 (test realizado cada 2 semanas) con ajuste Cosinor simultáneo en un periodo igual o inferior a 1 año. En el estudio de IDS, se exigen los 4 valores (test trimestral) en un año. Como los datos de cada paciente no se extienden, por lo general, a lo largo de exactamente 1 año, este umbral se traslada proporcionalmente a la duración que tenga el seguimiento del paciente en particular (esperado 18 meses).

3.4. Análisis estadístico

Con los pacientes que superan el cribado, se va a estudiar si existen diferencias en la variación circadiana (en términos de MESOR, amplitud y acrofase) de la FC entre pacientes con depresión severa y no severa., se considera que el paciente sufre una depresión severa cuando presenta un PHQ-8 ≥ 10 o un IDS ≥ 26 . Conocidos estos datos, puede llevarse a cabo un estudio de las diferencias significativas. Así, se busca determinar si existe una diferencia significativa en los parámetros Cosinor, sincronizados con el valor de PHQ-8 e IDS basal de los pacientes, dependiendo de la severidad de la depresión. Para ello, se ha utilizado la prueba U de Mann-Whitney. Si el p-valor es menor que 0.05, se considera evidencia suficiente para rechazar la hipótesis nula y concluir que hay una diferencia significativa entre los grupos.

4. Resultados y discusión

4.1. Definición de umbrales de pérdida de datos

Para el estudio de la robustez frente a la pérdida de datos en series de FC media de 1 día, se llevan a cabo 100 realizaciones. En los estudios de 1 y 2 semanas se estudian tantas realizaciones como localizaciones posibles de las ráfagas sin datos, calculando la media \pm desviación estándar del porcentaje de variación de los parámetros Cosinor, respecto a los obtenidos con la serie completa de datos. Además, se hace un conteo de realizaciones significativas, esto es, aquellas donde el p-valor obtenido en el ajuste Cosinor es < 0.05 . Este procedimiento se repite para ráfagas de diferente duración. La Tabla 1 muestra

los umbrales elegidos para cada ventana temporal.

4.2. Relación del ajuste Cosinor con la severidad de la depresión

Del subconjunto de 203 pacientes con el que se trabaja han superado el cribado, según la calidad de sus datos de PHQ-8, 109 y 108 pacientes, para ventanas temporales de 1 y 2 semanas, respectivamente. En el caso de IDS, han sido 107 y 101.

La prueba U de Mann-Whitney se ha realizado para los parámetros obtenidos del análisis del Cosinor (MESOR, amplitud y acrofase) de los ajustes sincronizados a los valores de PHQ-8 e IDS, con ventanas temporales de 1 semana, 2 semanas, y el promedio de 7 y 14 días. En el caso del IDS, como se mide cada 3 meses, se realiza un promediado de las ventanas temporales calculadas durante los 3 meses anteriores al indicador. De esta forma, se tiene en consideración todo el periodo englobado por el IDS. Se realiza únicamente con los datos basales del paciente, es decir, con el primer valor de PHQ-8/IDS del que se disponen datos de FC previos.

Como el número de pacientes con depresión severa es reducido, para que los datos estén balanceados, se consideran al azar el mismo número de pacientes no severos. Se llevan a cabo varias repeticiones para asegurar una mayor robustez estadística del resultado.

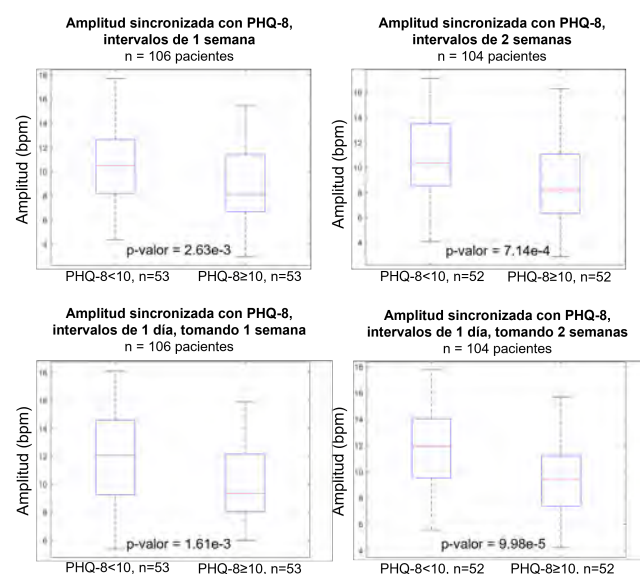


Figura 3. Distribución de amplitud obtenida del análisis Cosinor, para pacientes con depresión severa ($PHQ \geq 10$) y no severa ($PHQ < 10$).

La amplitud del Cosinor presenta diferencias significativas entre ambos grupos. En el caso de depresión severa, esta amplitud es significativamente menor y, por tanto, se reduce la oscilación noche-día de la frecuencia cardíaca cuando la depresión es severa. Esto ocurre tanto con la clasificación de pacientes en base al cuestionario PHQ-8 como cuando se usa el índice IDS. Las Figuras 3 y 4 muestran los diagramas de cajas de la distribución de los valores de amplitud para cada grupo de pacientes, junto con el p-valor.

Ventana temporal	Longitud de la ráfaga de muestras eliminadas	% variación MESOR	% variación amplitud	% variación acrofase	Realizaciones significativas
1 día (n=24)	4	-0.10 % ± 1,29	23,67 % ± 56,30	0,33 % ± 15,71	62/100
1 semana (n=168)	50	-0,26 % ± 0,59	7,08 % ± 17,72	0,04 % ± 3,40	119/119
2 semanas (n=336)	100	-0,20 % ± 0,56	5,86 % ± 13,96	0,34 % ± 1,61	237/237

Tabla 1. Umbral de ráfagas de pérdidas de datos seleccionados para distintas ventanas temporales.

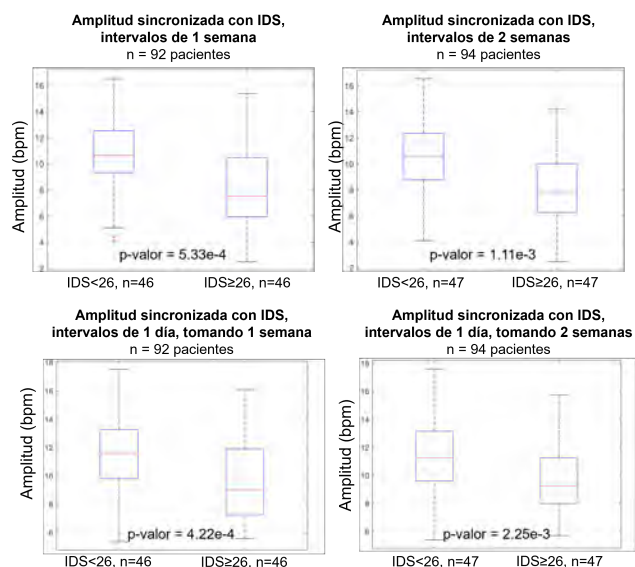


Figura 4. Distribución de amplitud obtenida del análisis Cosinor, para pacientes con depresión severa ($IDS \geq 26$) y no severa ($IDS < 26$).

Frente a la posibilidad de trabajar con distintas ventanas temporales, se ha observado que tanto 1 como 2 semanas dan resultados significativos en cuanto a la separación de pacientes según la severidad de la depresión. Bien es cierto que aplicar ventanas temporales de 1 día y promediar 7 y 14 días da resultados aún mejores, por lo que existe un compromiso entre el coste computacional y la precisión de los resultados.

La Figura 5 muestra gráficamente un ejemplo de ajuste Cosinor de las 2 semanas anteriores al valor basal de PHQ-8 en un paciente con depresión severa y en otro con depresión no severa. Se observa como en el caso de depresión severa, la amplitud de la oscilación es menor.

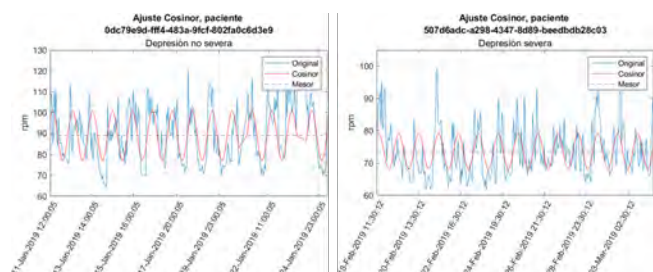


Figura 5. Ajuste Cosinor sincronizado con PHQ-8, ventana temporal de 2 semanas. Depresión no severa, amplitud=11.82 bpm. Depresión severa, amplitud=5.43 bpm.

5. Conclusiones

En este estudio se ha evaluado el efecto de la modulación circadiana en la FC en pacientes con diferentes grados de depresión. Los resultados obtenidos permiten afirmar que existe una pérdida de la circadianidad en la frecuencia cardíaca, inducida por la gravedad de la depresión: aquellos pacientes cuyos cuestionarios determinaron un estado de depresión severa, tienen un ajuste Cosinor con amplitud significativamente inferior.

La alteración del ritmo circadiano en la frecuencia cardíaca crea nuevas posibilidades. Como apunta [7], un estudio de trastornos del estado de ánimo (como el depresivo mayor), la mayoría de investigaciones se han centrado en las emociones, prestando menos importancia a los ciclos del sueño y alteraciones circadianas, a pesar de ser indicadores habituales de la enfermedad.

Agradecimientos

Este trabajo ha sido parcialmente financiado por los proyectos TED2021-131106B-I00, PID2021-126734OB-C21 (Ministerio de Ciencia e Innovación y Fondo Social Europeo), España, el grupo BSICoS T39-23R (Gobierno de Aragón y Fondo Social Europeo). El proyecto RADAR-CNS ha recibido financiación de Innovative Medicines Initiative 2 Joint Undertaking mediante el acuerdo No 115902. Los cálculos fueron realizados por el ICTS NANBIOSIS (HPC de la Universidad de Zaragoza).

Referencias

- [1] K. Kroenke et al. The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, 114:163–173, 04 2009.
- [2] M. Trivedi et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR) in public sector patients with mood disorders. *Psychological Medicine*, 34:73–82, 01 2004.
- [3] J. Octavio et al. Circadian Profiles of Heart Rate and its Instantaneous Variability in Patients With Chronic Chagas' Disease. 02 2004.
- [4] E. Hernandez et al. Alteraciones circadianas del sistema cardiovascular. *Revista Española de Cardiología*, 53:117–122, 01 2000.
- [5] RADAR-CNS: Remote Assessment of Disease and Relapse – Central Nervous System. <https://www.radar-cns.org/>.
- [6] G. Cornelissen. Cosinor-based rhythmometry. *Theoretical Biology and Medical Modelling*, 11, 04 2014.
- [7] I. B. Hickie et al. Manipulating the sleep-wake cycle and circadian rhythms to improve clinical management of major depression. *BMC Medicine*, 11, 03 2013.

Aplicación de inteligencia artificial para la predicción del estrés en cirugía robótica mínimamente invasiva

D. Caballero¹, J.A. Sánchez-Margallo¹, M.J. Pérez-Salazar¹, M. Kappel^{1,2}, F.M. Sánchez-Margallo¹

¹ Centro de cirugía de mínima invasión Jesús Usón, Cáceres, España {dcaballero, jasanchezm, mjperez, msanchez}@ccmijesususon.com

² Escuela de Ingeniería Superior, Universidad Sigma-Clermont, Aubière, Francia, mathilde.kappel@sigma-clermont.fr

Resumen

La monitorización de ciertos parámetros, tanto fisiológicos, como ergonómicos, durante la realización de cirugías robóticas es crucial para asegurar la salud de los cirujanos. Por ello, se observa la necesidad de estudiar varios factores relacionados con la cirugía robótica de manera predictiva. Por tanto, el objetivo de este estudio es predecir el valor de las características relacionadas con la fisiología (actividad electrodermal (EDA), temperatura corporal y presión arterial) y la ergonomía (cinemática) del cirujano a partir de los valores recopilados en el instante inmediatamente previo. Para ello, se recopilaban datos durante 26 sesiones quirúrgicas completadas por 11 cirujanos. Estos datos pertenecen a los parámetros ergonómicos y fisiológicos del cirujano (EDA, temperatura corporal, presión arterial y cinemática del cirujano). Una vez generado el conjunto de datos, se aplicaron técnicas de preprocesado, obteniendo dos subconjuntos de datos, los cuales fueron divididos en tres subconjuntos: con un 80% de los datos para la calibración, un 10% para la validación cruzada y un 10% para la validación externa. Sobre el subconjunto de calibración se aplicaron tres técnicas predictivas de inteligencia artificial para generar los modelos predictivos. Finalmente, estos modelos predictivos se validaron sobre los subconjuntos de validación y de test. Los resultados obtenidos mostraron que los modelos lineales eran ligeramente mejores que los modelos no lineales y las redes neuronales. Estos modelos fueron validados correctamente ($R^2 > 0,6$ y error relativo cuadrático medio (ERMC) $< 0,13$), mostrando la posibilidad de predecir estos factores a partir de los sucesos monitorizados en el instante inmediatamente previo.

1. Introducción

La monitorización de ciertos parámetros durante la realización de cirugías robóticas es crucial, tanto desde el punto de vista de la ergonomía como del fisiológico del cirujano [1]. Para evaluarlas, se han venido valorando un gran número de características que pueden ser extraídas desde diferentes dispositivos, y que se encuentran ampliamente definidos en la literatura científica [2].

Entre estos dispositivos para evaluar estos factores, se encuentran los sensores de actividad electrodermal (EDA). La utilización de dispositivos EDA es excelente para evaluar estos parámetros fisiológicos y ergonómicos del cirujano durante la cirugía robótica, ya que permite cuantificar pequeños cambios producidos, tanto ergonómicos como en la presión arterial o en la temperatura corporal. Estos dispositivos proporcionan información detallada del cirujano durante todo el desarrollo de la cirugía.

En relación con las técnicas de cirugía robótica de mínima invasión, muchas de ellas están siendo adoptadas como las técnicas quirúrgicas estándar. Estas técnicas tienen numerosas ventajas entre las que destacan la disminución del dolor en el postoperatorio o la reducción del consumo de analgésicos, aunque, también tienen ciertas desventajas para los cirujanos como son las limitaciones ergonómicas o los altos niveles de estrés [3].

Adicionalmente, recientemente las aplicaciones de inteligencia artificial han visto aumentado exponencialmente su uso y desarrollo. Estos algoritmos se basan en procesos no triviales para extraer conocimiento e información potencialmente útil [4]. Entre las diferentes técnicas de inteligencia artificial, hay numerosos algoritmos que permiten realizar modelos predictivos. En el presente estudio, los algoritmos aplicados están centrados en el enfoque lineal, el enfoque no lineal y el aprendizaje máquina [5].

Por tanto, dada la falta de aplicaciones que cubran este nicho, se observa la necesidad de estudiar varios factores relacionados con las actividades quirúrgicas de manera predictiva.

Por consiguiente, el objetivo de este estudio es predecir el valor de varias características relacionadas con la ergonomía (cinemática del cirujano) y la fisiología (EDA, temperatura corporal y presión arterial) del cirujano a partir de los valores recopilados en la situación inmediatamente previa, con el fin de predecir y prevenir posibles riesgos para la salud del cirujano.

2. Material y métodos

2.1. Diseño experimental

Se han recopilado datos durante 26 sesiones de cirugía robótica completadas por 11 cirujanos, siendo 6 cirujanos con un nivel de experiencia experto y 5 cirujanos con un nivel de experiencia novel, siendo 2 mujeres y 9 hombres de entre 25 y 50 años de edad. Los datos recopilados son los referentes a los diversos sistemas como son el sensor EDA, la temperatura corporal del cirujano, a la presión arterial y a la cinemática del cirujano (posición, pasos y velocidad angular), los cuales se recopilan de la situación actual y de la inmediatamente anterior. Estos datos se integran en un único dispositivo y se extrajeron usando el software *Unisens Viewer (Movisens, Karlsruhe, Alemania)*. Sobre los datos registrados se han aplicado diferentes técnicas de preprocesado, escalado y escalado y normalizado, generando tres conjuntos de datos: un 80% de los datos para la calibración, un 10% para la validación

y un 10% para el test. Sobre los conjuntos de datos de calibración se han aplicado diferentes técnicas predictivas de inteligencia artificial para generar modelos predictivos. Finalmente, estos modelos predictivos fueron validados sobre los conjuntos de datos de validación y de test. La Figura 1 muestra el diseño experimental seguido en el estudio.

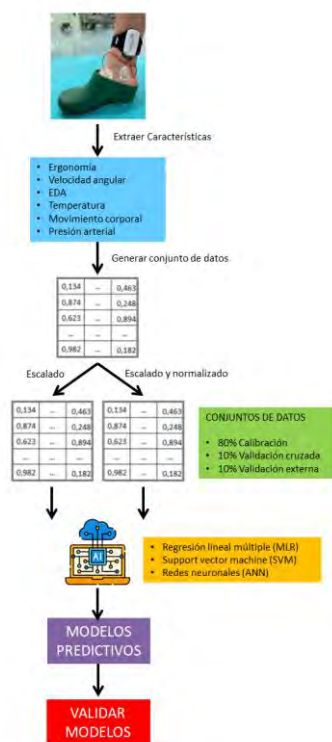


Figura 1. Diseño Experimental.

2.2. Robótica y Sensorización

Para la realización de las 26 sesiones quirúrgicas se utilizaron 2 plataformas robóticas para cirugía de mínima invasión: La plataforma SYMANI (MMI, Pisa, Italia) (<http://mmimicro.com/ourtechnology/how-it-works/>) para cirugía robótica microquirúrgica y la plataforma VERSIUS (CMR Surgical Ltd, Cambridge, Reino Unido) (<http://es.emrsurgical.com>) para cirugía robótica laparoscópica. Aplicandose en 14 sesiones la plataforma SYMANI y en 12 la plataforma VERSIUS.

Los sensores utilizados para la recopilación de los datos durante las cirugías robóticas fueron el sensor EDA, el cual recopiló datos fisiológicos y de la cinemática del cirujano. Todos ellos se encontraban integrados en un único dispositivo EdaMove 4 (Movisens, Karlsruhe, Alemania) (<http://movisens.com/en/products/eda-and-activity-sensor/>), el cual se ajusta al tobillo del cirujano (Figura 1).

2.3. Conjuntos de datos

A partir de los datos recopilados por el dispositivo EdaMove4, se generó un conjunto de datos constituido por 2860 tuplas. Para este estudio se utilizaron dos conjuntos de datos, generados a partir de aplicar diferentes técnicas de preprocesado, el de los datos escalados y el de los datos escalados y normalizados.

Una vez se han generado el conjunto de datos, estos fueron organizados en tres subconjuntos: el 80% (2288 tuplas) para generar los modelos predictivos; un 10% para la validación (286 tuplas); y un 10% para el test (286 tuplas).

2.4. Inteligencia artificial

En cuanto a las técnicas de inteligencia artificial, como técnicas predictivas se usaron 3 técnicas diferentes, una que genera modelos lineales (MLR; Multiple Linear Regression), una que genera modelos no lineales (SVM; Support Vector Machine) y otra relacionada con las redes neuronales (ANN; Artificial Neural Network).

MLR es una técnica predictiva que consiste en representar la relación lineal entre una variable dependiente y numerosas variables independientes. Con esta técnica se obtiene una ecuación que puede usarse para predecir valores futuros [5]. SVM se basa en la transformación del espacio de entrada en otro espacio de una dimensión superior en el que se puede resolver el problema. Para ello, se aplica un mapeado no lineal que depende de una función que es un producto interior al espacio de características [5]. ANN es una técnica de inteligencia artificial que intenta simular el funcionamiento del cerebro humano, mediante una combinación de modelos computacionales que se denominan neuronas artificiales, conectadas entre sí para transmitirse valores y señales. De tal manera que el modelo generado va aprendiendo basándose en los valores y señales que se transmiten por el sistema [5]. De todas las ANN se ha aplicado un Multilayer Perceptron ya que viene configurada en la herramienta suite de inteligencia artificial que se ha utilizado.

De manera previa a la aplicación de las técnicas predictivas, se han aplicado dos tipos diferentes de preprocesado como son el escalado y la normalización, dando lugar a dos conjuntos de datos distintos. El escalado permite que todos los datos para cada una de las características se escalen a valores que se encuentran entre 0 y 1 (ecuación 1). La normalización permite que los datos previamente escalados puedan ser transformados a otro conjunto de datos con mayor integridad y una menor redundancia de los datos (ecuación 2).

$$Dato_{escalado} = \frac{Dato_{inicial} - valor_{mínimo}}{valor_{máximo} - valor_{mínimo}} \quad (1)$$

$$Dato_{normalizado} = \frac{Dato_{escalado} - valor_{medio}}{Valor_{Desviación\ típica}} \quad (2)$$

La aplicación de las técnicas previamente descritas se realizó usando los algoritmos disponibles en la herramienta suite de inteligencia artificial Weka 3.8.6 (Universidad de Waikato, Hamilton, Nueva Zelanda). En todos los casos se aplicó la configuración por defecto de los algoritmos. Para validar los resultados de los diferentes modelos se utilizó el coeficiente R² y el error relativo cuadrático medio (ERCM).

3. Resultados y discusión

Con el objetivo de predecir las diferentes características durante las sesiones quirúrgicas, se comenzó el estudio

		MLR	SVM	ANN
Ergonomía X	R ²	0,9063	0,9061	0,8972
	ERCM	0,0449	0,0784	0,0806
Ergonomía Y	R ²	0,8458	0,8444	0,8302
	ERCM	0,0236	0,0478	0,0499
Ergonomía Z	R ²	0,9255	0,9248	0,9024
	ERCM	0,0361	0,0438	0,0497
Veloc. Angular X	R ²	0,5475	0,5156	0,3776
	ERCM	0,0238	0,0691	0,0862
Veloc. Angular Y	R ²	0,5117	0,3182	0,2786
	ERCM	0,0202	0,0543	0,0614
Veloc. Angular Z	R ²	0,3858	0,3444	0,3208
	ERCM	0,0207	0,0532	0,0633
EDA	R ²	0,9844	0,9825	0,9823
	ERCM	0,0144	0,0299	0,0317
T ^a	R ²	0,9985	0,9965	0,9956
	ERCM	0,0035	0,0061	0,0077
Mov.	R ²	0,7856	0,7744	0,7255
	ERCM	0,0343	0,0717	0,0831
Presión Arterial	R ²	0,9998	0,9981	0,9969
	ERCM	0,0009	0,0023	0,0038

Tabla 1. Resultados obtenidos para el conjunto de datos ESC (con escalado) y para realizar la calibración (CAL).

recopilando los datos y generando los dos conjuntos de datos a partir del conjunto de datos original, uno de ellos con los datos escalados para cada característica (cinemática, EDA, temperatura corporal y presión arterial) y otro con los datos escalados y normalizados para las mismas características. Una vez se crearon ambos conjuntos de datos, se le añadieron los datos de las situaciones inmediatamente previas. Por último, se dividieron ambos conjuntos de datos en 3 subconjuntos: el 80% para la calibración, el 10% para la validación y el 10% para el test.

3.1. Resultados de calibración y de validación

Una vez se aplicaron las tres técnicas de inteligencia artificial, se generaron los modelos para todas las características. La Tabla 1 muestra los resultados de la calibración de los modelos predictivos sobre el conjunto de datos escalado.

Los resultados de calibración de los modelos generados para el conjunto de datos escalado y normalizado son mostrados en la Tabla 2.

De los resultados obtenidos para la calibración de los modelos, se puede deducir que los resultados obtenidos son similares aplicando las 3 técnicas, aunque son ligeramente mejores aplicando MLR. Por consiguiente, la validación y el test se han evaluado únicamente aplicando MLR. La Tabla 3 muestra los resultados obtenidos para la

validación cruzada en los conjuntos con ambos tipos de preprocesado.

		MLR	SVM	ANN
Ergonomía X	R ²	0,9063	0,9060	0,8972
	ERCM	0,0423	0,0432	0,0445
Ergonomía Y	R ²	0,8458	0,8454	0,8302
	ERCM	0,0265	0,0538	0,0561
Ergonomía Z	R ²	0,9255	0,9248	0,9024
	ERCM	0,0378	0,0383	0,0435
Veloc. Angular X	R ²	0,5473	0,5165	0,4776
	ERCM	0,0840	0,0899	0,1221
Veloc. Angular Y	R ²	0,5117	0,3177	0,2786
	ERCM	0,0371	0,0995	0,1254
Veloc. Angular Z	R ²	0,3859	0,3432	0,3208
	ERCM	0,0366	0,0941	0,1426
EDA	R ²	0,9844	0,9825	0,9823
	ERCM	0,0759	0,0798	0,1053
T ^a	R ²	0,9985	0,9978	0,9966
	ERCM	0,0551	0,0653	0,0729
Mov.	R ²	0,7856	0,7739	0,7255
	ERCM	0,0621	0,0648	0,0822
Presión Arterial	R ²	0,9988	0,9980	0,9969
	ERCM	0,0059	0,0069	0,0117

Tabla 2. Resultados obtenidos para el conjunto de datos ESC NORM (con escalado y normalizado) y para realizar la calibración (CAL).

Analizando los resultados de calibración y de validación se puede observar que los resultados de movimiento corporal del cirujano, EDA, temperatura corporal y presión arterial son los que alcanzaron el coeficiente R² más alto y los valores más bajos para la velocidad angular.

3.2. Resultados de test

Para validar los modelos predictivos generados previamente, se aplicaron directamente los modelos generados sobre el subconjunto de datos para el test. La Tabla 4 muestra los resultados obtenidos.

Los resultados obtenidos vienen a confirmar los resultados obtenidos tanto en la calibración como como en la validación, mostrando valores del coeficiente R² superiores a 0,6 para todos los casos y un ERCM menor de 0,13 para todas las características, validando así los resultados obtenidos [6]. Además, estos resultados concuerdan con las actividades quirúrgicas realizadas, ya que la velocidad angular tiene una alta variación en función de la plataforma robótica utilizada y la consecuente mayor dificultad para realizar la predicción.

		ESC	ESC NORM
Ergonomía X	R ²	0,9074	0,8956
	ERCM	0,0493	0,0683
Ergonomía Y	R ²	0,8506	0,8505
	ERCM	0,0271	0,0232
Ergonomía Z	R ²	0,9265	0,9265
	ERCM	0,0409	0,0409
Veloc. Angular X	R ²	0,5635	0,5633
	ERCM	0,0318	0,0385
Veloc. Angular Y	R ²	0,5126	0,5125
	ERCM	0,0299	0,0299
Veloc. Angular Z	R ²	0,4961	0,4862
	ERCM	0,0272	0,0258
EDA	R ²	0,9846	0,9845
	ERCM	0,0144	0,0143
T ^a	R ²	0,9984	0,9984
	ERCM	0,0036	0,0035
Mov.	R ²	0,7599	0,7528
	ERCM	0,0461	0,0256
Presión Arterial	R ²	0,9998	0,9998
	ERCM	0,0010	0,0039

Tabla 3. Resultados para el conjunto de datos ESC (con escalado) y ESC NORM (con escalado y normalizado) para evaluar la validación (VAL) del modelo lineal (MLR).

4. Conclusiones

En este estudio se ha conseguido predecir todos los parámetros fisiológicos y ergonómicos del cirujano durante la práctica en robótica quirúrgica a partir de los valores registrados de las situaciones inmediatamente previas. Los mejores resultados han sido obtenidos con los modelos predictivos lineales. Estos resultados abren la puerta a novedosas aplicaciones para prevenir situaciones de riesgo para la salud del cirujano durante la práctica robótica quirúrgica con el consiguiente beneficio para la calidad asistencial.

Agradecimientos

Este estudio ha sido cofinanciado por la Junta de Extremadura (TA18023) y el Plan Complementario Biotecnología Aplicada a la Salud, cofinanciado por el Ministerio de Ciencia e Innovación con fondos de la Unión Europea NextGenerationEU, el Plan de Recuperación, Transformación y Resiliencia (PRTR-C17.I1) y el Programa Operativo FEDER Extremadura 2021-2027. Sin embargo, los puntos de vista y las opiniones expresadas son únicamente los de los autores. Agradecer también la realización del presente estudio a los compañeros que colaboraron en las sesiones de cirugía robótica (Elena Abellán, Laura Pires, Javier Vela, José

Luis Campos, Carlos Sánchez, David Durán, José Ramón Torres, Rubén Baltá y Pedro Olivares).

		ESC	ESC NORM
Ergonomía X	R ²	0,9077	0,9041
	ERCM	0,0858	0,0923
Ergonomía Y	R ²	0,8545	0,8531
	ERCM	0,0848	0,0887
Ergonomía Z	R ²	0,8715	0,8715
	ERCM	0,1277	0,1499
Veloc. Angular X	R ²	0,6195	0,6192
	ERCM	0,1263	0,0921
Veloc. Angular Y	R ²	0,7069	0,7061
	ERCM	0,0574	0,0644
Veloc. Angular Z	R ²	0,7120	0,6867
	ERCM	0,0850	0,0804
EDA	R ²	0,9547	0,9541
	ERCM	0,0778	0,0776
T ^a	R ²	0,9957	0,9961
	ERCM	0,0183	0,0179
Mov.	R ²	0,9975	0,9976
	ERCM	0,0281	0,0252
Presión Arterial	R ²	0,7402	0,7602
	ERCM	0,1068	0,0851

Tabla 4. Resultados obtenidos para el conjunto de datos ESC (con escalado) y ESC NORM (con escalado y normalizado) para realizar el test (TEST) del modelo lineal (MLR).

Referencias

- [1] Healey JA, Picard RW. Detecting stress during real-world driving tasks using physiological sensors. *IEEE transactions on Intelligent Transportation Systems*, vol 6, 2005, pp. 156-66. (ISSN: 1524-9050).
- [2] Saoughi F, Behmanesh A, Sayfour N. Internet of things in medicine: A systematic mapping study. *Journal of Biomedical Informatics*, vol 103 (103383), 2020, pp. 1-20 (ISSN: 1532-0480)
- [3] Ballantyne GH. The pitfalls of laparoscopic surgery, telepresence, and telemonitoring. Review of early clinical results. *Surgical endoscopy*, vol 16, 2002, pp. 1389-402. (ISSN: 0930-2794).
- [4] Ávila-Tomás JF, Mayer-Pujadas MA, Quesada-Varela VJ. La inteligencia artificial y sus aplicaciones en medicina I: introducción antecedentes a la IA y robótica. *Atención primaria*, vol 52, 2020, pp. 778-84. (ISSN: 0212-6567).
- [5] Rath M, Satpathy J, Oreku GS. Artificial Intelligence and machine learning applications in cloud computing and internet of things. *Artificial Intelligence to solve pervasive internet of things issues*. Academic Press, Chapter 6, 2021, pp. 103-23. (ISBN: 978-0-12-818576-6)
- [6] Colton T. Statistics in medicine. Ed. Little Brown and Co. 1974 (ISBN: 978-0-31-615249-5)

Imágenes Médicas (II)

Jueves 23 8:30 - 10:00

Paraninfo

Predicción automática del diagnóstico del sarcoma de Ewing en una serie extensa de tumores de células redondas incluidas en micromatrices de tejidos

Pablo Meseguer^{1,2}, Ana Rubio¹, Rocío del Amor¹, Jose Antonio Lopez-Guerrero³, Samuel Navarro^{4,5}, Katia Scotlandi⁶, Antonio Llombart-Bosch⁴, Isidro Machado^{4,5,7,8}, Valery Naranjo¹

¹ Instituto Universitario de Investigación e Innovación en Tecnología Centrada en el Ser Humano, HUMAN-tech, Universitat Politècnica de València, Valencia, España. pabmees@upv.es

² valgrAI - Valencian Graduate School and Research Network of Artificial Intelligence, Valencia, España

³ Laboratorio de Biología Molecular, Instituto Valenciano de Oncología, Valencia, España

⁴ Departamento de Patología, Universidad de Valencia, Valencia, España

⁵ Cancer CIBER (CIBERONC), Madrid, España

⁶ Laboratorio de Patología Experimental. Instituto Ortopédico de Rizzoli, Bologna, Italia

⁷ Departamento de Patología, Instituto Valenciano de Oncología, Valencia, España

⁸ Laboratorio Patologika, Hospital QuironSalud, Valencia, España

Resumen

El sarcoma de Ewing (SE) es una neoplasia maligna que afecta principalmente a personas jóvenes y se origina en el tejido óseo o en los tejidos blandos circundantes. El estudio histopatológico de las biopsias es crucial en el proceso diagnóstico del SE, pero está limitado debido a su diversidad histológica y a la similitud con otros tumores de células redondas y pequeñas. Además, resulta fundamental realizar un diagnóstico diferencial del SE puesto que requiere de un tratamiento particular. En este contexto, la implementación de sistemas automáticos de ayuda al diagnóstico basados en inteligencia artificial (IA) puede contribuir a orientar a los médicos en la selección de técnicas diagnósticas adicionales para confirmar el diagnóstico. El presente trabajo explora la implementación de algoritmos basados en dos paradigmas de aprendizaje profundo (supervisado y débilmente supervisado) para la predicción del sarcoma de Ewing sobre micromatrices de tejido digitalizadas. Los resultados muestran que ambos permiten la clasificación del SE frente al rhabdomyosarcoma con un desempeño mayor del 95 % en términos de exactitud demostrando la capacidad de los algoritmos basados en IA para asistir a los patólogos en el proceso diagnóstico.

1. Introducción

Los sarcomas son un tipo de neoplasias que se originan en los huesos o los tejidos blandos y que suponen más del 20 % de las neoplasias malignas pediátricas [1]. En concreto, el sarcoma de Ewing (SE) es un tipo de tumor óseo maligno que se caracteriza por su rápida propagación y agresividad. Aunque se trata de un cáncer con una baja incidencia a nivel global, su impacto social es considerable puesto que afecta mayoritariamente a niños y adolescentes [2]. La tasa de supervivencia a 5 años del SE depende en gran medida de la extensión del tumor puesto que se reduce desde un 68 % en tumores localizados hasta un 39 % cuando existe metástasis [3]. Estas estadísticas resaltan la necesidad de conseguir un diagnóstico precoz y preciso del tumor.

El análisis histológico de las biopsias es una herramienta fundamental en el diagnóstico del SE. La histología del SE está constituida mayoritariamente por células redondas y en menor cuantía por ovoides y fusiformes de pequeño tamaño distribuidas en un patrón ho-

mogéneo, características también presentes en la familia de tumores de células redondas y pequeñas (TCRP). El estudio histopatológico se puede realizar mediante micromatrices de tejido (TMA) que incluyen múltiples cortes micrométricos circulares de biopsias cilíndricas permitiendo así su análisis comparativo. Esta técnica es diferente a las secciones de biopsias completas que posibilita la observación detallada de todo el tejido. Con los avances en patología digital y la aparición de escáneres especializados, se ha conseguido la digitalización de las muestras de tejido histológico de las biopsias permitiendo así la implementación de algoritmos de visión por computador.

Anteriormente, se ha investigado la implementación de algoritmos de aprendizaje profundo para la detección del sarcoma de Ewing en radiografías pediátricas [4]. Sin embargo, hasta donde conocemos, ningún trabajo ha propuesto la aplicación de algoritmos de inteligencia artificial (IA) para la predicción del diagnóstico del SE sobre imágenes histológicas. En este trabajo, se propone la comparación de diferentes paradigmas de aprendizaje profundo para la predicción del diagnóstico de sarcomas de Ewing sobre micromatrices de tejido. En concreto, se explora el aprendizaje supervisado con una reducción de la resolución asumible y el aprendizaje débilmente supervisado basado en múltiples instancias, paradigma ampliamente utilizado para el análisis automático de imágenes histológicas digitalizadas.

2. Materiales

El presente trabajo se centra en el diseño de un sistema de análisis automático de muestras histológicas con la finalidad de investigar si esta herramienta pudiera asistir a los patólogos en el diagnóstico diferencial del SE. La base de datos se constituye a partir de cortes micrométricos circulares de biopsias cilíndricas extraídas a pacientes con una confirmación genética del diagnóstico [5]. Estas muestras siguieron las técnicas histológicas, fueron colocadas ordenadamente en micromatrices de tejido y digitalizadas con escáneres de patología digital obteniendo así imágenes histológicas digitalizadas

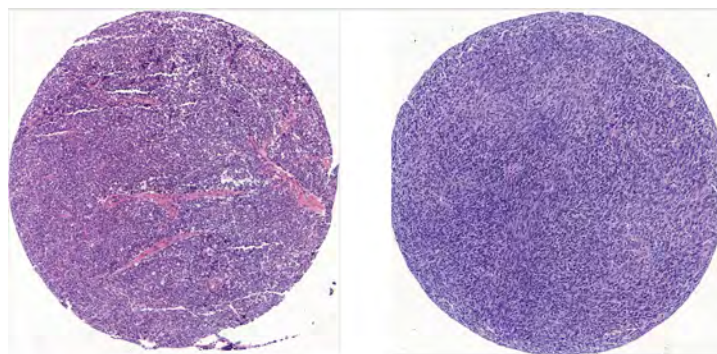


Figura 1. Biopsia de sarcoma de Ewing (izquierda) y de rhabdomiosarcoma (derecha).

(WSI). Aunque la base de datos incluyó diferentes tipos de TCRP, este trabajo se centra en la diferenciación del SE con el rhabdomiosarcoma (RMS). El RMS es un tumor maligno con diferenciación miogénica y es el sarcoma de tejidos blandos más común en pacientes pediátricos [6]. La Figura 1 muestra dos secciones circulares (*cores*) en que se puede comparar las características histológicas del SE y del RMS.

El resumen de la base de datos en términos de número de TMA, secciones y parches por sección se presenta en la Tabla 1. Cabe destacar que algunos *cores* de los TMA con muestras de SE fueron considerados no informativos debido a la presencia de artefactos de descalcificación o necrosis u otros tejidos no tumorales por lo que el número de secciones por TMA varía notablemente entre el SE y el RMS. El tamaño de las imágenes de secciones se sitúa sobre los 3000^2 píxeles. Para la implementación del aprendizaje débilmente supervisado, se extraen parches recortados de cada biopsia con un tamaño de 512^2 píxeles con un solapamiento del 50%.

	S. de Ewing	Rabdomiosarcoma
TMA	26	6
Secciones	203	257
Parches ($\mu \pm \sigma$)	92 ± 15	101 ± 24

Tabla 1. Resumen de la base de datos.

3. Métodos

Un resumen de la metodología propuesta para el diseño de un sistema de predicción automática del sarcoma de Ewing se presenta en la Figura 2 y cada uno de sus componentes se detalla a continuación.

3.1. Aprendizaje supervisado

El bloque de aprendizaje supervisado toma como datos de entrada cada una de las secciones cilíndricas pertinentemente anotada en función de la confirmación genética del paciente. Con estos algoritmos se pretende investigar si es factible la clasificación de los *cores* empleando imágenes a menor resolución debido a las limitaciones computacionales asociadas a imágenes gigapixel. Así se

pretende investigar diferentes arquitecturas de red neuronal convolucional (CNN) para determinar cual presenta la capacidad de extracción de características más discriminatorias en este tipo de imágenes. Se exploran tres CNN ampliamente utilizadas en visión por computador como VGG-16 [7], ResNet-50 [8] e InceptionV3 [9].

Cabe destacar que el aprendizaje de transferencia puede colaborar mejorando la resolución de problemas con escasa disponibilidad de datos empleando el conocimiento adquirido en grandes bases de datos. Además, las imágenes naturales e histológicas presentan características notablemente diferentes por lo que cabe determinar el método de transferencia óptimo para ajustarse a las particularidades de estas imágenes. Se exploran tres aproximaciones de aprendizaje de transferencia con el objetivo de determinar la óptima a aplicar en imágenes histológicas. Estas son la extracción de características (FE), el ajuste fino (FT) y la inicialización de los pesos (WI). Todos los modelos han sido inicializados en ImageNet [10] y en el caso del ajuste fino se han congelado los pesos de los dos primeros bloques convolucionales de cada red.

3.2. Aprendizaje de múltiples instancias

El aprendizaje de múltiples instancias (MIL) se caracteriza por la formación de bolsas que incluyen múltiples parches de cada imagen. En problemas bajo el paradigma MIL, se considera una bolsa positiva si contiene al menos una instancia de la clase positiva y el objetivo es predecir la etiqueta global de la biopsia, no la de cada una de sus instancias. La formulación de MIL basada en características se divide en tres etapas: (1) extracción y (2) agregación de características a nivel de instancia y (3) clasificación. La primera etapa se realiza con el método de extracción de características optimizado en el aprendizaje supervisado mientras que en esta sección se pretende investigar diferentes algoritmos de agregación de características para la obtención del vector global de la biopsia. Inicialmente, se propusieron funciones como la agregación de máximos (BGMP) o la media (BGAP) para la obtención directa del vector global representativo de la biopsia. También se han propuesto diferentes módulos de agregación entrenables

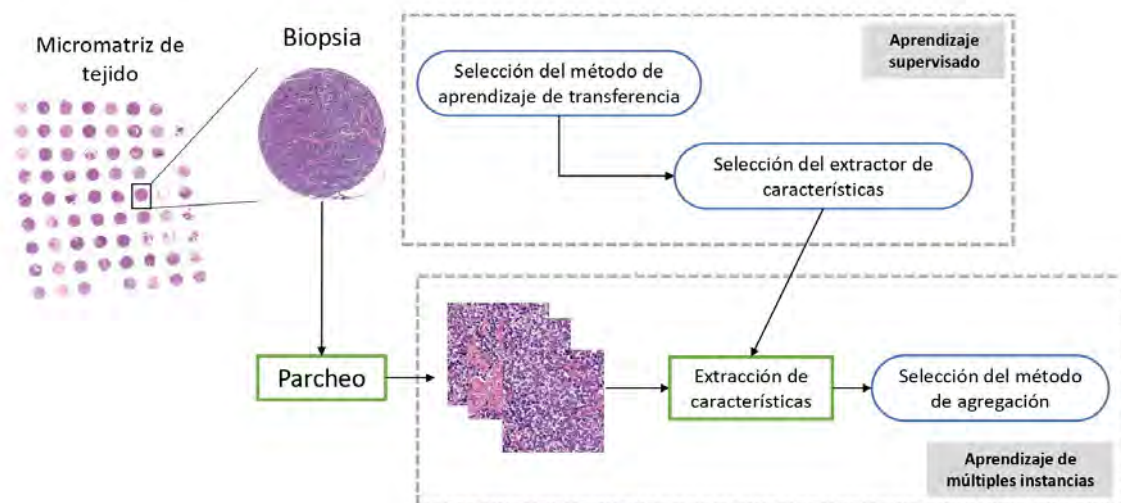


Figura 2. Resumen de la metodología propuesta.

que pretenden encontrar las instancias más significativas mediante atención (MILAtt) [11] o determinar la relación entre las mismas empleando mecanismos de auto-atención (TransMIL) [12].

4. Resultados y discusión

4.1. Detalles de implementación Todos los algoritmos fueron entrenados para minimizar la entropía cruzada binaria con el optimizador Adam y una tasa de aprendizaje $\eta = 10^{-4}$. Los algoritmos de aprendizaje supervisado se entrenan durante 15 épocas y un tamaño de lote igual a 8, mientras que los de MIL lo hacen durante 25 épocas con un tamaño de lote unitario. Debido a limitaciones computacionales, se reducen las secciones completas a un tamaño de 750^2 píxeles y los parches a 224^2 píxeles. Los modelos se evalúan en función de la especificidad (ESP), sensibilidad (SEN), exactitud (ACC) y la métrica F1 (F1S) considerando el sarcoma de Ewing la clase positiva. La partición de la base de datos en los conjuntos de entrenamiento, validación y test se realiza con los siguientes porcentajes: 70/10/20.

4.2. Experimentos de ablación

En esta sección, se presenta una serie de experimentos de ablación diseñado evaluados en el conjunto de validación con el objetivo de validar los diferentes componentes del sistema.

1) Ablación de la red convolucional y el método de aprendizaje de transferencia. La Tabla 2 incluye las métricas de evaluación de la combinación de las redes convolucionales y los métodos de transferencia. Estas demuestran un desempeño inferior de la extracción de características lo que manifiesta que las características obtenidas a partir de redes entrenadas en imágenes naturales no son suficientemente discriminatorias para la clasificación de imágenes histológicas. Por su parte, el ajuste fino mejora un 64.5% y un 3.8% en términos de F1S a la inicialización de los pesos para

las arquitecturas ResNet-50 y InceptionV3. Los resultados destacan la utilidad de mantener la extracción de características congelada en los primeros bloques convoluciones con la finalidad de extraer características más generales de las imágenes y conseguir una convergencia más rápida de la red. Aunque el ajuste fino presente el mismo desempeño en las redes VGG16 e InceptionV3, se escoge la arquitectura VGG16 para la extracción de características de MIL debido a que ya ha demostrado su desempeño superior en diferentes configuraciones.

		ESP	SEN	F1S	ACC
VGG16 [7]	FE	0,920	0,400	0,533	0,689
	FT	1	0,900	0,947	0,956
	WI	1	0,900	0,947	0,956
ResNet-50 [8]	FE	0,840	0,450	0,545	0,667
	FT	0,960	0,850	0,895	0,911
	WI	0,960	0,150	0,250	0,600
InceptionV3 [9]	FE	0,880	0,800	0,821	0,844
	FT	1	0,900	0,947	0,956
	WI	0,840	1	0,909	0,911

Tabla 2. Ablación de la red convolucional y el método de aprendizaje de transferencia.

2) Ablación de los métodos de agregación en aprendizaje de múltiples instancias. La Tabla 3 muestra las figuras de mérito de los diferentes métodos de agregación de características. Aunque la agregación de máximos (BGMP) demuestre un mejor desempeño global alcanzando una exactitud $ACC = 95,6\%$ con excelente especificidad ($ESP = 100\%$), destaca la capacidad del método MILAtt [11] de determinar las instancias más significativas para determinar el sarcoma de Ewing con una sensibilidad superior ($SEN = 95\%$). Estudios previos han demostrado que la incorporación de información contextual en TransMIL [12] mejora la predicción a nivel de biopsia, cabe la posibilidad que en estas secciones circulares pierda relevancia al encontrarse el tumor en toda la biopsia.

	ESP	SEN	F1S	ACC
BGAP	0.920	0.800	0.863	0.867
BGMP	1	0.900	0.955	0.956
MILAtt [11]	0.920	0.950	0.933	0.933
TransMIL [12]	0.960	0.850	0.909	0.911

Tabla 3. Ablación de los métodos de agregación de características.

4.3. Resultados de test

Finalmente, se presenta como validación final en la Tabla 4 una comparativa del desempeño de ambos algoritmos en el conjunto de test. De este modo, se comparan tanto el aprendizaje supervisado con el ajuste fino de la red VGG-16 como el aprendizaje de múltiples instancias con este mismo extractor y BGMP y MILAtt [11] como métodos de agregación de características. Los resultados demuestran la capacidad de ambos paradigmas de aprendizaje de diferenciar correctamente el SE y el RMS consiguiendo resultados un 1 % y 3.2 % superiores en términos de exactitud mediante el paradigma más sencillo como es el aprendizaje supervisado en comparación con ambos métodos de MIL.

	ESP	SEN	F1S	ACC
Sup: VGG + FT	1	0.976	0.988	0.989
MIL: VGG + BGMP	0.981	0.976	0.978	0.979
MIL: VGG + MILAtt	0.925	1	0.957	0.957

Tabla 4. Resultados de test comparativos del entrenamiento supervisado y de múltiples instancias.

5. Conclusiones

En el manejo del sarcoma de Ewing, el diagnóstico diferencial y temprano es fundamental puesto que la tasa de supervivencia está asociada a la presencia de metástasis. Por ello, es necesario investigar en herramientas que asistan y guíen a los patólogos en el proceso diagnóstico. La digitalización de las micromatrices de tejido permite la implementación de diferentes paradigmas de aprendizaje debido al tamaño de las secciones y a las limitaciones computacionales implícitas en las imágenes histológicas. Este trabajo ha demostrado la viabilidad de implementación de paradigmas de aprendizaje tanto supervisado como basados en múltiples instancias. Se ha comprobado que la arquitectura VGG16 consigue un desempeño superior en la extracción de características de imágenes histológicas y que la agregación basada en atención determina las instancias más significativas para la predicción del sarcoma de Ewing. Cabe destacar que únicamente se ha investigado la predicción del SE frente al rhabdomyosarcoma por lo que futuros trabajos deberían explorar otros subtipos de sarcomas similares a nivel histológico.

Agradecimientos

Este trabajo ha recibido financiación del Ministerio de Economía y Competitividad a través de los proyectos PID2019-105142RB-C21 (AI4SKIN) y PID2022-140189OB-C21 (AS-SIST) y de la GVA a través del proyecto CIPROM/2022/20

(COMTACTS2). Además, fue llevado a cabo por los proyectos PROTHETS (no. 503036) y EuroBonet (no. 018814) financiados por la Unión Europea.

Referencias

- [1] Zachary Burningham, Mia Hashibe, Logan Spector, and Joshua D Schiffman. The epidemiology of sarcoma. *Clinical sarcoma research*, 2(1):1–16, 2012.
- [2] Michael Paulussen, Birgit Fröhlich, and Herbert Jürgens. Ewing tumour: incidence, prognosis and treatment options. *Paediatric drugs*, 3:899–913, 2001.
- [3] Natia Esiashvili, Michael Goodman, and Robert B Marcus. Changes in incidence and survival of ewing sarcoma patients over the past 3 decades: Surveillance epidemiology and end results data. *Journal of pediatric hematology/oncology*, 30(6):425–430, 2008.
- [4] Sarah Consalvo, Florian Hinterwimmer, Jan Neumann, Marc Steinborn, Maya Salzman, Fritz Seidl, Ulrich Lenze, Carolin Knebel, Daniel Rueckert, and Rainer HH Burgkart. Two-phase deep learning algorithm for detection and differentiation of ewing sarcoma and acute osteomyelitis in paediatric radiographs. *Anticancer Research*, 42(9):4371–4380, 2022.
- [5] Antonio Llombart-Bosch, Isidro Machado, Samuel Navarro, Franco Bertoni, Patrizia Bacchini, Marco Alberghini, Apollon Karzeladze, Nikita Savelov, Semyon Petrov, Isabel Alvarado-Cabrero, et al. Histological heterogeneity of ewing’s sarcoma/pnet: an immunohistochemical analysis of 415 genetically confirmed cases with clinical support. *Virchows Archiv*, 455:397–411, 2009.
- [6] Roshni Dasgupta, Jörg Fuchs, and David Rodeberg. Rhabdomyosarcoma. In *Seminars in pediatric surgery*, volume 25, pages 276–283. Elsevier, 2016.
- [7] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [9] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.
- [10] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.
- [11] Maximilian Ilse, Jakub Tomczak, and Max Welling. Attention-based deep multiple instance learning. In *International conference on machine learning*, pages 2127–2136. PMLR, 2018.
- [12] Zhuchen Shao, Hao Bian, Yang Chen, Yifeng Wang, Jian Zhang, Xiangyang Ji, et al. Transmil: Transformer based correlated multiple instance learning for whole slide image classification. *Advances in neural information processing systems*, 34:2136–2147, 2021.

Evaluation of machine learning algorithms and relevant biomarkers for the diagnosis of multiple sclerosis based on optical coherence tomography

Pablo García Mesa¹, Pilar Rojas Lozano², Nuria Díaz Gutiérrez², Manuel Cadena Santoyo², Alberto J. Beltrán Carrero¹, Juan J. Gómez-Valverde^{1,3}

¹Biomedical Image Technologies, ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain.

²Instituto Provincial de Oftalmología, Hospital General Universitario Gregorio Marañón, Madrid, Spain.

³Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Madrid, Spain.

Abstract

Multiple sclerosis (MS) is a prevalent neurodegenerative disease with significant visual pathway-related symptoms. Optical coherence tomography (OCT) has emerged as a valuable tool, and machine learning (ML) techniques hold promise for MS diagnosis. However, existing studies often lack comprehensive feature exploitation and require interpretable model analysis to improve clinical insights and diagnostic criteria. This study evaluates machine learning models for classification of healthy controls and MS patients using a comprehensive set of macular and optic-disc parameters from OCT imaging. The study included a dataset of 77 MS eyes and 54 control eyes, obtained by ophthalmic examination and OCT measurements from Optic Disc and Macular Cube scan protocols of a Cirrus HD-OCT 5000 (Carl Zeiss, Meditec, Dublin, CA, USA). Our results identified 19 features, validated by p -values ($p < 0.001$), as effective discriminators between MS patients and healthy controls. Patient-wise cross-validation is used to evaluate the performance of five ML algorithms. Gaussian Naive Bayes achieved the best AUC ($87.9\% \pm 7.7\%$), while SHAP analysis reinforced the alignment with clinical observations of MS-related visual pathway changes and ganglion cell layer degeneration, with minimum ganglion cell thickness being the feature with the highest impact on classification. These findings underscore the potential of OCT-ML for early diagnosis and personalized treatment of MS.

1. Introduction

Multiple sclerosis (MS) is an autoimmune and neurodegenerative disease in which the myelin sheath surrounding nerve cells in the brain and spinal cord is damaged. This condition causes the improper transmission of nerve impulses, leading to various potential disabilities, with partial or total blindness, sensory loss, and motor disorders being the most common [1]. It is estimated that 2.8 million people worldwide are living with MS, with a higher prevalence of 140 cases per 100,000 population in Europe and the Americas [2].

Diagnosis and monitoring of MS rely on the integration of clinical, imaging and laboratory evidence, which may involve invasive procedures such as contrast-enhanced magnetic resonance imaging (MRI) and lumbar punc-

ture [3]. Several important symptoms of MS are related to visual pathway disorders, such as: optic neuritis, diplopia and oscillopsia. The assessment of ophthalmic patients has been revolutionized by optical coherence tomography (OCT), a rapid and reproducible imaging technique that employs low-coherence interferometry to generate cross-sectional images of the retina and optic nerve head (ONH) [4]. In recent years, several studies have revealed the presence of biomarkers associated with MS as well as other neurodegenerative diseases such as Parkinson's and Alzheimer's in the retina and optic disc. These biomarkers include specific structures like the ganglion cell layer (GCL), the retinal nerve fiber layer (RNFL) of the optic disc, and the optic nerve head (ONH). [5, 6, 7]. Furthermore, with the increasing popularity of artificial intelligence (AI) techniques, several studies have employed parameters from OCT to train machine learning (ML) algorithms for the diagnosis of MS [8, 9, 10]. However, most studies do not use a wide range of features extracted from both the macular and optic-disc regions. Furthermore, an analysis of the interpretability of the resulting models is crucial to properly understand the importance of each feature. In addition, this type of in-depth analysis can help clinicians to gain insights into the disease and improve diagnostic criteria [11].

This study focuses on the evaluation of a set of ML models for the classification of healthy controls and patients with MS based on a comprehensive collection of macular and optic-disc parameters obtained through an OCT imaging device. The study is composed of several key steps. First, a feature selection process is performed using a ML approach to identify the most crucial features for discerning between the two groups. This feature selection process is further supported by a statistical analysis to ensure the selection of relevant attributes. Next, a set of ML models is trained using the selected features and their performances are compared. This step aims to determine the most effective model for accurately classifying healthy controls and MS patients based on the OCT data. Finally, to improve the interpretability of the models' predictions, an explainable AI technique known as Shapely Additive Explanations (SHAP) is used, which allows for a detailed analysis of the importance of each individual feature during the in-

ference process. This step provides a rich breakdown of the models' decision criteria, which might be of great interest in clinical practice.

2. Materials

2.1. Dataset

This study included 40 patients with MS and 27 healthy controls recruited from the Neuro-ophthalmology department from Gregorio Marañón University Hospital in Madrid, Spain. The protocol was approved by the Ethics Committee of the hospital and all subjects were informed of the purpose and procedures of the study, and provided written consent to be included. The following inclusion criteria were applied: a refractive error within ± 5.00 dioptres for the equivalent sphere, and ± 2.00 dioptres for astigmatism. Exclusion criteria included prior intraocular surgery or any other conditions that could impact the visual field or nervous system, and the presence of pharmacological treatment that could affect visual function.

The participants received a complete ophthalmological examination and OCT measurements were obtained. The dataset included OCT volumes acquired using a Cirrus HD-OCT 5000 (Carl Zeiss, Meditec, Dublin, CA, USA), employing the Optic Disc Cube 200x200 and Macular Cube 512x128 scan protocols (6x6 mm). Only images with a signal-to-noise ratio of 7 or higher were included to ensure adequate acquisition quality, resulting in a final dataset comprising 77 MS eyes and 54 control eyes.

2.2. OCT Protocols

The Cirrus HD-OCT 5000 device provides a comprehensive evaluation of both retinal and optic-disc regions through a combination of quantitative and qualitative analyses. This study employs three distinct categories of analysis: Ganglion Cell Analysis (GCA), which assesses the condition of the GCL and the Inner Plexiform Layer (IPL); Macular Thickness Analysis (MTA), which measures the thickness of the macular region; and Optic-disc RNFL and ONH analyses. Figure 1 shows all the layers considered in this study. GCL and IPL are combined (GCIPL) in the GCA report of the device.

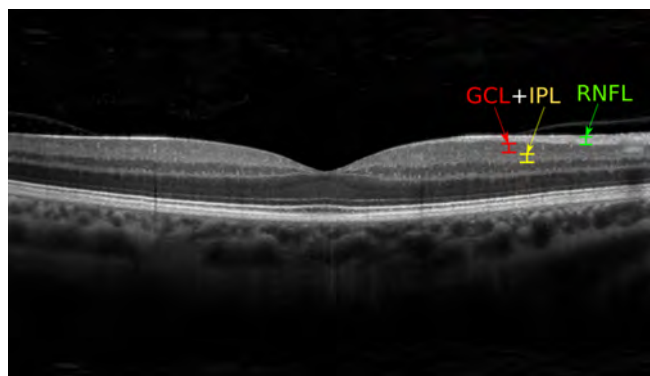


Figure 1: Macular OCT image example of a healthy control indicating the most important retinal layers, according to the feature selection process: RNFL and GCIPL, which combines GCL and IPL.

For GCA, the grid relies on six sector maps (Figure 2a), positioned at the fovea. The MTA analysis employs the early treatment diabetic retinopathy study (ETDRS) grid, automatically placed at the center of the fovea, effectively partitioning the macular area into nine distinct regions (Figure 2b). Lastly, the optic-disc analysis employs a clock-based grid (Figure 2c), dividing the area into twelve sectors and four regions corresponding to the anatomical locations (superior, temporal, inferior, nasal).

3. Methods

3.1. Feature selection

A feature selection process was performed to filter out the most relevant parameters and reduce overfitting. The aforementioned analyses (GCA, MTA, optic disc RNFL, and ONH) provided a total of 67 features. A tree-based method was used to remove the least relevant parameters, based on the feature importance values provided by the estimator [12]. More precisely, features that were assigned an importance value below the global mean feature importance were discarded. Significance analyses were performed using Student's t-tests to compare healthy controls and MS patients and validate the selected features.

3.2. Machine learning models and performance metrics

A comprehensive performance comparison of five different ML classifiers was conducted for this study. The selected models included Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), Bagging Classifier, Random Forest, and Gaussian Naive Bayes (NB), similarly to other reference studies [8, 9, 10]. For each model, the default combination of hyperparameters given by [12] was employed. The most relevant ones are summarized below:

- **SVM:** $C = 1$, $kernel = RBF$, $\gamma = 1/(N_{feat} \cdot \sigma_X^2)$
- **k-NN:** $num. neighbours = 5$.
- **Random Forest:** $num. trees = 100$, $max. features = \sqrt{N_{feat}}$.
- **Bagging Classifier:** $estimator = Decision Tree Classifier$, $num. estimators = 10$.

Several metrics were calculated to evaluate their effectiveness, including accuracy, F1 score, and area under the curve (AUC) [13]. We divided the dataset into a 70/30 percent split, ensuring that both eyes from the same patient were grouped together within each split to maintain inter-eye correlations during model training. To assess the model's performance, we employed 5-fold cross-validation.

3.3. SHAP analysis

We used the SHAP method [14] to improve our understanding of the impact of each feature on the model. SHAP values, which are rooted in cooperative game theory, provide insight into the contribution of each

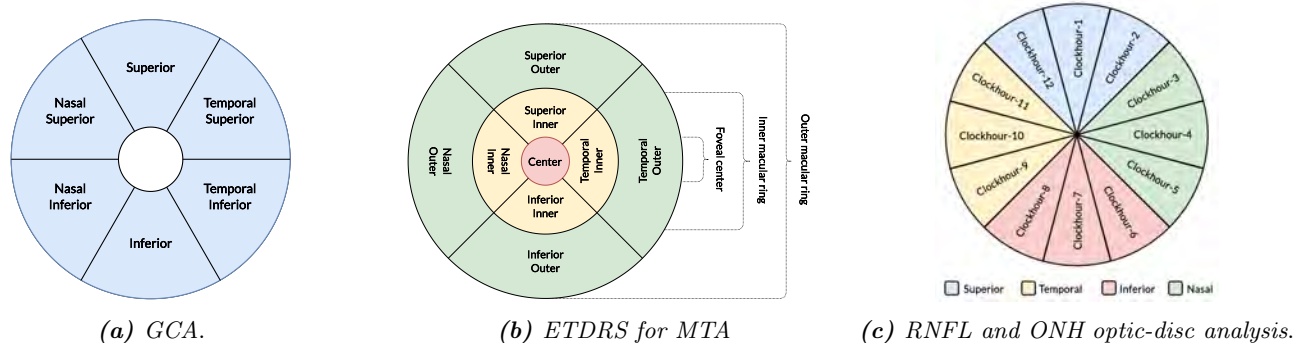


Figure 2: Sectors employed in the analysis of the Cirrus HD-OCT 5000 protocols.

feature to the model's predictions. For the validation set of each cross-validation iteration, we computed the SHAP values for all features in the best model, and then averaged the results from the five iterations to derive an estimate of feature importance.

4. Results and discussion

Macula features

GCA analysis

Ganglion cells Nasal Superior thickness
 Ganglion cells Nasal Inferior thickness
 Ganglion cells Superior thickness
 Ganglion cells Temporal Superior thickness
 Ganglion cells Temporal Inferior thickness
 Ganglion cells minimum thickness
 Ganglion cells average thickness
 Macular RNFL Nasal Inferior thickness
 Macular RNFL Nasal Superior thickness
 Macular RNFL Inferior thickness
 Macular RNFL Superior thickness
 Macular RNFL average thickness

MTA analysis

Macular cube average thickness - ILM to RPE

Table 1: Selected macular features from GCA and MTA analysis. Ganglion cells refer to the combined layers GCL and IPL. All the presented features exhibit a high level of significance $p < 0.001$ in differentiating between MS patients and healthy controls.

The feature selection step resulted in a final set of 19 features, which are presented in Tables 1 and 2. Notably, all of the p-values obtained from Student's t-tests of the selected features were below the standard significance threshold ($p < 0.05$), confirming that all of them are effective discriminators between the MS patients and healthy controls. In addition, previous studies [5, 6, 9] have consistently identified most of the selected pa-

rameters as important MS biomarkers, specifically the measurements related to the nasal, superior and inferior sectors of the macula and the temporal, superior and inferior sectors of the optic-disc.

Optic-disc features

ONH analysis

ONH RNFL Inferior Temporal thickness (Clockhour-7)
 ONH RNFL Temporal central thickness (Clockhour-9)
 ONH RNFL Temporal Superior thickness (Clockhour-10)
 ONH RNFL Superior Temporal thickness (Clockhour-11)
 ONH RNFL Temporal Quadrant thickness
 ONH RNFL average thickness

Table 2: Selected optic-disc features from RNFL and ONH analysis. All features have $p < 0.001$ significance between MS patients and healthy controls.

The results corresponding to the performance metrics comparison of the ML models are shown in Table 3. The mean and the standard deviation for each metric are presented and the best metric is highlighted in bold. The Gaussian NB obtained the best AUC with $87.9\% \pm 7.7\%$, and SVM obtained the best accuracy with $87.3\% \pm 10.5\%$ and F1 score with $87.3\% \pm 10.5\%$.

Classifier	Accuracy %	F1 score %	AUC %
SVM	87.3 ± 10.5	87.3 ± 10.5	87.6 ± 10.2
k-NN	83.4 ± 9.4	83.4 ± 9.4	84.3 ± 9.7
Gaussian NB	87.2 ± 8.1	87.2 ± 8.0	87.9 ± 7.7
Random Forest	86.5 ± 10.2	86.5 ± 10.2	86.4 ± 10.4
Bagging Classifier	84.9 ± 5.8	84.8 ± 5.7	84.3 ± 5.4

Table 3: Computed metrics for the ML classifiers after training using 5-fold cross validation. Mean and standard deviation are shown.

Figure 3 illustrates the results of the SHAP analysis applied to the Gaussian NB model. The SHAP values confirm their correspondence with clinically observed variations in the visual pathway in individuals with MS. This

alignment is particularly pronounced in the nasal sector of the macular region and the temporal sector of the optic disc, consistent with previous studies[5, 15].

In addition, our results are consistent with previous research showing progressive degeneration of the GCIPL in MS[15]. This progressive thinning of the GCIPL is underscored by the classifier’s emphasis on five specific features closely associated with the nasal sectors of the GCIPL. Notably, the model places primary importance on the feature indicating minimum GCIPL thickness.

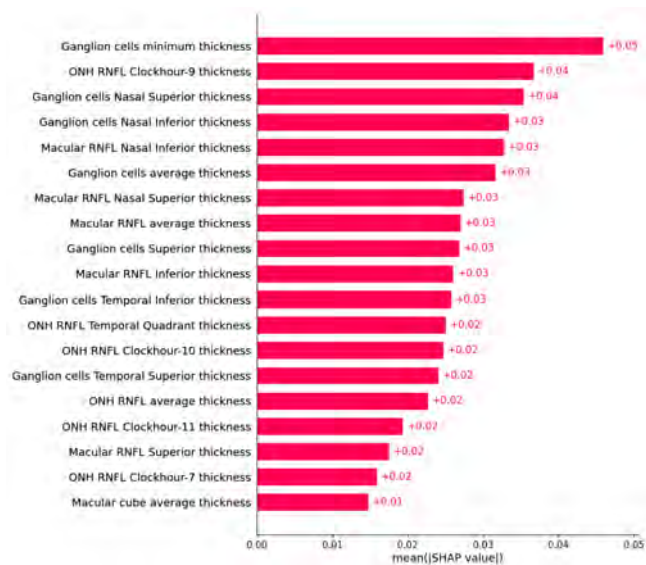


Figure 3: Average SHAP values calculated across the 5-fold validation set for the Gaussian NB model.

5. Conclusions

This study highlights OCT and ML’s potential for enhancing MS diagnosis by evaluating comprehensive features from routine ophthalmologic exams. Our results reveal key features from GCA, MTA, and ONH analyses that significantly contribute to the highest-performing classification model (Gaussian NB) with an AUC of 87.9%. However, limitations include a single-center dataset with one OCT device, potentially affecting result generalizability. Future research should explore longitudinal studies to demonstrate OCT’s value in patient follow-up and its potential integration into standard MS care criteria. Additionally, expanding the dataset to multiple centers and considering alternative modalities, such as OCT angiography images, are promising avenues for future investigation.

Acknowledgments

This research is part of the grant TED2021-131951B-I00 funded by MCIN/AEI/10.13039/501100011033 and by the the “European Union NextGenerationEU/PRTR” and the grant “Primeros Proyectos” funded by the ETSI Telecomunicación (Universidad Politécnica de Madrid).

References

- [1] Daniel S. Reich, Claudia F. Lucchinetti, and Peter A. Calabresi. “Multiple Sclerosis”. In: *The New England Journal of Medicine* 378.2 (Jan. 2018), pp. 169–180. DOI: 10.1056/nejmra1401483.
- [2] Clare Walton et al. “Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition”. In: *Multiple Sclerosis Journal* 26.14 (Nov. 2020), pp. 1816–1821. DOI: 10.1177/1352458520970841.
- [3] Marisa P McGinley, Carolyn H Goldschmidt, and Alexander D Rae-Grant. “Diagnosis and treatment of multiple sclerosis: a review”. In: *Jama* 325.8 (2021), pp. 765–779.
- [4] Neda Minakaran et al. “Optical coherence tomography (OCT) in neuro-ophthalmology”. In: *Eye* 35.1 (Nov. 2020), pp. 17–32. DOI: 10.1038/s41433-020-01288-x.
- [5] Friedemann Paul et al. “Optical coherence tomography in multiple sclerosis: A 3-year prospective multicenter study”. In: *Annals of clinical and translational neurology* 8.12 (Nov. 2021), pp. 2235–2251. DOI: 10.1002/acn3.51473.
- [6] Ricardo N. Alonso, Dolores González-Morón, and Orlando Garcea. “Optical coherence tomography as a biomarker of neurodegeneration in multiple sclerosis: A review”. In: *Multiple sclerosis and related disorders* 22 (May 2018), pp. 77–82. DOI: 10.1016/j.msard.2018.03.007.
- [7] Maria Satue et al. “Optical coherence tomography as a biomarker for diagnosis, progression, and prognosis of neurodegenerative diseases”. In: *Journal of Ophthalmology* 2016 (2016).
- [8] Elena Garcia-Martin et al. “Early diagnosis of multiple sclerosis by OCT analysis using Cohen’s d method and a neural network as classifier”. In: *Computers in Biology and Medicine* 129 (2021), p. 104165.
- [9] Alberto Montolio et al. “Comparison of machine learning methods using Spectralis OCT for diagnosis and disability progression prognosis in multiple sclerosis”. In: *Annals of Biomedical Engineering* 50.5 (Feb. 2022), pp. 507–528. DOI: 10.1007/s10439-022-02930-3.
- [10] Zahra Khodabandeh et al. “Discrimination of Multiple Sclerosis using multicenter OCT images”. In: *Multiple Sclerosis and Related Disorders* (2023), p. 104846.
- [11] Cristina González-Gonzalo et al. “Trustworthy AI: closing the gap between development and integration of AI systems in ophthalmic practice”. In: *Progress in retinal and eye research* 90 (2022), p. 101034.
- [12] F. Pedregosa et al. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.
- [13] Yasen Jiao and Pufeng Du. “Performance measures in evaluating machine learning based bioinformatics predictors for classifications”. In: *Quantitative Biology* 4 (2016), pp. 320–330.
- [14] Scott M Lundberg and Su-In Lee. “A Unified Approach to Interpreting Model Predictions”. In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon et al. Vol. 30. Curran Associates, Inc., 2017.
- [15] Ce Shi et al. “Visual function and disability are associated with focal thickness reduction of the ganglion cell-inner plexiform layer in patients with multiple sclerosis”. In: *Investigative ophthalmology & visual science* 60.4 (2019), pp. 1213–1223.

Enhancing Ulcerative Colitis Histological Image Segmentation with an Active Learning Framework

Fernando García-Torres¹, Giovanni Santacroce², Irene Zammarchi², Pablo Meseguer^{1,3}, Rocío del Amor¹, Brian Hayes⁴, Rory Cotty⁴, Subrata Ghosh², Marietta Iacucci², Valery Naranjo¹

¹ Instituto Universitario de Investigación e Innovación en Tecnología Centrada en el Ser Humano, HUMAN-tech, Universitat Politècnica de València, Valencia, España. fergart1@upv.es.

² APC Microbiome Ireland, College of Medicine and Health, University College of Cork, Cork T12 EC8P, Ireland

³ valgrAI - Valencian Graduate School and Research Network of Artificial Intelligence, Valencia, España

⁴ Department of Histopathology, Cork University Hospital, Cork, Ireland

Abstract

In the management of UC, a chronic inflammatory bowel disease, histological evaluation plays an essential role in targeting histological remission (HR) of intestinal inflammatory activity as the primary aim of treatment. Recent work in the field has proposed a novel index called PHRI that quantifies UC activity based on the presence or absence of neutrophils in different cellular compartments: lamina propria, cryptal epithelium, surface epithelium and cryptal lumen. Digitization of tissue samples into whole-slide images (WSI) has permitted the implementation of computer vision algorithms. These can perform tasks such as semantic segmentation but they require detailed pixel-level annotations, which are laborious and time-consuming in gigapixel images such as WSI. To address this limitation, this work introduces an active learning (AL) algorithm, an innovative approach to alleviating the workload of pathologists by improving the efficiency in identifying crucial regions in these images for UC diagnosis. Leveraging a diverse database of annotated WSI and an iterative training process, the aim is to optimize the performance of the encoder-decoder-based segmentation model.

1. Introduction

Ulcerative colitis (UC) is a chronic inflammatory bowel disease characterized by inflammation and ulcers in the lining of the large intestine and rectum. The motivation for studying UC is rooted in its significant impact on the quality of life of affected individuals [3]. UC has seen an increasing incidence worldwide [5]. It affects individuals of all ages but often manifests in young adulthood. Histological remission, beyond clinical remission, is a popular treatment target in UC. It involves the healing of the mucosal lining of the colon, which can reduce the risk of relapse and complications [6]. PHRI [2] is a simplified and neutrophil-based histological scoring system designed to track histological mucosal healing. It consists of the identification of the neutrophils in four key cellular structures: lamina propria, cryptal epithelium, surface epithelium, and cryptal lumen.

In recent years, biopsy samples have been digitalized into whole slide images (WSI) using specialised digital scanners. WSI allow the remote examination of tissue slides and their automated analysis by computer-aided diagnosis systems. These can incorporate computer vision algorithms to perform tasks such as semantic segmentation, a crucial problem in automatic PHRI scoring due to the nature of the index. However, performing de-

tailed pixel-level annotations of cellular compartments is a laborious and time-consuming process for the pathologists, but crucial for segmentation model development. To address this limitation, this work presents a novel algorithm based on AL, applied for the first time to the segmentation of regions of interest in the context of UC. This algorithm aims to significantly reduce the workload of pathologists by improving the efficiency of identifying these critical regions in histological images. Through an uncertainty-based selection of data samples and an iterative training process, the algorithm allows for optimising the performance of the segmentation model, rooted in encoder-decoder techniques, which in turn results in an objective and more accurate diagnostic process for this disease. This framework strikes a balance between acquiring a diverse database of annotated WSIs while managing the annotation time cost for pathologists. The project relies on an extensive private database comprising WSIs from UC patients, along with pixel-level annotations for selected biopsies.

2. Materials

The database of histological images used in this project was obtained from an international real-life prospective study about UC [4]. A total of 529 biopsies were extracted and followed traditional histological procedures before digitalisation thus creating high-resolution giga-pixel images of tissue samples. Background pixels in WSIs were automatically detected using the Otsu threshold and were not considered for model evaluation and uncertainty estimation.

The pathologists collaborating on the project were tasked with performing pixel-level annotations on only a small subset of 52 WSIs from the entire database. The four regions of interest (lamina propria, cryptal epithelium, surface epithelium, and cryptal lumen) were annotated at the pixel level using in-house software. The aim of requesting these annotations was to have an initial set to train the baseline segmentation model and a permanent set for continuous model evaluation. From the whole database, fifteen WSIs were selected as a WSI database without annotation for further use in the active learning (AL) framework to improve the performance of the segmentation model. The distribution of the data is shown in Table 1. 33 of these images constitute the initial training set with which the initial model is built,

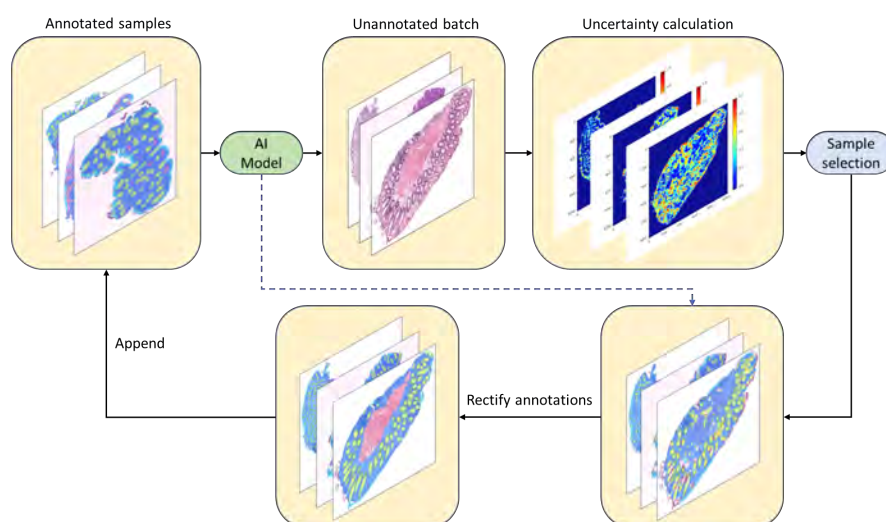


Figure 1. Active Learning and Human-in-the-loop proposed framework.

while the remaining 19 are kept out of the training as a test set for the whole process.

Database	Number of images
<i>Project database</i>	529
Annotated WSI	52
Train	33
Test	19
Unannotated WSI	15

Table 1. Description of the WSI database.

3. Methodology

The proposed AL approach, shown in Figure 1 aims to enhance the performance of the segmentation model by iteratively selecting and annotating samples with the highest uncertainty, thereby improving the model’s ability to accurately identify the four structures of interest. In the first step, a segmentation algorithm is trained and evaluated with the initial set of annotated WSIs. The proposed model consists of a U-Net [7], an encoder-decoder-based architecture, to perform WSI segmentation. Once the initial model has been built, we obtain the prediction of new images from a set of 15 unannotated images which cannot be evaluated in terms of segmentation performance as the GT is not available. Based on these predictions, the uncertainty of each of them will be estimated as a measure of the information they can contribute to the next model. From this set, the most informative samples will be selected, and pathologists will be required to correct the annotations predicted by the model in a human-in-the-loop process. Finally, the segmentation model will be retrained by adding the new actively selected and refined-annotated samples, and its performance will be evaluated with the test set.

3.1. Uncertainty calculation

Uncertainty has been proposed as a measure of the informativeness of the samples [1]. One can assert that

when a prediction carries higher uncertainty, a more significant opportunity exists to enhance our knowledge by incorporating the actual ground truth of that particular sample into the training set. Three different measures for uncertainty estimation have been proposed.

- **Least Confident Sampling (LC):** It is calculated as one minus the maximum predicted probability. One limitation of LC is that it solely considers the information regarding the most likely label while disregarding information about the distribution of the remaining labels.
- **Margin Sampling (M):** The difference between the highest and the second-highest predicted class probability is computed. This sampling method also does not consider the probabilities of all classes.
- **Shanon Entropy (SE):** Entropy serves as a metric to quantify the information needed for encoding a distribution. Entropy’s versatility in extending to probabilistic multiclass annotations surpasses the limitations present in LC and M metrics. Consequently, it is commonly regarded as a means to assess uncertainty within machine learning tasks:

$$x_E = \arg \max_x - \sum_i P(y_i|x) \log P(\hat{y}_i|x) \quad (1)$$

where x is the input image, y_i and \hat{y}_i are the GT and predicted segmentation covering all classes, respectively.

Each of these metrics was calculated at a pixel-level at first to draw uncertainty maps. Subsequently, the overall uncertainty of each WSI was calculated as the average uncertainty of all tissue pixels.

3.2. Human-in-the-loop process

Human-in-the-loop consists of the iterative interaction of humans with the model outputs to guide the models

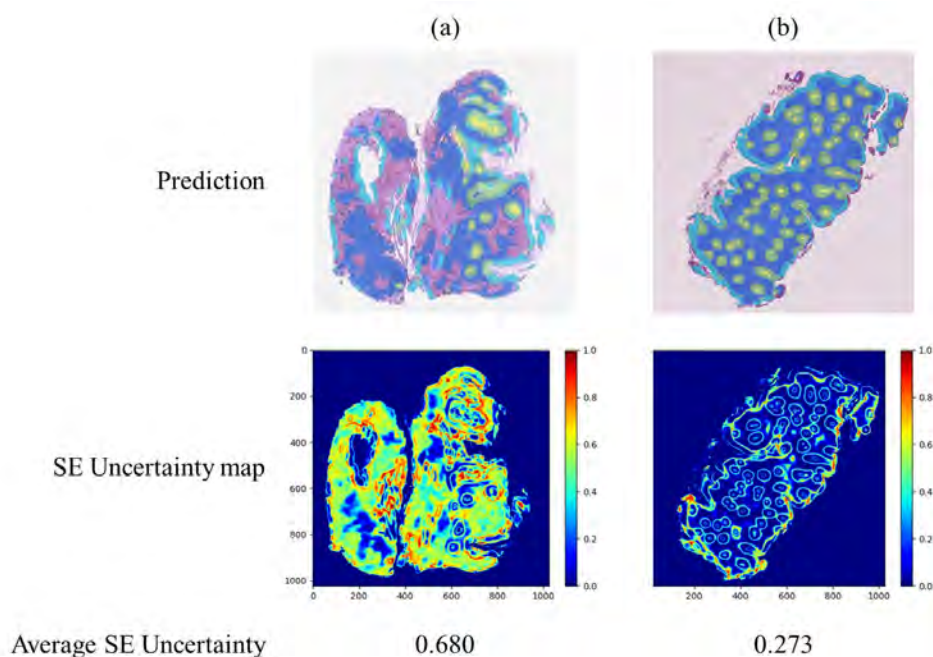


Figure 2. WSI with the highest (a) and the lowest (b) level of uncertainty.

to an optimum. Once the uncertainties of the images were calculated, those with the highest level of uncertainty were selected for correction by the pathologists. Model prediction constitutes the base annotation on which the pathologists had to correct those areas where the model shows more uncertainty. For this purpose, the images with the highest uncertainty and their respective prediction were uploaded to the annotation software to be refined by the pathologist. Then, the corrected segmentation is considered as an annotated WSI and incorporated in the next iteration of model retraining. In this way, the complete segmentations of the images with the highest uncertainty were obtained to provide the model with more information in the subsequent retraining process.

In the AL process, out of the 15 unannotated WSIs we set a specific number of 5 images to comprise each batch to be annotated, starting with those with the highest uncertainty. This decision was driven by the need to find a balance between several crucial factors: the workload for pathologists in annotating samples, the frequency of annotation requests and the time needed for training at each iteration.

4. Results and discussion

To facilitate the comparison of various methods, we choose Dice score and Intersection over Union (IoU) as the performance metrics to evaluate the segmentation model. After training the baseline model in the annotated training set, model performance of $DICE = 0.622$ and $IoU = 0.386$ was obtained in the hold-out test set. Following the proposed framework, we calculate the pixel-level and overall uncertainty for each of the three

measures in the unlabelled set of images. Although all three uncertainty measures provide correlated information about the model prediction, we argue that the most informative metric is the Shannon entropy as it considers the output probabilities for all the classes.

The samples with the highest and the lowest uncertainty are shown in Figure 2 with their associated uncertainty maps according to the SE method, calculated at the pixel level. Looking at the resulting uncertainty maps qualitatively and the prediction of the sample, the model seems to have difficulties distinguishing the lamina propria from other similar tissue regions considered as non-interest like background pixels. As a consequence, the model shows larger uncertainty in non-predicted regions that are similar to lamina propria (see Figure 2a). In contrast, Figure 2b shows the sample with the lowest level of uncertainty according to the SE method. On a qualitative level, it can be observed that most of the pixel uncertainty values are very close or directly 0. A higher level of uncertainty seems to be found at the edges of cryptal epithelium and cryptal lumen structures.

After requesting the pathologist the annotation refinement starting from the model prediction, the model was re-trained first with the five samples with the highest uncertainty to determine the number of samples needed to find a significant improvement in network performance. The model was retrained jointly with the training set and the five new annotated images (AL w/ 5) in the next iteration step. Lastly, we perform the external evaluation with the annotated hold-out test set to extract the test results of this iteration. As shown in

Table 2, the new model trained by adding the five most uncertain samples improves the Dice metric concerning the preliminary model, while the IoU metric decreases slightly.

WSI set	Number of WSIs	Dice	IoU
Training set	33	0.622	0.386
AL w/ 5	33 + 5	0.636	0.384
AL w/ 10	33 + 10	0.651	0.415
AL w/ 15	33 + 15	0.651	0.416

Table 2. Test results for the preliminary model and the models retrained within the AL method.

Based on this result, there was no significant improvement, so the pathologists were asked to correct five more images to add the ten total images to the initial training set and retrain the model. Thus, a new iteration was performed within the framework of AL, evaluating the performance of the retrained model with the test images at the end of the process. As a result, the test results of this iteration outperform the results of the previous models, increasing both Dice and IoU.

Finally, despite the apparent improvement obtained by training with ten additional corrected images, we requested specialists to correct the last five images to check if training with 15 images would result in an even more significant improvement than with 10. This way, we could assess the optimal number of images to correct in each iteration. As can be seen, the last model (AL w/ 15) does not improve the previous model trained (AL w/ 10). Therefore, we argue that the optimal number of annotation corrections required to improve the prediction of regions of interest in biopsies with UC through an AL and human-in-the-loop approach corresponds to 10 WSI.

5. Conclusions

The rising global incidence of UC poses a growing challenge for clinical centres, straining their resources and management. This project addresses the issue by developing an AL-based framework for region segmentation aimed at predicting histologic remission in whole-slide images, a crucial aspect of UC assessment. Pixel-level annotations for selected WSI play a pivotal role in model development, guiding the identification of regions of interest crucial for classifying UC according to the PHRI score. An AL method was developed to optimise the model’s performance in identifying relevant areas. The AL framework involves the calculation of uncertainty maps, pathologist correction, and iterative model refinement. The results indicate that the model’s ability to segment key regions of interest improves significantly with the addition of ten corrected annotations by pathologists to the training set.

The main limitations of the current project include the lack of annotations for non-interest regions, coupled with their similarity to the regions of interest, hinder-

ing accurate segmentation. Efforts to address this emphasize the need for future research to improve neural network differentiation of non-interest and lamina propria regions. As this study represents an initial exploration of iterative techniques for UC segmentation, a larger pool of annotated images is essential for improving performance, making the proposed AL framework crucial. Ultimately, this project successfully achieved its goal of designing an AL-based system for segmenting histologic images before PHRI classification and represents a novel approach in the field of UC diagnosis. This research highlights the promising role of AL in significantly reducing image annotation times and enhancing the precision of ulcerative colitis detection.

Acknowledgements

We gratefully acknowledge the PICaSSO Consortium for the acquisition of images.

Fundings

This work has received funding from the Spanish Ministry of Economy and Competitiveness through projects PID2019-105142RB-C21 (AI4SKIN) and PID2022-140189OB-C21 (ASSIST) and GVA through the project CIPROM/2022/20 (CONTACTS2).

References

- [1] Samuel Budd et al. “A survey on active learning and human-in-the-loop deep learning for medical image analysis”. In: *Medical Image Analysis* 71 (2021), p. 102062 (cit. on p. 2).
- [2] Xianyong Gui et al. “PICaSSO Histologic Remission Index (PHRI) in ulcerative colitis: development of a novel simplified histological score for monitoring mucosal healing and predicting clinical outcomes and its applicability in an artificial intelligence system”. In: *Gut* 71.5 (2022), pp. 889–898 (cit. on p. 1).
- [3] Kathleen A Head et al. “Inflammatory Bowel Disease Part I: Ulcerative Colitis – Pathophysiology and Conventional and Alternative Treatment Options”. In: *Altern Med Rev* 8.3 (2003) (cit. on p. 1).
- [4] Marietta Iacucci et al. “An international multicenter real-life prospective study of electronic chromoendoscopy score PICaSSO in Ulcerative Colitis”. In: *Gastroenterology* 160.5 (2021), pp. 1558–1569 (cit. on p. 1).
- [5] Natalie A Molodecky et al. “Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review”. In: *Gastroenterology* 142.1 (2012), pp. 46–54 (cit. on p. 1).
- [6] Sunhee Park et al. “Histological disease activity as a predictor of clinical relapse among patients with ulcerative colitis: systematic review and meta-analysis”. In: *Official journal of the American College of Gastroenterology—ACG* 111.12 (2016), pp. 1692–1701 (cit. on p. 1).
- [7] Olaf Ronneberger et al. “U-net: Convolutional networks for biomedical image segmentation”. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18*. Springer. 2015, pp. 234–241 (cit. on p. 2).

Discriminación Histológica mediante Espectroscopia Óptica de Onda Parcial de Muestras Tisulares de Cerebro

J.L. Ganoza-Quintana¹, J.L. Arce-Diego¹, F. Fanjul-Vélez¹

¹ Grupo de Ingeniería Biomédica, Departamento TEISA, Universidad de Cantabria, Av. de los Castros 46, 39005 Santander (Cantabria), España

arcedj@unican.es, fanjulf@unican.es

Resumen

Las técnicas de imagen ópticas existentes ofrecen herramientas para visualizar la estructura celular a microescala; sin embargo, su capacidad de sensibilidad a nanoescala está restringida por el límite de difracción. El análisis de la teoría del transporte de luz mesoscópica de los espectros de ondas parciales que se propagan dentro de un medio débilmente desordenado, como las células biológicas [esto es, la espectroscopia de onda parcial (PWS, Partial Wave Spectroscopy)] cuantifica las fluctuaciones del índice de refracción en escalas de longitud subdifraccionales. En este sentido, PWS permite detectar estos cambios nanoestructurales midiendo un parámetro estadístico llamado intensidad de desorden (L_d). PWS utiliza una combinación de un filtro sintonizable y una cámara para adquirir los espectros de retroesparcimiento de cada píxel de la imagen, o de manera más compacta, un espectrómetro en fibra. En este trabajo se estudia y se evalúa la posibilidad de obtener una medida cuantitativa del desorden utilizando el espectro de la información espacial promediada. En lugar de utilizar información espacial y medir rangos espectrales secuencialmente, medimos la señal retroesparcida recogida por una fibra óptica mediante un espectrógrafo. Se aplicaron y evaluaron varios algoritmos de clasificación basados en inteligencia artificial. Se observó que para el tejido de cerebro con magnificaciones de 1X, 4X y 10X con el clasificador de árbol de decisiones se obtienen errores de resubstitución y de correlación cruzada muy bajos. Estos resultados muestran el potencial de proporcionar un diagnóstico significativo automático sin etiquetas de muestras histológicas frescas con el parámetro de desorden obtenido mediante PWS.

1. Introducción

La espectroscopia de esparcimiento elástico se utiliza habitualmente para estudiar la morfología del tejido. Sin embargo, la sensibilidad de una señal de esparcimiento de la luz a las fluctuaciones del índice de refracción se reduce significativamente cuando el tamaño de las estructuras de esparcimiento cae por debajo de la longitud de onda. Recientemente, ha habido un gran interés en comprender los sistemas biológicos a nanoescala, lo que requiere la medición de fluctuaciones del índice de refracción por debajo de la longitud de onda. Según la teoría del transporte de luz mesoscópica, para un objeto ligeramente desordenado y con esparcimiento débil, es posible sondear fluctuaciones del índice de refracción de cualquier escala de longitud, incluidas aquellas muy por debajo de la longitud de onda, si se analiza una señal generada por la interferencia múltiple de ondas que se propagan unidimensionalmente (1D), reflejadas por las fluctuaciones

del índice de refracción dentro del objeto [1]. PWS permite detectar estos cambios nanoestructurales midiendo un parámetro estadístico llamado intensidad de desorden (L_d). Este trabajo está dividido en las siguientes secciones: en la sección 2 se presenta el marco teórico de PWS. En la sección 3 se muestra el método de medición empleado y se detallan los materiales utilizados. En la sección 4 se obtienen los resultados de la investigación. Finalmente, en la sección 5 se discuten las conclusiones.

2. Espectroscopia de onda parcial

Se utiliza la técnica de espectroscopia de onda parcial (PWS, *Partial Wave Spectroscopy*) para medir las fluctuaciones espectrales en los espectros de retroesparcimiento con el fin de obtener el parámetro estadístico denominado intensidad de desorden (L_d) [1,2]. PWS se basa en la idea de que el espectro de retroesparcimiento de un medio débilmente desordenado contiene fluctuaciones espectrales que pueden usarse para cuantificar la intensidad de desorden en función de las fluctuaciones del índice de refracción dentro de la célula, n . En una ubicación dada dentro de la célula, n es proporcional a la densidad local de macromoléculas con un incremento de su índice de refracción, Δn , prácticamente independiente de la composición química. Por tanto, las variaciones espaciales de la densidad macromolecular se pueden medir una vez que se conocen las fluctuaciones de n , suponiendo que estas fluctuaciones son aleatorias y dentro de una distancia de correlación l_c .

El espectro de retroesparcimiento $I(\lambda; x, y)$ (para $\lambda=390-750$ nm, con una resolución espectral ~ 3 nm, muestreo espectral $\sim 0,25$ nm; habitualmente sólo se analiza la señal de $\lambda_1=450$ nm a $\lambda_2=600$ nm debido a la baja iluminación y eficacia de transmisión del instrumento) se registra para cada píxel de celda (x, y) . Todos los espectros obtenidos de cada píxel celular son normalizados por los espectros de la luz incidente utilizando la reflexión especular. Para cada (x, y) , $I(\lambda) (\equiv I(\lambda; x, y))$ se procesa adicionalmente para eliminar el ruido espectral de alta frecuencia utilizando un filtro de Butterworth paso bajo de sexto orden con una frecuencia de corte normalizada de 0,08, y las variaciones en el espectro de la lámpara utilizando un polinomio de orden bajo $I_p(\lambda)$ ajustado a $I(\lambda)$. La frecuencia de corte normalizada (frecuencia de corte/frecuencia de muestreo) se elige de tal manera que elimine todas las oscilaciones por debajo de la resolución espectral del espectrómetro. El

espectro resultante $R(\lambda) = I(\lambda) - I_p(\lambda)$ se conoce como la parte fluctuante del coeficiente de reflexión. Es importante darse cuenta de que, a diferencia de los experimentos tradicionales de esparcimiento de la luz, donde una señal de esparcimiento está formada por todas las ondas que se propagan dentro de una partícula de esparcimiento e interfieren en el campo lejano, el espectro de retroesparcimiento analizado en PWS está formado por un subconjunto de ondas, en particular las ondas parciales 1D. Entonces, $R(\lambda)$ se relaciona con las propiedades del objeto: espesor L , índice de refracción promedio n_0 , o la varianza y la longitud de correlación de las fluctuaciones del índice de refracción ($\langle \Delta n^2 \rangle$ y l_c). En el régimen donde la aproximación de canales independientes 1D es válida, $R(k)$ (donde $k = 2\pi/\lambda$ es el número de onda) se puede caracterizar usando la teoría del transporte de luz mesoscópica 1D. En consecuencia, el promedio rms de $R(k)$ se puede escribir como [1],

$$\langle R \rangle \equiv \frac{L}{\xi} \quad (1)$$

Donde ξ^{-1} es el coeficiente de esparcimiento de un canal 1D,

$$\xi^{-1} \cong \frac{2k^2 L_d}{n_0^2} = \frac{2k^2 \langle \Delta n^2 \rangle l_c}{n_0^2} \quad (2)$$

Mediante la función de autocorrelación, $\langle C(\Delta k) \rangle$

$$\langle C(\Delta k) \rangle = \frac{\langle R(k)R(k+\Delta k) \rangle}{\langle R(k)R(k) \rangle} \quad (3)$$

$\langle C(\Delta k) \rangle$ se calcula para todo el espectro con una longitud de onda central $\lambda_{mean} = 550$ nm para cada píxel de celda. Por tanto, conociendo las cantidades obtenidas experimentalmente $\langle R \rangle$ y $C(\Delta k)$ para un píxel de celda dado (x, y) y n_0 (se supone que es ~ 1.38), se puede calcular L_d ,

$$L_d = -\beta \frac{n_0^2}{2k^2} \langle R \rangle \frac{(\Delta k)^2}{\ln \langle C(\Delta k) \rangle} \quad (4)$$

Donde β es una constante de calibración determinada experimentalmente y $(\Delta k)^2 / \ln \langle C(\Delta k) \rangle$ se obtiene ajustando una pendiente lineal a $\ln \langle C(\Delta k) \rangle$ frente a $(\Delta k)^2$.

3. Materiales y métodos

En primer lugar se mide la señal retroesparcida recogida por una fibra óptica por medio de un espectrómetro. Se evalúa la posibilidad de utilizar una configuración basada en un espectrómetro para calcular el parámetro L_d directamente a partir de la señal de retroesparcimiento recogida por una fibra óptica. Se implementa haciendo uso de un divisor de haz conectado a la salida del microscopio. La fibra óptica se utiliza para dirigir la luz de retroesparcimiento, recogida por un colimador de fibra. El esquema del montaje aparece en la Figura 1.

Para aplicar la teoría PWS, la reflectancia media utilizada para extraer el parámetro L_d se calculó para un solo valor de intensidad media para cada longitud de onda en lugar de calcularlo para cada píxel. Por tanto, se ha denominado a este parámetro de desorden como \bar{L}_d , que es un parámetro de desorden de un solo valor que representa la intensidad

de desorden media de la muestra. Aunque no es posible construir un mapa 2D, se obtiene un parámetro que representa el desorden medio de la muestra. Para extraer este parámetro, se aplicaron las ecuaciones desarrolladas anteriormente. Esta información cualitativa, proporcional al tamaño y densidad de la arquitectura macromolecular, permitió la separación de diferentes muestras de tejido y diferentes áreas de un mismo tejido.

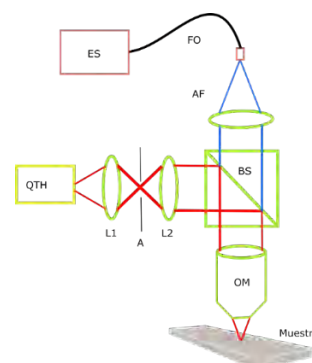


Figura 1. Esquema del sistema PWS: QTH: lámpara de cuarzo-tungsteno-halógeno; L1, L2: lentes de colimación; A: apertura; OM: objetivo de microscopio; BS: divisor de haz; AF: acoplador en fibra óptica; FO: fibra óptica; ES: espectrómetro.

Esta técnica hace uso de una fibra óptica con el fin de guiar la luz retroesparcida de una muestra a un espectrómetro para así analizar todo el espectro y proporciona un valor promediado espacialmente proporcional al campo de visión actual desordenado. Esta configuración hace posible adquirir el espectro de retroesparcimiento en una cámara CCD y en un espectrómetro, a través de una fibra óptica, [2].

Se emplearon muestras frescas de tejidos sanos y tumorales del biobanco de tejidos biológicos del Hospital Universitario Marqués de Valdecilla. Particularmente las muestras biológicas procedían de cerebro. Las muestras fueron congeladas y cortadas por microtomo, con un espesor menor a 6 μm . No se aplicó ningún procedimiento de fijación o tinción a las muestras, ya que estos procesos podrían alterar las propiedades del tejido, en particular las propiedades del índice de refracción. Las muestras se recogieron sobre portaobjetos de vidrio de microscopía tratados para su medición. Las muestras se conservaron a 4°C durante dos horas antes de la medición [3].

En la Figura 2 se observa una imagen digital de un tejido al que posteriormente se aplica la espectroscopia óptica de onda parcial.

4. Resultados

Se han empleado varios algoritmos de clasificación de inteligencia artificial para realizar la discriminación histológica de muestras tisulares de cerebro [4]. Los algoritmos de clasificación empleados fueron: Análisis discriminante lineal (LDA) y cuadrático (QDA), normal Naïve Bayes (NBn), kernel Naïve Bayes (kNB), k vecinos

más próximos (kNN), máquina de vectores de soporte (SVM) y árbol de decisión (DT).

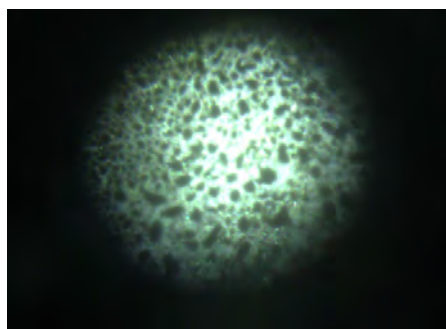


Figura 2. Imagen digital de un tejido al que se aplica la espectroscopia óptica de onda parcial.

En la Tabla 1 se muestran las tasas de errores de resubstitución de diferentes clasificadores de tejido de cerebro sano y tumoral a diferentes magnificaciones.

Algoritmo	Magnificación	Falso Positivo	Falso Negativo
LDA	1X	0.0000	0.0417
	4X	0.3333	0.0000
	10X	0.2083	0.0000
QDA	1X	0.0000	0.0417
	4X	0.3333	0.0000
	10X	0.2083	0.0000
NBn	1X	0.0000	0.0417
	4X	0.3333	0.0000
	10X	0.2083	0.0000
AD	1X	0.0000	0.0417
	4X	0.0833	0.0833
	10X	0.1250	0.0000
NBk	1X	0.0000	0.0417
	4X	0.3333	0.0000
	10X	0.2083	0.0000
KNN	1X	0.0000	0.0417
	4X	0.1667	0.0417
	10X	0.2083	0.0000
SVM	1X	0.3333	0.0000
	4X	0.3333	0.0000
	10X	0.3333	0.0000

Tabla 1. Tasas de errores de resubstitución de diferentes clasificadores de tejido de cerebro sano y tumoral a diferentes magnificaciones.

En la Tabla 2 se muestran las tasas de errores de correlación cruzada de diferentes clasificadores de tejido de cerebro sano y tumoral a diferentes magnificaciones.

Algoritmo	Magnificación	Falso Positivo	Falso Negativo
LDA	1X	0.0000	0.0417
	4X	0.3333	0.0000
	10X	0.2083	0.0000
QDA	1X	0.0000	0.0833
	4X	0.2917	0.0833
	10X	0.1667	0.1250
NBn	1X	0.0417	0.0417
	4X	0.3333	0.0417
	10X	0.2083	0.1667
AD	1X	0.0000	0.0417
	4X	0.3333	0.0000
	10X	0.2083	0.3333
NBk	1X	0.0833	0.0417
	4X	0.3333	0.0000
	10X	0.2083	0.0000
KNN	1X	0.0000	0.0417
	4X	0.2083	0.2083
	10X	0.2500	0.0000
SVM	1X	0.3333	0.0000
	4X	0.3333	0.0000
	10X	0.3333	0.0000

Tabla 2. Tasas de errores de correlación cruzada de diferentes clasificadores de tejido de cerebro sano y tumoral a diferentes magnificaciones.

Se observa que para el error de resubstitución, en la magnificación de 4X, se obtienen las tasas de falsos positivos y falsos negativos más bajas para los algoritmos de clasificación del árbol de decisión y de los k vecinos más próximos. La tasa de falsos positivos y falsos negativos más baja en la magnificación 10X se obtiene para el algoritmo de clasificación del árbol de decisión.

Se observa que para el error de correlación cruzada las tasas más bajas de falsos positivos y falsos negativos se obtienen en la magnificación de 1X.

Se observa que para todos los algoritmos de clasificación excepto SVM para una magnificación de 1X, tienen una tasa de falsos positivos muy pequeña de errores de resubstitución y de correlación cruzada, por lo que tendrán una especificidad alta, por tanto, menos probabilidades de

dar resultados positivos falsos en muestras de tejido de cerebro que no son tumorales.

Se observa que para todos los algoritmos de clasificación para las magnificaciones de 1X, 4X y 10X, tienen unas tasas de falsos negativos muy pequeñas de errores de resubstitución y de correlación cruzada, por lo que tendrán una sensibilidad alta, por tanto, menos probabilidades de perder muestras de tejido de cerebro que son tumorales.

Por tanto, se observa que una relación de alta especificidad y sensibilidad para el parámetro de intensidad de desorden del PWS se tendría para los algoritmos de clasificación de LDA, AD y kNN a una magnificación de 1X, teniendo en cuenta los errores de resubstitución y de correlación cruzada, por tanto, un alto rendimiento de estos algoritmos de clasificación en la discriminación histológica de muestras de tejidos de cerebro sano y tumoral.

5. Conclusiones

Se han obtenido muestras tisulares de tejido sano y tumoral de cerebro de un biobanco. Se ha utilizado la técnica PWS mediante fibra óptica para obtener el parámetro de desorden (L_d) a las magnificaciones de 1X, 4X y 10X. Se aplicaron y evaluaron varios algoritmos de clasificación basados en inteligencia artificial. Se obtuvieron para algunos algoritmos de clasificación errores de resubstitución y de correlación cruzada muy bajos. Se ha observado que para los algoritmos de clasificación LDA, AD y kNN en 1X, se obtiene una relación de alta especificidad y sensibilidad, lo cual aumenta el rendimiento de discriminación histológica de tejido cerebral sano y tumoral. Esto brinda una alta expectativa de extender esta técnica que tiene en cuenta la intensidad de desorden en PWS, a muestras de tejido de otros tipos de órganos y realizar la discriminación histológica de tejido sano y tumoral en esas muestras de tejido de otros tipos de órganos. Estos resultados muestran el potencial de

proporcionar un diagnóstico significativo automático sin etiquetas de muestras histológicas frescas con el parámetro de desorden obtenido mediante PWS.

Agradecimientos

Este Trabajo ha sido parcialmente financiado por el proyecto del Plan Nacional de I+D+i “Cribado diagnóstico de microorganismos mediante microscopia avanzada e inteligencia artificial en patologías humana” (PID2021-127691OB-I00), del Ministerio de Ciencia e Innovación, cofinanciado con fondos FEDER.

Referencias

- [1] Subramanian H, Pradhan P, Liu Y, Capoglu IR, Rogers JD, Roy HK, Brand RE, Backman V. Partial-wave microscopic spectroscopy detects subwavelength refractive index fluctuations: an application to cancer diagnosis. *Opt Lett*. 2009 Feb 15;34(4):518-20. doi: 10.1364/ol.34.000518. PMID: 19373360; PMCID: PMC2701738.
- [2] Fernández R, Marcos-Vidal A, Gallego S, Beléndez A, Desco M, Ripoll J. Qualitative disorder measurements from backscattering spectra through an optical fiber. *Biomed Opt Express*. 2020 Oct 1;11(10):6038-6048. doi: 10.1364/BOE.396013. PMID: 33150004; PMCID: PMC7587252.
- [3] Ganoza-Quintana JL, Fanjul-Vélez F, Arce-Diego JL. Digital Histology by Phase Imaging Specific Biomarkers for Human Tumoral Tissues Discrimination. *Applied Sciences*. 2021; 11(13):6142. <https://doi.org/10.3390/app11136142>.
- [4] Ganoza-Quintana JL, Arce-Diego JL, Fanjul-Vélez F. Digital Histopathological Discrimination of Label-Free Tumoral Tissues by Artificial Intelligence Phase-Imaging Microscopy. *Sensors*. 2022; 22(23):9295. <https://doi.org/10.3390/s22239295>.

Estudio comparativo entre los modelos geométricos de córneas 3D paciente-específico generados a partir de los topógrafos Sirius y MS-39

J. González-Cabrero^{1*}, C. Gómez¹, J. Alió², J. Sornichero³, F. Cavas⁴

¹ Escuela Internacional de Doctorado, Universidad Politécnica de Cartagena, Cartagena, España, jose.gonzalezc@edu.upct.es, gomezgarciacarmelo@gmail.com

² Unidad de Queratocono de la Corporación Vissum Alicante, C/ Cabañal, nº1, Alicante, España, jlalio@vissum.com

³ Servicio de Oftalmología. Hospital General Universitario Santa Lucia, Cartagena, España, jsornichero@gmail.com

⁴ Departamento de Estructuras, Construcción y Expresión Gráfica – Universidad Politécnica de Cartagena, Cartagena, España, francisco.cavas@upct.es

Resumen

La córnea es una estructura transparente, avascular y de geometría semiesférica que se encuentra en la parte frontal del ojo. Esta estructura juega un papel crucial en la refracción y transmisión de luz hacia la retina. Diversas afecciones pueden llegar a alterar su morfología lo que conlleva una disminución de la calidad visual de los pacientes. Los topógrafos corneales son unas tecnologías basadas en técnicas de imagen que permiten evaluar la arquitectura corneal. En función de la técnica de imagen en la que se fundamentan, los topógrafos operan con unos algoritmos internos cuyo funcionamiento detallado a menudo es opaco. En el presente estudio, se modela geoméricamente la misma córnea sana paciente-específico utilizando los datos (nubes de puntos) generados por los topógrafos Sirius y MS-39. El modelo tridimensional generado está basado en las nubes de puntos de las superficies corneales generadas por los dos topógrafos. El objetivo principal es comparar los dos modelos geoméricos de una misma córnea, obtenidos a partir de los datos proporcionados por cada topógrafo para identificar posibles discrepancias en las representaciones geométricas y determinar si dichas variaciones pueden influir en el diagnóstico clínico de enfermedades corneales.

1. Introducción

La córnea es una de las estructuras más vitales del ojo humano. Transparente y avascular, esta parte frontal del ojo tiene la función primordial de refractar y transmitir la luz hacia el interior del ojo, específicamente hacia la retina, facilitando así el proceso de la visión. [1,2] Esta estructura de naturaleza convexa o semiesférica se compone de dos caras distintas: la cara anterior, que es la superficie externa, y la cara posterior o interna, que se encuentra en contacto directo con el humor acuoso, una sustancia líquida que se encuentra en la cámara anterior del ojo, a partir de la cual la córnea obtiene los nutrientes necesarios para mantener su funcionalidad [3]. Hay diversas afecciones que se manifiestan por alteraciones en la forma y geometría de la córnea [4], como es el queratocono, el edema corneal, el astigmatismo, la distrofia corneal, entre otras. Afortunadamente, con el avance de la tecnología, existen equipos topográficos de última generación de imagen que

ofrecen al oftalmólogo un análisis detallado, tanto cualitativo como cuantitativo, de la morfología corneal [5]. A pesar de esta ventaja, un factor limitante es que los sistemas diagnósticos basan sus resultados en unos índices específicos para evaluar la regularidad de la córnea. Estos índices se derivan de algoritmos internos, propios de cada topógrafo, y cuyo funcionamiento detallado suele ser desconocido [5].

En este estudio, se propone modelar virtualmente la córnea sana de un mismo paciente-específico a partir de los datos obtenidos de los topógrafos avanzados: Sirius y MS-39. A partir de estos datos se generarán los modelos corneales virtuales 3D, y sobre estos sólidos se evaluarán varios parámetros geométricos. El estudio propuesto pretende evaluar si existen diferencias entre las mismas medidas obtenidas entre dos tecnologías para un mismo paciente sano. Esta distinción es fundamental, pues según el topógrafo utilizado en la práctica clínica, los diagnósticos podrían variar, sobre todo en los estados iniciales de enfermedad, donde la deformación y, por lo tanto, la variación geométrica puede ser muy pequeña. De este modo, es esencial entender si las diferencias tecnológicas llevan a variaciones significativas en el modelado geométrico y, en consecuencia, si pueden influir en el diagnóstico de las patologías corneales o si, por el contrario, son tan mínimas que pueden ser obviadas.

2. Estado del arte topógrafo corneal. Sirius y MS-39

La historia de los topógrafos corneales es una crónica de avances tecnológicos que se han desarrollado con el objetivo de entender y analizar con mayor precisión la superficie de la córnea. Su evolución ha sido crucial para la oftalmología y, en particular, para la cirugía refractiva.

Antes del desarrollo de la topografía corneal, la única manera de estudiar la superficie del ojo era a través de la queratometría, que sólo proporcionaba medidas de los principales meridianos de la córnea, no una vista completa de toda la superficie [6].

A partir de los años 80, el panorama cambió con la aparición de la topografía de disco de Placido. Usando un patrón de anillos de luz reflejados en la córnea, esta técnica permitió capturar imágenes que mostraban deformidades en los anillos, indicando variaciones en la curvatura corneal. Gracias a ello, los topógrafos de esa época pudieron ofrecer un mapa más completo de la córnea [7]. La llegada de la era digital en los años 90 potenció aún más estas herramientas. La capacidad de digitalizar imágenes y el desarrollo de software especializado permitieron a los topógrafos ofrecer mapas digitales más precisos de la córnea. El software podía interpretar los datos de formas novedosas, ofreciendo distintos tipos de mapas, como los de elevación, potencia y curvatura.

Más adelante, la introducción de la tecnología de Scheimpflug [5,7,8] permitió capturar imágenes de la córnea en un plano focal óptimo, lo que resultó en imágenes más nítidas y precisas. Estos dispositivos utilizan una cámara inclinada que rota alrededor del ojo para obtener múltiples imágenes de la córnea.

Y en los últimos tiempos, se ha visto cómo se combinan tecnologías para ofrecer una visión aún más detallada de la córnea. En este contexto, destacan topógrafos como el Sirius y el MS-39. Este último ha sido desarrollado recientemente, a través de un proyecto de investigación del año 2020, con el objetivo de introducir en el mercado las últimas innovaciones en técnicas de imagen, como la Optical Coherence Tomography (OCT). Tanto Sirius como MS-39 han sido desarrollados por CSO, Italia. El topógrafo Sirius combina la topografía de reflexión por anillos de Plácido con imágenes de Scheimpflug para proporcionar un análisis detallado y completo de la superficie corneal. El MS-39 integra la topografía corneal de disco de Plácido con tomografía de segmento anterior basada en OCT de alta resolución. Otros topógrafos cuyo uso está muy extendido y se basan también en el principio de fotografía Scheimpflug son Pentacam (Oculus, USA) y Galilei (Ziemer, Suiza).

2.1. Topógrafo Sirius

El topógrafo corneal Sirius [9] utiliza la tecnología de imagen Scheimpflug en combinación con el sistema de disco de Plácido. La técnica Scheimpflug consiste en capturar imágenes de la córnea utilizando una cámara rotativa que se coloca a un ángulo específico. Esta combinación permite obtener un análisis detallado de la córnea y otros componentes del ojo. Durante el proceso, las imágenes Scheimpflug capturan la superficie anterior y posterior de la córnea. El sistema de disco de Plácido captura la reflexión de un patrón de anillos en la superficie corneal. Estas imágenes son procesadas y convertidas en un conjunto discreto y finito puntos que representan las superficies corneales, posteriormente se obtienen las superficies corneales utilizando métodos de reconstrucción modal (polinomios de Zernike), la topografía corneal se obtiene comparando la superficie corneal obtenida con una superficie de referencia, normalmente una esfera. La Figura 1. Izquierda, muestra el topógrafo Sirius.

2.2. Topógrafo MS-39

El MS-39 [10] combina la topografía corneal de disco de Plácido con tomografía de segmento anterior basada en OCT de alta resolución, proporcionando imágenes y mapas detallados de la superficie corneal. Durante su funcionamiento, el sistema de disco de Plácido proyecta un patrón de anillos luminosos sobre la córnea y detecta su reflexión. Estos reflejos de los anillos se registran mediante un conjunto finito y discreto de datos, y posteriormente se obtienen las topografías corneales siguiendo el mismo procedimiento que el topógrafo Sirius. La tomografía basada en OCT proporciona imágenes de alta resolución del segmento anterior del ojo, incluyendo la córnea, para un análisis más profundo. La Figura 1. Derecha, muestra el topógrafo MS-39.

2.3. Diferencias entre Sirius y MS-39

Ambos dispositivos emplean la topografía de disco de Plácido para el análisis corneal. Sin embargo, se diferencian en la tecnología complementaria que emplean: el Sirius incorpora la tecnología de imagen Scheimpflug para un análisis detallado de la córnea, mientras que el MS-39 se apoya en la tomografía de segmento anterior basada en OCT de alta resolución. Esta divergencia tecnológica puede generar diferencias en la precisión, resolución y representación tridimensional de la córnea, lo que puede influir en la interpretación y análisis de los datos recabados.



Figura 1. Izquierda: Topógrafo Sirius [11]. Derecha: Topógrafo MS-39 [12].

3. Material y metodología

En las mediciones oftalmológicas utilizando topógrafos como Sirius y MS-39, es común que existan zonas sin datos en las nubes de puntos que representan las superficies anterior y posterior de la córnea. Esta ausencia de datos generalmente se debe a obstrucciones o interferencias durante la medición, los denominados errores intrínsecos y extrínsecos [5]. Por ejemplo, las pestañas del paciente pueden bloquear parte del láser o la luz utilizada para la medición, resultando en áreas sin datos. Además, la película lagrimal, que es la fina capa de lágrimas que cubre la superficie del ojo, puede causar reflejos o difracciones que impiden una lectura clara en ciertas áreas. También es posible que el ángulo de incidencia del rayo de medición en ciertas áreas de la córnea no sea el óptimo para obtener una medición, especialmente en los bordes del ojo. Para abordar este problema, antes de intentar crear un mapeo o representación continua de las superficies anterior y

posterior, es esencial tener un conjunto completo de datos discretos representativos de las superficies corneales. Para solucionar este problema, en nuestro estudio hemos empleado una técnica de interpolación basada en funciones de base radial (RBF) [13].

3.1. Interpolación de los datos de Sirius y MS-39

Antes de iniciar el proceso de interpolación, es fundamental identificar exactamente dónde están los datos faltantes del conjunto de datos, para ello se genera una máscara booleana, esta máscara se encarga de marcar con "Verdadero" los lugares donde faltan datos y con "Falso" donde los datos están presentes. Una vez identificados estos puntos, se procede a la interpolación y estimación de los valores en los puntos donde faltaban datos a través de la RBF [13] de tipo multicuadrática, para las superficies anterior y posterior de la córnea.

Después de la interpolación, los valores estimados se combinan con los valores originales para obtener conjuntos discretos completos de datos, es decir, matrices que no tengan ningún dato faltante. Este proceso se denomina concatenación de datos.

La elección de usar RBF multicuadrática para la interpolación proporciona a los datos interpolados una forma específica y característica, adecuada para una superficie convexa como es la superficie corneal. Las Figuras 2 y 3 muestran las nubes de puntos generadas por los topógrafos Sirius y MS-39, respectivamente. A la izquierda se muestran los datos en bruto que muestran nubes de puntos incompletas y que no caracterizan toda la superficie corneal. A la derecha se muestran las nubes de puntos interpoladas con los datos faltantes rellenos.

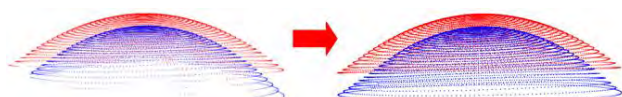


Figura 2. Izquierda: Nube de puntos de Sirius de las superficies anterior (rojo) y posterior (azul) con datos faltantes. Derecha: Nube de puntos de Sirius con datos faltantes interpolados.

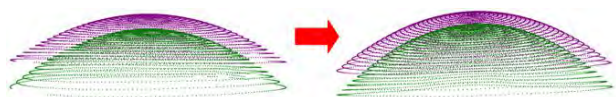


Figura 3. Izquierda: Nube de puntos de MS-39 de las superficies anterior (lila) y posterior (verde) con datos faltantes. Derecha: Nube de puntos de MS-39 con datos faltantes interpolados.

3.2. Generación de volumen corneal Sirius y MS-39

La generación de mallas es fundamental para el análisis y visualización figuras geométricas, y la triangulación de Delaunay [14,15] es una técnica ampliamente reconocida por su eficacia en la representación de superficies. En el caso de las superficies corneales, la naturaleza curvada de estas estructuras plantea desafíos específicos que hacen que la triangulación de Delaunay sea particularmente adecuada.

La triangulación de Delaunay se basa en el principio de convertir un conjunto de puntos, en este caso, la nube de puntos de las superficies corneales, en una malla de

triángulos de tal manera que ningún punto del conjunto original esté dentro del círculo circunscrito de cualquier triángulo en la malla. Este círculo circunscrito es el círculo que puede tocarse exactamente en los tres vértices del triángulo. La elección de este mallado es crucial para la modelización de las superficies corneales. Estas superficies tienen una topología que no es fácilmente representable mediante mallas basadas en triángulos regulares o patrones repetitivos. Delaunay, con su propensión a maximizar ángulos y generar triángulos lo más equiláteros posible, ofrece una malla que se adapta con precisión a las estructuras corneales, respetando y reflejando sus detalles y curvaturas. Una vez que se ha establecido la malla, es posible realizar una serie de operaciones y análisis en ella, como calcular áreas y volúmenes. Las Figuras 4 y 5 muestran el modelo virtual de la córnea completa generadas a partir de los datos de Sirius (Figura 4) y de MS-39 (Figura 5), gracias el mallado de las dos superficies corneales anterior y posterior, y su unión.

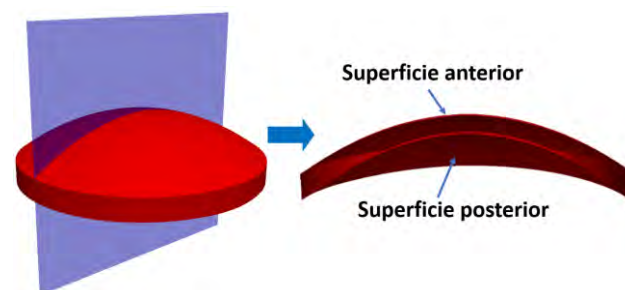


Figura 4. Córnea virtual generada con datos Sirius y vista en corte sagital.

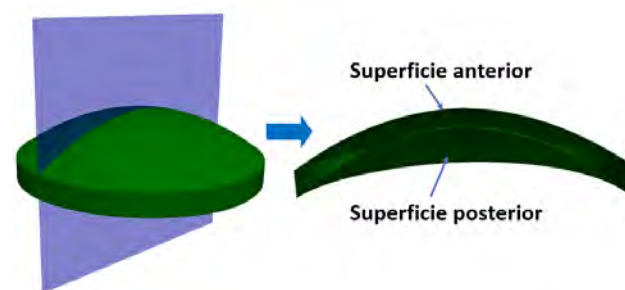


Figura 5. Córnea virtual generada con datos MS-39 y vista en corte sagital.

4. Resultados

Para poder comparar los dos modelos corneales creados a partir de datos generados por Sirius y MS-39, se calculan tres parámetros geométricos: área de la superficie corneal anterior, área de la superficie corneal posterior y volumen corneal, y por último obtenemos el porcentaje de error relativo (%error) entre los dos modelos. En nuestro estudio hemos usado la librería de Python PyVista. Empleando la triangulación Delaunay, el área total se calcula sumando las áreas de todos estos triángulos individuales a través del atributo `.area`. Para el cálculo del volumen se empleó la librería Trimesh, el volumen de un objeto 3D se determinó a partir de sus vértices y caras triangulares. Se descompuso el objeto en tetraedros formados por sus triángulos superficiales y un punto de referencia. Posteriormente se

calculó el volumen de cada tetraedro y se sumaron todos juntos para obtener el volumen total del objeto. Con esta técnica matemática se aprovecha el hecho de que el volumen de un tetraedro es computable si se conocen sus vértices. Los cálculos son realizados directamente por la librería Trimesh al usar el atributo *.volume*. Los resultados geométricos de una córnea sana, medida por los dos dispositivos, se presentan en la Tabla 1.

Parámetro geométrico	Sirius	MS-39	%Error
Área anterior (mm ²)	87.13	87.29	0.18
Área posterior (mm ²)	90.40	90.64	0.27
Volumen (mm ³)	56.22	55.60	1.12

Tabla 1. Resultados geométricos de los modelos corneales Sirius y MS-39 y correlación de resultados

Los parámetros geométricos de las córneas virtuales, basados en los datos del Sirius y del MS-39 para un mismo paciente sano y obtenidos en las mismas condiciones de entorno, muestran una coherencia notable entre ambos conjuntos de datos. Los resultados fueron obtenidos para un radio máximo de 5 mm.

Las pequeñas diferencias observadas pueden ser debido al ruido inherente a los propios dispositivos topográficos. Estos dispositivos están diseñados para ser precisos, pero no están exentos de ruido que puede introducirse durante la medición. La tecnología subyacente en Sirius y MS-39, al ser diferente, puede llevar a diferencias sutiles en los resultados obtenidos.

5. Conclusiones

Cualquier alteración en la forma y geometría de la córnea puede indicar ciertas afecciones como el queratocono o el edema corneal, lo que resalta la importancia de examinar con precisión la morfología corneal. Para ello, el topógrafo corneal se utiliza como dispositivo esencial para caracterizar la forma real de la córnea del paciente. En este estudio se ha realizado una comparación de parámetros geométricos (área y volumen) entre dos modelos corneales virtuales creados a partir de los datos generados por los topógrafos Sirius y MS-39 de una córnea sana paciente-específico. Esta comparación ha demostrado que, a pesar de unas discrepancias mínimas en los resultados, estas pueden ser obviadas debido a la pequeña variación. Adicionalmente, este estudio puede abrir un camino interesante en la homogeneización del modelado corneal, independientemente del topógrafo empleado para la adquisición de datos.

Financiación

Esta publicación se realizó en el marco del proyecto “Desarrollo y validación de un nuevo concepto de caracterización biomecánica-morfofuncional de la córnea” número de referencia DTS21/00103. Este proyecto ha

contado con el apoyo del Instituto de Salud Carlos III y cofinanciado por la Unión Europea.

Referencias

- [1] Delmonte, D.; Kim, T. Anatomy and physiology of the cornea. *Journal of cataract and refractive surgery* **2011**, *37*, 588-598. DOI: 10.1016/j.jcrs.2010.12.037.
- [2] Mansoor, H.; Ong, H. S.; Riau, A. K.; Stanzel, T. P.; Mehta, J. S.; Yam, G. H. Current Trends and Future Perspective of Mesenchymal Stem Cells and Exosomes in Corneal Diseases. In *International Journal of Molecular Sciences*, 2019; Vol. 20.
- [3] McLaren, J. W. Control of Aqueous Humor Flow. In *Encyclopedia of the Eye*, Dartt, D. A. Ed.; Academic Press, 2010; pp 389-397.
- [4] Hess, R. F.; Carney, L. G. Vision through an abnormal cornea: a pilot study of the relationship between visual loss from corneal distortion, corneal edema, keratoconus, and some allied corneal pathology. *Investigative Ophthalmology & Visual Science* **1979**, *18* (5), 476-483. (accessed 9/10/2023).
- [5] Cavas-Martínez, F.; De la Cruz Sánchez, E.; Nieto Martínez, J.; Fernández Cañavate, F. J.; Fernández-Pacheco, D. G. Corneal topography in keratoconus: state of the art. *Eye and Vision* **2016**, *3* (1), 5. DOI: 10.1186/s40662-016-0036-8.
- [6] Gutmark, R.; Guyton, D. L. Origins of the Keratometer and its Evolving Role in Ophthalmology. *Survey of Ophthalmology* **2010**, *55* (5), 481-497. DOI: <https://doi.org/10.1016/j.survophthal.2010.03.001>.
- [7] Martin R. Cornea and anterior eye assessment with placido-disc keratometry, slit scanning evaluation topography and scheimpflug imaging tomography. *Indian J Ophthalmol.* 2018; 66(3):360-366. doi: 10.4103/ijo.IJO_850_17
- [8] Fan R, Chan TC, Prakash G, Jhanji V. Applications of corneal topography and tomography: a review. *Clin Exp Ophthalmol.* 2018;46(2):133-146. doi:10.1111/ceo.13136
- [9] Sahu, J. Corneal Topography : Sirius. *Delhi Journal Of Ophthalmology* **2022**, *32*, 119-124. DOI: 10.7869/djo.731.
- [10] Mohammadpour, M.; Heidari, Z. MS-39®. 2021; pp 265-284.
- [11] Página web de la empresa Costruzione Strumenti Oftalmici (CSO Italia): <https://www.csoitalia.it/it/pagina/10-rd> (Consultada: Septiembre 2023).
- [12] Página web de la empresa Costruzione Strumenti Oftalmici (CSOItalia): https://www.csoitalia.it/app/public/files/prodott_o/Sirius_ENG_bianco_v1.pdf (Consultada: Septiembre 2023).
- [13] Combe, R.; Arrocha, I. Métodos de Interpolación Basados en Funciones de Base Radial con Aplicaciones a la Reconstrucción de Imágenes. *KnE Engineering* **2018**, *3*, 563. DOI: 10.18502/keg.v3i1.1460.
- [14] Dey, Tamal K.. “Delaunay Mesh Generation of Three Dimensional Domains *.” (2007).
- [15] Ortiz, S.; Siedlecki, D.; Remon, L.; Marcos, S. Three-dimensional ray tracing on Delaunay-based reconstructed surfaces. *Applied optics* **2009**, *48*, 3886-3893. DOI: 10.1364/AO.48.003886.

Aprendizaje auto-supervisado para mejorar el diagnóstico del síndrome de Sjögren con ultrasonografía

F. Herrero Quevedo¹, L. Cubero¹, O. Olivas-Vergara², E. Naredo Sánchez², J. Pascau^{1,3}

¹ Departamento de Bioingeniería, Universidad Carlos III, Madrid, España

² Servicio de Reumatología, Hospital Universitario Fundación Jiménez Díaz, Madrid, España

³ Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España

Resumen

La distinción precisa entre glándulas salivales saludables y aquellas afectadas por el síndrome de Sjögren (SS) en imágenes por ultrasonido es compleja. Esto resulta en un subdiagnóstico y diagnóstico erróneo significativo del SS, afectando severamente la calidad de vida de los pacientes y conduciendo a graves complicaciones clínicas. Los modelos de aprendizaje profundo han mostrado resultados prometedores en el diagnóstico de SS, pero se ven limitados por la gran cantidad de datos anotados manualmente que requieren. Este artículo investiga enfoques alternativos para entrenar dichos modelos con cantidades limitadas de datos anotados, centrándose específicamente en el aprendizaje por transferencia y el aprendizaje auto-supervisado. El estudio explora la efectividad del modelo Bootstrap Your Own Latent y demuestra su superioridad sobre los métodos de aprendizaje por transferencia en una base de datos multicéntrica de 305 imágenes de ultrasonido. Además, esta investigación proporciona valiosas perspectivas sobre la selección de modelos y estrategias de entrenamiento de modelos de aprendizaje profundo para mejorar la clasificación de glándulas salivales de acuerdo con la escala consensual Outcome Measures in Rheumatology.

1. Introducción

El síndrome de Sjögren (SS) es una enfermedad autoinmune, crónica y sistémica caracterizada por la inflamación de las glándulas exocrinas, especialmente las glándulas salivales y lagrimales [1]. A pesar de la naturaleza no específica y leve de sus síntomas, el SS puede evolucionar hacia complicaciones graves, incluyendo el desarrollo de linfomas en hasta el 5% de los pacientes [2]. Estos riesgos, junto a la ausencia de un conjunto único de signos y síntomas, resaltan la importancia del diagnóstico temprano para el tratamiento del SS [3].

Actualmente, el diagnóstico del SS se basa en el cumplimiento de criterios de clasificación consensuales basados en pruebas histológicas, inmunológicas y de función glandular [4]. Sin embargo, estas solo están disponibles en entornos clínicos especializados [5]. Como resultado, el SS se diagnostica con frecuencia de manera insuficiente o errónea [6]. Debido a esto, encontrar una herramienta de diagnóstico fiable y específica para el SS, cuya aplicabilidad pueda extenderse fuera de los niveles de atención especializada, es de suma importancia.

En las últimas dos décadas, se ha demostrado que la ultrasonografía de las glándulas salivales (SGUS) es una técnica precisa y fiable para la evaluación de pacientes con sospecha de SS [7]. Esto ha llevado a la propuesta de incluir la SGUS en los criterios de clasificación del SS [8]. En este contexto, un panel de expertos en ultrasonografía y SS del grupo *Outcome Measures in Rheumatology*

(OMERACT) ha desarrollado una escala semicuantitativa de cuatro clases (0-3) con alta fiabilidad inter- e intra-observador. Esta escala categoriza el compromiso glandular en el SS en función del aspecto morfológico y estructural de las principales glándulas salivales: glándulas parótidas (PG) y glándulas submandibulares (SMG) [9].

Aunque la SGUS tiene el potencial de convertirse en una práctica aceptada en el diagnóstico del SS, la interpretación y evaluación de imágenes de ultrasonido requieren una curva de aprendizaje pronunciada [10]. Por ello, ha habido un creciente interés en aplicar técnicas de aprendizaje profundo para esta tarea. Sin embargo, la mayoría de los enfoques exitosos se basan en la anotación manual de expertos, que a menudo es laboriosa y dificulta la capacidad de entrenar modelos de manera robusta.

Tradicionalmente, el aprendizaje por transferencia (TL) se basa en utilizar datos escasamente etiquetados y optimizar los entrenamientos mediante la adopción de las capas inferiores de modelos entrenados en conjuntos de datos con millones de anotaciones, como ImageNet. Estas capas inferiores adquieren características genéricas aplicables a diversas tareas y conjuntos de datos. Sin embargo, asumir que las características aprendidas en el dominio original son representativas del dominio objetivo no siempre es correcto. Para superar esta limitación, se han desarrollado técnicas para entrenar estas capas inferiores con datos sin anotar más específicos del dominio objetivo [11].

En este contexto, se ha propuesto el aprendizaje auto-supervisado (SSL). A diferencia de TL, SSL se centra en entrenar las capas de nivel inferior de la red utilizando datos sin etiquetar a través de tareas auto-supervisadas como el relleno de imágenes, la eliminación de ruido o la predicción de rotaciones [11]. Cuando el modelo se entrena con datos del mismo dominio que la tarea objetivo, las características de nivel inferior aprendidas representan mejor dicho dominio. Entre los métodos populares de SSL, destaca *Bootstrap Your Own Latent* (BYOL), basado en métodos de SSL de aprendizaje de contraste [12].

El objetivo de este estudio es entrenar y optimizar un modelo de aprendizaje profundo para clasificar imágenes de SGUS de las glándulas parótidas y submandibulares según su grado de SS correspondiente con el sistema de puntuación OMERACT. Para ello, identificamos el modelo más preciso en la clasificación en cuanto a arquitectura e hiperparámetros, pre-entrenado con TL. Después, evaluamos y comparamos estos resultados contra el rendimiento del mismo modelo pero pre-entrenado con BYOL.

2. Métodos

Con el objetivo de entrenar y optimizar un modelo de aprendizaje profundo en la clasificación del SS en US, se utilizó un conjunto de datos de imágenes de ultrasonido de glándulas salivales que combinó muestras de PG y SMG de un repositorio de acceso público [10] y de pacientes del Hospital Universitario Fundación Jiménez Díaz. El conjunto de datos consta de 305 imágenes, cada una puntuada según la escala OMERACT, y está descompensado hacia la clase 0, con 113 muestras, en comparación con las 52, 67 y 73 imágenes anotadas como clases 1, 2 y 3, respectivamente. Se utilizaron 245 imágenes para el entrenamiento, mientras que los 60 casos restantes se dividieron por igual entre los conjuntos de validación y prueba (30 imágenes cada uno). La tarea de clasificación se centró en asignar con precisión cada imagen a una de las cuatro clases definidas por la escala OMERACT.

El proceso se dividió en tres etapas: (i) seleccionar la arquitectura óptima, (ii) evaluar varias alternativas de aumento de datos y (iii) evaluar la viabilidad del aprendizaje auto-supervisado (SSL) como alternativa al aprendizaje por transferencia (TL) para el pre-entrenamiento del modelo.

2.1. Selección de la arquitectura óptima

La primera decisión fue la selección de la arquitectura de modelo óptima siguiendo un pre-entrenamiento de TL. En esta etapa, se evaluaron múltiples arquitecturas utilizando pesos pre-entrenados proporcionados por la biblioteca torchvision. Entre todas las arquitecturas, se seleccionó finalmente un subconjunto de 11 comúnmente utilizados para la clasificación de imágenes médicas y con un número de parámetros que no supusiera restricciones significativas de memoria en la GPU (Fig. 1).

El proceso de evaluación para cada modelo candidato siguió el siguiente flujo de trabajo. En primer lugar, se transfirieron los pesos pre-entrenados de ImageNet [13] a cada modelo. A continuación, se reemplazó la cabeza de clasificación del modelo con una nueva cabeza con pesos aleatorios. El modelo se afinó (*fine-tuning*) en la tarea objetivo durante un total de 100, utilizando el optimizador de descenso de gradiente estocástico (SGD) con inercia, y se empleó un programador de tasa de aprendizaje. A su vez, se evaluaron varios hiperparámetros para optimizar el rendimiento. Finalmente, se evaluó el rendimiento de cada modelo en el conjunto de validación y se seleccionaron el modelo de mejor rendimiento y el modelo de mejor rendimiento con menos de 10 millones de parámetros, con el fin de explorar la aplicación de esta solución en un régimen de tiempo real.

2.2. Evaluación de alternativas de aumento de datos

Tras seleccionar los modelos con mejor rendimiento, se evaluó el impacto de diferentes estrategias de aumento de datos durante la fase de entrenamiento. Se siguió un enfoque de aumento de datos sobre la marcha, en el que las transformaciones se aplicaron a las imágenes a medida que se cargaban en la memoria. Cada transformación era equiprobable y su probabilidad se ajustó de modo que solo el 5% de las muestras permanecieran sin modificar. Dado que

la literatura disponible sobre técnicas de aumento de datos para imágenes de SGUS es limitada, se adaptaron cinco *pipelines* del campo de detección de melanoma.

También se evaluó un enfoque de aumento de datos en tiempo de prueba (TTA) con el objetivo de mejorar la robustez de las predicciones del modelo durante las etapas de validación y prueba. Específicamente, para cada muestra, se realizó un recorte de imagen en el centro y en las cuatro esquinas, resultando en cinco versiones de la imagen, y por tanto cinco predicciones. Para cada muestra, se agregaron las cinco predicciones mediante promedio o *max-pooling* para obtener la predicción final.

2.3. Evaluación del SSL como alternativa al TL

Finalmente, se evaluó la posibilidad de reemplazar la estrategia de TL con el método de SSL denominado BYOL [12] para aprender características específicas del dominio de US a partir de datos no etiquetados. Este método emplea dos redes casi simétricas, conocidas como red *en línea* y red *objetivo*. La única diferencia es la presencia de una capa de predicción final en la red *en línea*.

El objetivo de BYOL es maximizar la similitud entre las representaciones devueltas por ambas redes de manera iterativa. Para ello, se generan dos vistas diferentes a partir de la misma imagen de entrada aplicando distintas transformaciones. Ambas redes obtienen una representación de menor dimensión al propagar ambas vistas a través de las capas compartidas. Finalmente, la representación obtenida red *en línea* se propaga por la capa de predicción de la red *en línea*, que intenta acercarla a la representación de la red *objetivo*. El error entre estas dos se propaga para ajustar los pesos de la red *en línea*, que luego son copiados a la red *objetivo*. Este proceso se repite para cada iteración. Al converger, el codificador de la red *en línea* se conserva y se utiliza para tareas posteriores, i.e. tareas de clasificación.

Análogo al caso de TL, las capas inferiores de este modelo *en línea* pre-entrenado elegido se volvieron a entrenar y la cabeza de clasificación se reemplazó por una con pesos aleatorios. Finalmente, esta última se afinó a la tarea de clasificación (sección 2.1).

3. Resultados y Discusión

Se evaluó el rendimiento de cada modelo en la tarea de clasificación calculando el área bajo la curva (AUC) en una modalidad de uno contra todos, para aliviar el desequilibrio de clases. Este enfoque permitió seleccionar los modelos y estrategias de aumento de datos en función de su rendimiento en el conjunto de validación durante las últimas 10 épocas.

3.1. Evaluación de diferentes arquitecturas

Se evaluó y comparó la evolución del rendimiento de cada una de las arquitecturas pre-entrenadas en ImageNet a través del *fine-tuning* de varios hiperparámetros. La Figura 1 ilustra la progresión del AUC a través de diversos valores de varios hiperparámetros. A partir de este gráfico, podemos extraer varias conclusiones principales.

En primer lugar, en cuanto a la tasa de aprendizaje, se observa una clara tendencia que indica que valores más altos

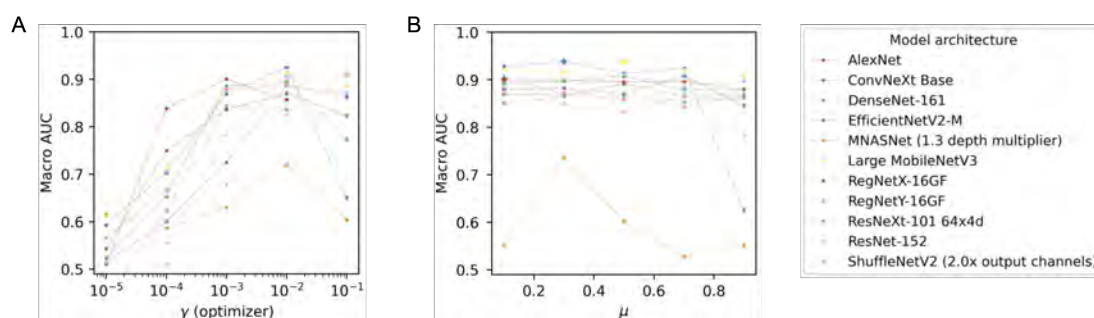


Figura 1. Progresión del AUC para múltiples valores de hiperparámetros para cada uno de los modelos evaluados. **A.** Tasa de aprendizaje. **B.** Inercia.

de la misma son más adecuados para el problema y el conjunto de datos (Fig. 1A). En particular, con una tasa de aprendizaje de 10^{-2} , todos los modelos con la excepción de MNASNet logran un AUC macro mayor de 0.8. Estos resultados sugieren la presencia de un componente inherente en el aprendizaje que es independiente de la arquitectura específica del modelo.

En cuanto a la inercia, los modelos generalmente mantienen un rendimiento relativamente alto en diferentes valores, lo que indica su robustez con respecto a este hiperparámetro (Fig. 1B). Notablemente, durante esta etapa, los modelos ConvNeXt y MobileNet logran el mejor rendimiento general, lo que sugiere su ventaja en configuraciones de optimización más simples.

La Figura 2 describe la relación entre la complejidad y el rendimiento de cada modelo. Su análisis indica que, con la excepción del modelo MNASNet, el número de parámetros no parece tener un impacto significativo en el rendimiento. Modelos como MobileNet obtienen mejores resultados que algunas arquitecturas más complejas en el mejor escenario (Fig. 2A). Esto sugiere que, para esta tarea de clasificación, el rendimiento del modelo también depende de otros factores como el diseño de la arquitectura y la optimización.

Sin embargo, que los modelos más complejos muestran una mayor robustez a lo largo del proceso de validación de hiperparámetros. En términos de rendimiento mediano, los modelos más pequeños tienden a clasificarse más bajo, mientras que los modelos más grandes mantienen sus posiciones relativas en la clasificación (Fig. 2B). Esto destaca la robustez.

Finalmente, todas estas consideraciones nos llevaron a elegir a la MobileNet y a la ConvNeXt como las redes óptimas para la tarea de clasificación de las SGUS para el diagnóstico del SS.

3.2. Evaluación de estrategias de aumento de datos

Para evaluar el rendimiento de diferentes estrategias de aumento de datos durante el entrenamiento, comparamos el rendimiento de los dos mejores modelos seleccionados en la etapa anterior, en los diferentes escenarios de aumento de datos.

Los resultados se resumen en la Figura 3. Mientras que para el modelo ConvNeXt el aumento de datos no mejora el rendimiento base, es observable que con MobileNet to-

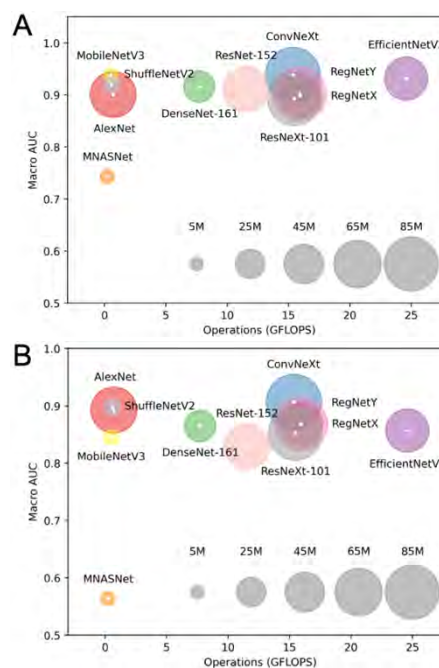


Figura 2. Rendimiento **A.** máximo y **B.** mediano frente a número de operaciones por segundo de cada modelo, a lo largo de la etapa de validación de hiperparámetros. El tamaño de los puntos indica el número de hiperparámetros.

dos los pipelines de aumento de datos conducen a una mejora en él. Es importante destacar que la estrategia 3 logra el mejor rendimiento, superando incluso el mejor caso para el modelo ConvNeXt. Dicho pipeline, de una complejidad moderada, incluía transformaciones afines y elásticas, así como fluctuaciones aleatorias en el brillo y contraste de la imagen.

Con respecto a TTA, observamos una mejora sobre el rendimiento base en el caso de MobileNet. Esto indica que el modelo se beneficia al ver diferentes vistas de la misma imagen durante la inferencia. En concreto, la opción de *max-pooling* supera a hacer el promedio de todas las predicciones. Esto sugiere que, en algunos casos, la información significativa se localiza en regiones específicas de la imagen en lugar de estar dispersa por toda la vista.

3.3. Evaluación del SSL como alternativa al TL

En cuanto al uso de SSL como alternativa al pre-entrenamiento con TL, los resultados son alentadores. La Figura 4 ilustra el mejor rendimiento alcanzado por los modelos con pesos pre-entrenados en ImageNet y BYOL para diferentes proporciones de datos anotados disponibles para el fine-

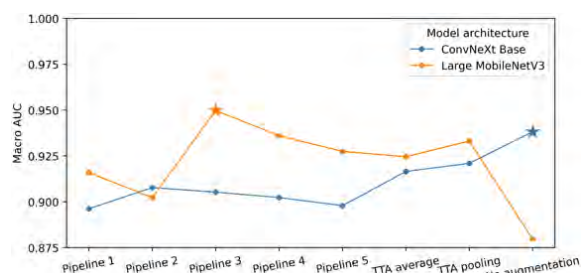


Figura 3. Rendimiento de los dos mejores modelos seleccionados antes y después de ser entrenados con cada uno de los pipelines de aumento de datos implementados.

tuning. Los resultados demuestran que, cuando todos los datos están disponibles, el modelo pre-entrenado con BYOL demuestra una ligera ventaja tanto para ConvNeXt como para MobileNet con respecto a los modelos pre-entrenados con TL. Este hallazgo parece respaldar el argumento de que las características aprendidas durante el pre-entrenamiento con BYOL son más representativas del dominio de imágenes de ultrasonido. La superioridad de BYOL se confirma con el crecimiento significativo en la diferencia de rendimiento observado entre los modelos pre-entrenados utilizando ambas estrategias a medida que disminuye la disponibilidad de datos anotados.

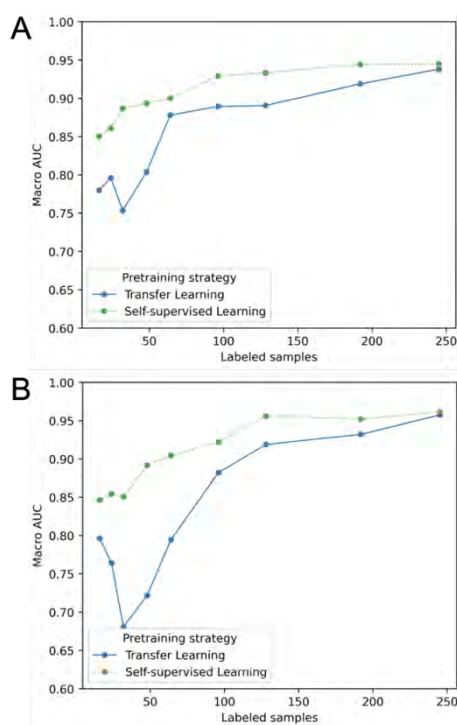


Figura 4. Rendimiento de los modelos seleccionados para diferentes proporciones de datos anotados en métodos pre-entrenados con TL y SSL. **A.** ConvNeXt y **B.** MobileNet.

4. Conclusiones y Líneas Futuras

En este estudio, se ha desarrollado un marco exitoso para clasificar imágenes de ultrasonido de las glándulas salivales principales con metodologías de aprendizaje profundo, de acuerdo con la escala OMERACT, para el diagnóstico de SS. Al mismo tiempo, se han investigado y optimizado diferentes estrategias de entrenamiento de los modelos. En primer lugar, hemos identificado la arquitectura MobileNet

como óptima. En segundo lugar, hemos demostrado la superioridad de los métodos de SSL, en particular de BYOL, sobre un enfoque TL estándar como estrategia de pre-entrenamiento, incluso con datos anotados limitados.

Una posible línea de investigación futura podría involucrar la evaluación y comparación de métodos alternativos de SSL para esta tarea. Además, la validación de los resultados en una cohorte más grande y variada de pacientes con SS, adquirida con múltiples ecógrafos, y la comparación del rendimiento entre médicos experimentados y aquellos asistidos por esta herramienta, permitiría un análisis más profundo de la precisión de la metodología propuesta.

Además, trabajos futuros deberían confirmar la superioridad del uso de la escala OMERACT y SSL para el diagnóstico de SS a partir de imágenes. En la actualidad, esto se ve mermado debido a la escasez de *datasets* anotados de manera experta en esta escala.

Agradecimientos

Estudio financiado por PI122/00601 y AC20/00102 (Ministerio de Ciencia, Innovación y Universidades, Instituto de Salud Carlos III, AECC y European Regional Development Fund “Una manera de hacer Europa”), PerPlanRT (ERA PerMed), TED2021-129392B-I00 y TED2021-132200B-I00 (MCIN/AEI/10.13039/501100011033 y Unión Europea “NextGenerationEU”/PRTR).

Referencias

- [1] R. I. Fox, «Sjögren’s syndrome», *The Lancet*, vol. 366, n.º 9482, pp. 321-331, jul. 2005, doi: 10.1016/S0140-6736(05)66990-5.
- [2] G. Nocturne y X. Mariette, «Sjögren Syndrome-associated lymphomas: an update on pathogenesis and management», *Br J Haematol*, vol. 168, n.º 3, pp. 317-327, feb. 2015.
- [3] M. Ramos-Casals *et al.*, «EULAR recommendations for the management of Sjögren’s syndrome with topical and systemic therapies», *Ann Rheum Dis*, vol. 79, n.º 1, pp. 3-18, ene. 2020.
- [4] C. H. Shiboski *et al.*, «2016 American College of Rheumatology/European League Against Rheumatism Classification Criteria for Primary Sjögren’s Syndrome: A Consensus and Data-Driven Methodology Involving Three International Patient Cohorts», *Arthritis & Rheumatology*, vol. 69, n.º 1, pp. 35-45, ene. 2017.
- [5] D. Parisi, C. Chivasso, J. Perret, M. S. Soyfoo, y C. Delporte, «Current State of Knowledge on Primary Sjögren’s Syndrome, an Autoimmune Exocrinopathy», *JCM*, vol. 9, n.º 7, p. 2299, jul. 2020.
- [6] S. S. Kassan y H. M. Moutsopoulos, «Clinical Manifestations and Early Diagnosis of Sjögren Syndrome», *Arch Intern Med*, vol. 164, n.º 12, p. 1275, jun. 2004.
- [7] A. Hočvar, A. Ambrožič, B. Rozman, T. Kveder, y M. Tomšič, «Ultrasonographic changes of major salivary glands in primary Sjögren’s syndrome. Diagnostic value of a novel scoring system», *Rheumatology*, vol. 44, n.º 6, pp. 768-772, jun. 2005.
- [8] S. Jousse-Joulin *et al.*, «Weight of salivary gland ultrasonography compared to other items of the 2016 ACR/EULAR classification criteria for Primary Sjögren’s syndrome», *J Intern Med*, vol. 287, n.º 2, pp. 180-188, feb. 2020.
- [9] S. Jousse-Joulin *et al.*, «Video clip assessment of a salivary gland ultrasound scoring system in Sjögren’s syndrome using consensual definitions: an OMERACT ultrasound working group reliability exercise», *Ann Rheum Dis*, vol. 78, n.º 7, pp. 967-973, jul. 2019.
- [10] A. Zabotti *et al.*, «Salivary Gland Ultrasonography in Sjögren’s Syndrome: A European Multicenter Reliability Exercise for the Harmonic-SS Project», *Front. Med.*, vol. 7, p. 581248, nov. 2020.
- [11] D. Anand, D. Tank, H. Tibrewal, y A. Sethi, «Self-Supervision vs. Transfer Learning: Robust Biomedical Image Analysis Against Adversarial Attacks», en *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, Iowa City, IA, USA: IEEE, abr. 2020, pp. 1159-1163.
- [12] J.-B. Grill *et al.*, «Bootstrap your own latent: A new approach to self-supervised Learning». arXiv, 10 de septiembre de 2020.
- [13] A. Krizhevsky, I. Sutskever, y G. E. Hinton, «ImageNet classification with deep convolutional neural networks», *Commun. ACM*, vol. 60, n.º 6, pp. 84-90, may 2017, doi: 10.1145/3065386.

Neurotecnologías (I)

Jueves 23 8:30 - 10:00
Aula Sebastián Ferignan

Metodología de análisis de la actividad cerebral en pacientes con ictus en fase de rehabilitación

M. Martín García de la Vega ¹, R. Pardo García ^{2,3}, E. M. Moreno Montes ^{2,3,4}, M. A. Pozo García ^{2,3}, G. Kontaxakis¹

¹ ETSI Telecomunicación, Universidad Politécnica de Madrid (UPM), Madrid, España;
m.mgarciadelavega@alumnos.upm.es, g.kontaxakis@upm.es

² Instituto Pluridisciplinar, Unidad de Cartografía Cerebral, Universidad Complutense de Madrid (UCM), Madrid, España;
rebeca.pardo96@gmail.com, emmoreno@pluri.ucm.es, pozo@ucm.es,

³ Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, España

⁴ Facultad de Lenguas y Educación, Universidad de Nebrija, Madrid, España

Resumen

Se ha desarrollado una plataforma de procesamiento y análisis de señales electroencefalográficas (EEG) para la evaluación de la efectividad de una nueva técnica de rehabilitación con realidad virtual en pacientes con ictus isquémico en fase crónica. Para ello, se han utilizado datos EEG de 64 canales de sujetos que realizaban tareas relacionadas con el movimiento de la mano derecha. Se ha elaborado un protocolo de preprocesamiento y post-procesamiento de señales EEG utilizando la plataforma EEGLAB. El preprocesado incluyó la eliminación de ruido a 50Hz, la re-referenciación, el rechazo de artefactos mediante Análisis de Componentes Independientes (ICA), y un filtro de paso banda de 0.5 a 50Hz. En el procesamiento, se analizaron los Potenciales Relacionados con Eventos (ERP). Este protocolo ha permitido extraer características relevantes de las señales EEG y se han encontrado diferencias entre los sujetos sanos y los pacientes con ictus.

1. Introducción

El ictus o accidente cerebrovascular es la interrupción de la circulación de sangre que llega al cerebro debido a la rotura (ictus hemorrágico) u obstrucción (ictus isquémico) de un vaso sanguíneo, lo cual provoca la muerte de las neuronas afectadas. Cabe destacar que, de todos los ictus, un 85% son isquémicos, y solo un 15% son hemorrágicos.

Actualmente, esta enfermedad es la segunda causa de muerte en España y la primera en el caso de mujeres, además de ser la primera causa de discapacidad. Por otro lado, la incidencia de esta enfermedad aumenta con la edad y, debido a que la población europea se está enfrentando a un envejecimiento generalizado (se estima que para 2050 el 35% de la población europea será mayor de 65 años), es inevitable que el número absoluto de casos de ictus aumente en los próximos años. Por ello, además de los indudables costos personales y emocionales, se trata de una enfermedad cuyo tratamiento y atención a largo plazo genera un gasto sanitario y social significativo [1].

Además, a pesar de los avances tecnológicos en el diagnóstico y tratamiento del ictus, se desconoce cuál es el proceso de recuperación post ictus, sobre todo en lo referente a la recuperación de la actividad motora ya que

en la mayoría de pacientes no sigue un patrón homogéneo. Por esta razón, actualmente cobra gran importancia la investigación de los cambios que ocurren en el sistema nervioso central tras un ictus, para así poder establecer pronósticos más adecuados y desarrollar nuevas metodologías para la recuperación motora.

En este trabajo, se ha empleado la electroencefalografía (EEG), que es una técnica no invasiva y de bajo coste que registra la actividad eléctrica del cerebro mediante la colocación de electrodos en el cuero cabelludo. Se utiliza para investigar trastornos neurológicos y proporciona información sobre diferentes tipos de ondas cerebrales, como las bandas alfa, beta, theta, delta y gamma.

En el proceso, es esencial eliminar los artefactos, que son señales no cerebrales que interfieren en los datos. Estos pueden ser de origen fisiológico (como el movimiento ocular, parpadeo, actividad muscular o cardíaca) o no fisiológico (como problemas con los electrodos o interferencias eléctricas).

Una vez los datos están limpios, para estudiar eventos cognitivos específicos, se emplean los Potenciales Relacionados con Eventos (ERP). Esto implica el promedio de las señales EEG registradas después de varios eventos para resaltar patrones relacionados con esos eventos. Cuantos más ensayos se promedian, mejor es la calidad de los datos.

En el contexto del ictus, la EEG puede revelar cambios en las bandas alfa y beta, y estos cambios pueden diferir entre personas sanas y pacientes con ictus, así como entre aquellos que siguen terapias de rehabilitación tradicionales y los que utilizan realidad virtual. Estos análisis aportan pruebas objetivas que ayudan a comprender mejor los efectos del ictus en la función cerebral y a mejorar las estrategias de rehabilitación, siendo el objetivo central de este trabajo proporcionar una metodología de análisis que contribuya a esta comprensión [2] [3].

2. Metodología

Para el desarrollo del protocolo y la evaluación de su eficacia, se ha trabajado con los datos EEG de sujetos control y pacientes con ictus isquémico en fase crónica

que han sido atendidos en el Hospital Clínico San Carlos de dicha patología. Es importante destacar que, aunque el experimento incluyó un grupo de tres sujetos control y seis pacientes con ictus, el enfoque principal del presente artículo se centra en la comparación entre un sujeto sano y uno con ictus. De ese modo se pretende mostrar que la metodología empleada es efectiva para identificar las diferencias significativas entre estos dos grupos.

2.1. Diseño del experimento

El montaje de los electrodos se ha realizado siguiendo el sistema internacional 10-20, que posiciona los electrodos en relación con puntos en el cráneo para mantener la consistencia en los estudios. Cada electrodo se etiqueta con una letra y número, denotando la región cerebral y lado del cerebro. Se han utilizado 64 electrodos incluyendo el mastoides (A2 y A3) como referencia. También se ha obtenido la señal de electrodos para medir movimientos oculares horizontales (HEOG) y verticales (VEOG), importantes en el estudio de artefactos oculares.

Los datos pertenecen a pacientes con ictus isquémico en el hemisferio dominante, en todos el izquierdo, que controla el lado derecho del cuerpo. El experimento, por lo tanto, involucraba la realización de movimientos con la mano derecha reflejados en la Figura 1. Se mostraba a los pacientes primero una imagen del movimiento a imitar y tras dos segundos escuchaban un pitido que les indicaba que debían hacerlo. Los eventos de aparición de imagen y de sonido del pitido quedaban registrados en los datos EEG para marcar el instante de cada estímulo.



Figura 1. Movimientos que deben realizar los sujetos

2.2. Preprocesado de los datos

En cuanto al protocolo de preprocesado, se ha seguido el modelo para la plataforma Brain Vision, con algunos cambios y adaptaciones. A continuación, se describen las técnicas utilizadas para procesar datos de EEG en un orden cronológico utilizando EEGLAB [4].

- Edición de la localización de electrodos. Una vez los datos han sido importados, se ha corregido la información de la ubicación espacial de los electrodos utilizando una base de datos estándar de etiquetas de canales (Standard-10-5-Cap385.sfp).
- Filtro Notch: Se ha aplicado un filtro Notch, que se caracteriza por dejar pasar todas las frecuencias excepto aquellas pertenecientes a una banda estrecha de frecuencias. Este se ha utilizado para eliminar la interferencia de la frecuencia de alimentación eléctrica, que en Europa se da a 50Hz. De esta manera, se ha aplicado el filtro Notch entre 49 y 51Hz.

A partir de este punto, los análisis no se han realizado sobre los canales oculares, ya que su objetivo es solo identificar parpadeos.

- Re-referenciación de canales: Se ha utilizado el método de referencia media común, que consiste en referenciar de nuevo cada canal a la media de todos ellos. El fin de este paso es reducir el ruido espacial que pudiese generar la referencia original y así asegurar una representación equilibrada de la actividad cerebral en todos los electrodos.
- Rechazo de canales ruidosos: Se han identificado y eliminado canales defectuosos utilizando análisis espectral. Para ello, se ha calculado el espectro promedio de todos los canales seleccionados y se ha establecido un umbral en el espectro promedio. En este caso se ha elegido umbral 3, en términos de desviación estándar; de manera que se identifican y eliminan los canales cuyo espectro esté por encima del umbral establecido.
- Eliminación de artefactos a través de Análisis de Componentes Independientes (ICA): Se aplicó ICA para separar señales independientes en el EEG, incluyendo componentes cerebrales y artefactos de movimientos oculares, parpadeos, actividad muscular, actividad cardíaca y ruidos. Se marcaron y eliminaron los componentes de ICA que se debían en menos del 10% a la actividad cerebral.
- Filtrado en banda: Se ha aplicado un filtro de paso banda de 0.5 Hz a 50Hz para resaltar la información relevante sobre la actividad cerebral y atenuar artefactos de alta frecuencia.
- Segmentación: Los datos se han dividido en épocas desde tres segundos antes hasta cinco segundos después del sonido del pitido, como queda reflejado en la Figura 2. De esta manera, se incluye la aparición de la imagen, que ocurría dos segundos antes del pitido. Se ha utilizado la extensión ERPLAB para este proceso. Por lo tanto, en este trabajo el estudio se ha centrado en la ejecución del movimiento, más que en su preparación.

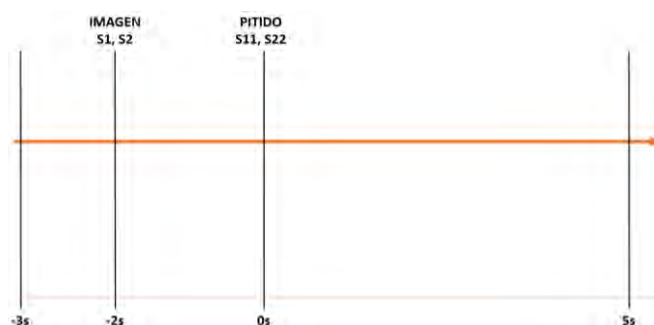


Figura 2. Esquema segmentación de datos en torno a eventos

- Eliminación de línea de base: Se ha eliminado la línea de base en cada época para evitar sesgos en el análisis.

- Acotado de amplitud (rechazo de artefactos): Se han rechazado segmentos de datos donde las amplitudes estaban fuera del rango $[-200\mu\text{V}, 200\mu\text{V}]$, ya que las amplitudes de los ERP suelen ser mucho menores.

2.3. Post-procesado de los datos

Los ERP se calculan a partir del promedio de la señal en torno a un evento específico, de manera que permiten identificar y aislar potenciales eléctricos relacionados con esos eventos específicos (el sonido del pitido). Para su estudio, EEGLAB permite representar estos potenciales en las denominadas imágenes ERP, que ofrecen una visión detallada de la actividad cerebral en diferentes momentos del tiempo. Este estudio se ha realizado principalmente en el electrodo C3 (hemisferio izquierdo) al investigar el movimiento de la mano derecha debido a la organización contralateral de la corteza motora en el cerebro.

Dentro del análisis ERP se ha empleado el análisis wavelet, que identifica las diferentes frecuencias en el momento en que ocurren en la señal EEG, para clasificar ensayos por fase espectral en ventanas de tiempo/frecuencia específicas. Además, este análisis ha permitido obtener métricas como la potencia espectral relacionada con eventos y la sincronización de fase (ITC, debido a sus siglas en inglés inter-trial coherence). Esta última mide la sincronización entre los distintos ensayos, de manera que permite determinar estadísticamente si la fase de la actividad α o β se resincroniza parcialmente por los estímulos. Se ha establecido un nivel de significancia del 0.01 para determinar la relevancia estadística de los resultados. En este punto, es importante destacar la diferencia entre la sincronización entre ensayos y la sincronización de ritmo en las imágenes ERP. La primera hace referencia a la ITC y la segunda comprende el estudio de la potencia en bandas de frecuencia para analizar la desincronización (ERD) y sincronización (ERS) de los ritmos cerebrales [5].

Por último, se han realizado comparaciones de ERP entre electrodos en diferentes hemisferios para analizar cambios relacionados con eventos durante el movimiento de la mano derecha [6].

3. Resultados

Las Figura 3 y Figura 4 reflejan la misma señal EEG representada a igual escala, antes y después de aplicar todo el protocolo explicado anteriormente.

En relación con el post-procesado, los resultados más relevantes se han obtenido al analizar las imágenes ERP. En el estudio del electrodo C3 del sujeto sano, se observó un aumento tanto en la potencia α como en la β a unos 150 ms después del estímulo, mientras que antes del estímulo permaneció constante. La ITC también mostró un aumento significativo alrededor de los 150 ms, lo que refleja que efectivamente tras el estímulo del pitido en todos los ensayos ocurre algo parecido. Ambos análisis indicaron que lo observado en la imagen ERP

correspondía a una ERS en ambas bandas. Por otro lado, en el sujeto con ictus, la coherencia de fase fue satisfactoria. Sin embargo, el aumento en la potencia α y β tras 150ms es menos evidente que en el sujeto sano y la potencia no se mantuvo constante antes y después del estímulo, de manera que no se puede concluir la presencia de ERD ni ERS en ninguna banda.

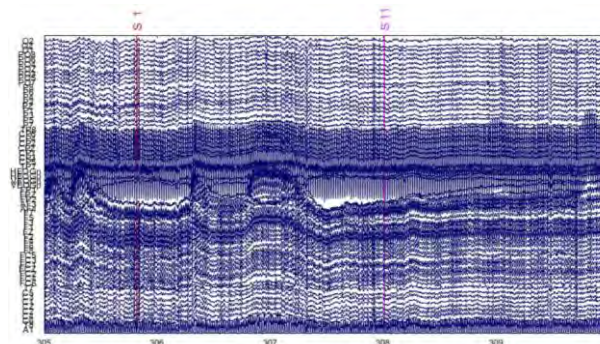


Figura 3. Señal electroencefalográfica cruda de un sujeto sano

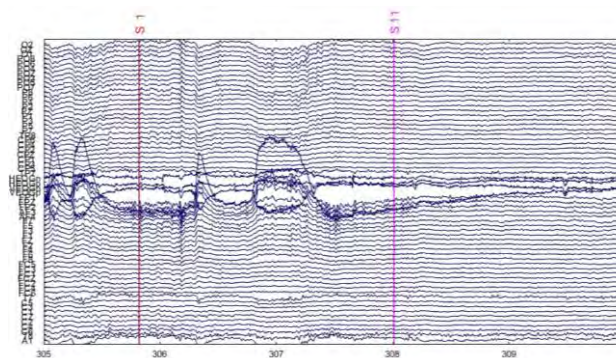


Figura 4. Señal electroencefalográfica de un sujeto sano tras aplicar filtro Notch a 50Hz, re-referenciar canales a la media, eliminar artefactos con ICA y filtrar paso banda (0.5-50Hz).

Por otro lado, se han observado diferencias relevantes entre sujeto sano y sujeto que ha padecido ictus en la comparación de imágenes ERP entre hemisferios. En este tipo de imagen, el voltaje negativo se indica en azul y el voltaje positivo en rojo.

En el sujeto sano, reflejado en la Figura 5, se observó que la disminución de potencial en el hemisferio izquierdo fue más pronunciada después del estímulo, con una magnitud de alrededor de $-2.5 \mu\text{V}$, mientras que en el hemisferio derecho no llegó a $-1 \mu\text{V}$. Además, el aumento de potencial después de la disminución fue más prolongado en el hemisferio derecho. En cuanto a los electrodos C5 y C6, no mostraron un patrón claro.

En cuanto al sujeto de ictus, mostrado en la Figura 6, el C1 fue rechazado por ser ruidoso, de modo que se pueden realizar observaciones limitadas sobre el hemisferio izquierdo. La despolarización que se observaba claramente en el canal C3 del sujeto sano apenas se aprecia en el hemisferio afectado (Figura 6 b). Además, su magnitud es mucho menor, sin alcanzar $-1 \mu\text{V}$, lo que es coherente con el hecho de ser el hemisferio afectado.

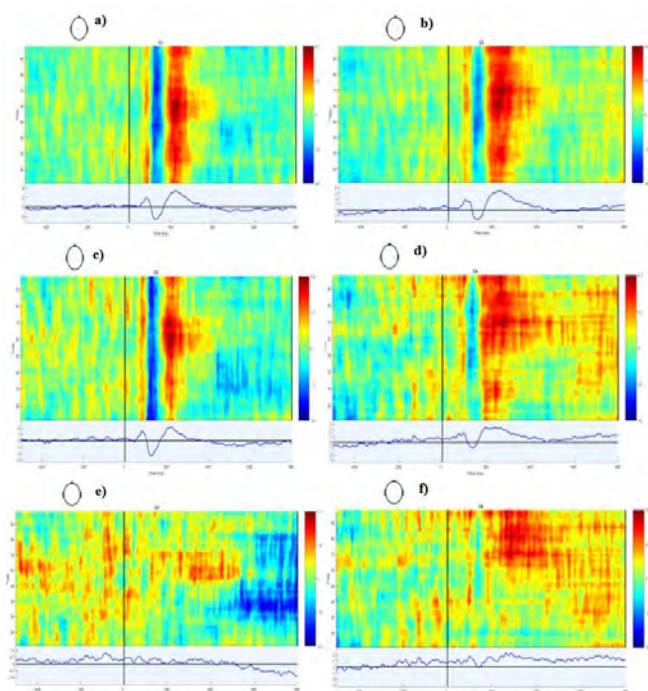


Figura 5. ERP del canal en las posiciones de los electrodos (a) C1, (b) C2, (c) C3, (d) C4, (e) C5 y (f) C6 de un sujeto sano durante el movimiento de la mano derecha.

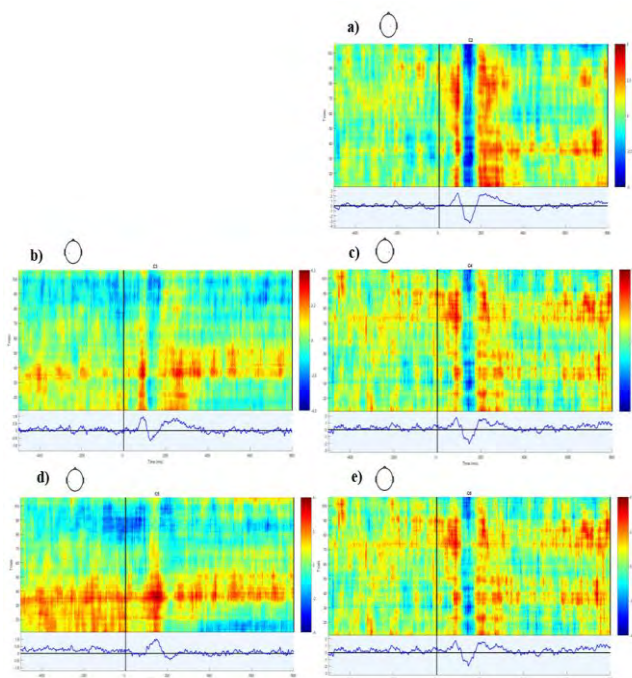


Figura 6. ERP del canal en las posiciones de los electrodos (a) C2, (b) C3, (c) C4, (d) C5 y (e) C6 de un paciente de ictus durante el movimiento de la mano derecha.

Por otro lado, en el hemisferio derecho se observa un aumento en la magnitud de la despolarización en comparación con el sujeto sano y el hemisferio izquierdo. También se nota una línea vertical azul más marcada, que era característica del hemisferio izquierdo en el sujeto sano. Estos hallazgos coinciden con estudios que indican que los pacientes con ictus crónico muestran un aumento significativo de los ERPs sobre el hemisferio sano [7].

4. Conclusiones

El enfoque principal del trabajo fue el procesamiento de señales EEG para obtener información relevante. Los resultados obtenidos revelan diferencias entre sujetos sanos y no sanos durante la realización de actividades motoras. En concreto, en el paciente sano se observa una ERS en α y β , mientras que en el paciente de ictus no se puede afirmar que exista una sincronización. Por otro lado, la disminución de voltaje es más marcada en el hemisferio izquierdo en el sujeto sano, mientras que en el paciente de ictus es propia del hemisferio sano.

Por lo tanto, este análisis es un primer paso importante en la comprensión de la respuesta cerebral en el contexto de la rehabilitación post-ictus, pues aporta datos objetivos de la alteración funcional y permite evaluar las diferentes intervenciones en la rehabilitación de estos pacientes.

5. Agradecimientos

Este trabajo ha sido parcialmente financiado por el proyecto S2022/BMD-7236 MINA-CM de la Comunidad de Madrid y el proyecto IEL.GKA.100 DIANIA de FUNDETEL-UPM.

6. Bibliografía

- [1] Soto et al., «Prevalencia e incidencia de ictus en Europa: revisión sistemática y metaanálisis,» *Anales del Sistema Sanitario de Navarra*, vol. 45, n° 1, 2022.
- [2] F. H. Lopes da Silva, *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, Florida: Lippincott Williams & Wilkins, 2005.
- [3] G. Pfurtscheller, M. Pregenzer y C. Neuper, «Visualization of sensorimotor areas involved in preparation for hand movement based on classification of mu and central beta rhythms in single EEG trials in man,» *Neuroscience Letters*, vol. 181, n° 1-2, pp. 43-46, 1994.
- [4] A. Delorme, «EEGLAB Wiki,» [En línea]. Available: <https://eeglab.org/tutorials/>.
- [5] G. Pfurtscheller y F. Lopes da Silva, «Event-related EEG/MEG synchronization and desynchronization: basic,» *Clinical Neurophysiology*, vol. 110, pp. 1842-1857, 1999.
- [6] M. Malini, K. E. Vidyasagar, D. Saritha y B. Sailalitha, «Offline Analysis of Motor Imagery EEG Signals,» *International Journal of Engineering Research & Technology (IJERT)*, vol. 4, n° 26, 2016.
- [7] E. Monge-Pereira, F. Molina-Rueda, F. Rivas-Montero, J. Ibáñez, J. Serrano, I. Alguacil-Diego y J. Miangolarra-Page, «Electroencefalografía como método de evaluación tras ictus. Una revisión actualizada,» *Neurología*, vol. 32, n° 1, pp. 40-49, 2017.

Análisis de señales EEG en movimientos de flexión plantar y dorsal mediante el empleo de un exoesqueleto de bajo coste para la caracterización de la acción motora

C. Polo-Hortigüela^{1,2}, J.V. Juan^{1,2,3}, M. Ortiz^{1,2}, E. Iáñez^{1,2}, J.M Azorín^{1,2,4}

¹ Brain-Machine Interface Systems Lab, Universidad Miguel Hernández de Elche, Elche, España
{cpolo,javier.juanp,mortiz,eianez,jm.azorin}@umh.es

² Instituto de Investigación en Ingeniería de Elche-I3E, Universidad Miguel Hernández de Elche, Elche, España

³ Center for Clinical Neuroscience Hospital Los Madroños, Hospital Los Madroños, Brunete, España

⁴ ValGRAI: Valencian Graduated School and Research Network of Artificial Intelligence, España

Resumen

Entre las técnicas de neurorrehabilitación que se emplean actualmente se encuentra la aplicación conjunta de una interfaz cerebro-máquina (BMI, del inglés Brain-Machine Interface) y un dispositivo robótico como es el caso de un exoesqueleto. La interfaz cerebro-máquina involucra la adquisición y decodificación de señales electroencefalográficas (EEG) con el fin de generar comandos de control. Los patrones cerebrales, en este caso, vienen asociados a la tarea mental de imaginación del movimiento sin ejecutarlo. Esto último se conoce como el paradigma de imaginación motora (MI, del inglés Motor Imagery). Durante la ejecución del movimiento se generan potenciales asociados a la acción motora. Para poder lograr un óptimo algoritmo de control de la activación y desactivación del exoesqueleto se debe caracterizar la acción motora frente a la imaginación motora. Por ello, en este trabajo se plantea una primera caracterización de la acción motora mediante el estudio de la señal EEG asociada a movimientos de flexión dorsal y flexión plantar de la articulación del tobillo. Las señales EEG se han obtenido a partir de dos pruebas realizadas a dos sujetos sanos empleando un exoesqueleto de bajo coste. Para llevar a cabo el análisis de dichas señales se ha empleado la transformada en frecuencia-tiempo de Stockwell. Sin embargo, no se han hallado diferencias significativas entre las señales de ambos movimientos, pero sí ha supuesto el punto de partida para varias líneas de investigación futuras.

1. Introducción

Los accidentes cerebrovasculares o las lesiones de médula espinal son dos de los principales agentes causantes de la pérdida de funcionalidad motora reduciendo así la capacidad en la ejecución de movimientos como puede ser la marcha humana. Actualmente, entre las terapias empleadas en la neurorrehabilitación de este tipo de enfermedades se incluye un sistema conformado por una interfaz cerebro-máquina no invasiva y un exoesqueleto enfocado en la región anatómica que se pretende rehabilitar [1]. Se ha demostrado que este tipo de técnicas tiene un efecto positivo en la mejora de la neuroplasticidad [2].

Los exoesqueletos son dispositivos robóticos que permiten la asistencia en el movimiento de una determinada articulación. Una de las problemáticas de estos dispositivos es su elevado coste [3], por ello en este estudio se emplea un dispositivo de bajo coste el cual no puede ser mostrado ya que se encuentra bajo proceso de patente. Este

dispositivo ha sido conformado para la rehabilitación de tobillo ya que es una articulación fundamental en la marcha humana durante los movimientos de flexión plantar y flexión dorsal. El movimiento de flexión plantar supone el descenso del pie con un rango máximo de 45°. La flexión dorsal es el movimiento contrario de elevación reduciéndose el rango de movimiento a 20° [4].

Una interfaz cerebro-máquina registra y procesa señales EEG para evaluar la activación de conjuntos neuronales en diferentes áreas y ritmos cerebrales. El objetivo principal de esta técnica es generar comandos de control a partir de tareas mentales, siendo la imaginación motora una solución prometedora[5]. Este paradigma implica imaginar el movimiento de una extremidad sin ejecutarlo[6]. Para lograr un control óptimo de los dispositivos es importante caracterizar la acción motora (potencial del movimiento del pie generado por exoesqueleto) y la imaginación motora [7].

En este trabajo se ha realizado un estudio preliminar para caracterizar la acción motora diferenciando la señal EEG en un movimiento de flexión dorsal y flexión plantar. Para ello se ha utilizado un enfoque distinto a los métodos de existentes de extracción de características espaciales [8] por un análisis de transformadas en frecuencia-tiempo [9]. Con esta técnica se pretende decodificar la señal EEG con mayor precisión temporal.

2. Materiales y métodos

2.1. Participantes

En esta prueba han participado dos mujeres (S1 y S2) sin ningún tipo de disfunción motora o neuronal diagnosticada. Los dos sujetos presentan una edad de 26 años. Cabe mencionar que los dos usuarios presentan lateralidad derecha. Los sujetos fueron informados previamente sobre el ensayo a realizar y firmaron el correspondiente consentimiento informado de acuerdo con la declaración de Helsinki. Todos los procedimientos que se llevaron a cabo fueron aprobados por la Oficina de Investigación Responsable de la Universidad Miguel Hernández de Elche (España) con referencia DIS.JAP.01.22.

2.2. Protocolo experimental

El protocolo que se ha definido para la adquisición de señales EEG en bucle abierto tiene como fin ser válido para diseñar un paradigma de control del exoesqueleto de tobillo basado en imaginación motora y máquina de estados. Para ello, se define dos modelos: modelo en estático y modelo en movimiento en combinación con dos tareas mentales, relax e imaginación motora (Figura 1). El exoesqueleto se mantendrá inmóvil a lo largo de todas las repeticiones del modelo en estático a diferencia de su estado en el modelo en movimiento que se encontrará asistiendo al sujeto en la realización de los movimientos de flexión plantar y flexión dorsal.

Al inicio de cada repetición se mantienen 15 segundos en los que el sujeto no está realizando ninguna actividad

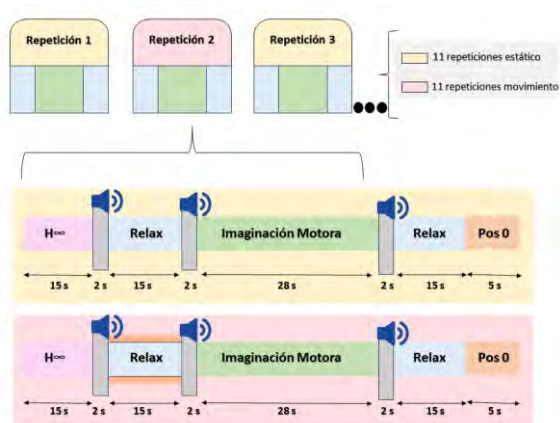


Figura 1. Protocolo experimental: modo estático (color amarillo), modo movimiento (color rosa). Queda resaltado en color naranja el tramo de relax que ha sido analizado en dicho estudio.

mental y que son necesarios para la convergencia del filtro que elimina los artefactos oculares. Después de dicho tiempo se activa durante 2 segundos un indicador auditivo por voz que indica la preparación del estado de “relax”. En esta actividad mental el sujeto debe mantener un estado de relajación y transcurre durante 15 segundos. Una vez finalizado ese tiempo vuelve a activarse el indicador auditivo “imagina” que señala el inicio de la actividad mental de imaginación motora. La imaginación motora que debe realizarse es un movimiento constante de flexión dorsal y flexión plantar del pie dominante. Esta tarea tiene una duración de 28 segundos. Nuevamente, tras este periodo se inicia otros 15 segundos de relax con su correspondiente indicador auditivo. Al finalizar ese último periodo se añaden otros 5 segundos en los que el exoesqueleto vuelve a la posición de inicio con el fin de no terminar la repetición en una postura desfavorable para el sujeto. Este tiempo se ha añadido tanto en el modelo estático como en movimiento para balancear los tiempos de las repeticiones de los dos modelos. En total se realizan 22 repeticiones de los dos modelos.

Previamente a al inicio del ensayo se realiza una calibración de los rangos de movimientos de flexión dorsal y flexión plantar que presenta cada sujeto con el objetivo

de caracterizar las posiciones máximas y mínimas que debe alcanzar el exoesqueleto empleado.

2.3. Material empleado

Para el registro de señales EEG se utiliza el equipo y gorro de electrodos g.Nautilus (g.tec medical engineering GmbH, Austria). El gorro está compuesto por un total de 32 electrodos húmedos no invasivos y 3 acelerómetros. De los 32 electrodos se utilizaron 28 siguiendo el sistema internacional 10-10: AF3, F3, FZ, FC3, FC1, FCZ, C5, C3, C1, CZ, CP3, CP1, CPZ, P3, PZ, PO3, AF4, F4, FC2, FC4, C2, C4, C6, CP2, CP4, P4, POZ, PO4. El electrodo AFZ se utiliza como GND y la referencia se localiza en el lóbulo de la oreja derecha del sujeto. Los cuatro electrodos restantes se utilizan para la captación de señales electrooculográficas (EOG) tales como parpadeos o movimientos oculares. Dicho gorro se comunica con el PC vía inalámbrica mediante un amplificador, que se localiza en la parte posterior del gorro y a su vez está conectado a los electrodos, y un receptor de señales conectado vía USB al mismo.

Para llevar a cabo los movimientos de flexión plantar y flexión dorsal se hace uso de un exoesqueleto de bajo coste que se localiza en el tobillo del sujeto. A grandes rasgos dicho dispositivo consta de un diseño mecánico y una electrónica que permiten la realización de los movimientos mencionados mientras el sujeto se encuentra sentado en una camilla manteniendo las dos piernas en voladizo tal y como se recrea en la Figura 2.

Para conocer la posición del pie del sujeto en cada repetición se ha hecho uso de cuatro unidades de medición inercial (IMUs) de la marca WIT Motion localizadas dos a dos en el eje rotor y en la plantilla del dispositivo. Dichos dispositivos situadas dos a dos en la plantilla del dispositivo y en los ejes. La sincronización de estos tres dispositivos que se han descrito se ha llevado a cabo mediante una arquitectura conformada en la filosofía de BMI desarrollada en Matlab.

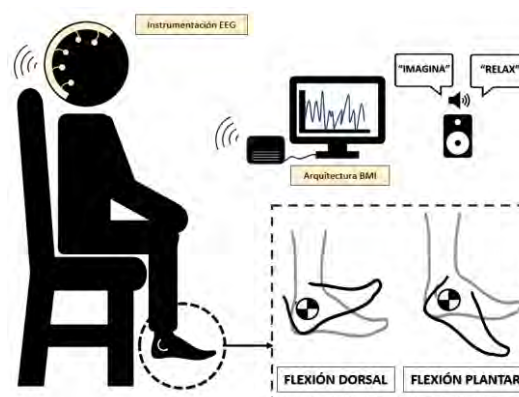


Figura 2. Representación de la posición del sujeto durante la prueba. En el recuadro inferior derecho se indican los movimientos de flexión plantar (FP) y flexión dorsal (FD). En los dos movimientos siempre se vuelve a una posición intermedia que es la resaltada en color gris claro.

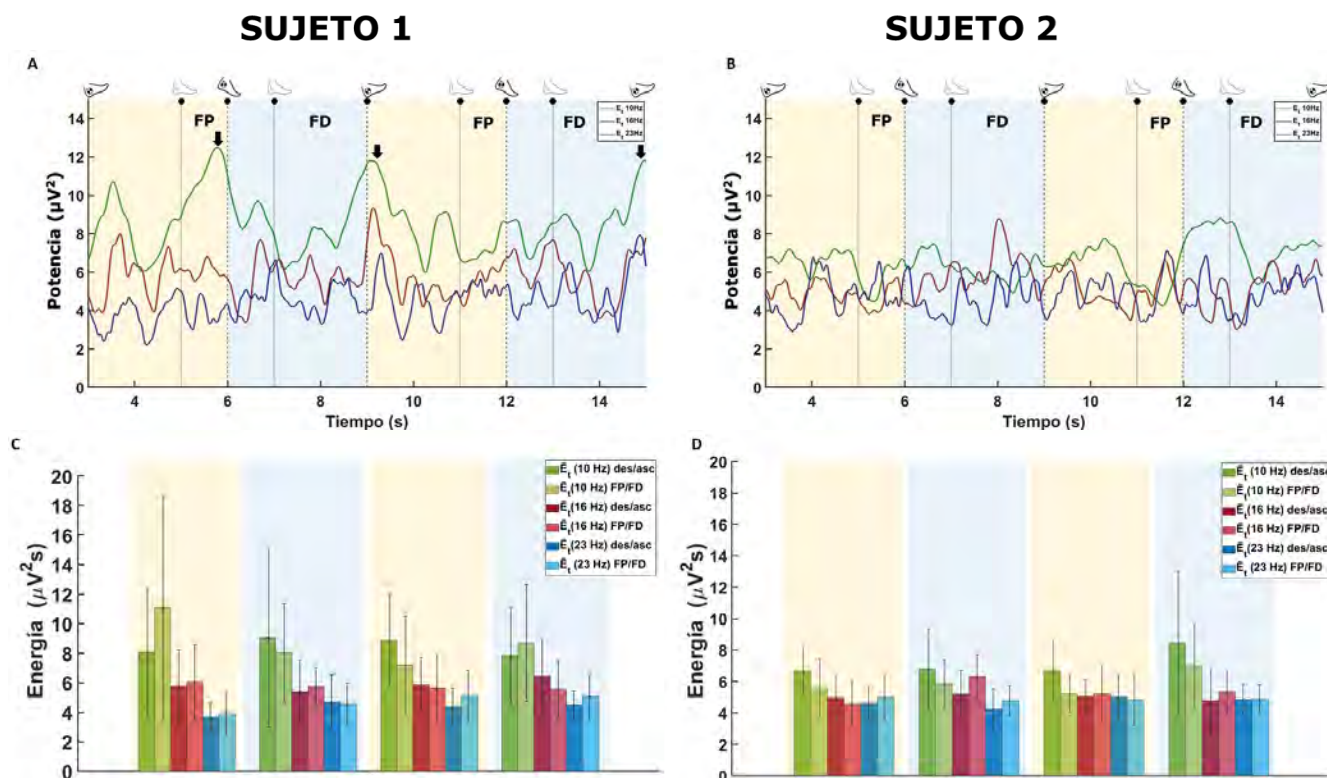


Figura 3. Las figuras 3A y 3B representan las potencias promedio suavizadas de las repeticiones de cada usuario en los tonos centrales de cada banda de estudio. Dichas gráficas están divididas en función de la actividad que está realizando el pie: los tramos azules representan el estado de descenso del pie (des) y los tramos amarillos en los que se encuentra en ascenso (asc). Dentro de dichos tramos encontramos los movimientos propios de flexión plantar (FP) y flexión dorsal (FD). También se indica mediante las verticales punteadas y pie en tono gris claro los instantes en los que el pie alcanza la posición intermedia. Por otro lado, las figuras 3C y 3D representan las energías promedio de cada tramo de las señales suavizadas de la parte superior. Dichas gráficas se encuentran divididas de la misma forma que las anteriores. Los tonos oscuros representan los movimientos anteriores de descenso o ascenso a los movimientos de flexión plantar y flexión dorsal que se representan en tonos claros.

2.4. Procesamiento señal EEG

Las señales EEG y EOG son registradas en la arquitectura a una frecuencia de muestreo de 250 Hz que previamente han sido filtradas mediante un filtro hardware paso banda de 0.5Hz-250Hz. Dicha señal EEG lleva consigo los artefactos de parpadeos y movimientos oculares por lo que es necesario la aplicación del filtro H_∞ [10]. Tras dicho proceso se observa que la señal presenta una componente de baja frecuencia que dificulta el análisis de la misma por lo que se aplica un filtro paso alto de 1Hz que elimina de forma satisfactoria dicha componente.

2.5. Análisis frecuencia-tiempo

El presente estudio se centra el análisis en las señales EEG, concretamente en el electrodo CZ durante la tarea del primer relax en las repeticiones de movimiento. Esto es debido a que es el tramo de señal óptimo para evaluar la acción motora que no lleva implícito ninguna tarea mental.

Para analizar dicha señal en frecuencia-tiempo se hace uso de la transformada de Stockwell [11]. Mediante dicha técnica se obtiene la potencia instantánea para cada una de las bandas. Integrando en las bandas de frecuencia 8-13Hz, 14-19Hz y 20-25Hz se obtiene el espectro marginal asociado a los ritmos alfa y beta de la actividad cerebral. Además, para limitar la oscilación por periodo dominante de cada tono de la potencia instantánea se realiza un suavizado en los tonos centrales de cada banda: 10, 16 y 23

Hz. Una segunda integración del espectro marginal en el tiempo nos permite obtener la energía de la señal promedio. Dicha señal se divide en los tramos de interés de análisis para el presente estudio. Dichos intervalos son los correspondientes a los momentos de ascenso y descenso del pie y dentro de estos, los propios a la flexión plantar y dorsal. Este proceso se realiza para todas las repeticiones de cada usuario, por lo que obtendremos la potencia instantánea y la energía promedio por tramo de todas las repeticiones para cada banda de frecuencia de estudio. A continuación, se procede a realizar un promedio de todas las señales mediante el cual obtendremos la energía promedio por ritmo para todos los registros. Por otro lado, para conocer la desviación típica de los valores de las energías para cada tramo se parte del promedio del suavizado de las energías para todas las repeticiones y para cada banda de frecuencia.

3. Resultados y discusión

La Figura 3A y la Figura 3B representan las potencias promedio instantáneas suavizadas para cada tono central de cada banda y para cada usuario. Si se atiende a la correspondiente para el sujeto 1 se observa que la señal en un tono de 10 Hz tiene un comportamiento creciente en los tramos propios de flexión plantar (FP) y flexión dorsal (FD). Dicha situación es contraria cuando el usuario está realizando los movimientos de retorno a la posición intermedia en los cuales la señal toma un valor decreciente.

Además, se observa que los picos máximos de energía de cada tramo se localizan muy próximos a los puntos en los que se realiza la máxima amplitud dorsal y plantar por el sujeto (flecha negra). En el caso del sujeto 2 también se observa una misma similitud con el sujeto 1 pero en este caso solo ocurre el comportamiento creciente en la flexión plantar. La Figura 3C y la figura 3D representan las energías promedio de cada tramo, así como las desviaciones típicas entre todas las repeticiones de cada usuario. En el caso del sujeto 1 se observa que la señal presenta mayor energía promedio en el tono de 10Hz y que dicha energía disminuye gradualmente a medida que se aumenta el tono de frecuencia. En cuanto a los resultados del sujeto 2 se observa que en el tono de 10Hz se requiere más energía en los movimientos hacia la posición intermedia. Además de esto, se distingue que la flexión dorsal es el movimiento que mayor energía promedio requiere en los tonos de 10Hz y 16Hz. En cuanto a las desviaciones de los valores promedios en el caso de los dos usuarios toman valores mayores en el tono de 10 Hz.

4. Conclusiones

El estudio en cuestión es una investigación inicial que utiliza transformadas frecuencia-tiempo para analizar la acción motora caracterizando los movimientos de flexión dorsal y flexión plantar. Sin embargo, los resultados todavía no son lo suficientemente sólidos para diferenciar de manera definitiva dicha señales. No obstante, puede concluirse que la banda que más información ha aportado para ello ha sido la correspondiente a la banda alfa. Además, en el sujeto 1 existe un comportamiento similar de la onda en los movimientos de ascenso del pie al igual que el sujeto 2 en el movimiento previo a la flexión dorsal. Esto último supone un punto de partida para verificarse en un estudio futuro. También se ha podido distinguir como existen diferencias en los valores de las energías promedio entre los tonos de 10Hz, 16 Hz y 23Hz.

Con el fin de poder mejorar este estudio preliminar será necesario obtener datos para un mayor número de sujetos. Además, se plantea el uso de otras transformadas frecuencia-tiempo. A posteriori será necesario el estudio de los tramos de imaginación motora para comparar si dicho comportamiento también se registra en dichos intervalos y poder así caracterizar dicha acción mental con la acción motora.

Agradecimientos

Esta publicación forma parte del programa "Yo Investigo" en el marco del Plan de Recuperación por la Unión Europea-NextGenerationEU. Está relacionada con el proyecto de investigación PID2021-124111OB-C31, financiado por MCIN/AEI y FEDER. Asimismo, ha recibido apoyo para la formación predoctoral en colaboración con empresas por parte de la Universidad Miguel Hernández de Elche y financiación de la Generalitat Valenciana para llevar a cabo una tesis industrial. La investigación se enfoca en el proyecto "Neurokit" financiado por el Centro Internacional para la Investigación del Envejecimiento de la Fundación de la Comunitat Valenciana (ICAR).

Referencias

- [1] Cramer, S. C. (2008). Repairing the human brain after stroke. II. Restorative therapies. In *Annals of Neurology* (Vol. 63, Issue 5). <https://doi.org/10.1002/ana.21412>
- [2] He, Y., Eguren, D., Azorín, J. M., Grossman, R. G., Luu, T. P., & Contreras-Vidal, J. L. (2018). Brain-machine interfaces for controlling lower-limb powered robotic systems. In *Journal of Neural Engineering* (Vol. 15, Issue 2). <https://doi.org/10.1088/1741-2552/aaa8c0>
- [3] Pinto, D., Garnier, M., Barbas, J., Chang, S. H., Charlifue, S., Field-Fote, E., Furbish, C., Tefertiller, C., Mummidisetty, C. K., Taylor, H., Jayaraman, A., & Heinemann, A. W. (2020). Budget impact analysis of robotic exoskeleton use for locomotor training following spinal cord injury in four SCI Model Systems. *Journal of NeuroEngineering and Rehabilitation*, 17(1). <https://doi.org/10.1186/s12984-019-0639-0>
- [4] Calderón Coy, M. V., Coronel Segrera, C. A., & Camargo Casallas, E. (2015). Prototipo de emulación de movimiento humano mediante una plataforma bípeda humanoide. *Revista Tecnura*, 18. <https://doi.org/10.14483/udistrital.jour.tecnura.2014.se1.a16>
- [5] Vasilyev, A., Liburkina, S., Yakovlev, L., Perepelkina, O., & Kaplan, A. (2017). Assessing motor imagery in brain-computer interface training: Psychological and neurophysiological correlates. *Neuropsychologia*, 97. <https://doi.org/10.1016/j.neuropsychologia.2017.02.005>
- [6] Ferrero, L., Quiles, V., Ortiz, M., Iáñez, E., & Azorín, J. M. (2021b). A BMI Based on Motor Imagery and Attention for Commanding a Lower-Limb Robotic Exoskeleton: A Case Study. *Applied Sciences*, 11(9), 4106. <https://doi.org/10.3390/app11094106>
- [7] Maksimenko, V. A., Pavlov, A., Runnova, A. E., Nedaivozov, V., Grubov, V., Koronovslii, A., Pchelintseva, S. V., Pitsik, E., Pisarchik, A. N., & Hramov, A. E. (2018). Nonlinear analysis of brain activity, associated with motor action and motor imaginary in untrained subjects. *Nonlinear Dynamics*, 91(4). <https://doi.org/10.1007/s11071-018-4047-y>
- [8] Ferrero, L., Ortiz, M., Quiles, V., Iáñez, E., & Azorín, J. M. (2021). Improving motor imagery of gait on a brain-computer interface by means of virtual reality: A case of study. *IEEE Access*, 9. <https://doi.org/10.1109/ACCESS.2021.3068929>
- [9] Ortiz, M., Rodríguez-Ugarte, M., Iañez, E., & Azorin, J. M. (2018). Comparison of different EEG signal analysis techniques for an offline lower limb motor imagery brain-computer interface. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2018-July*. <https://doi.org/10.1109/EMBC.2018.8512256>
- [10] Kilcarslan, A., Grossman, R. G., & Contreras-Vidal, J. L. (2016). A robust adaptive denoising framework for real-time artifact removal in scalp EEG measurements. *Journal of Neural Engineering*, 13(2). <https://doi.org/10.1088/1741-2560/13/2/026013>
- [11] Stockwell, R. G., Mansinha, L., & Lowe, R. P. (1996). Localization of the complex spectrum: The S transform. *IEEE Transactions on Signal Processing*, 44(4). <https://doi.org/10.1109/78.492555>

Aplicación de modelo de Deep Learning IFNet en decodificación de imaginación motora del pedaleo para señales EEG

J. V. Juan^{1,2}, R. Martínez^{2,3,4}, E. Iáñez^{1,5}, M. Ortiz^{1,5}, J. Tornero^{2,4}, J. M. Azorín^{1,5,6}

¹ Brain-Machine Interface Systems Lab, Universidad Miguel Hernández de Elche,

Elche, España, {javier.juanp,eianez,mortiz,jm.azorin}@umh.es, <http://bmi.edu.umh.es/>

² Center for Clinical Neuroscience HLM, Hospital Los Madroños, Brunete, España, jesus.tornero@lmh.es

³ Universidad Autónoma de Madrid, Madrid, España, ruben.martinez01@estudiante.uam.es

⁴ INNTEGRA, Hospital Los Madroños, Brunete, España

⁵ Instituto de Investigación en Ingeniería de Elche-I3E, Universidad Miguel Hernández de Elche, Elche, España

⁶ ValGRAI: Valencian Graduated School and Research Network of Artificial Intelligence, España

Resumen

El principio fundamental para decodificar la Imaginación Motora (MI, Motor Imagery) en interfaces cerebro-máquina (BMIs, Brain-Machine Interfaces) basadas en electroencefalogramas (EEG) radica en la extracción eficiente de características discriminativas de tareas, abarcando los dominios espectrales, espaciales y temporales de manera conjunta. Sin embargo, las señales EEG presentan desafíos debido a su no estacionalidad y a la fuerte presencia de ruido habitual en los registros, lo que dificulta el diseño de algoritmos de decodificación altamente efectivos. Estos algoritmos son esenciales para el control de dispositivos en tareas de neurorrehabilitación, ya que activan la corteza motora del paciente y contribuyen a su recuperación. Debido a estas razones, este estudio plantea un algoritmo de decodificación de MI en señales EEG fundamentado en características espaciales, y una mejora del mismo basada en técnicas de aprendizaje profundo o Deep Learning. Estos métodos han propiciado un acierto promedio del 72.5% y de hasta un 85% en algunos sujetos sanos. Estos avances son prometedores para el control preciso de dispositivos en tareas de neurorrehabilitación y podrían tener un impacto positivo en la recuperación de pacientes.

1. Introducción

Las Interfaces Cerebro-Máquina (BMIs) permiten registrar la actividad neuronal del cerebro del usuario y emplearla como elemento de control de dispositivos [1]. Se pueden utilizar métodos invasivos, como la electrocorticografía (ECoG) [2], o no invasivos, como las señales electroencefalográficas (EEG) [3-4] o la Imagen por Resonancia Magnética Funcional (fMRI, *functional Magnetic Resonance Image*) [5].

En el campo de la rehabilitación motora, las BMIs desempeñan un papel crucial al detectar la intención de movimiento del paciente [6], lo que facilita el control de dispositivos clínicos, como exoesqueletos de miembros inferiores [7] o cicloergómetros [8-9]. Los primeros aportan una experiencia más cercana a la marcha natural del individuo, mientras que los segundos son más seguros, fáciles de usar y considerablemente más baratos. La Imaginación Motora (MI) es una técnica comúnmente utilizada en estas BMIs para activar el área motora cerebral

y mejorar los resultados de la rehabilitación. Esto es debido principalmente a dos razones. La primera es que ciertos estudios sugieren que la corteza motora humana se haya particularmente activa durante la marcha [9]. La segunda es que una mayor inmersión del paciente en sus tareas puede propiciar una activación superior de esta área motora afectada, contribuyendo no solo a su terapia muscular mediante el movimiento real de sus miembros (apoyado por un cicloergómetro, un exoesqueleto, etc.), sino también a su rehabilitación neuronal. Esta mayor implicación mental se propicia al aplicar un control basado en MI.

Los enfoques para decodificar la MI de las señales EEG suelen implicar la extracción de características discriminantes, ya sean temporales, espaciales o frecuenciales, y alimentar un clasificador con ellas [10]. También se han explorado las redes neuronales convolucionales basadas en aprendizaje profundo o *Deep Learning* (DL) como posibles soluciones, pero todavía se necesita más evidencia para determinar si superan a los métodos basados en características [11].

En el caso específico de la decodificación de MI a partir de EEG, se ha observado que las características espaciales basadas en el algoritmo de Patrones Espaciales Comunes (*Common Spatial Patterns*, CSP) han dado buenos resultados [4]. Además, la división por bandas de frecuencia ha demostrado ser efectiva para mejorar los modelos decodificadores de MI [4]. Aún más, en el caso de [4] parece demostrarse que estas características casan especialmente bien con algoritmos de clasificación lineales bayesianos como el Análisis Discriminante Lineal o LDA (*Linear Discriminant Analysis*). Por su parte, existen también algunas redes neuronales optimizadas para la decodificación de MI en señales EEG, como [12], basada en características espaciales, o [13], la cual aplica el concepto de banco de filtros sumado a un análisis basado en características espaciales.

Por todas estas razones, en este estudio se ha comparado el rendimiento del método basado en clasificación de características espaciales desarrollado en [14] con el que ofrece otra aproximación basada en DL, desarrollada bajo el marco de la red neuronal IFNet [13].

2. Materiales y métodos.

2.1. Sujetos

Para este caso de estudio han sido empleado los datos de cuatro sujetos sanos, dos mujeres, U01 y U04, y dos hombres, U02 y U03, todos sin ningún tipo de disfunción motora ni neuronal diagnosticada, comprendidos entre los 18 y los 51 años (32 ± 14). Todos los sujetos recibieron información clara sobre el estudio, y firmaron el consentimiento informado planteado en el comité de ética aprobado por el Comité de Ética de la Investigación con medicamentos del Hospital Universitario Severo Ochoa de Leganés (Madrid) con código HLM-CYCLING-EEG.

2.2. Protocolo experimental

El protocolo fue diseñado de manera estratégica para entrenar dos tipos de control: el control en reposo (con el cicloergómetro inactivo) y el control en movimiento (con el cicloergómetro activo). Esto implicó tener fases de relajación (sin Imaginación Motora, MI) y fases de MI, tanto en reposo como durante el pedaleo (con el dispositivo asistiendo al usuario como en una terapia real).

Cada sesión de pruebas consistió en 22 repeticiones, donde cada repetición incluía 15 segundos de relajación, seguidos de 30 segundos de MI durante el pedaleo, y luego otros 15 segundos de relajación. De las 22 repeticiones, las 11 impares se realizaron con el cicloergómetro inactivo, y las 11 pares con el cicloergómetro activo en modo pasivo. De esta manera, se obtuvieron datos sobre los comandos de inicio (MI en reposo) y parada (relajación en movimiento) del dispositivo rehabilitador en una posible terapia. La Figura 1 muestra un esquema del protocolo empleado.

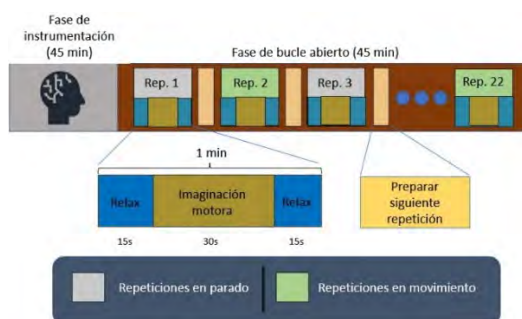


Figura 1. Protocolo experimental. Cada usuario realizó cuatro sesiones de este protocolo.

Se eligió un total de 22 repeticiones para poder descartar la peor de cada paradigma y quedarse con un número más redondo, es decir, 10 de cada tipo de repetición; y para poder ejecutar una validación cruzada 10+1, generando así 11 modelos decodificadores distintos mediante el método I (apartado 2.4) por paradigma y sesión de pruebas y pudiendo descartar la repetición de esas 11 usada como test que ha desencadenado en los peores resultados. Cada usuario realizó cuatro sesiones de prueba, esperando obtener una cantidad de datos aceptable para poder entrenar los modelos convolucionales.

Previamente a la primera sesión de cada usuario, se les dio unas directrices sobre cómo realizar la imaginación motora, tratando de guiarles hacia una imaginación cinestésica y homogénea a lo largo de las sesiones de

pruebas. Además, completaron el Movement Imagery Questionnaire-3 en su versión en castellano [15]. U01 obtuvo un 7.0 en el índice para MI cinestésica, mientras que U02 obtuvo un 4.8, U03 un 5.3 y U04 un 4.5.

2.3. Equipos empleados

El equipo de registro de EEG usado en las sesiones experimentales de este proyecto está formado por un gorro g.NAUTILUSPRO Flexible de 32 canales con la distribución de electrodos g.SCARABEO; junto con el emisor Wi-Fi HEADSET y el receptor BASE STATION, todo el conjunto del fabricante g.tec medical Engineering GmbH. La referencia fue colocada mediante pinza en el lóbulo de la oreja izquierda, previamente limpiado con gel dermoabrasivo, y todos los electrodos fueron acompañados de gel conductor hasta lograr un nivel de impedancia adecuado para el registro que genere una calidad de señal limpia ($\leq 15k\Omega$). Las señales fueron registradas a 500Hz.

Respecto al cicloergómetro, el modelo empleado es el CycleMotus™ A4, de Fourier Intelligence ©. La velocidad de giro fue ajustada al gusto de cada usuario, ya que se desean plantear modelos robustos adaptables a las necesidades de cada paciente de cara a una posible futura implementación en terapia real. El sujeto U01 empleó una velocidad de giro de 30 revoluciones por minuto, mientras que los usuarios U02 y U04 emplearon velocidad de 20 revoluciones por minuto, y U03 de 25 revoluciones.

La Figura 2 muestra el montaje experimental.



Figura 2. Montaje experimental. Se tapó la pantalla del cicloergómetro para evitar estímulos visuales externos en los usuarios.

2.4. Método de decodificación de MI I: CSP+LDA

▪ Preprocesamiento

Para este primer método de análisis, se aplicó una serie de filtros hardware a todos los canales durante la adquisición de las señales, implementados por el fabricante. Esta está formada por el algoritmo de reducción de ruido que incorpora el dispositivo, un filtro Notch de orden 4 entre 48Hz y 52Hz para eliminar el componente de 50Hz de la alimentación eléctrica, y un filtro pasa-banda de orden 8 con frecuencias de corte de 0.5Hz y de 100Hz, para acotar la señal al rango de frecuencias deseado. Tras ello, se ha realizado una preselección de canales del área motora, formada por 19 electrodos, los cuales, de acuerdo con el estándar 10/20, son: FZ, F4, F3, FC5, FC1, FC2, FC6, T7, C3, CZ, C4, T8, CP5, CP1, CP2, CP6, P3, PZ y P4 (Figura 3).

A continuación, se aplicó una división por bandas de frecuencia para el filtrado, formada por 10 bandas diferenciadas: 2-5Hz, 5-10Hz, 10-15Hz, 15-20Hz, 20-25Hz, 25-35Hz, 35-40Hz, 40-45Hz, 45-50Hz, 50-60Hz. Para cada uno de los canales seleccionados, se filtró la señal paralelamente mediante filtros Butterworth de orden 2 en cada uno de esos rangos, sumando todas las señales obtenidas para cada canal al final y obteniendo así la señal para el análisis, emulando el método empleado en [13], como ilustra la Figura 4a.

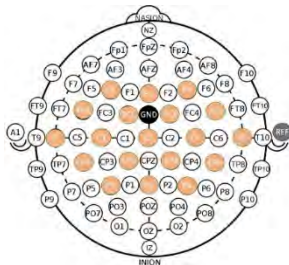


Figura 3. Selección de electrodos. La tierra (GND, Ground) va situada en FCZ.

▪ **Extracción de características**

Esta parte del análisis se ha fundamentado en el filtro CSP. Se ha aplicado para cada uno de los 19 canales seleccionados ya preprocesados, aplicando previamente una división en ventanas temporales, siendo las ventanas de dos segundos de duración y con un solape de segundo y medio, lo que significa que se obtienen características cada medio segundo.

▪ **Clasificación**

Esta se ha llevado a cabo mediante el algoritmo LDA, alimentándolo con las características obtenidas en el paso previo. Este algoritmo devuelve una salida (MI o relax) cada medio segundo, la cual puede compararse con la etiqueta real para comprobar si ha acertado. Se ha aplicado una validación cruzada 10+1 para obtener datos de desviación estándar, entrenando el modelo con 10 registros y testeando con el restante (tanto para pedaleo como para estático) en 11 ocasiones distintas por paradigma y sesión de pruebas, pudiendo probar así 11 modelos distintos.

2.5. Método de decodificación de MI II: IFNet

▪ **Preprocesamiento**

La adquisición de las señales se ha realizado de la misma forma que con el método I, aplicando los mismos filtros hardware y la misma preselección de canales (Figura 3); son los mismos registros. Respecto al banco de filtros, se ha efectuado uno diferente para este modelo, más enfocado al análisis en tiempo real. En este caso, ha consistido en un filtro online triple Butterworth de orden 8 en el espacio de estados, filtrando ventanas de un segundo en serie, paralelamente en estas tres bandas de frecuencia: 2-25Hz, 25-50Hz, y 50-75Hz. Tras el filtrado de cada ventana, se suma el resultado de las tres ramas para obtener la señal final para el análisis (Figura 4b).

▪ **Clasificación**

En este método, se ha efectuado la misma división en ventanas que en el método I (ventanas de dos segundos con

segundo y uno y medio de solape, una salida del clasificador cada medio segundo). Con esta división, se ha alimentado la red IFNet [13], generando un modelo por usuario, dividiendo todos sus datos ya preprocesados en 80% para entrenamiento, 15% para validación y 5% para test. A nivel funcional y de hiperparámetros, no se efectuó ningún cambio en la red por defecto, más allá de ajustar los filtros para que coincidan con los usados en el preprocesamiento. Se ha eliminado el peor trial de cada sesión de registros (quedándose con 10 en estático y 10 durante el pedaleo con sesión), elegido aplicando el método I.

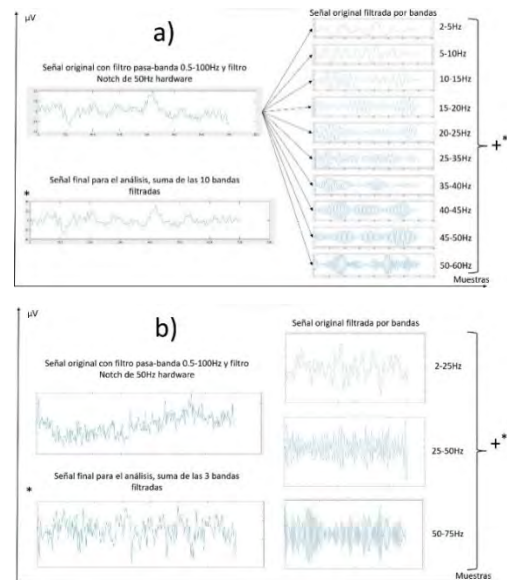


Figura 4. Bancos de filtros empleados en el preprocesamiento. a) se empleó en el método I, b) en el método II.

3. Resultados y discusión

La Tabla 1 muestra un resumen de los valores de acierto obtenidos con el método I. La Tabla 2, los obtenidos con el método II de decodificación. Salta rápidamente a la vista la superioridad en el acierto lograda por la red neuronal, casi un 12% de media. Pese a ello, el relax funciona mejor aplicando la extracción de características convencional en dos de los cuatro usuarios. Sí que se aprecia una importante superioridad del usuario U01 en ambos casos, y las tendencias parecen tener relación directa con los índices de MI cinestésicos obtenidos en el cuestionario MIQ-3 [15].

Tabla 1. Medidas de acierto obtenidas en la decodificación de MI mediante el método I (CSP+LDA).

USUARIO		U01	U02	U03	U04
ACIERTO [%]	Relax	72.0	51.6	66.9	57.7
	MI	74.9	47.0	67.8	56.6
	Total	73.5	49.3	67.4	57.2

Tabla 2. Medidas de acierto obtenidas en la decodificación de MI mediante el método II (IFNet).

USUARIO		U01	U02	U03	U04
ACIERTO [%]	Relax	100.0	38.3	55.5	60.1
	MI	79.1	74.7	78.9	74.1
	Total	85.4	66.3	70.9	67.1

Como gran pega a los modelos generados por la red neuronal, se debe comentar el alto desbalanceo existente en los resultados de acierto entre clases. Los modelos parecen polarizarse hacia una de las clases, obteniendo un acierto bastante más amplio en ella, disminuyendo así el acierto en la otra clase.

4. Conclusiones

El DL se está consolidando como respuesta a la mayoría de los retos que propone la computación moderna, y la decodificación de MI a partir de señales EEG no parece ser una excepción. Este estudio deja constancia de las mejoras que ofrece frente a métodos de decodificación convencionales, mejorando los resultados casi un 12% y alcanzando índices acordes a la literatura previa [4], [8], [10] con un acierto medio superior al 72%, mejorando también la investigación previa [14]. No obstante, se trata de un estudio preliminar con pocos usuarios y será necesaria más investigación para poder llegar a conclusiones más sólidas. El problema del desbalanceo entre clases puede ser la causa de que el relax mejore en algunos usuarios al emplear CSP+LDA, y este es un problema que puede abordarse asegurando un buen ajuste de los parámetros de la red neuronal y probando técnicas de entrenamiento inter-sujetos (*Transfer Learning*).

También parece clara la relación entre la calidad de la ejecución de la MI de los usuarios, y los resultados que obtienen los modelos al decodificarla. Esto refuerza la adecuación de efectuar una preparación o entrenamiento previo de los usuarios en la evocación de MI.

Agradecimientos

Esta investigación ha sido llevada a cabo en el Center for Clinical Neuroscience HLM del Hospital Los Madroños, Madrid, y ha sido financiada por una ayuda para la formación predoctoral en colaboración con empresas – modalidad A: incentivos para contratos predoctorales suscritos por la Universidad Miguel Hernández de Elche y financiados por la Generalitat Valenciana para la elaboración de una tesis con mención industrial.

Referencias

- [1] M. W. Slutzky, “Brain-Machine Interfaces: Powerful Tools for Clinical Treatment and Neuroscientific Investigations” *Neuroscientist*, vol. 25, no. 2, pp. 139–154, Apr. 2019, doi: 10.1177/1073858418775355.
- [2] E. C. Leuthardt, G. Schalk, D. Moran, and J. G. Ojemann, “The emerging world of motor neuroprosthetics: a neurosurgical perspective” *Neurosurgery*, vol. 59, no. 1, pp. 1–13, Jul. 2006, doi: 10.1227/01.NEU.0000221506.06947.AC.
- [3] G. Garipelli, R. Chavarriaga, and J. Del R Millán, “Single trial analysis of slow cortical potentials: a study on anticipation related potentials” *J. Neural Eng.*, vol. 10, no. 3, Jun. 2013, doi: 10.1088/1741-2560/10/3/036014.
- [4] M. Ortiz, L. Ferrero, E. Iáñez, J. M. Azorín, and J. L. Contreras-Vidal, “Sensory Integration in Human Movement: A New Brain-Machine Interface Based on Gamma Band and Attention Level for Controlling a Lower-Limb Exoskeleton” *Front. Bioeng. Biotechnol.*, vol. 0, p. 735, Sep. 2020, doi: 10.3389/FBIOE.2020.00735.
- [5] M. Misaki *et al.*, “Beyond synchrony: The capacity of fMRI hyperscanning for the study of human social interaction” *Soc. Cogn. Affect. Neurosci.*, vol. 16, no. 1–2, 2021, doi: 10.1093/scan/nsaa143.
- [6] J. N. Mak and J. R. Wolpaw, “Clinical Applications of Brain-Computer Interfaces: Current State and Future Prospects” *IEEE Rev. Biomed. Eng.*, vol. 2, p. 187, 2009, doi: 10.1109/RBME.2009.2035356.
- [7] J. Contreras-Vidal *et al.*, “Powered exoskeletons for bipedal locomotion after spinal cord injury” *J. Neural Eng.*, vol. 13, no. 3, Apr. 2016, doi: 10.1088/1741-2560/13/3/031001.
- [8] M. Ortiz, M. Rodríguez-Ugarte, E. Iáñez, and J. M. Azorín, “Analysis of algorithms for detection of pedaling intention in brain-machine interfaces” *RIAI - Rev. Iberoam. Autom. e Inform. Ind.*, vol. 16, no. 2, pp. 222–231, 2019, doi: 10.4995/RIAI.2018.9861.
- [9] T. Castermans and M. Duvinage, “Corticomuscular coherence revealed during treadmill walking: further evidence of supraspinal control in human locomotion” *J. Physiol.*, vol. 591, no. 6, pp. 1407–1408, Mar. 2013, doi: 10.1113/JPHYSIOL.2012.247593.
- [10] F. Lotte *et al.*, “A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update” *J. Neural Eng.*, vol. 15, no. 3, p. 031005, Apr. 2018, doi: 10.1088/1741-2552/AAB2F2.
- [11] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, “EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces” *J. Neural Eng.*, vol. 15, no. 5, 2018, doi: 10.1088/1741-2552/aace8c.
- [12] P. Authasan *et al.*, “MIN2Net: End-to-End Multi-Task Learning for Subject-Independent Motor Imagery EEG Classification” *IEEE Trans. Biomed. Eng.*, vol. 69, no. 6, 2022, doi: 10.1109/TBME.2021.3137184.
- [13] J. Wang, L. Yao, and Y. Wang, “IFNet: An Interactive Frequency Convolutional Neural Network for Enhancing Motor Imagery Decoding From EEG” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1900–1911, 2023, doi: 10.1109/TNSRE.2023.3257319.
- [14] J. V. Juan, E. Iáñez, M. Ortiz, J. Tornero, and J. M. Azorín, “Análisis de la Imaginación Motora durante pedaleo a partir de señales EEG” *Congr. Anu. la Soc. Española la Ing. Biomédica, CASEIB, 2022*, no. 8.5.2017, pp. 2003–2005, 2022.
- [15] S. Trapero-Asenjo, T. Gallego-Izquierdo, D. Pecos-Martín, and S. Nunez-Nagy, “Translation, cultural adaptation, and validation of the Spanish version of the Movement Imagery Questionnaire-3 (MIQ-3)” *Musculoskelet. Sci. Pract.*, vol. 51, 2021.

Evaluación de una interfaz cerebro-ordenador basada en potenciales relacionados con eventos para la detección de estímulos visuales en una tarea de vigilancia

Á. Fernández-Rodríguez¹, R. Ron-Angevin¹, F. Velasco-Álvarez¹, J. Díaz-Pineda², T. Letouzé³, J.-M. André³

¹ Departamento de Tecnología Electrónica, Instituto Universitario de Investigación en Telecomunicación de la Universidad de Málaga (TELMA), Universidad de Málaga, 29071 Malaga, Spain, afernandezrguez@uma.es, raron@uma.es, fvelasco@dte.uma.es

² Thales Avionics France, Bordeaux, France, jaimediazpineda@fr.thalesgroup.com

³ Bordeaux University, INP Bordeaux-ENSC, Laboratoire IMS - UMR CNRS 5218, 33400 Talence, France, tletouze@ensc.fr, jean-marc.andre@ensc.fr

Resumen

Las interfaces cerebro-ordenador (BCI) permiten establecer un canal de comunicación entre un usuario y un dispositivo a través de su actividad cerebral. El presente trabajo explora el uso de una BCI basada en potenciales relacionados con eventos (ERP) para la detección automática de nuevos estímulos durante una tarea de vigilancia. La tarea de vigilancia consistió en detectar la presencia de nuevos aviones en un mapa. Asimismo, se estudió el impacto de las siguientes variables sobre el rendimiento: el tamaño de la superficie a vigilar, y la saliencia de la aparición de nuevos aviones manipulada a través del color. Para ello, 10 participantes controlaron tres condiciones: aparición de un avión rojo, diferente a los amarillos ya presentes, en un área pequeña a vigilar (RP); aparición de un avión amarillo, similar a los ya presentes, en un área pequeña a vigilar (AP); y aparición de un avión rojo, diferente a los amarillos ya presentes, en un área grande a vigilar (RG). La accuracy promedio de cada condición fue la siguiente: RP, 64.5%; AP, 67.5%; RG, 41%. Los resultados mostraron que el tamaño de la superficie a vigilar tenía un impacto negativo en el rendimiento del sistema, pero no la saliencia manipulada a través del color. Por tanto, futuros estudios podrían explorar cómo solucionar el problema del área a vigilar, o estudiar qué otras variables medien en el rendimiento de una BCI para la detección de estímulos en tareas de vigilancia.

1. Introducción

Las interfaces cerebro-ordenador (BCI, por sus siglas en inglés) basadas en la electroencefalografía (EEG) son un tipo de tecnología que emplea la actividad eléctrica del cerebro para establecer un canal de comunicación entre un usuario y un dispositivo [1]. Estudios previos han demostrado cómo las BCI pueden utilizarse de manera efectiva para ayudar en la toma de decisiones o monitorear el estado de un usuario durante una tarea de vigilancia. Un controlador de tráfico aéreo (CTA) es un escenario en el que un operador profesional dirige aviones en una sección determinada del espacio aéreo. Por lo tanto, un CTA representa una situación en la cual un usuario debe estar atento a diferentes señales y reaccionar en consecuencia [2]. Este trabajo se centra en el uso de BCI para su aplicación a un CTA.

El objetivo de una BCI adaptada a un CTA debe ser mejorar la seguridad y precisión del sistema que se está

controlando. En general, se pueden distinguir dos tipos de sistemas BCI para cumplir estos objetivos: pasivos y activos. Por un lado, una BCI pasiva tiene como finalidad reconocer el estado cognitivo del usuario durante la ejecución de una tarea (p. ej., cansancio) con la finalidad de ayudar a la prevención de incidentes [3, 4]. Por otro lado, una BCI activa tendría como objetivo asistir en la toma de decisiones a través, de por ejemplo, la detección de elementos tales como un mensaje de alerta. Hasta donde sabemos, no se ha realizado ningún trabajo que haya utilizado una BCI activa para la detección de nuevos elementos en un CTA. Por lo tanto, el presente estudio se centra en las BCI activas.

El estímulo visual a atender por los usuarios será la aparición de nuevos aviones en un mapa. Por tanto, el presente trabajo usará los potenciales relacionados con eventos (ERP, por sus siglas en inglés) visuales como señal de control EEG para la BCI. Los ERP visuales son cambios de potencial en la actividad eléctrica del cerebro provocados por la presentación de estímulos visuales [5]. Por lo tanto, el objetivo de una BCI basada en ERP (BCI-ERP) es detectar el estímulo atendido basándose en la señal EEG del usuario. Hay varias diferencias entre las aplicaciones comúnmente controladas por una BCI-ERP (p. ej., un teclado virtual [6]), y un CTA. Dos de las diferencias más relevantes son: (i) el número de presentaciones del estímulo objetivo y (ii) la ubicación de la aparición del estímulo objetivo. Por un lado, en la mayoría de las aplicaciones de BCI-ERP, los estímulos objetivo se presentan varias veces para maximizar la probabilidad de que se seleccionen correctamente. Sin embargo, para un CTA y, en general, para cualquier aplicación en la que se presenten mensajes de alerta, es importante que el estímulo objetivo pueda ser reconocido con solo una presentación. Esto supone un desafío, ya que las BCI-ERP generalmente necesitan varias presentaciones del estímulo para obtener un rendimiento satisfactorio. Algunas propuestas previas de BCI-ERP enfocadas en el uso de presentación única han mostrado un rendimiento adecuado (*accuracy* del ~80% [7, 8, 9]). No obstante, estos trabajos no presentaron las características que podrían dificultar el rendimiento de un CTA, tales como el uso de un mapa cargado de información

visual, aviones en movimiento, o estímulos objetivo de tamaño reducido. Por otro lado, en la mayoría de las BCI-ERP visuales, los estímulos objetivo suelen presentarse en una ubicación específica previamente conocida por el usuario. Por el contrario, en un CTA, los nuevos aviones pueden aparecer en una ubicación aleatoria. Por lo tanto, sería interesante estudiar si el tamaño de la superficie de aparición (área a vigilar) influye en el rendimiento de la BCI-ERP.

El presente trabajo tiene como objetivo evaluar el efecto de diferentes variables visuales en el rendimiento de una BCI-ERP visual durante una tarea de vigilancia en un CTA. Específicamente, se explorará el efecto en el rendimiento de la detección de nuevos aviones con (i) el uso de un color diferente (o similar) a los demás aviones ya presentes en la interfaz y (ii) el tamaño del área que debe ser vigilada por el usuario.

2. Método

2.1. Participantes

El experimento involucró a 10 participantes (24.44 ± 2.01 años, seis hombres y cuatro mujeres). La experiencia de los participantes en el control del BCI-ERP fue variable. Todos ellos dieron su consentimiento para el uso anónimo de sus datos. Asimismo, declararon tener una visión normal o corregida. El estudio fue aprobado por el Comité Ético de la Universidad de Málaga y cumplió con los estándares de la Declaración de Helsinki.

2.2. Adquisición de datos y procesamiento de la señal

Las señales se registraron a través de ocho electrodos activos siguiendo el sistema internacional 10/10: Fz, Cz, Pz, Oz, P3, P4, PO7 y PO8. El electrodo de referencia fue colocado en el mastoide izquierdo, y el electrodo de tierra en AFz. Se utilizó un amplificador actiCHamp (Brain Products GmbH, Gilching, Alemania), con una frecuencia de muestreo de 250 Hz. Los datos se recopilaron con BCI2000 [10]. Para reducir el impacto del ruido del EEG, se aplicó un filtro paso banda entre 0.1 y 30 Hz, y un filtro *notch* a 50 Hz.

Para permitir el control *online* de la BCI-ERP, era necesario realizar una calibración que determinase los parámetros específicos del usuario. Para lograr esto, se analizó el EEG de cada uno de ellos utilizando el análisis discriminante lineal paso a paso (SWLDA, por sus siglas en inglés) con la herramienta P300Classifier de BCI2000. El intervalo de tiempo analizado – es decir, la longitud de la época – se estableció en el valor predefinido de 0-800 ms después de la presentación del estímulo.

Tanto el procesado como la sesión experimental fueron llevados a cabo en un portátil HP Envy 15-j100 (2.20 GHz, 16 GB, Windows 10), con una pantalla Acer P224W de 46.47×31.08 cm (relación de aspecto 16:10), conectada a través de HDMI, con una resolución de 1680×1050 píxeles. La frecuencia de actualización de la pantalla fue de 60.014 Hz. La distancia entre el punto de vista del usuario y la pantalla fue de aproximadamente 60 cm.

2.3. Condiciones experimentales

El paradigma consistió en presentar un vídeo de fondo (extraído de <https://www.fliht radar24.com>, consultado el 10 de julio de 2023) sobre el cual aparecía el estímulo objetivo al que el usuario debía prestar atención. El vídeo de fondo mostraba un mapa que cubría toda la pantalla y donde se podían ver diferentes aviones amarillos desplazándose. El lugar elegido para la grabación del vídeo fue un área de 129×80 km que contenía dos aeropuertos de París (Charles de Gaulle y Paris-Orly) alrededor de las 11:00-12:00 hora local.

El estímulo objetivo al que el usuario debía prestar atención era la aparición de nuevos aviones en una ubicación aleatoria. El área completa donde podían aparecer los aviones era de 42.93×27.74 cm (es decir, toda la pantalla excluyendo 1.67 cm de márgenes horizontales y 1.77 cm de márgenes verticales). Posteriormente, este rectángulo interno se dividió en nueve celdas, formando una matriz de 3×3 , de 14.31×9.25 cm cada una. Dependiendo de la condición experimental, el usuario era informado (o no) sobre en qué celda aparecería el próximo avión objetivo; dentro de esta celda, la posición del avión sería aleatoria. Después de que apareciera el avión objetivo, éste se movía en una dirección y velocidad constantes pero aleatorias, simulando a los aviones ya presentes en el vídeo. El tamaño de los aviones a los que debía prestar atención el participante era de 1×1 cm ($0.95^\circ \times 0.95^\circ$ a 60 cm). Específicamente, se probaron tres condiciones experimentales:

- Condición RP. El estímulo a atender en esta condición era la aparición de un avión de un color diferente (rojo) de los aviones que ya estaban en la pantalla (amarillos) y en un área relativamente pequeña e indicada (una de las nueve celdas posibles mencionadas anteriormente). Para indicar en qué celda específica aparecería el avión, se mostraba un rectángulo rojo semitransparente sobre esa celda (14.31×9.25 cm) durante 1 s antes del comienzo del ensayo. El avión se mostraba inicialmente en rojo pero cambiaba gradualmente a amarillo durante 2 s.
- Condición AP. Esta condición fue similar a RP; la única diferencia fue que los aviones objetivo se presentaron utilizando el mismo color de los aviones que ya estaban en la pantalla (amarillo).
- Condición AG. Esta condición fue similar a RP; la única diferencia fue que el rectángulo rojo semitransparente que indicaba la aparición de los aviones ocupaba el área completa correspondiente a las nueve celdas (42.93×27.74 cm).

2.4. Procedimiento

El experimento empleó un diseño intrasujeto, por lo que todos los usuarios pasaron por las tres condiciones. Cada condición consistió en dos tareas: (i) una tarea de calibración para adaptar el sistema al usuario y (ii) una tarea *online* en la que el sistema intentaba detectar la

aparición de estímulos objetivo. La principal diferencia entre las tareas de calibración y *online* fue que, en la tarea *online*, el usuario recibía retroalimentación sobre si el sistema había detectado correctamente el estímulo objetivo (figura 1). La tarea de calibración consistió en 32 selecciones repartidas en 4 bloques de 108 s cada uno. La tarea *online* consistió en 20 selecciones en un bloque de una duración de 263 s. Para cada selección, se elegía al azar la celda – de entre las nueve posibles – donde se mostraría el próximo avión objetivo. La ubicación exacta del avión dentro de la celda también era aleatoria.

El orden de las condiciones fue contrabalanceado entre los usuarios. La duración aproximada de la sesión fue de 60 minutos desde el momento en que el participante llegó al laboratorio hasta el final de las tareas. Antes del inicio de cada bloque, había un tiempo de espera de 5 s, después del cual comenzaban las diferentes selecciones. Además, al comienzo de cada selección, ya sea en la tarea de calibración u *online*, se presentaba un rectángulo rojo semitransparente (1 s), que indicaba el área de aparición del avión objetivo (restringido a una de las nueve celdas posibles en RP y AP, o la combinación de todas ellas en RG). Después de la aparición del rectángulo rojo, había una pausa de 2849 ms, tras la cual había un período de 4032 ms durante el cual podía aparecer el nuevo avión. Después de este período, había una pausa posterior a la selección de 5 s. En la tarea *online*, dentro de esta pausa posterior a la selección, se proporcionaba retroalimentación al usuario sobre si el sistema había detectado correctamente la aparición del nuevo avión. Esta retroalimentación se indicaba mediante una imagen, presentada durante 3 s, de un pulgar hacia arriba en verde si el sistema había detectado el nuevo avión, o un pulgar hacia abajo en rojo si el sistema no había detectado el nuevo avión. Tanto el rectángulo rojo semitransparente como las imágenes de retroalimentación iban acompañados de una señal de audio para facilitar la atención del participante a la tarea.

3. Resultados y discusión

La *accuracy* promedio obtenida para cada condición fue la siguiente: RP, $64.5 \pm 24.77\%$; AP, $67.5 \pm 10.69\%$; y RG, $41 \pm 20.25\%$. El ANOVA de medidas repetidas indicó que existían diferencias significativas entre las condiciones evaluadas ($F(2) = 9.368$, $p = 0.002$). Específicamente, se encontraron diferencias entre RP y RG ($p = 0.024$) y entre AP y RG ($p = 0.011$), pero no entre RP y AP ($p = 1$). Por lo tanto, los resultados relacionados con la *accuracy* demostraron que el tamaño del área de visualización tiene una influencia negativa en el rendimiento para detectar la aparición de nuevos aviones utilizando una BCI-ERP. Sin embargo, a pesar de lo esperado, parece que, en un área relativamente pequeña, el color del avión no tuvo un efecto significativo en la *accuracy* del sistema.

Una de las peculiaridades de un CTA o cualquier aplicación que requiera la detección rápida de un estímulo específico – como mensajes de advertencia o error – es que los estímulos deben detectarse con una sola presen-

tación. En comparación con los estudios previos de una BCI-ERP que emplease clasificación bajo presentación única [7, 8, 9], los resultados del presente trabajo son inferiores. Estos estudios previos presentaron una *accuracy* de alrededor del 80%; sin embargo, en el presente estudio, la mejor *accuracy* fue del $\sim 66\%$ (RP y AP, ambas condiciones con una superficie de estímulo reducida). No obstante, como se describió en la introducción, los estudios anteriores no presentaron las características propias de un CTA. Por ejemplo, los resultados del presente trabajo podrían verse afectados por factores relacionados con (i) el uso de un fondo complejo (el mapa), (ii) la presencia de estímulos distractores mientras se espera la aparición del estímulo objetivo (otros aviones que se mueven a través del escenario) o (iii) no conocer la posición exacta donde aparecería el estímulo objetivo. Otros factores que se deben considerar para su modificación, ya que se ha demostrado que afectan el rendimiento, son, por ejemplo, el tamaño del estímulo [11], su color [12] o luminosidad [13].

4. Conclusiones

El presente trabajo supone un acercamiento a la implementación de una BCI-ERP para la detección de nuevos aviones en el contexto de un CTA. Asimismo, se ha mostrado la importancia del tamaño del área de vigilancia en el control de estos escenarios. El rendimiento obtenido confirma el desafío que representa este tipo de aplicaciones para las BCI que empleen una única presentación del estímulo objetivo. Sin embargo, dado que la combinación de un CTA y una BCI activa es un área relativamente nueva, hay un considerable margen de acción para nuevas propuestas. Futuros trabajos pueden centrarse en cómo mejorar el rendimiento de estos sistemas a través de lo que se ha estudiado previamente en otros tipos de dispositivos BCI (p. ej., los teclados virtuales [6]). Algunas de estas propuestas de mejora incluyen aquellas relacionadas con los factores humanos [14], así como con diferentes técnicas de procesamiento y clasificación de señales [15]. Los sistemas BCI se han utilizado anteriormente en el campo de los CTA. Sin embargo, solo se han utilizado con el propósito de evaluar el estado cognitivo de los usuarios (p.ej., evaluación de la carga mental [4] o la presencia de estados de microsueño [3]). Por lo tanto, podría ser interesante que futuras propuestas utilizaran una BCI con el doble propósito de (i) medir el estado cognitivo del usuario y (ii) asistir la detección de estímulos en la interfaz.

Agradecimientos

Esta investigación forma parte del proyecto SICODIS (PID2021-127261OB-I00), que ha sido financiado conjuntamente por el Ministerio de Ciencia, Innovación y Universidades (MCIU); la Agencia Estatal de Investigación (AEI); el Fondo Europeo de Desarrollo Regional (FEDER) y la Universidad de Málaga (UMA). Asimismo, los autores quieren agradecer la colaboración de todos los participantes durante las pruebas.

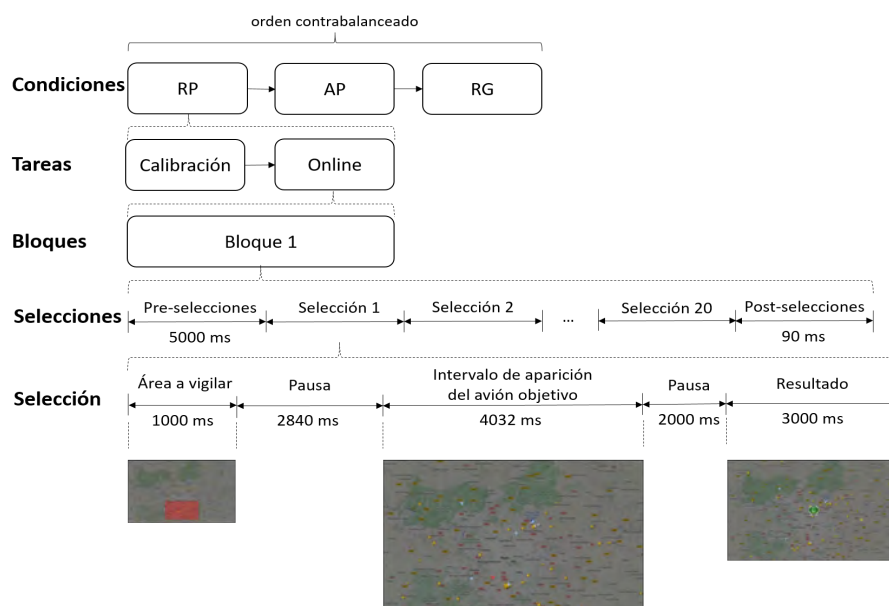


Figura 1. Procedimiento y temporización del experimento. Específicamente, la figura muestra la ejecución de la primera selección de la condición RP durante la tarea online.

Referencias

- [1] Jonathan R. Wolpaw, Niels Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, and Theresa M. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6):767–791, 2002.
- [2] Gianluca Di Flumeri, Francesca De Crescenzo, Bruno Berberian, Oliver Ohneiser, Jan Kramer, Pietro Aricò, Gianluca Borghini, Fabio Babiloni, Sara Bagassi, and Sergio Piastra. Brain-Computer Interface-Based Adaptive Automation to Prevent Out-Of-The-Loop Phenomenon in Air Traffic Controllers Dealing With Highly Automated Systems. *Frontiers in Human Neuroscience*, 13(September), 2019.
- [3] Linda Ng Boyle, Jon Tippin, Amit Paul, and Matthew Rizzo. Driver performance in the moments surrounding a microsleep. *Transportation Research Part F: Traffic Psychology and Behaviour*, 11(2):126–136, 2008.
- [4] Wenbin Li, Rong Li, Xiaoping Xie, and Yaoming Chang. Evaluating mental workload during multitasking in simulated flight. *Brain and Behavior*, 12(4):1–11, 2022.
- [5] Emily S. Kappenman and Steven J. Luck. *The Oxford Handbook of Event-Related Potential Components*. 2012.
- [6] Sarah S. Alrumiah, Leena A. Alhajjaj1, Jowharah F. Alshobaili, and Dina M. Ibrahim. A Review on Brain-Computer Interface Spellers: P300 Speller. *Biomedical Communication*, 13(3):1191–1199, 2020.
- [7] Hubert Cecotti and Anthony J. Ries. Best practice for single-trial detection of event-related potentials: Application to brain-computer interfaces. *International Journal of Psychophysiology*, 111:156–169, 2017.
- [8] Yin Tian, Huiling Zhang, Yu Pang, and Jinzhao Lin. Classification for Single-Trial N170 During Responding to Facial Picture With Emotion. *Frontiers in Computational Neuroscience*, 12(September), 2018.
- [9] Anahita Goljahani, Costanza D’Ávanzo, Stefano Silvoni, Paolo Tonin, Francesco Piccione, and Giovanni Sparacino. Preprocessing by a Bayesian single-trial event-related potential estimation technique allows feasibility of an assistive single-channel P300-based brain-computer interface. *Computational and Mathematical Methods in Medicine*, 2014, 2014.
- [10] Gerwin Schalk, Dennis J. McFarland, Thilo Hinterberger, Niels Birbaumer, and Jonathan R. Wolpaw. BCI2000: A general-purpose brain-computer interface (BCI) system, 2004.
- [11] Álvaro Fernández-Rodríguez, Aube Darves-Bornoz, Francisco Velasco-Álvarez, and Ricardo Ron-Angevin. Effect of Stimulus Size in a Visual ERP-Based BCI under RSVP. *Sensors*, 22(23), 2022.
- [12] D. B. Ryan, G. Townsend, N. A. Gates, K. Colwell, and E. W. Sellers. Evaluating brain-computer interface performance using color in the P300 checkerboard speller. *Clinical Neurophysiology*, 128(10):2050–2057, 2017.
- [13] Yueqing Li, Sangwoo Bahn, Chang S. Nam, and Jungnyun Lee. Effects of Luminosity Contrast and Stimulus Duration on User Performance and Preference in a P300-Based Brain-Computer Interface. *International Journal of Human-Computer Interaction*, 30(2):151–163, 2014.
- [14] Andrea Kübler, Elisa M. Holz, Angela Riccio, Claudia Zickler, Tobias Kaufmann, Sonja C. Kleih, Pit Staiger-Sälzer, Lorenzo Desideri, Evert Jan Hoogerwerf, and Donatella Mattia. The user-centered design as novel perspective for evaluating the usability of BCI-controlled applications. *PLoS ONE*, 9(12):1–22, 2014.
- [15] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger. A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update. *Journal of Neural Engineering*, 15(3), 2018.

Propuesta de BCI basado en RSVP con número reducido de estímulos usando el modo de escritura T9

F. J. Vizcaíno-Martín, F. Velasco-Álvarez, Á. Fernández-Rodríguez, R. Ron Angevin

Dpto. de Tecnología Electrónica, Universidad de Málaga, Málaga, España, {fjvizcaino, fvelasco, afernandezrguez, rron}@uma.es

Resumen

Se propone un sistema de comunicación para personas con severas deficiencias motoras, como pacientes afectados de esclerosis lateral amiotrófica, usando una interfaz cerebro computadora. El sistema se basa en el paradigma Rapid Serial Visual Presentation que tiene la ventaja de no requerir control oculomotor. Para reducir el tiempo de escritura, hace uso de solo 12 estímulos, es decir, aproximadamente la mitad de los estímulos que usan sistemas para escritura con interfaces cerebro-computadora tradicionales. De forma similar a los sistemas de escritura con teclados predictivos T9, cada estímulo se asoció con varios caracteres por lo que el sistema es soportado por un programa de análisis de palabras para componer la palabra final deseada de entre las posibles alternativas en función de los estímulos seleccionados. El sistema propuesto se ha comparado con un sistema de escritura tradicional de 30 estímulos usando el mismo paradigma de presentación de estímulos. Los resultados obtenidos, aunque con pruebas realizadas sobre un solo usuario sano y, por tanto, muy preliminar, han permitido comprobar la viabilidad del sistema, verificándose que, efectivamente, esa reducción del número de estímulos mejora las prestaciones del sistema de escritura llegándose a doblar su velocidad. Como contrapartida, el sistema propuesto requiere un entrenamiento para su uso al tener que agrupar diferentes caracteres con cada estímulo presentado, pero la propuesta de agrupación realizada permite pensar que el tiempo de aprendizaje será reducido.

1. Introducción

Personas con graves deficiencias en sus funciones motoras, como por ejemplo pacientes que sufren de esclerosis lateral amiotrófica (ELA), esclerosis múltiple, isquemia o parálisis cerebral, pueden tener enormes impedimentos para la comunicación con otras personas y con el entorno. Para estos casos, dotarles de un medio para comunicarse puede mejorar de forma significativa su apreciación de su calidad de vida [1]. Las interfaces cerebro-ordenador (BCI, de *brain-computer interface*), que usan la electroencefalografía (EEG) como técnica de registro de señales de entrada, se han postulado como el medio más versátil posible para establecer ese canal de comunicación [2]. Mediante estos sistemas y provocando la estimulación apropiada del usuario, se puede conseguir modular los potenciales de las señales EEG del cerebro, potenciales relacionados con eventos (ERP, de *event-related potential*) y, como consecuencia, codificar esas variaciones para controlar la respuesta del sistema [3].

De estos potenciales, destaca la componente P300, que ocurre en el cerebro humano como una desviación positiva

sobre el potencial base con una latencia de 300 ms después de que haya ocurrido un evento específico [4-5].

Una gran parte de los diferentes sistemas BCI para comunicación y control desarrollados usan la componente P300 de los ERP generada a través del paradigma *oddball*, presentación de estímulos al usuario de forma aleatoria entre los que el usuario debe prestar atención al deseado (objetivo), poco frecuente, en comparación con el total de estímulos presentados [5]. La principal diferencia entre ellos es la forma de presentar los estímulos al usuario, destacándose dos paradigmas: el paradigma fila-columna (RCP, de *row-column paradigm*) y el de presentación visual serial rápida (RSVP, de *rapid serial visual presentation*). En el primero, los estímulos se distribuyen de forma matricial ocupando la mayor parte de la pantalla de presentación y, en el segundo, los estímulos se presentan de uno en uno de forma consecutiva en el centro de la pantalla. Entre ellos, el paradigma RSVP es el más indicado para pacientes que tienen limitación a la hora de poder dirigir su mirada por tener reducida su capacidad de movimiento ocular [6]; su principal desventaja es el mayor tiempo requerido para seleccionar un carácter (estímulo) entre los presentados ya que, para cada selección, es necesario mostrar, uno detrás de otro, cada uno de los caracteres que forman el alfabeto.

Diferentes propuestas se han desarrollado para mejorar las prestaciones de los sistemas BCI para escritura. Dichas propuestas están enfocadas, principalmente, en diferentes modos de presentar los estímulos y en el número de estímulos empleados usando como base alguno de los dos paradigmas indicados [7]. Usando el paradigma RCP, en [8] se puso a prueba sobre un paciente con ELA el uso de una matriz reducida de 4×3 similar a la interfaz T9 desarrollada para teléfonos móviles. Hasta nuestro conocimiento, no se ha propuesto ningún sistema RSVP usando 12 estímulos seleccionables en un formato análogo al T9.

El objetivo del presente trabajo es el desarrollo de un teclado virtual basado en el paradigma RSVP con un número reducido de estímulos, 12, con el fin de aumentar la velocidad de escritura. La propuesta se comparará con otro teclado virtual convencional también basado en el paradigma RSVP y que incluirá 30 estímulos, similar al propuesto en [6], pero contemplando todos los caracteres alfabéticos del español.

2. Método

2.1. Participantes

La finalidad de este trabajo era comparar las prestaciones del sistema propuesto con un sistema de escritura RSVP que use todas las letras del alfabeto español. Este trabajo es un estudio preliminar en el que ha participado una persona sana (60 años, hombre, con experiencia previa en el manejo de sistemas BCI), permitiendo, de este modo, analizar la viabilidad de la propuesta presentada.

2.2. Adquisición de datos y procesamiento de la señal

Para la adquisición y amplificación de las señales del EEG se usó un amplificador acti-CHamp de BrainVision (Brain Products GmbH, Gilching, Alemania). Se usaron 9 electrodos para 8 canales en las posiciones Fz, Cz, Pz, Oz, P3, P4, PO7 y PO8, y el noveno conectado a tierra en la posición AFz, según el sistema internacional 10/10. Todos los canales se referenciaron al mastoide izquierdo. La frecuencia de muestreo fue de 250 Hz y las señales fueron procesadas mediante filtros banda eliminada de 50Hz, paso alto de 0.1 Hz y paso bajo de 9 Hz. Para el procesamiento de los datos del EEG se usó el software BCI2000 [9]. Dicho software empleó el análisis lineal discriminante de pasos sucesivos (SWLDA, de *stepwise linear discriminant analysis*) de los datos obtenidos a través del EEG para crear el clasificador que permitiría realizar la selección de caracteres en la fase online.

2.3. Implementación del sistema

El objetivo del sistema propuesto era dar la posibilidad a un paciente de suministrar a un ordenador las palabras o frases que desee que el ordenador, mediante una interfaz de audio, transmita a un posible destinatario, consiguiendo, por tanto, una comunicación “oral” entre el paciente y su interlocutor.

Para ello, se usó un ordenador HP Laptop con procesador Intel®Core™ i5-1035G1 CPU@1.00GHz 1.20 GHz, con RAM instalada: 16.6 GB. Sistema Operativo de 64 bits y Windows 11, conectado a través de HDMI con un monitor Lenovo C24-20, conectada a Intel®UHD Graphics 630, con resolución de 1920×1080 a 59.94 Hz y 8 bits de profundidad, donde se presentaron los estímulos que visualizaría el usuario.

En el ordenador se ejecutaban, por un lado, programas desarrollados en Python que permitían gestionar los caracteres seleccionados y sugerir, a partir de ellos, la palabra más probable, realimentando al usuario, a través de una interfaz de audio, los resultados obtenidos. Por otro lado, se ejecutaba UMA-BCI Speller [10] para la presentación de los estímulos según el paradigma RSVP y la llamada a BCI2000, para registrar y analizar las señales EEG generadas.

Las palabras más probables se obtienen usando el corpus de palabras en español de la Real Academia de la Lengua Española (RAE), CREA_total [11].

2.4. Paradigma de control

Para el proceso de escritura de una palabra o frase con el sistema, los usuarios tenían que seleccionar los caracteres

apropiados de entre los mostrados en la pantalla. Estos aparecen en la pantalla siguiendo un paradigma RSVP.

Siguiendo un modelo inspirado en el teclado T9, donde cada número lleva asociado una serie de caracteres alfabéticos que se pueden seleccionar pulsando repetidamente la tecla correspondiente, nueve de los caracteres del teclado del sistema propuesto presentaban triple funcionalidad. Así, el sistema propuesto, al que llamamos *RSVP-T9 Modificado (RT9M)*, usa 12 elementos o caracteres a seleccionar. Estos elementos son: “A”, “D”, “G”, “J”, “M”, “O”, “R”, “U”, “X”, “*”, “_” y “#” (figura 1). Los caracteres alfabéticos permiten al usuario seleccionar la terna de caracteres formados por el carácter mostrado y los dos siguientes en orden alfabético (“A”→ABC, “D”→DEF, “G”→GHI, “J”→JKL, “M”→MÑÑ, “O”→OPQ, “R”→RST, “U”→UVW, “X”→XYZ); por otro lado, los símbolos “*”, “_” y “#” dan la posibilidad al usuario de finalizar la escritura (“*”), añadir un espaciador de palabras en la frase (“_”) o borrar el carácter previo introducido, si se ha cometido error, o seleccionar la palabra correcta de entre las propuestas por los programas de análisis de palabras (“#”).

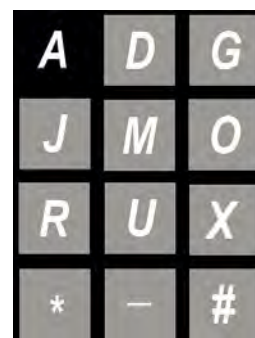


Figura 1: Representación matricial (4×3) de los 12 elementos que se usan en el sistema propuesto, RT9M.

Para seleccionar un elemento del conjunto, los usuarios tenían que prestar atención al elemento deseado (p. ej., el elemento “A” para seleccionar alguno de los caracteres A, B o C) según aparecía en pantalla y contar mentalmente el número de veces que aparece en cada bloque de selección, que estará determinado por el número de secuencias que se estimen necesarias después del proceso de calibración.

El paradigma RSVP se configuró para que cada carácter se presentara en el centro de la pantalla durante un tiempo de 192 ms y el tiempo entre estímulos (carácter) consecutivos fuera de 96 ms. Antes de cada secuencia, la pantalla permanece en negro por un tiempo de 3968 ms.

Para evaluar comparativamente las prestaciones del sistema propuesto, se desarrolló otro sistema similar en el que se contempla la presentación del alfabeto español completo, desde la A hasta la Z, incluyendo la Ñ, y los mismos tres símbolos de control de finalización de escritura (“*”), espaciado de palabras en la frase (“_”) y borrado (“#”), con un total de 30 elementos. A este sistema lo llamaremos *RSVP Tradicional Modificado (RTrM)*.

En ambos sistemas, los estímulos se presentaron al usuario en el mismo formato, tamaño y color: *Arial Bold* 300; 600×600 píxeles, fondo negro y carácter en blanco.

2.5. Procedimiento

Se habilitó una sala aislada con una mesa donde se instaló la instrumentación y equipos necesarios. El usuario se sentó cómodamente en una silla frente a la mesa a una distancia de 60 cm de la pantalla donde se presentaron los estímulos.

Antes de comenzar los experimentos se procedió a la instalación y configuración de la instrumentación y se colocó apropiadamente el gorro con sus electrodos en la cabeza del usuario, al que se le informó del proceso que debía seguir, que variaba entre la fase de calibración y la fase online, fases que se llevaron a cabo independientemente para cada sistema.

En la fase de calibración, similar para cada sistema, el usuario tenía que prestar atención a un total de 18 caracteres, agrupados en tres conjuntos de 6 caracteres: “AJR*_#”, “DMU_*#” y “GOX#_*”. Se emplearon 6 secuencias para cada grupo de caracteres, lo que quiere decir que cada carácter aparece en el centro de la pantalla 6 veces para cada grupo. Los tiempos empleados en la calibración se muestran en la tabla 1, donde se observa que el tiempo para la calibración en el sistema tradicional duplicó el tiempo empleado en la calibración del sistema propuesto. La exactitud indicada por el clasificador permitió seleccionar el número de secuencias que, posteriormente, se usaron en la fase online.

Sistema	Tiempo (minutos)
<i>RT9M</i>	9:30
<i>RTrM</i>	18:55

Tabla 1. Tiempo empleado en la calibración con cada sistema.

Durante la fase online, que se realizó en dos pasos, se solicitó al usuario, en el primer paso, que escribiera, tres textos cortos: dos palabras sueltas, “HOLA” y “AYUDA”, y una combinación de dos palabras, “POR FAVOR”. En un segundo paso, se solicitó la escritura de un texto largo: “EXPERIMENTO EN LA UNIVERSIDAD DE MÁLAGA”.

Para escribir un texto con cualquiera de los dos sistemas, se deben seleccionar los caracteres que componen el texto, deletreándolo convenientemente. En el sistema *RTrM*, el usuario únicamente debe saber deletrear las palabras usadas. En el sistema *RT9M*, el usuario debe saber deletrear las palabras y, además, conocer las letras asociadas a cada uno de los estímulos presentados por el sistema. Así, para poder escribir, por ejemplo, “HOLA” en el sistema *RT9M* se deben seleccionar los estímulos: “G” (GHI), “O” (OPQ), “J” (JKL), “A” (ABC), “_”, “#” y “*”. “G”, “O”, “J” y “A” corresponden a las selecciones a realizar para la palabra “HOLA”, “_”, para indicar fin de palabra o espaciado si hubiese más palabras en el texto a escribir. Una vez finalizada la palabra, el sistema propone, y se lo indica al usuario a través de la interfaz de audio, la palabra más probable listada en el corpus de palabras de la

RAE empleado. Sucede que, para esta palabra, la palabra más probable es “HOJA”, por lo que, para elegir la siguiente más probable habrá que seleccionar como siguiente carácter “#”. Finalmente habrá que seleccionar “*”, para indicar que se ha terminado la escritura del texto. Una vez finalizada la escritura del texto, el sistema transmitirá, a través de la interfaz de audio, el texto final escrito, “HOLA” en este caso, al usuario.

Para comparar los dos sistemas de escritura mediante RSVP sin que sea necesario que el usuario tenga un dominio de los diferentes caracteres que pueden escribirse con cada estímulo presentado en el sistema propuesto, se ha ideado que los programas de gestión de las palabras a escribir indiquen al usuario la selección de estímulos que debe hacer en cada momento. Así, estos programas desarrollados en Python, una vez conocida la palabra o frase que el usuario debiera escribir en la fase online, indicándosela al principio de su ejecución, le va diciendo al usuario, también mediante la interfaz de audio, qué estímulo debe seleccionar en cada momento, ya sea una letra o un símbolo de espaciado, fin de texto o borrado, ya que el sistema testea si el estímulo seleccionado es el correcto.

El experimento tuvo una duración total de alrededor de 103 minutos, con intervalos de descanso de 3 minutos entre calibración y cada paso de la fase online.

3. Resultados

En la fase de calibración el usuario alcanzó el 100% de *accuracy* con un máximo de 3 secuencias para los dos sistemas, por lo que se decidió realizar la fase online con solo tres secuencias por bloque.

En la fase online, el usuario no cometió ningún error de escritura con ninguno de los dos sistemas ni con ninguna de las tres secuencias de palabras en el primer paso de esta fase, empleándose los tiempos de escritura que se muestran en la tabla 2. Obsérvese que, para escribir la palabra “HOLA” en el sistema *RT9M*, hay que seleccionar 7 estímulos, ya que obliga a seleccionar la segunda palabra más probable propuesta por el sistema de análisis de palabras, por eso se tarda el mismo tiempo en escribir “HOLA” que “AYUDA”, para la que también hay que seleccionar 7 estímulos (cinco letras, “_” y “*”). Los tiempos obtenidos con el sistema *RT9M* están entre el 65.38% y el 54.9% de los tiempos obtenidos con el sistema *RTrM*.

Sistema	Tiempo (minutos)		
	“Hola”	“Ayuda”	“Por favor”
<i>RT9M</i>	2:16	2:16	3:30
<i>RTrM</i>	3:28	4:06	6:25

Tabla 2. Tiempos en minutos empleados por cada sistema en la escritura de los textos: “Hola”, “Ayuda” y “Por favor”.

En el segundo paso, relativo a la escritura de la frase larga (“EXPERIMENTO EN LA UNIVERSIDAD DE MÁLAGA”, 41 caracteres = 39 de la frase, “_” y “*”), el

usuario cometió 5 errores con el sistema *RTrM* y 2 errores con el sistema *RT9M*. La tabla 3 indica los tiempos empleados en la escritura de la frase por los dos sistemas, teniendo en cuenta que, con el sistema *RTrM* hubo que escribir 51 caracteres al cometer 5 errores y, con el sistema *RT9M*, solo se tuvieron que escribir 45 caracteres, ya que solo se cometieron 2 errores (dos caracteres adicionales por cada error, “#”, borrar, y otra vez el carácter deseado).

Sistema	Tiempo (minutos)
	“Experimento en la Universidad de MÁLAGA”
<i>RT9M</i>	2 errores → 13:55 minutos
<i>RTrM</i>	5 errores → 28:24 minutos

Tabla 3. *Tiempos en minutos empleados por cada sistema en la escritura de la frase larga: “Experimento en la Universidad de MÁLAGA”.*

En este segundo paso se observa que el tiempo empleado con el sistema *RT9M* es el 49% del empleado con el sistema *RTrM*.

La velocidad de escritura del sistema propuesto (*RT9M*) fue considerablemente más alta que la del sistema tradicional (*RTrM*). Asimismo, la ventaja del *RT9M* parece mayor conforme más caracteres se pretenden escribir, llegándose a reducir los tiempos de escritura a más de la mitad.

4. Discusión y conclusiones

En este trabajo se ha presentado un procedimiento que permite reducir el número de estímulos a mostrar al usuario para su selección en un método de escritura mediante BCI siguiendo el paradigma RSVP. Con ello, se pretendía reducir el tiempo de escritura en comparación con otros métodos disponibles. Aunque las pruebas se han limitado a un solo usuario, este primer análisis de resultados ha permitido observar que esta reducción de estímulos consigue un aumento en la velocidad de escritura de casi el doble con respecto al tradicional de 30 estímulos, donde se alcanzan ratios de 1.43 símbolos/minuto [6], frente a los 2.87 símbolos/minuto que se llegan a alcanzar con el procedimiento propuesto. Se observa también que la ventaja es mayor en la escritura de textos largos. Esto se puede deber a dos motivos principales, la fatiga creciente del usuario conforme usa el sistema y el hecho de que el sistema tradicional, al tener que presentar más estímulos, alargando el proceso de selección, es más fatigoso y cansado para el usuario y, por tanto, más propenso a errores. No obstante, el sistema propuesto tiene la contrapartida de que el usuario debe aprender a asociar los estímulos seleccionados con los posibles caracteres que puede escribir con los mismos. Por tanto, puede ser potencialmente más propenso a errores de escritura si no se invierte el tiempo de entrenamiento necesario que permita un manejo ágil del sistema. Asimismo, sería conveniente que futuros trabajos verificaran los resultados obtenidos con más usuarios, incluyendo los potenciales objetivos de este tipo de

sistemas: pacientes con severas limitaciones motoras. Estos estudios debieran evaluar profundamente la usabilidad de estos sistemas para ofrecer unos resultados más concluyentes sobre la apreciación que los usuarios tienen sobre ellos y sus posibilidades de uso.

Agradecimientos

Este trabajo ha sido financiado conjuntamente por el Ministerio de Ciencia, Innovación y Universidades, el Fondo Europeo de Desarrollo Regional (FEDER) a través del proyecto SICODIS, referencia: PID2021-127261OB-I00 y por la Universidad de Málaga.

Referencias

- [1] A. Londral, A. Pinto, S. Pinto, L. Azevedo and M. De Carvalho, "Quality of life in amyotrophic lateral sclerosis patients and caregivers. doi:10.1002/mus.24659. Epub 2015 Jun 30. PMID:25808635.
- [2] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, 2002 Jun; 113(6): 767-91. doi: 10.1016/s1388-2457(02)00057-3 PMID: 12048038.
- [3] A. Kübler, "The history of BCI: From a vision for the future to real support for personhood in people with locked-in syndrome", *Neuroethics* 13, 163-180 (2020), <https://doi.org/10.1007/s12152-019-09409-4>.
- [4] R.F. Helfrich and R.T. Knight, "Cognitive neurophysiology: Event Related potentials", *Handb Clin Neurol*, 2019; 160:543-558. doi: 10.1016/B978-0-444-64032-1.00036-9. PMID:31277875.
- [5] L.A. Farwell, E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials", *Electroencephalogr Clin Neurophysiol*. 1988 Dec;70(6): 510-23. doi: 10.1016/0013-4694(88)90149-6. PMID: 2461285.
- [6] L. Acqualagna, B. Blankertz, "Gaze-independent BCI-spelling using rapid serial visual presentation (RSVP)", *Clinical Neurophysiology*, Volume 124, Issue 5, 2013, Pages 901-908, ISSN 1388-2457, <https://doi.org/10.1016/j.clinph.2012.12.050>.
- [7] A. Rezeika, M. Benda, P. Stawicki, F. Gembler, A. Sabour and I. Volosyak (2018), "Brain-Computer Interface Spellers: A Review", *Brain Sciences*. 8. 57. 10.3390/brainsci8040057.
- [8] R. Ron-Angevin, S. Varona-Moya, L. da Silva-Sauer, "Initial test of a T9-like P300-based speller by an ALS patient", *J Neural Eng.* 2015 Aug;12(4):046023. doi: 10.1088/1741-2560/12/4/046023. Epub 2015 Jun 17. PMID: 26083683.
- [9] G. Schalk, D.J. McFarland, T. Hinterberger, N. Birbaumer and J.R. Wolpaw. 2004, "BCI2000: A general-purpose brain-computer interface (BCI) system", *IEEE Trans. Biomed. Eng.* 51 1034-43.
- [10] F. Velasco-Alvarez, S. Sancha-Ros, E. García-Garaluz, Á. Fernández-Rodríguez, M. T. T. Medina-Juliá, and R. Ron-Angevin, "UMA-BCI Speller: an Easily Configurable P300 Speller Tool for End Users," *Comput. Methods Programs Biomed.*, vol. 172, pp. 127–138, 2019.
- [11] Real Academia Española. Corpus de Referencia del Español Actual (CREA) – Listado de Frecuencias. Lista total de frecuencias. Recurso en línea disponible en: <http://corpus.rae.es/lfrecuencias.html>. Fecha de consulta: 13/07/2023.

Evaluación del Impacto del Aprendizaje Auto-Supervisado en la Precisión de Interfaces Cerebro-Ordenador basadas en Imaginación Motora

S. Pérez Velasco^{1,2}, D. Marcos-Martínez^{1,2}, E. Santamaría-Vázquez^{1,2}, R. Ruiz-Gávez¹, R. Hornero^{1,2}

¹ Grupo de Ingeniería Biomédica (GIB), Universidad de Valladolid, Valladolid, España, sergio.perezv@uva.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

Resumen

Las interfaces cerebro-ordenador (BCIs) buscan proporcionar vías de comunicación directas entre el cerebro humano y dispositivos externos. No obstante, la decodificación precisa de las intenciones del usuario es todavía un desafío, en parte por las limitaciones inherentes a la electroencefalografía (EEG), como son una resolución espacial limitada y relación señal-ruido reducida. Este estudio aborda estos desafíos mediante la aplicación de técnicas de aprendizaje auto-supervisado (SSL) en el preentrenamiento de una red basada en la arquitectura de transformer. Nuestra aproximación descompone la señal EEG en segmentos y utiliza un enmascarado y reconstrucción para obtener representaciones más robustas y efectivas.

Evaluamos el impacto de estas técnicas en la mejora de la clasificación de un sistema BCI basado en imaginación motora (MI) con una base de datos pública de 109 sujetos, utilizando un esquema de validación cruzada inter-sujeto k-fold (k=5). Comparamos tres escenarios diferentes: un modelo sin preentrenamiento SSL, un modelo con sondeo lineal de las características extraídas, y un modelo con fine tuning de toda la red. Nuestros resultados indican que el preentrenamiento con técnicas de SSL mejora significativamente la precisión de la clasificación de MI. Concretamente, la precisión se incrementa desde un 78.4%±12.2%, hasta un 79.9%±11.9% utilizando sondeo lineal, y alcanza un 82.4%±11.7% cuando se aplica fine tuning a toda la red. Este trabajo demuestra el potencial del SSL aplicado a redes basadas en transformer para avanzar en la interpretación de señales EEG.

1. Introducción

La Ingeniería Biomédica (IB) supone una sinergia entre medicina e ingeniería, orientada hacia el desarrollo de herramientas innovadoras para abordar enfermedades y trastornos humanos [1]. Este estudio se focaliza en una herramienta particularmente prometedora: la interfaz cerebro-ordenador (*brain-computer interface*, BCI) [2]. Estas interfaces ofrecen un camino de comunicación alternativo y directo entre el cerebro y el entorno [2]. Entre los métodos para captar la actividad cerebral, la electroencefalografía (EEG) destaca por su alta resolución temporal, facilidad de uso y coste relativamente bajo, convirtiéndola en una de las mejores elecciones de registro para los investigadores BCI [2]. No obstante, la EEG también presenta desafíos, como su baja resolución espacial y una relación señal-ruido (*signal to noise ratio*, SNR) reducida. Estas limitaciones subrayan la necesidad de perfeccionar tanto los sistemas de registro como las estrategias de procesamiento de señales para interpretar de forma precisa las intenciones del usuario BCI.

Paralelamente, el campo del aprendizaje profundo (*deep learning*, DL) ha experimentado avances significativos en la última década, algunos de los cuales se han aplicado con éxito en el análisis de señales EEG para una variedad de aplicaciones médicas y de investigación [3]. Gran parte de estos avances han sido posibles debido a la capacidad que tiene DL de transferir el aprendizaje de unas tareas a otras y a la disponibilidad de grandes conjuntos de datos etiquetados. No obstante, el aprendizaje auto-supervisado (*self-supervised learning*, SSL) ofrece una oportunidad adicional para optimizar las capacidades del DL, al permitir el preentrenamiento en datos no etiquetados, más abundantes y menos costosos de adquirir. Este método ya ha sido usado en estudios de señales EEG, según las referencias [4], [5]. En el primer estudio, se adapta el exitoso proceso de SSL anteriormente empleado al reconocimiento de habla [4]. En este enfoque, la red diseñada primero comprime la señal con una red de capas convolucionales (CNN) en un espacio latente, que luego se enmascara. Después, se intenta reconstruir este espacio latente tras introducirlo en una arquitectura basada en *transformer* [6]. Un segundo estudio se basa en el primero mencionado [5]. Al igual que en el estudio anterior, la red simplifica la señal en un espacio latente y luego lo enmascara. Sin embargo, la diferencia clave aquí es que la red intenta recuperar la señal original, en lugar del espacio latente enmascarado. Si bien estos trabajos han tenido éxito en probar la utilidad de SSL, la tarea que realizan limita el aprendizaje de la red puesto que se basa en el enmascarado del espacio latente y no el enmascarado de la señal original.

En el presente estudio, abordamos estos desafíos al aplicar técnicas de SSL en el preentrenamiento de una red de aprendizaje profundo basada en la arquitectura de *transformer*. Descomponemos la señal EEG en segmentos, similar al enfoque utilizado en las redes *transformer* de visión (ViT) [7], y empleamos una técnica de enmascarado para ocultar y posteriormente reconstruir fragmentos de la señal original. Esta tarea, con un alto ratio de enmascarado (75%), fuerza a la red a aprender representaciones más robustas y efectivas de la señal EEG [8]. Además, una mayor relación de enmascarado acelera la fase de preentrenamiento SSL, ya que solo se procesan los segmentos no enmascarados de la señal.

El objetivo principal de esta investigación es evaluar el impacto del preentrenamiento con técnicas de SSL, basadas en enmascarado y reconstrucción de señal EEG para la mejora de la clasificación en sistemas BCI basados

en imaginación motora. Este incremento se evaluará en una base de datos pública de 109 sujetos [9]. El estudio prueba por primera vez el uso de redes basadas completamente en la arquitectura *transformer* para la extracción de características relevantes del EEG.

2. Métodos

2.1. Dataset

Se ha examinado la base de datos pública Physionet [9]. Esta base de datos incluye registros EEG a 160 Hz de 109 participantes, con un rango de 42 a 46 eventos de MI por individuo. Después de una señal auditiva, una flecha señalaba qué tipo de MI se debía efectuar (i.e., imaginar abrir o cerrar la mano izquierda o derecha) durante un periodo continuo de tres segundos. La actividad se llevó a cabo en una única sesión y los participantes no recibieron retroalimentación acerca de la tarea mental que estaban ejecutando. Esta base de datos se utilizará tanto para la tarea de preentrenamiento como para la clasificación.

Aplicamos el siguiente preprocesamiento sobre la señal EEG [10]: (1) se extraen los electrodos F7, F3, T7, C3, P7, P3, O1, PZ, CZ, F8, F4, T8, C4, P8, P4 y O2; (2) se aplica un filtro notch para eliminar la señal de línea eléctrica; (3) se realiza un filtrado espacial de referencia promedio común (CAR) a estos 16 electrodos; (4) se aplica un remuestreo a 128 Hz evitando aliasing con un filtrado pasabajo a 63 Hz; (5) se extraen los ensayos con una ventana

de tiempo de 1.5 segundos después de la señal auditiva; y (6) se aplica estandarización z-score por canal y ensayo.

2.2. Aprendizaje auto-supervisado (SSL)

En este trabajo adaptamos de forma rigurosa el SSL basado en la arquitectura de *autoencoders* con enmascaramiento (*masked auto-encoders*, MAE) originalmente aplicada en el análisis de imágenes [8]. Este proceso se articula en cuatro etapas fundamentales, esquematizadas en la *Figura 1*. En la primera etapa, diseñamos una capa neuronal específica que fragmenta la señal EEG en segmentos temporales de 250 ms, equivalente a 32 muestras a una frecuencia de 128 Hz. En la segunda etapa, introducimos una capa de *embedding*, que cumple una doble función: añadir información posicional a cada segmento y enmascarar un subconjunto de segmentos seleccionados aleatoriamente siguiendo una distribución uniforme. La necesidad de añadir información posicional surge debido a la naturaleza invariante al orden de las redes *transformer*, que no reconocen el orden secuencial de las entradas. En la tercera etapa, un *encoder* compuesto por una red *transformer* de seis capas se encarga de procesar los segmentos de señal que permanecen visibles. Finalmente, en la cuarta etapa, un *decoder* más ligero de una sola capa *transformer* se encarga de la tarea de reconstrucción de la señal EEG original, utilizando tanto los segmentos codificados como los tokens asociados a los segmentos enmascarados.

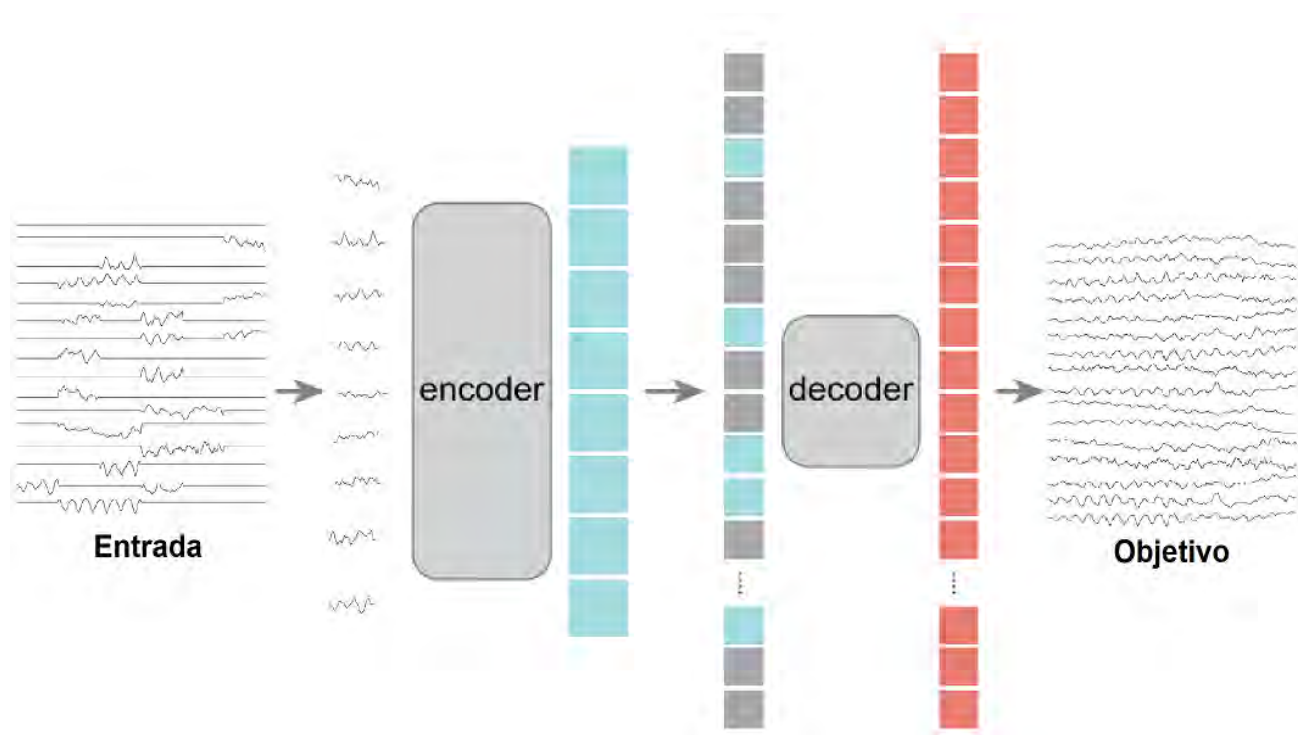


Figura 1. Arquitectura de autoencoder con enmascaramiento (*masked auto-encoder*, MAE) optimizada para aplicaciones en electroencefalografía (EEG). En la fase de preentrenamiento, una proporción significativa de la señal EEG (e.g., 75%) se enmascara intencionadamente. El encoder se encarga de procesar únicamente los segmentos de señal que permanecen visibles. A continuación, el decoder opera tanto sobre los segmentos codificados como sobre los tokens de enmascaramiento, con el objetivo de reconstruir fielmente la señal EEG original. Tras finalizar el preentrenamiento, se prescinde del decoder, y el encoder se emplea en el procesamiento de señales EEG sin enmascarar para llevar a cabo la clasificación de eventos de imaginación motora (MI). Figura adaptada [8].

Los hiperparámetros utilizados en la fase de preentrenamiento se detallan a continuación: (1) optimizador *AdamW*, (2) *learning rate* de 1×10^{-3} , (3) *weight decay* de 1×10^{-5} , (4) 500 épocas de entrenamiento, (5) detención temprana (*early stopping*) con un margen de espera de 25 épocas, (6) tamaño de *embedding* de 256, (7) 6 cabezas del mecanismo de atención, y (8) función de pérdidas basada en el error medio cuadrado (*mean squared error*, MSE).

2.3. Clasificación de imaginación motora (MI)

Tras completar la fase de preentrenamiento, descartamos el *decoder* y utilizamos exclusivamente el *encoder* para procesar la señal EEG completa sin enmascaramiento, que abarca un intervalo temporal de 1.5 segundos, con el objetivo de clasificar MI. Para evaluar la eficacia de nuestro enfoque, se establecen tres escenarios comparativos: (1) un modelo que no ha sido preentrenado con técnicas de SSL; (2) la aplicación de un sondeo lineal utilizando las características extraídas por el *encoder* en el preentrenamiento; y (3) *fine tuning* de toda la arquitectura de la red a partir del *encoder* preentrenado.

Optamos por medir el rendimiento mediante un esquema de validación cruzada inter-sujetos *k-fold* ($k=5$). En otras palabras, en cada uno de los tres escenarios mencionados, el modelo se entrena para clasificar MI en el 80% de los participantes, y posteriormente se valida en el 20% de los sujetos que no fueron parte del conjunto de entrenamiento. Repetimos este proceso cinco veces, lo que nos permite calcular la precisión de clasificación en todos los sujetos incluidos en la muestra de entrenamiento. Al comienzo de cada iteración los parámetros del *encoder* se inicializarán aleatoriamente en el escenario (1), mientras que se usarán los obtenidos durante el preentrenamiento en los escenarios (2) y (3).

3. Resultados

3.1. Reconstrucción de la señal

Si bien la tarea de reconstrucción de la señal no es el objetivo final que se busca con la aplicación del SSL, es de gran interés observar qué grado de éxito alcanza en esta tarea el conjunto *encoder/decoder*. La *Figura 2* ilustra la

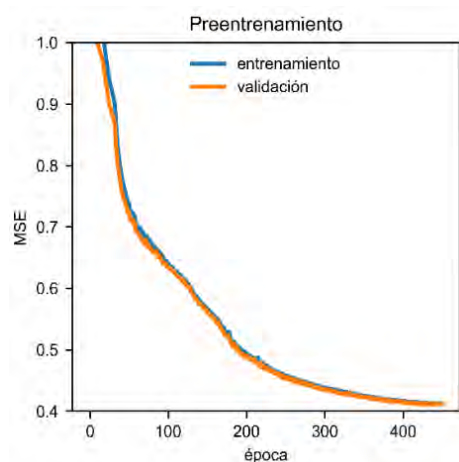


Figura 2. Evolución del valor del error medio cuadrado (MSE) de la tarea de reconstrucción durante el aprendizaje auto-supervisado (SSL).

evolución del valor del MSE entre la señal reconstruida y la señal original a lo largo del proceso de entrenamiento del SSL. El gráfico evidencia una evolución paralela en el MSE, tanto para el conjunto de entrenamiento como para el conjunto de validación. El entrenamiento se detiene antes de llegar a las 500 épocas debido a la activación del mecanismo de *early stopping*. La *Figura 3*, por otro lado, muestra el proceso de enmascaramiento y reconstrucción llevado a cabo por nuestro conjunto *encoder/decoder* tras el preentrenamiento basado en MAE.

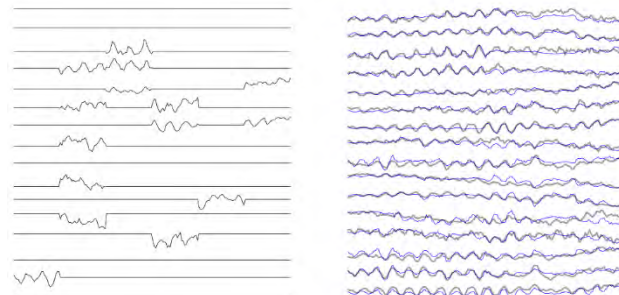


Figura 3. Representación del enmascaramiento y reconstrucción de una señal tras el preentrenamiento. A la izquierda la señal enmascarada. A la derecha en gris transparente la señal original y en azul la señal reconstruida a partir de la señal no enmascarada.

3.2. Clasificación en Physionet

En la *Tabla 1* se pueden observar los resultados de precisión para los 109 sujetos de Physionet [9] en cada una de las tres condiciones: (1) sin preentrenamiento; (2) preentrenamiento con sondeo lineal; y (3) preentrenamiento con *fine tuning* de la arquitectura completa. Se presenta la media de clasificación y la desviación estándar de las precisiones de clasificación de la imaginación motora inter-sujeto. Se observa como la condición de preentrenamiento SSL basado en MAE junto con *fine tuning* de toda la red logran los mejores resultados de precisión. Este comportamiento es análogo al observado en el trabajo original aplicado a imágenes [8].

	$\mu \pm \sigma$
Sin preentrenamiento	78.4±12.2
SSL + sondeo lineal	79.9±11.9
SSL + <i>fine tuning</i>	82.4±11.7

Tabla 1. Comparación de precisiones para los tres escenarios. $\mu \pm \sigma$: precisión media y desviación estándar obtenida para todos los sujetos siguiendo un *k-fold* con 5 iteraciones. Los mejores resultados están en **negrita**.

4. Discusión

El presente estudio aborda el desafío persistente en el ámbito BCI referente a la decodificación precisa de las intenciones del usuario. La combinación de técnicas de SSL basadas en MAE y la arquitectura de *transformer* representa un avance significativo en el procesamiento de señales EEG. El uso de un enmascaramiento de alta proporción (75%) directamente sobre la señal EEG es novedoso,

acelera el entrenamiento, y se demostró eficaz en la extracción representaciones relevantes de la señal. Además, la utilización de un esquema de validación cruzada inter-sujeto k -fold proporciona un marco riguroso para evaluar el rendimiento del modelo.

Los resultados sugieren que el uso de SSL podría tener aplicaciones más amplias en el campo de la IB, particularmente en la interpretación y el análisis de datos EEG. Sin embargo, hay algunas limitaciones que deben considerarse. Por ejemplo, aunque el dataset Physionet incluye gran cantidad de usuarios, la inclusión de más bases de datos podría ofrecer un preentrenamiento más potente y mejorar las representaciones obtenidas. Sobre todo, bases de datos con miles de usuarios y horas de registros EEG como la base de datos de Temple University Hospital (TUH) [11]. Aunque esta metodología ha mostrado avances al eliminar completamente el uso de capas convolucionales, los resultados obtenidos aún no superan a los trabajos que se consideran el estado del arte y que utilizan CNNs para la clasificación.

Comparado con trabajos anteriores que emplean SSL en el análisis de EEG [4, 5], este estudio supera ciertas limitaciones al centrarse no solo en el enmascarado del espacio latente creado por una etapa convolucional, sino también en el enmascarado de la señal original. Este enfoque permite a la red componerse solo de módulos *transformer*, que permiten acelerar el preentrenamiento al solo introducirse los segmentos de señal visibles.

Las futuras investigaciones podrían explorar el *fine tuning* del modelo a otros paradigmas BCI. Además, se podría emplear mayor cantidad de datos de EEG con gran variedad de paradigmas como datos de preentrenamiento, y no solo la misma base de datos que se usará para evaluar.

5. Conclusión

Este trabajo ha investigado la aplicabilidad y eficacia de técnicas de SSL para mejorar la precisión en la decodificación de las intenciones del usuario en BCIs basadas en EEG. Hemos demostrado que la arquitectura de *transformer*, que está preentrenada mediante técnicas SSL, permite segmentar y procesar eficazmente las señales EEG.

La validación de nuestro modelo en la clasificación de MI sobre una base de datos pública de 109 sujetos, mediante un enfoque de validación cruzada inter-sujeto k -fold, ha revelado una mejora significativa en precisión con respecto a la misma red sin este preentrenamiento. Los resultados indican que el preentrenamiento con técnicas de SSL no solo mejora la precisión de la clasificación, sino que también ofrece un nuevo camino hacia la interpretación más precisa de las señales EEG, lo cual es crucial para el avance de las BCIs.

Agradecimientos

Este estudio ha sido financiado por los proyectos TED2021-129915B-I00, PID2020-115468RB-I00 y RTC2019-007350-1 financiadas por el Ministerio de Ciencia e Innovación/Agencia Estatal de investigación/10.13039/501100011033/, FEDER Una forma de hacer

Europa; y por ‘Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBERBBN)’ a través de ‘Instituto de Salud Carlos III’. S. Pérez-Velasco y D. Marcos-Martínez son beneficiarios de una ayuda PIF de la Consejería de Educación de la Junta de Castilla y León.

Referencias

- [1] J. Enderle *et al.*, *Introduction to biomedical engineering*. Academic press, 2012.
- [2] J. R. Wolpaw and E. W. Wolpaw, *Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 2012.
- [3] A. Craik *et al.*, “Deep learning for electroencephalogram (EEG) classification tasks: a review,” *J Neural Eng*, vol. 16, no. 3, p. 031001, Jun. 2019.
- [4] D. Kostas *et al.*, “BENDR: Using Transformers and a Contrastive Self-Supervised Learning Task to Learn From Massive Amounts of EEG Data,” *Front Hum Neurosci*, vol. 15, no. June, pp. 1–15, Jun. 2021.
- [5] H.-Y. S. Chien *et al.*, “MAEEG: Masked Auto-encoder for EEG Representation Learning,” no. NeurIPS, Oct. 2022.
- [6] A. Vaswani *et al.*, “Attention Is All You Need,” *Adv Neural Inf Process Syst*, vol. 2017-Decem, no. Nips, pp. 5999–6009, Jun. 2017.
- [7] A. Dosovitskiy *et al.*, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” *ICLR 2021 - 9th International Conference on Learning Representations*, Oct. 2020.
- [8] K. He *et al.*, “Masked Autoencoders Are Scalable Vision Learners,” in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2022, pp. 15979–15988.
- [9] A. L. Goldberger *et al.*, “PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals,” *Circulation*, vol. 101, no. 23, Jun. 2000.
- [10] S. Perez-Velasco *et al.*, “EEGSym: Overcoming Inter-Subject Variability in Motor Imagery Based BCIs With Deep Learning,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 1766–1775, 2022.
- [11] A. Harati *et al.*, “The TUH EEG CORPUS: A big data resource for automated EEG interpretation,” in *2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, IEEE, Dec. 2014, pp. 1–5.

Impresión 3D

Jueves 23 8:30 - 10:00

Aula PB3

Metodología de planificación, diseño y fabricación de modelos anatómicos y guías quirúrgicas a medida mediante impresión 3D en caso de hemipelvectomía por osteosarcoma

E. Gómez Larrén¹, S. Gómez de los Infantes Peña¹, E. Aguilera Jiménez¹, M. Rodríguez Rodríguez¹, R. Pérez-Mañanes¹

¹ Unidad de Planificación Avanzada y Manufactura 3D, Hospital Universitario Gregorio Marañón, Madrid, España, 3dlab.hgugm@salud.madrid.org

Resumen

Se presenta un caso clínico de hemipelvectomía II - III por osteosarcoma realizado por el Servicio de Cirugía Ortopédica y Traumatología en colaboración con los ingenieros de Unidad de Impresión 3D (UPAM3D) del Hospital Universitario Gregorio Marañón. Se describe el flujo de trabajo aplicado para la planificación previa a la cirugía, así como el uso de la tecnología 3D en el diseño y fabricación a medida de modelos anatómicos, guías quirúrgicas de corte e implantes personalizados.

1. Introducción

La aparición de la industria 4.0 y el uso de las tecnologías de impresión 3D ha supuesto un cambio de paradigma en el campo de la medicina, impulsando una transformación digital de los procesos y orientando la atención médica hacia una medicina personalizada. Esto ha permitido ofrecer diagnósticos y tratamientos más precisos y específicos para cada paciente, llevando a una atención más individualizada y efectiva.

Por ello, el Hospital General Universitario Gregorio Marañón ha apostado por la creación de una unidad multidisciplinar de diagnóstico y planificación terapéutica avanzada que presta servicios de diseño biomédico y fabricación de producto sanitario personalizado a partir de tecnologías de imagen e impresión 3D, la UPAM3D.

La Unidad de Planificación Avanzada y Manufactura 3D (UPAM3D) permite la integración de las tecnologías 3D en la práctica clínica asistencial estimulando el trabajo colaborativo entre clínicos e ingenieros, generando conocimiento en equipo y estableciendo controles de calidad y vigilancia. La UPAM3D representa un espacio de innovación asistencial que supone la consolidación del proyecto del Laboratorio de Impresión 3D en funcionamiento en el hospital desde el año 2015, con un histórico acumulado de más de 1.400 casos clínicos tratados mediante tecnología 3D, lo que le ha permitido llegar a ser referente a nivel nacional e internacional. [1]

Se presenta un caso clínico de un paciente de 76 años con osteosarcoma en la hemipelvis derecha, un subgrupo de tumor maligno óseo que se origina en los huesos de gran tamaño [], que requiere ser intervenido para evitar un mal pronóstico.

La creación de una unidad multidisciplinar dentro del hospital, y por tanto la integración de tecnologías 3D y de la figura del ingeniero en la práctica clínica asistencial, tiene como objetivo principal mejorar la atención sanitaria de los pacientes y los resultados obtenidos tras los procedimientos quirúrgicos.

En el caso de la hemipelvectomía, el objetivo es realizar una planificación prequirúrgica avanzada, incluyendo la obtención de un biomodelo a escala real de la pelvis del paciente (Producto Sanitario a Medida Clase I) y del aloinjerto del Banco de Huesos que se va a emplear en la cirugía; así como el diseño y fabricación de las guías quirúrgicas de corte personalizadas (PSM Clase IIa) para mejorar la precisión y disminuir el tiempo quirúrgico.

2. Metodología

Para llevar a cabo la planificación quirúrgica y fabricación de producto sanitario a medida se lleva a cabo el flujo de trabajo indicado a continuación.

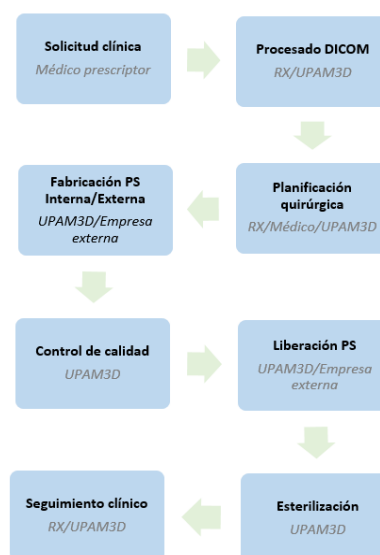


Figura 1. Flujo de trabajo

El proceso comienza con una solicitud clínica a la UPAM3D por parte del facultativo prescriptor, en la que se indica la patología del paciente y la necesidad.

2.1. Procesamiento de imágenes médicas

Partiendo de la imagen radiológica del paciente, en este caso una Tomografía Axial Computarizada (TAC); los ingenieros, el Servicio de Radiología y los cirujanos trabajan de forma colaborativa obteniendo una reconstrucción tridimensional de la pelvis, osteosarcoma y vasculatura de la región afectada. A partir de la segmentación del TAC se seleccionan las zonas anatómicas de interés para un posterior procesado de las mismas.

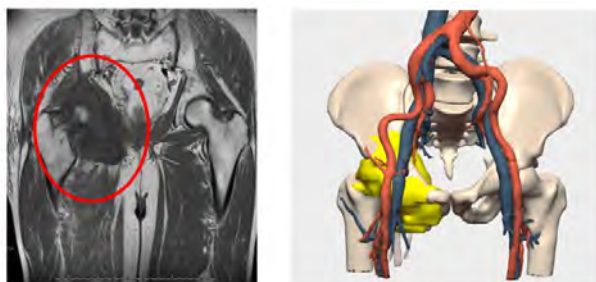


Figura 2. Imagen radiológica del paciente y reconstrucción tridimensional

Por otro lado, se obtiene una reconstrucción tridimensional de los diferentes aloinjertos de pelvis almacenados en el Banco de Huesos del propio Hospital. De entre todas las pelvis escaneadas, se selecciona la que mejor se adapta a la anatomía del paciente.

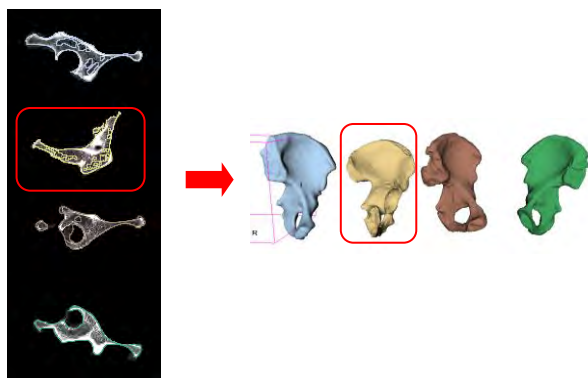


Figura 3. Segmentación TAC aloinjertos. Reconstrucción 3D de los aloinjertos

2.2. Planificación quirúrgica

Gracias a la colaboración entre ingenieros y cirujanos se planifica la cirugía de forma virtual, determinando la correcta posición del aloinjerto sobre el paciente, los planos de osteotomía y las guías quirúrgicas.

En primer lugar, se superponen los aloinjertos sobre el modelo 3D de la pelvis del paciente, determinando qué hueso es el óptimo para la reconstrucción de la anatomía del paciente.

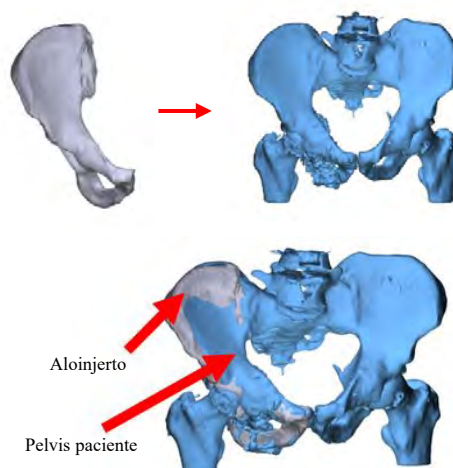


Figura 4. Superposición virtual del aloinjerto y de la pelvis del paciente

Una vez se establece la posición del aloinjerto sobre la pelvis, se determinan los planos de osteotomía de ambas piezas con la finalidad de diseñar guías de corte (PSM Clase IIa) tanto para la pelvis del paciente como para el aloinjerto. Estas guías tienen alojamientos para ser ancladas al hueso y un plano de corte para guiar la osteotomía. Se fabrican en resina biocompatible mediante tecnología SLA, y tras el proceso de curado, verificación y liberación del producto sanitario se lleva a cabo el proceso de esterilización.

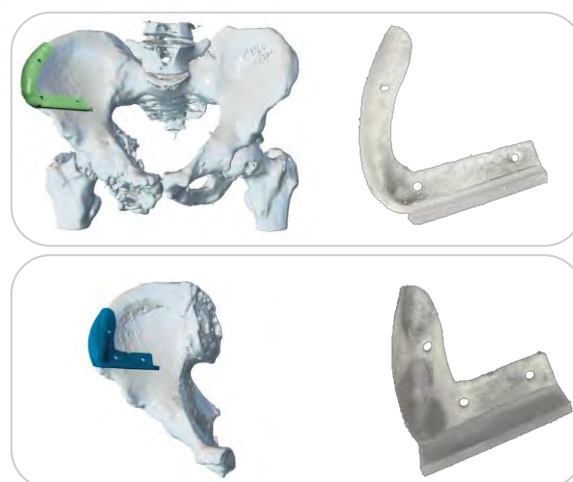


Figura 5. Diseño de las guías. Guías fabricadas mediante impresión 3D en tecnología SLA

Para esta cirugía, se diseñaron placas de fijación a medida y fabricadas en resina, que sirvieron para moldear las placas que se encuentran en stock en el propio Hospital y que finalmente son implantadas en el paciente.

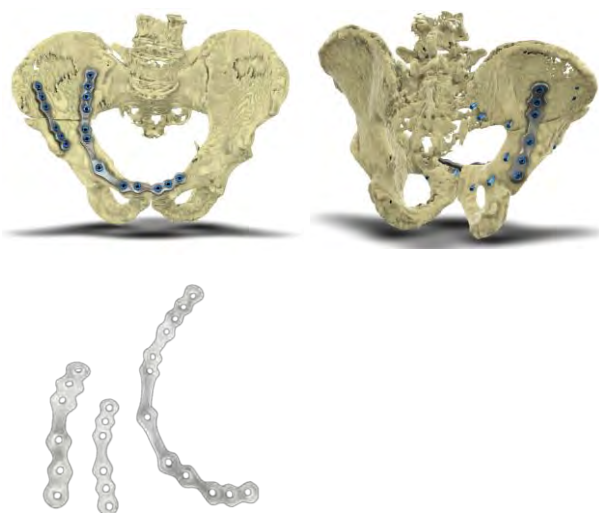


Figura 5. Placas Diseño virtual de placas a medida sobre la pelvis del paciente. Placas impresas en resina

2.3. Intervención quirúrgica

Durante la cirugía, parte del equipo quirúrgico lleva a cabo el abordaje sobre el paciente para realizar la osteotomía en la pelvis y lograr extraer el osteosarcoma. De forma paralela se trabaja sobre el aloinjerto, usando la guía quirúrgica según lo planificado.



Figura 6. Planificación virtual de la guía de corte. Guía sobre el hueso del paciente

Durante la resección oncológica, se reseca la lesión teniendo en cuenta los márgenes de tejido sano adyacentes al tumor.

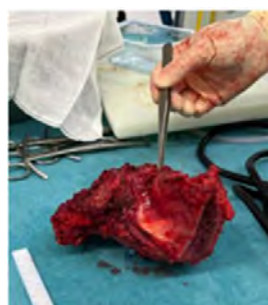


Figura 7. Pieza resección oncológica incluyendo márgenes de tejido sano

De forma paralela, se premoldean las placas estándar sobre las placas diseñadas a medida y fabricadas en el propio Hospital.

Una vez el aloinjerto está cortado y la placa moldeada, se comprueban sus posiciones junto con el biomodelo (Producto Sanitario a medida Clase I).



Figura 8. Placa premoldeada sobre biomodelo y aloinjerto

Finalmente, una vez finalizada la cirugía se realiza una radiografía al paciente, en la que se observan las placas colocadas según lo planificado.

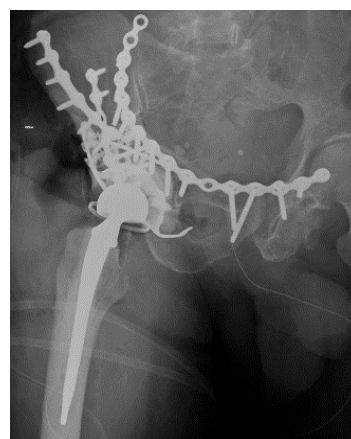


Figura 9. Imagen radiológica postquirúrgica

2.4. Marco regulatorio

Siguiendo el Reglamento de Producto Sanitario, los modelos anatómicos y las guías quirúrgicas son Productos Sanitarios a Medida, ya que están destinados únicamente a un paciente determinado y son empleados para tratar una necesidad de ese paciente. Los productos sanitarios se clasifican en función de la finalidad prevista y los riesgos siguiendo las reglas del Anexo VIII Reglas de Clasificación del MDR. [] . De esta forma, los modelos anatómicos se incluyen dentro de la norma como Producto Sanitario a Medida Clase I y las guías quirúrgicas como Producto Sanitario a Medida Clase IIa. El fabricante de estos productos, en este caso el Hospital Universitario Gregorio Marañón, posee la licencia de fabricante otorgada por la Comunidad de Madrid, cumpliendo así con los requisitos que establece el reglamento europeo 745/2017 [3].

Además, la UPAM3D tiene certificado su sistema de gestión de calidad bajo la norma internacional ISO 13485:2016 para dispositivos médicos.

3. Resultados

Como resultado del trabajo colaborativo entre los ingenieros y los cirujanos, se ha planificado el abordaje del caso y se han obtenido las guías de corte a medida prescritas para su uso en el procedimiento quirúrgico.

Durante la intervención se confirmaron las similitudes entre la planificación prequirúrgica y la anatomía del paciente. Asimismo, las guías quirúrgicas se adaptaban de forma correcta tanto al hueso del paciente como al aloinjerto, logrando así una reconstrucción óptima. Durante la cirugía, el equipo de cirujanos trabaja de forma paralela para atender al paciente y preparar el aloinjerto, permitiendo una disminución del tiempo quirúrgico y una mayor precisión.

La reconstrucción en 3D del tumor difiere de la resección que los cirujanos realizaron durante la cirugía, ya que en la intervención quirúrgica se establecen ciertos márgenes de seguridad que incluyen tejido sano.

El resultado final de la intervención quirúrgica fue satisfactorio, concluyendo con éxito el procedimiento.

4. Conclusiones

La complejidad de los problemas médicos exige la síntesis de conocimientos tecnológicos e ingenieriles para desarrollar soluciones. Por ello, la ingeniería se ha convertido en un pilar esencial para un cambio de paradigma en la medicina actual, más orientada a un tratamiento personalizado y basado en las necesidades tanto del paciente, como de los profesionales sanitarios, y, por tanto, a la mejora de la atención sanitaria.

El uso de las tecnologías 3D y el trabajo multidisciplinar es fundamental para garantizar el éxito de la intervención, reduciendo los tiempos quirúrgicos y los riesgos durante la cirugía.

Como se ha observado durante este trabajo, los productos diseñados a medida son complementarios de los productos estándar.

La planificación preoperatoria avanzada permite una mejor comprensión del caso, entendiendo las relaciones anatómicas y planificando los pasos quirúrgicos. El flujo de trabajo y las tecnologías aplicadas se pueden usar en cirugías de otras áreas anatómicas, permitiendo una personalización del tratamiento clínico del paciente.

Agradecimientos

El trabajo llevado a cabo en la UPAM3D se lleva a cabo gracias a la implicación del Hospital Universitario Gregorio Marañón, a los profesionales sanitarios y a las empresas Osteobionix, Cella Medical Solutions, Mizar Health, Serveo y Bitfab.

Referencias

- [1] Página web de la UPAM3D. <https://www.comunidad.madrid/hospital/gregoriomaranon/profesionales/idades-multidisciplinares/unidad-3d-upam3d> (Consultada: Agosto 2023).
- [2] Sánchez-Torres LJ, Santos-Hernández M. Osteosarcoma. *Rev Mex Ortop Ped.* 2011;13 (1):10-19.
- [3] Página web de Eur-Lex. <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32017R0745&from=D A> (Consultada: Agosto 2023)

3D printed implant structure for breast reconstruction and systematic drug delivery for preventing the reoccurrence of cancer

C. Ludusan¹, D. Ojados González², R. Pacurar¹, J. F. Roca González³, M. García Fuentes², L. Martínez Abellán²

¹ Industrial Engineering Department, Technical University of Cluj-Napoca, Cluj-Napoca, Romania, cristinna.ludusan@yahoo.com, razvan.pacurar@tcm.utcluj.ro

² Technological Research Support Service (SAIT), Polytechnic University of Cartagena, Cartagena, Spain, {lola.ojados, marta.garcia, lucia.mabellan}@upct.es

³Department of Automation, Electrical Engineering and Electronic Technology, Polytechnic University of Cartagena, Cartagena, Spain, jroca.gonzalez@upct.es

Abstract

Breast cancer is one of the most common and deadly cancers among women worldwide. The treatment of breast cancer often involves mastectomy, which can cause physical and psychological distress to the patients, affecting their quality of life and self-esteem [1]. Therefore, the surgical intervention of breast cancer should not include only the removal of a tumor, but also psychological support by restoring the natural shape of the breast and achieving proper aesthetic for each patient. Conventional breast reconstruction techniques have several limitations, such as donor site morbidity, implant failure, infection, and poor cosmetic outcome. Moreover, breast cancer patients often require systemic chemotherapy or hormone therapy after surgery, which can cause severe side effects and complications.

In this article, it is proposed a novel implant based 3d printed structure for breast reconstruction and systematic drug delivery. The implant consists of a biodegradable scaffold that is 3d printed with a customized shape, coated with Polydopamine and Alginate, and sized to fit the patient's chest. The implant provides mechanical support, aesthetic improvement and can also enhance the efficacy and safety of chemotherapy or hormone therapy by delivering treatment directly to the target site through a channel system. The aim of the article presented here is a proposed process protocol for obtaining this prosthesis.

1. Introduction

Due to tissue engineering, breast reconstruction surgeries and clinical treatment strategies gain an encouraging perspective when it comes to overcoming the limitations of their development and innovation [2]. Using scaffolds to guide the formation of adipose tissue offers an optimistic view regarding the control of regeneration of soft tissue and a new tactic for obtaining the desired shape and surface features for each patient. The implant opens the possibility for the scaffold to be seeded with adipose-derived stem cells (ADSCs) that can differentiate into adipocytes and form new fat tissue [3]. Moreover, choosing the right unit cell, density and distribution in the scaffold, an efficient system for delivering targeted treatment could be obtained. The goal is to design a framework that can withstand the above-mentioned requirements and provide resistance to

mechanical forces, stresses and strains and offer comfort and aesthetic satisfaction to the patient.

2. Objectives

The main objective of this study is to develop a protocol for the design of this type of prosthesis, which is a biodegradable implant with a customized shape for each patient. This implant must provide a structure where new fat tissue is generated. This also requires the choice of a material suitable for this purpose.

3. Methodology

3.1. Model

In order to obtain the desired protocol, this study follows the case of a patient who will undergo a tumor removal surgery as a support to elaborate a guide of steps to follow in this process set as an objective.

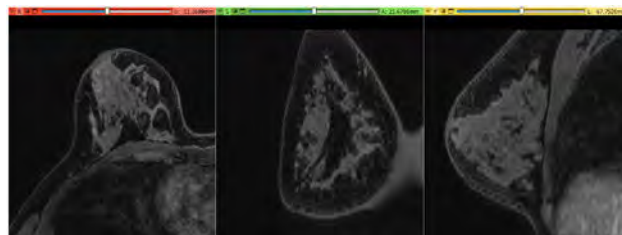


Figure 1. DICOM images

For this purpose, the image used from this particular case is a specific MRI (Figure 1) of a patient whose 3D model and geometry were generated using Slicer 5.2.2. (Figure 2) which is a free, open source software for visualization, processing, segmentation, registration, and analysis of medical, biomedical, and other 3D images [4]. In this case, the images have been obtained from the radiology repository “cancerimagingarchive.net” [5].

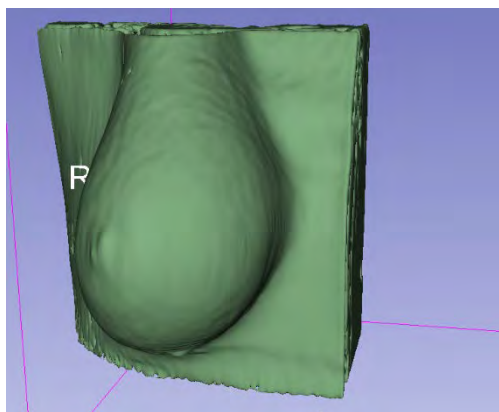


Figure 2. Example of a 3D model generated from a specific MRI scan

Then, the surface mesh (.STL model) (Figure 3) was imported into SolidWorks 2022-2023, where the model was converted into a solid and therefore used as a template for the design of the following structures. In the preliminary state, a prototype of the implant was designed, to represent the layout of the scaffold's architecture and the channel in the structure. The pre-designed model is then customized to the patient's breast parameters and needs (Figure 4), being able to withstand the sequelae required for tumor removal. This opens the possibility to regulate the topology of the model and apply different patterns with variable mesh density in each section. The final CAD design (Figure 5) will then be used for finite element analysis (FEA) and finally exported for fused deposition modeling (FDM).

Therefore, the protocol steps can be extracted from the methodology used in this specific case.

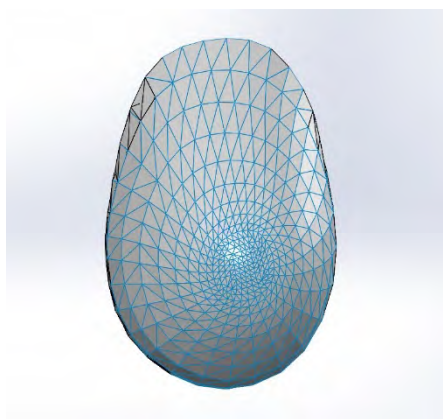


Figure 3. Surface mesh

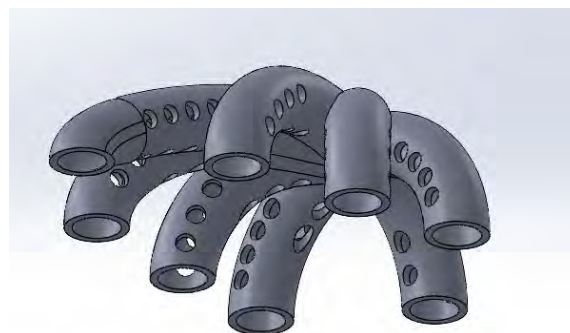


Figure 4. Example of distribution channels

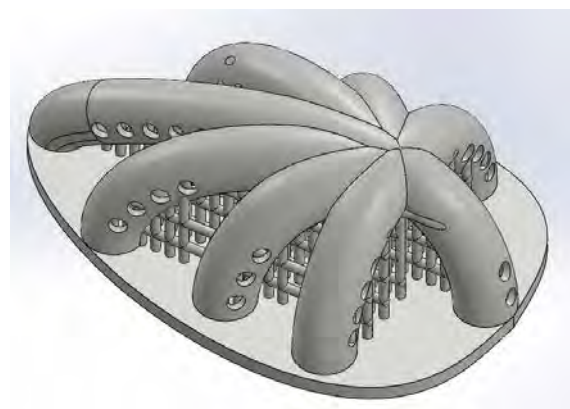


Figure 5. Example of final implant structure

3.2. Materials

The materials required to make this type of prosthesis can be estimated after the completion of this study, in which it is proposed that the scaffold consists of three materials: a base made of polycaprolactone, coated with polydopamine and alginate (Figure 6).

Coating PCL with a PDA layer followed by an alginate layer can be a multi-step process and has various purposes: PCL is often used as a scaffold material in tissue engineering. Coating it with PDA and alginate can enhance cell adhesion [6], proliferation, and tissue integration, making them more suitable for cell attachment and growth [7]. Also, PCL can be used as a drug delivery vehicle, and PDA and alginate coating can further control pharmacokinetics and improve drug-loading capacity.

PDA and alginate/gelatine coatings can also be used to immobilize bioactive molecules, growth factors, or signaling molecules onto the PCL surface, further enhancing its functionality. The addition of alginate can also introduce some viscoelastic behavior to the composite material [8], which can be relevant for simulating the mechanical behavior of natural tissues.

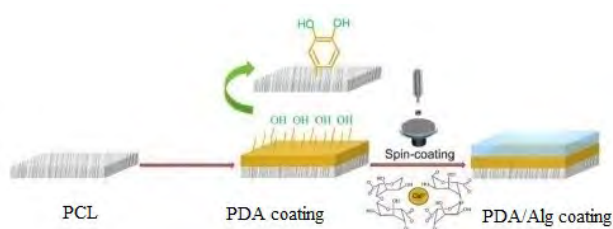


Figure 6. PCL scaffold with PDA and alginates coatings

3.3. Scaffolds

The scaffold (Figure 7) should enable the formation of tissue that closely resembles the patient's natural tissue in terms of its biological, physical, and chemical characteristics. It should offer three-dimensional structural stability, exhibit a low elastic modulus and include proliferation [9]. The construction and distribution of the unit cell is vital.

As for the printing parameters, the pattern should be zig zag with a 45 degrees infill direction 0.2 mm layer height and 70% infill density. These parameters would assure an ideal environment for cell growth [8].

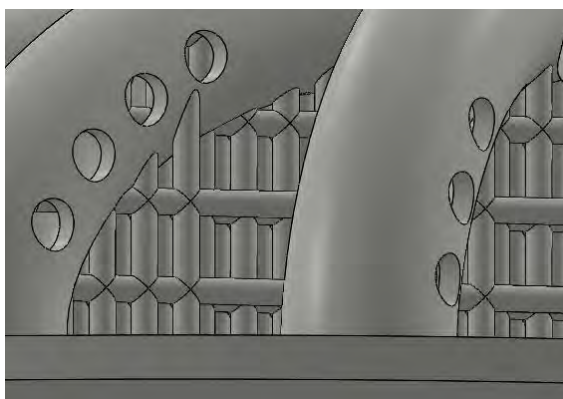


Figure 7. Close up of an example of the scaffold

3.4. Finite element analysis

The finite element analysis is done in Solidworks 2022-2023 Simulation by applying mechanical loads on the external part of the structure (Figure 8).

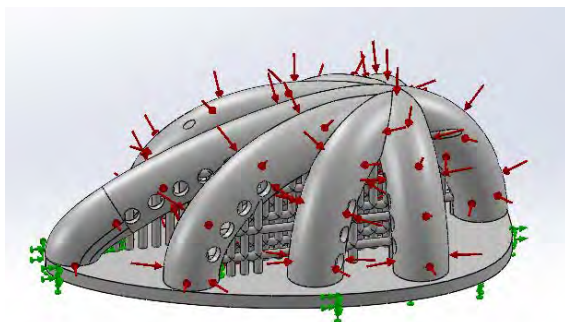


Figure 8. FEA of the structure

The stresses and strains that appear during the analysis are in correlation with the volume, structure and figure of the scaffold, as well as the patient's habits and amount of physical activities.

The first step is to set a static analysis. After that, the properties of the material are introduced and a mesh control is projected on the whole model. The parameters and mechanical forces are applied to obtain visual results of the strains that occur in the solid model, as following: a normal force of 5 N that comes in contact with the implant, the tensile strength and Young's modulus of 59 ± 11.3 KPa, respectively 0.53 ± 0.24 MPa [9].

3.5. Optimization

Further studies are required in order to obtain realistic outcomes of the influence of the above-mentioned parameters.

Optimization in this case would constitute modifying the unit cell of the scaffold, diameter, density and their distribution and topography. Moreover, it could also imply a change in the number of channels or their volume. One important mention is that the mechanical properties of the PCL-PDA/Alg scaffolds could be tailored by the thickness of the PDA coatings and the density infill.

4. Results

The result to be obtained from this article is the following protocol for the development of this type of prosthesis:

- Firstly, the .stl will be extracted from a DICOM image of a patient by using specific software for this purpose.
- The mesh will be obtained through a CAD software.
- The mesh will be used as a template to obtain a 3D model of the prosthesis that is the target of this article.
- Once the 3D model is obtained, a finite element analysis is carried out.
- Finally, it will go through the usual FDM printing process.

5. Conclusion

Breast scaffolds with distinct structures have been created to meet the complex biomechanical demands of large-volume and long-term tissue regeneration. Furthermore, a channel structure is included to enable precise drug delivery to prevent the reoccurrence of cancer. The process including MRI scans and producing patient-specific model, adapting the dimensions and overall form of the implant would solve the aesthetic and comfortability problems mentioned in the beginning of his article.

Depending on the viscosity of the fat, a gradient of pore sizes can be applied to minimize fat displacement and prevent leakage effectively. Therefore, these scaffolds should be capable of topical stimulus-responsive treatment, long-term drug delivery and overall encouraging the fat distribution inside the structure.

The study examined in this context illustrates the potential of 3D printing for breast reconstruction but, nevertheless,

significant efforts are still required to validate this technique and facilitate its broader clinical application.

Aknowledgements

The author would like to express their gratitude to the organizing committee of CASEIB for the permission to use their style guides as a reference for the realization of this document and also thank the staff of SEDIC (Industrial Design and Scientific Calculation Service) of the Technical University of Cartagena for their support in the development of this work and for the possibility of having scientific-technical equipment.

References

- [1] Fernández-Delgado, J., López-Pedraza, M. J., Blasco, J. A., Andradas-Aragones, E., Sánchez-Méndez, J. I., Sordo-Miralles, G., & Reza, M. M. (2008a). Satisfaction with and psychological impact of immediate and Deferred Breast Reconstruction. *Annals of Oncology*, 19(8), 1430–1434. doi:10.1093/annonc/mdn153
- [2] Dang, W., Wu, J., Wang, G., Zhen, Y., & An, Y. (2023). Role of adipose-derived stem cells in breast cancer. *Chinese Journal of Plastic and Reconstructive Surgery*, 5(2), 73–79. doi:10.1016/j.cjprs.2023.02.002
- [3] Teixeira, A. M., & Martins, P. (2023). A review of bioengineering techniques applied to breast tissue: Mechanical properties, tissue engineering and finite element analysis. *Frontiers in Bioengineering and Biotechnology*, 11. doi:10.3389/fbioe.2023.1161815
- [4] Web 3D Slicer (2022). URL: <https://www.slicer.org/> (Visited September 2023)
- [5] Web cancerimagearchive (2022). URL: <https://www.cancerimagingarchive.net/> (Visited: September 2023)
- [6] Khan, Ziyauddin & Shanker, Ravi & Um, Doo-Seung & Jaiswal, Amit & Ko, Hyunhyub. (2018). Bioinspired Polydopamine and Composites for Biomedical Applications. 10.1002/9783527807918.ch1.
- [7] Mu X, Zhang J and Jiang Y, 3D Printing in Breast Reconstruction: From Bench to Bed. *Front. Surg.* 8:641370. doi: 10.3389/fsurg.2021.641370
- [8] Mina Mohseni, Onur Bas, Nathan J. Castro, Beat Schmutz, Dietmar W. Huttmacher, Additive biomanufacturing of scaffolds for breast reconstruction, *Additive Manufacturing*, Volume 30, 2019, 100845, pp 4, (ISSN 2214-8604).
- [9] Rabionet M, Polonio E, Guerra AJ, Martin J, Puig T, Ciurana J. Design of a Scaffold Parameter Selection System with Additive Manufacturing for a Biomedical Cell Culture. *Materials (Basel)*. 2018 Aug 14;11(8):1427. doi: 10.3390/ma11081427. PMID: 30110889; PMCID: PMC6119890.
- [10] Chunyang Liu, Zhiyong Wang, Xiaoyue Wei, Birui Chen, Yongxiang Luo, 3D printed hydrogel/PCL core/shell fiber scaffolds with NIR-triggered drug release for cancer therapy and wound healing. *Acta Biomaterialia*, Volume 131, 2021, Pages 314-325, ISSN 1742-7061, <https://doi.org/10.1016/j.actbio.2021.07.011>.

Simulador de broncoscopias mediante fabricación aditiva para la exploración médica personalizada

J. Hurtado Fuentes¹, Á. Hurtado Fuentes², D. Ojados González³, R. Pamies Porras¹

¹ Departamento de Ingeniería Mecánica, Materiales y Fabricación, Universidad Politécnica de Cartagena, Cartagena, Murcia, España {javier.hurtado@edu.upct.es}, {ramon.pamies@upct.es}

² Complejo Hospitalario Universitario de Albacete, Albacete, Castilla la Mancha, España

{alvarohurtadofuentes@gmail.com}

³ Servicio de Apoyo a la Investigación Tecnológica (SAIT), Universidad Politécnica de Cartagena, Cartagena, Murcia, España {lola.ojados@upct.es}

Resumen

Se ha fabricado mediante FDM un prototipo de simulador de broncoscopias a partir de la adquisición de imágenes de TAC. Se ha realizado un diseño inicial con ayuda del software 3D Slicer, elaborando un modelo tridimensional del segmento traqueobronquial que se ha exportado en formato STL (estereolitografía) al software Autodesk Meshmixer. Se ha añadido un espesor a la superficie de malla del archivo STL, se ha realizado una operación de molde y se ha realizado un corte para originar una entrada y salida del prototipo hueco. Con el programa CAD de diseño paramétrico Solidworks se han diseñado piezas adicionales, en concreto unos soportes y una glotis que se han ensamblado al prototipo ajustando la unión entre las piezas. Finalmente se ha realizado la impresión 3D del prototipo, donde con el programa Ultimaker Cura, se han configurado los parámetros de impresión.

De esta forma se ha obtenido un primer prototipo en PLA, que posteriormente se ha modificado, introduciendo en el diseño modificaciones y correcciones a raíz de las sugerencias del equipo médico del hospital donde fue probado. Para evitar la rigidez del simulador, se ha seleccionado como material para la impresión 3D del segundo prototipo, un elastómero termoplástico basado en poliuretano de la marca comercial FilaFlex®. Las modificaciones en el caso de estudio del prototipo final, han alcanzado a todos los niveles de obtención del mismo, no sólo han afectado a su fase de diseño, sino que se ha llegado a realizar una optimización en todo el proceso, usando todos los programas para rediseñar y ajustar el modelo 3D diseñado con 3D Slicer, realizar un corte más superficial y modificar el espesor impuesto inicialmente, mediante el programa Autodesk Meshmixer y finalmente, variando los parámetros del programa de configuración del proceso de fabricación aditiva, pasando de usar Ultimaker Cura a la extensión BCN3D Cura.

1. Introducción

Fused Deposition Modelling (FDM) es una de las tecnologías para la impresión 3D más extendida que da la posibilidad de obtener piezas de geometría compleja. Además, esta técnica puede proporcionar prototipos y componentes adaptados a demandas muy singulares [1]. Comparada con otras tecnologías 3D de fabricación aditiva, FDM es un método simple de coste competitivo que aporta resultados excelentes, por lo que supone un 60 % de los

equipos 3D utilizados en la actualidad [2]. La comunidad científica explora la impresión 3D para el diseño de piezas con aplicaciones tecnológicas como la electrónica [3], medicina [4], la industria espacial [5] entre otras [6].

En este trabajo se ha realizado la fabricación de un simulador para la realización de broncoscopias mediante tecnología de impresión 3D a partir de imágenes de tomografía axial computarizada (TAC) de pacientes. Para lo cual ha sido necesario realizar tareas de modelado 3D a partir de tratamiento de imágenes 2D y de selección de los materiales adecuados para la obtención de un primer prototipo y modificación del mismo, alcanzando un prototipo optimizado que recoge las sugerencias y necesidades del equipo médico encargado de probar y validar la funcionalidad y eficiencia del mismo.

2. Materiales y Métodos

Para el estudio del prototipo buscado se ha planteado como objetivo principal obtener la réplica del segmento traqueobronquial de un humano, de forma que se pueda simular en ella la realización de pruebas de broncoscopia, conformando así un simulador que, tras pruebas funcionales, requiera el visto bueno del personal médico, en este caso del Complejo Hospitalario Universitario de Albacete, implicado en este trabajo. Se persigue minimizar hasta anular defectos en su fabricación y que el resultado sea lo más realista posible.

Como materiales para la fabricación aditiva del prototipo del simulador se han empleado dos tipos diferentes de termoplástico. Por un lado, un bioplástico termoplástico basado en ácido poliláctico (PLA) de la marca BCN 3D Filaments y por otra parte FilaFlex®, un elastómero termoplástico basado en poliuretano, ambos comercializables en forma de filamento para impresión 3D.

Para llegar a la fabricación del prototipo se obtenido un modelo 3D digital del segmento traqueobronquial que se inspecciona mediante la prueba diagnóstica de broncoscopia. Para que este modelo sea fiel a la realidad se parte del escaneado mediante Tomografía Computarizada (TAC/TC) de esta parte de la anatomía de un paciente real y se procesa con herramientas de diseño CAD.

Software empleado. El procesado de imágenes 2D procedentes de TC se ha realizado mediante el programa Philips DICOM Viewer R3.0 SP15, incorporando las

imágenes con formato DICOM (Digital Imaging and Communications in Medicine). Posteriormente se ha utilizado el programa 3D Slicer 4.11 para obtener el modelo en tres dimensiones, que fue finalmente procesado con Autodesk Meshmixer. Para la realización del proceso de fabricación aditiva ha sido necesario utilizar el software Ultimaker Cura para fijar parámetros de impresión 3D y la opción Custom Supports para la inclusión y configuración de soportes.

Para la materialización física del prototipo se ha realizado un proceso de fabricación aditiva del diseño obtenido con herramientas CAD, empleándose equipos de impresión 3D.

Impresoras 3D. El simulador de PLA fue fabricado con una impresora Ender-3 Pro 3D, mientras que el simulador de poliuretano fue elaborado con el equipo BCN3D SigmaX R19, usando dos extrusoras debido a que los soportes fueron fabricados en PLA. Los detalles técnicos de las impresoras están recogidos en la Tabla 1.

Especificaciones	Ender-3 Pro 3D	BCN3D
		SigmaX R19
Espacio de impresión	220x220x250 mm	420x297x210mm
Extrusores	1	2
Espesor de capas recomendado	0,1mm-0,4mm	0,05mm–0,5mm
Diámetro de la boquilla	0,2mm–0,4mm	0,3mm–1mm
Diámetro del filamento	1,75mm	2,85mm
Temperatura de la cama de impresión	≤ 100°C	≤ 100°C

Tabla 1. Especificaciones básicas de las impresoras Ender-3 Pro y BCN3D SigmaX R19 Modelado 3D

3. Modelado

El punto de inicio del proyecto es la adquisición de las imágenes de TAC/TC gracias al escáner del Complejo Hospitalario Universitario de Albacete, centro promotor de este trabajo. Para la extracción de imágenes 2D segmentadas en tres planos usando el programa DICOM Viewer R3.0 SP15 (figura 1), fue necesario hacer una revisión de las imágenes obtenidas por tomografía para seleccionar y/o eliminar imágenes no válidas o formatos defectuosos.



Figura 1. Imagen DICOM en Philips DICOM Viewer

Cada secuencia de imágenes presenta cerca de 400 secciones. Tras esta verificación, se utilizó el software 3D Slicer para elaborar el modelo 3D. Se optó por una selección manual del volumen deseado para representar la tráquea y el árbol bronquial, para ello, en cada sección seleccionada se indicó la misma área visible desde los tres planos de representación perpendiculares a los ejes de observación, tal y como se muestra en las vistas 2a, 2b y 2c de la figura 2. El área mostrada en azul se selecciona inicialmente en el plano visible desde el eje axial y es el mismo programa el que identifica esa misma área en los otros dos planos visibles desde los otros dos ejes. Este proceso se ha repetido a lo largo de toda la serie de imágenes correspondientes a cada sección de lectura del escáner TAC. El resultado de procesado de estas imágenes ha dado lugar a un modelo tridimensional en geometría de superficies tipo malla, que se ha exportado posteriormente a formato STL para poder ser importado desde el siguiente software a emplear, Autodesk Meshmixer, cuya interfaz se puede ver en la figura 3. Mediante este programa se ha obtenido el primer modelo 3D del prototipo del simulador, seleccionando la superficie exterior de la malla del archivo STL y realizando una operación de extrusión hacia el exterior para dar espesor y de esta forma volumen al modelo convirtiéndolo en sólido. A posteriori, se ha ejecutado la operación de vaciado del modelo extruido, usando como límite la malla del modelo original importado.

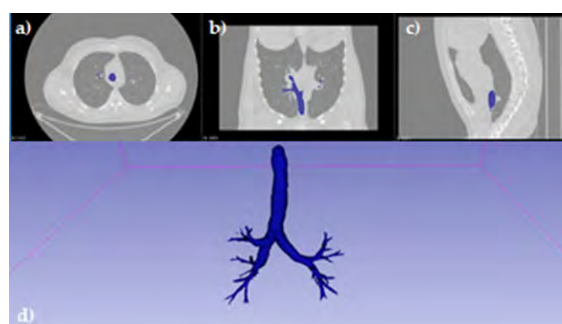


Figura 2. Vista en los tres planos (a, b y c) del área seleccionada para la construcción del modelo en 3D (d)



Figura 3. Interfaz de Autodesk Meshmixer

4. Impresión del primer prototipo

Para la obtención del primer prototipo estudiado, se ha realizado la impresión del simulador en 3D con el material PLA. Mediante el programa Ultimaker Cura se ha posicionado el prototipo en la base simulada en el software de la impresora y se han impuesto los parámetros para la configuración de su impresión. Para su correcto posicionamiento se ha comprobado que no existen soportes en el interior, puesto que su extracción habría sido muy complicada. Por este motivo en la configuración se ha optado por posicionar únicamente soportes que estén en contacto con la cama o base de la impresora, de esta manera se han podido controlar fácilmente los soportes que coloca el programa de forma automática. Esto ha provocado la necesidad de revisión del posicionamiento para que todos los lóbulos tengan su soporte y evitar así fallos en la impresión tales como la inexistencia de ciertas partes del prototipo por falta de soportes.

Antes de manipular el prototipo se ha revisado que el programa finalizara con éxito la impresión del modelo 3D, comparando éste con el archivo configurado en Ultimaker Cura. Una vez retirado el prototipo de la impresora, se inició la extracción manual de los soportes de forma mecánica, con la ayuda de diversas herramientas. Figura 4.



Figura 4. Primer prototipo con soportes (izquierda) y primer prototipo sin soportes (derecha)

Para el ensayo de funcionamiento de este prototipo, se ha procedido a practicar la simulación de la prueba de broncoscopia con la pieza impresa en PLA para verificar el modelo y obtener las primeras conclusiones. En este proceso se ha encontrado dificultad a la hora de realizar la introducción de la sonda de broncoscopia por el interior del prototipo obtenido, debido a la rigidez del material con el que se ha fabricado, que no ha resultado muy fiel a la realidad, y no ha permitido acceder a todos los orificios interiores diseñados en el modelo 3D del prototipo.

Adicionalmente, se ha apreciado que en algunas zonas han quedado residuos de material de impresión y acabados no deseados, visibles en la figura 5. Esto se ha podido originar por arrastre de errores de aproximación en la ejecución del procedimiento de obtención del modelo 3D mediante la secuencia de uso de las herramientas CAD empleadas y por la existencia de secreciones mucosas en el paciente que no se pudieron distinguir en el archivo DICOM de partida.

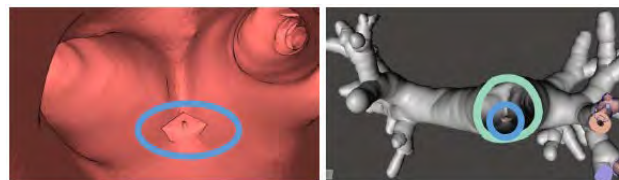


Figura 5. Errores detectados con Sketch up y Meshmixer

5. Iteración del proceso y optimización del diseño

Los fallos detectados en el ensayo de funcionamiento realizado sobre el primer prototipo han sido corregidos en una segunda iteración, para la que se ha decidido utilizar un filamento de poliuretano termoplástico mucho más flexible, cuyo comportamiento de deformación al ser sometido a tensión puede observarse, en comparación con el del PLA, en la gráfica de la figura 6.

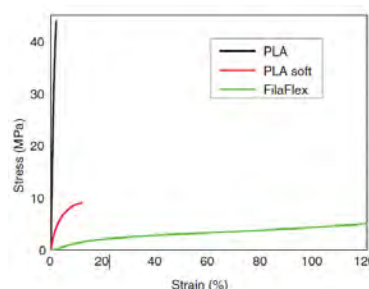


Figura 6. Gráfica tensión-deformación del PLA, PLA suave y FilaFlex®

En esta ocasión, antes de llevar a cabo la impresión del modelo 3D del prototipo en la impresora 3D, se ha realizado una simulación computacional del proceso de broncoscopia, realizándose una inspección por el interior del modelo con el software SketchUp Pro-2021, que ha garantizado que los residuos interiores hayan sido correctamente eliminados del diseño. Mediante este método de realización de una broncoscopia virtual en el modelo CAD tridimensional previo a su fabricación, se ha conseguido visualizar en forma de video de este proceso de inspección, en la figura 7 se muestra uno de los fotogramas en la interfaz del programa.

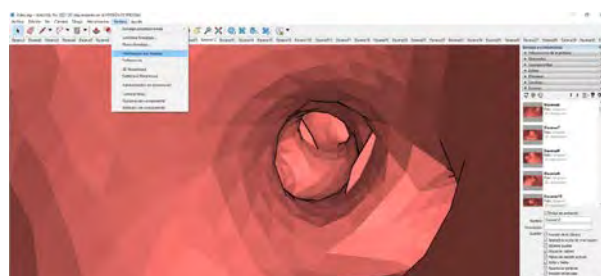


Figura 7. Inspección virtual en interfaz de SketchUp Pro 2021

En el diseño de este segundo prototipo se han incluido unas nuevas piezas diseñadas como sujeción de la pieza principal del segmento traqueobronquial que conforma el simulador y se ha diseñado un acople exterior al prototipo para incluir la réplica de una glotis, de esta forma el simulador quedaría completo y más cercano a lo que pueda encontrar el personal sanitario que vaya a utilizar este prototipo para su formación, antes de enfrentarse a pacientes reales. En el diseño de estos soportes se ha tenido en cuenta la posición anatómicamente correcta de los pulmones a la hora de realizar la prueba de broncoscopia, de forma que el paciente está situado en posición decúbito supino o boca arriba. También se encuentra a una cierta altura respecto de la base para que sólo estén las piezas o soportes de esta sujeción en contacto con la mesa o lugar de la simulación, el ensamblaje de todas las piezas ha quedado como se puede observar en la figura 8.



Figura 8. Glotis sujeta al soporte (a) y ensamblaje de todas las piezas que conforma el simulador completo final (b)

6. Resultados

El resultado final obtenido tras este trabajo de diseño y fabricación de prototipo de simulador de broncoscopia ha sido el ensamblaje de las piezas diseñadas, evaluadas mediante previa simulación virtual del proceso de broncoscopia e impresas en 3D mediante técnicas de fabricación aditiva por tecnología FDM con material elastómero termoplástico de poliuretano. Con este simulador se ha realizado una serie de experiencias en el hospital, mediante el equipo de broncoscopia mostrado en las figuras 9 y 10, observándose que los usuarios del simulador destacan que el prototipo ha causado muy buena impresión y se evalúa muy positivamente como una herramienta de enseñanza muy útil.



Figura 9. Equipo Ambu^R a ScopeTM 4 Broncho large



Figura 10. Pruebas de broncoscopia en hospital

7. Conclusiones

El prototipo definitivo del simulador de broncoscopia obtenido, tras el análisis de los resultados, mediante un estudio estadístico para el que se usó un modelo de cuestionario de evaluación y grado de satisfacción del usuario, confeccionado para tal fin en este trabajo, se puede considerar válido como herramienta útil y efectiva de formación para facultativos médicos que precisen realizar este tipo de prueba invasiva a modo de entrenamiento antes de someter al paciente a la misma. Y, por tanto, se puede alcanzar de base la conclusión de que este tipo de tecnologías son de valiosa ayuda en el campo de la medicina y concretamente en la obtención de metodologías alternativas para la realización de pruebas diagnósticas y la formación de profesionales de la salud.

Agradecimientos

Agradecer al Servicio de Apoyo a la Investigación Tecnológica (SAIT) de la Universidad Politécnica de Cartagena y a sus técnicos por la ayuda y atención recibida y por poner a disposición de los investigadores participantes en este trabajo, los equipos necesarios para su desarrollo.

Referencias

- [1] R. E. Przekop, M. Kujawa, W. Pawlak, M. Dobrosielska, B. Sztorch, and W. Wieleba, "Graphite Modified Poly lactide (PLA) for 3D Printed (FDM/FFF) Sliding Elements," *Polymers (Basel)*, vol. 12, no. 6, 2020, doi: 10.3390/polym12061250
- [2] R. Melnikova, A. Ehrmann, and K. Finsterbusch, "3D printing of textile-based structures by Fused Deposition Modelling (FDM) with different polymer materials," 2014 GLOBAL CONFERENCE ON POLYMER AND COMPOSITE MATERIALS (PCM 2014), vol. 62, no. Global Conference on Polymer and Composite Materials (PCM). Niederrhein Univ Appl Sci, Fac Text & Clothing Technol, D-41065 Monchengladbach, Germany PU - IOP PUBLISHING LTD PI - BRISTOL PA - DIRAC HOUSE, TEMPLE BACK, BRISTOL BS1 6BE, ENGLAND, 2014, doi: 10.1088/1757-899X/62/1/012018
- [3] G. L. Goh, H. N. Zhang, T. H. Chong, and W. Y. Yeong, "3D Printing of Multilayered and Multimaterial Electronics: A Review," *Adv. Electron. Mater.*, vol. 7, no. 10, 2021, doi: 10.1002/aelm.202100445
- [4] W. Jiang, H. Y. Mei, and S. Y. Zhao, "Applications of 3D Bio-Printing in Tissue Engineering and Biomedicine," *J. Biomed. Nanotechnol.*, vol. 17, no. 6, pp. 989–1006, 2021, doi: 10.1166/jbn.2021.3078
- [5] G. Saccone et al., "Performance Improvement of Piston Engine in Aeronautics by Means of Additive Manufacturing Technologies," *J. Aersp. Eng.*, vol. 34, no. 5, 2021, doi: 10.1061/(ASCE)AS.1943-5525.0001305
- [6] N. Li, D. Qiao, S. Zhao, Q. Lin, B. Zhang, and F. Xie, "3D printing to innovate biopolymer materials for demanding applications: A review & nbsp;," *Mater. TODAY Chem.*, vol. 20, 2021, doi: 10.1016/j.mtchem.2021.100459

Obtención de modelos 3D de nuevos prototipos de hisopos nasofaríngeos mediante la tecnología de impresión 3D en materiales biocompatibles

H. Flores Aparicio^{1,2}, D. Ojados González^{1,2}, J.F. Roca González^{2,3}, C. Cinesi Gómez⁴

¹ Servicio de Diseño Industrial y Cálculo Científico (SEDIC), Universidad Politécnica de Cartagena, Cartagena, España, {hector.flores, lola.ojados}@upct.es

² Laboratorio de Investigación, Desarrollo e innovación de Tecnologías Biomédicas (LIDiTeB), Hospital General Universitario Santa Lucía - Universidad Politécnica de Cartagena, Cartagena, España, {jroca.gonzalez}@upct.es

³ Departamento de Automática, Ingeniería Eléctrica y Tecnología Electrónica, Universidad Politécnica de Cartagena, Cartagena, España

⁴ Servicio de Urgencias, Hospital General Universitario Reina Sofía, Murcia, España, cesar.cinesi{@carm.es}

Resumen

Este estudio nace de la necesidad de dar respuesta a la falta de suministro de material sanitario durante las etapas más críticas de la pandemia de Covid-19. En el panorama mundial de escasez de material para hacer frente a la lucha contra la enfermedad de Covid-19 durante la etapa más agresiva de la pandemia desencadenada por la presencia del virus SARS-CoV-2, surgió la imperiosa necesidad de buscar alternativas de fabricación in situ de material requerido tanto para el tratamiento, como para diagnóstico de pacientes, al igual que ocurrió con elementos de seguridad o equipos de protección individual para personal clínico y de a pie. En el caso base de estudio de este trabajo se precisó en concreto de elementos para toma de muestras a analizar mediante técnicas de PCR, ya que hubo una falta de stock a nivel mundial de este tipo de material sanitario. Por este motivo desde muchos centros de investigación e instituciones se emprendió una búsqueda de soluciones a partir de iniciativas públicas y privadas que permitieran obtener elementos eficaces para la toma de muestras con las que determinar el contagio de este virus. Desde el SEDIC de la UPCT, teniendo en cuenta estudios previos, se obtuvieron diseños de hisopos nasofaríngeos como material de último recurso útil para estos fines, que podían ser fabricados mediante técnicas de impresión 3D.

1. Introducción

La aplicación de la impresión 3D en el campo de la medicina, ganó gran relevancia durante la etapa de la pandemia COVID-19, siendo la impresión 3D de hisopos de los prototipos más fabricados mediante esta tecnología.

Los factores que se dieron durante esa etapa y que fomentaron la inclusión de este sistema de fabricación en la cadena de suministros de material de último recurso se debieron a:

- La escasez durante la etapa de pandemia COVID-19 de suministro de material sanitario esencial sobre todo en lo referente a pruebas de diagnóstico, por lo que la impresión 3D se convirtió en ese momento en una posible solución. En el caso de la fabricación de hisopos, esta tecnología, permitió fabricar una gran cantidad de estos elementos en unas pocas horas, lo que supuso un abastecimiento rápido de estos

elementos, sobre todo a nivel local, en aquellas pedanías que contasen con laboratorios, centros de investigación o fabricas con equipos adaptados a la impresión 3D de material biocompatible.

- Los diseños de los hisopos realizados mediante impresión 3D, debían cumplir con una serie de estándares y requisitos que garantizaran la calidad y la seguridad de las pruebas de diagnóstico, por lo que en los momentos de necesidad, fabricantes y entidades de investigación proporcionaron diseños validados y aprobados para ser usados con técnicas de fabricación aditiva mediante impresión 3D.
- Una de las ventajas principales que promovió este movimiento, fue una fuerte presencia de un entorno colaborativo y con resultados de código abierto. Las comunidades de impresión 3D se unificaron para compartir diseños y conocimientos, estableciendo redes de colaboración en proyectos de código abierto para la producción de hisopos y otros suministros de carácter médico. Esto permitió generar una respuesta rápida y efectiva ante la fuerte crisis de suministros presente en aquella época. [1][2][3]

Se ha de tener en cuenta que, aunque la impresión 3D de hisopos y suministros médicos puede ser una solución útil en situaciones de emergencia, en ningún momento reemplaza la fabricación y regularización tradicional de los dispositivos médicos. Las calidades, esterilización y aprobación son aspectos fundamentales que han de garantizar la eficacia y la seguridad de estos elementos en los entornos médicos.

2. Materiales

Por lo general, los hisopos que son usados para la realización de las pruebas de diagnóstico COVID-19 están compuestos por diferentes materiales, según las diferentes zonas que tiene por diseño.

- Parte de la varilla o asta: esta parte de los hisopos, por norma general está hecha de polipropileno, un material de tipo polímero plástico que confiere

seguridad en aplicaciones de uso médico, pues el polipropileno es un material flexible y fácil de esterilizar, lo cual lo hace óptimo para esta clase de usos.

- Punta del hisopo: la punta de los hisopos puede diferir en tamaño y forma dependiendo de los fabricantes además de que puede estar fabricadas con diferentes materiales dependiendo de su uso específico. Por ejemplo, para los hisopos de pruebas PRC (Reacción en Cadena de la Polimerasa), en donde se recogen las muestras de la parte posterior de la nariz y la garganta, por norma general se usan materiales para las puntas de tipo poliéster o flocadas con nylon. El uso de estos materiales favorece la retención eficiente de las células y las partículas virales para su posterior análisis tras la extracción.

Durante la etapa de pandemia, en la que se dieron problemas de extrema escasez de materiales, la impresión 3D fue solución para la fabricación de hisopos en las regiones con escasez de suministros. Sin embargo, seguía siendo necesario asegurar que los materiales usados para la fabricación de estos hisopos cumplieren una serie de estándares.

Por estos motivos se optó por la fabricación de los hisopos nasofaríngeos a través de la fabricación aditiva mediante impresión 3D de resinas. En este caso en concreto, mediante el uso de impresoras de resinas biocompatibles.

De entre las resinas biocompatibles, la seleccionada para llevar a cabo la fabricación de estos elementos de diagnóstico, es la conocida como Surgical Guide.

La resina Surgical Guide es un tipo específico de resina usada principalmente en el campo de la odontología para la fabricación de guías quirúrgicas dentales, para llevar a cabo procedimientos quirúrgicos de forma precisa y segura, como es el caso de la colocación de implantes dentales.

Como la resina Surgical Guide está destinada a la fabricación de guías quirúrgicas, esta resina debe cumplir con ciertas características para garantizar su eficacia en los procedimientos quirúrgicos, es por tanto lo que la hace favorable también para la impresión de los hisopos para diagnósticos COVID-19.

Las principales características que aporta esta resina son:

- **Transparencia:** en el caso de su uso en ortodoncia, es un factor importante para el dentista, para que pueda ver la anatomía y las estructuras dentales en la guía, lo que le permite una colocación precisa del implante.
- **Precisión:** la resina debe contar con una alta precisión en la impresión 3D para que los modelos impresos puedan ajustarse perfectamente a las anatomías de los pacientes
- **Biocompatibilidad:** es esencial, puesto que las resinas han de ser seguras para el uso que se les va a asignar y no causen reacciones adversas en los tejidos con los que entren en contacto.

- **Estabilidad:** la resina debe contar con estabilidad y resistencia para soportar el proceso al que se someta el elemento impreso, sin deformarse ni romperse.
- **Facilidad de manipulación:** debe permitir una manipulación y colocación sencilla en la impresora, para asegurar que se fabriquen los elementos deseados de forma segura y eficiente.

3. Metodología

El diseño de hisopos para ser fabricados mediante impresión 3D debe considerar una serie de aspectos que garanticen su funcionalidad, seguridad y precisión en la recolección de las muestras usadas para el diagnóstico.

- Se ha de tener en consideración el uso de materiales seguros y adecuados, en este caso en concreto la impresión de hispo con resina Surgical Guide [4].
- Debe considerarse la forma y el tamaño de la punta, factores relevantes en la recolección de muestras y en el caso de hisopos para PCR, la punta debe tener la forma adecuada para llegar a las áreas de la garganta y la nariz de forma efectiva.
- Ha de asegurarse que la punta del hisopo sea estéril para evitar posibles contaminaciones cruzadas entre muestras y mantener la precisión de los resultados.
- La estructura del mango del hisopo deber aportar ergonomía y fácil manipulación hacia el personal sanitario que es el encargado de la realización de la recolección de las muestras.
- Se han de evitar los bordes afilados o superficies rugosas en el diseño para no causar mas incomodidad de la innecesaria sobre el sujeto o paciente durante la realización de la prueba.
- Se han de realizar pruebas y validaciones del diseño para asegurar de que se cumplen con los estándares.
- Por norma general, los diseños de este tipo de hisopos realizados en la etapa de la pandemia COVID-19 se consideraron para su uso a código abierto para que cualquier grupo o persona capaz de trabajar con ellos pudiera contribuir, mejorar y validar su diseño.

3.1. Modelos de referencia

Se parte de una selección de cuatro diseños diferentes de modelos 3D de hisopos para uso médico (Figura 1), destinados a poder ser reproducidos mediante técnicas de impresión 3D, a partir de resinas biocompatibles o aptas para uso médico.

En la selección de estos cuatros modelos de partida, se ha tenido en cuenta, que ya han sido reproducidos en 3D anteriormente, testeados y llevados a su aplicación final con miles de unidades usadas y buenos resultados.

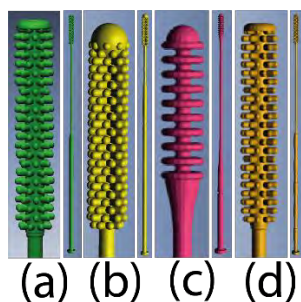


Figura 1. Modelos 3D de hisopos de referencia

Los Modelos dereferencia mostrados en la Figura 1 corresponden a :

- (a) Modelo de hisopo Northwell [5]
Dimensiones 5x5x151,15 mm
- (b) Modelos de la Universidad de Florida (USF) [6].
dimensiones 5x5x154,34 mm
- (c) Modelo hisopo de Valdecillas [7]
Dimensiones 6x6x120,694 mm
- (d) Modelo hisopo FormLabs [8]
Dimensiones 5x5x151,15 mm

3.2. Modelos creados en UPCT

Según se fueron testeando los diferentes modelos con sus modificaciones, se fueron extrayendo resultados y feedback sobre el funcionamiento de cada uno de ellos obteniendo de ello resultados en modificaciones como puede ser la modificación del modelo de Valdecillas readaptado para una mejor manipulación del mango:

Modelo de Valdecillas modificado en dimensiones de mango (Figura 2):

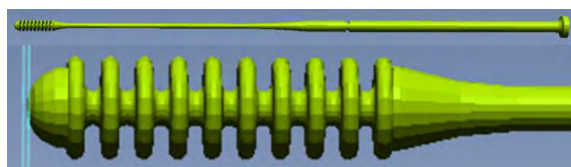


Figura 2. Modificación de hisopo Valdecillas

Dimensiones aproximadas de (6x6x150,69) mm.

O bien con el rediseño del modelo basado en Valdecillas en todas sus dimensiones, con mayor número de elementos de rascado y ensanche del tamaño de la cabeza, aumento de la longitud del mango, dando lugar al modelo que nace de este rediseño modificado:

Modelo UPCT1 (Figura 3):

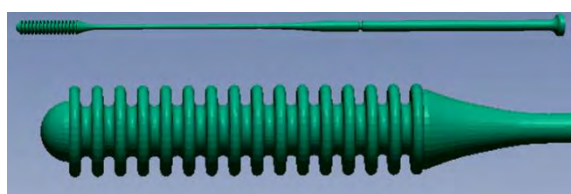


Figura 3. Hisopo UPCT1

Dimensiones aproximadas de (6,05x6,05x150) mm.

Pero no solamente se estuvo trabajando con la modificación y rediseños de hisopos a partir de modelos funcionales ya probados, sino que en paralelo se iban

realizando diseños nuevos que pudieran mejorar la funcionalidad de los hisopos a la hora de la recogida de las muestras.

Las creaciones de estos nuevos diseños se llevaron a cabo mediante el uso del software de diseño SolidWorks, haciendo uso de sus herramientas de diseño y modelado, dando lugar a la generación de croquis con sus acotaciones completas (Figura 4).

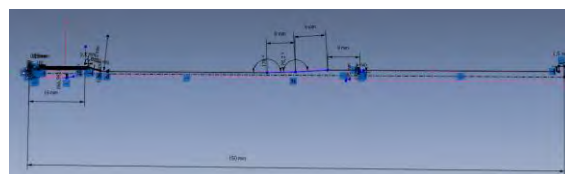


Figura 4. Croquis dimensionado del hisopo UPCT

A partir de estos croquis y con las operaciones de modelado, como la de revolución (Figura 5), se da lugar al modelo sólido, que será el resultado del nuevo modelo de prototipo de hisopo.



Figura 5. Operación de revolución Solidworks

De esta manera, nacen dos nuevos modelos:

Modelo UPCT2 (Figura 6): cabeza trenzada con cavidades.

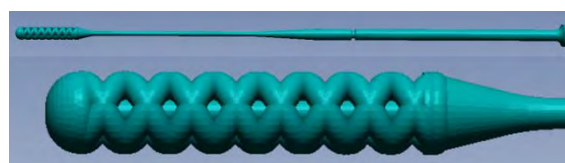


Figura 6. Modelo hisopo UPCT2

Dimensiones aproximadas de (6,04x6,04x150) mm.

Modelo UPCT3 (Figura 7): cabeza doble helicoidal.

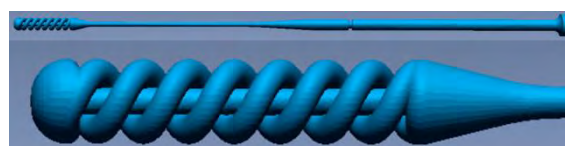


Figura 7. modelo hisopo UPCT3

Dimensiones aproximadas de (6,04x6,04x150) mm.

Mediante un cuerpo de revolución apoyado en el diseño antecesor (Figura 8).



Figura 8. Sección bastoncillo hisopo UPCT

Sobre el que se ha generado la nueva cabeza para el hisopo a través del uso de nuevas operaciones desde este punto de partida (Figura 9).

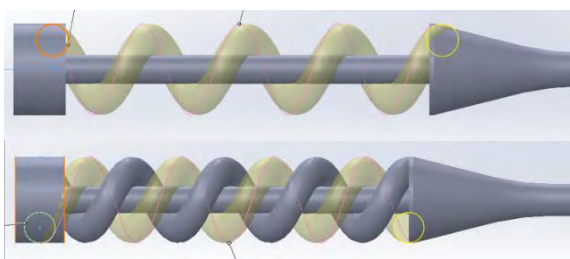


Figura 9. Operaciones modelado cabeza de hispo

Nuevamente a través de las pruebas, aparece a partir del modelo de hispo UPCT2, una modificación conocida como: Modificación UPCT2, modelo no impreso con menor número de revolución de hélice para la cabeza de trenza (el cual no llegó a ser impreso, Figura 10).



Figura 10. Modificación del modelo UPCT2

Dimensiones aproximadas de (6,04x6,04x150) mm.

4. Resultados

Fabricación y obtención de los prototipos físicos.

Con los diferentes modelos creados con SolidWorks, se procede a la generación de los archivos de cada uno de los prototipos de hispos en formato stl, necesario para su uso en la fabricación mediante impresión 3D en resina.

Se importa el archivo stl resultado del prototipo al software laminador de la impresora 3D, colocando el modelo sobre la plataforma de impresión en la orientación adecuada y configurando los parámetros óptimos para su impresión (Figura 11).

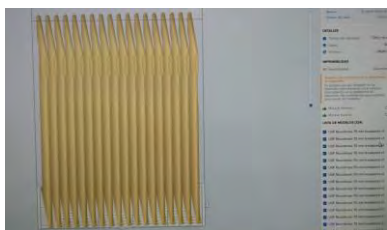


Figura 11. Configuración impresión 3D Form 3B [9]

Para la configuración de la resolución de la capa y del tipo de resina, para este caso, se usan los indicados para la resina Surgical Guide, de carácter biocompatible. Configurado todo, se envía la orden de impresión a la impresora 3D (Figura 12).

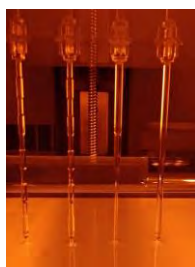


Figura 12. Impresión 3D hispos

Terminada la impresión se pasan los modelos con la bandeja de impresión a la máquina de lavado mediante alcohol isopropílico. Unos 20 minutos de limpieza.

Tras la limpieza, se pasan los modelos al horno de curado, unos 30 minutos a 60 grados centígrados (Figura 13).

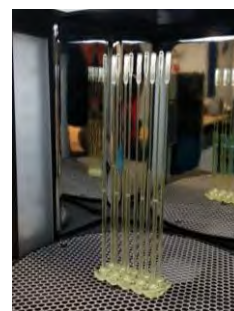


Figura 13. Curado de hispos tras fabricación

Tras el curado, los hispos quedan correctamente impresos y listos para su utilización.

Como dato para la impresión según el tamaño de la bandeja de la impresora usada, modelo Form 3B, se pueden obtener unos 324 hispos, tardando unas 135 HORAS con un gasto de material en resina aproximado de 246 gramos.

Agradecimientos

Al Departamento de microbiología del Hospital Reina Sofía de Murcia, a Rosa Blázquez, al SMS, UPCT y todas las personas, grupos de investigación y entidades que lo hicieron posible y que formaron parte del entorno colaborativo con sus pequeñas y grandes aportaciones.

Referencias

- [1] Página web de 3Dnatives (3D natives: abril 3, 2020) <https://www.3dnatives.com/es/empresas-crean-hispos-pruebas-covid-19-03042020/>
- [2] Página web de 3Printig Design (3D Printing Design: abril 24, 2020) <http://www.3dprintingdesign.es/es/noticia/alianza-para-la-creacion-de-hispos-nasofaringeos-impresos-en-3d-para-la-prueba-covid-19>
- [3] Página web formlabs Cómo ROE Dental Laboratory ha pasado a fabricar suministros médicos para combatir la COVID-19 (Formlabs: mayo 13, 2020) <https://formlabs.com/es/blog/roe-dental-laboratory-fabrica-suministros-medicos-combatir-covid-19/>
- [4] Página web Surgical Guide Resign (Formlabs 2023) <https://formlabs.com/es/materials/dental/#surgical-guide>
- [5] Página web Northwell Health (Northwell 2023) <https://www.northwell.edu/?locale=es>
- [6] Página web repositorio USF Nasal Swab (USF Nasal Swab por Jonathan Ford: mayo 21, 2020) <https://usf.app.box.com/s/wxmlj0r66vp8bzei6o7sur1kq1jr8oli>
- [7] Página web del Hospital Virtual Valdecilla <https://www.hvvaldecilla.es/proyectos/recogidas-de-muestras-hispos/> (Hospital Virtual Valdecilla 2020).
- [8] Página web de Formlabs (FORMLABS 2023) <https://formlabs.com/es/covid-19-response/covid-test-swabs/>
- [9] Página web Form 3B (Formlabs 2023) <https://formlabs.com/es/3d-printers/form-3b/>

Adaptación de pieza para tubos de respirador: tubuladora segura para la ventilación no invasiva

D. Ojados González^{1,3}, J. Roca González^{1,2}, C. Cinesi Gómez⁴, H. Flores Aparicio^{1,3}

¹ Laboratorio de Investigación, Desarrollo e innovación de Tecnologías Biomédicas (LIDiTeB), Hospital General Universitario Santa Lucía - Universidad Politécnica de Cartagena, Cartagena, España, {lola.ojados, jroca.gonzalez, hector.flores}@upct.es

² Departamento de Automática, Ingeniería Eléctrica y Tecnología Electrónica, Universidad Politécnica de Cartagena, Cartagena, España

³ Servicio de Apoyo a la Investigación Tecnológica (SAIT), Universidad Politécnica de Cartagena, Cartagena, España

⁴ Servicio de Urgencias, Hospital General Universitario Reina Sofía, Murcia, España, {cesar.cinesi}@carm.es

Resumen

El estudio llevado a cabo por el Servicio de Diseño Industrial y Cálculo Científico SEDIC del Servicio de Apoyo a la Investigación Tecnológica SAIT de la Universidad Politécnica de Cartagena UPCT, responde a la necesidad de disponer de conexiones de tipo estándar para equipos de ventilación asistida. Esta demanda parte de personal médico, que en periodo de pandemia (Covid-19) identificó un déficit en cuanto a falta de stock de estas válvulas de conexión y un problema en cuanto a la ausencia de estandarización que las hicieran útiles para cualquier tipo de ventilador fuera de la marca y modelo que fuera. Se ha llevado a cabo para ello, el diseño y obtención de un prototipo, mediante técnicas de fabricación aditiva con materiales biocompatibles, de una pieza de adaptación para tubos de respirador para la ventilación por Terapia de presión positiva continua en las vías respiratorias (CPAP), que además ofrece la ventaja de mejorar los requerimientos que son de aplicación para el tratamiento de enfermos que necesitan ventilación asistida a la vez que posibilitar el suministro de medicación desde la misma válvula. El prototipo final ha sido revisado por especialistas clínicos del Sistema de Salud y propuesto para su evaluación mediante investigación clínica.

1. Introducción

Puesto en marcha el decreto de estado de alarma, en marzo de 2020, debido a la pandemia de Covid-19 derivada de la aparición del virus SARS-CoV-2, surgió la necesidad de realizar estudios para obtener nuevos prototipos de material sanitario que poder fabricar a nivel local y supliera la falta de recursos de los centros sanitarios y hospitalarios. En el caso particular estudiado en este trabajo, se pretendía mejorar los requerimientos de aplicación para el tratamiento de pacientes afectados en ese momento por la enfermedad de Covid-19 que necesitaban ventilación a la vez que suministro de medicación.

Atendiendo a las necesidades expresadas por el Servicio Murciano de Salud (SMS) de contar con dispositivos para el tratamiento de oxigenoterapia de pacientes en los primeros estadios de la enfermedad del Covid-19, los esfuerzos realizados por el equipo de investigadores de la Universidad Politécnica de Cartagena (UPCT), se centraron en el desarrollo de adaptadores, válvulas y accesorios de material sanitario para la ventilación por Terapia de presión positiva continua en las vías respiratorias (CPAP).

Entre los problemas detectados para la aplicación de los dispositivos existentes en pacientes afectados por Covid-19, el más preocupante era el de la propagación del virus a través del aerosol que se produce en la fase espiratoria del ciclo respiratorio.

Con el fin de resolver este problema, los investigadores de la Universidad propusieron el uso de máscaras faciales completas, descartando ya el uso de las nasales, y la inclusión de filtros antibacterianos/antivirales. En este sentido, el trabajo se orientó en el diseño y perfeccionamiento de un elemento necesario para conectar dicha máscara al respirador.

Se diseñó una interfaz que permite la conexión de los elementos del circuito de paciente habitualmente utilizados en clínica, posibilitando la administración de terapias CPAP a pacientes afectados por Covid-19.

El diseño se ha perfeccionado hasta en tres iteraciones, dando como resultado la confección de un prototipo mediante impresión SLA, empleando una resina con biocompatibilidad certificada, capacidad de interconexión de los elementos utilizados con los equipos CPAP disponibles en el SMS de distintos fabricantes, posibilidad de conexión de filtro antiviral en la línea de evacuación espiratoria y disponibilidad de punto de conexión de la línea para la toma de muestras del analizador de gases, fueron los requisitos de diseño.

Como resultado, el prototipo final fue aceptado por el especialista clínico del SMS, el Dr. César Cinesi y propuesto para su evaluación mediante investigación clínica, una vez autorizada esta por el Comité Ético de Investigación con medicamentos (CEIm), del Servicio Murciano de Salud.

Superada la aprobación del Comité Ético, el siguiente paso es la solicitud de validación por parte de la AEM Agencia Española de Medicamentos. Superar esta fase posibilitaría la fabricación del producto a demanda.

2. Recursos empleados

El proceso de toma de mediciones se ha llevado a cabo mediante la aplicación de técnicas de digitalización 3D

como el escaneado 3D, siendo complementado con tomas de medidas con aparatos de precisión manual.

El tratamiento del resultado del escaneado 3D (nube de puntos y malla), se ha realizado mediante el uso del software de ingeniería inversa Geomagic Design X.

En cuanto a lo referente al modelado y diseño de los prototipos desarrollados, se ha optado por el uso de SolidWorks, software especializado en este campo.

Para la fabricación de los prototipos realizados, se han empleado técnicas de fabricación aditiva o impresión 3D, como la Fabricación por Fusión de Filamento (FFF) para los prototipos iniciales, y la tecnología de estereolitografía (SLA) mediante láser de baja potencia (LFS) para la fabricación en resina biocompatible de los prototipos finales con carácter funcional.

3. Punto de partida

El punto de partida para el rediseño adaptable de la pieza se encuentra en los modelos comerciales (Figura 1) y bajo estándar para sistemas de respiración asistida y tubos de respiración. [1]



Figura 1. Mascarilla con adaptadores CPAP

Generalmente estos modelos están conformados por el ensamblaje de un par de piezas, en este caso una en forma de cono en “L” de color azul transparente y otra en forma de “T” incolora y transparente [2][3]. (Figura 2)



Figura 2. Modelos comerciales codos CPAP

Partiendo de estas piezas y mediante la toma de mediciones y técnicas de escaneado 3D para aplicar los conceptos de ingeniería inversa, se ha generado un primer modelo

basado en las dimensiones aproximadas del ensamblaje, comenzándose a trabajar en primer lugar con pieza en “L”. (Figura 3)



Figura 3. Modelado de codo. Pieza en “L”

Y posteriormente aplicando las mismas técnicas a la pieza transparente en forma de “T”, para obtener como resultado otro modelo con las dimensiones similares y listo para comenzar a trabajar con diferentes configuraciones de las modificaciones propuestas para el diseño. (Figura 4)



Figura 4. Modelado de conexión. Pieza en “T”

4. Primeras modificaciones según los requisitos de diseño propuestos

Se ha comenzado a trabajar con la parte superior de la primera pieza añadiendo una primera modificación lateral para el acople de un nuevo conducto para inyectables. (Figura 5)



Figura 5. Modificación de diseño para pieza en “L”

Respecto a la pieza de conexión en “T”, se ha decidido no realizar ninguna modificación inicial. Por lo que el resultado del conjunto es el mostrado en la Figura 6.

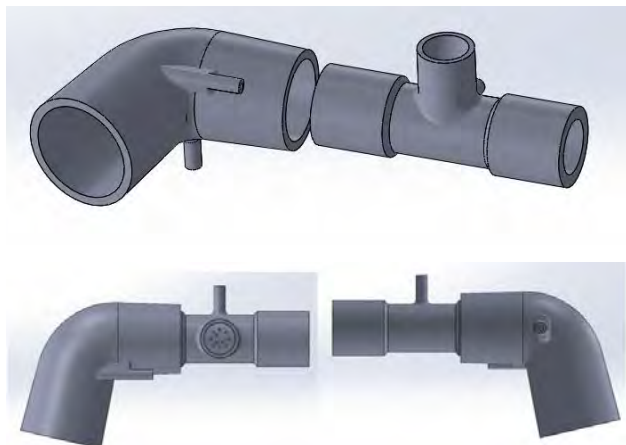


Figura 6. Modelado del ensamblaje tras primera modificación

Para la comprobación de este primer resultado de diseño inicial se ha realizado su impresión 3D con material PLA con la impresora 3D Ender 3 Pro disponible en el SEDIC/SAIT de la UPCT, por ser un proceso rápido y económico que ha permitido verificar de la geometría y dimensiones de las piezas del conjunto antes de continuar con otras modificaciones que se consideren necesarias y de esta forma asegurar la obtención de un modelo correcto antes de proceder con la fabricación de un prototipo optimizado final, cuyo coste puede ser más elevado al emplear para que sea funcional otra tecnología y otro tipo de materiales. (Figura 7)



Figura 7. Prototipo en PLA del ensamblaje del adaptador

5. Segunda iteración de diseño del conjunto

La posterior mejora planteada sobre el modelo 3D de ensamblaje inicial obtenido por modificación de diseño a partir de la solución comercial, e impreso en PLA, ha consistido en la unificación en una sola pieza de las dos piezas ensambladas que componen el diseño del adaptador, esto se ha conseguido haciendo uso de herramientas CAD, en concreto se ha usado el software SolidWorks, en su módulo de diseño mecánico aplicando diferentes operaciones de modelado para obtener el diseño de un modelo 3D de una sola pieza del adaptador con la misma funcionalidad del conjunto de dos piezas de partida. (Figura 8)

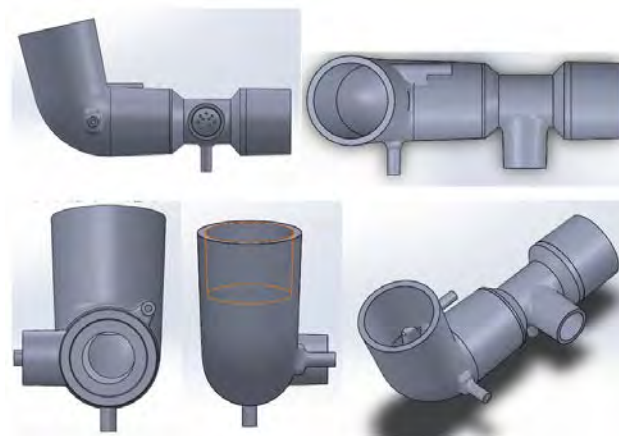


Figura 8. Modelo 3D de pieza única de conexión

Del mismo modo que en el paso anterior, el resultado de esta segunda iteración de diseño ha sido fabricado mediante impresión 3D con material PLA en la misma impresora para verificar que esta modificación cumple con los requisitos planteados. (Figura 9)



Figura 9. Prototipo en PLA de la pieza única

Superadas las pruebas geométricas y dimensionales de este prototipo en PLA, se ha procedido a la obtención de un primer prototipo de carácter funcional mediante impresión 3D con tecnología de estereolitografía láser de baja fuerza (LFS) con el equipo Form 3 de Formlabs, haciendo uso de la resina biocompatible “Surgical Guide”. (Figura 10)



Figura 10. Prototipo en material biocompatible

6. Resultados

Para la optimización del prototipo, tras las primeras pruebas de funcionamiento exitosas sobre la segunda iteración, se ha mejorado el diseño del adaptador mediante la incorporación de una toma adicional de inyectable para suministro de medicamentos, se ha planteado como preferencia que esta toma esté ubicada en la parte superior del prototipo. Como resultado ante este requisito, se han obtenido dos opciones posibles para el prototipo final.

Una primera solución, ha sido diseñada con la toma del inyectable incrustada internamente. (Figura 11)

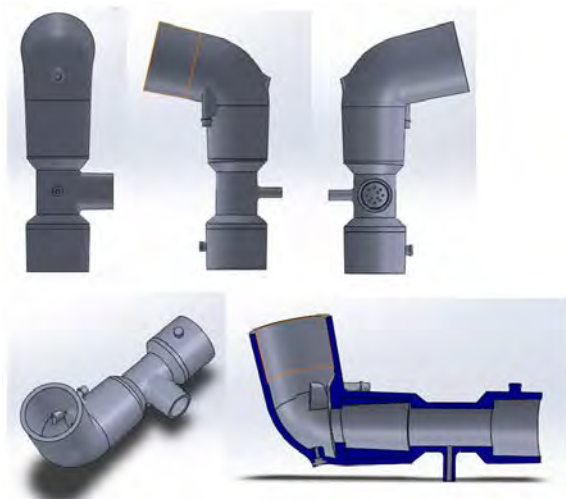


Figura 11. Modelo 3D del prototipo con toma inyectable interna

Pero también se realizó una segunda versión con la toma para inyectable exterior. (Figura 12)



Figura 12. Modelo 3D del prototipo con toma inyectable externa

Como mejora adicional a ambos prototipos finales se les ha dotado de tapones para proteger las diferentes vías de suministro de medicamentos. (Figura 13)

Una vez obtenido el prototipo final (Figura 14), se ha procedido con el siguiente paso, la aprobación del Comité Ético, que tiene que ver con el proceso administrativo de validación de cualquier nuevo producto sanitario, y posteriormente es preciso cursar la solicitud de validación por parte de la AEM, Agencia Española de Medicamentos. Superar esta fase conllevaría la fabricación aditiva mediante impresión 3D del producto a demanda en casos de necesidad.

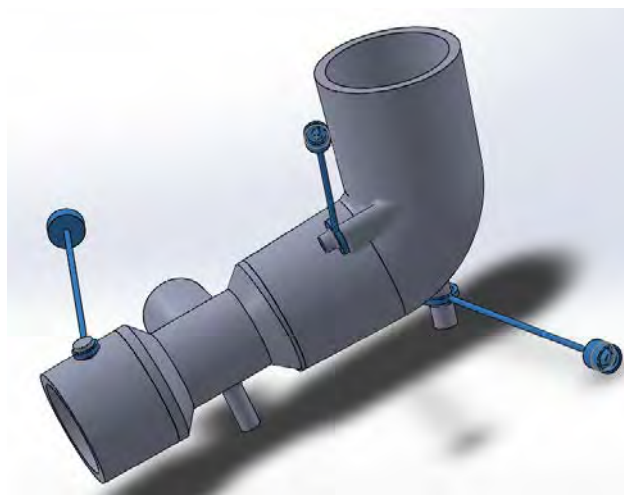


Figura 13. Modelo 3D de prototipo final con elementos de sellado

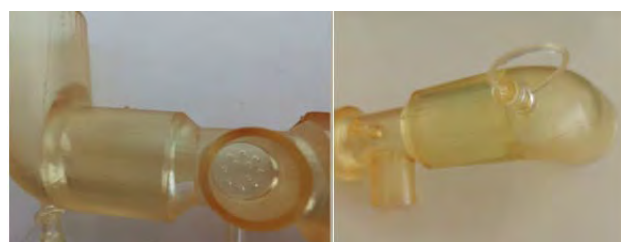


Figura 14. Prototipo impreso en resina biocompatible para comprobación de sellado y pruebas funcionales

7. Conclusiones

Validar el prototipo por la AEM, Agencia Española de Medicamentos, y poder realizar la fabricación aditiva del producto a demanda supondría un gran avance para el sistema sanitario nacional si se tiene en cuenta que se podría disponer de todo este desarrollo, preparado para posibles eventualidades y emergencias futuras como la propiciada por la Covid-19, pero también un gran avance para la sociedad. Y es que disponer de un producto de estas características, validado por la Agencia de Medicamentos, tras un proceso de investigación conjunta y colaborativa entre entidades educativas, sanitarias y empresas privadas sería todo un logro a nivel nacional.

Referencias

- [1] Página web: Suministros Propato (Propato Hnos. S.A.I.C, 2023).
<https://propato.com.ar/producto/mascara-vni-desc-oron-codo-se-arns-strap-m-af531/>
- [2] Página web: Material sanitaria Philips Healthcare (Koninklijke Philips N.V., 2004 - 2023).
<https://www.philips.es/healthcare>
- [3] Repositorio web: Philips, Meeting the big challenge of little faces (Koninklijke Philips N.V., 2004 - 2023)
https://www.documents.philips.com/doclib/enc/9027750/Pediatric_Mask_Brochure_EMAIL_8_12.pdf

Asiento o lecho de posición infantil personalizado mediante técnicas de escaneo 3D y fabricación aditiva

H. Flores Aparicio^{1,2}, D. Ojados González^{1,2}, G. Salcedo Eugenio²

¹ Laboratorio de Investigación, Desarrollo e innovación de Tecnologías Biomédicas (LIDiTeB), Hospital General Universitario Santa Lucía - Universidad Politécnica de Cartagena, Cartagena, España, {hector.flores, lola.ojados}@upct.es

² Servicio de Apoyo a la Investigación Tecnológica (SAIT), Universidad Politécnica de Cartagena, Cartagena, España {gustavo.salcedo}@upct.es

Resumen

El trabajo plasmado en este artículo tiene su origen en la solicitud de un sujeto en edad infantil con capacidades especiales, para buscar una mejora que le ayude a cubrir la necesidad de conseguir obtener un lecho de posición o asiento ergonómico, adaptado a su constitución, que le permita desarrollar sus labores de carácter cotidiano, dentro de sus limitaciones y le ayude a adoptar una postura correcta evitando posibles desviaciones en la columna provocadas por malos hábitos y de esta manera ayudar a favorecer el desarrollo correcto de la columna durante las etapas de crecimiento.

1. Introducción

Los lechos o asientos de posición infantil son un útil valioso sobre todo en aquellos sujetos que están en fase de crecimiento con limitaciones de movilidad y permiten garantizar la seguridad, comodidad y el desarrollo adecuado de los mismos.

Es por tanto que este tipo de utensilios deben estar diseñados con algunos requisitos que le confieran cualidades necesarias como son:

Ergonomía: La posición debe ser ergonómica y apoyar su desarrollo adecuado. El asiento debe proporcionar un soporte adecuado para la columna vertebral y el cuello en las fases crecimiento.

Materiales: Los materiales deben adecuados, seguros, no causar molestias al usuario y libres de productos químicos dañinos. En caso necesario, se puede implementar un acolchado adecuado para proporcionar soporte y comodidad.

Adaptabilidad: debe ser ajustable para adaptarse al crecimiento y desarrollo del sujeto. Esto puede ser solventado con la capacidad de ajustar la altura, el ángulo de inclinación y el tamaño del asiento.

Limpieza fácil: el lecho o asiento debe ser fácil de limpiar. Los materiales deben ser lavables y resistentes a las manchas ya que los sujetos en fases de crecimiento, sobre todo si están en edades tempranas, suelen ser propensos a ensuciar con facilidad.

Durabilidad: debe ser duradero y resistente, ya que será utilizado a lo largo de las etapas de crecimiento con alta actividad del sujeto.

Facilidad de uso: son también usuarios del lecho o asiento las personas que ofrezcan su apoyo o ayuda al sujeto, por tanto, deben poder usar y ajustar el lecho o asiento de manera fácil y rápida. Los mecanismos de ajuste y sujeción deben ser intuitivos y seguros.

A pesar de todos estos puntos mencionados, se ha de tener en cuenta que en esta clase de útiles de lecho o asiento adaptable al sujeto, se dependerá sobre todo de las necesidades específicas de cada usuario, por lo que aparte de intentar cubrir el máximo de estándares posibles, es recomendable investigar y buscar diferentes opciones apoyándose en los consejos de los padres o tutores de los sujetos, así del personal especializado o sanitario que les preste asistencia, para así poder cubrir las necesidades más específicas y personales de cada sujeto. [1]

2. Materiales

En lo referente a los materiales, para la elección del material para el prototipo definitivo, se ha de tener en consideración que al optarse por un proceso de fabricación aditiva mediante impresión 3D, según el tipo de tecnología de aditiva, se van a tener a disposición un abanico limitado de materiales con los que trabajar.

En este caso en concreto, se optó por el uso de un proceso de fabricación aditiva mediante impresión 3D de termoplásticos conocida como FFF (Fused Filament Fabrication) también conocida como FDM (Fused Deposition Modeling).

Este proceso de fabricación implica el alimentar el equipo de impresión 3D mediante un filamento termoplástico (de entre los más comunes el ABS o el PLA) a través de un cabezal de impresión caliente a un cierto rango de temperatura, que depende del material empleado y que lo reblandece, a su estado viscoso para que pueda fluir mientras el cabezal se mueve por las coordenadas específicas depositando así el material capa por capa hasta construir el objeto deseado.

La ventaja del uso de la tecnología FFF, es que está ampliamente desarrollada, extendida y relativamente simple, lo que le permite alcanzar una buena versatilidad y obtener bajos costes de fabricación. Es muy útil para los procesos que requieren de la fabricación de prototipos, elemento o piezas personalizadas e incluso piezas o productos finales en ciertos sectores. (Figura 1) [2]

Es por tanto que, seleccionada la tecnología de fabricación, se accede al abanico de los posibles materiales limitados con los que se pueden trabajar en esta tecnología, y por características y uso que se le va a conceder al prototipo con él se va a trabajar, se van a reducir a tres posibles candidatos:

- **PLA (Poly-Lactic Acid):** es un material de carácter biodegradable, pues es capaz de descomponerse de forma natural haciéndolo menos perjudicial para el medio ambiente que otros plásticos. Presenta facilidad de impresión debido a las bajas temperaturas de extrusión. Presenta poca deformación. Tiene una resistencia moderada para muchas aplicaciones. Además, hay una amplia gama de colores para imprimir en 3D con este material y el olor durante el proceso de impresión es muy reducido.
- **ABS (Acrylonitrile Butadiene Styrene):** es un material con muy buena resistencia y tenacidad, por lo que es conocido por su resistencia a impactos, durabilidad y resistencia. Presenta una buena tolerancia a la temperatura en comparación con otros filamentos como por ejemplo el PLA. Es un poco más complicado su uso en cuanto al proceso de impresión 3D y presenta olor, por lo que es importante trabajar en una atmósfera ventilada. Es menos biodegradable y también presenta un amplio abanico de colores para imprimir.
- **PETG (Polyethylene Terephthalate Glycol):** presenta muy Buena resistencia y tenacidad al igual que el ABS. Posee una mayor flexibilidad que el PLA por lo que es capaz de absorber mejor los impactos. En cuanto a su uso en el proceso de fabricación mediante impresión, es más sencillo que el uso de ABS, pero más complejo que el PLA. Tiene buena tolerancia a las temperaturas. En cuanto a la gama de colores, actualmente presenta más limitaciones.

El material definitivo para la realización del prototipo quedó fijado en el uso de ABS por su resistencia y tenacidad ante impactos, lo cual es excelente para el uso que se le va a dar, porque el sujeto para el cual el lecho está diseñado está en la etapa infantil, por lo que presenta bastante energía, lo que favorece que el asiento pueda sufrir impactos mientras se hace uso de él. Además, el ABS presenta buena resistencia a las temperaturas, lo que permite que el asiento pueda quedarse expuesto a la intemperie, bajo condiciones de fuerte sol, sin sufrir deformaciones. Adicionalmente, este material permite ser impreso en varios colores, permite ser pintado, permite el uso de adhesivos y ser mecanizado en caso de que requiera de modificaciones posteriores.

El uso del PETG se contempló como alternativa por presentar características similares al ABS, aunque tiene una peor respuesta ante post-procesados de pintura, uso de adhesivos y mecanizado.

En cuanto el PLA, es opción más desfavorable ante los dos anteriormente mencionados, pero puede presentar una mejor opción económica y al tener la mayor facilidad de impresión, permite realizar prototipos más rápidos y facilitar la impresión de posibles prototipos con

modificaciones, que también sean funcionales, aunque menos duraderos a exposiciones a intemperie e impactos.

3. Metodología

El uso de la técnica de escaneo 3D para la adquisición de la postura ergonómica del individuo con capacidades especiales, constituye un método muy ágil y versátil a la hora de obtener los datos de forma rápida y con precisión, obteniendo como resultado un conjunto de información digitalizada en forma de nube de puntos o de malla que puede ser post- procesada y usada para la obtención de los modelos superficiales ergonómicos, la verificación y corrección de ángulos de posicionamiento y la personalización del prototipo.



Figura 1. Lecho fabricado tradicionalmente frente a lecho fabricado por impresión 3D

Todo lo referente al postprocesado, es llevado a cabo mediante el uso de software de modelado y diseño 3D con el tratamiento de los datos obtenidos mediante las nubes de puntos y mallas, permitiendo trabajar y generar las diferentes superficies de contacto asiento-individuo, que actúen como ayuda a mantenerlo en la postura correcta para evitar que sufra dolores o desviaciones ante posibles malas posiciones o hábitos de postura durante el desarrollo de sus actividades.

En este caso, la aplicación de las técnicas de escaneo 3D nos permiten trabajar en dos posibles líneas de trabajo.

3.1. Línea de trabajo 1

El flujo de trabajo con esta línea tiene como origen el lecho fabricado en escayola mediante el moldeo de la escayola sobre el cuerpo del sujeto usando así el método tradicional de fabricación para este tipo de asientos.

Sobre este modelo de escayola se procede aplicar las técnicas de digitalización para obtener como resultado un modelo de malla totalmente cerrado que es igual tanto en tamaño como en forma al asiento fabricado en escayola, lo que generalmente se conoce como la obtención del gemelo digital.

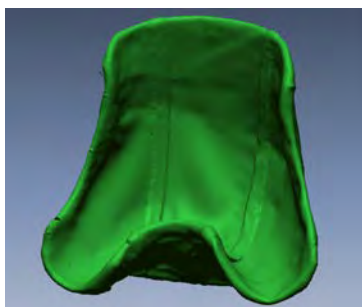


Figura 2. Resultado del escaneado 3D del lecho

Este gemelo digital puede ser usado como elemento de documentación dentro del propio archivo del sujeto, como elemento para realizar estudios o comprobaciones sobre las superficies de contacto, ángulos de inclinación de las superficies, redondeos de curvaturas, elementos de ergonomía, etc. O bien para fabricar réplicas similares en caso de que el asiento de escayola quede destruido o se extravíe, sin necesidad de volver a tener que fabricar nuevamente otro sometiendo al sujeto al proceso de moldeado de escayola. (Figura 2)

Este gemelo, también puede ser usado para comenzar a trabajar en diferentes prototipos mejorados y actualizados, manteniendo las superficies de apoyo ergonómicas que son las que corrigen la postura del sujeto durante el uso del asiento. El abanico de opciones que se abren a la hora de trabajar las actualizaciones y mejoras de este tipo de asiento son enormes, desde la creación de un nuevo modelo con superficies más optimizadas para la transpiración, el uso de accesorios para el acople de reposabrazos, elementos de prolongación o sujeción de apoyo para la cabeza, acople de cinchas para poder movilizar mejor al sujeto en el asiento, etc.

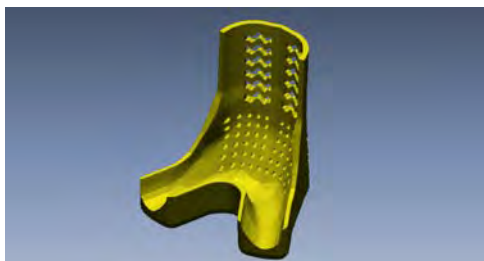


Figura 3. Prototipo de lecho a partir del modelado 3D

Todos estos prototipos o modificaciones, totalmente compatibles con los actuales procesos de fabricación aditiva mediante impresión 3D, que permiten la obtención directa de prototipos funcionales con la ventaja adicional del bajo coste. (Figura 3)

3.2. Línea de trabajo 2

Es la óptima ya que no solamente aporta una mejora al proceso de obtención del asiento ergonómico, sino que mejora la calidad del sujeto al minimizar el tiempo que ha de estar en la posición para la toma de mediciones y sin la necesidad de tener que realizar un molde de escayola, por lo que desaparece el tiempo de espera de secado del molde durante el cual el sujeto ha de estar inmóvil para no deformarlo durante el proceso.

El punto de partida de esta línea comienza cuando el sujeto se posiciona sobre una camilla, boca abajo, con las piernas flexionadas en “L”, adoptando la postura natural que adquirimos al usar un asiento y apoyándose en una base que le ayude a aguantar su peso mientras se realice el proceso de medición.



Figura 4. Resultado del escaneado 3D de la postura del sujeto

Cuando el sujeto se encuentra totalmente asentado en la posición correcta, se comienza con el proceso de digitalización de toda la parte posterior del cuerpo que actúa directamente sobre el asiento realizado a medida. Abarcando la captura de los datos desde la cabeza, parte más alta, hasta la región un poco más debajo de las corvas, parte posterior de las rodillas, asegurando que los ángulos de flexión de esa postura son los debidos. (Figura 4)

El proceso de digitalización de esta región, haciendo uso de un escáner adecuado, en este caso un escáner de luz estructurada EVA de Artec, es bastante ágil, durando unos pocos minutos. Tras la finalización de los barridos necesarios, el sujeto ya puede moverse y la captura de datos realizada quedaría almacenada en un proyecto del que se puede extraer la nube de punto o la malla de las zonas lumbares y de contacto con el asiento a desarrollar. [3][4]

Es entonces cuando comienza el trabajo computacional, en el que se lleva a cabo el post- procesamiento preliminar de la malla obtenida, en este caso mediante el software Artec Studio, propio del Escáner usado y su exportación a otros softwares de modelado más especializados.

Concretamente, esta primera malla post- procesada se exporta a Geomagic Design X, software de ingeniería inversa, que permite realizar un mejor seccionamiento de la malla y procesamiento para reducir las superficies a las necesarias para comenzar con el proceso de modelado superficial.

Con la malla totalmente preparada, con este mismo software se puede comenzar con las primeras etapas del modelado del asiento, hasta obtener como resultado una estructura principal del asiento con la ergonomía de las superficies de contacto de los respaldos ya creadas.

Con esta base creada, a fin de poder elaborar un prototipo mucho más adaptable y con elementos más optimizados, la base obtenida se exporta al software SolidWorks, software de modelado y diseño 3D con carácter industrial y que trabajar el desarrollo del prototipo realizando operaciones de diseño más complejas e incluso trabajar con ensamblajes para crear diferentes partes o mejoras que puedan ser modulares. (Figura 5)

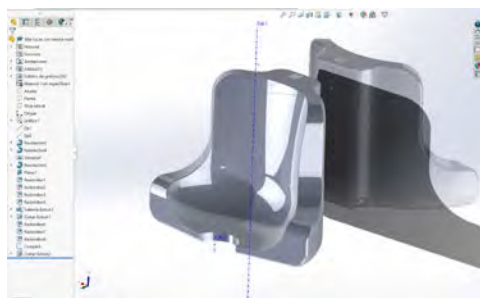


Figura 5. Prototipo de lecho en Solidworks

Con SolidWork se obtienen los primeros diseños para el prototipo y mediante un feedback por parte del sujeto, se establece la adición de elementos modulares de mejora, como son un distanciador para la zona de la entepierna y un elemento para servir de punto de apoyo para la cabeza. De esta manera se incorporan las primeras mejoras al diseño, dando lugar a un primer prototipo funcional.

4. Resultados

El uso de las técnicas de escaneado 3D como complemento para la toma de mediciones de un sujeto para la fabricación de asientos ergonómicos combinado con las técnicas de diseño 3D y de fabricación aditiva, muestran que hoy en día son capaces de obtener resultados mediante prototipos totalmente funcionales y que son posibles de obtener a un coste reducido.

El sujeto presenta una mejor adaptabilidad al asiento además de generar una mayor seguridad con su uso, ya que al sustituir el material de fabricación el prototipo resultante gana considerablemente robustez y resistencia frente a impactos.

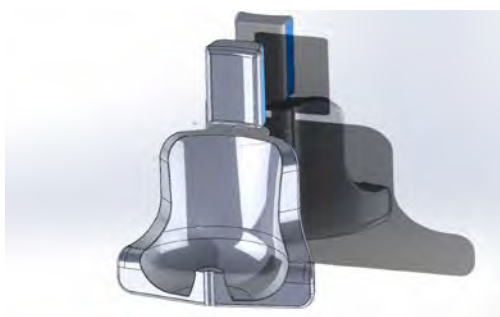


Figura 6. Prototipo de lecho mejorado

El trabajar con este tipo de asientos permite también la posibilidad de añadirle varias opciones de acolchado o incluso pueden plantearse mejoras para incluir elementos tapizados.

Al usar directamente las superficies obtenidas desde el escaneado 3D, el nivel de adaptabilidad del respaldo a la forma del cuerpo es óptima. (Figura 6)

Al tener digitalizadas las superficies, estas pueden trabajarse sobre el software para realizar las modificaciones oportunas sobre estas antes de fabricar un prototipo funcional, a diferencia que, con el método tradicional de obtención de un asiento de escayola, sobre el que se trabaja directamente con el modelo físico obtenido del moldeo sobre el sujeto.

5. Conclusiones

El uso de las nuevas tecnologías aplicadas en este campo y sobre todo con sujetos de necesidades especiales, favorece la comodidad del sujeto a la hora de participar en las etapas del diseño para su asiento adaptado, sobre todo en las etapas iniciales en las que se requiere de el que aguante en ciertas posturas de forma inmóvil durante unos tiempos de espera, al ser reducidas notablemente al orden de pocos minutos, en vez de unos 20min o 30min que se requieren con los métodos tradicionales que siguen siendo aplicados.

Así mismo, las técnicas de digitalización suponen una ventaja a la hora de adquirir la información sobre los sujetos, ya que toda la información almacenada puede ser guardada para consultas posteriores, para formar parte de estudios, para replicar procesos o modelos, etc. (Figura 7)



Figura 7. Verificación de la superficie de contacto

Los costes para la fabricación de los prototipos obtenidos mediante este flujo de trabajo son relativamente bajos, ya que no requieren de la implicación de numerosos procesos de fabricación ni de la implicación de un número muy elevado de departamentos o grupos de trabajo altamente especializado.

Es un proceso cómodo que puede repetirse cómodamente en caso de sujetos que estén en estas de crecimiento o desarrollo.

Referencias

- [1] Página web. Asiento pélvico niño. <https://fiixit.es/asiento-pelvipedico-impreso-en-3d/> (Consultada: septiembre 2023)
- [2] Página web de interempresas. AJIU silla para niña con necesidades especiales. <https://www.interempresas.net/Fabricacion-aditiva/Articulos/486502-Mas-200-empresas-descubren-ventajas-tecnologia-fabricacion-aditiva-mano-Formlabs-Grupo.html> (Consultada: septiembre 2023)
- [3] Página web Artec3D. Escáner eva. <https://www.artec3d.com/es/portable-3d-scanners/artec-eva> (Consultada: octubre 2023)
- [4] Página web Artec3D. Escáner Spider. <https://www.artec3d.com/es/portable-3d-scanners/artec-spider> (Consultada: octubre 2023)

Premios José María Ferrero Corral

Jueves 23 12:00 - 13:30

Paraninfo

Very low elastic modulus Ti alloys obtained by Laser Directed Energy Deposition to avoid bone resorption in bone implants

F. Arias-González ¹, A. Rodríguez-Contreras ^{2,3,4}, M. Punset ^{2,3,4,5}, J.M. Manero ^{2,3,4}, Ó. Barro ¹,
M. Fernández-Arias ¹, F. Lusquiños ^{1,6}, J. Gil ⁷, J. Pou ^{1,6}

¹LaserOn Research Group, CINTECX, School of Engineering, Universidade de Vigo (UVIGO), Vigo, Spain
felipeag@uvigo.es (F.A.-G.); obarro@uvigo.es (Ó.B.); monfernandez@uvigo.es (M.F.-A.); flusqui@uvigo.es (F.L.);
jpou@uvigo.es (J.P.)

²Biomaterials, Biomechanics and Tissue Engineering Group, Materials Science and Engineering Department, and Research Center for Biomedical Engineering, Universitat Politècnica de Catalunya (UPC), Barcelona, Spain
alejandra.maria.rodriguez@upc.edu (A.R.-C.); miquel.punset@upc.edu (M.P.); jose.maria.manero@upc.edu (J.M.M.)

³Institut de Recerca Sant Joan de Déu (IRSJD), Barcelona, Spain

⁴Barcelona Research Center in Multiscale Science and Engineering, Universitat Politècnica de Catalunya (UPC), Barcelona, Spain

⁵UPC Innovation and Technology Center (CIT-UPC), Universitat Politècnica de Catalunya (UPC), Barcelona, Spain

⁶Galicia Sur Health Research Institute (IIS Galicia Sur), SERGAS-UVIGO, Vigo, Spain

⁷School of Dentistry, Universitat Internacional de Catalunya (UIC), Sant Cugat del Vallès, Spain; xavier.gil@uic.cat

Abstract

The elastic modulus of the cortical bone is below 30 GPa, whereas biomedical titanium implants exhibit an elastic modulus above 100 GPa. This mismatch in the elastic modulus can lead to bone resorption caused by the stress-shielding effect and poor osseointegration of the implant. This study aimed to determine whether the intense <100> fiber texture developed in Laser Directed Energy Deposition of beta-type Ti alloy ingots, results in a significant reduction in the elastic modulus. We demonstrated that laser-deposited beta-type Ti-42Nb (wt%) alloy ingots exhibit anisotropic mechanical properties. A low elastic modulus (below 50 GPa) and a high yield strength (above 700 MPa) were obtained in the building direction because of the intense <100> fiber texture. The novel laser deposited Ti-42Nb alloy also shows excellent biological performance in vitro, which suggests its suitability for biomedical applications.

1. Introduction

Titanium and its alloys are highly regarded as the most suitable metallic materials for biomedical applications due to their exceptional comprehensive properties [1]. Commercially pure titanium (cp-Ti) and Ti-6Al-4V have established themselves as the predominant choices in orthopedic and dental applications, thanks to their impressive attributes including toughness, excellent biocompatibility, strong corrosion resistance, and a relatively low elastic modulus compared to other metallic materials [2]. However, cp-Ti cannot meet the requirements for high-strength applications, while Ti-6Al-4V can release cytotoxic ions, such as Al and V [3, 4]. Moreover, the higher elastic modulus of these materials compared to cortical bone can lead to stress shielding, hindering bone regeneration and potentially causing complications like bone fractures and implant loosening [3, 5].

Ongoing research focuses on developing new biomedical titanium alloys, particularly β -type titanium alloys, using elements with low toxicity like Nb, Ta, Zr, Sn, and Hf, which offer a favorable combination of low elastic modulus, high strength, corrosion resistance, and biocompatibility [4, 6]. Binary Ti-Nb alloys have gained attention as implant materials due to their non-toxicity, biocompatibility, and corrosion resistance [7, 8]. A critical concentration of 36 wt% Nb is needed to retain the β -phase at room temperature, and the elastic modulus reaches a minimum (~ 62 GPa) at 40–45 wt% Nb [9, 10]. Although lower than cp-Ti and Ti-6Al-4V alloy, the elastic modulus remains higher than cortical bone, emphasizing the desire for further reduction.

A promising approach to reduce the elastic modulus of biocompatible β -type Ti alloy implants involves controlling the crystallographic texture to influence the orientation-dependent elastic modulus [11]. It has been shown that β -type Ti alloys exhibit anisotropic elastic modulus, with the highest value along the <111> orientation and the lowest value along the <100> orientation. By manipulating the crystallographic texture, the elastic modulus of β -type Ti alloys can be effectively lowered. However, the challenge lies in developing new manufacturing processes to achieve controlled crystallographic texture in biomedical implants. Recent research has demonstrated successful texture control in β -type Ti alloys using selective laser melting [12] and cold groove rolling techniques [13], resulting in low elastic modulus values that potentially can minimize stress shielding and improve implant performance.

In this study, Laser Directed Energy Deposition (LDED) was used to develop a <100> fiber texture in β -type Ti-42Nb alloy ingots, aiming to achieve anisotropic

mechanical properties and a very low elastic modulus. LDED is an additive manufacturing process that utilizes a laser beam to selectively deposit melted material onto a surface [14-20]. This process has been successfully applied to produce various titanium and titanium alloy components. In previous studies, LDED-generated β type Ti alloys exhibited a retained metastable β -phase and a distinctive elongated columnar β grain structure with a prominent $\langle 100 \rangle$ fiber texture [21]. Building upon this background, we aimed to manufacture biocompatible β -type Ti-42Nb alloy ingots with a strong $\langle 100 \rangle$ fiber texture using LDED and evaluate their crystallographic texture, mechanical properties, and cytocompatibility. This study provides insights into the elastic modulus of laser-deposited biocompatible β -type Ti alloy ingots and their potential applications in orthopedic and dental fields.

2. Materials and methods

The Laser Directed Energy Deposition (LDED) technique was used to create Ti-42Nb specimens on flat cp-Ti grade 2 substrates. The LDED set-up included a High Power Diode Laser (HPDL) with a wavelength of 915-976 nm and a maximum power of 1600 W, a pneumatic powder feeder, a coaxial injection powder system, and a CNC controlled three-axis positioning system. The processing zone was enclosed in an inert chamber using argon gas with oxygen levels below 50 ppm. A lens focused the laser beam onto the substrate, creating a circular spot with a diameter of around 3 mm. Process parameters included a laser power of 1000 W, processing speed of 6 mm/s, and Ti-42Nb powder flow rate of 2.2 g/min. The scanning path followed a back-and-forth pattern with a 90° rotation between layers. Spherical pre-alloyed Ti-42Nb powder with particle sizes of 63-105 μm was used as the precursor material.

The Ti-42Nb samples were prepared for examination by cutting them, embedding them in resin, and polishing them. The crystallographic texture of the samples was then analyzed using Electron Backscatter Diffraction (EBSD) with the help of ATEX software [22]. To evaluate their mechanical properties, uniaxial tensile tests were conducted following the ISO 6892 1:2020 standard [23], using a Zwick Z100 universal machine. The tests measured the values of elastic modulus (E) and yield strength (Y).

Human osteoblast-like cells SaOS-2 were used to assess the in vitro cytocompatibility of the laser-deposited Ti-42Nb alloy compared to the cp-Ti grade 2 control. A cellular density of 20 000 cells/well was cultured on both types of samples in Falcon 12-well plates. Adhesion studies were conducted for 4 hours, while proliferation studies were performed for 3, 7, and 14 days. The cells were examined using a field emission scanning electron microscope (FESEM). For quantitative analysis, ten random images were taken at 300x magnification from the surface of four samples.

3. Results and discussion

Through the precise adjustment of process parameters, we achieved successful fabrication of solid β type Ti-42Nb alloy ingots using Laser Direct Energy Deposition (LDED) technique. To examine the crystallographic texture of the

laser-deposited Ti-42Nb, we conducted an analysis using Electron Backscatter Diffraction (EBSD). Our findings revealed a notable $\langle 100 \rangle$ fiber texture in the laser-deposited Ti-42Nb samples (Fig. 1), with a considerable number of β -phase grains aligning one of their $\langle 100 \rangle$ axes nearly parallel to the building direction (Z-axis).

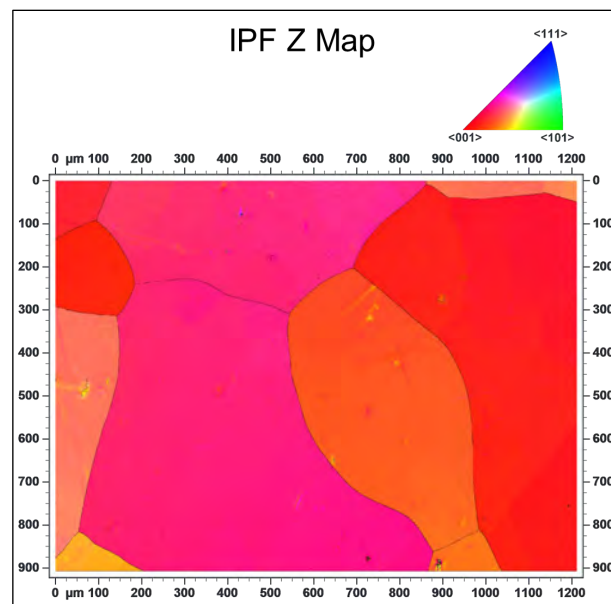


Figure 1. Electron backscatter diffraction (EBSD) results at XY cross section (normal to Z-axis): inverse pole figure Z map.

The intense $\langle 100 \rangle$ fiber texture present in these laser-deposited ingots gives rise to anisotropic mechanical properties. In Fig. 2, we provide a comparison of the elastic modulus (E) and yield strength (Y) of the laser deposited Ti-42Nb alloy (X-axis and Z-axis) with selected biomedical titanium alloys: cp-Ti grade 2, cp-Ti grade 4, Ti-6Al-4V and Ti-6Al-4V ELI [24]. In the scanning direction (X-axis), the laser-deposited Ti-42Nb exhibits a relatively low modulus of elasticity ($E_x = 59.4 \pm 3.0$ GPa), and reasonably high yield strength ($Y_x = 735 \pm 22$ MPa). On the other hand, the elastic modulus of the laser-deposited Ti-42Nb in the building direction (Z-axis) is even lower ($E_z = 47.9 \pm 3.9$ GPa), while maintaining comparable yield strength ($Y_z = 715 \pm 41$ MPa). Notably, the elastic modulus along the Z-axis is very low compared to the commercially available biomedical titanium alloys (cp-Ti grade 2, cp-Ti grade 4, Ti-6Al-4V and Ti-6Al-4V ELI).

To evaluate the in vitro cytocompatibility of the laser-deposited Ti-42Nb alloy compared to cp-Ti grade 2 control, human osteoblast-like cells SaOS-2 were employed. The cell count analysis revealed significant differences in cell adhesion after 4 hours of incubation (Fig. 3), with the Ti-42Nb substrate exhibiting more favorable results. Over a period of 14 days, cell proliferation was observed on both substrates; however, the laser-deposited Ti-42Nb surface promoted a higher degree of proliferation. Overall, the laser-deposited Ti-42Nb alloy demonstrated enhanced osteoinductive effects by facilitating the adhesion and proliferation of SaOS-2 cells.

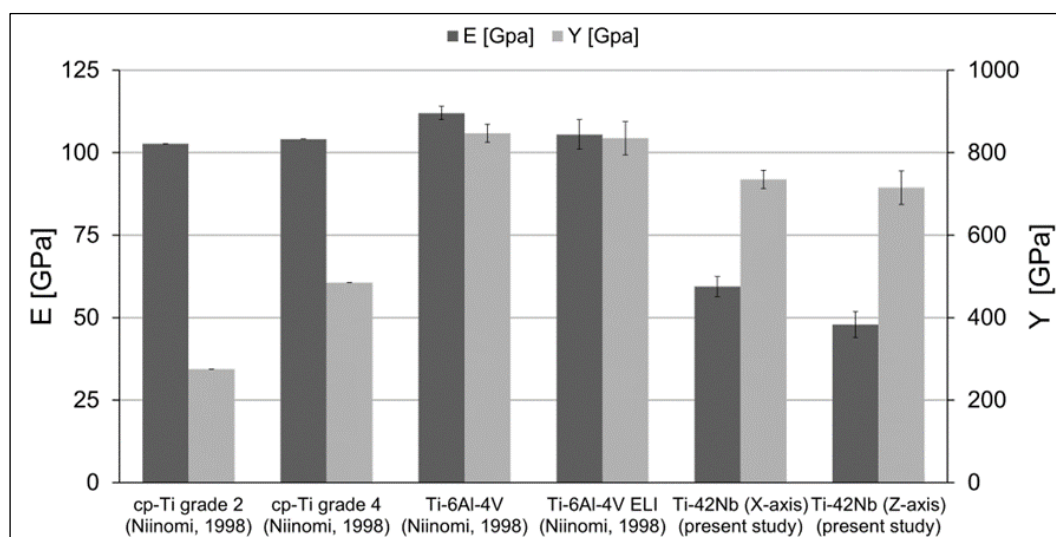


Figure 2. Mechanical properties of selected biomedical Ti alloys determined in tensile tests: elastic modulus (E), yield strength (Y).

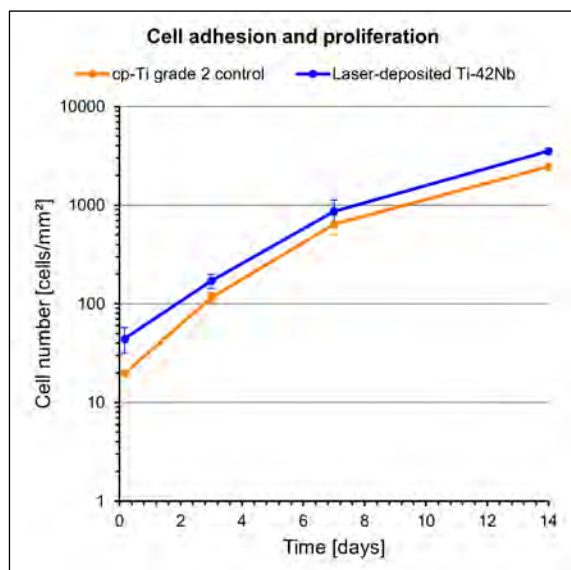


Figure 3. Cell adhesion (4h) and proliferation (3, 7 and 14 days) on cp-Ti grade 2 and laser-deposited Ti-42Nb samples. Error bars represent standard deviation.

The findings from this study suggest that the laser-deposited β -type Ti-42Nb alloy holds significant potential for the development of innovative biomedical implants characterized by a remarkably low elastic modulus, which can help mitigate stress shielding effects. The achievement of such a low elastic modulus in the building direction (Z-axis) is attributed to the pronounced $\langle 100 \rangle$ fiber texture, wherein a substantial number of β -phase grains align one of their $\langle 100 \rangle$ axes parallel to the Z-axis. Previous studies by Tane et al. [25] and Hermann et al. [11] had already proposed that biocompatible β -type titanium alloy polycrystals featuring a texture aligning the crystallographic $\langle 100 \rangle$ directions or single crystals with the $\{100\}$ direction oriented along the loading direction in human bones show promise as biomedical implant materials with an elastic modulus comparable to that of natural bones. However, a major challenge lies in establishing new processes for producing large single-

crystals or textured polycrystal ingots and subsequently precisely machining the implants. This paper successfully demonstrates the feasibility of Laser Directed Energy Deposition for generating a β -type Ti-42Nb alloy textured polycrystal ingot with an exceptionally low modulus, opening doors for the development of novel biomedical implants.

4. Conclusions

Biocompatible β -type Ti-42Nb alloy ingots were successfully produced using Laser Directed Energy Deposition (LDED) technique. The laser-deposited Ti-42Nb samples exhibited a notable $\langle 100 \rangle$ fiber texture, with a significant number of β -phase grains aligning one of their $\langle 100 \rangle$ axes parallel to the building direction (Z-axis). This intense texture resulted in the specimens displaying anisotropic mechanical properties. Notably, the laser-deposited Ti-42Nb demonstrated a remarkably low elastic modulus in the building direction ($E_z = 47.9 \pm 3.9$ GPa) and high yield strength ($Y_z = 715 \pm 41$ MPa). Furthermore, the laser-deposited Ti-42Nb alloy exhibited enhanced osteoinductive effects promoting the adhesion and proliferation of Saos-2 cells. These findings suggest that the laser-deposited β -type Ti-42Nb alloy is a promising candidate for the development of novel biomedical implants with a very low elastic modulus, capable of mitigating stress shielding effects.

Acknowledgements

The authors wish to thank the technical staff from Barcelona Research Center in Multiscale Science and Engineering (Universitat Politècnica de Catalunya) and Light Alloys and Surface Treatments Design Centre (Universitat Politècnica de Catalunya) for their technical assistance with the characterization techniques. Technical assistance of Mr. A. Abalde (Universidade de Vigo) is also acknowledged. This research was partially funded by SOADCO S.L. (Escaldes Engordany, Andorra), by the Spanish Government (This publication is part of the

TED2021-131721B-I00 project, funded by MCIN/AEI/10.13039/501100011033 and European Union “NextGenerationEU”/PRTR), by Xunta de Galicia (ED431C 2023/25; ED481B-2021-113), by Generalitat de Catalunya (2017 SGR-1165).

References

- [1] Niinomi M, Nakai M, Hieda J. Development of New Metallic Alloys for Biomedical Applications. *Acta Biomaterialia*, vol 8, sup 11, 2012, pp 3888-3903 (ISSN 1742-7061).
- [2] Szczyński G, Kopec M, Politis DJ, Kowalewski ZL, Łazarski A, Szolc T. A Review on Biomaterials for Orthopaedic Surgery and Traumatology: From Past to Present. *Materials*, vol 15, sup 10, 2022, pp 3622 (ISSN 1996-1944).
- [3] Li Y, Yang C, Zhao H, Qu S, Li X, Li Y. New Developments of Ti-Based Alloys for Biomedical Applications. *Materials* vol 7, sup 3, 2014, pp 1709-1800 (ISSN 1996-1944).
- [4] Biesiekierski A, Wang J, Abdel-Hady Gepreel M, Wen C. A New Look at Biomedical Ti-Based Shape Memory Alloys. *Acta Biomaterialia*, vol 8, sup 5, 2012, pp 1661-1669 (ISSN 1742-7061).
- [5] Brizuela A, Herrero-Climent M, Rios-Carrasco E, Rios-Santos JV, Pérez R, Manero JM, Gil Mur J. Influence of the Elastic Modulus on the Osseointegration of Dental Implants. *Materials*, vol 12, sup 6, 2019, pp 980 (ISSN 1996-1944).
- [6] Kim KM, Kim HY, Miyazaki S. Effect of Zr Content on Phase Stability, Deformation Behavior, and Young's Modulus in Ti-Nb-Zr Alloys. *Materials*, vol 13, sup 2, 2020, pp 476 (ISSN 1996-1944).
- [7] Godley R, Starosvetsky D, Gotman I. Corrosion Behavior of a Low Modulus β -Ti-45%Nb Alloy for Use in Medical Implants. *Journal of Materials Science: Materials in Medicine*, vol 17, sup 1, 2006, pp 63-67 (ISSN 0957-4530).
- [8] Bai Y, Deng Y, Zheng Y, Li Y, Zhang R, Lv Y, Zhao Q, Wei S. Characterization, Corrosion Behavior, Cellular Response and In Vivo Bone Tissue Compatibility of Titanium-Niobium Alloy with Low Young's Modulus. *Materials Science and Engineering: C*, vol 59, 2016, pp 565-576 (ISSN 0928-4931).
- [9] Hanada S, Matsumoto H, Watanabe S. Mechanical Compatibility of Titanium Implants in Hard Tissues. *International Congress Series*, vol 1284, 2005, pp 239-247 (ISSN 0531-5131).
- [10] Ozaki T, Matsumoto H, Watanabe S, Hanada S. Beta Ti Alloys with Low Young's Modulus. *Materials Transactions*, vol 45, sup 8, 2004, pp 2776-2779 (ISSN 1345-9678).
- [11] Hermann R, Hermann H, Calin M, Büchner B, Eckert J. Elastic Constants of Single Crystalline β -Ti70Nb30. *Scripta Materialia*, vol 66, sup 3-4, 2012, pp 198-201 (ISSN 1359-6462).
- [12] Ishimoto T, Hagihara K, Hisamoto K, Sun SH, Nakano T. Crystallographic Texture Control of Beta-Type Ti-15Mo-5Zr-3Al Alloy by Selective Laser Melting for the Development of Novel Implants with a Biocompatible Low Young's Modulus. *Scripta Materialia*, vol 132, 2017, pp 34-38 (ISSN 1359-6462).
- [13] Shinohara Y, Matsumoto Y, Tahara M, Hosoda H, Inamura T. Development of $\langle 001 \rangle$ -Fiber Texture in Cold-Groove-Rolled Ti Mo-Al-Zr Biomedical Alloy. *Materialia*, vol 1, 2018, pp 52-61 (ISSN 2589-1529).
- [14] Saboori A, Gallo D, Biamino S, Fino P, Lombardi M. An Overview of Additive Manufacturing of Titanium Components by Directed Energy Deposition: Microstructure and Mechanical Properties. *Applied Sciences*, vol 7, sup 883, 2017, pp 883 (ISSN 2076-3417).
- [15] Arias-González F, del Val J, Comesaña R, Penide J, Lusquiños F, Quintero F, Riveiro A, Boutinguiza M, Gil J, Pou J. Microstructure and Crystallographic Texture of Pure Titanium Parts Generated by Laser Additive Manufacturing. *Metals and Materials International*, vol 24, sup 1, 2018, pp 231-239 (ISSN 1598-9623).
- [16] Barro Ó, Arias-González F, Lusquiños F, Comesaña R, del Val J, Riveiro A, Badaoui A, Gómez-Baño F, Pou J. Effect of Four Manufacturing Techniques (Casting, Laser Directed Energy Deposition, Milling and Selective Laser Melting) on Microstructural, Mechanical and Electrochemical Properties of Co-Cr Dental Alloys, Before and After PFM Firing Process. *Metals*, vol 10, sup 10, 2020, pp 291 (ISSN 2075-4701).
- [17] Barro Ó, Arias-González F, Lusquiños F, Comesaña R, del Val J, Riveiro A, Badaoui A, Gómez-Baño F, Pou J. Improved Commercially Pure Titanium Obtained by Laser Directed Energy Deposition for Dental Prosthetic Applications. *Metals*, vol 11, sup 1, 2021, pp 291 (ISSN 2075-4701).
- [18] Arias-González F, Rodríguez-Contreras A, Punset M, Manero JM, Barro Ó, Fernández-Arias M, Lusquiños F, Gil J, Pou J. 2021. In-Situ Laser Directed Energy Deposition of Biomedical Ti-Nb and Ti-Zr-Nb Alloys from Elemental Powders. *Metals*, vol 11, 2021, pp 1205 (ISSN 2075-4701).
- [19] Arias-González F, Barro O, del Val J, Lusquiños F, Fernández-Arias M, Comesaña R, Riveiro A, Pou J. 2021. Laser-Directed Energy Deposition, in “Additive Manufacturing”, Elsevier, 2021 (ISBN 978-0-12-818411-0).
- [20] Arias-González F, Rodríguez-Contreras A, Punset M, Manero JM, Barro Ó, Fernández-Arias M, Lusquiños F, Gil J, Pou J. Laser-Deposited Beta Type Ti-42Nb Alloy with Anisotropic Mechanical Properties for Pioneering Biomedical Implants with a Very Low Elastic Modulus. *Materials*, vol 15, sup 20, 2022, pp 7172 (ISSN 1996-1944).
- [21] Banerjee R, Nag S, Samuel S, Fraser HL. Laser-Deposited Ti-Nb-Zr-Ta Orthopedic Alloys. *Journal of Biomedical Materials Research Part A*, vol 78, sup 2, 2006, pp 298-305 (ISSN 1549-3296).
- [22] B. Beausir, J.-J. Fundenberger, Analysis Tools for Electron and X-ray diffraction, ATEX - software, www.atex-software.eu, Université de Lorraine - Metz, 2017
- [23] ISO 6892-1:2019 Metallic materials — Tensile testing — Part 1: Method of test at room temperature.
- [24] Niinomi M. Mechanical properties of biomedical titanium alloys. *Materials Science and Engineering: A*, vol 243, sup 1-2, 1998, pp 231-236 (ISSN 0921-5093).
- [25] Tane M, Akita S, Takayoshi N, Hagihara K, Umakoshi Y, Niinomi M, Nakajima H. Peculiar elastic behavior of Ti-Nb-Ta-Zr single crystals. *Acta Materialia*, vol 56, 2008, pp 2856-2863 (ISSN 359-6454)

Comparación *in-silico* de la eficacia de los fármacos flecainida y vernakalant en fibrilación auricular paroxística

V. Puche-García¹, L Martínez Mateu², D. Filgueiras-Rama³, L. Romero¹, J. Saiz¹

¹ Centro de Investigación e Innovación en Bioingeniería (Ci2B), Universitat Politècnica de València, España

² Grupo de investigación de alto rendimiento en Ingeniería Biomédica y Ciencia de Datos (BigMed+), Universidad Rey Juan Carlos, España

³ Centro Nacional de Investigaciones Cardiovasculares (CNIC), España

jsaiz@ci2b.upv.es

Resumen

La terapia farmacológica tiene un gran potencial para el tratamiento de la fibrilación auricular paroxística (pAF); aunque su eficacia es dependiente de las características de cada paciente. Este estudio busca analizar qué perfiles iónicos y mecanismos electrofisiológicos modulan la efectividad de dos fármacos antiarrítmicos (flecainida y vernakalant) contra los episodios fibrilatorios. En los modelos auriculares paroxísticos de aurícula izquierda 2D, los bloqueos G_{K1} , G_{NaK} , G_{to} y $G_{K,ACH}$ son importantes para reducir la iniciación y el mantenimiento de la AF a través del aumento de la duración del potencial y la refractariedad junto con la reducción de la frecuencia dominante del rotor (DF). Cuando se parte de pAFs con DFs bajos, los medicamentos flecainida y vernakalant tienen una mayor eficacia. Las características de las corrientes iónicas influyen en la iniciación, el mantenimiento y la respuesta de la AF al tratamiento farmacológico.

1. Introducción

La fibrilación auricular (Atrial Fibrillation, AF) es la arritmia cardíaca más común afectando casi al 2% de la población general. Puede progresar desde episodios cortos (≤ 7 días) de AF paroxística (pAF) hasta cuadros de mayor duración y frecuencia de AF persistente. Se ha demostrado que pacientes con pAF sufren cambios en el tejido auricular izquierdo (LA, *Left Atrium*) que incluyen alteraciones de diferentes corrientes iónicas (remodelado eléctrico), una conducción más lenta y periodos refractarios cortos que favorecen el inicio y mantenimiento de circuitos de reentrada propios de la AF [1].

La administración de fármacos antiarrítmicos ofrece un gran potencial para el tratamiento de pacientes con pAF, pero su eficacia es dependiente de las características de cada paciente [1]. Por ello, es necesario tener en cuenta la heterogeneidad presente entre pacientes y comprender el impacto de la variabilidad electrofisiológica sobre los mecanismos de acción y efectividad de los fármacos.

En este estudio, se hace una aproximación *in-silico* usando una malla 2D de tejido de aurícula izquierda con patología pAF sobre el qué se aplican dos fármacos antiarrítmicos ampliamente usados, flecainida y vernakalant, con el objetivo de estudiar que perfiles iónicos y mecanismos bioeléctricos modulan la efectividad de estos fármacos y su impacto sobre la vulnerabilidad a sufrir episodios fibrilatorios.

2. Métodos

2.1. Modelo de miocito humano auricular

Para reproducir el potencial de acción (*Action Potential*, AP) de un miocito humano auricular se utilizó el modelo Courtemanche-Ramirez-Nattel [2]. Se incluyó la corriente de K^+ activada por acetilcolina ($I_{K,ACH}$) propuesta por Grandi et al. [3] ($[ACH] = 0.005 \mu M$). Este modelo de partida reproduciría el AP sano de la pared de la aurícula derecha (*Right Atrium*, RA).

El remodelado eléctrico y la heterogeneidad auricular se implementaron mediante modificaciones en las conductancias iónicas (G_i) para reproducir las condiciones de pAF en LA [4]. Los factores de escala aplicados se resumen en la Tabla 1. La Figura 1-A muestra los APs para el modelo de partida (sano-RA) y una vez aplicados los factores para reproducir la heterogeneidad de la región LA (sano-LA) y el remodelado por pAF en LA (pAF-LA). Este último fue el modelo de referencia en este estudio.

Tabla 1. Factores de escala aplicados para conseguir la heterogeneidad de la región LA y el remodelado por pAF.

	G_{Kr}	G_{CaL}	G_{K1}	$G_{K,ACH}$
LA	2	0.9	-	-
pAF	-	-	2	2

2.2. Modelado del fármaco

El efecto de los fármacos se reprodujo mediante el modelo de poro simple por el cual las conductancias iónicas se reducen de acuerdo con la ecuación de Hill:

$$g_{i,fármaco} = G_i \cdot \left[1 + \left(\frac{D}{IC_{50,i}} \right)^{nH} \right]^{-1} = G_i \cdot F_{bloqueo} \quad (1)$$

donde para cada corriente i , $IC_{50,i}$ es la concentración inhibitoria media del fármaco, nH es el coeficiente de Hill y D es la concentración de fármaco libre (en este estudio, las concentraciones terapéuticas: $1.5 \mu M$ para flecainida y $10 \mu M$ para vernakalant). Esto se traduce en un factor de bloqueo que multiplica la conductancia máxima (G_i) reduciendo la conductancia final en presencia del fármaco a $g_{i,fármaco}$. Debido a la variabilidad de los IC_{50} observada en experimentos [5], se hizo un estudio preliminar examinando diferentes versiones virtuales de cada fármaco (diferentes IC_{50s}) de manera que la respuesta simulada

fuera similar a los estudios experimentales encontrados en la literatura (datos no incluidos). Los factores de bloqueo obtenidos se resumen en la Tabla 2 y su efecto se muestra en la Figura 1-B donde se observan los APs para pAF-LA en ausencia de fármacos (control) y bajo los efectos de flecainida y vernakalant.

Tabla 2. Factores de bloqueo en función de la concentración terapéutica de cada fármaco.

	G_{CaL}	G_{Kr}	G_{Ks}	G_{Kur}	G_{to}	G_{Na}	$G_{K,ACh}$
Flecainida	0.94	0.50	0.93	0.66	0.78	0.75	-
Vernakalant	0.81	0.67	-	0.60	0.60	0.83	0.50

2.3. Simulación en tejido 2D

Las simulaciones computacionales se realizaron usando una malla 2D que replica un parche de tejido auricular para provocar y estudiar las propiedades de los circuitos de reentrada. El problema monodominio se resolvió utilizando el software Elvira [6].

Previo a la simulación 2D, el modelo celular se estabilizó aplicando 50 pulsos de 2 ms de duración y de 1.5 veces la corriente umbral con un periodo de estimulación de 400 ms. Las versiones del modelo celular incluyen los factores para reproducir las condiciones de pAF-LA en control (pAF-control) e incluyendo los fármacos flecainida (pAF-flecainida) y vernakalant (pAF-vernakalant).

En las simulaciones 2D se utilizó una malla de 5x5 cm² con 300 μm de resolución espacial, conductividad longitudinal igual a 0.0022 S/cm·pF y una relación de anisotropía de 0.35. La malla se estimuló aplicando una corriente de excitación de 2 ms con una amplitud de 100 pA/pF. Para inducir los rotors, se aplicó un protocolo de campo cruzado S1-S2. Primero, se realizó una simulación de estabilización de 10 pulsos planos (S1) con un periodo de 400 ms. A continuación, se aplicó un extra estímulo rectangular (S2) en la esquina inferior izquierda variando el intervalo de acoplamiento entre S1-S2 con una precisión de 1 ms. La vulnerabilidad temporal a la generación de reentradas se cuantificó como el ancho de la ventana de tiempo durante la cual un estímulo S2, aplicado a la cola refractaria del décimo S1, inicia una reentrada que duran al menos 1 s. Se computó el correspondiente pseudo-EGM y la frecuencia dominante (DF, *Dominant Frequency*) asociada a este en cada reentrada obtenida para los últimos 5 s de la simulación para estudiar únicamente la actividad autosostenida del rotor [4].

Los biomarcadores celulares referentes al potencial de reposo (RMP, *Resting Membrane Potential*), a la duración del AP (APD₉₀) y a la velocidad máxima de despolarización (dV/dt_{max}) se calcularon para el AP del nodo central de la malla en el último pulso de estabilización S1. También se obtuvieron biomarcadores tisulares: ventanas vulnerables (VW, *Vulnerable Window*) y DFs; así como el periodo refractario efectivo (ERP, *Effective Refractory Period*), la velocidad de conducción (CV, *Conduction Velocity*) y longitud de onda (WL, *Wave-Length*). El ERP se calculó como el intervalo de acoplamiento más corto que permite la transmisión del segundo estímulo. El biomarcador WL representa la

distancia más corta necesaria para que una onda de reentrada se propague y se calculó de manera directa a partir del ERP y el CV como WL = ERP · CV.

2.4. Análisis de sensibilidad

Se pretende investigar el impacto de la variabilidad en las corrientes iónicas (I_i) y de la conductividad longitudinal del tejido (σ_L) sobre los biomarcadores preclínicos de riesgo arrítmico. Para ello, 10 conductancias iónicas, así como la conductividad longitudinal, se variaron de manera individual en un ±30% y ±40% para pAF-control, pAF-flecainida y pAF-vernakalant con el objetivo de estudiar la dependencia de los fármacos a las variaciones iónicas.

Previamente, los resultados de la variación de ±40% para el modelo pAF-control se usaron para conocer su sensibilidad antes de aplicar los fármacos. Se calcularon los índices porcentuales de cambio (D_{b,p}) y las sensibilidades absolutas (S_{b,p}) y relativas (R_{b,p}) de cada biomarcador “b” para la variación en el parámetro “p”:

$$D_{b,p,\pm 40\%} = 100 \cdot \frac{C_{b,p_0 \pm 40\%} + C_{b,p_0}}{C_{b,p_0}} \quad (2)$$

$$S_{b,p} = \frac{D_{b,p_0+40\%} - D_{b,p_0-40\%}}{0.8} \quad (3)$$

$$R_{b,p} = \left| \frac{S_{b,p}}{\max_p \{S_{b,p}\}} \right| \quad (4)$$

3. Resultados y discusión

3.1. Características de pAF en la aurícula izquierda

Los valores de APD₉₀ para sano-RA, sano-LA y pAF-LA son 229 ms, 180 ms, 113 ms, respectivamente (Figura 1-A), por lo que se consigue reproducir la disminución del APD₉₀ debido a la heterogeneidad auricular y el remodelado eléctrico [4].

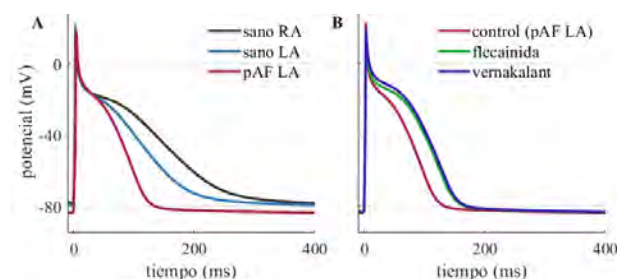


Figura 1. APs de los modelos para el último pulso de estabilización celular. (A) Modelos sano en RA, sano en LA y pAF en LA. (B) Modelos pAF-LA en control y bajo la aplicación de flecainida y vernakalant.

La Figura 2 muestra el impacto de la variación en las corrientes iónicas y en la conductividad longitudinal sobre los biomarcadores estudiados.

Observando los biomarcadores celulares, la inhibición de G_{K1} produce un alargamiento del APD₉₀ y un incremento en el RMP. Junto con G_{K1} las variaciones en G_{CaL}, G_{to} y G_{K,ACh} suponen los mayores cambios sobre el APD₉₀. El aumento del APD₉₀ constituye un mecanismo antiarrítmico pues permite mitigar la reducción causada por el remodelado eléctrico relativo a la pAF. En el modelo, este aumento es considerable cuando G_{K1}, G_{to} y G_{K,ACh} están reducidas (corrientes de K⁺ salientes) y G_{CaL} está aumentada (corriente de Ca²⁺ entrante). Respecto a

dV/dt_{max} , G_{Na} tiene el mayor peso ya que es la responsable de la despolarización y del inicio del AP.

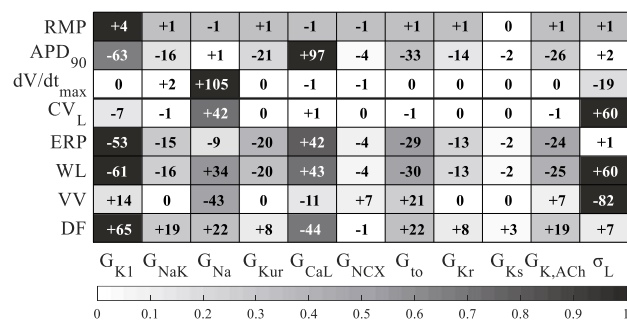


Figura 2. Análisis de sensibilidad del modelo. Valores absolutos (valor numérico) y relativos (escala de grises) en condiciones de control sin la aplicación de ningún fármaco.

Respecto a los biomarcadores tisulares, el ERP está altamente influenciado por G_{K1} , G_{CaL} y G_{to} que puede estar debido a la sensibilidad de estas conductancias sobre el APD₉₀. Se observa que tanto inhibiciones de G_{Na} (ligada especialmente con la excitabilidad tisular) como σ_L ralentizan la CV. Aparte, sólo el bloqueo de G_{K1} produce cambios significativos en CV probablemente a causa de su efecto en la despolarización de RMP. En cuanto a la WL, su dependencia es notable respecto de G_{K1} , G_{CaL} y G_{to} por su dependencia sobre el ERP. La variación de G_{Na} es menor ya que induce cambios opuestos en ERP y CV.

Las conductancias que mayor influencia tienen sobre las dinámicas fibrilatorias son G_{K1} , G_{Na} , G_{NaK} , G_{Ca} y G_{to} . La VW se ve reducida cuando G_{Na} , G_{CaL} y σ_L están aumentadas y G_{K1} reducida. Los bloqueos en G_{K1} , G_{NaK} , G_{Na} , G_{to} y $G_{K,ACh}$ y aumentos en G_{CaL} implican una reducción en la DF del rotor (Figura 4-A).

De este análisis destaca el papel crucial de G_{K1} , y en menor medida de G_{to} y $G_{K,ACh}$, en la estabilidad del rotor reentrante y el riesgo arritmico pues su bloqueo actúa como mecanismo antiarritmico aumentando APD₉₀, ERP y WL lo que ha demostrado detener con éxito las arritmias en diferentes especies animales [7, 8].

La influencia de la inhibición de I_{Na} sobre las propiedades del tejido en cuanto a la reducción de dV/dt_{max} y CV junto con el aumento del ERP favorecen la reducción antiarritmica de la excitabilidad celular y de la automaticidad ectópica auricular. Esto hace que los fármacos que incluyen este bloqueo sean utilizados ampliamente para el tratamiento de la AF [5].

3.2. Efecto farmacológico sobre biomarcadores celulares: APD₉₀ y dV/dt_{max}

En pAF-LA, los valores de APD₉₀ para control y en presencia de flecainida y vernakalant son 113 ms, 142 ms y 145 ms, respectivamente (Figura 1-B), por lo que se consigue reproducir el incremento del APD₉₀ característico de estos fármacos antiarrítmicos [5].

Atendiendo a los resultados de sensibilidad, la Figura 3 muestra la variación de APD₉₀ entre control y bajo efecto de fármacos en presencia de alternaciones de $\pm 30\%$ y $\pm 40\%$ en las conductancias que mayor impacto tienen sobre este biomarcador: G_{K1} , G_{CaL} , G_{to} y $G_{K,ACh}$, con el

objetivo de estudiar los efectos de los fármacos para variaciones individualizadas de cada corriente. De manera general, vernakalant consigue mayores incrementos del APD₉₀ respecto a flecainida. Sin embargo, se observan diferentes mecanismos para cada fármaco excepto para la variación de APD₉₀ en función de G_{to} , que sí sigue la misma tendencia para ambos (mayor aumento a mayor G_{to}). Respecto al modelo base, flecainida es más efectiva cuando G_{K1} y $G_{K,ACh}$ están inhibidas y G_{CaL} está aumentada; mientras que vernakalant aumenta su efectividad para este biomarcador cuando G_{K1} y $G_{K,ACh}$ están aumentadas. Vernakalant es menos efectivo cuando G_{CaL} está alterada en comparación con el modelo sin modificar.

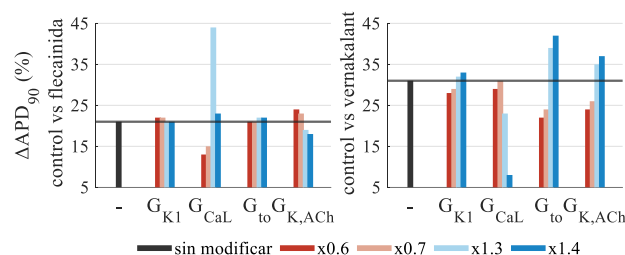


Figura 3. Variación de APD₉₀ (%) entre simulaciones control y con fármaco para variaciones en G_{K1} , G_{CaL} , G_{to} y $G_{K,ACh}$.

En la Tabla 3 se muestra cómo ambos fármacos consiguen reducir dV/dt_{max} siendo mayor el efecto de la flecainida y siendo mayor la diferencia cuando G_{Na} esta inhibida frente a cuando está aumentada.

Tabla 3. Variación de dV/dt_{max} (%) entre simulaciones control y con fármaco para la variación de G_{Na} .

G_{Na}	x0.6	x0.7	-	x1.3	x1.4
Flecainida	-33	-31	-27	-25	-25
Vernakalant	-22	-20	-18	-17	-16

3.3. Efecto farmacológico sobre la actividad fibrilatoria en tejido: DF y VW

La Figura 4-A muestra los valores de DF de los rotors obtenidos en las simulaciones control sin fármaco; mientras que las Figuras 4-B y 4-C muestran la variación en la VW entre las simulaciones control y en las que se aplican flecainida y vernakalant, respectivamente, para las variaciones individualizadas de cada parámetro. Como se adelantaba en los resultados de sensibilidad, en la Figura 4-A se observa cómo incrementos en G_{K1} y reducciones en G_{CaL} producen un aumento considerable en las DFs.

Para DF altas en control, los fármacos presentan mecanismos diferentes en relación con la VW. Flecainida tiende a aumentar ligeramente la VW a DFs altas en control, mientras que vernakalant reduce de manera general la VW independientemente de los valores de DF en control. Esto revela que, el uso de la flecainida estaría desaconsejado para pacientes con DF elevadas, ya que podría favorecer la vulnerabilidad en LA. Por el contrario, para pacientes con DF menores, ambos fármacos serían óptimos siendo más efectivo el vernakalant. Este mecanismo resulta opuesto para el caso de G_{Na} y σ_L , donde DFs altas en control implican una mayor eficacia de los

fármacos (mayor en el caso del vernakalant) y DFs bajas en control llevan asociado un ligero aumento de la VW.

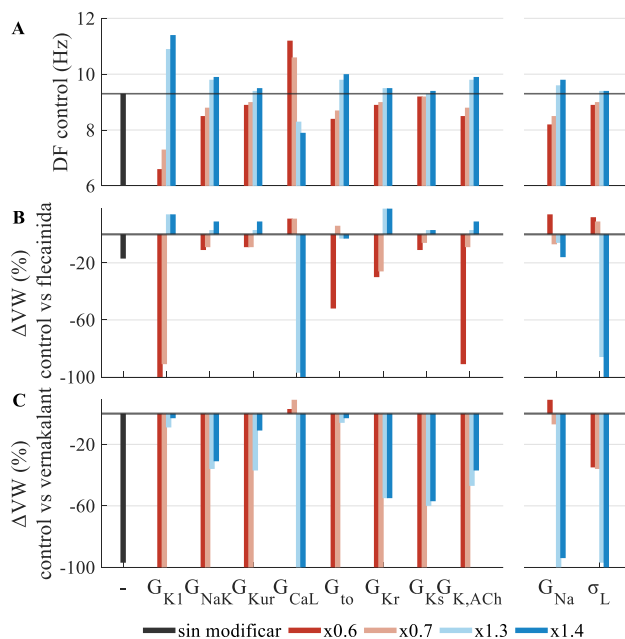


Figura 4. Relación entre la DF en control y la variación de la VW al aplicar fármacos para cada variación de los parámetros. (A) Valores de DF. (B, C) Variación de la VW (%) al aplicar flecainida y vernakalant respecto con las simulaciones control.

Con todo ello, los fármacos resultan considerablemente efectivos cuando son aplicados sobre modelos con bloqueos en G_{K1} , G_{to} y $G_{K,ACh}$ disminuyendo la vulnerabilidad a generar rotos fibrilatorios (reducen VW) pues provocan una reducción de la DF además de aumentar el APD_{90} , ERP y WL (mecanismos antiarrítmicos) [7].

3.4. Líneas futuras

En próximos estudios, se podrían incluir fármacos que incluyan un bloqueo de I_{K1} como la cloroquina que ha dado buenos resultados en estudios ovinos [9], ya que el aumento de I_{K1} ha demostrado ser esencial para la estabilización del rotor [7]. Además, estudios previos han demostrado que existe una correlación clínica entre las características de frecuencia del ECG (ondas f) y las de las señales intracardiacas (valores de DF) [10]. Por tanto, el enfoque del presente trabajo podría utilizarse para ayudar a decidir el tratamiento farmacológico óptimo para cada paciente en base al ECG, señal ampliamente utilizada en la práctica clínica. Ello requería extender el análisis a una población de modelos que considere la variabilidad inter-sujeto e incluir diferentes protocolos de inducción de AF. Esto permitiría tener unos resultados más realistas en cuanto a la efectividad del fármaco en función de la DF.

4. Conclusiones

El análisis realizado en este estudio es un nuevo avance en la comprensión de los mecanismos iónicos subyacentes a la AF a nivel tisular y de su respuesta al tratamiento farmacológico con flecainida y vernakalant. Nuestros resultados sugieren que la iniciación, el mantenimiento y la respuesta de la AF al tratamiento farmacológico dependen en gran medida de las propiedades de las

corrientes iónicas. En concreto, los bloqueos de G_{K1} , G_{NaK} , G_{to} y $G_{K,ACh}$ ayudan a prevenir la iniciación y el mantenimiento de la AF en los modelos auriculares paroxísticos de aurícula izquierda a través del aumento del APD_{90} , ERP, WL y reducción de DF. Además, para aquellos modelos que parten de DFs bajas, los fármacos flecainida y vernakalant consiguen reducir en mayor medida la VW. Este trabajo demuestra la utilidad de los estudios *in-silico* en la medicina personalizada, ya que pueden ayudar a predecir y comprender el efecto de los fármacos en la reducción de la vulnerabilidad a sufrir AF.

Agradecimientos

Este trabajo ha sido financiado por el Programa Prometeo de la Generalitat Valenciana (2020/043) y por el programa de investigación e innovación Horizon 2020 de la Unión Europea bajo subvención No 101016496 (SimCardioTest). Los autores agradecen también los recursos informáticos de MareNostrum y el soporte técnico prestado por Barcelona Supercomputing Center (IM-2021-1-0003, IM-2022-1-0002, IM-2023-1-0002), y los créditos aportados por Oracle Cloud y recursos relacionados de Oracle for Research.

Referencias

- [1] J. Andrade *et al.*, "The 2020 Canadian Cardiovascular Society/Canadian Heart Rhythm Society comprehensive guidelines for the management of atrial fibrillation," *Canadian Journal of Cardiology*, vol. 36, no. 12, pp. 1847-1948, 2020.
- [2] M. Courtemanche *et al.*, "Ionic mechanisms underlying human atrial action potential properties: insights from a mathematical model," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 275, no. 1, pp. H301-H321, 1998.
- [3] E. Grandi *et al.*, "Human atrial action potential and Ca^{2+} model: sinus rhythm and chronic atrial fibrillation," *Circulation research*, vol. 109, no. 9, pp. 1055-1066, 2011.
- [4] L. Martinez-Mateu, Mapping of the Electrical Activity of Human Atria Multiscale Modelling and Simulations, 2018.
- [5] M. Geng *et al.*, "Revisiting antiarrhythmic drug therapy for atrial fibrillation: reviewing lessons learned and redefining therapeutic paradigms," *Frontiers in Pharmacology*, vol. 11, no. 581837, 2020.
- [6] E. Heidenreich *et al.*, "Adaptive macro finite elements for the numerical solution of monodomain equations in cardiac electrophysiology," *Annals of biomedical engineering*, vol. 38, pp. 2331-2345, 2010.
- [7] S. Pandit *et al.*, "Ionic determinants of functional reentry in a 2-D model of human atrial cells during simulated chronic atrial fibrillation," *Biophysical journal*, vol. 88, no. 6, pp. 3806-3821, 2005.
- [8] R. Sato *et al.*, "A new class III antiarrhythmic drug, MS-551, blocks the inward rectifier potassium channel in isolated guinea pig ventricular myocytes," *Journal of Pharmacology and Experimental Therapeutics*, vol. 274, no. 1, pp. 469-474, 1995.
- [9] D. Filgueiras-Rama *et al.*, "Chloroquine terminates stretch-induced atrial fibrillation more effectively than flecainide in the sheep heart," *Circulation: Arrhythmia and Electrophysiology*, vol. 5, no. 3, pp. 561-570, 2012.
- [10] S. Uetake *et al.*, "Frequency analysis of surface electrocardiograms (ECGs) in patients with persistent atrial fibrillation: Correlation with the intracardiac ECGs and implications for radiofrequency catheter ablation," *Journal of Arrhythmia*, vol. 30, no. 6, pp. 453-459, 2014.

Análisis wavelet de la señal de oximetría cerebral para el cálculo de la frecuencia de las compresiones del masaje de resucitación cardiopulmonar

A. Sanz-Pescador^{1,2}, E. Alonso², I. Isasi², E. Aramendi¹, A. Larrea³, R. Salaberria⁴

¹Departamento de Ingeniería de Comunicaciones, Universidad del País Vasco (UPV/EHU), Bilbao, España, {amaia.sanzp}@ehu.eus

²Departamento de Matemática Aplicada, Universidad del País Vasco (UPV/EHU), Bilbao, España.

³Emergentziak-Osakidetza, Bilbao, España.

⁴Emergentziak-Osakidetza, Donostia, España.

Resumen

La señal de oximetría cerebral de alta resolución temporal obtenida a través de espectroscopia del infrarrojo cercano refleja fluctuaciones debidas a las compresiones torácicas (CT) administradas durante el masaje de resucitación cardiopulmonar. El objetivo de este estudio fue desarrollar un método para detectar la frecuencia de las CT basado en el procesado de la señal de oximetría cerebral de pacientes en parada cardiorrespiratoria extrahospitalaria (PCREH). La base de datos de estudio constó de 284 segmentos extraídos de 30 pacientes de PCREH. Los segmentos incluían las señales de impedancia torácica y oximetría. El método propuesto analizó ventanas no-solapadas de 10 s de la señal de oximetría. Primero, se extrajo la componente inducida por las CT mediante la transformada wavelet estacionaria. Después, se realizó el control de calidad para evitar artefactos y/o señales de baja calidad. Finalmente, se calculó la frecuencia de las compresiones, f_{CT} , como el argumento máximo del espectro. La evaluación del método se realizó en términos de: mediana y rango interdecil (RID) del error absoluto entre la f_{CT} y la frecuencia de las CT de referencia calculada en la impedancia torácica en compresiones por minuto (cpm), y análisis de Bland-Altman y su rango de confianza del 90%. La mediana (RID) del error absoluto fue 0.62 (0.10 – 3.945) cpm con un intervalo de confianza de -2.72– 4.74 cpm para el 90.27% del total de las ventanas que superaron el control de calidad. Estos resultados evidencian la robustez y precisión del método que podría ser integrado en sistemas de monitorización.

1. Introducción

Las actuales guías de resucitación resaltan la importancia de la aplicación de un masaje de resucitación cardiopulmonar (RCP) de alta calidad en los episodios de parada cardiorrespiratoria extrahospitalaria (PCREH) [1]. El masaje RCP consta de series de 30 compresiones torácicas (CT) intercaladas con dos ventilaciones. El componente clave de una RCP efectiva y de alta calidad son las CT que deben aplicarse de manera ininterrumpida a una frecuencia de 100-120 compresiones por minuto (cpm), con una profundidad de entre 5-6 cm y garantizando la completa relajación del pecho del paciente tras cada compresión [2].

Existen dispositivos de ayuda a la RCP para monitorizar la calidad de las CT. Dichos dispositivos se basan en señales biomédicas registradas durante el tratamiento del episodio

PCREH, como por ejemplo, en la impedancia torácica [3] o en la profundidad de compresión basada en acelerómetros [4]. Sin embargo, estos dispositivos se limitan a cuantificar y reportar el rendimiento del rescatador teniendo como valores objetivo la frecuencia y profundidad de las CT recomendados por las guías de resucitación. No miden, por tanto, el efecto de la RCP en la hemodinámica del paciente, es decir, no proporcionan información clínica relevante sobre el estado del paciente ni el efecto de la terapia aplicada.

La espectroscopia del infrarrojo cercano (NIRS) se ha presentado recientemente como alternativa no invasiva para monitorizar la saturación regional de oxígeno en el cerebro (rSO₂) en episodios de PCREH. La tecnología NIRS permite calcular los cambios de concentración en oxihemoglobina (ΔO_2Hb), deoxihemoglobina (ΔHHb) y hemoglobina total (ΔcHb), a partir de los cuales se estima la saturación de oxígeno en los tejidos.

La oximetría cerebral ha sido mayoritariamente empleada en el ámbito de la PCREH para monitorizar la respuesta del paciente al tratamiento ya que el aumento del nivel de oxigenación cerebral está asociado al retorno de la circulación sanguínea espontánea [5]. Sin embargo, estudios recientes también muestran que las señales de oximetría de alta resolución temporal reflejan fluctuaciones debidas a las CT del masaje RCP y, por tanto, son capaces de monitorizar la frecuencia de las CT [6].

El objetivo de este estudio fue desarrollar un método basado en el procesado de la señal de oximetría cerebral para estimar la frecuencia de las CT durante el masaje RCP. Para ello, primero se realizó el análisis wavelet de la señal de oximetría cerebral con el objetivo de aislar la componente correspondiente a las CT y después, se desarrolló un algoritmo en el dominio de la frecuencia para la detección de la frecuencia fundamental de las CT.

2. Materiales

Los datos usados en este estudio provienen de 30 pacientes de PCREH tratados por Emergentziak-Osakidetza (Bilbao, Bizkaia) a través del monitor/desfibrilador LifePak 15 (Stryker, Kalamazoo, MI, EEUU) y el monitor de oximetría cerebral NIRO-200NX (Hamamatsu Photonics,

Hamamatsu, Japón) entre mayo de 2018 y septiembre de 2019. Para cada paciente se registraron las señales de impedancia torácica (IT) con una frecuencia de muestreo de 250 Hz y los cambios en la concentración de hemoglobina total (ΔcHb) en el lóbulo frontal izquierdo muestreada a 20 Hz. Las señales de cada paciente fueron visualizadas en una interfaz gráfica de usuario desarrollada en Matlab (Mathworks, Natick, MA, EEUU) que permitió la alineación temporal de las señales de los distintos equipos y la extracción de segmentos con una duración mínima de 30 s durante las series de CT. Así, la base de datos usada en este estudio se compuso de un total de 284 segmentos con una duración mediana (rango intercuartílico, RIC) de 83 (13-153) s con a un total de 49491 CT y a una mediana (RIC) de 1546 (1410) por paciente.

3. Métodos

La Figura 1 muestra el diagrama de bloques del método propuesto que analiza ventanas no-solapadas de 10 s de la señal de oximetría a lo largo de sus tres etapas. Primero, la componente debida a las CT, $s_{CT}(n)$, se extrae de la señal de oximetría, $s_{oxi}(n)$, a través de la transformada wavelet estacionaria (TWE). Posteriormente, se realiza el control de calidad de la señal para garantizar los estándares mínimos de calidad requeridos para que en la última etapa el algoritmo espectral realice el cómputo de la frecuencia de las compresiones, f_{CT} , de forma fiable.

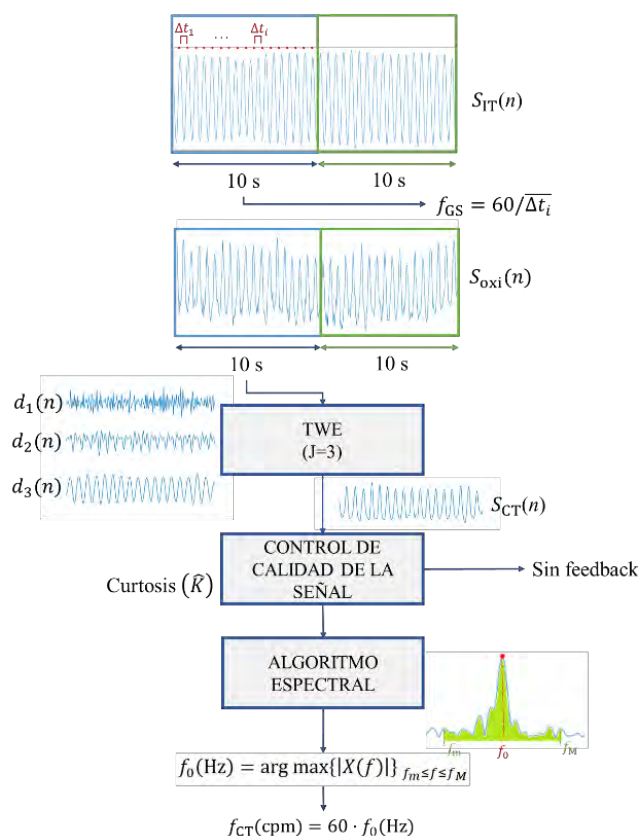


Figura 1. Diagrama de bloques del método para el cálculo de la frecuencia de las CT.

3.1. Análisis wavelet

La extracción de la $s_{CT}(n)$ se realizó mediante la TWE que consiste en la aplicación recursiva de dos filtros: un paso-bajo, $h_{j-1}(n)$, y un paso-alto, $g_{j-1}(n)$, para obtener los coeficientes de aproximación, $a_j(n)$ y detalle, $d_j(n)$, del siguiente nivel de descomposición:

$$a_j(n) = h_{j-1}(n) * a_{j-1}(n) \quad (1)$$

$$d_j(n) = g_{j-1}(n) * a_{j-1}(n) \quad (2)$$

donde $a_0(n) = s_{oxi}(n)$, y $*$ denota la convolución. En la etapa j los filtros son los mismos que en la etapa 0 pero incrementando su longitud a través de la inserción de $2^j - 1$ ceros entre muestras.

En concreto, se utilizó una wavelet madre Coiflet-3 para realizar la descomposición de $s_{oxi}(n)$ en $J = 3$ niveles que dieron lugar a los coeficientes de detalle $d_1 - d_3$ y el coeficiente de aproximación a_3 . La síntesis de $s_{CT}(n)$ se realizó usando los coeficientes d_2 y d_3 (banda 1.25 - 5 Hz). El resto de los coeficientes detalle y aproximación se igualaron a cero. Dicha síntesis se llevó a cabo mediante la evaluación recursiva de la siguiente expresión:

$$a_{j-1}(n) = \frac{1}{2}(g'_j(n) * a_j(n) + h'_j(n) * d_j(n)) \quad (3)$$

para $j = J, \dots, 1$, donde $g'_j(n)$ y $h'_j(n)$ son los denominados filtros de síntesis y el coeficiente sintetizado $a_0(n)$ se corresponde con $s_{CT}(n)$.

3.2. Control de calidad

La señal $s_{CT}(n)$ fue sometida a un control de calidad para evitar errores sustanciales en el cómputo de la f_{CT} originados por la baja calidad de señal y/o alta presencia de ruido. El índice de calidad utilizado para realizar dicho control fue la curtosis, el cuarto momento estandarizado de la distribución de $s_{CT}(n)$, que describe la forma de las colas de una distribución en relación con su forma general. Así, valores elevados de curtosis corresponden a una distribución con colas más gruesas, es decir, una mayor concentración de valores de $s_{CT}(n)$ muy lejos de su media. La estimación empírica de la curtosis, \hat{K} , de la señal $s_{CT}(n)$ viene dada por:

$$\hat{K} = \frac{1}{N} \sum_{n=1}^N \left[\frac{s_{CT}(n) - \hat{\mu}}{\hat{\sigma}} \right]^4 \quad (4)$$

donde N representa el número de muestras de la señal $s_{CT}(n)$, y $\hat{\mu}$ y $\hat{\sigma}$ son las estimaciones empíricas de su media y desviación estándar, respectivamente. La Figura 2 muestra, de izquierda a derecha, una señal $s_{CT}(n)$ de calidad suficiente ($\hat{K}=1.68$) y de calidad insuficiente ($\hat{K}=10.42$), respectivamente. La señal $s_{CT}(n)$ de suficiente calidad debería tener una naturaleza sinusoidal quasi-periódica (panel izquierdo de la Figura 2) por lo que se espera una $\hat{K} \approx 1.5$. Valores muy superiores de \hat{K} indicarán presencia de valores atípicos en $s_{CT}(n)$ presumiblemente debido a la alta presencia de ruido (panel derecho de la Figura 2).

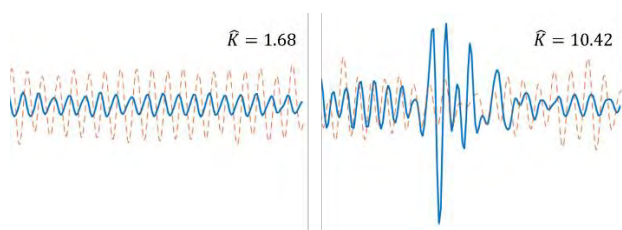


Figura 2. Ejemplos de $s_{CT}(n)$ (línea azul continua) con alta (izquierda) y baja (derecha) calidad de señal. La línea roja discontinua muestra la señal de impedancia torácica concurrente, señal utilizada para marcar los instantes de tiempo de las CT de referencia.

3.3. Algoritmo espectral

En la tercera etapa se estimó la f_{CT} en el dominio de la frecuencia. Se calculó primero la transformada rápida de Fourier de 4096 puntos, $X(f)$, de la señal $s_{CT}(n)$ que previamente había superado el control de calidad. Después, la frecuencia de las CT se calculó como $f_{CT}(\text{cpm}) = 60 \cdot f_0(\text{Hz})$ siendo f_0 :

$$f_0(\text{Hz}) = \arg \max\{|X(f)|\}_{f_m \leq f \leq f_M} \quad (5)$$

que representa el argumento del máximo de $|X(f)|$ en la banda de frecuencias de interés donde $f_m = 1 \text{ Hz}$ (60 cpm) y $f_M = 4 \text{ Hz}$ (240 cpm).

En ocasiones, la amplitud espectral más grande puede corresponder erróneamente al segundo armónico de $X(f)$ como se aprecia en el ejemplo de la Figura 3. Para evitar estos errores se introdujo la siguiente modificación en el método cuando en primera instancia $f_{CT}(\text{cpm}) \geq 170$. Primero, se calculó, R_A , la razón de las amplitudes máximas alrededor de $f_{CT}/2$ y f_{CT} :

$$R_A = \frac{\max_{|f-f_{CT}/2| < \Delta f} |X(f)|}{\max_{|f-f_{CT}| < \Delta f} |X(f)|} \quad (6)$$

Después, R_a , la razón de las áreas alrededor de $f_{CT}/2$ y f_{CT} :

$$R_a = \frac{\sum_{|f-f_{CT}/2| < \Delta f} |X(f)|}{\sum_{|f-f_{CT}| < \Delta f} |X(f)|} \quad (7)$$

donde $\Delta f = 5 \text{ cpm}$. Cuando ambas razones, R_A y R_a , superaron el umbral establecido de 0.3, se aplicó la corrección oportuna de la f_{CT} , pasando ésta a ser la mitad del valor inicialmente computado.

3.4. Evaluación

La evaluación del método se realizó en términos de mediana y rango interdecil (RID) del error absoluto entre la f_{CT} y la frecuencia de las CT de referencia, f_{GS} . El cálculo de la f_{GS} se realizó a través de la siguiente ecuación:

$$f_{GS}(\text{cpm}) = 60/\overline{\Delta t_i} \quad (6)$$

donde $\overline{\Delta t_i}$ denota la media de la duración (en segundos) de los intervalos entre los instantes de las CT en la señal IT, $s_{IT}(n)$ (línea roja discontinua en la Figura 1). Finalmente, como complemento para evaluar la fiabilidad y robustez del método, se calculó el gráfico de Bland-Altman, así como su rango de confianza del 90%.

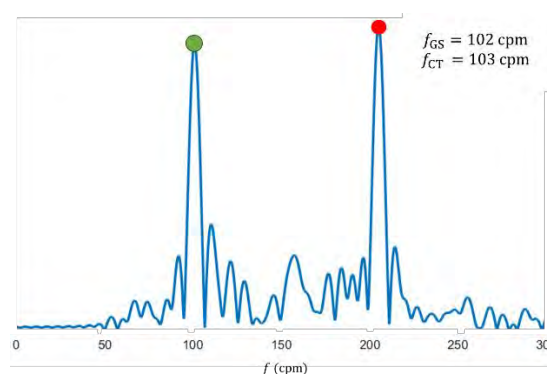


Figura 3. Ventana analizada en la que la f_{CT} fue inicialmente erróneamente calculada debido a que el argumento máximo del espectro estaba situado en el segundo armónico de la frecuencia fundamental (206 cpm, punto rojo) y en la que posteriormente se aplica correctamente la corrección a la mitad del valor (103 cpm, punto verde) dado que $f_{GS} = 102 \text{ cpm}$.

4. Resultados y discusión

La Figura 4 muestra la mediana (p_{50}), percentil 10 (p_{10}) y percentil 90 (p_{90}) del error absoluto en función del porcentaje de ventanas analizadas según el control de calidad aplicado. Se aprecia una disminución de la mediana del error absoluto a medida que el índice de calidad comienza a ser efectivo en la detección de ventanas con baja calidad en las que se cometen errores sustanciales. Dicha disminución se debe mayoritariamente a la caída del p_{90} del error que es muy pronunciada hasta la inclusión del 90.27% (umbral de \hat{K} de 6.5) de ventanas analizadas donde la mediana (RID) del error absoluto fue de 0.64 (0.10 – 3.945) cpm. La magnitud de la caída del p_{90} se estabiliza para un porcentaje de ventanas analizadas inferior al 90%.

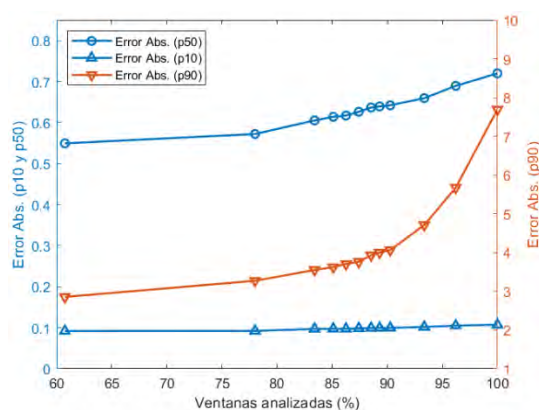


Figura 4. Mediana (RID) del error absoluto en función del porcentaje de ventanas que superaron el control de calidad. La mediana (p_{50} , en círculos azules) y el percentil 10 (p_{10} , triángulos azules) corresponden en amplitud al eje e izquierdo (en azul), mientras que el percentil 90 (p_{90} , triángulos rojos invertidos) corresponden en amplitud al eje y derecho.

En la Figura 5 se representa el gráfico de Bland-Altman junto con su rango de confianza del 90% para un porcentaje de ventanas analizadas del 90.27%. El error, $f_{GS} - f_{CT}$, medio fue de 0.31 cpm, mientras que el rango de confianza del 90% fue de -2.72– 4.74 cpm. Puede apreciarse como un muy alto porcentaje de las frecuencias se encuentran

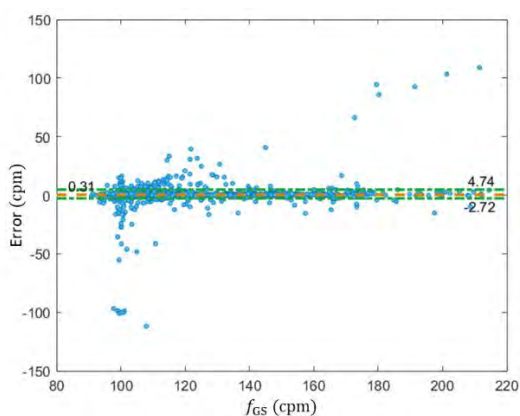


Figura 5. Gráfico de Bland-Altman. La línea roja discontinua representa la media del error, mientras que las líneas verdes discontinuas representan el intervalo de confianza del 90%.

concentradas en el intervalo 100 – 120 cpm recomendado por las guías de resucitación, y es por ello que en esta franja se encuentran también los errores más sustanciales. Existen ventanas analizadas con errores de hasta -100 cpm en esta franja como la mostrada en el panel izquierdo de la Figura 6 donde $f_{GS} = 98$ cpm, pero la $s_{CT}(n)$ muestra una primera componente armónica muy marcada que lleva al algoritmo espectral a la detección de $f_{CT} = 195$ cpm, y, por tanto, a cometer un error de -97 cpm, al no cumplirse las condiciones ($R_A < 0.3$ y $R_a < 0.3$) para aplicar su corrección a 97.5 cpm. En el panel derecho de la Figura 6 se muestra una ventana analizada que presenta un error de 109 cpm ($f_{CT} = 212$ cpm y $f_{CT} = 103$ cpm) debido a que, aunque el argumento del máximo del espectro se encuentra en 211 cpm, el algoritmo espectral aplica erróneamente la corrección a 103 cpm ya que tanto R_A como R_a superan el umbral de 0.3 establecido para llevar dicha corrección a cabo.

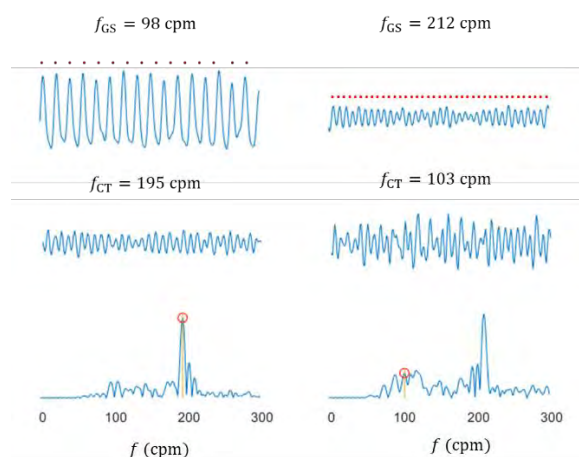


Figura 6. Dos ejemplos de ventanas de 10 s con errores muy elevados. De arriba abajo, se representan la $s_{IT}(n)$, la $s_{CT}(n)$ y el módulo de la transformada de Fourier correspondiente a la $s_{CT}(n)$, $|X(f)|$.

5. Conclusiones

Este estudio ha propuesto un método basado en el análisis wavelet de la señal de oximetría cerebral para la detección de la frecuencia de las CT durante la RCP en casos de PCREH. En concreto, el método utiliza primero la TWE para extraer la componente relativa a las CT, $s_{CT}(n)$, de la

señal de oximetría. Después, realiza un control de calidad de $s_{CT}(n)$ para evitar errores importantes en el cálculo de la frecuencia de las CT originados por la baja calidad de señal y/o alta presencia de ruido. En su última etapa el método consta de un algoritmo espectral que computa la frecuencia de las CT como el argumento del máximo del espectro de $s_{CT}(n)$ en la banda de frecuencias correspondiente a 60 – 240 cpm. El bajo error absoluto mediano (0.62 cpm) y el estrecho intervalo de confianza del 90% (-2.72– 4.74 cpm) reportados por el método (ambos calculados para el 90.27% de ventanas analizadas) evidencian la robustez y precisión del método propuesto. El método podría ser implementado en los equipos de oximetría cerebral de alta resolución temporal, lo cual requeriría únicamente de modificaciones software, sin necesidad de realizar cambios en su hardware. Esto ayudaría al rescatador a recibir feedback sobre los cambios hemodinámicos que el paciente presenta debido a la calidad de las CT, pudiendo así ajustar la terapia a las necesidades del paciente en tiempo real.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el Ministerio de Ciencia, Innovación y Universidades a través del proyecto PID2021-122727OBI00, conjuntamente con el Fondo Europeo de Desarrollo Regional (FEDER), y en parte por el Gobierno Vasco por medio del proyecto IT1717-22.

Referencias

- [1] J. Soar, J. P. Nolan and B. W. Böttiger *et al.*, “European Resuscitation Council Guidelines for Resuscitation 2015,” *Resuscitation*, vol. 95, pp. 100–147, Oct. 2015.
- [2] P. A. Meaney, B. J. Bobrow and M. E. Mancini *et al.*, “Cardiopulmonary Resuscitation Quality: Improving Cardiac Resuscitation Outcomes Both Inside and Outside the Hospital,” *Circulation*, vol. 128, no. 4, pp. 417–435, Jul. 2013.
- [3] F. S. Stecher, J.-A. Olsen, R. E. Stickney, and L. Wik, “Transthoracic impedance used to evaluate performance of cardiopulmonary resuscitation during out of hospital cardiac arrest,” *Resuscitation*, vol. 79, no. 3, pp. 432–437, Dec. 2008.
- [4] S. O. Aase, Myklebust, et al. Compression depth estimation for CPR quality assessment using DSP on accelerometer signals. *IEEE Trans. on Biomedical Engineering*, vol. 49, no. 3, March 2002.
- [5] C. Genbrugge, I. Meex and W. Boer *et al.*, “Increase in cerebral oxygenation during advanced life support in out-of-hospital patients is associated with return of spontaneous circulation,” *Crit Care*, vol. 19, no. 1, p. 112, 2015.
- [6] Y. Koyama, T. Wada and B. D. Lohman *et al.*, “A new method to detect cerebral blood flow waveform in synchrony with chest compression by near-infrared spectroscopy during CPR,” *The American Journal of Emergency Medicine*, vol. 31, no. 10, pp. 1504–1508, Oct. 2013.

Contacto

Amaia Sanz
 Escuela de Ingeniería
 Departamento de Ingeniería de Comunicaciones,
 Alameda Urquijo, s/n 48013 Bilbao, España
 Email: amaia.sanzp@ehu.eus

Development of a neuroblastoma-on-a-chip bioengineered platform enabling the translational study of emerging nanotherapies

Gema Quiñonero¹, Juan Gallo², Andreia Magalhaes², Josep Samitier^{1,3,4}, Aránzazu Villasante^{1,3}

¹Institute for Bioengineering of Catalonia (IBEC), Barcelona, Spain

²International Iberian Nanotechnology Laboratory (INL), Braga, Portugal

³Department of Electronic and Biomedical Engineering, University of Barcelona, Barcelona, Spain.

⁴Biomedical Research Networking Center in Bioengineering, Biomaterials, and Nanomedicine (CIBER-BBN), Spain.

Summary

Neuroblastoma (NB) is a developmental cancer with a high level of vascularization. Around 50% of the affected children will develop high-risk NB, and more than half will relapse. The inability to represent complexity and interactions of the human NB hampers clinical translation. In this project, we developed an NB-on-a-chip that can partially recapitulate tumor's behavior in vitro, allowing testing the efficiency of novel hyperthermia-based nanotherapies. Hyperthermia treatments demonstrated a reduction of cell proliferation specifically in the models treated with nanoparticles. For the biomimetic nanoparticles characterization, two polydisperse populations (exosomes and microvesicles) of EVs/EV-NPs were found, presenting negative ζ -potential. Perfusion system in the NB-on-a-chip allowed the systemic administration of nanotherapies mimicking tumor vasculature overnight.

1. Introduction

Neuroblastoma is a developmental cancer with high level of vascularization. It usually arises as a defect in the development of neural crest cells and is normally located in the adrenal glands. Unfortunately, neuroblastoma is the most common extracranial solid tumor in children under five [1-5]. Neuroblastoma cells can present three different phenotypic variants: type N, of neuroblastic character, characterized by forming small aggregates; type S, which is substrate adherent and present contact inhibition; and type I, which has mixed phenotype [2, 6, 7]. Approximately 50% of the children coursing this disease will develop high-risk NB. This tumor is characterized by its resistant character which leads to only a 50% of survival rate; and accounts for 15% of all pediatric oncology deaths [3-5, 8, 9].

Current models, such as 2D conventional cultures or animals, present some limitations: one way or another, they fail in the representation of human NB. The key feature to be represented is tumor microenvironment in all their aspects such as composition, cell niche... but, importantly, in mechanical stimuli [8, 10-14]. For that reason, recent studies focus on the tumor-on-a-chip approach. This microfluidic technique provides not only structural but also functional recapitulation of the human NB tumor [10, 13, 15, 16]. Besides, bioengineering also pays attention to emerging therapies based on nanomedicine principles. It consists of the use of nanoscale or nanostructured products

as diagnostic or therapeutic medical tools by exploiting their unique physicochemical features [17]. Particularly, hyperthermia must be highlighted, which exposes the tumor to temperatures over the physiological limits, inducing cell death or sensitization [18-23]. In these treatments, nanoparticles (NPs) play an essential role since they are able to transform external stimuli into heat. Besides, their size provides an additional level of selectivity thanks to enhanced permeability and retention effect [24, 25]. Among all the suitable molecules for hyperthermia treatments, superparamagnetic iron oxide nanoparticles (SPIONs) are worth to be mentioned; specially, due to their double function of magnetism and NIR (near-infrared) light absorption [22, 26]. Regarding hyperthermia treatments, we should highlight two of the techniques: magnetic hyperthermia (MH), in which the whole tumor is irradiated with alternating magnetic fields SPIONs transform the magnetic energy into heat [21, 27]; and photothermal therapy (PTT), in which a specific region of the tumor is locally irradiated with NIR laser and SPIONs transform optical energy into heat [18, 19, 24, 28].

However, nanomedicine has tried to go further. They propose the customization of the nanoparticles according to the characteristics of the native tumor [25, 29]. So, biomimetics arise as the attempt of combining the functionality of biological materials with the adaptability of synthetic ones [30, 31]. As potential candidates, we should mention extracellular vesicles (EVs), which are key players in cross communication between tumor cells, acting as specific messengers [2, 31, 32]. They can be classified into exosomes or microvesicles depending on their size [33, 34]. Thus, the ultimate proposal is to use these extracellular vesicles to camouflage the nanoparticles (EV-NPs), which will provide specific active targeting of tumor cells [32, 34, 35]. It has been demonstrated that the size and the content of released EVs highly depend on the dimensionality, composition and stiffness forming the TME [36]. Therefore, the above-mentioned necessity for reliable NB-tissue engineered (TE) models present a highlighted limitation for the development of emerging nanotherapies.

The hypothesis of the project states that an NB-on-a-chip will allow a superior recapitulation of NB biological behavior in vitro and predict therapy response. Therefore,

the main objective is to develop a bioengineered platform based on the tumor-on-a-chip technology which enables nanotherapies testing in neuroblastoma tumors *in vitro*. The project is organized around three sub-objectives: (1) to study the effect of the addition of nanoparticles to 3-Dimensional biomaterials mimicking the composition of NB tumors when subjected to hyperthermia-based nanotherapies (magnetic hyperthermia, MH, and photothermal therapy, PTT), (2) to fabricate and characterize biomimetic iron-based nanoparticles, and (3) to develop an NB-on-a-chip which enables perfusion of the model so that it is suitable for testing systemic nanotherapies.

2. Methodology

2.1. Culture of NB cell lines

Neuroblastoma cell lines used in this project are SK-N-BE(2) (Type N), LAN-1 (Type N), and SK-N-LP (Type I) [2, 6, 7]. They are all commercial and obtained from ATCC human cell culture collection. Cells were cultured with RPMI-1640 (with L-glutamine and sodium bicarbonate, Sigma-Aldrich, UK) with 10 % FBS (Fetal Bovine Serum, Gibco, Brazil) and 1% PS (Penicillin Streptomycin with 10000 units/ml Penicillin and 10000 µg/ml Streptomycin, Gibco, USA). Cells are incubated at 37°C and 5% CO₂. Live/dead assays have been used to measure cell viability. 16.2 mM Hoechst 33342 (Life Technologies, USA), 2 µM Calcein and 4 µM Ethidium homodimer-1 (LIVE/DEAD Kit, Life Technologies, USA) stained nuclei, live and dead cells, respectively.

2.2. Col1-HA 3D porous scaffolds

Three-dimensional scaffolds were fabricated using freeze-drying technique with collagen I (Col1) (Collagen Type I 9.48 mg/ml, Corning, USA) and hyaluronic acid (HA) (Sodium Hyaluronate low molecular weight, Cantipro a.s., Czech Republic) as materials. Optimal concentration is 1*10⁶ cells/scaffold in 2 ml exo-free culture medium. Scaffolds are dried and placed in a falcon with the culture medium and the cells. They are incubated at 37°C 5% CO₂ for 4 hours in a rotatory platform. Ultimately, each scaffold is incubated in a well with 2 ml exo-free culture medium for one week.

NPs used in this project are magnetite-based SPIONs (FDA approved). At day 7 of seeding (day 0 of treatment from now on), 50 µg Fe/ml are added to the scaffolds. Scaffolds will be tissue-engineered neuroblastoma models (TE-NB models) from now on. They are subjected to magnetic hyperthermia or photothermal therapy in the International Iberian Nanotechnology Laboratory (Braga, Portugal) at day 1 or 2 after the addition of NPs. Magnetic Resonance Imaging is used to check internalization. Histological slides are prepared. Supernatant is collected at days 0, 1 and 5 after adding NPs, centrifuged, and again collected and stored at -80°C for a posterior evaluation of the products. For the isolation and characterization, Total Exosome Isolation (from cell culture media) (Invitrogen, Lithuania) is used with the collected supernatant. Samples are incubated at 4°C overnight. Samples are centrifuged at 4°C, and pellet is resuspended in PBS. Dynamic Light

Scattering (DLS) (Zetasizer Nano ZS, Malvern) is used to analyze the population size (ZEN0040 cell, Zetasizer nano series, Malvern, Germany) and the ζ-potential (DTS1070 cell, Zetasizer nano series, Malvern, Germany) of the EVs or EV-NPs collected from the TE-NB models.

At day 5, TE-NB models are fixed to obtain histological sections (hematoxylin-eosin and blank). Prussian blue staining (HEMATOGNOST Fe - staining kit for the detection of free ionic iron Fe³⁺ in cells, Sigma-Aldrich, Germany) is used to study internalization of the NPs inside the NB cells. For the samples treated with hyperthermia, histological slides were obtained at 24 and 48h and stained for ki67 antibody (kit R.T.U. VECTASTIN, USA), to visualize cell proliferation in brown.

2.3. Microfluidic chips

Two fabrication techniques were explored: home-made chips made from PDMS (Polydimethylsiloxane) (SYLGARD 184 Silicone Elastomer Kit, Dow, Germany); and commercial chips (idenTx 3 chip DAX-1, AimbioTech, Singapore), that were the ones finally used for the experiments. Cells are seeded in a fibrinogen/collagen-based hydrogel in the middle channel. Desired concentration is 30000 cells/chip. Components of the hydrogel are: 9.4% M199 culture medium (10x, With Hanks salts, without L-glutamine and sodium bicarbonate, Sigma-Aldrich, UK), 0.3% HEPES (Sigma, USA) and NaCHO₃ (Sigma, USA) buffers, 6% aprotinin 7500 KIU/ml (Wako Chemical, Japan), 20% fibrinogen (Fibrinogen from bovine plasma, Sigma, USA) and 24% Col1. Then, cells are added. Finally, 20% diluted thrombin (Sigma, USA) is added as polymerization agent. Just before polymerization, the hydrogel is introduced in the middle channel. Reservoirs are filled with culture medium.

Chips are connected to a perfusion system (IPC High Precision Multichannel Pump, ISMATEC) at day 1 after seeding. It provides flux to the cells in the hydrogel at 0.143 ml/min. Reservoirs are filled with medium without NPs (ctl) or with 50 µg Fe/ml (+NPs). Administration is maintained overnight. Samples are observed with confocal microscopy (Dragonfly 505, Andor Technologies Inc).

3. Results and discussion

3.1. Scaffold-based TE-NB models

The extracellular matrix (ECM) in NB is enriched in Col1 and HA [2], so we used those materials to fabricate the scaffolds in this project. The estimation of the diameter of the models is 4.97 ± 0.35 mm. Live/dead assays of the TE-NB models at day 7 of seeding (day 0 of treatment) show high percentage of the cells stained in green for all the cell lines, in comparison to the low percentage of cells in red. This suggests that they enabled the culture of NB cells for at least 7 days. Internalization of the NPs was studied by three different procedures: (1) intrinsic magnetic characteristic of the NPs let us use MRI images to show that the NPs were inside the scaffolds at 48h; (2) fluorescence microscopy took advantage of the Rhodamine B functionalization of the NPs to confirm that they were trapped in the scaffold-based TE-NB models at day 5; and

(3) Prussian Blue staining confirmed that the NPs were, in fact, effectively internalized by NB cells.

After subjecting the models to hyperthermia treatments, samples were stained for ki67 (proliferation marker) to assess the level of malignant cells destruction. For both treatments (MH and PTT), cell proliferation is reduced in the cases with NPs and not affected in the control conditions. Samples subjected to MH present a more homogeneous reduction of the proliferative cells while samples subjected to PTT present clearly an absence of cell proliferation only in the irradiated zone. However, tumor is not completely removed with any of the hyperthermia treatments, highlighting the need for long-term studies and combination of techniques. Lastly, hematoxylin-eosin-stained histological sections of these models were analyzed to evaluate the effect of the increment of temperature in the surroundings (in the biomaterial). None of the treatments affect the structure of the scaffold-based TE-NB models in any of the conditions, indicating selective effect of the treatment in the tumor cells with the internalized NPs.

3.2. Biomimetic NPs

Regarding size, both EVs/EV-NPs present two polydisperse populations corresponding to exosomes (the ones with smaller diameter), and microvesicles (the ones with bigger diameter). Normally, the diameter of the vesicles with the camouflaged NPs tends to be bigger than the corresponding control condition. ζ -potential is in negative in all cases, being less negative in the vesicles with the camouflaged NPs inside them, as expected (the positive potential of the NPs, when camouflaged, increases the negative potential of the EVs).

3.3. NB-on-a-chip

This time, the ECM composition to embed NB cells presented different composition: a fibrin network was provided by enzymatically transforming the fibrinogen with thrombin. This network will provide the basis for adding endogenous vasculature to the model in the future. The NB-on-a-chip demonstrated to be viable for the culture of NB cells under static condition for 1 day. Perfusion system allowed the systemic administration of nanotherapies mimicking tumor vasculature. Besides, the flux in the pump simulates capillary flux [37, 38] so that the representation is closer to the real systemic administration. After overnight perfusion of the treatment, cell viability is not qualitatively affected by the addition of NPs, which are effectively internalized by NB cells and accumulated in the cytoplasm.

The key ultimate objective of the project is translationality. Current preclinical models provide inadequate results which cannot be further translated in the research chain. Then, these models will provide not only the reduction of development time and expenses but, most importantly, the improvement of the patients' quality of life.

4. Conclusions

We successfully developed an NB-on-a-chip that can partially recapitulate tumor's behavior in vitro, allowing

testing of the efficiency of novel hyperthermia-based nanotherapies.

Acknowledgements

The Nanobioengineering Lab is a member of the Biomedical Research Networking (CIBER), Spain, funded by the VI National R&D&I Plan 2008-2011 and the Instituto de Salud Carlos III (RD16/0006/0012; RD16/0011/0022) with the support of the European Regional Development Fund (ERDF), and Generalitat de Catalunya (2023SGR 015452). AV is supported by AECC, the NEN Association, and the R+D+i plan PID2020-117977RA-I00.

References

- [1] Le Grand M, Mukha, A, Püschel J, Valli E, Kamili A, Vittorio O, Dubrovskaya A, Kavallaris M. Interplay between MycN and c-Myc regulates radioresistance and cancer stem cell phenotype in neuroblastoma upon glutamine deprivation. *Theranostics*, vol 10, sup 14, 2020, pp 6411-6429 (ISSN: 1838-7640).
- [2] Villasante A, Godier-Furnemont A, Hernandez-Barranco A, Le Coq J, Boskovic J, Peinado H, Mora J, Samitier J, Vunjak-Novakovic G. Horizontal transfer of the stemness-related markers EZH2 and GLI1 by neuroblastoma-derived extracellular vesicles in stromal cells. *Translational research: the journal of laboratory and clinical medicine*, vol 237, 2021, pp 82–97 (ISSN 1931-5244).
- [3] Kamijo T. Role of stemness-related molecules in neuroblastoma. *Pediatric Research*, vol 71, sup 4 Pt 2, 2012, pp 511–515 (ISSN: 1530-0447).
- [4] Villasante A, Sakaguchi K, Kim J, Cheung NK, Nakayama M, Parsa H, Okano T, Shimizu T, Vunjak-Novakovic G. Vascularized Tissue-Engineered Model for Studying Drug Resistance in Neuroblastoma. *Theranostics*, vol 7, sup 17, 2017, pp 4099–4117 (ISSN: 1838-7640).
- [5] Chakrabarti L, Abou-Antoun T, Vukmanovic S, Sandler AD. Reversible adaptive plasticity: a mechanism for neuroblastoma cell heterogeneity and chemo-resistance. *Frontiers in Oncology*, vol 2, sup 82, 2012 (ISSN: 2234-943X).
- [6] Mitchell CB, O'Neill GM. Cooperative cell invasion: matrix metalloproteinase-mediated incorporation between cells. *Molecular Biology of the Cell*, vol 27, sup 21, 2016, pp 3284–3292 (ISSN: 1939-4586).
- [7] Acosta S, Lavarino C, Paris R, Garcia I, de Torres C, Rodríguez E, Beleta H, Mora J. Comprehensive characterization of neuroblastoma cell line subtypes reveals bilineage potential similar to neural crest stem cells. *BMC Developmental Biology*, vol 9, sup 1, 2009, pp 12 (ISSN:1471-213X).
- [8] Ning L, Shim J, Tomov ML, Liu R, Mehta R, Mingee A, Hwang B, Jin L, Mantalaris A, Xu C, Mahmoudi M, Goldsmith KC, Serpooshan V. A 3D Bioprinted in vitro Model of Neuroblastoma Recapitulates Dynamic Tumor-Endothelial Cell Interactions Contributing to Solid Tumor Aggressive Behavior. *Advanced Science*, vol 9, sup 23, 2022, pp 2200244 (ISSN: 2198-3844).
- [9] Quinn CH, Beierle AM, Hutchins SC, Marayati R, Bownes LV, Stewart JE, Markert HR, Erwin MH, Aye JM, Yoon KJ, Friedman GK, Willey CD, Markert JM, Beierle EA. Targeting High-Risk Neuroblastoma Patient-Derived Xenografts with Oncolytic Virotherapy. *Cancers*, vol 14, sup 3, 2022, pp 762 (ISSN: 2072-6694).
- [10] Marzagalli M, Pelizzoni G, Fedi A, Vitale C, Fontana F, Bruno S, Poggi A, Dondero A, Aiello M, Castriconi R, Bottino C, Scaglione S. A multi-organ-on-chip to

- recapitulate the infiltration and the cytotoxic activity of circulating NK cells in 3D matrix-based tumor model. *Frontiers in Bioengineering and Biotechnology*, vol 10, 2022, pp 945149 (ISSN: 2296-4185).
- [11] Marrella A, Dondero A, Aiello M, Casu B, Olive D, Regis S, Bottino C, Pende D, Meazza R, Caluori G, Castriconi R, Scaglione S. Cell-Laden Hydrogel as a Clinical- Relevant 3D Model for Analyzing Neuroblastoma Growth, Immunophenotype, and Susceptibility to Therapies. *Frontiers in Immunology*, vol 10, 2019 (ISSN: 1664-3224).
- [12] Eichler M, Jahnke H, Krinke D, Müller A, Schmidt S, Azendorf R, Robitzki AA. A novel 96-well multielectrode array based impedimetric monitoring platform for comparative drug efficacy analysis on 2D and 3D brain tumor cultures. *Biosensors & Bioelectronics*, vol 67, 2015, pp 582–589 (ISSN: 1873-4235).
- [13] Sontheimer-Phelps A, Hassell B, Ingber D. Modelling cancer in microfluidic human organs-on-chips. *Nature Reviews Cancer*, vol 19, sup 2, 2019 (ISSN: 1474-1768).
- [14] Monferrer E, Martín-Vañó S, Carretero A, García-Lizarribar A, Burgos-Panadero R, Navarro S, Samitier J, Noguera R. A three-dimensional bioprinted model to evaluate the effect of stiffness on neuroblastoma cell cluster dynamics and behavior. *Scientific Reports*, vol 10, sup 1, 2020, pp 6370 (ISSN: 2045-2322).
- [15] Wang Y, Cuzzucoli F, Escobar A, Lu S, Liang L, and Wang S. Tumor-on-a-chip platforms for assessing nanoparticle-based cancer therapy. *Nanotechnology*, vol 29, sup 33, 2018, pp 332001 (ISSN: 1361-6528).
- [16] Nothdurfter D, Ploner C, Coraça-Huber DC, Wilflingseder D, Müller T, Hermann M, Hagenbuchner J, Ausserlechner M. 3D bioprinted, vascularized neuroblastoma tumor environment in fluidic chip devices for precision medicine drug testing. *Biofabrication*, vol 14, sup 3, 2022 (ISSN: 1758-5090).
- [17] Guin D, Thakran S, Singh P, Ramachandran S, Hasija Y, Kukreti R. Chapter 1 - Translational biotechnology: A transition from basic biology to evidence-based research.. Academic Press, 2021.
- [18] Fernandes N, Rodrigues C, Moreira A, Correia I. Overview of the application of inorganic nanomaterials in cancer photothermal therapy. *Biomaterials Science*, vol 8, sup 11, 2020, pp 2990–3020 (ISSN: 2047-4849).
- [19] Estelrich J, Busquets MA. Iron Oxide Nanoparticles in Photothermal Therapy. *Molecules*, vol 23, sup 7, 2018, pp 1567 (ISSN: 1420-3049).
- [20] Altanerova U, Babincova M, Babinec P, Benejova K, Jakubecova J, Altanerova V, Zduriencikova M, Repiska V, Altaner C. Human mesenchymal stem cell-derived iron oxide exosomes allow targeted ablation of tumor cells via magnetic hyperthermia. *International Journal of Nanomedicine*, vol 12, 2017 (ISSN: 1176-9114).
- [21] Parsa Montazersaheb, Elahe Pishgahzadeh, Vahid Bayrami Jahani, Raheleh Farahzadi, and Soheila Montazersaheb. Magnetic nanoparticle-based hyperthermia: A prospect in cancer stem cell tracking and therapy. *Life Sciences*, vol 323, 2023, pp 121714 (ISSN: 1879-0631).
- [22] Wlodarczyk A, Gorgon S, Radon A, Bajdak-Rusinek K. Magnetite Nanoparticles in Magnetic Hyperthermia and Cancer Therapies: Challenges and Perspectives. *Nanomaterials*, vol 12, sup 11, 2022 (ISSN: 2079-4991).
- [23] Crezee J, Franken N, Oei A. Hyperthermia-Based Anti-Cancer Treatments. *Cancers*, vol 13, sup 6, 2021, pp 1240 (ISSN 2072-6694).
- [24] Liang C, Xu L, Song G, Liu Z. Emerging nanomedicine approaches fighting tumor metastasis: animal models, metastasis-targeted drug delivery, phototherapy, and immunotherapy. *Chemical Society Reviews*, vol 45, sup 22, 2016, pp 6250–6269 (ISSN: 1460-4744).
- [25] Swetha KL, Roy A. Tumor heterogeneity and nanoparticle-mediated tumor targeting: the importance of delivery system personalization. *Drug Delivery and Translational Research*, vol 8, sup 5, 2018, pp 1508–1526 (ISSN: 2190-3948).
- [26] Revia R, Zhang M. Magnetite nanoparticles for cancer diagnosis, treatment, and treatment monitoring: recent advances. *Materials Today*, vol 19, sup 3, 2016, pp 157–168 (ISSN: 1369-7021).
- [27] Ahmed S, Rajak BL, Gogoi M, Sarma HD. Chapter 16 - Magnetic nanoparticles mediated cancer hyperthermia. Academic Press, 2020.
- [28] Li X, Lovell J, Yoon J, Chen X. Clinical development and potential of photothermal and photodynamic therapies for cancer. *Nature Reviews Clinical Oncology*, vol 17, sup 11, 2020, pp 657–674 (ISSN: 1759-4782).
- [29] Wang H, Liu Y, He R, Xu D, Zang J, Weeranoppanant N, Dong H, Li Y. Cell membrane biomimetic nanoparticles for inflammation and cancer targeting in drug delivery. *Biomaterials Science*, vol 8, sup 2, 2020 (ISSN: 2047-4849).
- [30] Jin J, Bhujwala Z. Biomimetic Nanoparticles Camouflaged in Cancer Cell Membranes and Their Applications in Cancer Theranostics. *Frontiers in Oncology*, vol 9, 2020, pp 1560 (ISSN: 2234-943X).
- [31] Sancho-Albero M, Encabo-Berzosa M, Beltrán-Visiedo M, Fernández-Messina L, Sebastián V, Sánchez-Madrid F, Arruebo M, Santamaría J, Martín-Duque P. Efficient encapsulation of theranostic nanoparticles in cell-derived exosomes: leveraging the exosomal biogenesis pathway to obtain hollow gold nanoparticle-hybrids. *Nanoscale*, vol 11, sup 40, 2019, pp 18825–18836 (ISSN: 2040-3372).
- [32] Wang Y, Zhang Y, Cai G, Li Q. Exosomes as Actively Targeted Nanocarriers for Cancer Therapy. *International Journal of Nanomedicine*, vol 15, 2020 (ISSN: 1176-9114).
- [33] Mathieu M, Martin-Jaular L, Lavieu G, Théy C. Specificities of secretion and uptake of exosomes and other extracellular vesicles for cell-to-cell communication. *Nature Cell Biology*, vol 21, sup 1, 2019, pp 9–17 (ISSN: 1476-4679).
- [34] Bunggulawa E, Wang W, Yin T, Wang N, Durkan C, Wang Y, Wang G. Recent advancements in the use of exosomes as drug delivery systems. *Journal of Nanobiotechnology*, vol 16, sup 1, 2018, pp 81 (1477-3155).
- [35] Luan X, Sansanaphongpricha K, Myers I, Chen H, Yuan H, Sun D. Engineering exosomes as refined biological nanoplatfroms for drug delivery. *Acta Pharmacologica Sinica*, vol 38, sup 6, 2017, pp 754–763 (ISSN: 1745-7254).
- [36] Villasante A, Marturano-Kruik A, Ambati SR, Liu Z, Godier-Furnemont A, Parsa H, Lee BW, Moore M, Vunjak-Novakovic G. Recapitulating the Size and Cargo of Tumor Exosomes in a Tissue-Engineered Model. *Theranostics*, vol 6, sup 8, 2016, pp 1119–1130 (ISSN:1838-7640).
- [37] Müller B, Lang S, Dominietto M, Rudin M, Schulz G, Deyhle H, Germann M, Pfeiffer F, David C, Weitkamp T. High-resolution tomographic imaging of microvessels. Optical Engineering + Applications Conference, San Diego, California, USA, 2008, pp 70780B.
- [38] Webpage of Elsevier Connect. <https://www.elsevier.com/es-es/connect/medicina/principios-basicos-de-la-funcion-circulatoria-distribucion-y-presiones-sanguineas> (consulted: June 2023).

Augmented reality simulator for laparoscopic surgery on a realistic phantom

V. López-González¹, A. Pose-Díez-de-la-Lastra^{1,2}, J. de León-Luis^{3,4}, J. Pascau^{1,2}

¹ Departamento de Bioingeniería, Universidad Carlos III de Madrid, Spain, vlopezg@pa.uc3m.es, apose@ing.uc3m.es, jpascau@ing.uc3m.es

² Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain

³ Servicio de Obstetricia y Ginecología, Hospital General Universitario Gregorio Marañón, Madrid, Spain, jaleon@med.ucm.es

⁴ Departamento de Salud Pública y Materno-Infantil, Universidad Complutense de Madrid, Madrid, Spain

Abstract

In recent years, emulation-based environments have become increasingly important in medical education. This progress is particularly beneficial for skills acquisition regarding minimally invasive surgery (MIS). In this work, we present a novel augmented reality (AR) simulator based on a 3D printed realistic phantom for laparoscopic surgery training in gynecology. Our approach was designed to facilitate anatomical referencing skills acquisition through detailed visualization and feedback. The system was validated in an evaluation session at Hospital General Universitario Gregorio Marañón (HGUGM). A total of 8 healthcare professionals assessed the laparoscopic simulator prototype. They also fulfilled a questionnaire on aspects such as the quality of the system as training tool in MIS, potential inclusion in students' curricula, realism and intuitiveness and their interest in additional practice. They assessed the simulator with very positive feedback. These results, combined with the low cost of the system, support its use as a self-training tool on human laparoscopic anatomy.

1. Introduction

The acquisition of surgical skills is traditionally based on the 'see one, do one, teach one' approach. In this scheme, the trainee observes an experienced surgeon executing a task, then actively performs the task and finally, takes the instructor role [1]. Although effective in open surgery, this principle does not apply successfully to minimally invasive procedures [2]. These procedures involve long instrument operation, reduced haptic feedback, complex visual scenes, 3D manipulation with 2D visualization, fulcrum effect, intensified physiological tremor and limited degrees of freedom [3]. MIS demands a higher level of dexterity from surgeons and introduces increased cognitive and physical stress. Consequently, it presents a steeper learning curve and greater difficulty compared to open surgery.

When it comes to acquiring surgical skills for laparoscopic surgery execution, different training simulators with varying complexity and realism are available. MIS simulators are often the first training step for novices before performing human laparoscopic surgeries in the operating room (OR). They can be classified as video-box simulators, which include an illuminated box with holes for trocars insertion to acquire general technical skills; organic simulators, where MIS is performed on human cadavers or animals; virtual reality (VR) simulators, that offer an immersive 3D virtual environment;

and augmented reality (AR) laparoscopic simulators. AR simulators benefit from the superimposition of 3D computer-generated elements onto physical objects in the real world. This enables concurrent interactions with virtual and real information, providing convincing feedback during the entire simulation [4, 5]. The realistic environment provided by AR simulators differentiates them from VR alternatives [6, 7]. In addition, their interactivity distinguishes them from video-box trainers. Moreover, AR laparoscopic simulators can provide real-time information about the tasks to be practiced, as well as feedback during and after their execution. Therefore, they offer an objective evaluation of trainee performance without requiring expert supervision.

In this work, we present an AR laparoscopic simulator prototype that utilizes a real-size 3D printed adult female phantom integrated with an AR application. The AR application displays anatomical information perfectly overlaid onto the phantom, offering the trainee an intricate visual representation of both the physical and virtual organs, blood vessels, and the spaces between visceral structures. In addition, a laparoscopic tool inserted through the phantom holes is tracked during the simulation. This tool, manually controlled by the user, enables interacting with virtual models during phantom exploration. When reached, virtual organs are illuminated and their names appear on the screen. Overall, the AR laparoscopic simulator presented focuses on the acquisition of surgical skills with anatomical referencing during laparoscopic surgery. The laparoscopic simulator was evaluated with clinical users who provided feedback on a user-experience questionnaire.

2. Materials

2.1. 3D printed phantom

The head of gynecology and obstetrics service at Hospital General Universitario Gregorio Marañón (HGUGM), in Madrid, designed and 3D printed the real-size adult female phantom employed in our simulator (Figure 1). Its dimensions are 470 mm in height, 370 mm in width and 290 mm in depth. It encompasses thirty different anatomical structures from the base of the neck to the upper region of the femur bone. It includes organs, blood vessels and gaps between viscera. Additionally, the anterior wall contains four different holes for laparoscopic tool insertion. Apart from the

physical mannequin, the thirty anatomical structures were available in stereolithography format (STL).

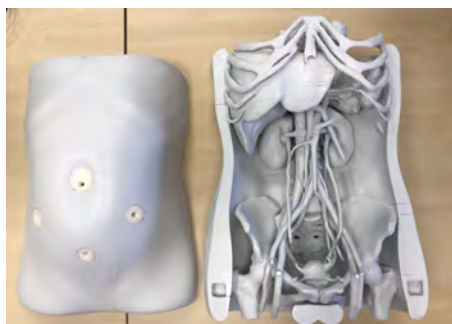


Figure 1. 3D printed adult female phantom.

2.2. AR markers

The AR laparoscopic simulator prototype tracks two AR markers. The first one contains an identifiable 2D image (image-based AR marker) and was printed on paper (Figure 2a). The second AR marker, presented in [8], contains specific patterns easily recognizable by the AR application (pattern-based AR marker). It was 3D printed in polylactic acid (PLA) using fused deposition modelling technique with a double extruder *Ultimaker 3 Extended* 3D printer (Figure 2b).



(a) AR image-based marker



(b) AR pattern-based marker

Figure 2. AR markers employed in the AR laparoscopic simulator prototype.

2.3. Endoscopic and web cameras

The AR system includes two cameras. The first one is a web cam, connected to the computer via USB and its main function is pointing at the AR markers. Once the AR markers are within its field of view (FOV), the AR software detects them and their position is tracked in real time. Vuforia SDK is an AR platform that enables the recognition of specific features from AR markers. Therefore, identifying these features triggers the projection of virtual models in the correct position with respect to the AR markers.

The second camera is an endoscopic camera inserted through the mannequin's holes to visualize the interior of the 3D printed adult female phantom. It is a cylinder-shaped camera of 45 mm in length, 8 mm in lens diameter and a view angle of 67°. It is also connected to the computer via USB and includes LED illumination with different levels for a proper visualization of the inner anatomical structures.

3. Methods

3.1. Design and 3D printing of AR markers guides

Two different structures have been designed for the correct attachment of AR markers to the 3D printed human phantom and endoscopic camera. For the proper design of both guides, Meshmixer and Autodesk Tinkercad softwares were utilized. They enable the creation of personalized sketches according to the surfaces and measurements of the 3D printed phantom and the endoscopic camera.

After completing the design step, both 3D guides were printed. The first one holds the AR pattern-based marker (Figure 3a) and it attaches over the anterior wall and heart surfaces of the 3D phantom. The second one is connected to the endoscopic camera and supports the AR image-based marker (Figure 3b). Both guides were printed in PLA using FDM with a *Crealitty Ender-3* 3D printer.



(a) Structure holding the AR image-based marker



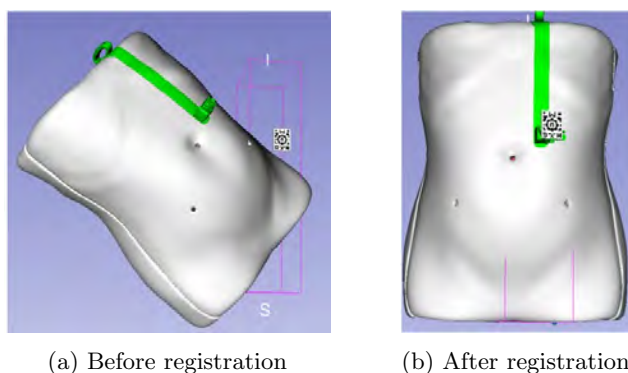
(b) Structure holding the AR pattern-based marker

Figure 3. 3D printed structures holding the AR markers.

3.2. Models registration

Registration was calculated with 3D Slicer software. The spatial transformation that best aligns two objects was computed. Firstly, the relative position between the phantom and the AR image-based marker was calculated. In order to do it, a series of translations and rotations were performed to the phantom with respect to the AR marker (Figure 4). Secondly, the relative position between the endoscopic camera and the AR pattern-based marker was determined following the same process.

As a consequence, when AR system software detects and recognizes each AR maker, the virtual models are displayed in the predefined positions. This way, an automatic registration step is accomplished between real and virtual worlds.



(a) Before registration (b) After registration

Figure 4. AR marker registration

3.3. Augmented reality software

The AR application was developed using Unity3D platform (version 2021.3.1f1), a software tool that enables the creation of interactive 2D, 3D, AR and VR solutions. The AR application is responsible for different functionalities. Firstly, it provides a detailed visualization of the phantom thanks to the endoscopic camera inserted through the holes of the phantom. Secondly, it displays virtual anatomical structures and their analogous in the 3D printed phantom in adjacent windows. This can be observed in Figure 5; in the right side of the screen, an inner visualization of the 3D printed phantom is provided, while the left side of the screen provides exactly the same viewpoint of the virtual phantom. Endoscopic camera tracking is the third functionality deployed in the AR software. If the endoscopic camera moves, the screen views move, and if the endoscopic camera rotates, the physical and virtual perspectives rotate as well. The fourth functionality of the AR software allows the user to interact with the virtual phantom. Once the AR image-based marker is recognized, the software shows a virtual sphere located at 10 cm from the tip end of the endoscopic camera. While manually moving the endoscopic camera, the virtual sphere interacts with the different organs. Once touched by the sphere, the organ changes to a realistic color and its name appears on the screen. Additionally, user interface (UI) buttons allow modifying the sphere distance from the tip end of the camera by 0.5cm/click in the Y direction to reach the deepest anatomical structures of the phantom.



Figure 5. Laparoscopy simulator: hardware and software.

3.4. Validation

Our laparoscopic simulator was validated in a clinical session at HGUGM. Before this validation session, a workshop was organized to receive feedback about the simulator and improve the initial version of the training software and hardware.

The evaluation session at HGUGM was organized by the head of gynecology and obstetrics service and it was coordinated by the *Image-Guided Therapy* group from UC3M. This assembly was aimed at medical students, midwife and medical residents and clinical staff. Eight professionals participated in the session: one medical student, five medical residents and two physicians.

The validation protocol was composed of three stages. The first phase of the session lasted two minutes and was dedicated to user familiarization with laparoscopic simulator hardware and software.

The second part was the key-point of the validation protocol. The users were requested to identify and touch with the virtual sphere ten anatomical structures in ascending order of difficulty according to the criteria of the head of gynecology and obstetrics service from HGUGM. The easiest structures (difficulty 1) were the right kidney and the liver. The right ureter and the sacrum were evaluated as difficulty 2. Difficulty 3 was assigned to the portal venous system and the uterus and vagina. The left iliac vein and the aorta together with the left iliac artery had difficulty 4. The most difficult structures were the right gonadal vein and the arterial celiac trunk. The maximum time allowed to identify and touch each structure was one minute. After this time slot, the anatomical structure was labelled as ‘found’ or ‘not found’ and the next anatomical structure was requested.

To conclude the assessment, the apprentices fulfilled a user-experience questionnaire. It started with general yes/no questions about their previous experience with AR and laparoscopic simulators. Next, they answered several statements in a range from 1 to 5 following *Likert scale* [9] with ‘1’ corresponding to ‘strongly disagree’ and ‘5’ to ‘strongly agree’. These affirmations are related to different competences of the laparoscopic simulator prototype: good training tool regarding MIS, inclusion in students curricula, realistic training environment, additional practice opportunities and an intuitive tool to handle with.

4. Results and discussion

The results obtained in the validation session at HGUGM are summarized in the following figures.

Figure 6 represents the performance of clinical users during phantom exploration with AR system software.

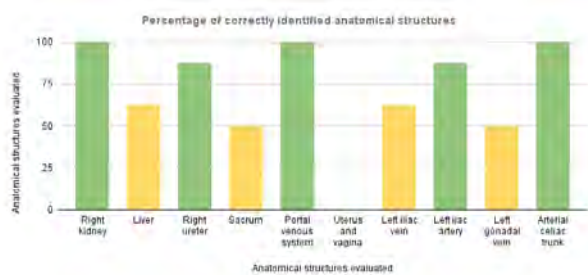


Figure 6. Percentage of anatomical structures correctly identified by users at HGUGM. Accurate recognition of organs ranges from 50% - 75% in yellow color and from 75% - 100% in green color.

The difficulty scale and the percentage of correctly identified anatomical structures are not linearly correlated (Table 1 and Figure 6). One potential reason might be the feedback provided by the laparoscopy simulator during the evaluation step. When the user touched an organ, its name and its real color were provided on the screen. Therefore, the participants had knowledge on the different organs encountered, regardless if they were the requested or not.

As it can be observed in Figure 6, uterus and vagina were not accessible to the participants. The underlying explanation is the severe endoscopic camera tilt to reach these organs. During camera tilting, the AR image-based marker got out the FOV of the web cam limiting the tracking system functionality, preventing users from completing this particular task.

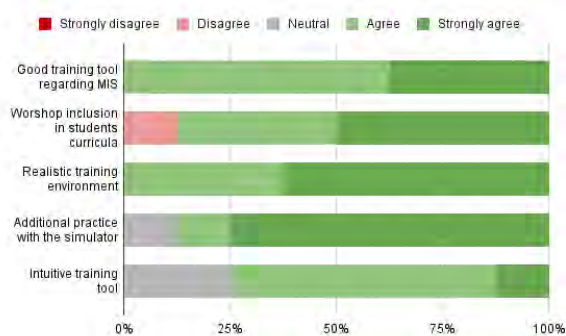


Figure 7. Questionnaire results regarding statements about laparoscopy simulator competences during validation session 2.

Figure 7 displays the feedback provided by the clinicians concerning laparoscopic simulator competences following Likert scale. No ‘strongly disagree’ punctuation was recorded in any of the answers.

Generally, positive feedback was received from the users during the validation session. This is endorsed by the results obtained in the questionnaire answers shown in Figure 7. 87.5% of the healthcare professionals agreed that they would like to have additional training sessions with the laparoscopic simulator prototype developed in this paper. Moreover, 100% of healthcare staff stated

that it is good training tool regarding MIS and a realistic environment to train with.

5. Conclusions

The laparoscopic simulator described in this paper is the first step towards self-training on human laparoscopic anatomy. In addition, the low-cost system described above contributes to home learning and skills acquisition in a safe atmosphere.

The AR simulator prototype has been validated by 8 users with clinical backgrounds at HGUGM. Their positive reviews endorse the AR laparoscopy simulator as a self-learning strategy to acquire laparoscopic skills in anatomical referencing.

Acknowledgments

Research supported by projects PI122/00601 and AC20/00102 (Ministerio de Ciencia, Innovación y Universidades, Instituto de Salud Carlos III, Asociación Española Contra el Cáncer and European Regional Development Fund “Una manera de hacer Europa”), project PerPlanRT (under the frame of ERA PerMed), TED2021-129392B-I00 and TED2021-132200B-I00 (MCIN/AEI/10.13039/501100011033 and European Union “NextGenerationEU”/PRTR) and EU Horizon 2020 research and innovation programme Conex plus UC3M (grant agreement 801538).

References

- [1] J. Vozenilek. See one, do one, teach one: Advanced technology in medical education. *Academic Emergency Medicine*, 11(11):1149–1154, November 2004.
- [2] WS. Melvin *et al.* Laparoscopic skills enhancement. *The American Journal of Surgery*, 172(4):377–379, October 1996.
- [3] FH. Halvorsen *et al.* Simulators in surgery. *Minimally Invasive Therapy & Allied Technologies*, 14(4-5):214–223, January 2005.
- [4] R. Moreta-Martinez *et al.* Combining augmented reality and 3d printing to improve surgical workflows in orthopedic oncology: Smartphone application and clinical evaluation. *Sensors*, 21(4):1370, February 2021.
- [5] JH. Shuhaiber. Augmented reality in surgery. *Archives of Surgery*, 139(2):170, February 2004.
- [6] E. Yiannakopoulou *et al.* Virtual reality simulators and training in laparoscopic surgery. *International Journal of Surgery*, 13:60–64, January 2015.
- [7] SMBI. Botden *et al.* ProMIS augmented reality training of laparoscopic procedures face validity. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, 3(2):97–102, 2008.
- [8] R. Moreta-Martinez *et al.* Augmented reality in computer-assisted interventions based on patient-specific 3d printed reference. *Healthcare Technology Letters*, 5(5):162–166, September 2018.
- [9] AT. Jebb *et al.* A review of key likert scale development advances: 1995–2019. *Frontiers in Psychology*, 12, May 2021.

Sesión Póster II

Jueves 23 10:00 - 11:00

Claustro

Análisis del Volumen de Ablación en Diagnóstico Médico mediante Espectroscopía de Plasma Inducido

R. Sosa-Santos^{1,2}, J.L. Arce-Diego¹, F. Fanjul-Vélez¹

¹ Grupo de Ingeniería Biomédica, Departamento TEISA, Universidad de Cantabria, Av. de los Castros 46, 39005 Santander (Cantabria), España

arcedj@unican.es, fanjul@unican.es

² Centro de Investigaciones Ópticas y Energía, Universidad Privada Boliviana, Colcapirhua 3967, Cochabamba, Bolivia

Resumen

El análisis de la generación de plasma inducido electromagnéticamente y el procesamiento de inteligencia artificial constituyen una aproximación de interés en el diagnóstico médico. De acuerdo con su alta sensibilidad y capacidad de análisis espectral, la espectroscopía de emisión atómica por plasma surge como respuesta de interés diagnóstico sobre tejidos biológicos. En particular, la espectroscopía de ruptura inducida por láser (LIBS) constituye una herramienta empleada en la caracterización de tejidos biológicos. Esta técnica presenta como efecto la ablación de una parte del tejido bajo análisis, lo que en ocasiones podría generar efectos colaterales no deseados in vivo. En este trabajo se analiza el volumen de ablación inducido electromagnéticamente mediante un modelo computacional basado en el método de Monte Carlo (MC) para estudiar la propagación de la radiación en materiales biológicos. Con esta información es posible estudiar las interacciones radiación-tejido en la ingeniería biomédica. En particular, se utiliza una implementación de un algoritmo de Monte Carlo basado en mallas (MMC) como método para mejorar la precisión en el modelado de tejidos no homogéneos. Se cuantifica el volumen y extensión del tejido ablacionado bajo diferentes parámetros del sistema empleado, fundamentalmente la energía, tiempo de pulso y tiempo de exposición. Los resultados pueden contribuir a la limitación de efectos no deseados en la planificación diagnóstica.

1. Introducción

La espectroscopía de descomposición por ablación láser (LIBS) es una técnica analítica que ha probado ser muy útil para la caracterización de diversos materiales, enfatizando muestras biológicas. El procedimiento se basa en la creación de un plasma inducido por láser en la muestra que emite radiación electromagnética de diferentes longitudes de onda.

LIBS se utiliza para varios tejidos biológicos como tejido blando [1], tejido rígido [2], tejido tumoral en tejidos blandos [3], y bacterias [4], sólo por mencionar algunas [5]. LIBS utiliza luz láser pulsada que se enfoca sobre la superficie de la muestra para generar una nube de plasma que contiene las especies altamente excitadas de la muestra, según un montaje similar al de la Figura 1. Existe un valor umbral de densidad de energía para producir plasma. El nivel umbral vendrá determinado por el coeficiente de absorción de la superficie de la muestra y la longitud de onda del láser, que variará mucho en función de la fase del experimento [6].

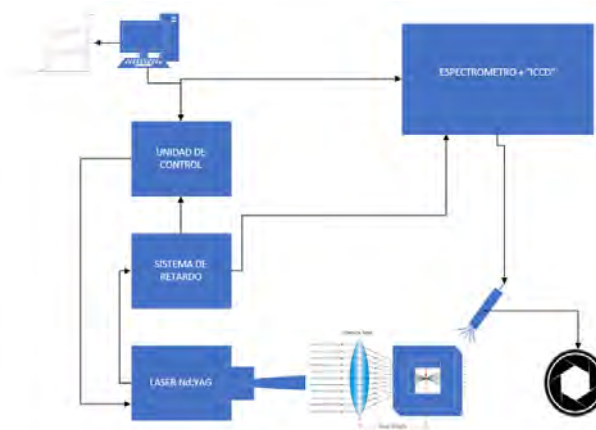


Fig. 1. Representación de un montaje de experimentación de LIBS.

LIBS se basa en un proceso que, aunque los principios que se puede encontrar en la interacción de láser-materia son actualmente complejos y aun no se comprenden del todo, se trata de procesos físicos que tienen lugar durante la formación y expansión del plasma inducido por láser, teniendo en cuenta pulsos de láser cortos y múltiples para excitar dicho plasma [7]

1.1. Modelo de Monte Carlo orientado a tejidos biológicos

El método Monte Carlo (MC) se utiliza con frecuencia en aplicaciones biofotónicas para modelar la migración de fotones en el interior de los tejidos humanos [8]; en otras palabras, el modelo Monte Carlo del transporte de fotones en tejido biológico es un método de simulación numérica utilizado para modelar el comportamiento de los fotones a medida que viajan a través del tejido biológico. Este método se basa en el algoritmo de Monte Carlo, que trae consigo simular una gran cantidad de eventos aleatorios para estimar el comportamiento de un sistema. Por lo tanto, se modela la trayectoria de un fotón como una caminata aleatoria persistente, con la dirección de cada paso dependiendo del paso anterior [9].

El método de Monte Carlo es particularmente útil para modelar el transporte de fotones en tejido biológico porque permite tener en cuenta geometrías y propiedades ópticas complejas. Por ejemplo, se puede utilizar para modelar el esparcimiento y absorción de luz por diferentes tipos de estructuras de tejido, como vasos sanguíneos o membranas celulares. Esta información puede utilizarse para

desarrollar nuevas técnicas de imagen o para comprender mejor cómo la luz interactúa con el tejido biológico.

Una aplicación importante de la modelización de Monte Carlo del transporte de fotones en tejido biológico es en la investigación en óptica biomédica. La óptica biomédica implica el uso de tecnologías basadas en luz para estudiar sistemas biológicos a nivel microscópico. Al comprender cómo la luz interactúa con el tejido biológico, los investigadores pueden desarrollar nuevas técnicas de imagen que se pueden utilizar para el diagnóstico y el tratamiento [10].

La modelización de Monte Carlo se ha utilizado ampliamente en la investigación en óptica biomédica para estudiar varios aspectos del transporte de fotones en tejido biológico. Por ejemplo, se ha utilizado para estudiar cómo la luz se dispersa en diferentes tipos de células y tejidos, lo que puede ayudar a los investigadores a comprender mejor cómo se desarrollan y se propagan enfermedades como el cáncer [11].

A pesar de su utilidad, la modelización de Monte Carlo tiene algunas limitaciones. Una limitación es que requiere importantes recursos informáticos debido al gran número de simulaciones requeridas. Además, puede que no modele con precisión todos los aspectos del transporte de fotones en tejido biológico, como los efectos del movimiento del tejido o los cambios en la estructura del tejido con el tiempo.

1.2. Algoritmo de Monte Carlo basado en vóxeles (mmc)

MMC es la base para las nuevas técnicas de modelado biofotónico en tejidos no homogéneos tridimensionales. Prácticamente MMC es la norma del sector para simular la propagación de la luz en un medio turbio. Se tienen criterios para determinar el tamaño del vóxel y el número de fotones en simulaciones tridimensionales con una precisión y un tiempo de cálculo aceptable para reducir el coste computacional. Además el tamaño del vóxel puede elegirse en función de la geometría del tejido y de las características ópticas, teniendo una relación determinada por el entorno de simulación, siendo que el número de fotones debe ser al menos cinco veces el número total de vóxeles para tener un método de división de haces de fotones para una malla local [12].

MMC puede utilizar mallas tetraédricas para obtener una mayor precisión anatómica siendo indiferente a la unidad de procesamiento gráfico comparando con simulaciones procesadas con un único subproceso de la CPU. La implementación de OpenCL permite ejecutar simulaciones de fotones MC de alto rendimiento en una amplia gama de CPU y GPU, mostrando una excelente escalabilidad para adaptarse a arquitecturas de GPU cada vez más potentes. Describimos los conocimientos que hemos adquirido sobre la utilización de la memoria de la GPU y las diferencias entre fabricantes. Además, proporcionamos pruebas de velocidad que abarcan desde dominios homogéneos sencillos hasta modelos muy sofisticados del mundo real. [13]

1.3. Colon

El colon es la parte principal que compone el intestino grueso. Es uno de los últimos tramos del sistema digestivo y tiene una dimensión de unos 1.5 metros de longitud y 6.5 centímetros de diámetro. Entre sus funciones principales están la extracción de agua, el almacenamiento de residuos y la absorción de vitaminas antes de ser eliminadas del cuerpo. El colon cuenta con cuatro capas, que son las que definiremos para la formación de la malla, mucosa, submucosa, muscular y serosa [5].

2. Materiales y Métodos

Buscamos definir los parámetros con los que lograr la ablación que genere el plasma y obtener el volumen afectado por dicha interacción; es por tal motivo que tenemos que definir lo siguiente.

2.1. Definición de los tejidos multicapa

El tejido que emplearemos para esta simulación será de colon, el cual consta de cuatro capas. Los introduciremos brevemente y definiremos el tamaño de la muestra a tratar, así como su espesor entre capas y, por supuesto, los valores de sus propiedades ópticas. Cada una de estas capas cuenta con un espesor determinado que será necesario conocer para su correcta simulación, que aparece en la Tabla 1.

	Mucosa	Submucosa	Muscular	Serosa
Espesor (mm)	1.658	0.626	1.8	0.076

Tabla 1. Espesor en mm de las capas que componen el colon

A partir de estos espesores simularemos la malla con sus cuatro capas, que será un cubo de 4.16x4.16x4.16 mm.

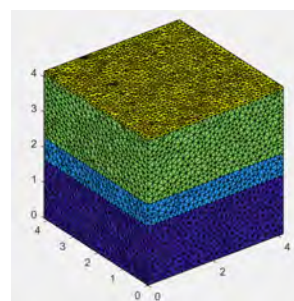


Fig. 2. Representación de las distintas capas del colon

En la Figura 2 se observan las cuatro capas diferenciadas con distintos colores, también del uso de los tetraedros para la generación de la malla, como ya comentamos. La capa azul oscuro corresponde con la mucosa, la capa azul claro con la submucosa, la verde con la muscular y la amarilla con la serosa. Estas son las dimensiones, otro dato necesario del tejido son sus propiedades ópticas, las cuales varían su valor en función de la longitud de onda del láser. Las propiedades necesarias para esta simulación son los coeficientes de absorción, de esparcimiento, la anisotropía y los índices de refracción de cada capa.

2.2. Definición de la fuente láser

En la presente investigación emplearemos un láser con una longitud de onda de $\lambda=1064$ nm, tipo Nd:YAG, con el que realizaremos distintos análisis. El láser además tendrá una

frecuencia de repetición fija en 15 Hz. La zona de impacto del haz de fotones sobre el tejido se situará en el punto medio de sus superficies superiores. Posteriormente se realizó la variación de distintos parámetros del láser para observar cómo lograr una mejor ablación. Se simularon los tres casos primero para un tiempo de pulso de 150 ns y luego para un tiempo de pulso de 10 μs, como refleja la Tabla 2.

	Energía de pulso	Tiempo de exposición	Tiempo de pulso
Caso 1	50 mJ	180 s	150 ns
Caso 2	85 mJ	320 s	
Caso 3	120 mJ	720 s	10 μs

Tabla 2. Casos que analizar en las simulaciones

Para poder observar si se ha logrado una ablación se debe establecer un umbral a partir del cual se logra dicho fenómeno, cuyo valor será de $5 \frac{J}{cm^2}$.

3. Aplicación y resultados

De esta forma observamos, y realizaremos una serie de análisis variando energías de pulso, tiempos de exposición y tiempos de pulso, para observar cómo influye esto en el proceso de la ablación (Tabla 2). En los casos donde se supere el umbral antes mencionado, calcularemos además el volumen de la muestra que ha experimentado esta ablación. A la hora de realizar las simulaciones en el colon, estableceremos un tiempo de simulación de $3 \cdot 10^{-10}$ s y un tiempo de paso de $5 \cdot 10^{-12}$ s.

El transcurso aplicado consiste en obtener la energía acumulada en el tejido para cierta duración de pulso y, a través de una representación de la sección de dicho tejido, comparar los valores obtenidos con el umbral. Una vez obtenido, se realiza el cálculo según la energía de pulso deseada, se ajusta la frecuencia de repetición fijada en 15 Hz, y con el tiempo de exposición. Para fines del artículo se presentarán 2 casos.

3.1. Caso 1 ($t_{pulso} = 150$ ns)

Energía de pulso	Tiempo de exposición	Tiempo de pulso
50 mJ	180 s	150 ns

Tabla 3. Datos de los parámetros para el Caso 1 ($t_{pulso}=150$ ns)

Con los datos de la Tabla 3, los resultados obtenidos son los siguientes:

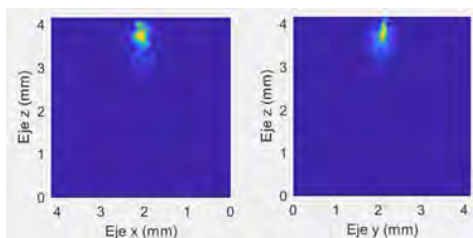


Fig. 3. Energía acumulada por centímetro cuadrado (J/cm^2) en secciones del colon para el Caso 1 ($t_{pulso}=150$ ns)

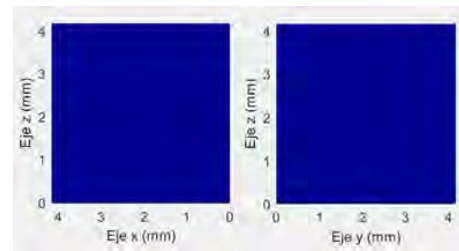


Fig.4. Zonas que experimentan ablación (rojo) y zonas que no (azul), en secciones del colon para el Caso 1 ($t_{pulso}=150$ ns)

En la figura 3. se observa la energía por centímetro cuadrado acumulada en las secciones transversales $y=2.08$ mm y $x=2.08$ mm del tejido. Además, como se puede apreciar claramente en la figura 4, para este caso no ocurre la ablación en ningún punto, ya que el valor máximo obtenido ha sido $3.89 J/cm^2$, menor que el umbral de $5 J/cm^2$.

3.2. Caso 2 ($t_{pulso} = 10$ μs)

Energía de pulso	Tiempo de exposición	Tiempo de pulso
120 mJ	720 s	10 μs

Tabla 4. Datos de los parámetros para el Caso 2 ($t_{pulso}=10\mu$ s)

Con los datos de la Tabla 4, los resultados obtenidos son los siguientes:

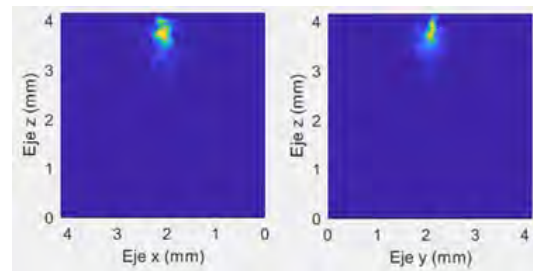


Fig. 5. Energía acumulada por centímetro cuadrado (J/cm^2) en secciones del colon para el Caso 3 ($t_{pulso}=10\mu$ s)

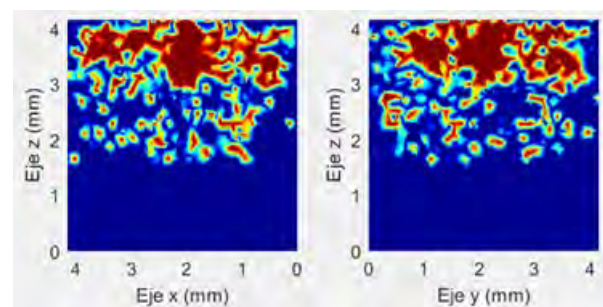


Fig. 6. Zonas que experimentan ablación (rojo) y zonas que no (azul), en secciones del colon para el Caso 3 ($t_{pulso}=10\mu$ s)

Para este último caso de estudio del colon, la ablación ha continuado su expansión por el tejido. La ablación ocupa una mayor extensión en las capas del colon, sin llegar a la perforación. El valor máximo de profundidad alcanzada desde superficie es de 2.531 mm. Para este caso, el volumen total de ablación tiene un valor de $0.4179 mm^3$.

Entre las limitaciones se puede mencionar las suposiciones de homogeneidad en las capas de tejido; en adición, el uso

de un umbral específico de energía para definir la ablación podría variar en función de las características del tejido. Y por último, aunque, el estudio se centró en las propiedades ópticas del tejido y los parámetros del láser, no se tuvo en cuenta factores biológicos adicionales, como la irrigación sanguínea, la temperatura del tejido y la variabilidad anatómica entre pacientes.

4. Conclusiones

En esta investigación sobre la ablación de tejidos multicapa, nos enfocamos en el colon humano, estimamos cómo varían los efectos de la ablación en función de distintos parámetros láser y los rasgos del tejido. Comprendemos que la ablación ocurre cuando se supera un umbral específico de energía, y las capas afectadas varían en función de las propiedades de la fuente. Por tanto, observamos que el tiempo de pulso láser tiene un impacto significativo en la extensión de la ablación, dando lugar a pulsos más largos, provocando así una mayor expansión de la ablación. Son estos hallazgos cruciales para comprender y observar la ablación en tejidos multicapa, con posibles aplicaciones en el desarrollo de procedimientos médicos que utilizan láseres para el diagnóstico y tratamiento de enfermedades.

Agradecimientos

Este Trabajo ha sido parcialmente financiado por el proyecto del Plan Nacional de I+D+i “Cribado diagnóstico de microorganismos mediante microscopia avanzada e inteligencia artificial en patologías humana” (PID2021-127691OB-I00), del Ministerio de Ciencia e Innovación, cofinanciado con fondos FEDER, y por la Fundación Carolina a través de la beca concedida a R. Sosa-Santos.

Referencias

[1] P. Winnand, K. O. Boernsen, G. Bodurov, M. Lammert, F. Hölzle, y A. Modabber, «Evaluation of electrolyte element composition in human tissue by laser-induced breakdown spectroscopy (LIBS)», *Sci. Rep.*, vol. 12, n.º 1, p. 16391, 2022, doi: 10.1038/s41598-022-20825-0.

[2] M. A. Kasem, J. J. Gonzalez, R. E. Russo, y M. A. Harith, «LIBS analysis of artificial calcified tissues matrices», *Talanta*, vol. 108, pp. 53-58, abr. 2013, doi: 10.1016/j.talanta.2013.02.062.

[3] Y. Chu *et al.*, «Discrimination of nasopharyngeal carcinoma serum using laser-induced breakdown spectroscopy combined with an extreme learning machine and random forest method», *J. Anal. At. Spectrom.*, vol. 33, n.º 12, pp. 2083-2088, nov. 2018, doi: 10.1039/C8JA00263K.

[4] S. J. Rehse, «A review of the use of laser-induced breakdown spectroscopy for bacterial classification, quantification, and identification», *Spectrochim. Acta Part B At. Spectrosc.*, vol. 154, pp. 50-69, abr. 2019, doi: 10.1016/j.sab.2019.02.005.

[5] F. Fanjul-Vélez, S. Pampín-Suárez, y J. L. Arce-Diego, «Application of Classification Algorithms to Diffuse Reflectance Spectroscopy Measurements for Ex Vivo Characterization of Biological Tissues», *Entropy*, vol. 22, n.º 7, Art. n.º 7, jul. 2020, doi: 10.3390/e22070736.

[6] Q. Wang, W. Xiangli, G. Teng, X. Cui, y K. Wei, «A brief review of laser-induced breakdown spectroscopy for human and animal soft tissues: pathological diagnosis and physiological detection», *Appl. Spectrosc. Rev.*, vol. 56, n.º 3, pp. 221-241, mar. 2021, doi: 10.1080/05704928.2020.1791151.

[7] S. J. Rehse y A. W. Miziolek, «Recent advances in the use of laser-induced breakdown spectroscopy (LIBS) as a rapid point-of-care pathogen diagnostic», en *Sensing Technologies for Global Health, Military Medicine, Disaster Response, and Environmental Monitoring II; and Biometric Technology for*

Human Identification IX, SPIE, may 2012, p. 837106. doi: 10.1117/12.919256.

[8] M. Burmen, F. Pernus, y P. Naglic, «MCDataset: a public reference dataset of Monte Carlo simulated quantities for multilayered and voxelated tissues computed by massively parallel PyXOpto Python package», *J. Biomed. Opt.*, vol. 27, n.º 8, p. 083012, ago. 2022, doi: 10.1117/1.JBO.27.8.083012.

[9] I. Salas-García, F. Fanjul-Vélez, y J. L. Arce-Diego, «Superficial radially resolved fluorescence and 3D photochemical time-dependent model for photodynamic therapy», *Opt. Lett.*, vol. 39, n.º 7, pp. 1845-1848, abr. 2014, doi: 10.1364/OL.39.001845.

[10] V. V. Tuchin, S. R. U. M.d, y I. V. Yaroslavsky, «Tissue optics, light distribution, and spectroscopy», *Opt. Eng.*, vol. 33, n.º 10, pp. 3178-3188, oct. 1994, doi: 10.1117/12.178900.

[11] A. V. Bykov, M. Y. Kirillin, y A. V. Priezzhev, *Monte Carlo Simulation of Light Propagation in Human Tissues and Noninvasive Glucose Sensing*. en Series in Medical Physics and Biomedical Engineering. Boca Raton: CRC Press-Taylor & Francis Group, 2009, pp. 65-95. Accedido: 28 de agosto de 2023. [En línea]. Disponible en: <https://www.webofscience.com/wos/woscc/full-record/WOS:000266998200004>

[12] D. Li, B. Chen, W. Y. Ran, G. X. Wang, y W. J. Wu, «Selection of voxel size and photon number in voxel-based Monte Carlo method: criteria and applications», *J. Biomed. Opt.*, vol. 20, n.º 9, p. 095014, sep. 2015, doi: 10.1117/1.JBO.20.9.095014.

[13] N. Ren, J. Liang, X. Qu, J. Li, B. Lu, y J. Tian, «GPU-based Monte Carlo simulation for light propagation in complex heterogeneous tissues», *Opt. Express*, vol. 18, n.º 7, p. 6811, mar. 2010, doi: 10.1364/OE.18.006811.

Calibración de simulaciones de fibrilación auricular para reproducir biomarcadores clínicos

M. Termenón-Rivas, J. Barrios-Álvarez de Arcaya, G. S. Romitti, P. Romero, D. Serra, I. García-Fernandez, M. Lozano, R. Sebastian, A. Liberos, M. Rodrigo

CoMMLab, Departament d'Informàtica, Departament d'Enginyeria Electrònica, Universitat de València, València, España
{maría.termenon, alejandro.liberos, miguel.rodrigo}@uv.es

Resumen

Los Gemelos Digitales capaces de reproducir patrones eléctricos específicos del paciente durante fibrilación auricular (FA) podrían ayudar a identificar la terapia óptima. La calibración de las simulaciones biofísicas es necesaria para que estas reproduzcan biomarcadores clínicos observados en pacientes, tales como la duración de ciclo (CL) o la velocidad de conducción (CV).

Se construyó una anatomía auricular 3D realista que incluía orientación de fibras y anisotropía en la conducción de todas las regiones auriculares para simular patrones de FA. Se simularon reentradas funcionales bajo diferentes grados de remodelado eléctrico y difusión tisular (desde sano a FA crónica) y, a continuación, se midieron la CV y la CL en diferentes regiones auriculares.

Como cabía esperar, la CV en las regiones auriculares aumenta con el incremento de la difusión eléctrica, de 232,1 a 570,1 mm/s en la aurícula izquierda (AI), mientras que la CL disminuye con el remodelado eléctrico, desde el sustrato sano al sustrato de FA crónica (de 217,4 a 147,9 ms en la AI).

Los factores de remodelado eléctrico y difusión pueden utilizarse de forma independiente para calibrar los modelos de FA a los biomarcadores de CV y CL medidos en cada paciente. Esto permitirá a los Gemelos Digitales reproducir los patrones electrofisiológicos específicos del paciente para ensayar posibles terapias.

1. Introducción

Las arritmias cardíacas son una de las principales causas de mortalidad y morbilidad en los países desarrollados, siendo la fibrilación auricular (FA) la de mayor prevalencia [1]. Aproximadamente, un 1% de la población sufre esta patología [2]. Sin embargo, los mecanismos específicos de la iniciación y el mantenimiento de la FA todavía no se entienden por completo, lo que hace que los tratamientos tanto farmacológicos como quirúrgicos no tengan el éxito requerido. El modelado computacional de la electrofisiología auricular permite comprender mejor los datos experimentales y las patologías, facilitando el diagnóstico y la selección de tratamiento.

Obtener un Gemelo Digital de las aurículas de cada paciente permite reproducir su fenotipo específico de FA y predecir el resultado potencial de cada terapia. Sin embargo, esto requiere adaptar los modelos matemáticos para reproducir la actividad eléctrica del paciente, ya que se ha reportado una gran variabilidad en los biomarcadores clínicos como la duración de ciclo (CL) y la velocidad de

conducción (CV), que pueden variar desde 130 a 300 ms y de 0,37 a 1,33 m/s, respectivamente [2].

El objetivo de este trabajo es estudiar la relación entre los parámetros de simulación y los marcadores clínicos para adaptar modelos computacionales realistas. Esto se llevará a cabo en un modelo 3D detallado que incluye la heterogeneidad del sustrato, la dirección de fibras y el remodelado eléctrico para reproducir fielmente la actividad eléctrica de la FA en sustratos realistas.

2. Materiales y métodos

2.1. Modelo auricular heterogéneo

Se ha utilizado el modelo 3D de aurícula humana mostrado en la Figura 1a, que incluye dirección de fibras y división de la aurícula en diferentes regiones [3]: aurícula derecha (AD), aurícula izquierda (AI), nodo sinoauricular (NSA), crista terminalis (CT), músculos pectíneos (MP), Haz de Bachmann (HB), orejuela derecha (OD), orejuela izquierda (OI), vena cava inferior (VCI), vena cava superior (VCS), istmo cavotricuspidé (ICT), venas pulmonares (VP) y fosa ovalis (FO).

Utilizando como base el trabajo de Krueger [3], la anisotropía se incluyó variando la difusión longitudinal y transversal de cada región como se muestra en la Tabla 1. La propagación tisular se basó en el modelo monodominio y se resolvió con una discretización temporal de 10 μ s en un solver de GPU propio.

Región auricular	Difusión transversal ($\mu\text{m}^2/\text{ms}$)	Difusión longitudinal ($\mu\text{m}^2/\text{ms}$)
AD, AI	0,1200	0,4500
NSA	0,4392	0,4392
CT	0,1200	0,7872
MP	0,0480	1,1160
HB	0,2892	1,1221
OD, OI, VCI, VCS, ICT	0,1200	0,1200
VM, VT, VP, FO	0,0840	0,3049

Tabla 1. Valores de difusión transversal y longitudinal

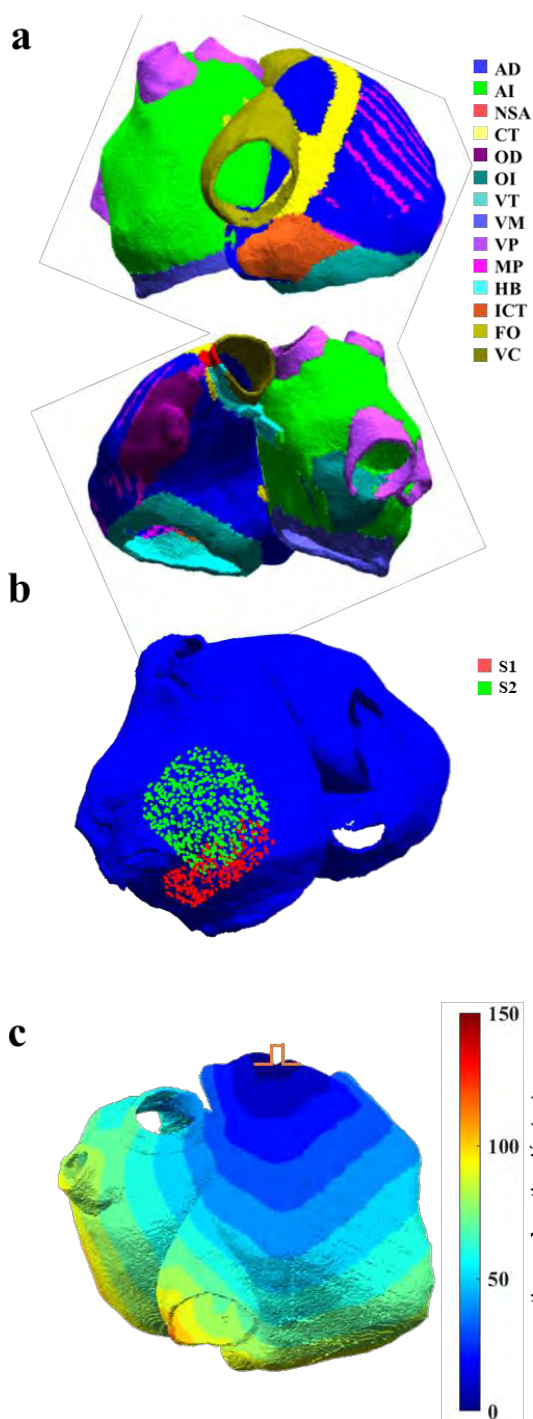


Figura 1. a) Regiones auriculares. b) Zonas de estimulación del protocolo S1-S2. c) Mapa de tiempos de activación en una estimulación periódica a 1Hz.

A partir del modelo basal se incluyeron variaciones en los porcentajes de remodelado y difusión entre el 25 y el 150%, para simular distintas condiciones del sustrato. Un porcentaje de remodelado del 0% y una difusión del 100% representan las condiciones de una aurícula sana. Un aumento del porcentaje de remodelado implica una progresión desde el caso sano hasta fibrilación auricular crónica (FAc), para un remodelado del 100%, y FA de larga duración por encima de este porcentaje. Para crear el modelo de FAc se modificaron, en todas las regiones de la aurícula, las siguientes corrientes a partir del modelo de Koivumäki: expresión de la bomba SERCA (-16%), ratio

PLB a SERCA (+18%), ratio SLN a SERCA (-40%), I_{NCX} máxima (+50%), sensibilidad de RyR a $[Ca^{2+}]_{SR}$ (+100%), conductancia de I_{CaL} (-59%), conductancia de I_{to} (-44%), conductancia de I_{Kur} (-22%) y conductancia de I_{K1} (+100%). Para los modelos con un porcentaje de remodelado entre el 0 y el 100%, los valores de las corrientes modificadas se obtuvieron como una interpolación lineal entre el caso sano y el modelo de FAc, mientras que las corrientes de los modelos de remodelado superior al 100% se calcularon extrapolando linealmente.

2.2. Simulaciones del modelo heterogéneo

Después de estabilizar el modelo siguiendo un protocolo de estimulación periódica a 5 Hz durante 4 s, se generó una actividad reentrante en el septo de la aurícula izquierda mediante un protocolo S1-S2 (fig 1b). Se realizó un conjunto de 36 simulaciones combinando variaciones en el porcentaje de difusión y de remodelado ($\{25, 50, 75, 100, 125, 150\}$ %, para ambos parámetros). Por último, se descartaron los escenarios en los que la reentrada funcional no se auto-sostenía (manteniéndose en 26 casos).

2.3. Medida de la duración de ciclo y la velocidad de conducción

De la superficie de ambas aurículas se seleccionaron siete puntos de los que se midieron los marcadores clínicos: AI, VM, VP, CT, OD y dos en la pared de la AD. Estos puntos, representados en la Figura 2a, se utilizaron para medir la CL y la CV utilizando 2,5 s de cada simulación después de que la reentrada funcional se estabilizase.

Para medir la velocidad de conducción (CV), se seleccionaron 4 puntos a 5 mm de distancia alrededor de cada uno de los 7 puntos de medida (fig 2b). Se ajustó una conducción de onda planar por mínimos cuadrados a cada uno de los instantes de activación, a partir de la cual se extrajeron la dirección de propagación y la CV. Finalmente, la CV se obtuvo como el valor medio de todas las activaciones.

La duración de ciclo (CL) se obtuvo a partir de la curva de potencial transmembrana (PT) en cada uno de estos 7 puntos y promediando en el tiempo entre activaciones consecutivas (fig 2c).

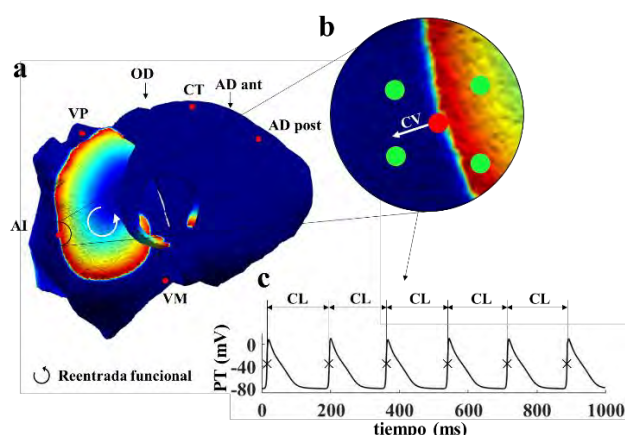


Figura 2. a) Puntos de medida de la CL y la CV. b) Detalle de los 4 puntos utilizados para medir la CV. c) Potencial transmembrana para calcular la CL.

3. Resultados

3.1. Evaluación del modelo en ritmo sinusal

La Figura 1c muestra el mapa de activación local para una simulación de estimulación periódica a 1Hz desde el NSA para condiciones sanas. La aurícula tarda 112 ms en despolarizarse por completo. El efecto de la anisotropía se puede observar como una zona de conducción más rápida correspondiente al CT que provoca un perfil triangular en las isócronas.

3.2. Caracterización de la velocidad de conducción

La Figura 3a muestra la CV obtenida en la zona de medida de la AI para todas las combinaciones de porcentaje de difusión y remodelado en las que la reentrada funcional se mantenía. Se observa que, la variación del remodelado no afecta significativamente a la medida obtenida de velocidad. Sin embargo, como se esperaba, se observó que un aumento de la difusión provoca un incremento de la CV: de alrededor de 400 mm/s para un porcentaje de difusión del 50% a alrededor de 625 mm/s para una difusión del 150%.

3.3. Caracterización de la duración de ciclo

Los resultados obtenidos de duración de ciclo para las simulaciones analizadas se muestran en la Figura 3b. En este caso, se obtuvo que la CL depende de ambos parámetros de simulación, siendo el remodelado el que provoca un mayor cambio en la medida obtenida. Un aumento en la difusión o el remodelado provoca una disminución en la CL. Por ejemplo, en el caso de la AI, un aumento del porcentaje de remodelado del 25% al 150% provoca una disminución de la CL de 217,4 a 147,9 ms, para una difusión del 50%. Sin embargo, un incremento de la difusión del 25% al 125% reduce la CL de 174,0 a 153,4 ms, para un remodelado del 75%.

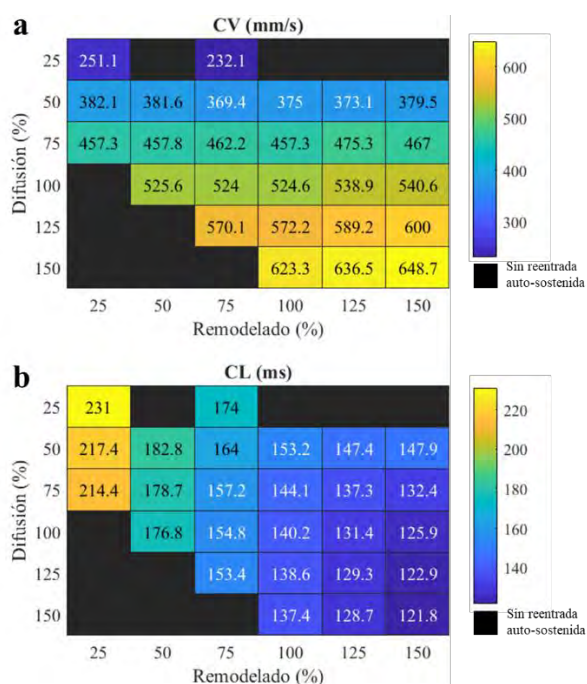


Figura 3. a) CV para distintos porcentajes de difusión y remodelado. b) CL para distintos porcentajes de difusión y remodelado.

3.4. Evaluación de las heterogeneidades del modelo

En la Figura 4 se muestra la CV medida en las 7 regiones anatómicas de las aurículas, para un remodelado del 75%.

En todas las regiones se obtuvo un aumento de la CV con la difusión. Sin embargo, la CV medida varió en función de la zona. Las regiones de conducción más rápida fueron el CT y la pared anterior de la AD, con un aumento del 417,1 a 902,7 mm/s y 368,8 a 937,9 mm/s, respectivamente. Por el contrario, la zona en la que se obtuvo una menor velocidad fue la OD, en la que la CV aumentó de 132,1 a 390,7 mm/s.

La variación de la CL en función del remodelado para los distintos puntos de medida se muestra en la Figura 5a, obtenida con para una difusión del 50%. En los puntos en las VP y VM, la CL varía de la misma forma que en la AI, disminuyendo al aumentar el remodelado. Hay que tener en cuenta que, la ratio de conducción de la AI a la AD varió en función del porcentaje de remodelado, comenzando en, aproximadamente, 5:4 para el 25% (fig 5b), disminuyendo a 1:1 para remodelados del 50% al 75% (fig 5c) y volviendo a aumentar hasta una ratio de 2:1 a partir del 125% (fig 5d), causando una mayor CL en la AD respecto a la AI en los casos de ratio mayor a 1:1. Esto indica que aunque el modelo no incluya variaciones específicas en el modelo iónico de cada región, la diferencia en la conducción provoca patrones de bloqueo de conducción típicos de FA.

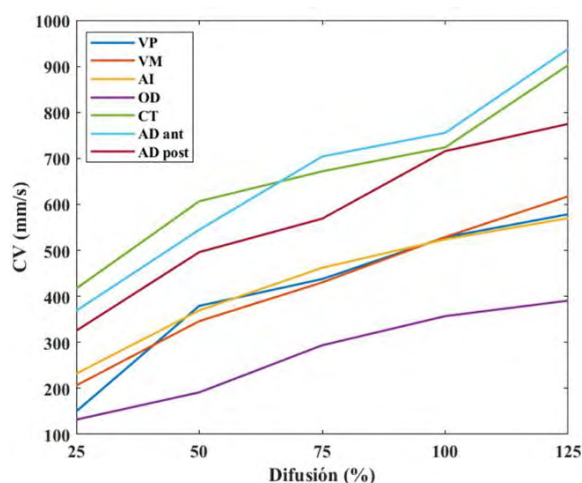


Figura 4. CV en función de la difusión en las distintas regiones auriculares para un remodelado del 75%.

4. Discusiones y conclusión

Este estudio demuestra que las simulaciones detalladas del substrato auricular, incluyendo dirección de fibras, anisotropía de cada región y remodelado eléctrico, pueden reproducir marcadores clínicos como la CV y la CL en los rangos observados en la práctica clínica. Las tablas de calibración obtenidas pudieron relacionar de forma independiente los biomarcadores medibles clínicamente, CV y CL, como función de los parámetros utilizados para llevar a cabo las simulaciones: remodelado eléctrico y difusión. Se ha obtenido que la CV aumenta al incrementar el porcentaje de difusión cubriendo el rango desde 232,1 mm/s a 648,7 mm/s en la AI, y de 100 a 900 mm/s en la

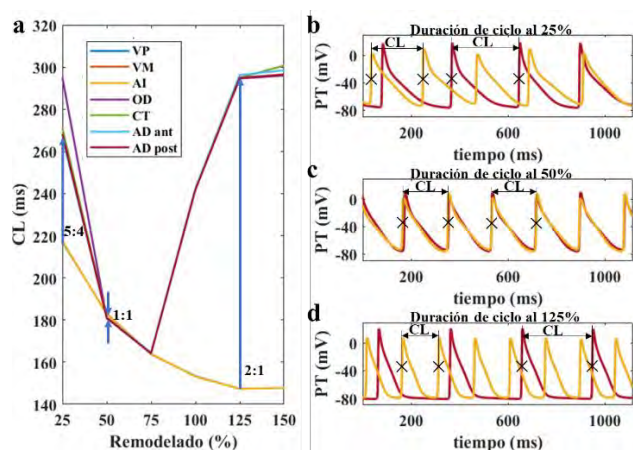


Figura 5. a) CL en función del remodelado para las distintas regiones auriculares para una difusión del 50%. b) Curvas de potencial transmembrana en los puntos de la AI (amarillo) y AD (rojo) para un remodelado del 25% (b), 50% (c) y 125% (d).

aurícula completa, y manteniéndose invariable frente al remodelado. Por el contrario, la CL disminuyó al aumentar el remodelado eléctrico y, en menor medida, al incrementar la difusión, produciéndose una disminución de la CL al variar el remodelado, de 217,4 a 147,9 ms en la AI para una difusión del 50%. En la AD la ratio de conducción varió en función del remodelado obteniéndose CLs incluso mayores, aunque siempre ligados a bloqueos de conducción.

Este estudio se centra en la condición arrítmica en la AI, ya que se conoce la relevancia de esta región en el mantenimiento de la FA. Sin embargo, deben considerarse otros mecanismos de mantenimiento de la FA, como las ondas múltiples o las microrreentradas, y otras regiones dominantes. Además, las simulaciones de este estudio se han realizado utilizando un único modelo auricular. Es necesario ampliar el trabajo a un mayor número de geometrías auriculares para validar y generalizar los resultados obtenidos, así como extraer una relación analítica entre los parámetros del modelo y los marcadores clínicos.

Los resultados obtenidos en esta investigación demuestran que la difusión y el remodelado eléctrico pueden adaptarse en modelos computacionales para reproducir los valores de CL y CV medidos en pacientes. Esto permitirá crear Gemelos Digitales adaptando el porcentaje de remodelado y difusión de las simulaciones realizadas a los rangos correspondientes con los biomarcadores clínicos medidos en el propio paciente. Incluso realizar barridos de simulaciones para reproducir el dinamismo temporal de los marcadores. De esta forma, se podrá reproducir no sólo la anatomía sino también el fenotipo específico de FA de cada paciente y así, identificar sus mecanismos específicos y predecir la eficacia de cada terapia en cada paciente.

Agradecimientos

Este trabajo fue financiado por la subvención de la Generalitat Valenciana AICO/2021/318 (consolidables 2021) y subvención PID2020-114291RB-I00 financiado por MCIN/ 10.13039/501100011033 y por “ERDF A way of making Europe” INVEST/2022/232.

Referencias

- [1] Jalife J, Delmar M, Anumonwo J, Berenfeld O, Kalifa J. Basic cardiac electrophysiology for the clinician. John Wiley & Sons, 2011 (ISBN: 9781405183338).
- [2] Dössel O, Krueger MW, Weber FM, Wilhelms M, Seemann G. Computational modeling of the human atrial anatomy and electrophysiology. *Med Biol Eng Comput*, no. 50, 2012, pp. 773–99. (ISSN: 0140-0118).
- [3] Krueger MW et al. Modeling Atrial Fiber Orientation in Patient-Specific Geometries: A Semi-automatic Rule-Based Approach. *Functional Imaging and Modeling of the Heart (FIMH)*, New York City, 2011, vol. 6666, pp. 223-32 (ISBN: 978-3-642-21028-0).
- [4] Feng J, Yue L, Wang Z, Nattel S. Ionic mechanisms of regional action potential heterogeneity in the canine right atrium. *Circulation research*, vol. 7, no 83, 1998, pp. 541-51 (ISSN: 0009-7330).
- [5] Ehrlich JR., Cha TJ, Zhang L, Chartier D, Melnyk P, Hohnloser SH, Nattel S. Cellular electrophysiology of canine pulmonary vein cardiomyocytes: action potential and ionic current properties. *The Journal of physiology*, vol. 551, no. 3, 2003, pp. 801-13 (ISSN: 0022-3751).
- [6] Aslanidi OV, Colman MA, Stott J, Dobrzynski H, Boyett MR, Holden AV, Zhang H. 3D virtual human atria: A computational platform for studying clinical atrial fibrillation. *Progress in biophysics and molecular biology*, vol. 107, no. 1, 2011, pp. 156-68 (ISSN: 0079-6107).
- [7] Calvo CJ, Deo M, Zlochiver S, Millet J, Berenfeld O. Attraction of Rotors to the Pulmonary Veins in Paroxysmal Atrial Fibrillation: A Modeling Study. *Biophysical Journal*, vol. 106, 2014, pp. 1811-21 (ISSN: 0006-3495).
- [8] Ferrer A, Sebastián R, Sánchez-Quintana D, Rodríguez JF, Godoy EJ, Martínez L, Saiz J. Detailed Anatomical and Electrophysiological Models of Human Atria and Torso for the Simulation of Atrial Activation. *PLoS One*, vol. 10, no 11, 2015 (ISSN: 1932-6203).
- [9] Krueger MW, Dorn A, Keller DU, Holmqvist F, Carlson J, Platonov PG, Rhode KS, Razavi R, Seemann G, Dössel O. In-silico modeling of atrial repolarization in normal and atrial fibrillation remodeled state. *Med Biol Eng Comput*, no. 51, 2013, pp. 1105–19 (ISSN: 0140-0118).
- [10] Sánchez C, Bueno-Orovio A, Pueyo E, Rodríguez, B. Atrial Fibrillation Dynamics and Ionic Block Effects in Six Heterogeneous Human 3D Virtual Atria with Distinct Repolarization Dynamics. *Frontiers in Bioengineering and Biotechnology*, vol. 5, 2017 (ISSN: 2296-4185).
- [11] Koivumäki JT, Seemann G, Maleckar MM, Tavi P. In silico screening of the key cellular remodeling targets in chronic atrial fibrillation. *PLoS Computational Biology*, vol. 10, no. 5, 2014 (ISSN: 1553-734X).
- [12] Garcia-Molla VM, Liberos A, Vidal A, Guillem MS, Millet J, Gonzalez A, Martinez-Zaldivar FJ, Climent AM. Adaptive step ODE algorithms for the 3D simulation of electric heart activity with graphics processing units. *Computers in biology and medicine*, vol. 44, 2014, pp. 15-26 (ISSN: 0010-4825).

Cinemática del movimiento del miembro superior durante una tarea funcional realizada en entorno virtual y entorno físico real en población sana y con lesión medular cervical

L. Blanco Coloma¹, V. Herrera Tirado², R. Perales Gómez³, M. Alcobendas Maestro⁴, S. Ceruelo Abajo⁴, V. Lozano Berrio¹, J. Albusac Jiménez², C. González Morcillo², A. de los Reyes Guzmán^{1,5,*}

¹ Unidad de Biomecánica y Ayudas Técnicas, Hospital Nacional de Paraplégicos (SESCAM), Toledo, España, lbcoloma@externas.sescam.jccm.es, vlozanob_adlos@sescam.jccm.es

² Departamento de Tecnologías y Sistemas de Información, UCLM, Ciudad Real, España, vanesa.herrera_javieralonso.albusac@uclm.es, carlos.gonzalez@uclm.es

³ Unidad de Terapia Ocupacional, Hospital Nacional de Paraplégicos (SESCAM), Toledo, España, rp.gomez@sescam.jccm.es

⁴ Servicio de Rehabilitación, Hospital Nacional de Paraplégicos (SESCAM), Toledo, España, malcobendas_sceruelo@sescam.jccm.es

⁵ Unidad de Neurorehabilitación, Biomecánica y Función Sensitivo-Motora (HNP-SESCAM), Unidad Asociada de I+D+I al CSIC

Resumen

Hay un interés creciente en nuevas terapias por medio de tecnología, como la Realidad Virtual (RV). La RV ofrece un gran potencial para mejorar la función motora de miembro superior (MS) en paciente neurológicos. No obstante, se desconoce acerca de las características de los patrones de movimiento ejecutados en entornos virtuales en contexto rehabilitador. Por ello, este estudio analiza las características cinemáticas de los movimientos de MS durante la ejecución de una tarea funcional, el test clínico Box and Block (BBT), en pacientes con lesión medular (LM) cervical y en población sana en entornos virtuales inmersivos (3D), no inmersivos (2D) y en el entorno físico real. El análisis principal se centra en comparar la cinemática de los pacientes con la de los sujetos sanos para conocer la funcionalidad global del MS durante la ejecución del BBT en los tres entornos de experimentación. Para ello se estudian medidas relativas al desempeño de la tarea y medidas relativas al patrón de movimiento. Finalmente, existen diferencias entre los patrones cinemáticos ejecutados en el BBT de entorno real y los de RV. En ambas poblaciones existen diferencias en los rangos de movimiento (ROM) articular de codo y muñeca. En pacientes con LM, además, varían el ROM de hombro y el desplazamiento de tronco. En ambos grupos, la velocidad y suavidad disminuyen en entornos de RV respecto al entorno físico real. La longitud de trayectoria aumenta en sanos en el entorno de RV 3D, mientras que, en el caso de pacientes con LM, disminuye.

1. Introducción

La pérdida de la función motora es una de las consecuencias más comunes tras una lesión neurológica como la Lesión Medular (LM), cuya incidencia mundial varía desde 12.1 a 57.8 casos por millón de habitantes [1]. Más del 50% de lesionados medulares presentan alterada la función de los miembros superiores (MS), experimentando limitaciones en la realización de tareas funcionales y actividades básicas de la vida diaria (ABVD). En este contexto, uno de los principales objetivos de la rehabilitación es lograr el máximo nivel de

independencia del paciente en la realización de ABVD [2-3].

Así, actualmente, se trabaja para poner a punto tecnología que permita mejorar la calidad de vida de las personas con LM, favoreciendo la rehabilitación y la recuperación en la medida de lo posible de aquellas funciones perdidas a causa de la lesión. En este sentido está habiendo un interés creciente en nuevas terapias por medio de tecnología, siendo una de ellas mediante técnicas de Realidad Virtual (RV). Está bien estudiado el potencial que la RV ofrece como herramienta terapéutica, fomentando el grado de interacción del paciente con la terapia minimizando aspectos como el aburrimiento y la falta de motivación, que pueden afectar negativamente a la recuperación motora.

La eficacia de la RV como herramienta para mejorar la función motora de los MS ha sido analizada en pacientes neurológicos. Se tiene evidencia de estudios realizados en ictus [4,5] y en LM [6]. Sin embargo, no se tiene tanto conocimiento acerca de las características cinemáticas de los patrones de movimiento ejecutados en entornos físicos reales y en entornos virtuales en contexto rehabilitador. Estos estudios enfocados en los patrones cinemáticos se han realizado fundamentalmente en población sana [7] y en pacientes con ictus [8-11], siendo escasos actualmente en LM. Por ello, el objetivo de este estudio es analizar las características cinemáticas de los movimientos de MS en pacientes con LM cervical realizados en entornos virtuales inmersivos, no inmersivos y en el entorno físico real, con el fin de proponer entornos virtuales inmersivos adaptados a las necesidades de los pacientes. Se incluye también población sana para caracterizar, desde el punto de vista cinemático, la misma tarea funcional en los diferentes entornos experimentales.

2. Metodología

2.1. Procedimiento experimental

La valoración cinemática del MS, establecida como prueba clínica biomecánica de análisis del movimiento, proporciona datos específicos y objetivos acerca de la funcionalidad del MS durante la ejecución de una tarea funcional. En este caso, se analiza la cinemática del movimiento del MS dominante en población sana y en pacientes con LM cervical durante la ejecución del test clínico Box and Block (BBT) realizado en tres entornos cuyo orden se aleatoriza:

- Entorno físico real, mediante la ejecución del BBT, escala clínica validada y estandarizada para valorar la destreza manipulativa.
- Entorno virtual no inmersivo, mediante la ejecución del BBT implementado para las gafas de RV Oculus Quest 2, cuya visualización es retransmitida en 2D en la pantalla de un PC.
- Entorno virtual inmersivo, ejecución del BBT implementado para las gafas de RV Oculus Quest 2 en 3D.

El movimiento del MS se registra con el sistema optoelectrónico Vicon (Vicon Motion System, Oxford, Reino Unido). Vicon es un sistema de captura del movimiento (mocap) óptico-pasivo que utiliza marcadores retroreflectantes que son rastreados por cámaras de infrarrojos. Actualmente, es el sistema más preciso utilizado en la industria. En esta prueba, se emplean un total de 21 marcadores pasivos repartidos entre el tronco y la extremidad superior dominante con la que se realiza la tarea (Figura 1).

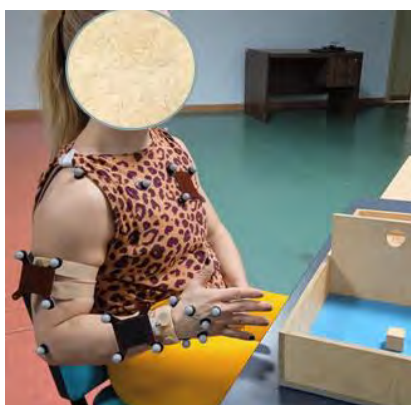


Figura 1. Participante sano instrumentado con los marcadores del equipo de fotogrametría Vicon durante una prueba experimental

El protocolo experimental está compuesto por una sola sesión, con una duración aproximada de una hora, que se divide en tres partes: (i) anamnesis, exploración física y funcional del paciente, (ii) instrumentación, (iii) desarrollo de la prueba biomecánica de MS y adquisición de datos. Previo a realizar la prueba se añade una breve fase de acomodamiento y aclimatación al entorno.

El usuario realiza la prueba en sedestación en una silla convencional (siempre la misma) o en su propia silla de ruedas, frente a una mesa cuya altura es regulable hasta

conseguir una flexión de codo de 90° con la palma de la mano apoyada sobre la superficie de la mesa. Esta posición determina la posición de inicio. Durante la prueba biomecánica se registra la trayectoria del movimiento asociado a coger un cubo del BBT y pasarlo al otro lado de la partición de la caja que conforma el test, partiendo de la posición de inicio y finalizando en la misma. Se registran un total de 10 patrones en cada entorno experimental.

El tratamiento de los datos adquiridos se realiza con el software Visual3D, que permite la obtención de la información cinemática asociada a la funcionalidad del MS. Se obtuvo la conformidad del Comité de Ética para realizar el estudio.

2.2. Análisis de datos

El análisis principal se centra en comparar la cinemática de los pacientes con LM cervical con la de los sujetos sanos para conocer la funcionalidad global del MS durante la ejecución de la tarea correspondiente al BBT en los tres entornos de la experimentación. Con respecto a la cinemática, las variables registradas son:

- Medidas relativas al desempeño de la tarea: velocidad pico y velocidad media de la mano durante el movimiento realizados, y longitud y suavidad en la trayectoria realizada.
- Medidas relativas al patrón de movimiento: rango de movimiento (ROM) articular de hombro, codo y muñeca, desplazamiento de tronco, y coordinación entre articulaciones de hombro y codo.

En segundo plano, se realiza también un análisis de la funcionalidad del MS en los pacientes con LM cervical que se relaciona con la puntuación obtenida en las escalas clínicas del Test de Edinburg, Medida de Independencia en Lesión Medular (SCIM II) y el Índice Motor del Miembro Superior (UEMS).

El análisis descriptivo se realiza con SPSS 17.0 para Windows (SPSS Inc., Chicago, IL). En cuanto a las variables cinemáticas, se comprueba su normalidad con la prueba de Kolmogorov-Smirnov. En caso de no seguir una distribución normal, se estudian mediante el test no paramétrico de Wilcoxon y las posibles diferencias existentes entre el grupo de sanos y el grupo de pacientes con LM cervical se analizan mediante el test Mann-Whitney.

La relación entre las variables cinemáticas obtenidas y las relativas a la valoración funcional se evalúan por medio de los coeficientes de correlación de Pearson y Spearman.

3. Resultados

Los resultados que se presentan son aquellos obtenidos hasta el momento tras realizar las pruebas experimentales con tres sujetos sanos y tres pacientes con LM cervical. El análisis de todas las variables se ha realizado con el test no paramétrico de Wilcoxon.

En cuanto a los resultados obtenidos tras el análisis de las medidas relativas al desempeño de la tarea, se presentan

la Tabla 1 para el grupo de sujetos sanos y la Tabla 2 para el grupo de pacientes con LM cervical.

En el grupo de sanos (Tabla 1), se observa una trayectoria significativamente más larga en el entorno de RV inmersivo 3D, además de un patrón de movimiento menos suave en ambos entornos con RV. En cuanto al grupo de pacientes (Tabla 2), se aprecia una trayectoria más larga en el entorno físico real respecto a los entornos con RV, siendo la trayectoria más corta la obtenida para el entorno de RV no inmersiva 2D. Los movimientos son menos suaves y más lentos en los entornos de RV.

Sanos (n=3)	BBT Real	RV no inmersiva	RV inmersiva
Velocidad pico (m/s)	0.88 (0.09)a,d	0.65(0.05)a,b	0.79(0.09)b,d
Velocidad media (m/s)	0.55 (0.04)a,c	0.32(0.02)a,b	0.39(0.02)b,c
Longitud de trayectoria (m)	1.44 (0.12)d	1.13 (1.45)a	1.56(2.07)a,d
Número de picos (unidades)	3.20 (0.77)a,b	5.46 (1.35)a	5.46 (1.40)b

Tabla 1. Resultados estadísticos de las medidas relativas al desempeño de la tarea en el grupo de sujetos sanos (a,b,c: $p < 0.01$; d: $p < 0.05$)

LM (n=3)	BBT Real	RV no inmersiva	RV inmersiva
Velocidad pico (m/s)	0.89 (0.07)a,c	0.79(0.09)a,d	0.83(0.10)c,d
Velocidad media (m/s)	0.53 (0.07)a,b	0.35 (0.06)a,d	0.38 (0.03)b,d
Longitud de trayectoria (m)	2.00 (0.20) a	1.56 (0.29)a,d	1.78 (0.33)d
Número de picos (unidades)	2.40 (1.29)	5.26 (1.16)	4.80 (1.74)

Tabla 2. Resultados estadísticos de las medidas relativas al desempeño de la tarea en el grupo de pacientes con LM cervical (a,b,c: $p < 0.01$; d: $p < 0.05$)

En las Tablas 3 y 4 se muestran los rangos de movimiento articular, y en las Tablas 5 y 6 los valores de desplazamiento del tronco y la coordinación interarticular, para sanos y pacientes, respectivamente.

Sanos (n=3)	BBT Real	RV no inmersiva	RV inmersiva
ROM total (°)	274.55(21.84)	282.62(35.51)	296.37(24.12)
ROM hombro (°)	158.76(15.27)	156.52 (28.89)	168.33 (36.85)
ROM codo (°)	94.27(11.59)a,b	112.19(11.61)a	111.97(18.12)b
ROM muñeca (°)	21.52(12.78)a,b	13.90 (7.03)a	16.06 (9.04)b

Tabla 3. Resultados estadísticos de las medidas relativas al patrón de movimiento (ROM articular) en el grupo de sujetos sanos (a,b,c: $p < 0.01$; d: $p < 0.05$)

En la Tabla 3, se aprecia como en el grupo de sanos el ROM de codo es significativamente mayor en los

entornos de RV que en el BBT de entorno real. Sin embargo, el ROM de muñeca se observa lo opuesto.

Respecto al grupo de pacientes, en la Tabla 4, el ROM de codo es mayor en el entorno de RV 3D que en RV 2D. En ambos entornos virtuales el ROM de muñeca es mayor que en BBT real, siendo significativo en RV 2D. Por último, el ROM de hombro es significativamente menor en RV no inmersiva en comparación con el ROM obtenido en las versiones real e inmersiva.

En referencia al desplazamiento de tronco (flexión-extensión) y a la coordinación, la Tabla 5 muestra que, para sujetos sanos, hay pequeñas desviaciones de tronco sin significación estadística entre los tres entornos experimentales, así como una elevada coordinación, significativamente mayor en el BBT real.

LM (n=3)	BBT Real	RV no inmersiva	RV inmersiva
ROM total (°)	238.92(18.79)a,c	188.91(50.19)a,b	263.29(21.83)b,c
ROM hombro (°)	133.74(13.14)a	77.28(22.82)a,b	128.43(38.49)b
ROM codo (°)	90.92(14.20)	90.80(31.15)d	115.41(33.23)d
ROM muñeca (°)	14.25 (4.32)a	20.82 (4.00)a	19.44 (4.40)

Tabla 4. Resultados estadísticos de las medidas relativas al patrón de movimiento (ROM articular) en el grupo de pacientes con LM cervical (a,b,c: $p < 0.01$; d: $p < 0.05$)

Por otro lado, en el grupo de pacientes con LM (Tabla 6) la desviación de tronco es significativamente menor en ambos entornos de RV en comparación con en el entorno físico real. Igualmente, la coordinación es mayor en entornos de RV, especialmente en el entorno de RV inmersivo.

Sanos (n=3)	BBT Real	RV no inmersiva	RV inmersiva
ROM tronco (°)	5.95 (2.59)	2.89 (2.55)	6.82 (3.05)
Coordinación (Pearson* 100)	94.04 (1.94)a,b	87.37 (8.03)a	87.13 (5.45)b

Tabla 5. Resultados estadísticos de las medidas relativas al patrón de movimiento (ROM de tronco y coordinación interarticular) en el grupo de sujetos sanos (a,b,c: $p < 0.01$; d: $p < 0.05$)

LM (n=3)	BBT Real	RV no inmersiva	RV inmersiva
ROM tronco (°)	7.77 (4.36)a,c	3.14 (1.39)a,b	4.27 (1.12)b,c
Coordinación (Pearson* 100)	27.97 (19.56)	37.12 (21.76)a	54.52 (34.24)a

Tabla 6. Resultados estadísticos de las medidas relativas al patrón de movimiento (ROM de tronco y coordinación interarticular) en el grupo de pacientes con LM cervical (a,b,c: $p < 0.01$; d: $p < 0.05$)

4. Discusión

La metodología propuesta ha permitido obtener indicios acerca del nivel de similitud entre los patrones de movimientos ejecutados para la tarea del BBT.

En cuanto a los recorridos articulares se han medido diferencias entre los distintos entornos experimentales, en principio no coincidentes con la bibliografía previa que sugiere que las diferencias se deben a haber utilizado un entorno RV 2D [8]. En pacientes con LM se observan también diferencias a nivel proximal en la articulación del hombro.

En cuanto al desplazamiento de tronco, no se detectan diferencias significativas en el grupo de sanos. Sin embargo, en el grupo de pacientes sí existe una compensación de flexión-extensión de tronco, que presenta menor desviación en los entornos de RV que en el entorno real, lo que coincide con estudios previos [9].

Ambas poblaciones exhiben un patrón de movimiento menos suave y con velocidad media más lenta en entornos de RV que en el entorno físico. Según la literatura en la patología del ictus, los movimientos son más curvados en los entornos de RV [8-11]. En este estudio realizado con pacientes neurológicos con LM cervical, se observa una trayectoria más larga en entorno físico real respecto a las trayectorias seguidas en entornos de RV. Sin embargo, la longitud de trayectoria en ambas poblaciones presenta una disminución notable en el entorno de RV no inmersivo.

La variabilidad que presentan estas tres últimas medidas relativas al desempeño de la tarea puede verse influenciada por la no manipulación de objetos reales en RV. Esto supone, entre otras cosas, la ausencia de retroalimentación háptica, lo que puede llevar a realizar movimientos menos naturales.

Hasta el momento, todos los participantes han podido realizar las tareas virtuales gracias al óptimo tracking de manos, que permite adaptar el sistema a las necesidades reales del paciente.

Entre las limitaciones del estudio cabe destacar que no hay consenso en cuanto a las tareas a realizar para analizar la ejecución de patrones de movimiento de MS en los diferentes entornos, así como las evidencias bibliográficas reducidas a población con ictus.

5. Conclusión

Los resultados obtenidos sugieren que la metodología propuesta es adecuada para analizar la posible similitud entre los patrones de movimiento en ambos entornos virtuales y el entorno físico real.

Actualmente se continúa aumentando la muestra experimental para comprobar la consistencia de los resultados mostrados.

No obstante, la futura investigación en esta línea va a requerir la incorporación de Realidad Mixta que permita incluir objetos reales en los entornos virtuales.

Agradecimientos

Proyecto REHAB-IMMERSIVE PID2020-117361RB-C22 financiado por MCIN/AEI/10.13039/501100011033.

Referencias

- [1] Van den Berg MEL, Castellote JM, Mahillo-Fernandez I, de Pedro-Cuesta J. Incidence of spinal cord injury worldwide: a systematic review. *Neuroepidemiology*, vol 34, sup 3, 2010, pp 184-192 (ISSN: 1423-0208).
- [2] Beninato M, O'Kane KS, Sullivan PE. Relationship between motor FIM and muscle strength in lower cervical-level spinal cord injuries. *SpinalCord*, vol 42, sup 9, 2004, pp 533-540 (ISBN: 1362-4393).
- [3] Schmitz TJ. Traumatic spinal cord injury. In: Sullivan SB, Schmitz TJ (eds.) *Physical Rehabilitation: Assessment and Treatment*, 4th edn. F.A. Davis Company: Philadelphia, 2001, pp 873-923.
- [4] Maggio MG, Latella D, Maresca G, et al. Virtual Reality and Cognitive Rehabilitation in People with Stroke: An Overview. *Journal of Neuroscience Nursing*, 51(2), 2019, pp 101-105. doi:10.1097/JNN.0000000000000423
- [5] Laver KE, Lange B, George S, Deutsch JE, Saposnik G, Crotty M. Virtual reality for stroke rehabilitation. *Cochrane Database of Systematic Reviews*, 11(11), 2017, CD008349. doi:10.1002/14651858.CD008349.pub4
- [6] de Araújo AVL, Neiva JFO, Monteiro CBM, Magalhães FH. Efficacy of Virtual Reality Rehabilitation after Spinal Cord Injury: A Systematic Review. *Biomed Research International*, 2019, 7106951. Published 2019 Nov 13. doi:10.1155/2019/7106951
- [7] Magdalon EC, Michaelsen SM, Quevedo AA, Levin MF. Comparison of grasping movements made by healthy subjects in a 3-dimensional immersive virtual versus physical environment. *Acta psychologica*, 138(1), 2011, pp 126-134.
- [8] Viau A, Feldman AG, McFadyen BJ, Levin MF. Reaching in reality and virtual reality: a comparison of movement kinematics in healthy subjects and in adults with hemiparesis. *Journal of neuroengineering and rehabilitation*, 1(1), 2004, 1-7.
- [9] Knaut LA, Subramanian SK, McFadyen BJ, Bourbonnais D, Levin MF. Kinematics of pointing movements made in a virtual versus a physical 3-dimensional environment in healthy and stroke subjects. *Archives of physical medicine and rehabilitation*, 90(5), 2009, pp 793-802.
- [10] Levin MF, Magdalon EC, Michaelsen SM, Quevedo AA. Quality of grasping and the role of haptics in a 3-D immersive virtual reality environment in individuals with stroke. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 23(6), 2015, pp 1047-1055.
- [11] Arlati S, Keijsers N, Ferrigno G, Sacco M. A protocol for the comparison of reaching gesture kinematics in physical versus immersive virtual reality. *IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, 2020, pp 1-6.

Data augmentation study for rare diseases assessment with Deep Learning: Confocal Imaging analysis of Congenital Muscular Dystrophy

M. Frías¹, C. Jiménez-Mallebrera^{2,3,4}, C. Badosa², J.M. Porta*⁵, M. Roldán *¹

¹ Unitat de Microscòpia Confocal i Imatge Cel·lular, Departament de Medicina Genètica i Molecular, Institut Pediàtric de Malalties Rares, Hospital Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain. marcos.frias@sjd.es, monica.roldan@sjd.es

² Laboratorio de Investigación Aplicada en Enfermedades Neuromusculares, Unidad de Patología Neuromuscular, Servicio de Neuropediatría, Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain. cecilia.jimenez@sjd.es, mariacarmen.badosa@sjd.es

³ Centro de Investigaciones Biomédicas en Red de Enfermedades Raras (CIBERER), Madrid, Spain.

⁴ Departamento de Genética, Microbiología y Estadística, Universitat de Barcelona, Barcelona, Spain.

⁵ Institut de Robòtica i Informàtica Industrial, UPC-CSIC, Barcelona, Spain. porta@iri.upc.edu

*Both authors contributed equally to this work

Summary

Artificial Intelligence (AI) algorithms are widely used in healthcare nowadays. However, there are still fields where the application of these technologies could be challenging, such as rare diseases. In these cases, the main challenge arises from the reduced size of the available data sets. This paper proposes a data augmentation pipeline to address this challenge when using a Deep Learning (DL) algorithm to assess fibroblast cultures from skin biopsies to diagnose Collagen VI-related Congenital Muscular Dystrophy (COL6-CMD). Different data augmentation schemes are described in the literature. However, they must be used cautiously since they might result in overfitting. The results presented in this paper demonstrate that the right combination of data augmentation techniques results in a high diagnostic accuracy (up to 75.35% for the best approach) even with a scarce amount of data.

1. Introduction

Collagen VI-related Congenital Muscular Dystrophy (COL6-CMD) is a rare disease¹ that produces deficiencies in the structure of the protein Collagen VI [1]. Depending on the severity, different manifestations can be found. Together with the low prevalence and the typically limited understanding of rare diseases, it results very hard to achieve the correct diagnosis [2]. The visual analysis of fibroblast cultures with optical and photonic microscopy technologies is one of the most relevant techniques to reach the precise diagnosis of this disease [3]. More specifically, confocal microscopy provides a high signal sensibility able of extracting detailed information even in intermediate and mild manifestations of the disease [4]. Microscopy images present Collagen VI networks such as

the ones shown in Figure 1. Different aspects are considered to assess them, such as the collagen fibres distribution and intensity or the nuclei shape, which might appear altered in pathological cases. Nevertheless, they are always inspected from the subjective point of view of the professional. For that reason, novel Artificial Intelligence (AI) techniques can provide more efficient and objective procedures to perform this task.

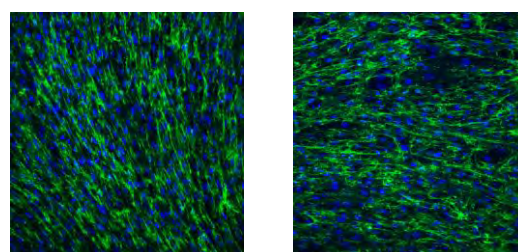


Figure 1 Example of control (left) and patient (right) confocal images of primary fibroblasts cultures. The collagen fibres are shown in green and the cell nuclei in blue

Some studies have demonstrated the viability of implementing AI approaches for microscopy samples and rare disease diagnosis [5]. However, the reduced number of patients suffering from these types of diseases generates a clear limitation regarding the amount of data available, which represents the main difficulty when using these technologies. This is especially true in Deep Learning (DL), where samples must be as representative as possible so that the algorithm can determine the key features to correctly perform the task [6]. Thus, there is a critical demand for novel data augmentation approaches in this context.

A few AI techniques have been applied to help on automatizing and accelerating the diagnosis process more quantitatively and objectively in COL6-CMD. In 2018,

¹ In Europe, a disease is considered as rare when its prevalence is below 1 in 2000.

Bazaga et al. developed the first Computer-Aided Diagnosis (CAD) system for COL6-CMD based on a handcrafted DL algorithm trained with a small database formed by confocal microscopy acquisitions of primary fibroblast cultures extracted from skin biopsies. The homogeneous nature of the samples allowed for the generation of small, fixed-size image patches to generate a larger dataset without affecting the biological structure for identifying the disease. However, the technique implications in the algorithm performance were not analysed in detail [7].

This paper explores which are the key factors to consider when implementing a data augmentation pipeline to use DL technologies in a rare disease context. The COL6-CMD is analysed as an example in a limited-available dataset scenario. Three different data augmentation techniques are explored. First, the patches generation approach from Bazaga et al. is extended by trying different patch sizes. Additionally, different degrees of overlap between patches are considered. Finally, standard data augmentation transformations are applied.

2. Methods

2.1. Dataset

The amount of available data is crucial for an AI model to succeed. In 2018, the Neuromuscular Unit at the Research Institute of the Hospital Sant Joan de Déu together with the Confocal Microscopy and Cellular Imaging Unit started a database with fibroblasts culture confocal images acquired with a Leica TCS SP8 equipped with a white light laser and hybrid spectral detectors (Leica Microsystems GmbH, Mannheim, Germany). Confocal images, each consisting of 1024×1024 pixels, was accomplished using an HCX PL APO 20x/0.75 dry objective, with the confocal pinhole set to 1 Airy unit. Collagen VI fluorescence was excited by an argon laser (488 nm) and detected within the 500–560 nm range. Simultaneously, nuclei were excited with a blue diode laser (405 nm) and detected within the 420–460 nm range. To eliminate any potential channel interference, we employed sequential acquisition settings. Different images, called fields, are extracted from a single fibroblasts' culture. For each field, a series of ten sections, spaced 1.5 μm apart along the focal axis (Z-stack), were acquired. These individual sections were subsequently integrated into a maximum intensity projection, resulting in a single comprehensive image.

In 2022, this dataset contained 411 TIFF images, 221 were labelled as controls and 190 as patients. They are extracted from 24 control biopsies and 19 patient biopsies. This images set is divided in training and testing subsets, with an 80-20% ratio respectively. It is crucial to remark that there is a dependency between the different fields extracted from the same cells' culture. Thus, they cannot be considered as independent (i.e., all the fields from the same cells' culture can only be assigned to either the training or the test set, but not split between them).

2.2. Data augmentation

The dataset size was not enough for training a Convolutional Neural Network (CNN) satisfactorily. For that reason, a data augmentation pipeline was developed. First, each image was split into small patches. In that way, the different patches could be classified individually to be posteriorly grouped at image level following a majority voting system. Different patch sizes were explored: 64x64x, 128x128 and 224x224 pixels. Moreover, partial overlapping was considered to further increase the dataset dimension. The degree of overlap considered were 0%, 25%, 50%, and 75%. Table 1 shows the resulting number of patches generated when applying the different combinations of data augmentation. Additionally, other transformations were performed on-the-fly. Such transformations are random rotations and horizontal and vertical flips, which were applied each time data was set into the network with a 0.5 probability.

		Patch size (in pixels)		
		64x64	128x128	224x224
Overlap	0%	105.216	26.304	10.275
	25%	180.840	49.320	14.385
	50%	394.560	92.064	32.880
	75%	1.528.920	345.240	104.805

Table 1 Number of total patches for the different patch sizes and the degree of overlap between patches, after manual data augmentation

2.3. Convolutional Neural Network

To classify the available images, we used the EfficientNetB0 network since it outperforms other alternatives with a relatively small number of parameters [8]. The network is initialized with pre-trained weights that are latter fine-tuned for the considered problem [9]. In this way, 5 epochs (with a batch size of 64) are enough for the training algorithm to converge.

The reduced dataset size in these type of pathologies causes the model to be very dependent on the split between training and testing. For that reason, we follow the K-fold cross-validation technique: the results are averaged over five different training-test partitions [10]. Moreover, an accumulative confusion matrix for the five partitions is compute to generate more robust results.

3. Results

Both the training and test accuracies and the Area Under the Curve (AUC) were used quantify the model performance when applying different data augmentation techniques. Tables 2 and 3 summarize the obtained results.

		Patch size (in pixels)			
		64x64	128x128	224x224	
Overlap	0%	Train	95.44 ± 0.27	97.59 ± 0.24	97.91 ± 0.12
		Test	60.40 ± 2.08	62.54 ± 4.91	74.53 ± 3.79
	25%	Train	97.00 ± 0.16	97.96 ± 0.15	98.52 ± 0.14
		Test	60.04 ± 4.48	67.12 ± 6.78	73.31 ± 4.86
	50%	Train	98.25 ± 0.18	99.12 ± 0.18	99.17 ± 0.04
		Test	57.28 ± 5.26	64.57 ± 5.47	72.77 ± 4.76
	75%	Train	99.32 ± 0.03	99.64 ± 0.01	99.61 ± 0.03
		Test	55.83 ± 4.62	66.8 ± 3.08	72.60 ± 2.94

Table 2 Average train and test accuracies obtained with each combination of patch size and degree of patch overlap over the 5-fold cross-validation. The standard error is also shown.

		Patch size (in pixels)		
		64x64	128x128	224x224
Overlap	0%	0.671 ± 0.036	0.659 ± 0.078	0.811 ± 0.041
	25%	0.612 ± 0.061	0.706 ± 0.105	0.809 ± 0.054
	50%	0.546 ± 0.092	0.733 ± 0.063	0.767 ± 0.061
	75%	0.562 ± 0.066	0.705 ± 0.047	0.792 ± 0.042

Table 3 Average AUC obtained with each combination of patch size and degree of patch overlap over the 5-fold cross-validation. The standard error is also shown.

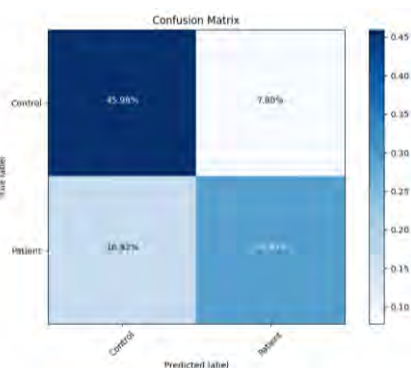


Figure 2 Accumulative confusion matrix at patch level obtained with patches of 224x224 pixels, with no overlap between patches, and with standard data augmentation techniques applied on-the-fly.

4. Discussion

The implementation of DL techniques in healthcare is flourishing, largely due to the increasingly available medical data and to the more powerful computer resources. The main factor for a DL algorithm to succeed is to have enough available data. Otherwise, the learned model would not properly generalize the desired task. For this reason, the application of this technology to a rare disease, where data is severely limited, is a very challenging task that requires the development of data augmentation techniques [11].

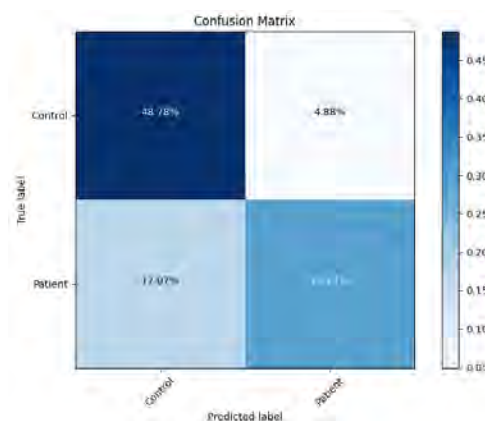


Figure 3 Accumulative confusion matrix at fibroblasts' culture level obtained after majority voting of patches of 224x224 pixels, with no overlap between patches, and with standard data augmentation techniques applied on-the-fly.

Different approaches have been explored in this paper to deal with data limitations in a COL6-CMD clinical scenario. This disease presents alterations over a homogeneous collagen network, which allowed the generation of smaller patches to increase the number of samples, while preserving the complete biological structure [12]. However, the balance between the size and the overlap of these patches is key for the technique to succeed, since very small regions might not include enough biological information while big patches might not generate enough samples for effectively training the network.

Results showed that it is possible to achieve a good performance when implementing a DL algorithm if the adequate data augmentation techniques are applied. The best results (with a 74.53% of testing accuracy) were achieved with the 224x224 patch size with no overlap. The results indicate that the smaller the patch, the worst the result. Regarding the overlap between patches, in general, no improvement appears when increasing the percentage of overlap. The higher the overlap, the larger the data set, but also the larger the possibility of overfitting the model, since the network is trained repetitively with the same sub-patches. This causes the testing metrics to decrease while the training metrics become higher [13]. This behaviour can be observed in Table 2, which suggests that overlapping patches might not be a relevant data augmentation technique for this type of biological samples and clinical context. The high AUC values confirm the capability of the model to correctly discern between classes.

Standard on-the-fly data augmentation techniques in DL slightly improve the model performance, as expected [14].

Control samples are more coherent between them, but patient samples are more diverse since the disease can appear in different manifestations and degrees of severity. Therefore, not both classes are classified equally correctly. The algorithm learnt to recognize patches coming from control cultures almost perfectly since only

14% of the total healthy patches were misclassified over the 5-fold cross-validation process. On the contrary, the patches extracted from patient biopsies are more complex to identify. In this case, 37% of the total patches were misclassified as control. At the biopsy level, using the majority voting approach, similar results were obtained. This confirms the effective extrapolation of this data augmentation technique at the person level, in order to give an indicative diagnosis to the person suffering from this disease.

Alternative data augmentation techniques could be explored to complement this pipeline. One option could be to explore generative models such as Generative Adversarial Networks (GANs), which aim to learn to generate synthetic realistic images to enlarge the database. This is a complex task prone to mode collapse and the production of hallucinations. Nevertheless, it is a very promising field of work in which many advances have been recently done [15]. Additionally, shared protocols between different centres could be established to allow for sharing their data and increment the size of the image data sets.

In conclusion, many aspects have to be considered for developing an effective data augmentation pipeline when applying DL techniques to a rare disease. The dataset size and organization, the knowledge about the underlying biologic structure, and the photonic techniques used for the image acquisitions requires professionals from multiple fields to work together. In this way, the challenges appearing when using DL techniques on the small datasets available in the context of rare and ultra-rare diseases can be successfully overcome.

Acknowledgments

Funded by the European Union. We are indebted to the HORIZON–MSCA–2022–DN, Improving BioMedical diagnosis through LIGHT-based technologies and machine learning “BE-LIGHT” (GA n° 101119924 – BE-LIGHT).

References

- [1] Natera-de Benito D et al. Association of Initial Maximal Motor Ability With Long-term Functional Outcome in Patients With COL6-Related Dystrophies. *Neurology*. 2021 Mar 9;96(10):e1413-e1424.
- [2] Wakap SN; Lambert DM; Olry A; Rodwell C, Gueydan C, Lanneau V, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *Eur J Hum Genet*. 28:165–173. (2020).
- [3] Jimenez-Mallebrera, C., Maioli, M. A., Kim, J., Brown, S. C., Feng, L., Lampe, A. K., Bushby, K., Hicks, D., Flanigan, K. M., Bönnemann, C. G., Sewry, C. A., & Muntoni, F. (2006). A comparative analysis of collagen VI production in muscle, skin and fibroblasts from 14 Ullrich Congenital Muscular Dystrophy patients with dominant and recessive COL6A mutations. *Neuromuscular Disorders*, 16(9-10), 571-582. <https://doi.org/10.1016/j.nmd.2006.07.015>
- [4] Nwaneshiudu, A., Kuschal, C., Sakamoto, F. H., Rox Anderson, R., Schwarzenberger, K., & Young, R. C. (2012). Introduction to confocal microscopy. *Journal of Investigative Dermatology*, 132(12), 1-5. <https://doi.org/10.1038/jid.2012.429>
- [5] Lee, J. H., Liu, C., Kim, J., Chen, Z., Sun, Y., Rogers, J. R., Chung, W. K., & Weng, C. (2022). Deep Learning for Rare Disease: A scoping review. *Journal of Biomedical Informatics*, 135, 104227. <https://doi.org/10.1016/j.jbi.2022.104227>
- [6] Yu, K. H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature Biomedical Engineering*, 2(10), 719–731. <https://doi.org/10.1038/s41551-018-0305-z>
- [7] Bazaga, A., Roldán, M., Badosa, C., Jimenez-Mallebrera, C., & Porta, J. M. (2019). A convolutional neural network for the automatic diagnosis of collagen VI-related muscular dystrophies. *Applied Soft Computing*, 85, 105772. <https://doi.org/10.1016/j.asoc.2019.105772>
- [8] Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning* (pp. 6105-6114). PMLR
- [9] Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A Survey on Deep Transfer Learning. *Lecture Notes in Computer Science*, 270-279. https://doi.org/10.1007/978-3-030-01424-7_27
- [10] Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L., & Ridella, S. (2012, April). The 'K' in K-fold Cross Validation. In *ESANN* (pp. 441-446).
- [11] Alsentzer, E., Li, M. M., Kobren, S. N., Undiagnosed Diseases Network, Kohane, I. S., & Zitnik, M. (2022). Deep learning for diagnosing patients with rare genetic diseases. *medRxiv*, 2022-12.
- [12] Garcea, F., Serra, A., Lamberti, F., & Morra, L. (2023). Data Augmentation for Medical Imaging: A Systematic Literature Review. *Computers in Biology and Medicine*, 152, 106391. <https://doi.org/10.1016/j.compbimed.2022.106391>
- [13] Ying, X. (2019, February). An overview of overfitting and its solutions. In *Journal of physics: Conference series* (Vol. 1168, p. 022022). IOP Publishing.
- [14] Chlap, P., Huang, M., Vandenberg, N., Dowling, J., Holloway, L., & Haworth, A. (2021). A review of medical image data augmentation techniques for deep learning applications. *Journal of Medical Imaging and Radiation Oncology*, 65(5), 545-563. <https://doi.org/10.1111/1754-9485.13261>
- [15] Zhao, S., Liu, Z., Lin, J., Zhu, J. Y., & Han, S. (2020). Differentiable augmentation for data-efficient gan training. *Advances in neural information processing systems*, 33, 7559-7570.

Evaluating Techniques for Neuron Identification in Complex Cultures: A Deep Learning Approach ¹

Paula Puerta ¹, Berke Öztürk ¹, Víctor M. González ², José R. Villar ¹, Esther Serrano-Pertierra ³, Antonello Novelli ⁴, M. Teresa Fernández-Sánchez ³, and Ángel Río-Álvarez ¹

¹ Computer Science Department, University of Oviedo, Spain, rioangel@uniovi.es.

² Electrical Engineering Department, University of Oviedo, Spain

³ Biochemistry and Molecular Biology Department; University Institute of Biotechnology of Asturias (IUBA), University of Oviedo, Spain

⁴ Psychology Department; University Institute of Biotechnology of Asturias (IUBA), University of Oviedo, Spain

Abstract

Microscopy image analysis of neurons cultures represents a formidable challenge due to their complex structure because of the dynamic nature of neurite tissue development, the neuron movement, the morphological changes, and the presence of many elements in the cultures apart from neurons such as glial cells, dead cells, vesicles, etc. A rigorous evaluation of deep learning techniques to address this intricate problem is undertaken in this study. Several methodologies, including Instance Segmentation and Object Detection models, are scrutinized within a comprehensive experimental framework. The efficacy of the Instance Segmentation model is underscored by the findings, demonstrating superior quantitative results. Precise neuron quantification is facilitated by this model through the detection of bounding boxes in images, thereby enabling the automation of tasks such as morphological and size analysis of neuronal cells and tracking individual neurons across sequential images from the same culture. This paper serves as a robust foundation for future investigations aimed at expanding the understanding of neurons cultures.

1. Introduction

The development of quantitative methods for the rapid and accurate evaluation of neuronal cell viability and morphological changes is a critical aspect of drug screening and the determination of neuronal loss associated with neuropathologies [1, 2]. Neurotoxicity, a well-accepted *in vitro* marker in neuronal models of stroke and neurodegenerative diseases such as Alzheimer's and Parkinson's, is a key factor in these assessments [3]. Furthermore, the evaluation of neurotoxicity is crucial in the safety evaluation of drugs and environmental chemicals [4].

Traditional biochemical approaches to measuring neurotoxicity rely on the observation that cell death cascades alter enzymatic activity, cell membrane permeability, and neuronal mitochondrial activity. However, these methods often lack single-cell resolution and can involve complex and costly methodologies.

In contrast, phase contrast optics allows for the qualitative detection of early and late signs of toxicity in cultured neurons, including changes in morphology,

and neurite outgrowth and integrity. However, direct quantitative assessment of neuronal viability from phase-contrast images is not possible without additional methodologies.

One such method involves differential staining with fluorescein diacetate or Calcein AM and ethidium bromide, which can be used to visualize live cells and nuclei of dead neurons [5, 2]. This approach, while inexpensive and reliable, has significant drawbacks. It requires the permeabilization of the neuronal cell membrane with organic solvents, preventing long-term studies, and necessitates the precise tracking of individual viable neurons. Traditionally, this tracking has been performed manually, a process that is labor-intensive, time-consuming, and potentially inconsistent due to variability in counting criteria. This variability often necessitates the re-counting of the same data set by different individuals, further adding to the complexity of the process.

Due to the limitations of these traditional techniques, the use of Deep Learning methods for neuronal viability analysis is gaining significant attention. However, the application of intelligent systems in our specific case is difficult because of several factors, necessitating a comparison of different Deep Learning-based techniques. The challenges and peculiarities of the neuron cultures include:

1. Variability of elements: Our cultures contain dead neurons, glial cells, neurites, and other bodies.
2. Neuronal mobility and changes: Neurons move and change in appearance, size, and coloration.
3. Development-induced changes: As the culture develops, neurites grow and invade the background of the images. Meanwhile, neurons tend to stabilize, and other bodies appear in the images.
4. Neuronal superposition: Neurons may overlap with each other, creating a new element that is difficult to recognize.

Due to these peculiarities, an image of a culture can be completely different from another obtained just a few

¹This research has been partially funded by the Instituto Universitario de Tecnología Industrial de Asturias (IUTA), grant SV-22-GIJÓN-1-22; by the Agencia Estatal de Investigación, grant MCI-21-PID2020-119087RB-I00, and GRUPIN SV-PA-21-AYUD/2021/52132; by the Spanish Ministry of Economics and Industry, grant PID2020-112726RB-I00; by the Ministry of Science and Innovation under CERVERA Excellence Network project CER-20211003 (IBERUS) and Missions Science and Innovation project MIG-20211008 (INMERBOT); and by Principado de Asturias, grant SV-PA-21-AYUD/2021/50994.

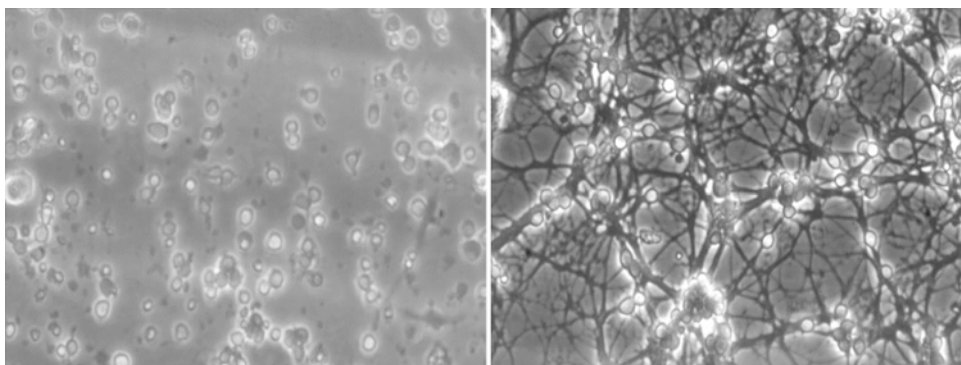


Figure 1. Comparison between the same area of the same culture, in an early stage of the culture (left) and in a later stage (right).

hours later as is shown in Figure 1. Therefore, it is necessary to compare different Computer Vision techniques to determine which is best suited to such a dynamic and complex environment.

Primarily, there are two Deep Learning-based techniques that can assist us in recognizing neurons: Object Detection and Instance Segmentation [6]. Object Detection identifies the presence and location of objects within an image, while Instance Segmentation not only identifies objects but also delineates their boundaries [6], providing morphological information about the neurons. Although Instance Segmentation could provide additional morphological insights, in this study, our primary aim is to determine which of these two techniques is better suited for recognizing neurons in our cultures.

The remainder of this paper is organized as follows: In Section 2, "Materials", we elucidate the instrumental resources employed in the development of the project. Section 3, "Methods", describes the techniques subjected to testing during our experimentation. The paper culminates with Section 4, "Experimentation", where we present our experimental findings, and Section 5, "Conclusions", where we summarize our insights and the implications of our work.

2. Materials

For the observation of neurons cultures, a phase-contrast microscope was employed, specifically the Olympus IMT-2 model. This microscope allows for the qualitative detection of early and late signs of toxicity, including changes in morphology, and neurite outgrowth and integrity. To capture the images, a photographic camera was required. The device used was a black and white video camera, manufactured by Sony, model SPT-M308CE. This camera facilitated the capture of high-quality images of the neurons cultures, essential for the subsequent analysis.

3. Methods

Primarily, there are two Deep Learning-based techniques that can assist us in recognizing neurons: Object Detection and Instance Segmentation. For the task of Object Detection, we have employed the YOLO v5

model, which has demonstrated superior performance in various applications [7]. On the other hand, for the task of Instance Segmentation, we have opted for the YOLO v7 model. This model has shown promising results in the accurate segmentation of complex objects such as trees, even in the presence of class imbalance in the dataset [8]. While Instance Segmentation could potentially provide additional morphological information about the neurons, our primary focus in this study is to ascertain which of these two techniques is better suited for neuron recognition in our cultures.

3.1. YOLO v5. Object Detection

YOLO v5, standing for You Only Look Once version 5, is a real-time Object Detection algorithm that has gained popularity due to its efficiency and accuracy. In the context of our study, this algorithm is particularly relevant as it provides a robust solution for the detection of neuronal cells in our cultures.

The architecture of YOLO v5 is composed of a backbone network for feature extraction, a neck network for feature fusion, and a head network for Object Detection. The backbone network, typically a deep convolutional network, is responsible for extracting high-level features from the input images, in our case, the images of neuronal cultures. The neck network fuses the high-level features with low-level features to enhance the feature representation, thereby capturing the intricate details of neuronal morphology. Finally, the head network uses these fused features to predict the class probabilities and bounding box coordinates for each object in the image, enabling the identification and localization of individual cells in multiple scenarios [9].

YOLO v5 also incorporates several advanced techniques to improve its performance. For instance, it uses anchor boxes to handle objects of different shapes and sizes, a feature that is crucial in dealing with the diverse morphology of neuronal cells. It applies non-maximum suppression to eliminate duplicate detections, ensuring that each neuron is only counted once. Furthermore, it employs a multi-scale prediction strategy to detect objects at different scales, accommodating for the varying sizes of neurons. Lastly, it includes a loss function

that balances the trade-off between localization and classification errors, optimizing the accuracy of neuron detection.

3.2. YOLO v7. Instance Segmentation

YOLO v7, an optimized version of YOLO v5, is a state-of-the-art model for Instance Segmentation tasks. The architecture of YOLO v7 is similar to YOLO v5, but it incorporates additional features that make it more suitable for Instance Segmentation. These features include the use of anchor boxes for better object localization, and the use of a multi-scale feature pyramid network for detecting objects at different scales.

In the context of our study, YOLO v7 can be used to segment individual neurons in our cultures, providing not only the location of the neurons but also their shape and size. This is particularly useful for studying the morphology of neurons, which can provide valuable insights into their health and function.

While there is a growing body of literature on the use of deep learning for cell segmentation, the application of YOLO v7 for neurons segmentation is still relatively unexplored.

The use of YOLO v7 for neurons segmentation in our study could therefore represent a novel and valuable contribution to the field. By comparing the performance of YOLO v5 for Object Detection and YOLO v7 for Instance Segmentation, we aim to identify the most suitable approach for analyzing neurons cultures.

3.3. Metrics

In order to compare the performance of Object Detection and Instance Segmentation models, it is essential to consider metrics at the bounding box level for the latter. This is due to the fact that Instance Segmentation not only identifies the presence of an object, but also delineates its precise spatial extent within the image [10].

The metrics that will be utilized in this study are Precision, Recall, and Mean Average Precision (mAP) at an Intersection over Union (IoU) threshold of 0.5. Precision is the proportion of true positive detections among all positive detections, while Recall is the proportion of true positive detections among all actual positives in the image. The mAP, on the other hand, provides a single score that encapsulates the precision-recall trade-off across different IoU thresholds, and is particularly useful for comparing the performance of different models [11].

A Precision-Recall curve is a suitable method for comparing two models of this kind. It plots the relationship between Precision and Recall for different threshold values, providing a comprehensive view of the model's performance across all possible classification thresholds. This allows for a more nuanced understanding of the model's performance, beyond what a single metric can provide [12].

3.4. Dataset

In this study, we utilized 292 images derived from pri-

mary cultures of rat cerebellar neurons. Of these images, 205 were used for training, 58 for validation, and 29 for testing. These images were captured at various stages of culture evolution using a conventional phase-contrast microscope. Each of these images was meticulously annotated by subject matter experts, thereby ensuring the quality of the samples. These experts are biochemists specialized in the analysis of neuronal cultures. The annotations are in mask format, where the presence of neurons is identified in white against a black background. In order to train the Object Detection model, these masks were transformed into bounding boxes. The average number of neurons per image is 50. The correspondence between an image and its mask can be seen in Figure 2.

5. Results and discussion

In this study, we evaluated two deep learning models for cell recognition in microscopy images: an Object Detection model and an Instance Segmentation model. As summarized in Table 1, both models demonstrated high performance across all metrics. The Object Detection model achieved a precision of 0.9217, a recall of 0.8950, and a mean average precision (mAP) at an Intersection over Union (IoU) threshold of 0.5 of 0.9451. The Instance Segmentation model, on the other hand, slightly outperformed the Object Detection model with a precision of 0.9236, a recall of 0.9108, and a mAP@0.5 of 0.9559.

Model	Precision	Recall	mAP@0.5
Object Detection	0.9217	0.8950	0.9451
Instance Segmentation	0.9236	0.9108	0.9559

Table 1. Main metrics for both models

In conclusion, our findings, as illustrated in Figure 3, demonstrate that the Instance Segmentation model consistently outperforms the Object Detection model in our specific context of neuron cultures. The precision-recall curve for the Instance Segmentation model is consistently above that of the Object Detection model, indicating a higher true positive rate at various threshold settings. This suggests that the Instance Segmentation model is more effective at correctly identifying and counting neurons in our cultures, despite the unique challenges and peculiarities associated with these cultures, as detailed in the relevant sections of this paper. Therefore, our results strongly advocate for the use of Instance Segmentation over Object Detection for neuron identification and counting in microscopy images of neurons cultures. These insights not only advance our understanding of effective techniques for neuron segmentation but also provide a robust methodological framework for future research in this area.

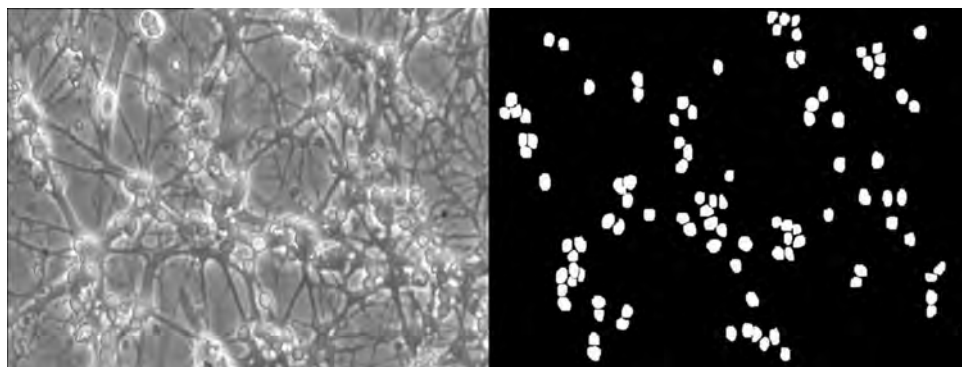


Figure 2. Comparison between an original image (left) and its segmentation mask (right).

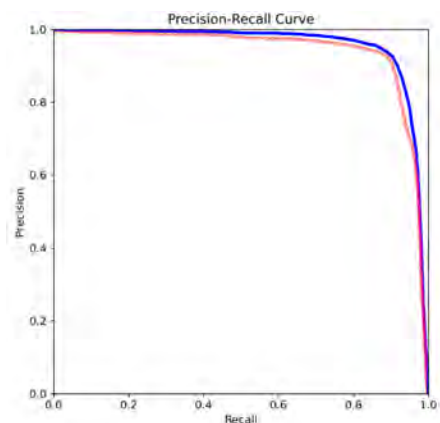


Figure 3. Precision-Recall curve comparing both models.

References

- [1] Richard Eglén, Annette Gilchrist, and Terry Reisine. An overview of drug screening using primary and embryonic stem cells. *Combinatorial Chemistry & High Throughput Screening*, 11(7):566–572, August 2008.
- [2] Levi B. Wood, Ashley R. Winslow, Elizabeth A. Proctor, Declan McGuone, Daniel A. Mordes, Matthew P. Frosch, Bradley T. Hyman, Douglas A. Lauffenburger, and Kevin M. Haigis. Identification of neurotoxic cytokines by profiling alzheimer’s disease tissues and neuron culture viability screening. *Scientific Reports*, 5(1), November 2015.
- [3] Sandra Cetin, Damijan Knez, Stanislav Gobec, Janko Kos, and Anja Pišlar. Cell models for alzheimer’s and parkinson’s disease: At the interface of biology and drug discovery. *Biomedicine & Pharmacotherapy*, 149:112924, May 2022.
- [4] Andreea Rosca, Raquel Coronel, Miryam Moreno, Rosa González, Andreea Oniga, Alberto Martín, Victoria López, María del Carmen González, and Isabel Liste. Impact of environmental neurotoxic: current methods and usefulness of human stem cells. *Heliyon*, 6(12):e05773, December 2020.
- [5] María Teresa Fernández, Vladimir Zitko, Santiago Gascón, and Antonello Novelli. The marine toxin okadaic acid is a potent neurotoxin for cultured cerebellar neurons. *Life Sciences*, 49(19):PL157–PL162, January 1991.
- [6] Renato Sortino, Daniel Magro, Giuseppe Fiameni, Eva Sciacca, Simone Riggi, Andrea DeMarco, Concetto Spampinato, Andrew M. Hopkins, Filomena Bufano, Francesco Schillirò, Cristobal Bordiu, and Carmelo Pino. Radio astronomical images object detection and segmentation: a benchmark on deep learning methods. *Experimental Astronomy*, May 2023.
- [7] N. Salahin, A. Rahman, M. Islam, M. Islam, and M. Islam. Yolo v5: An improved real-time object detection algorithm. *IOP Conference Series: Materials Science and Engineering*, 1234:012034, 2023.
- [8] Y. Cao, J. Wang, Y. Zhang, and Z. Li. Yolo v7: A new method for instance segmentation in high-resolution images. *Sensors*, 23:5138, 2023.
- [9] Shakir M. Abas, A. Abdulazeez, and D. Zeebaree. A yolo and convolutional neural network for the detection and classification of leukocytes in leukemia. *International Journal of Electrical and Computer Engineering*, 25(1):200–213, 2022.
- [10] Alberto Garcia-Garcia, Sergio Orts-Escolano, Sergiu Oprea, Victor Villena-Martinez, and Jose Garcia-Rodriguez. A review on deep learning techniques applied to semantic segmentation. *Pattern Recognition Letters*, 86:60–66, 2017.
- [11] Mark Everingham, Luc Van Gool, Christopher K.I. Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, 88:303–338, 6 2010.
- [12] Jesse Davis and Mark Goadrich. The relationship between precision-recall and roc curves. In *Proceedings of the 23rd international conference on Machine learning*, pages 233–240, 2006.

Un enfoque explicativo para el diagnóstico de la degeneración macular asociada a la edad mediante técnicas de deep learning

M. Herrero-Tudela¹, R. Romero-Oraá^{1,2}, R. Hornero Sánchez^{1,2}, Gonzalo C. Gutiérrez-Tobal^{1,2}, M. I. López Gálvez^{1,2}, M. García^{1,2}

¹ Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España,

{maria.herrero.tudela, roberto.romero, roberto.hornero, gonzalocesar.gutierrez, maria.garcia.gadanon}@uva.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

Resumen

La degeneración macular asociada a la edad (DMAE) es un trastorno que afecta a la mácula, una zona de la retina clave para la agudeza visual. La DMAE es una de las causas más frecuentes de ceguera en personas mayores de 60 años en los países desarrollados. Aunque se han propuesto tratamientos que frenan su desarrollo, su eficacia disminuye significativamente en las fases avanzadas. Por ello, son importantes los programas de cribado para la detección precoz. Sin embargo, implementar tales programas para enfermedades como la DMAE suele ser inviable debido a la extensa población de riesgo y a la necesidad de que los profesionales revisen y localicen manualmente lesiones en las retinografías. En este sentido, la principal contribución de este trabajo fue la aplicación de Explainable Artificial Intelligence como ayuda al diagnóstico de la DMAE. Para ello, desarrollamos un modelo de deep learning basado en RegNetY-320 con el que obtuvimos una precisión, sensibilidad y especificidad de 86.5%, 85.21% y 91.01%, respectivamente sobre la base de datos ADAM (1200 imágenes). Mediante la técnica integrated gradients, identificamos las áreas específicas en las retinografías que influyen en las decisiones del modelo. Esto permite la detección y localización de las lesiones asociadas a la DMAE, proporcionando una herramienta para la interpretación clínica y mejorando la confianza en el diagnóstico. Los resultados obtenidos indican que el método propuesto podría ser útil en los sistemas de ayuda al diagnóstico de la DMAE.

1. Introducción

La Degeneración Macular Asociada a la Edad (DMAE) se ha convertido en una causa principal de pérdida irreversible de la visión, especialmente en países desarrollados [1], [2]. Afecta principalmente a personas mayores de 60 años, y se prevé un aumento significativo de su prevalencia debido al envejecimiento demográfico, proyectándose más de 288 millones de casos para el año 2040 [2].

La DMAE se clasifica clínicamente en seca y húmeda [3]. La DMAE seca se caracteriza por la presencia de drusas de tamaño medio y cambios pigmentarios en la retina, mientras que la DMAE húmeda se caracteriza por la presencia de neovascularización y atrofia [3]. La retinografía y la tomografía de coherencia óptica son los exámenes auxiliares más utilizados en oftalmología. La fotografía del fondo de ojo es la prueba más económica y necesaria en la DMAE, ya que puede identificar de manera intuitiva lesiones y diagnosticar la enfermedad [3]. A pesar de su conveniencia, la implantación de programas de cribado de la DMAE a gran escala suele ser inviable, ya que la población de riesgo es grande y el análisis de las

imágenes de fondo de ojo es muy complicado. Todo ello motiva la investigación de métodos de diagnóstico automáticos [4]–[8]. En la literatura, el enfoque predominante para el diagnóstico automatizado de la DMAE es entrenar un clasificador de aprendizaje automático para discriminar entre dos clases: DMAE y no DMAE. En algunos trabajos se han empleado métodos clásicos basados en redes neuronales totalmente conectadas [4] o máquinas de vectores de soporte [5]. Por el contrario, la mayoría de los trabajos recientes se basan en redes neuronales convolucionales (*Convolutional Neural Networks*, CNNs) [6]–[8]. Sin embargo, hasta donde sabemos, ninguno de ellos incorpora mecanismos de explicabilidad para ayudar a los expertos a comprender mejor las predicciones de los modelos. Este problema es particularmente relevante, ya que la ausencia de tales mecanismos limita la aplicación de enfoques automáticos en escenarios reales [9]. En tareas de ayuda al diagnóstico, esta necesidad de explicabilidad es aún más pronunciada, ya que la decisión del modelo puede tener un impacto directo en la salud del paciente [9].

Por ello, en este estudio, proponemos un enfoque de aprendizaje profundo que busca mejorar la detección automatizada de la DMAE y proporcionar explicaciones detalladas basadas en la identificación de las lesiones asociadas. El objetivo es desarrollar un sistema que combine la detección precisa de la DMAE en imágenes de retinografía con la capacidad de explicar las lesiones relacionadas, mejorando la precisión diagnóstica y la comprensión clínica de esta enfermedad ocular. Utilizando la técnica *integrated gradients*, generamos mapas de atribución que permitieron evaluar el impacto de cada píxel en la clasificación final de cada imagen y destacar las zonas más relevantes para la clasificación, contribuyendo así al campo de *Explainable Artificial Intelligence* (XAI).

2. Base de datos

En este estudio se utilizó la base de datos pública ADAM. Este conjunto de datos consiste en 1,200 imágenes de fondo de ojo almacenadas en formato JPEG, con 8 bits por canal de color. Estas imágenes fueron proporcionadas por el Centro Oftalmológico Zhongshan de la Universidad Sun Yat-sen en China [7]. Las imágenes de fondo de ojo se capturaron con un retinógrafo Zeiss Visucam 500, con una resolución de 2,124 × 2,056 píxeles, y con un retinógrafo Canon CR-2, con una resolución de 1,444 × 1,444 píxeles.

El centro del campo de visión de las fotografías se ubicó en el disco óptico, la mácula o el punto medio entre el disco óptico y la mácula [7].

Todas las imágenes fueron etiquetadas indicando la presencia o no de DMAE. Las imágenes etiquetadas como DMAE abarcan casos de DMAE en sus etapas temprana, intermedia o avanzada, mientras que la categoría sin DMAE incluye muestras sin esta enfermedad pero que podrían presentar otros trastornos de retina [7].

Los creadores de la base de datos dividieron el conjunto de datos resultante en tres partes: un conjunto de entrenamiento (400 imágenes), un conjunto de validación (400 imágenes) y un conjunto de test (400 imágenes). En este estudio, se ha utilizado esta división.

3. Métodos

El método propuesto parte de una etapa de preprocesado para normalizar las imágenes de entrada. Después, se desarrolló un modelo CNN para detectar la DMAE haciendo uso de técnicas como *data augmentation*, *transfer learning* y *fine-tuning*. Finalmente, se utilizó la técnica *integrated gradients* para visualizar las zonas de la imagen que influyen en las predicciones del modelo.

3.1. Preprocesado

Como entrada al modelo, se utilizaron versiones preprocesadas de las imágenes originales. En primer lugar, se redimensionaron todas las imágenes a 512 x 512 píxeles con el objetivo de normalizar todas las imágenes de entrada [10]. A continuación, los valores de los píxeles en cada canal de color de las imágenes se normalizaron en el intervalo [0,1]. Esto asegura que los valores de los píxeles estén en una escala consistente, facilitando el proceso de entrenamiento del modelo [11].

3.2. Data augmentation

Para aumentar el número de imágenes con las que entrenar la CNN, se empleó la técnica *data augmentation* en tiempo real. Esta técnica permite generar nuevas imágenes aleatorias exclusivas para cada época de entrenamiento. Para ello, se aplicaron transformaciones simples sobre las 400 imágenes del conjunto de entrenamiento: rotaciones aleatorias en el rango [-90, +90] grados y volteos horizontales y verticales aleatorios [12].

3.3. Transfer learning y fine-tuning

El objetivo de aplicar *transfer learning* y *fine-tuning* es aprovechar el conocimiento adquirido por modelos previamente entrenados en tareas relacionadas. Esto permite acelerar el entrenamiento, mejorar el rendimiento y reducir los datos necesarios para lograr un modelo efectivo y preciso [13], [14].

La técnica *transfer learning* consiste en utilizar una red neuronal pre-entrenada para resolver un problema similar al que la red fue diseñada y entrenada inicialmente para resolver [13]. Para aplicar esta técnica, se utilizó un modelo base pre-entrenado con las imágenes del proyecto ImageNet, una base de datos que se compone de más de 14 millones de imágenes pertenecientes a más de 20,000 clases distintas [15].

A continuación, se aplicó la técnica *fine-tuning*. En los modelos basados en redes profundas, las primeras capas capturan características generales comunes a todas las imágenes, como bordes o patrones, mientras que las últimas capas identifican las características más relevantes para el problema a resolver [26]. La técnica de *fine-tuning* permite reentrenar todas o algunas de las capas finales de la CNN utilizando imágenes específicas del problema a resolver. De esta manera, el método resultante es más preciso en comparación con los modelos entrenados desde cero, ya que los pesos están mejor adaptados al objetivo del problema [14]. Si los datos son similares a los utilizados para pre-entrenar el modelo original, el *fine-tuning* de las últimas capas puede ser suficiente. Sin embargo, en este trabajo, se reentrenaron todas las capas del modelo ya que las imágenes tienen particularidades significativas. Reentrenar todas las capas permite una adaptación más efectiva, mejorando el rendimiento del modelo [14].

3.4. Arquitectura CNN

Las CNN son redes neuronales capaces de extraer características representativas de las imágenes de manera óptima y están formadas por distintos tipos de capas llamadas convolucionales, de *pooling* y *fully-connected*, entre otras [13]. En este trabajo se desarrolló una arquitectura CNN utilizando como modelo base la arquitectura RegNetY302 [16]. RegNet es un enfoque de diseño de redes que busca crear redes rápidas, simples y eficientes mediante una parametrización lineal. Estas redes superan a modelos populares como EfficientNet en rendimiento y velocidad en GPU. [16].

Para adaptar la arquitectura al problema de clasificación binaria de la DMAE, se han añadido una capa *average pooling*, una capa de *dropout* con un factor de 0.5, una capa densa de 2048 neuronas, otra capa de *dropout* con factor de 0.5 y una capa densa de 2 neuronas [17]. El número de neuronas de la última capa se corresponde con el número de clases que queremos discriminar: paciente sano y paciente con DMAE. En la primera capa densa se utilizó una función de activación tipo ReLU, mientras que en la última capa densa se utilizó la función de activación *softmax*. La función *softmax* devuelve una distribución de probabilidad sobre las clases. En este caso, la distribución de probabilidad se utiliza para indicar la probabilidad de que la imagen de entrada pertenezca a la clase sana o con DMAE [17]. En la fase de *fine-tuning*, se empleó la técnica *early-stopping* para minimizar el sobre-entrenamiento de la red [13]. De esta manera, el proceso de entrenamiento se detuvo automáticamente cuando la pérdida de validación no mejoraba durante 5 épocas. Se aplicó la entropía cruzada categórica como función de pérdida y Adam como algoritmo de optimización [17]. Asimismo, en épocas avanzadas, se redujo la tasa de aprendizaje en un factor de 10 cada vez que el error de validación alcanzase un mínimo y se mantuviese constante [13].

3.5. Integrated Gradients

Los modelos basados en *deep learning* son vistos como cajas negras que no explican cómo se realizan las predicciones. Esto ha llevado a una aceptación relativamente baja de estos modelos entre los profesionales

de la salud. Por lo tanto, se requiere una explicación e interpretación del funcionamiento del modelo en aplicaciones clínicas. En este sentido, en este trabajo proponemos un enfoque XAI que emplea el método de *integrated gradients* [18].

El método *integrated gradients* se basa en calcular la importancia de cada píxel en una imagen al sumar los gradientes de una serie de imágenes interpoladas entre una imagen de referencia y la imagen original [18]. La imagen de referencia representa la ausencia de las características que queremos analizar. Para elegir esta imagen de referencia nos basamos en estudios previos relacionados con el análisis de retinografías en retinopatía diabética (RD) puesto que, hasta donde tenemos conocimiento, no hay estudios que empleen este enfoque en el diagnóstico de DMAE. En los estudios de RD, es común utilizar una imagen negra como referencia [19]. En este estudio se utilizó la misma aproximación puesto que las imágenes de partida (retinografías) son iguales en ambos casos. Para cada retinografía, se generó una serie de 30 imágenes interpoladas entre la imagen de referencia negra y la imagen original. Estas imágenes son transiciones desde la imagen negra hasta la imagen de entrada original y representan diferentes niveles de importancia de las características [19].

Mediante *integrated gradients* obtenemos mapas de atención que indican la contribución de cada píxel en la imagen original al resultado del diagnóstico. Estas contribuciones se miden en comparación con la imagen de referencia, que no aporta información relevante al modelo [18]. Esta técnica nos permite comprender cómo el modelo procesa la información y qué características de la imagen considera más relevantes durante el proceso de predicción.

4. Resultados

El entrenamiento de la CNN se llevó a cabo con 400 imágenes de la base de datos ADAM. Las 400 imágenes del conjunto de validación permitieron monitorizar el error de aprendizaje en la aplicación de la técnica *early-stopping*. El método se evaluó sobre un conjunto de test de 400 imágenes de la misma base de datos. La Tabla 1 recoge los resultados en términos de precisión (PR), sensibilidad (SE), especificidad (ES), F1-score y área bajo la curva ROC (AUC). Además, en la Figura 1 se presentan varios ejemplos de mapas obtenidos con el método *integrated gradients* donde se puede apreciar la identificación de lesiones en la clase con DMAE.

5. Discusión

En este estudio, se presenta una herramienta de explicabilidad basada en *deep learning* y el método *integrated gradients* para mejorar el diagnóstico de la DMAE. Es importante destacar que, hasta donde sabemos, este enfoque de atención no se ha empleado en el contexto del diagnóstico de la DMAE. Este método permite determinar si los resultados del modelo son fiables al resaltar las regiones importantes que conducen a la

PR (%)	SE (%)	ES (%)	AUC
86.5	91.01	85.21	0.95

Tabla 2. Resultados sobre el conjunto de test.

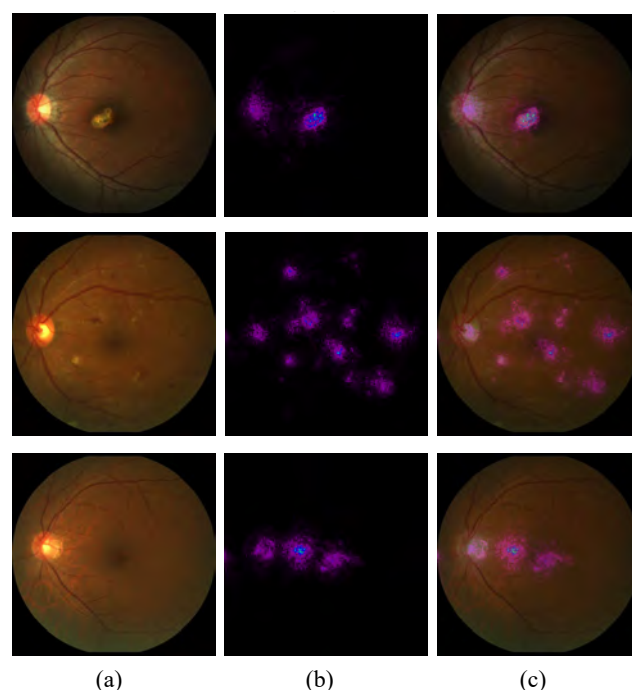


Figura 1. Ejemplos de mapas generados con *integrated gradients*: (a) imágenes originales, (b) mapas generados, (c) superposición de mapas generados e imágenes originales.

decisión del modelo. El método se ha desarrollado utilizando exclusivamente la base de datos ADAM. Los resultados de la Tabla 1 indican que el método propuesto ofrece un alto rendimiento para la detección de la DMAE. En la Tabla 2 se muestra que nuestros resultados son comparables a los obtenidos en estudios previos.

La contribución más importante, en el contexto de XAI, son los mapas de atención generados con *integrated gradients*, que se validaron de forma cualitativa. Estos mapas desempeñan un papel fundamental al representar la relevancia de cada región de una imagen en el proceso de toma de decisiones del modelo. En la Figura 1 se presentan varios ejemplos de estos mapas sobre retinografías. En la primera fila, se puede apreciar que el modelo identifica de manera precisa las drusas cercanas al área de la mácula, lo cual constituye un indicativo relevante de la presencia de DMAE. La segunda fila representa un caso en el que una imagen fue incorrectamente clasificada como DMAE, posiblemente debido a otras lesiones. La última fila corresponde a un caso en el que una imagen se etiquetó incorrectamente como no DMAE, a pesar de la presencia de la enfermedad, lo que podría deberse a otros datos clínicos. El método *integrated gradients* proporciona a los

Estudio	Conjunto de test (n)	AUC
VUNO EYE TEAM [10]	ADAM (n=400)	0.97
ForbiddenFruit [10]	ADAM (n=400)	0.96
ADAM-TEAM [10]	ADAM (n=400)	0.93
XxlzT [10]	ADAM (n=400)	0.91
TeamTiger [10]	ADAM (n=400)	0.91
Airamatrix [10]	ADAM (n=400)	0.88
Método propuesto	ADAM (n=400)	0.95

Tabla 1. Comparación de los resultados con otros estudios. n – Número de imágenes en el conjunto.

clínicos una mayor confianza en el diagnóstico al marcar claramente las lesiones asociadas a la enfermedad.

Este estudio presenta ciertas limitaciones que también es necesario mencionar. En primer lugar, el algoritmo ha sido desarrollado y validado únicamente en una base de datos. Sería deseable evaluar la capacidad de generalización de manera más exhaustiva, empleando un mayor número de imágenes, captadas con diferentes retinógrafos y procedentes de diversas localizaciones. Además, el método de visualización empleado representa una primera aproximación en términos de la interpretación y explicación de las predicciones del modelo. Aunque *integrated gradients* ofrece una visión de las áreas influyentes en las decisiones del modelo, existen otros mecanismos y enfoques de visualización basados en XAI que podrían ser explorados en investigaciones futuras. Abordar otros métodos podría ofrecer una comprensión más completa de cómo las CNNs toman decisiones.

6. Conclusiones

En este trabajo se ha propuesto un método automático basado en técnicas de *deep learning* para la detección de la DMAE que, además, incorpora un mecanismo de explicación visual para ayudar a los expertos a comprender mejor las predicciones del modelo. Los resultados obtenidos indican que método propuesto podría ser de utilidad en un entorno clínico como primera etapa en los sistemas de ayuda al diagnóstico de la DMAE.

Agradecimientos

Esta investigación se ha desarrollado en el marco de las ayudas TED2021-131913B-I00, PID2020-115468RB-I00 y PGC2018-098214-A-I00 financiadas por el 'Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación/10.13039/501100011033/' y el Fondo Europeo de Desarrollo Regional (FEDER). Una forma de hacer Europa; y por el 'CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)' a través del 'Instituto de Salud Carlos III' cofinanciado con fondos FEDER. M. Herrero Tudela cuenta con un contrato predoctoral de la Universidad de Valladolid.

Referencias

- [1] C. J. Flaxel *et al.*, “Age-Related Macular Degeneration Preferred Practice Pattern®,” *Ophthalmology*, vol. 127, no. 1, pp. P1–P65, Jan. 2020, doi: 10.1016/J.OPHTHA.2019.09.024.
- [2] W. L. Wong *et al.*, “Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis,” *Lancet Glob Health*, vol. 2, no. 2, pp. e106–e116, Feb. 2014, doi: 10.1016/S2214-109X(13)70145-1.
- [3] L. S. Lim, P. Mitchell, J. M. Seddon, F. G. Holz, and T. Y. Wong, “Age-related macular degeneration,” *The Lancet*, vol. 379, no. 9827, pp. 1728–1738, 2012.
- [4] E. Pead *et al.*, “Automated detection of age-related macular degeneration in color fundus photography: a systematic review,” *Surv Ophthalmol*, vol. 64, no. 4, pp. 498–511, Jul. 2019, doi: 10.1016/j.survophthal.2019.02.003.
- [5] M. R. K. Mookiah *et al.*, “Automated detection of age-related macular degeneration using empirical mode decomposition,” *Knowl Based Syst*, vol. 89, pp. 654–668, Nov. 2015, doi: 10.1016/J.KNOSYS.2015.09.012.
- [6] C. González-Gonzalo *et al.*, “Evaluation of a deep learning system for the joint automated detection of diabetic retinopathy and age-related macular degeneration,” *Acta Ophthalmol*, vol. 98, no. 4, pp. 368–377, Jun. 2020, doi: 10.1111/AOS.14306.
- [7] H. Fang *et al.*, “ADAM Challenge: Detecting Age-related Macular Degeneration from Fundus Images,” Feb. 2022, doi: 10.1109/TMI.2022.3172773.
- [8] X. Li, M. Jia, M. T. Islam, L. Yu, and L. Xing, “Self-Supervised Feature Learning via Exploiting Multi-Modal Data for Retinal Disease Diagnosis,” *IEEE Trans Med Imaging*, vol. 39, no. 12, pp. 4023–4033, Dec. 2020, doi: 10.1109/TMI.2020.3008871.
- [9] B. H. M. van der Velden, H. J. Kuijf, K. G. A. Gilhuijs, and M. A. Viergever, “Explainable artificial intelligence (XAI) in deep learning-based medical image analysis,” *Med Image Anal*, vol. 79, p. 102470, Jul. 2022, doi: 10.1016/J.MEDIA.2022.102470.
- [10] H. Fang *et al.*, “ADAM Challenge: Detecting Age-Related Macular Degeneration From Fundus Images,” *IEEE Trans Med Imaging*, vol. 41, no. 10, pp. 2828–2847, Oct. 2022, doi: 10.1109/TMI.2022.3172773.
- [11] S. Ioffe and C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.” 2015.
- [12] L. Perez and J. Wang, “The Effectiveness of Data Augmentation in Image Classification using Deep Learning,” Dec. 2017, Accessed: Jul. 12, 2023. [Online]. Available: <https://arxiv.org/abs/1712.04621v1>
- [13] L. Shao, F. Zhu, and X. Li, “Transfer Learning for Visual Categorization: A Survey,” *IEEE Trans Neural Netw Learn Syst*, vol. 26, no. 5, pp. 1019–1034, 2015, doi: 10.1109/TNNLS.2014.2330900.
- [14] C. Käding, E. Rodner, A. Freytag, and J. Denzler, “Fine-Tuning Deep Neural Networks in Continuous Learning Scenarios,” Jul. 2017, pp. 588–605. doi: 10.1007/978-3-319-54526-4_43.
- [15] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255. doi: 10.1109/CVPR.2009.5206848.
- [16] I. Radosavovic, R. P. Kosaraju, R. Girshick, K. He, and P. Dollár, “Designing Network Design Spaces,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 10425–10433, Mar. 2020, doi: 10.1109/CVPR42600.2020.01044.
- [17] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [18] M. Sundararajan, A. Taly, and Q. Yan, “Axiomatic Attribution for Deep Networks,” in *Proceedings of the 34th International Conference on Machine Learning*, D. Precup and Y. W. Teh, Eds., in Proceedings of Machine Learning Research, vol. 70. PMLR, Jul. 2017, pp. 3319–3328.
- [19] R. Sayres *et al.*, “Using a Deep Learning Algorithm and Integrated Gradients Explanation to Assist Grading for Diabetic Retinopathy,” *Ophthalmology*, vol. 126, no. 4, pp. 552–564, 2019, doi: <https://doi.org/10.1016/j.ophtha.2018.11.016>.

Aplicación de MobileNet para el diagnóstico temprano del glaucoma: Un enfoque binocular

O. Kovalyk Borodyak ¹, J. Morales Sánchez ¹, R. Verdú Monedero ¹, I. Sellés Navarro ², J.L. Sancho Gómez ¹

¹ Universidad Politécnica de Cartagena, Cartagena, España,
olekesandr.kovalyk@edu.upct.es, {juan.morales, rafael.verdu, josel.sancho}@upct.es

² Hospital Universitario Reina Sofía, Murcia, España, inmasell@um.es

Resumen

Este estudio presenta los resultados preliminares de la detección temprana del glaucoma utilizando la red neuronal MobileNet. MobileNet es una red eficiente en términos del uso de memoria y recursos computacionales. En este trabajo se comparan dos enfoques: monocular, que considera un ojo, y binocular, que integra información de ambos ojos. La investigación utiliza la base de datos PAPILA, con retinografías y datos clínicos de 244 pacientes. Los resultados muestran que el enfoque binocular supera al monocular, con un aumento del 9 % en el área bajo la curva (AUC) y un 15 % en sensibilidad a una especificidad del 90 %. Estos resultados indican una posible ventaja del uso de ambos ojos de un paciente para la mejora del poder de diagnóstico de las redes neuronales.

1. Introducción

El glaucoma es la principal causa de ceguera irreversible en todo el mundo [1]. Esta patología suele aparecer en edades avanzadas, generalmente causada por alta presión intraocular, afecta la cabeza del nervio óptico (CNO) y provoca una pérdida progresiva del campo visual. Se sabe que los indicadores del glaucoma aparecen años antes que las lesiones que causan pérdida de campo visual. En este contexto, la detección precoz es de vital importancia para minimizar los daños producidos por la enfermedad y los costes sanitarios [2]. Las principales pruebas que se realizan para detección del glaucoma son: la tomografía de coherencia óptica, la tonometría, la campimetría y la retinografía. La retinografía (Figura 1) es una imagen del fondo ocular, normalmente centrada en el disco óptico, que sirve para el propósito de estudiar la morfología de la CNO. Cabe destacar que la retinografía es una modalidad de imagen médica mínimamente invasiva.

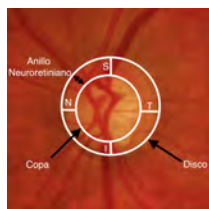


Figura 1: Datos morfológicos de una retinografía. Se han marcado las longitudes Inferior (I), Superior (S), Nasal (N) y temporal (T).

Con el avance de la tecnología en medicina, y particularmente con la irrupción de las técnicas de procesamiento de imágenes y la inteligencia artificial, las redes neuronales convolucionales han ganado un protagonismo consi-

derable en el diagnóstico de diversas patologías, incluido el glaucoma. Estas herramientas, se caracterizan por su capacidad para analizar grandes volúmenes de imágenes médicas, ofreciendo una respuesta en cuestión de segundos.

Por otro lado, la hipótesis sobre la asimetría entre las características anatómicas de la CNO de ambos ojos ha sido recientemente reconsiderada como un indicador para el diagnóstico temprano del glaucoma [3, 4, 5, 6]. Discrepancias significativas entre ambos ojos de un mismo paciente pueden ser indicativas de alteraciones patológicas. En este sentido, las técnicas avanzadas de aprendizaje automático podrían ofrecer la posibilidad de aprovechar estas diferencias para la detección temprana del glaucoma.

Dentro de este marco de trabajo, proponemos un método basado en la arquitectura MobileNet [7] para el diagnóstico automático del glaucoma. Se ha escogido esta red en particular dado que puede ser implementada en dispositivos compactos, como retinógrafos portátiles, simultáneamente, permitiría analizar la morfología del nervio óptico, eludiendo así las demoras que pudieran surgir debido a la necesidad de acceder a un retinógrafo tradicional. Esto facilitaría a los profesionales de la atención primaria realizar evaluaciones inmediatas para descartar la presencia de glaucoma. El modelo se evaluará considerando dos enfoques: el enfoque monocular y el enfoque binocular. En el enfoque binocular información de ambos ojos de forma simultánea, accediendo así a características a las que de otro modo no son accesibles. Los resultados del enfoque binocular se compararán con el modelo monocular, el cual predice el glaucoma a partir de una única imagen. Con todo ello se pretende diseñar una herramienta efectiva y eficiente para la detección temprana del glaucoma, mejorando así el pronóstico de los pacientes afectados por esta enfermedad.

2. Materiales

2.1. Conjunto de datos

En este trabajo, emplearemos la base de datos PAPILA [8], la cual es de acceso abierto, proporciona retinografías de ambos ojos por paciente, segmentaciones realizadas por dos expertos y datos clínicos. Elegimos PAPILA por diversos motivos: al ser de acceso abierto facilita la replicabilidad de resultados, ofrece información de ambos ojos de cada paciente, está estructurada como una muestra representativa de la población e in-

cluye el grado de progresión del glaucoma. PAPILA comprende datos de ambos ojos de 244 pacientes, es decir un total de 488 retinografías. Éstas se distribuyen en 333 etiquetadas como sanas, 87 diagnosticadas con glaucoma y 68 consideradas sospechosas de glaucoma.

2.2 Preparación del conjunto de datos

En primer lugar, se excluyeron de las pruebas los pacientes considerados por los expertos como sospechosos de padecer glaucoma. Los paciente sospechosos son aquellos cuyas pruebas no son concluyentes como para diagnosticarlos. En pruebas anteriores se ha observado que la etiqueta sospechosa no es una etiqueta intermedia, si no que algunos de los sospechosos realmente tienen glaucoma y otros no. Esto provocaba un impacto negativo en el entrenamiento del algoritmo. Por ello las etiquetas de pacientes sospechosos del glaucoma se han descartado para este estudio.

Debido al tamaño relativamente pequeño del conjunto de datos, se hace uso de técnicas de validación cruzada, concretamente el método *k-fold*. Este método consiste en dividir el conjunto de datos en *k* subconjuntos llamados *folds*. En nuestra investigación se ha considerado $k = 5$. Cada uno de estos *folds* sirve como conjunto de validación, mientras que el resto se emplea para el entrenamiento. En este trabajo en concreto se ha considerado dos particiones *k-fold*, una para el ajuste de hiperparámetros y otra para la presentación de resultados finales.

Para la segmentación de datos, adoptamos criterios específicos: la división se basa en pacientes y no en ojos individuales y se tiene en cuenta el nivel de progresión del glaucoma. Identificamos tres niveles de progresión: glaucoma temprano, glaucoma medio y glaucoma severo. En el modelo monocular, se descarta aleatoriamente la imagen de un ojo por paciente. Así, tanto en el modelo binocular como en el monocular contamos con 42 instancias para test y 168 para entrenamiento.

Antes del entrenamiento, las imágenes se normalizaron en un rango de 0 a 1 dividiendo las imágenes entre 255. Se ha observado en experimentos anteriores que los tiempos de entrenamiento se reducen de manera significativa al reducir el tamaño de las imágenes, sin que esto afecte de manera significativa a los resultados. Por lo que para este estudio las imágenes se han redimensionado a $224 \times 224 \times 3$. Durante el entrenamiento, se aplicaron técnicas de aumento de datos, con el fin de conseguir un modelo más robusto y obtener mejores resultados. Los aumentos de datos se han realizado de manera aleatoria (50% por para aplicar cada transformación) durante el entrenamiento y son los siguientes: volteo horizontal y vertical, rotación de 0 a 180 grados, transformación afín, transformación de la perspectiva, desenfoque Gaussiano y la alteración de color.

3. Métodos

3.1. MobileNet monocular

MobileNet[7] es una arquitectura de red neuronal convolucional diseñada para ser eficiente y ligera. Para con-

seguir esto los autores de MobileNet hacen uso de: convoluciones separables por profundidad, multiplicadores de ancho y resolución y eficiencia en parámetros. En las versiones *v2* y *v3* de MobileNet se introducen funciones de activación más eficientes tales como *ReLU6* y *h-swish*

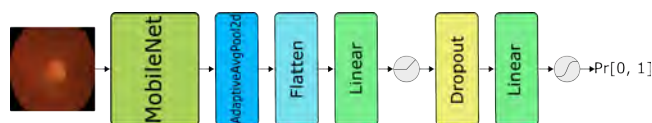


Figura 2: Esquema general del método con el enfoque monocular.

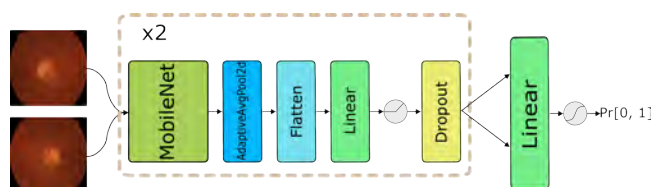


Figura 3: Esquema general del método con el enfoque binocular.

Arquitecturas como VGG, ResNet u otras arquitecturas que pueden ofrecer un rendimiento ligeramente superior en términos de precisión, pero también son mucho más pesadas en cuanto al número de parámetros y operaciones de cómputo. MobileNet, por otro lado, está diseñada con un equilibrio entre precisión y eficiencia, abriendo así la posibilidad de su utilización en dispositivos con recursos limitados, tales como smartphones, dispositivos móviles o dispositivos IoT.

En la Figura 2 podemos ver un esquema general del algoritmo en modo monocular. En la entrada tenemos una retinografía que tras el preprocesado mencionado anteriormente se le suministra a MobileNet. La red extrae las características obteniendo así un espacio latente. Este espacio latente se filtra pasándolo por una capa *Adaptive Average Pooling* y posteriormente se aplanan con una capa *flatten*, obteniendo así el vector latente. Tras ello, el vector resultante se administra a la capa lineal, a la cual le sigue de una función de activación *ReLU6*. Con el objetivo de evitar el sobreajuste se hace uso de una capa *dropout* (solo durante el entrenamiento). Finalmente, la capa de salida con activación *Softmax* devuelve la probabilidad de que una imagen tenga glaucoma o no.

3.2. MobileNet binocular

La Figura 3 muestra un esquema del algoritmo en su modo binocular. En el enfoque binocular la estructura del algoritmo es prácticamente la misma salvo por algunas modificaciones. La entrada pasa a ser dos imágenes. De estas imágenes se extrae un vector latente único que contiene información de ambos ojos. Este vector latente es el que se le administra a la capa lineal para establecer el diagnóstico.

3.3. MobileNet implementación

La implementación de la red está realizada en lenguaje Python, concretamente con la librería de aprendizaje

máquina PyTorch, de la mano de TorchVision [9]. La versión de la librería de PyTorch y TorchVision son 2.0.1 y 0.15.2 respectivamente. La versión de MobileNet que se ha utilizado es la *v3*.

3.4. Entrenamiento

Para ambos modos, binocular y monocular, la red se configuró de la siguiente manera: una tasa de aprendizaje establecida en $8,08 \times 10^{-5}$, un *weight decay* de $5,35 \times 10^{-6}$, y un tamaño de lote de 2. Se asignó un peso a cada clase basado en la frecuencia inversa de la misma. Las primeras 20 capas de la MobileNet se mantuvieron congeladas y se utilizaron pesos preentrenados de ImageNet [10]. Además, se aplicó suavizado de etiquetas (*label smoothing*) con un valor de 0,003 durante 30 épocas. La elección de los distintos valores de los hiperparámetros corresponden a las pruebas realizadas sobre un conjunto *k-fold* distinto al conjunto sobre el cual se han presentado los resultados. Los valores que se han elegido son aquellos que devolvían un menor error entre los 5 *folds*.

3.5. Hardware

Los experimentos se han realizado en un servidor de cómputo con sistema Ubuntu Linux 20.04.06 LTS, instalado sobre el siguiente hardware: procesador AMD EPIC 7643 de 48 núcleos, RAM de 256 GB y tarjeta gráfica NVIDIA 3090 de 24GB de VRAM.

4. Resultados y discusión

En esta sección, analizaremos los resultados obtenidos con el método propuesto, enfocándonos en su capacidad diagnóstica. Utilizaremos tres métricas ampliamente conocidas en el análisis de rendimiento de algoritmos de aprendizaje automático, especialmente en el ámbito del diagnóstico médico: sensibilidad, especificidad y el área bajo la curva, del acrónimo AUC en inglés. La sensibilidad computa el tanto por ciento (o tanto por uno) de pacientes que se clasifican como enfermos siendo estos realmente enfermos. La especificidad lo que computa es el tanto por ciento de los paciente diagnosticados como sanos de los que realmente lo son. En este trabajo, como en otros del mismo estilo [11], se fijará la especificidad al 80, 85 y 90 por ciento y se medirá la sensibilidad obtenida. Además, se destacará la eficiencia del método en términos de tiempo de cálculo.

4.1. Capacidad de diagnóstico

Los resultados obtenidos bajo el enfoque monocular se presentan en la Tabla 1 y la Figura 4. La Tabla 1 muestra la sensibilidad alcanzada para especificidades fijadas al 80 %, 85 % y 90 %. Esto se hace para obtener un valor en tanto por cien de enfermos que se detectan preservando las cantidades fijadas de tanto por cien de sanos detectados como tal. En lugar de promediar los resultados de cada *fold*, el total, se calcula uniendo en un único conjunto, los conjuntos (*folds*) de test de todos los 5 *folds*.

A pesar de las dificultades inherentes al conjunto de datos, que representa situaciones clínicas realistas, la sensibilidad para una especificidad del 80 % en cada pliegue se mantiene por encima del 50 %. En general, se logra una sensibilidad superior al 50 % para las tres especificidades mencionadas y un área bajo la curva (AUC) del 81 %, como se ilustra en la Figura 4. Al igual que para el total en la Tabla 1 y 2, el cálculo de la curva AUC se realiza tomando los datos de los 5 *folds* juntos.

Sensib. a la Especif. de	80 %	85 %	90 %
<i>Fold 1</i>	60	60	30
<i>Fold 2</i>	50	50	20
<i>Fold 3</i>	57,14	57,14	57,14
<i>Fold 4</i>	50	50	37,5
<i>Fold 5</i>	66,66	33,33	33,33
Total	65,9	63,63	54,54

Tabla 1: Resultados considerando un solo ojo de cada paciente.

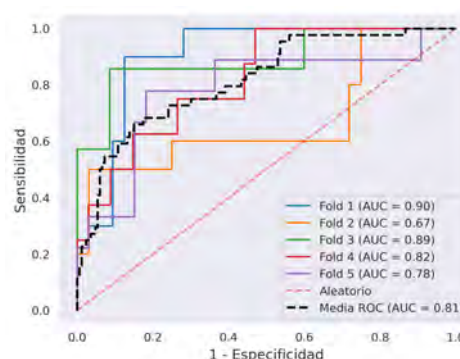


Figura 4: Curva Característica de Operación del Receptor para los 5 folds y el total con el enfoque monocular.

En la Tabla 2 podemos observar los resultados obtenidos al tener en cuenta ambos ojos de un mismo paciente. Se puede observar una notoria mejora en la sensibilidad. Esta supera el 66 % para especificidades fijadas al 80 % y 85 % en todos los *folds*. observando el total de la Tabla 2, la sensibilidad excede el 70 % para cada valor de especificidad. Además, el área bajo la curva (AUC), ilustrada en la Figura 5, alcanza el 90 % al considerar ambos ojos a la misma vez.

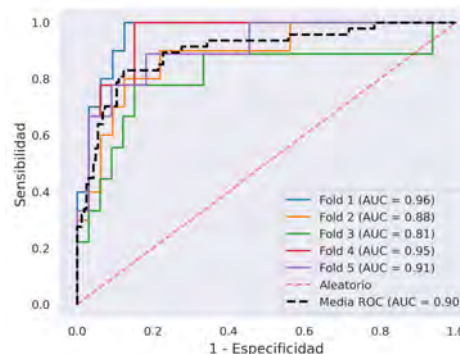


Figura 5: Curva Característica de Operación del Receptor para los 5 folds y el total con el enfoque binocular.

Sensib. a la Especif. de	80 %	85 %	90 %
<i>Fold 1</i>	90	90	80
<i>Fold 2</i>	80	70	60
<i>Fold 3</i>	66,6	66,6	44,4
<i>Fold 4</i>	77,7	77,7	66,6
<i>Fold 5</i>	77,7	77,7	66,6
Total	82,97	80,85	70,21

Tabla 2: Resultados considerando ambos ojos de cada paciente.

Como muestran las Tablas 1 y 2, además de las figuras 4 y 5, la incorporación de información de ambos ojos ha conducido a una mejora significativa en los resultados. En términos cuantitativos, esta mejora se traduce en incrementos del 17,07 %, 17,22 % y 15,67 % en la sensibilidad para especificidades del 80 %, 85 % y 90 %, respectivamente. Además, se observa un aumento del 9 % en la métrica de AUC promedio.

4.2. Tiempo de cómputo

La Tabla 3 muestra los tiempos de cálculo por pliegue. Es evidente que al tener en cuenta la información de ambos ojos el tiempo de cálculo aumenta ligeramente ($\times 1,3$), tanto en la fase de entrenamiento como en la de inferencia. Hay que remarcar que en la fase de inferencia el algoritmo es capaz de procesar 46 imágenes (o 92 si se consideran ambos ojos) en menos de un segundo. Esta eficiencia lo posiciona como un método altamente viable para aplicaciones en tiempo real en dispositivos con recursos limitados.

Tiempo de	Entrenamiento	Inferencia
<i>Un ojo</i>	80,32 \pm 1,051	0,65 \pm 0,02
<i>Ambos ojos</i>	96,64 \pm 0,44	0,86 \pm 0,03

Tabla 3: Tiempo promedio de ejecución entre folds del método propuesto en segundos (media \pm desviación).

5. Conclusiones

En este trabajo se ha explorado la influencia de considerar la información de ambos ojos en una red diseñada para el diagnóstico automático de glaucoma con recursos limitados. Los resultados indican que al tener en cuenta ambos ojos de un mismo paciente se logra una mejora significativa: un incremento del 9 % en el AUC y un aumento de más del 15 % en la sensibilidad para una especificidad del 90 %.

Este hallazgo sugiere que la consideración conjunta de ambos ojos podría ser valiosa en la detección del glaucoma en el marco del aprendizaje máquina. Una posible razón podría ser que, al procesar simultáneamente la información latente de ambos ojos, la red puede acceder a características morfológicas que de otro modo pasarían inadvertidas. Sin embargo, dado que estos son resultados preliminares, aún no es posible establecer conclusiones definitivas al respecto, por lo que se precisa de un análisis exhaustivo y pruebas complementarias de validación.

Referencias

- [1] Yahya Shaikh, Fei Yu, and Anne L. Coleman. Burden of undetected and untreated glaucoma in the united states. *American Journal of Ophthalmology*, 158(6):1121–1129.e1, 2014.
- [2] Alfred Sommer, Joanne Katz, Harry A Quigley, Neil R Miller, Alan L Robin, Ronald C Richter, and Kathe A Witt. Clinically detectable nerve fiber atrophy precedes the onset of glaucomatous field loss. *Archives of ophthalmology*, 109(1):77–83, 1991.
- [3] Rafael Berenguer-Vidal, Rafael Verdú-Monedero, Juan Morales-Sánchez, Inmaculada Sellés-Navarro, and Oleksandr Kovalyk. Analysis of the asymmetry in rNFL thickness using spectralis oct measurements in healthy and glaucoma patients, 2022.
- [4] Tahereh Mahmudi, Raheleh Kafieh, Hossein Rabbani, Alireza Mehri, and Mohammad-Reza Akhlaghi. Evaluation of asymmetry in right and left eyes of normal individuals using extracted features from optical coherence tomography and fundus images. *Journal of Medical Signals and Sensors*, 11(1):12–23, 2021.
- [5] Donald L. Budenz. Symmetry between the right and left eyes of the normal retinal nerve fiber layer measured with optical coherence tomography (an aos thesis). *Transactions of the American Ophthalmological Society*, 106:252–275, 2008.
- [6] L. S. Ong, P. Mitchell, P. R. Healey, and R. G. Cumming. Asymmetry in optic disc parameters: the blue mountains eye study. *Investigative Ophthalmology & Visual Science*, 40(5):849–857, Apr 1999.
- [7] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications, 2017.
- [8] Oleksandr Kovalyk, Juan Morales-Sánchez, Rafael Verdú-Monedero, Inmaculada Sellés-Navarro, Ana Palazón-Cabanes, and José-Luis Sancho-Gómez. Papi-la: Dataset with fundus images and clinical data of both eyes of the same patient for glaucoma assessment. *Scientific Data*, 9(1):1–12, 2022.
- [9] TorchVision maintainers and contributors. Torchvision: Pytorch’s computer vision library, 2016.
- [10] Jia Deng, R. Socher, Li Fei-Fei, Wei Dong, Kai Li, and Li-Jia Li. ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, volume 00, pages 248–255, 06 2009.
- [11] Rui Fan, Kamran Alipour, Christopher Bowd, Mark Christopher, Nicole Bbye, James A Proudfoot, Michael H Goldbaum, Akram Belghith, Christopher A Girkin, Massimo A Fazio, Jeffrey M Liebmann, Robert N Weinreb, Michael Pazzani, David Kriegman, and Linda M Zangwill. Detecting glaucoma from fundus photographs using deep learning without convolutions: Transformer for improved generalization. *Ophthalmology science*, 3(1):100233, March 2023.

Ensamblaje de redes neuronales para el establecimiento de una prioridad en la atención del paciente con lesión en la piel

M. Borrego Casado¹, Begoña Acha¹, Carmen Serrano¹

¹ Departamento de Teoría de la Señal y Comunicaciones, Universidad de Sevilla, Sevilla, España, migborcas@alum.us.es, {bacha, cserrano}@us.es

Resumen

El cáncer de piel supone un tercio de los casos de cáncer en todo el mundo y la celeridad en su detección es fundamental para su correcto tratamiento. Desde hace unos años se vienen proponiendo herramientas de asistencia al diagnóstico basadas en inteligencia artificial para reducir la carga de trabajo de los dermatólogos. En este trabajo proponemos una herramienta de triaje que nos permita clasificar las lesiones recogidas en la base de datos de ISIC-2019 en tres niveles de prioridad de tal modo que se puedan atender con más presteza las lesiones más graves. Con este fin, hemos desarrollado un sistema ensamblado de redes neuronales profundas de clasificación de imágenes y probado diferentes configuraciones de éste.

1. Introducción

De acuerdo con la Organización Mundial de la Salud (OMS) el cáncer de piel constituye un tercio de todos los casos de cáncer diagnosticados en el mundo, además se observa una tendencia en aumento para la ratio de prevalencia de esta enfermedad. En España, el cáncer de piel es el quinto tipo de cáncer más prevalente con una tasa de 47 casos por cada 100.000 [1].

El Servicio Andaluz de Salud cuenta con un servicio de teledermatología que da servicio a una media de 315 pacientes al mes en el Hospital Universitario Virgen Macarena (HUVVM) y 938 en el Hospital Universitario Virgen del Rocío (HUVVR)[2]. Diseñar herramientas de asistencia al diagnóstico nos permitiría reducir la carga de trabajo del dermatólogo y dar un seguimiento más frecuente a los pacientes que lo necesiten [3].

En este trabajo proponemos una herramienta de triaje que permita la clasificación de las lesiones en tres categorías de prioridad a partir de imágenes obtenidas en atención primaria. Para tareas de clasificación de imágenes cobran especial importancia las Redes Neuronales Convolucionales (CNN), por su capacidad de aprender características de la imagen clínica [4]. Existen varios estudios que concluyen que estos sistemas muestran un rendimiento equivalente o superior al de los clínicos a la hora de identificar las lesiones [5, 6].

En los últimos años, varios grupos han conseguido mejores resultados de especificidad utilizando modelos de CNN Ensamblados [7–9]. Estos permiten combinar las fortalezas de varios modelos de CNN lo que repercute en predicciones más enfocadas, a cambio tenemos que lidiar con sistemas más complejos que integran varias CNN.

Vélez et al. [2] realizó un estudio comparativo usando un sistema ensamblado de Resnet50, InceptionV3 y VGG16.

El objetivo de este proyecto era ver si una segmentación previa favorecía o no la predicción del sistema. Se concluyó que la piel que rodea a la lesión ayuda a la CNN a clasificar la lesión.

Lucieri et al. [7] diseñaron una herramienta de clasificación de lesiones cutáneas compuesta por dos CNN diferentes. La primera se centra en encontrar y clasificar conceptos dermatoscópicos, mientras que la segunda clasifica las imágenes a nivel de enfermedad. Ambas redes se influyen entre sí en diferentes momentos durante el proceso de toma de decisiones. Esta herramienta fue probada en varios repositorios, entre ellos ISIC-2019 [10–12], que fue el más extenso. En esta base de datos el sistema obtuvo una exactitud del 88,9%, una precisión del 88,2%, una exactitud del 84,9% y un Área Bajo la Curva (AUC) de 0,91.

Gong et al. [8, 9] desarrollaron un sistema en el que varias CNN realizan predicciones en paralelo sobre una misma imagen de cáncer de piel y después se analiza de manera independiente para seleccionar la predicción más fiable. Este trabajo también realiza un gran esfuerzo de preprocesado a nivel de base de datos, destacando el empleo de Redes GAN para aumento de datos y modificaciones en la función de pérdida.

2. Materiales y Métodos

2.1. Base de Datos

Es fundamental contar con una base de datos lo suficientemente extensa para poder entrenar correctamente estos sistemas [8]. El repositorio público de ISIC-2019 [10–12] es muy utilizado para el entrenamiento de algoritmos de clasificación de lesiones cutáneas [7–9, 13–15]. Esta base de datos cuenta con 25.331 imágenes de entrenamiento y 8238 imágenes de *test*. Para estas últimas no se proporciona la información de la clase por lo que no pueden utilizarse para los fines de este proyecto. La distribución por clase de las imágenes de *test* viene reflejada en la **Tabla 1**.

A partir de las lesiones que se contemplan en este repositorio, los especialistas del servicio de dermatología del HUVVM realizaron una clasificación de estas enfermedades en 3 niveles de prioridad, de tal manera que al final obtenemos un total de **8.473** imágenes de prioridad 1, **16.366** imágenes de prioridad 2 y **881** imágenes de prioridad 3. Las CNN son entrenadas sobre esta clasificación en 3 clases de prioridad, aunque también se hizo una prueba entrenando sobre las 8 categorías de ISIC-

2019 a modo de control y posteriormente mapeando la clasificación a las tres clases de prioridad siguiendo el criterio fijado por los dermatólogos.

Tipo de Lesión	Imágenes de Entrenamiento	Nivel de prioridad
Melanoma (MEL)	4522	1
Nevus melanocítico (NV)	12.875	2
Carcinoma Basal (BCC)	3323	1
Queratosis Actínica (AK)	867	2
Queratosis Benigna (BKL)	2624	2
Dermatofibroma (DF)	239	3
Lesión Vascular (VASC)	253	3
Carcinoma de células escamosas (SCC)	628	1
Total	25.331	

Tabla 1. Distribución de clases de ISIC-2019 [16].

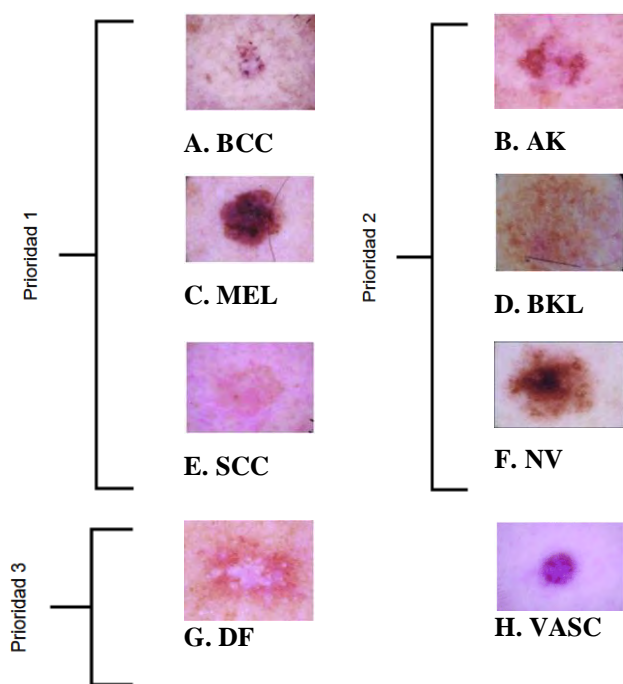


Figura 1. Ejemplos de imágenes organizados por prioridad y enfermedad. A. Carcinoma Basal, B. Queratosis Actínica, C. Melanoma, D. Queratosis Benigna, E. Carcinoma de Células Escamosas, F. Nevus Melanocítico, G. Dermatofibroma, H. Lesión Vascular

Estas imágenes pasan por un proceso básico de aumento de datos consistente en escalado y rotación de las imágenes, tras el cual se obtiene un total de **86.824** imágenes con un reparto porcentual del 40% en las dos primeras clases y un 20% en la última. Dentro de cada clase repartimos las

imágenes 70/20/10 entre imágenes de entrenamiento, de validación y de *test*, respectivamente.

2.2. Clasificador

La arquitectura propuesta consiste en un ensamblado de 3 CNNs: Resnet50, inceptionV3 y VGG16. Estas son entrenadas de manera independiente y realizan sus predicciones en paralelo. La predicción final del modelo ensamblado atiende a la siguiente fórmula:

$$f(y_{j,k}) = \arg_k \max(y_{j,k}),$$

donde $k=1, \dots, 3$ indexa las neuronas de salida de las red y $j=1, \dots, 3$ indexa cada una de estas redes.

Se probaron diferentes estructuras para las capas finales. En Resnet50 e InceptionV3 se probó a prescindir de las capas densas y hacer un entrenamiento integral de la red. A su vez, se varió el tamaño de las capas densas y se ajustó el número de capas convolucionales a entrenar. De esta manera los datos que presentamos en la siguiente sección pueden dividirse en 3 etapas de desarrollo.

Etapas	VGG16	InceptionV3	Resnet50
Control (8 clases)	2 bloques conv, capas densas 512	Todo entrenable, sin densas	Todo entrenable, sin densas
Primera (3 clases)	2 bloques conv, capas densas 512	Todo entrenable, sin densas	Todo entrenable, sin densas
Segunda (3 clases <i>fine tuning</i>)	2 bloques conv, capas densas 128	23 capas conv, Capas densas 128	3 bloques conv, capas densas 128

Tabla 2. Ajustes del sistema en las distintas fases del desarrollo

En la etapa de control y en la primera etapa el parámetro de tasa de aprendizaje se fijó dinámicamente, sin embargo, se obtuvieron mejores resultados con una tasa de aprendizaje estática de 0,00001 en la segunda etapa. También en la segunda etapa se hizo una aproximación a equilibrar la base de datos ajustando los pesos de las clases de la función de pérdida.

3. Resultados

Se probaron todas las redes contra una colección de 5066 imágenes de lesiones cutáneas; de entre las cuales 1700 son de prioridad 1, 3258 son de prioridad 2 y 108 de prioridad 3.

A continuación, expondremos las métricas que extraemos a partir de esta prueba por etapas de desarrollo. Adicionalmente a las métricas usuales hemos calculado la sensibilidad por clase de las CNNs.

3.1. Control

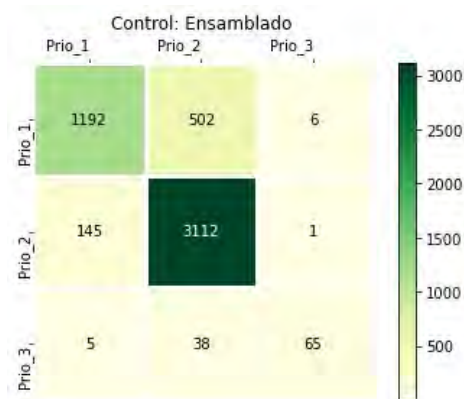


Figura 2. Matriz de confusión red ensamblada, Control. Vertical: Ground Truth, Horizontal: predicción.

	AUC	Sen 1	Sen 2	Sen 3
InceptionV3	0,760	0,568	0,936	0,546
Resnet50	0,800	0,695	0,933	0,546
VGG16	0,843	0,756	0,930	0,685
Ensamblado	0,818	0,701	0,955	0,601

Tabla 3. Métricas Control

3.2. Etapa 1

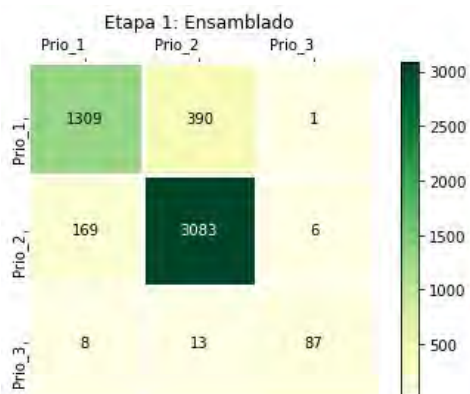


Figura 3. Matriz de confusión red ensamblada, Etapa 1. Vertical: Ground Truth, Horizontal: predicción.

	AUC	Sen 1	Sen 2	Sen 3
InceptionV3	0,785	0,641	0,917	0,592
Resnet50	0,824	0,684	0,944	0,685
VGG16	0,865	0,802	0,900	0,777
Ensamblado	0,874	0,77	0,946	0,805

Tabla 4. Métricas Etapa 1

3.3. Etapa 2

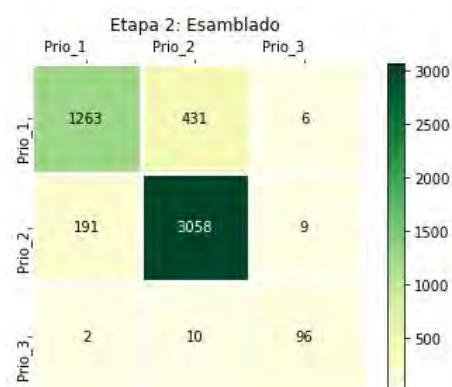


Figura 4. Matriz de confusión red ensamblada, Etapa 2. Vertical: Ground Truth, Horizontal: predicción.

	AUC	Sen 1	Sen 2	Sen 3
InceptionV3	0,800	0,746	0,871	0,666
Resnet50	0,877	0,746	0,918	0,916
VGG16	0,870	0,742	0,926	0,870
Ensamblado	0,877	0,742	0,938	0,888

Tabla 5. Métricas Etapa 2

4. Conclusiones

Se ha realizado un estudio comparativo para un sistema ensamblado de clasificación de imágenes clínicas de cáncer de piel, en el que se han probado diferentes configuraciones de la base de datos y de las CNNs.

Si comparamos la etapa 1 y la 2 podemos ver que el ensamblado no ha sido tan efectivo en la etapa 2, ya que pese a que las CNNs que componen este sistema presentan mejores métricas de forma individual, la diferencia no es tan amplia entre los sistemas ensamblados, lo que sugiere una pérdida de efectividad del ensamblado a altos niveles de precisión individual. Por otro lado, también es cierto que estamos empleando un ensamblado básico, arquitecturas más complejas podrían arrojar mejores resultados.

En general, hemos visto que las clases con más imágenes obtienen mejores resultados de sensibilidad. Esto nos es bastante perjudicial ya que la prioridad 2 es la clase más predominante y a nosotros nos interesa tener una sensibilidad alta para la prioridad 1, ya que son las lesiones más graves y sensibles al tiempo. Por ello, debemos profundizar en el tratamiento de nuestros datos para hacer nuestro sistema más sensible a las clases de interés.

Un fenómeno que hemos observado viendo la sensibilidad a las distintas clases de prioridad es que en ocasiones una CNN se vuelve más sensible a imágenes de una clase concreta comparándola con las otras pese a que todas reciben el mismo *input* y entrenamiento. Sería interesante controlar este fenómeno de cara a poder focalizar la sensibilidad del sistema a las clases con más relevancia

clínica. Con este objetivo nos proponemos desarrollar nuevas técnicas para tratar la distribución de las clases.

5. Agradecimientos

Este trabajo ha sido cofinanciado por el proyecto de Generación de Conocimiento 2021 del Ministerio de Ciencia e Innovación (PID2021-127871OB-I00) y el proyecto PAIDI 2021 de la Consejería de Universidad, Investigación e Innovación de la Junta de Andalucía (PROYEXCEL_00889).

6. Referencias

- [1] “Dimensiones del cáncer | AECC Observatorio.” <https://observatorio.contraelcancer.es/explora/dimensiones-del-cancer> (accessed Jul. 05, 2023).
- [2] Vélez P., Miranda M., Serrano C., Acha B., “Does a Previous Segmentation Improve the Automatic Detection of Basal Cell Carcinoma Using Deep Neural Networks?,” 2022, doi: 10.3390/app12042092.
- [3] M. Saarela and L. Geogieva, “Robustness, Stability, and Fidelity of Explanations for a Deep Skin Cancer Classification Model,” *Applied Sciences (Switzerland)*, vol. 12, no. 19, Oct. 2022, doi: 10.3390/app12199545.
- [4] D. Shen, G. Wu, and H.-I. Suk, “Deep Learning in Medical Image Analysis,” 2017, doi: 10.1146/annurev-bioeng-071516.
- [5] S. Haggemüller *et al.*, “Skin cancer classification via convolutional neural networks: systematic review of studies involving human experts,” *Eur J Cancer*, vol. 156, pp. 202–216, Oct. 2021, doi: 10.1016/j.ejca.2021.06.049.
- [6] P. Tschandl *et al.*, “Human–computer collaboration for skin cancer recognition,” *Nat Med*, vol. 26, no. 8, pp. 1229–1234, Aug. 2020, doi: 10.1038/s41591-020-0942-0.
- [7] A. Lucieri, M. N. Bajwa, S. A. Braun, M. I. Malik, A. Dengel, and S. Ahmed, “ExAID: A multimodal explanation framework for computer-aided diagnosis of skin lesions,” *Comput Methods Programs Biomed*, vol. 215, Mar. 2022, doi: 10.1016/j.cmpb.2022.106620.
- [8] A. Gong, X. Yao, and W. Lin, “Dermoscopy Image Classification Based on StyleGANs and Decision Fusion,” *IEEE Access*, vol. 8, pp. 70640–70650, 2020, doi: 10.1109/ACCESS.2020.2986916.
- [9] A. Gong, X. Yao, and W. Lin, “Classification for Dermoscopy Images Using Convolutional Neural Networks Based on the Ensemble of Individual Advantage and Group Decision,” *IEEE Access*, vol. 8, pp. 155337–155351, 2020, doi: 10.1109/ACCESS.2020.3019210.
- [10] M. Combalia *et al.*, “BCN20000: Dermoscopic Lesions in the Wild,” Aug. 2019, [Online]. Available: <http://arxiv.org/abs/1908.02288>
- [11] N. C. F. Codella *et al.*, “Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC),” Oct. 2017, [Online]. Available: <http://arxiv.org/abs/1710.05006>
- [12] P. Tschandl, C. Rosendahl, and H. Kittler, “Data descriptor: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions,” *Sci Data*, vol. 5, Aug. 2018, doi: 10.1038/sdata.2018.161.
- [13] M. A. Kassem, K. M. Hosny, and M. M. Fouad, “Skin Lesions Classification into Eight Classes for ISIC 2019 Using Deep Convolutional Neural Network and Transfer Learning,” *IEEE Access*, vol. 8, pp. 114822–114832, 2020, doi: 10.1109/ACCESS.2020.3003890.
- [14] D. M. Krishna, S. K. Sahu, and G. R. L. V. N. S. Raju, “DTLNet: Deep Transfer Learning-based Hybrid Model for Skin Lesion Detection and Classification,” in *2022 2nd Asian Conference on Innovation in Technology, ASIANCON 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, doi: 10.1109/ASIANCON55314.2022.9909252.
- [15] N. Nigar, M. Umar, M. K. Shahzad, S. Islam, and D. Abalo, “A Deep Learning Approach Based on Explainable Artificial Intelligence for Skin Lesion Classification,” *IEEE Access*, vol. 10, pp. 113715–113725, 2022, doi: 10.1109/ACCESS.2022.3217217.
- [16] B. Cassidy, C. Kendrick, A. Brodzicki, J. Jaworek-Korjakowska, and M. H. Yap, “Analysis of the ISIC image datasets: Usage, benchmarks and recommendations,” *Med Image Anal*, vol. 75, Jan. 2022, doi: 10.1016/j.media.2021.102305.

Comparación de múltiples redes neuronales convolucionales para el diagnóstico automático de la degeneración macular asociada a la edad usando retinografías

R. Romero-Oraá^{1,2}, M. Herrero-Tudela¹, R. Hornero^{1,2}, M. I. López Gálvez^{1,2}, M. García^{1,2}

¹ Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, {roberto.romero, maria.herrero.tudela, roberto.hornero, maria.garcia.gadanon}@uva.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

Resumen

La degeneración macular asociada a la edad (DMAE) es la principal causa de pérdida irreversible de visión en la población anciana. Los algoritmos de deep learning y, en particular, las redes neuronales convolucionales (CNN), han demostrado ser útiles en el diagnóstico precoz de la enfermedad. En este trabajo se compararon 12 modelos CNN preentrenados: VGG-19, ResNet, InceptionV3, InceptionResNetV2, Xception, DenseNet, NASNet, MobileNetV3, RegNet, EfficientNetV2, ResNet-RS y ConvNeXt. se utilizó la base de datos ADAM, compuesta por 1,200 retinografías. Los resultados obtenidos permitieron seleccionar el modelo basado en la arquitectura ResNet-RS como aquel más adecuado para el diagnóstico, alcanzando una precisión del 89.50%, una sensibilidad del 90.35%, una especificidad del 86.52%, un F1-score de 0.93 y un área bajo la curva ROC de 0.9497. Cabe destacar que la arquitectura ResNet-RS nunca había sido explorada en el contexto de la DMAE. El modelo propuesto podría ser de utilidad en el entorno clínico y servir de base para desarrollar modelos más complejos.

1. Introducción

La degeneración macular asociada a la edad (DMAE) es la principal causa de pérdida de visión entre las personas mayores de 50 años [1]. Esta pérdida de visión es irreversible, provocando con el tiempo una ceguera completa [2]. Como la mayoría de las enfermedades retinianas, la DMAE no presenta síntomas visibles en sus primeras etapas, por lo que su detección precoz requiere exámenes oftalmológicos periódicos [3]. En este sentido, las retinografías son la modalidad de imagen más rentable para el cribado de la enfermedad [4]. Sin embargo, la alta prevalencia de la DMAE y la falta de recursos sanitarios está sobrecargando el sistema de atención al paciente, haciendo que los tratamientos lleguen tarde [2]. Por ello, los algoritmos de inteligencia artificial han demostrado ser útiles en la ayuda al diagnóstico de la DMAE [2].

En los últimos años, los algoritmos de *deep learning* están superando los clasificadores tradicionales [2]. Además, permiten obtener representaciones relevantes de los datos sin la necesidad de extraer características de forma manual [2]. En este contexto, las redes neuronales convolucionales (*convolutional neural network*, CNN) son los modelos de *deep learning* más populares empleados para tareas de visión artificial, incluyendo el análisis de imágenes médicas [2]. Aunque los *transformers* han revolucionado el estado del arte de los modelos de clasificación de imagen

[5], las últimas arquitecturas CNN han logrado superar su rendimiento en términos de precisión y escalabilidad [6].

En la literatura podemos encontrar varios métodos de detección de la DMAE basados en CNNs. Se han explorado individualmente las arquitecturas OverFeat [7], [8] Alexnet [9], VGG-16 [10], ResNet50 [10], [11], ResNet101 [4] y EfficientNetB4 [4]. Tan et al. [12], por su parte, desarrollaron una arquitectura ad-hoc de 14 capas. Otros autores utilizaron estrategias combinadas. Por ejemplo, Grassman et al. [13] combinaron AlexNet, VGG-16, GoogLeNet, Inception-v3, ResNet101 e Inception-ResNet-v2. Govindaiah et al. [14], por otro lado, combinaron las arquitecturas Inception-ResNet-V2 y Xception. En algunos trabajos el entrenamiento del modelo se hizo en base a la detección de lesiones, para lo cual se utilizaron las etiquetas correspondientes manualmente anotadas [15]. En la literatura asociada al diagnóstico automático de la DMAE, tiene especial interés el desafío ADAM [4]. Cabe destacar que los 5 equipos mejor clasificados propusieron estrategias combinadas de CNNs.

La base de todas estas propuestas es una arquitectura CNN (o conjunto de ellas) a modo de *backbone* o extractor de características. Sin embargo, ningún estudio ha establecido una comparación directa que permita elegir la arquitectura más eficaz para detectar los signos de la DMAE. El objetivo de este trabajo ha sido comparar múltiples arquitecturas CNN para el diagnóstico de la DMAE. De esta manera, sería posible construir modelos más complejos basados en la arquitectura más adecuada. Además, se han incluido en la comparación arquitecturas modernas, como ConvNeXt [6] o ResNet-RS [16], que no han sido exploradas anteriormente en este contexto.

2. Base de datos de retinografías

En este trabajo se utilizó la base de datos ADAM [4], que consta de 1,200 retinografías almacenadas en formato JPEG y fue proporcionada por el Centro Oftalmológico de Zhongshan de la Universidad de Sun Yat-sen (China). Se utilizó el retinógrafo Zeiss Visucam 500 para capturar 824 de las imágenes a una resolución de $2,124 \times 2,056$ píxeles y el retinógrafo Canon CR-2 para capturar las 376 imágenes restantes a una resolución de $1,444 \times 1,444$ píxeles. El campo de visión de todas las imágenes se centró en la mácula. Los pacientes eran de nacionalidad china, con edades comprendidas en 53.19 ± 15.59 años (47%

mujeres). Se seleccionaron exclusivamente imágenes de calidad adecuada para su procesamiento. La base de datos resultante se dividió en 3 grupos de 400 imágenes cada uno: entrenamiento, validación y test. Todos mantenían la misma proporción de muestras patológicas (311 imágenes) y sanas (89 imágenes) [4]. La Figura 1 muestra un ejemplo de cada clase.

3. Métodos

Todos los modelos a comparar partían de una etapa de preprocesado para normalizar las imágenes de entrada. A continuación, incluían el *backbone* correspondiente a la CNN elegida y, finalmente, compartían la misma estructura de capas *fully-connected* al final de la red. Asimismo, para todos los modelos se aplicaron los mismos hiperparámetros y técnicas adicionales, como son *data augmentation*, *transfer learning* y *fine-tuning*.

3.1. Preprocesado

Para normalizar los datos de entrada y reducir el tiempo de procesado, todas las imágenes se redimensionaron a una resolución de 512×512 píxeles [4]. Además, para adaptar los valores de los píxeles a la entrada esperada por las redes, las imágenes se normalizaron en el rango $[-1,1]$ [16].

3.2. Data augmentation

Las redes neuronales profundas requieren un elevado número de datos de entrenamiento. Para aumentar el número de imágenes con las que entrenar los modelos de este trabajo, se empleó la técnica *online data augmentation* [2]. Esta técnica permite generar aleatoriamente nuevas imágenes sintéticas en cada época a partir del conjunto de entrenamiento. Para ello, se aplicaron transformaciones simples: rotaciones, desplazamientos, volteos y escalado.

3.3. Transfer learning y fine-tuning

Transfer learning es una solución adicional al escaso volumen de datos de entrenamiento disponible [2]. Para aplicar esta técnica, inicializamos los pesos de la red con los valores aprendidos al resolver un problema de otro dominio del que sí se dispone de un gran volumen de datos [2]. En este trabajo, se utilizaron modelos preentrenados con las imágenes del proyecto ImageNet [17].

A continuación, se aplicó *fine-tuning* [2]. Para ello, se utilizaron las imágenes del conjunto de entrenamiento para reentrenar todas las capas de cada modelo preentrenado. Esto permitió ajustar los pesos de las redes a la tarea de detección de la DMAE.

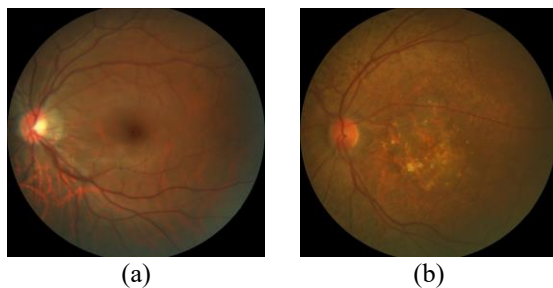


Figura 1. Imágenes de ejemplo de la base de datos ADAM. (a) Sin DMAE. (b) Con DMAE.

3.4. Backbone CNN

En este trabajo se han desarrollado múltiples modelos CNN con distintas arquitecturas en el *backbone* para su comparación. Las arquitecturas elegidas han demostrado gran rendimiento en la clasificación de imágenes genéricas y fácil acceso a su correspondiente modelo preentrenado. Se detallan a continuación [18]:

- VGG-19. Presenta una topología homogénea de núcleos de convolución pequeños que permitieron extraer características de baja resolución espacial. En este trabajo se utilizó la versión de 19 capas.
- ResNet. Introduce el concepto de conexiones residuales, que evita el problema de la degradación en redes muy profundas. En este trabajo se utilizó la versión de 152 capas.
- InceptionV3. Es una versión de bajo coste computacional de la arquitectura GoogleNet, que introdujo el bloque Inception. Este bloque permite abstraer características de distintas escalas espaciales.
- InceptionResNetV2. Combina las ventajas de Inception y ResNet.
- Xception. Se considera una versión extrema de Inception que explota la idea de convolución separable en profundidad. Esto hace que la red sea computacionalmente eficiente al desacoplar la correlación del espacio y el canal.
- DenseNet. Se propuso conectar cada capa hacia adelante con todas las demás capas. Esto alivia el problema del gradiente de fuga, fortalece la propagación y reutilización de características y reduce la cantidad de parámetros. En este trabajo se empleó la versión de 201 capas.
- NASNet. Introduce una técnica llamada *neural architecture search* (NAS), que automatiza el diseño de redes neuronales. Con ello, descubre nuevos bloques compuestos de celdas normales y de reducción que optimizan el aprendizaje.
- MobileNetV3. Utiliza NAS para introducir nuevos avances en la arquitectura manteniendo la eficiencia computacional.
- RegNet. Se diseñan espacios de búsqueda para mejorar la técnica NAS.
- EfficientNetV2. Se propone aplicar NAS con foco en el entrenamiento y escalado. Se trata de un modelo muy rápido y eficiente.
- ResNet-RS. Supone una mejora de la arquitectura ResNet aportando 2 nuevas estrategias de escalado.
- ConvNeXt. Se inspira en los *transformers* para obtener una versión mejorada de ResNet.

3.5. Capas fully-connected

Para cada modelo desarrollado se añadieron al final de la arquitectura 3 capas *fully-connected* de 1024, 512 y 1 neuronas, respectivamente, que hacían la función de clasificador. En las dos primeras capas se utilizó la función de activación ReLu y, en la última capa, la función de activación sigmoide. Entre cada capa *fully-connected* se

aplicó la técnica *dropout* con factor 0.25 para minimizar el sobreentrenamiento [2].

3.6. Entrenamiento

Cada modelo se entrenó hasta un máximo de 200 épocas aplicando *early stopping* cuando el error sobre el conjunto de validación no mejoraba durante 7 épocas. Además, se redujo la tasa de aprendizaje a la mitad cada vez que dicho error alcanzase un mínimo y se mantuviese constante durante 3 épocas [1]. Para lidiar con el desbalanceo de clases, se aplicó la entropía cruzada binaria ponderada como función de pérdidas. También se utilizó el algoritmo de optimización Adam, la tasa inicial de aprendizaje se fijó en 0.00005 y el tamaño de *batch* fue de 4 imágenes [1].

4. Resultados

Todos los modelos se evaluaron sobre las 400 imágenes del conjunto de test en términos de precisión (PR), sensibilidad (SE), especificidad (ES), F1-score y área bajo la curva ROC (AUC). La Tabla 1 recoge estos resultados.

5. Discusión

En este trabajo se han desarrollado 12 modelos CNN para la detección automática de la DMAE. Con cada modelo se exploró una arquitectura distinta en el *backbone*. El resto de la arquitectura, así como los hiperparámetros y técnicas utilizadas fueron idénticas en cada caso. Esto nos ha permitido construir modelos eficaces y establecer una comparación directa entre ellos. Analizando los resultados de la Tabla 1, podemos observar que todos los modelos superan el 85.00 % de PR excepto aquel basado en VGG-19. Esto puede deberse a que esta arquitectura no es muy profunda en comparación con el resto. Atendiendo exclusivamente a la PR, la arquitectura InceptionV3 es la que alcanza el valor más alto (91.00 %). Sin embargo, sus valores de SE=93.89% y ES=80.90% parecen bastante desbalanceados, tal y como sucede para otras arquitecturas (ResNet, NASNet, Xception y ConvNeXt). Por el

contrario, los modelos basados en las arquitecturas DenseNet, ResNet-RS y EfficientNetV2 alcanzaron valores de SE y ES mucho más balanceados. Además, el modelo de la arquitectura ResNet-RS arrojó el valor más alto de AUC (0.9497), lo que nos llevó a seleccionar este modelo como aquel más adecuado para la detección automática de la DMAE. Hasta donde alcanza nuestro conocimiento, esta arquitectura nunca se había empleado antes con este objetivo.

Nuestros resultados son directamente comparables con los publicados en el desafío ADAM [4], pues se utilizó la misma base de datos y la misma separación de grupos. La Tabla 2 recoge esta comparativa para la métrica AUC. Como se puede observar, nuestro modelo seleccionado, basado en la arquitectura ResNet-RS, alcanzaría el cuarto puesto en el desafío ADAM. Este trabajo demuestra que una elección apropiada del *backbone* junto con la optimización de hiperparámetros y aplicación de técnicas adecuadas permite desarrollar modelos CNN muy eficaces manteniendo una arquitectura simple. Cabe destacar que los métodos propuestos por los equipos finalistas se basan en redes combinadas bastante más complejas que nuestro modelo seleccionado. En concreto, el equipo ganador hizo uso de las anotaciones clínicas correspondientes a 15 tipos de lesiones para entrenar modelos independientes [4]. Por el contrario, nuestro método requiere exclusivamente las anotaciones a nivel de imagen correspondientes a la ausencia o presencia de enfermedad.

La mayoría de los modelos desarrollados en este trabajo ofrecieron resultados prometedores. Sin embargo, el estudio no está exento de limitaciones. En primer lugar, solamente se ha utilizado una base de datos. Sería deseable incorporar nuevas imágenes con distintas características para evaluar la robustez del método. En segundo lugar, la capacidad predictiva del modelo seleccionado se ve superada por otras alternativas. Aunque el objetivo aquí no era desarrollar un método complejo, en el futuro se deben explorar mecanismos de atención, *transformers*, nuevas funciones de pérdidas y estrategias combinadas.

CNN	PR (%)	SE (%)	ES (%)	F1	AUC
VGG-19	77.50	74.92	86.52	0.84	0.8117
ResNet	89.00	92.92	75.28	0.92	0.9392
InceptionV3	91.00	93.89	80.90	0.94	0.9239
InceptionResNetV2	89.00	91.96	78.65	0.93	0.9288
Xception	90.25	94.32	76.40	0.94	0.9294
DenseNet	89.25	90.35	85.39	0.93	0.9492
NASNet	89.50	93.57	75.28	0.93	0.9257
MobileNetV3	88.00	90.03	80.90	0.92	0.9330
RegNet	89.50	92.92	77.53	0.93	0.9260
EfficientNetV2	85.75	85.21	87.64	0.90	0.9417
ResNet-RS	89.50	90.35	86.52	0.93	0.9497
ConvNeXt	89.50	94.86	70.79	0.93	0.9491

Tabla 1. Resultados sobre el conjunto de test.

Equipo	AUC
VUNO EYE TEAM	0.9714
ForbiddenFruit	0.9592
Zasti_AI	0.9581
Nosotros	0.9497
Muenai_Tim	0.9399
ADAM-TEAM	0.9287
WWW	0.9178
XxlzT	0.9097
TeamTiger	0.9086
AiraMatrix	0.8847

Tabla 2. Comparación de nuestros resultados con los resultados obtenidos por los equipos finalistas del desafío ADAM.

6. Conclusiones

Este estudio expone con éxito una comparación de múltiples modelos CNN basados en distintas arquitecturas para la detección automática de la DMAE. Se seleccionó el modelo basado en la arquitectura ResNet-RS como aquel más adecuado para el diagnóstico de la enfermedad. Aunque se trata de un modelo simple, los resultados son muy satisfactorios, superando a varias alternativas basadas en estrategias combinadas. Por tanto, su uso podría ser de utilidad en el entorno clínico. Además, el modelo propuesto podría servir de base para desarrollar modelos más complejos.

Agradecimientos

Esta investigación se ha desarrollado en el marco de las ayudas TED2021-131913B-I00, PID2020-115468RB-I00 y PGC2018-098214-A-I00 financiadas por el 'Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación/10.13039/501100011033/' y el Fondo Europeo de Desarrollo Regional (FEDER). Una forma de hacer Europa; y por el 'CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)' a través del 'Instituto de Salud Carlos III' cofinanciado con fondos FEDER. M. Herrero Tudela cuenta con un contrato predoctoral de la Universidad de Valladolid.

Referencias

- [1] T. Li *et al.*, “Applications of deep learning in fundus images: A review,” *Medical Image Analysis*, vol. 69. Elsevier B.V., p. 101971, Apr. 01, 2021. doi: 10.1016/j.media.2021.101971.
- [2] S. Sengupta, A. Singh, H. A. Leopold, T. Gulati, and V. Lakshminarayanan, “Ophthalmic diagnosis using deep learning with fundus images-A critical review,” 2019, doi: 10.1016/j.artmed.2019.101758.
- [3] S. Iqbal, T. M. Khan, K. Naveed, S. S. Naqvi, and S. J. Nawaz, “Recent trends and advances in fundus image analysis: A review,” *Comput Biol Med*, vol. 151, p. 106277, Dec. 2022, doi: 10.1016/J.COMPBIOMED.2022.106277.
- [4] H. Fang *et al.*, “ADAM Challenge: Detecting Age-Related Macular Degeneration From Fundus Images,” *IEEE Trans Med Imaging*, vol. 41, no. 10, pp. 2828–2847, Oct. 2022, doi: 10.1109/TMI.2022.3172773.
- [5] A. Dosovitskiy *et al.*, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” *ICLR 2021 - 9th International Conference on Learning Representations*, Oct. 2020, Accessed: Sep. 04, 2023. [Online]. Available: <https://arxiv.org/abs/2010.11929v2>
- [6] Z. Liu, H. Mao, C. Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, “A ConvNet for the 2020s,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2022-June, pp. 11966–11976, Jan. 2022, doi: 10.1109/CVPR52688.2022.01167.
- [7] P. Burlina, D. E. Freund, N. Joshi, Y. Wolfson, and N. M. Bressler, “Detection of age-related macular degeneration via deep learning,” *Proceedings - International Symposium on Biomedical Imaging*, vol. 2016-June, pp. 184–188, Jun. 2016, doi: 10.1109/ISBI.2016.7493240.
- [8] A. Horta *et al.*, “A hybrid approach for incorporating deep visual features and side channel information with applications to AMD detection,” *Proceedings - 16th IEEE International Conference on Machine Learning and Applications, ICMLA 2017*, vol. 2017-December, pp. 716–720, 2017, doi: 10.1109/ICMLA.2017.00-75.
- [9] P. M. Burlina, N. Joshi, M. Pekala, K. D. Pacheco, D. E. Freund, and N. M. Bressler, “Automated Grading of Age-Related Macular Degeneration From Color Fundus Images Using Deep Convolutional Neural Networks,” *JAMA Ophthalmol*, vol. 135, no. 11, pp. 1170–1176, Nov. 2017, doi: 10.1001/JAMAOPHTHALMOL.2017.3782.
- [10] A. Govindaiah, M. A. Hussain, R. T. Smith, and A. Bhuiyan, “Deep convolutional neural network based screening and assessment of age-related macular degeneration from fundus images,” *Proceedings - International Symposium on Biomedical Imaging*, vol. 2018-April, pp. 1525–1528, May 2018, doi: 10.1109/ISBI.2018.8363863.
- [11] P. M. Burlina, N. Joshi, K. D. Pacheco, D. E. Freund, J. Kong, and N. M. Bressler, “Use of Deep Learning for Detailed Severity Characterization and Estimation of 5-Year Risk Among Patients With Age-Related Macular Degeneration,” *JAMA Ophthalmol*, vol. 136, no. 12, p. 1359, Dec. 2018, doi: 10.1001/JAMAOPHTHALMOL.2018.4118.
- [12] J. H. Tan *et al.*, “Age-related Macular Degeneration detection using deep convolutional neural network,” *Future Generation Computer Systems*, vol. 87, pp. 127–135, Oct. 2018, doi: 10.1016/J.FUTURE.2018.05.001.
- [13] F. Grassmann *et al.*, “A Deep Learning Algorithm for Prediction of Age-Related Eye Disease Study Severity Scale for Age-Related Macular Degeneration from Color Fundus Photography,” *Ophthalmology*, vol. 125, no. 9, pp. 1410–1420, Sep. 2018, doi: 10.1016/J.OPHTHA.2018.02.037.
- [14] A. Govindaiah, R. T. Smith, and A. Bhuiyan, “A New and Improved Method for Automated Screening of Age-Related Macular Degeneration Using Ensemble Deep Neural Networks,” *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2018-July, pp. 702–705, Oct. 2018, doi: 10.1109/EMBC.2018.8512379.
- [15] Y. Peng *et al.*, “DeepSeeNet: A Deep Learning Model for Automated Classification of Patient-based Age-related Macular Degeneration Severity from Color Fundus Photographs,” *Ophthalmology*, vol. 126, no. 4, pp. 565–575, Apr. 2019, doi: 10.1016/J.OPHTHA.2018.11.015.
- [16] I. Bello *et al.*, “Revisiting ResNets: Improved Training and Scaling Strategies”, Accessed: Sep. 04, 2023. [Online]. Available: <https://github.com/tensorflow/tpu/tree/>
- [17] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A Large-Scale Hierarchical Image Database”, Accessed: Sep. 05, 2023. [Online]. Available: <http://www.image-net.org>.
- [18] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artificial Intelligence Review 2020* 53:8, vol. 53, no. 8, pp. 5455–5516, Apr. 2020, doi: 10.1007/S10462-020-09825-6.

Métodos de ayuda al diagnóstico de glaucoma basado en diferencias de características extraídas de ambos ojos en tomografías de coherencia óptica

David Palazón Palau, Francisco Rodríguez Robles, Oleksandr Kovalyk,
Rafael Verdú Monedero, Juan Morales Sánchez

Universidad Politécnica de Cartagena, 30202 Cartagena, España
david.palazon@edu.upct.es, franroroba@gmail.com, oleksandr.kovalyk@edu.upct.es,
rafael.verdu@upct.es, juan.morales@upct.es

Resumen

Una de las enfermedades más relevantes que impacta en el nervio óptico es el glaucoma, provocando un deterioro progresivo e irreversible que disminuye el campo de visión del paciente. El espesor de la capa de fibras nerviosas de la retina se convierte en un indicador crucial para evaluar la condición y la evolución de esta afección. En el presente artículo se muestran resultados favorables para los métodos de ayuda al diagnóstico a partir de tomografías de coherencia óptica de las cuales se han extraído los grosores de cada sector de la capa de fibras, se le han aplicado métricas de asimetría entre ambos ojos y se han realizado varios estimadores basados en modelos de Machine Learning. Las asimetrías presentes en ambos ojos han resultado ser informativas respecto a la detección de pacientes con glaucoma. Los resultados muestran la efectividad de los métodos de ayuda al diagnóstico desarrollados y una mejora evidente en la clasificación de los pacientes con el uso de Random Forest respecto de Decision Tree, logrando clasificar correctamente en media a un 87,5 % de los pacientes con neuropatía óptica glaucomatosa.

1. Introducción

El glaucoma, una enfermedad que daña el nervio óptico, es una de las principales causas de ceguera en los países desarrollados, después de las cataratas. Se estima que en 2020, alrededor de 80 millones de personas padecían glaucoma [1], y se espera que esta cifra aumente a 111,8 millones para 2040 [2]. Esta enfermedad es crónica y generalmente no presenta síntomas en sus primeras etapas, lo que resalta la importancia del diagnóstico temprano para prevenir la pérdida permanente de la visión. Por ello, se busca desarrollar métodos automatizados y simplificados para el diagnóstico, así como mejorar los ya existentes [3], con el objetivo de hacerlos accesibles en los centros de atención primaria y mejorar las posibilidades de prevención y tratamiento temprano.

La neuropatía óptica por glaucoma es una condición que afecta al nervio óptico y puede tener diversas causas. Principalmente, esta afección se desarrolla en la parte frontal del ojo debido a desequilibrios en la producción y drenaje del humor acuoso, lo que aumenta la presión intraocular y, en última instancia, daña el nervio óptico produciendo una pérdida progresiva e irreversible del campo visual que deriva en ceguera [4]. El diagnóstico clínico del glaucoma se basa en la medición de la presión intraocular mediante tonometría, la evaluación del campo visual mediante campimetría y el análisis de

la anatomía ocular, utilizando técnicas como la retinografía y la tomografía de coherencia óptica (OCT).

Estudios previos han centrado su hipótesis acerca de las asimetrías interoculares de la capa de fibras nerviosas de la retina (retinal nerve fiber layer, RNFL) con umbrales estadísticos sobre pacientes sanos y pacientes con glaucoma [5], proponiendo a los algoritmos como potenciales métodos de ayuda al diagnóstico. Estudios como [3, 6, 7] han desarrollado y confirmado dicha propuesta.

Este artículo se centra en el avance de los métodos de ayuda al diagnóstico ya desarrollados desde un paradigma de la sencillez e interpretabilidad de los modelos basados en características de asimetría entre ambos ojos de un mismo paciente, así como del avance y el uso de modelos más complejos pero potentes.

2. Materiales y Métodos

Esta sección ofrece una breve descripción de los recursos utilizados en el estudio, específicamente, las imágenes de tomografía de coherencia óptica 2D (modo B-scan), que abarcan la región alrededor del nervio óptico y que conforman la base de datos, así como el proceso de obtención de estas imágenes y su posterior procesado. Este tipo de imágenes médicas se emplea de manera extensa en el diagnóstico de la enfermedad del glaucoma, ya que permite determinar los espesores de la RNFL.

2.1. Base de datos

La base de datos, utilizada en [3, 6, 8], está conformada por las imágenes de OCT de ambos ojos de 207 pacientes de los cuales 160 han sido etiquetados como sanos y 47 como patológicos. El etiquetado manual por parte de expertos del Servicio de Oftalmología del Hospital General Universitario Reina Sofía en Murcia, España, donde se utilizó un OCT Spectralis S2610-CB de Heidelberg Engineering GmbH, ha permitido realizar estudios sobre las características y asimetrías de la RNFL mediante algoritmos de aprendizaje máquina supervisado.

2.2. Adquisición de imágenes

Las imágenes OCT-SD en el dominio espectral, también conocidas como Fourier-domain OCT (FD-OCT), se han capturado utilizando un escaneo con un diodo láser superluminiscente (SLD) de longitud de onda de 870 nm. Estas imágenes registran una sección cilíndri-

ca alrededor del disco óptico de las capas de la retina (Figura 1).

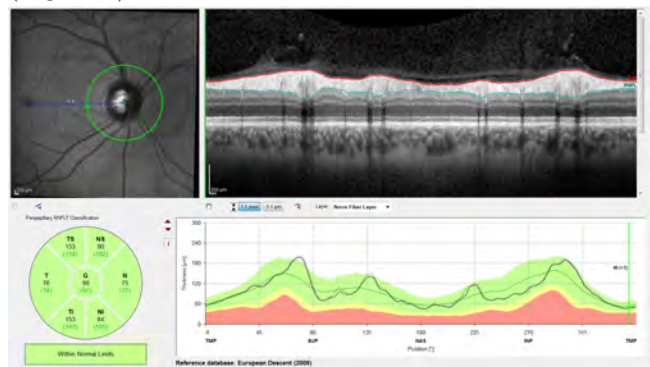


Figura 1. Captura de pantalla del espectrómetro Spectralis. En la parte superior, se muestra una imagen de la retina centrada en el disco óptico, y se indica la ubicación del barrido B peripapilar mediante un círculo verde. A la derecha, se muestra el barrido B peripapilar OCT en 2D con reconstrucción cartesiana. En las imágenes se incluyen coordenadas y valores medios estimados del grosor de la capa de fibras nerviosas de la retina (RNFL) para cada sector y el promedio general. También se muestra el contorno rectificado de la RNFL con el grosor estimado y los valores de referencia de la base de datos de ascendencia europea (2009).

Esta sección circular se divide en varios sectores, denominados temporal (T), temporal superior (TS), nasal superior (NS), nasal (N), nasal inferior (NI) y temporal inferior (TI). En cada sector se mide el grosor de la capa de fibras nerviosas de la retina.

Para facilitar el estudio y el procesamiento de estas imágenes, el B-scan circular se proyecta de coordenadas polares a coordenadas cartesianas en el eje x . En la parte inferior de la Figura 1 se detalla la correspondencia entre los valores medidos en grados en coordenadas polares y los valores medidos en píxeles en coordenadas cartesianas.

Las imágenes de OCT utilizadas en este estudio tienen una resolución de 768×496 píxeles, una profundidad de 8 bits por píxel en escala de grises y un escalado en z de $3,87 \mu\text{m}$ por píxel. El escalado en x depende del diámetro de la circunferencia peripapilar de la OCT y no es significativo para este estudio, ya que el objetivo principal es determinar el grosor vertical del RNFL en la imagen de la OCT. En la parte superior izquierda de la Figura , se muestra el trazado circular del haz del láser SLD posicionado sobre el centro del nervio óptico, también conocido como papila, en la imagen de fondo de ojo de la retina. La correspondiente proyección de coordenadas cartesianas del B-scan de OCT se presenta en la parte superior derecha de la Figura 1.

2.3. Segmentación de la RNFL y cálculo del grosor de los sectores en OCT

El siguiente paso se enfoca en la segmentación de las imágenes con el fin de estimar el grosor de la capa de fi-

bras nerviosas de la retina. La segmentación de la RNFL representa un gran desafío debido a las características de las imágenes, especialmente las propias de la capa retinal en análisis. Como se menciona en los estudios previos [9, 10], es común encontrar ruido especular, un bajo nivel de contraste y características morfológicas irregulares en los conjuntos de datos de OCT. Además, dado que la RNFL es la capa más interna de la retina, las imágenes pueden presentar artefactos debido al efecto de sombreado causado por las venas retinianas, lo cual dificulta definir con precisión las capas en sectores específicos del ojo. En [11] se realiza una revisión exhaustiva de los enfoques para la segmentación, algunos de los cuales son relevantes para el conjunto de datos utilizado en este estudio.

La segmentación utilizada en este trabajo es la proporcionada por el software Spectralis versión 6.9.4.0. Utilizando este algoritmo, se calcula el grosor de la RNFL a partir de las posiciones de los límites superior e inferior de la capa proporcionados por el algoritmo de segmentación. Estas posiciones se escalan de píxeles a micrómetros según la resolución de OCT y se proporciona el grosor promedio de la RNFL para cada sector del ojo (T, TS, NS, N, NI y TI), así como el valor promedio de grosor del contorno circumpapilar global (G).

2.4. Métricas de asimetría aplicadas a los grosores de la RNFL

Se recogen las métricas de asimetría que mejores resultados ofrecen, anteriormente propuestas en [3, 6] y se plantea una nueva en la que se tiene en cuenta una mayor implicación del sector G, el cual ha demostrado ser el sector con mayor importancia en la modalidad de imagen médica de OCT en los modelos de clasificación entrenados en otros trabajos como [6, 7]. Las métricas consideradas en este trabajo son:

$$|\Delta_{S,i}| = \frac{|w_{S,i}^r - w_{S,i}^l|}{w_{S,i}^r + w_{S,i}^l}, \quad (1)$$

$$|\bar{\Delta}_{S,i}| = \frac{|w_{S,i}^r - w_{S,i}^l|}{w_{G,i}^r + w_{G,i}^l}, \quad (2)$$

$$|\Delta_{S,i}|^{0,5} = \sqrt{\frac{|w_{S,i}^r - w_{S,i}^l|}{w_{S,i}^r + w_{S,i}^l}}, \quad (3)$$

$$|\tau_{S,i}| = \left| \frac{w_{S,i}^r + w_{G,i}^r}{w_{G,i}^l} - \frac{w_{S,i}^l + w_{G,i}^l}{w_{G,i}^r} \right|, \quad (4)$$

donde S indica el sector concreto (TS, T, TI, NS, N, NI, G) sobre el que se calcula la asimetría, i indica el paciente y r y l hacen referencia al ojo derecho e izquierdo, respectivamente [3].

2.5. Preprocesado y estimador

Una vez extraídas las medidas de los grosores de la RNFL y aplicada la métrica de asimetría, los datos de cada paciente están preparados para el entrenamiento

y evaluación del estimador. Dicho estimador está compuesto por un preprocesado y un algoritmo de aprendizaje máquina. La Figura 2 muestra el flujo que se ha utilizado en los datos para el entrenamiento y validación.

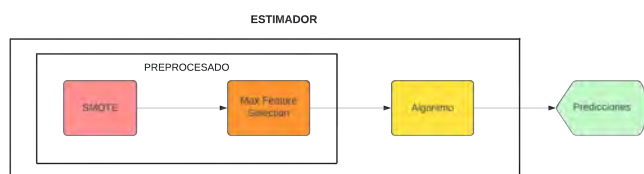


Figura 2. Flujo interno del estimador utilizado..

El flujo de ejecución es el siguiente. Inicialmente se parte de la base de datos original y se obtienen los valores de asimetría interocular de cada paciente para cada una de las métricas de asimetría propuestas. A partir de dichos valores se realiza una optimización en dos pasos del estimador completo, primero haciendo uso de una búsqueda aleatoria (*Random Search*) y posteriormente, realizando una búsqueda por cuadrícula (*Grid Search*) en un conjunto de valores definidos entorno a los mejores hiperparámetros del paso anterior. En todos los pasos del proceso se garantiza la independencia de la información del conjunto de datos de entrenamiento y de validación.

Debido al desbalanceo de la clase objetivo, donde solo hay 47 pacientes con glaucoma frente a los 160 pacientes sanos, es necesario aplicar técnicas de sobremuestreo de la clase minoritaria para realizar un balanceo de clases como SMOTE (*Synthetic Minority Over-sampling TEchnique*) [12] sobre el conjunto de entrenamiento. Además, se ha realizado una validación cruzada repetida estratificada y aleatoria (5 pliegues y 10 repeticiones), garantizando la distribución equitativa de clases y aumentando la calidad de los resultados. En adición a esta problemática, para reducir la dimensionalidad que evite el sobreajuste de los modelos utilizando un *Feature Selector*, se han seleccionado las características más importantes para la clasificación según el MDI (*Mean Decrease Impurity*) [13] obtenido de un Random Forest genérico.

3. Resultados

Esta sección presenta los mejores resultados obtenidos en ambos modelos junto con la métrica de asimetría asociada a dichos resultados, es decir, el mejor estimador basado en *Decision Tree* (DT) [14] y el mejor estimador basado en *Random Forest* (RF) [15]. El criterio de elección del mejor estimador para cada uno de los modelos utilizados se fundamenta principalmente en la métrica de clasificación conocida como *Recall*, que en este estudio denominamos *Sensitivity*.

A continuación se presenta en la Tabla 1 los mejores resultados obtenidos, mostrando el valor medio y dos veces la desviación típica, con el objetivo de tener en cuenta no solo el valor medio de dicha métrica de clasificación, sino también, la variabilidad de los resultados.

Por ello, el mejor resultado será el que tenga un mayor valor promedio de *Sensitivity* acompañado de la menor desviación típica.

ML Model	Decision Tree	Random Forest
Assymetry metric	$ \Delta $	$ \Delta ^{0,5}$
Sensitivity	0.810 ± 0.214	0.875 ± 0.154
Precision	0.619 ± 0.278	0.661 ± 0.187
Accuracy	0.822 ± 0.140	0.862 ± 0.098
Specificity	0.826 ± 0.208	0.858 ± 0.114

Tabla 1. Resultados de la validación cruzada de las métricas de asimetría destacadas para cada modelo y modalidad de imagen médica.

En primer lugar, podemos ver como las métricas de asimetría dominantes son distintas en cada modelo. Esto puede indicar que el rendimiento del modelo está ligado a la métrica de asimetría utilizada y que distintos modelos rinden mejor con distintas métricas de asimetría. Sin embargo, a la vista de la variabilidad de los resultados, este razonamiento aunque sea lógico carece de valor ya que no se puede demostrar. Por este motivo resulta interesante aumentar el número de muestras y reducir la varianza de los modelos pudiendo así ofrecer resultados más realistas y profundizar en dicha hipótesis.

En segundo lugar, se observa una mejora en el rendimiento de todas las métricas de clasificación al pasar de DT a RF tanto en promedio como en desviación típica. RF aumenta un 8% el valor promedio de *Sensitivity* y reduce su desviación típica en un 38%. La métrica de asimetría propuesta en este trabajo ($|\tau|$) no ha conseguido destacar en ninguno de los dos modelos.

En la Figura 3 se muestran las matrices de confusión de ambos estimadores. Estas se han obtenido con una validación cruzada aleatoria y estratificada de 5 pliegues donde en cada iteración se realizaba el entrenamiento y la posterior clasificación del conjunto de validación. Una vez completadas las iteraciones, mediante los identificadores de cada muestra se realizaba la comparativa entre la clasificación del modelo y la clasificación real para generar dicha matriz de confusión.

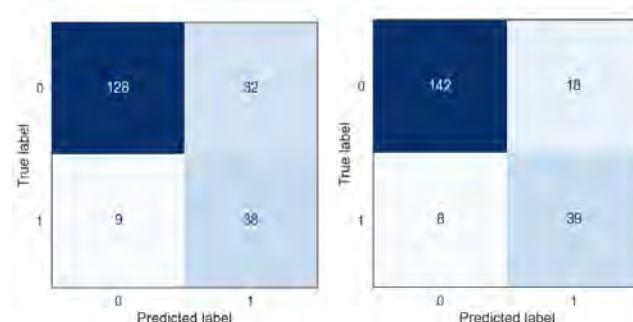


Figura 3. Matrices de confusión del mejor estimador de DT y RF.

En trabajos anteriores donde se ha utilizado la misma base de datos para entrenar y evaluar modelos basados en *Decision Trees* [3, 6], se han hallado los siguientes resultados *Sensitivity* 80.9%, *Specificity* 88.2%, *Precision* 66.7%, *Accuracy* 86.5%. Por otro lado, en [7] se obtienen los siguientes resultados en promedio para una base de datos distinta pero haciendo uso de *Random Forest*, 88.2% *Accuracy*, 91.7% *Sensitivity* y 85.1% *Specificity*.

Aunque no óptimos, los resultados obtenidos en este trabajo son favorables y evidencian que la asimetría en la morfología de la RNFL en ambos ojos junto con la ayuda del aprendizaje máquina son un potencial método de ayuda al diagnóstico precoz de glaucoma.

4. Conclusiones

Este estudio se enfocó en mejorar los métodos de ayuda al diagnóstico del glaucoma mediante el uso de técnicas de *Machine Learning*, utilizando una perspectiva diferente a la de investigaciones previas como [6] y [3]. Se emplearon diversas técnicas de análisis de datos y se llevaron a cabo en un entorno de desarrollo diferente.

El uso de la misma base de datos que los estudios anteriores destacó la importancia de las métricas de asimetría en la morfología de ambos ojos para diagnosticar el glaucoma. Sin embargo, este estudio logró resultados significativamente superiores a los previos, estableciéndose como un punto de referencia para investigaciones futuras en esta área.

La elección de árboles de decisión se basó en su sencillez y facilidad de interpretación, con la intención de consolidar y mejorar resultados anteriores. Además, se consideró el uso de *Random Forest* como una opción más compleja para mejorar aún más la precisión del diagnóstico.

Los resultados destacan la efectividad de los modelos basados en *Random Forest* como los principales métodos de apoyo en el diagnóstico de glaucoma.

La limitación de datos presentó un desafío, ya que intentar abordar el sesgo y la varianza simultáneamente resultaba complicado debido al compromiso sesgo-varianza. La implementación de *Random Forest* abordó este problema reduciendo la varianza, aunque sacrificando la interpretabilidad del modelo.

Para futuras mejoras, es esencial aumentar el tamaño de la muestra de datos disponible. También se pueden explorar estrategias de optimización de hiperparámetros y la inclusión de características adicionales para mejorar la representación de los datos. La investigación de modelos completamente nuevos, aunque menos interpretables, como *XGBoost* o *causalML*, podría conducir a avances significativos en el diagnóstico del glaucoma.

Referencias

- [1] Jesus Claramunt and Noelia Fernandez. Glaucoma y poblacion en 2020. *Gac Optom optica Oftalmica*, 2016.
- [2] Yih-Chung Tham, Xiang Li, Tien Y Wong, Harry A Quigley, Tin Aung, and Ching-Yu Cheng. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*, 121(11):2081–2090, 2014.
- [3] R. Berenguer, R. Verdu, J. Morales, I. Selles, O. Kovalyk, and J.L. Sancho. Decision trees for glaucoma screening based on the asymmetry of the RNFL in OCT. *Sensors*, 22(13), 2022.
- [4] Kwon YH, Figert JH, Kuehn MH, and Alward WL. Primary open-angle glaucoma. *N Engl J Med.*, 360(11):1113–1124, 2009.
- [5] Matthew G Field, Tarek Alasil, Neda Baniyasi, Christian Que, Huseyin Simavli, Doaa Sobeih, David Soladell Valle, Matthew J Best, and Teresa C Chen. Facilitating glaucoma diagnosis with intereye retinal nerve fiber layer asymmetry using Spectral-Domain OCT. *J Glaucoma*, 25(2):167–176, February 2016.
- [6] F. Rodriguez, R. Verdu, R. Berenguer, J. Morales, I. Selles. Analysis of the asymmetry between both eyes in early diagnosis of glaucoma combining features extracted from retinal images and octs into classification models. *Sensors*, 23(10), 2023.
- [7] Chao-Wei Wu, Hsiang-Li Shen, Chi-Jie Lu, Ssu-Han Chen, and Hsin-Yi Chen. Comparison of different machine learning classifiers for glaucoma diagnosis based on spectralis oct. *Diagnostics*, 11(9), 2021.
- [8] O. Kovalyk, J. Morales, R. Verdu, I. Selles, A. Palazon, and J.L. Sancho. Papila: Dataset with fundus images and clinical data of both eyes of the same patient for glaucoma assessment. *Scientific Data*, 9(1):1–12, 2022.
- [9] Eric A. Swanson and James G. Fujimoto. The ecosystem that powered the translation of oct from fundamental research to clinical and commercial impact. *Biomed. Opt. Express*, 8(3):1638–1664, Mar 2017.
- [10] David Huang, Eric A. Swanson, Charles P. Lin, Joel S. Schuman, William G. Stinson, Warren Chang, Michael R. Hee, Thomas Flotte, Kenton Gregory, Carmen A. Puliafito, and James G. Fujimoto. Optical coherence tomography. *Science*, 254(5035):1178–1181, 1991.
- [11] R. Berenguer, R. Verdú, J. Morales, I. Sellés, R. Del Amor, G. García, and V. Naranjo. Automatic segmentation of the retinal nerve fiber layer by means of mathematical morphology and deformable models in 2D OCT imaging. *Sensors*, 21(23):8027, 2021.
- [12] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.
- [13] Eugene Tuv, Alexander Borisov, George Runger, and Kari Torkkola. Feature selection with ensembles, artificial variables, and redundancy elimination. *The Journal of Machine Learning Research*, 10:1341–1366, 2009.
- [14] Leo Breiman, Jerome H Friedman, Richard A Olshen, and Charles J Stone. Classification and regression trees. *ISBN-13*, pages 978–0412048418, 1984.
- [15] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, Oct 2001.

Monitorización sin contacto en tiempo real: estudio piloto mediante vehículo aéreo no tripulado

L. Franco-Terriza¹, M. Elena², C. Romero-Olóriz³, M. L. Alcázar Laynez⁴, A. Barriga-Rivera⁵

¹ Asociación de Investigación y Cooperación Industrial de Andalucía (AICIA), Sevilla, España, lferriza19@gmail.com

² Departamento de Tecnología Electrónica, Universidad de Sevilla, Sevilla, España, maren@us.es

³ Centro de Emergencias Sanitarias 061 de Andalucía, Estepa, España, carlos.romero.epes@juntadeandalucia.es

⁴ Safran Electronics & Defense Spain S.L., Granada, España, manuel.a@fossa.systems

⁵ Departamento de Física Aplicada III, Universidad de Sevilla, Sevilla, España, abrivera@us.es

Resumen

Las cámaras de vídeo pueden ser empleadas como sensores para la monitorización sin contacto en tiempo real, aplicando sobre una secuencia de fotogramas distintos algoritmos de seguimiento y extracción de bioseñales. Las características de la cámara, como la resolución, la tasa de muestreo o las bandas espectrales de la luz susceptibles de captar, son determinantes en el rendimiento de los algoritmos. Este estudio pretende ser una prueba de concepto consistente en un vehículo aéreo no tripulado con cámara RGB integrada, cuyas imágenes captadas son procesadas por dos algoritmos, uno de detección de personas (YOLOv3) basado en redes neuronales y otro de medida de tasa cardiaca (Heartbeat) basado en fotoplestismografía remota. Las pruebas funcionales llevadas a cabo para validar el montaje, la detección y la medida de tasa cardiaca realizada por el prototipo muestran que la detección de personas sin contacto en tiempo real puede llevarse a cabo bajo las condiciones, dispositivos y algoritmos expuestos, pero la medida de la tasa cardiaca se ve afectada por la inestabilidad de la transmisión de fotogramas en tiempo real.

1. Introducción

De forma tradicional, las bioseñales del cuerpo humano se obtienen mediante el uso de distintos sensores y técnicas que requieren el contacto con el paciente y la asistencia de profesionales especializados. Con el avance de la tecnología y la medicina, numerosos estudios plantean la medición de bioseñales sin contacto en tiempo real [1]. Este nuevo enfoque ofrece numerosas posibilidades en entornos clínicos e incluso deportivos, ya que permite la monitorización remota de la población en diversos escenarios en los que el contacto no es posible. Así, este avance es de gran utilidad en situaciones de emergencia médica en las que la posibilidad de prestar una asistencia presencial es limitada, como es el caso de un accidente en un área de difícil acceso o en situaciones de aislamiento como la vivida durante la pandemia de COVID-19. Además, la monitorización sin contacto puede ser también de aplicación en salas de espera para discriminar casos de urgencia [2, 3].

Aunque existen distintas técnicas de monitorización de bioseñales sin contacto en tiempo real, muchas de ellas dependen de equipamiento especializado, como cámaras multispectrales, que utilizadas como sensores son capaces de medir algunas propiedades ondulatorias de la luz. Ciertas técnicas, como la fotoplestismografía remota (rPPG)

[4], son capaces de extraer parámetros como la frecuencia cardiaca, la presión arterial o el ritmo respiratorio. Esto se consigue midiendo la intensidad de luz reflejada por la piel en distintos instantes, que varía periódicamente a causa del cambio de volumen de los vasos sanguíneos. Estos cambios son más susceptibles de ser medibles en determinadas zonas del cuerpo, como la cara o el cuello.

En la actualidad existen diversos algoritmos validados para la medida de bioseñales mediante cámaras de vídeo, pero se requiere la previa identificación de la zona de interés (ROI). Para obtener la ROI puede ser útil la previa detección, identificación y seguimiento de las personas que aparezcan en la secuencia de imágenes. Por tanto, será útil implementar algoritmos de detección y seguimiento combinados con algoritmos de medida de parámetros biomédicos en un único dispositivo.

Como en cualquier aplicación tecnológica, las características del sensor determinarán la precisión y la calidad de los resultados. En el caso de emplear una cámara, la resolución de la imagen y la tasa de cuadro son parámetros fundamentales durante el procesado. También será determinante en los resultados la forma de programación de los algoritmos en función de si usan o no redes neuronales.

El objetivo de este estudio es realizar un prototipo capaz de detectar personas y medir su tasa cardiaca en tiempo real y sin contacto, a distancia. Para ello se ha llevado a cabo un estudio piloto basado en una cámara RGB, que tienen la limitación de medir un menor número de bandas espectrales como son los tres colores primarios de la luz visible (rojo, verde y azul), pero por otro lado tienen un menor coste y mayor resolución que las cámaras multispectrales. La cámara RGB usada se encuentra integrada en un vehículo aéreo no tripulado (VANT), por ello, la cámara podrá desplazarse en el espacio y variar así la distancia entre la misma y las personas que se encuentran en la escena. El vídeo se ha captado como una secuencia de fotogramas en tiempo real sobre las que se han ejecutado algoritmos de detección de personas, YOLOv3, y medida de tasa cardiaca, Heartbeat. Mediante distintas pruebas de funcionamiento se ha comprobado la validez del prototipo como una primera aproximación de un VANT con cámara integrada con la capacidad de detectar personas y medir bioseñales tanto en estacionamiento como en vuelo.

2. Materiales y métodos

Este estudio ha sido aprobado por el Comité de Ética de Investigación de la universidad de Sevilla. Todos los participantes han sido debidamente informados y han firmado el consentimiento informado previamente a su participación en este estudio. El VANT ha sido manejado durante el vuelo por uno de los investigadores, cumpliendo todas las condiciones de seguridad. Dadas las características del estudio se engloba dentro de las operaciones de bajo riesgo, en la categoría “Abierta” establecida por la Agencia Estatal de Seguridad Aérea (AESA) [5], por lo que no es requerida ningún tipo de licencia para el manejo del VANT.

2.1. Sistema de monitorización

Para la realización de las pruebas se ha empleado un VANT Mi Drone 4k (Xiaomi, Haidian District, Beijing, China) con cámara integrada, conectado mediante radiofrecuencia a un mando de control a distancia. A su vez, este se conecta mediante Wi-Fi a un smartphone Redmi Note 7 (Xiaomi, Beijing, China) y a un ordenador portátil Legion Y540-15IRH (Lenovo, Beijing, China) con Intel Core i7-9750H, 16GB RAM, Nvidia RTX2060-6GB (Intel, Quezon City, Philippines) a los que retransmite video en directo. La cámara integrada en el VANT consta de un sensor RGB con resolución de 1920 x 1080 píxeles², con una frecuencia de adquisición de 60 fotogramas por segundo. La cámara también cuenta con un gimbal y estabilizadores de imagen. La configuración de los dispositivos se muestra en la Figura 1.

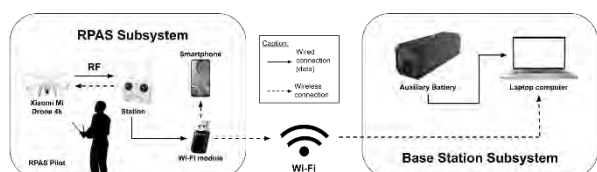


Figura 1. Conexión entre los distintos dispositivos. El dron se conecta por radiofrecuencia a una emisora, y la emisora se conecta por Wi-Fi a un smartphone; junto al piloto conforman un subsistema. Este subsistema está unido mediante Wi-Fi al subsistema formado por el portátil encargado del procesamiento de la imagen, sustentado por una batería auxiliar.

2.2. Procesado digital de video

Para llevar a cabo la detección, seguimiento e identificación de las RROOII en el cuerpo así como la medida de distintas bioseñales, se usan algoritmos de procesamiento digital de video. Para ejecutarlos se ha utilizado un sistema operativo de código abierto, Ubuntu 19.04. Entre distintos algoritmos de código abierto, se han priorizado aquellos con menor coste computacional, menor tiempo de respuesta y mayor precisión media.

Los algoritmos de detección han sido entrenados usando redes neuronales, que a partir de una imagen de entrada son capaces de ofrecer información específica de la imagen como salida. YOLOv3 destaca ante otros algoritmos de detección gracias a sus prestaciones, como se determinó en su proceso de validación, lo que lo convierte en una buena

elección para la detección de personas y objetos (valor medio de precisión 28’2, tiempo de respuesta aproximado de 20 ms, resolución 320 px²) [6]. La secuencia de imágenes de salida de este algoritmo sirve como entrada a otros algoritmos que realizan un seguimiento sobre cada persona detectada, como deepSort (valor medio de precisión 61’4, frecuencia de ejecución aproximada de 60 Hz) [7], y de estimación de la pose, como OpenPose (valor medio de precisión 61’8, tiempo de detección aproximado de 0’58 ms) [8]. Dada la correspondencia que se establece entre la salida de YOLOv3 y las entradas de deepSort y OpenPose, asegurar el correcto funcionamiento del primero es imprescindible antes de pasar a la etapa de seguimiento y detección de pose.

Para la medida de bioseñales se ha optado por monitorizar la tasa cardiaca mediante el algoritmo Heartbeat, basado en la técnica de la fotopleletismografía remota (rPPG), debido a su bajo coste computacional frente a otros algoritmos que implementan el uso de inteligencia artificial, más costosa computacionalmente. Heartbeat necesita una secuencia de video como entrada para ofrecer como salida el valor de la tasa cardiaca. Este algoritmo fue validado por Rouast et al. [9] empleando una cámara webcam (Logitech C270) a una resolución de 640x640 px² y 30 fps, colocándose a 0’5 m de distancia de los participantes.

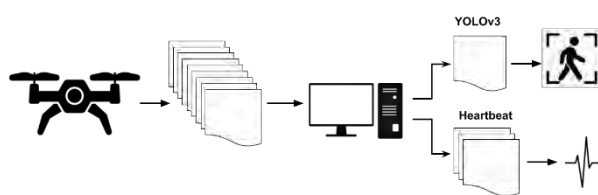


Figura 2. Flujo del proceso completo. La cámara del dron toma una secuencia de fotografías, que es procesada por un ordenador. YOLOv3 trabaja sobre cada fotografía de forma independiente mientras Heartbeat trabaja con una secuencia de fotografías.

2.3. Evaluación experimental

Los algoritmos seleccionados para la medida de bioseñales sin contacto y a tiempo real a través de la cámara del dron han sido probados en una muestra de n=3 sujetos (dos hombres de edad 18 y 25 años, y una mujer de 49 años).

Se ha diseñado un primer bloque de pruebas para validar el montaje del sistema (F1-F2). Estas pruebas consisten en comprobar la autonomía, estabilidad y recepción de video en directo durante un vuelo del dron. El segundo bloque de pruebas consiste en llevar a cabo la detección de personas y medir de frecuencia cardiaca (F3-F4), mediante la ejecución de YOLOv3 y Heartbeat sobre video tomado en tiempo real, y que ha sido almacenado para su posterior análisis.

Las pruebas llevadas a cabo se recogen en la Tabla 1, clasificadas en los cuatro grupos mencionados.

Código	Tipo	Prueba
--------	------	--------

F1	Vuelo con cámara a bordo	1.1. Vuelo de prueba 1.2. Prueba de autonomía 1.3. Prueba de gimbal
F2	Recepción de vídeo en directo	2.1. Desde smartphone 2.2. Desde PC
F3	Detección de personas	3.1. En estacionamiento 3.2. En vuelo
F4	Captación de la tasa cardíaca	4.1. En estacionamiento 3.2. En vuelo

Tabla 1. Matriz de pruebas funcionales.

2.3.1. Evaluación experimental de YOLOv3

Para comprobar el funcionamiento del algoritmo de detección de personas dentro de nuestro sistema, se va a analizar una muestra de video de 4 segundos (240 fotogramas) tomada con el dron estacionado, en la que aparecen los tres sujetos en movimiento, y sobre otra muestra de las mismas características con el dron en vuelo. Se van a recoger el número de detecciones realizadas sobre las muestras de video seleccionada en una matriz de confusión. A partir de los datos de esta matriz se podrá calcular la sensibilidad, la precisión y la tasa de falsos descubrimientos.

2.3.2. Evaluación experimental de Heartbeat

El algoritmo de medición de la tasa cardíaca ha sido probado sobre video en directo de cada uno de los sujetos de estudio, por lo que se va a trabajar con tres muestras de video en estacionamiento y otras tres en vuelo de un minuto de duración. Los valores de salida del algoritmo van a ser comparados con los valores aportados por un tensiómetro doméstico, Omron M3 Comfort Digital Tensiometer HEM-7155-E (Omron Corporate, Kyoto, Japan), clasificado como dispositivo médico.

3. Resultados

Las pruebas de validación del montaje, F1 y F2, fueron calificadas como exitosas. Una vez verificado el montaje, fue posible llevar a cabo las pruebas descritas para los algoritmos.

3.1. Detección de personas

En la Tabla 2 se recogen los resultados de las pruebas realizadas con YOLOv3 (F3) en forma de matriz de confusión, tanto en estacionamiento como en vuelo. En base a estos datos se ha calculado la sensibilidad, que indica el número de detecciones correctas en relación con el número total de personas susceptibles de ser detectadas (83.97% en estacionamiento frente a 83.54% en vuelo), la precisión, referida al número de detecciones correctas frente al total de detecciones de personas realizadas por el algoritmo (97.87% en estacionamiento frente a 98.38% en

vuelo) y la tasa de falsos descubrimientos, complementaria a la precisión, que mide el número de detecciones erróneas sobre el total de detecciones (2.13% en estacionamiento frente a 1.62% en vuelo).

	VP	FN	FP
Estacionado (E)	597	114	13
Vuelo (V)	548	108	9

Tabla 2. Matriz de confusión de los resultados de las pruebas de YOLOv3 con el dron estacionado y en vuelo.

Un ejemplo de la salida del algoritmo YOLOv3 para la detección de personas se muestra en la Figura 2.

3.2. Detección de la tasa cardíaca

En las pruebas de captación de tasa cardíaca en tiempo real, F4, se detectaron diferentes problemas en las modalidades de estacionamiento en vuelo. Cuando el dron estaba estacionado, la salida del algoritmo una vez estabilizado mostraba un valor constante y sin apenas variación de 60 latidos por minuto en todos los participantes, mientras el tensiómetro indicaba valores variables en el rango de 70 a 90 ppm como se observa en la Figura 3. Con el dron en vuelo, si bien la tasa cardíaca mostraba variación, este valor estaba muy lejos de la referencia.

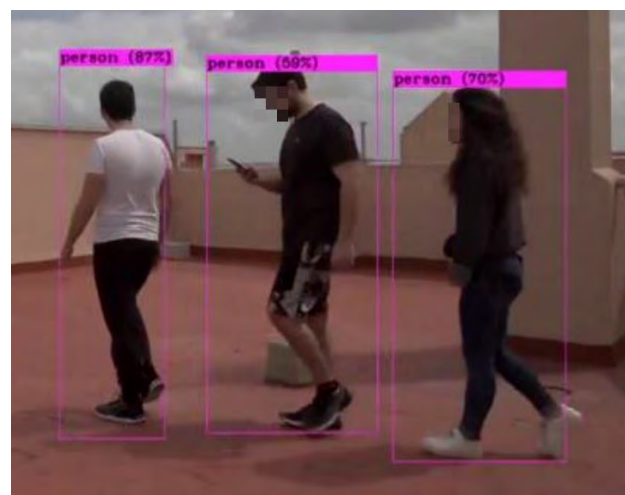


Figura 3. Detección de personas con YOLOv3. En la imagen se observa un fotograma en el que aparecen tres sujetos correctamente detectados, clasificados como verdaderos positivos (VP).



Figura 4. Comparación entre la salida del algoritmo Heartbeat en tiempo real con el dron estacionado (derecha, 59.5 LPM) frente a la salida del tensiómetro (izquierda, 87 LPM).

4. Discusión

En esta prueba de concepto se han llevado a cabo una serie de pruebas de funcionamiento utilizando una cámara integrada en un dron con algoritmos de procesamiento en tiempo real para la medida de bioseñales sin contacto. En primer lugar, se ha puesto a prueba el sistema diseñado para garantizar su estabilidad y autonomía, y a continuación se han puesto a prueba dos algoritmos, uno de detección de personas basado en redes neuronales (YOLOv3), y otro de medición de tasa cardíaca, basado en la fotoplethysmografía remota (Heartbeat).

Las pruebas de los bloques F1 y F2 han verificado el montaje con éxito y garantizado condiciones óptimas para la realización de las siguientes pruebas.

Las pruebas del bloque F3 arrojan valores superiores al 83% para la sensibilidad, superiores al 97% para la precisión e inferiores al 3% para la tasa de falsos descubrimientos, difiriendo en menos de un 1% al cambiar de la modalidad de estacionamiento a vuelo. Aunque solo se ha realizado una prueba, lo que no es suficiente para asegurar la detección de personas al variar las condiciones, estos valores son suficientemente buenos para validar una primera aproximación del sistema.

En cuanto a las pruebas del bloque F4, no se han conseguido resultados válidos del algoritmo Heartbeat en las condiciones establecidas. Se ha observado que este fallo de funcionamiento se debe, principalmente, a la inestabilidad de la frecuencia de transmisión de datos desde el dispositivo hasta el PC mediante conexión inalámbrica. Esto no ocurre en el algoritmo de seguimiento porque es aplicado a cada fotograma, mientras la tasa cardíaca requiere una secuencia de fotogramas con una tasa de transmisión constante para ser obtenida. Este problema puede solucionarse en las siguientes etapas de desarrollo mediante una pila que almacene los fotogramas de forma que se aplique un retraso de varios segundos, suficiente para ejecutar el algoritmo sobre la secuencia de fotogramas sin pérdidas. Si el problema persiste, puede ser posible que sean necesarias algunas mejoras en el algoritmo Heartbeat para conseguir que funcione en distintos escenarios de validación.

Agradecimientos

Queremos agradecer al Centro de Emergencias Sanitarias 061 de Andalucía por su apoyo en este proyecto.

Referencias

- [1] Molinaro, N., Schena, E., Silvestri, S., Bonotti, F., Aguzzi, D., Viola, E., ... & Massaroni, C. (2022). Contactless vital signs monitoring from videos recorded with digital cameras: an overview. *Frontiers in Physiology*, 13, 160.
- [2] Massaroni, C., Schena, E., Silvestri, S., Taffoni, F., & Merone, M. (2018, June). Measurement system based on RGB camera signal for contactless breathing pattern and respiratory rate monitoring. In *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)* (pp. 1-6). IEEE.
- [3] Yang, F., He, S., Sadanand, S., Yusuf, A., & Bolic, M. (2022). Contactless measurement of vital signs using thermal and RGB cameras: A study of COVID 19-related health monitoring. *Sensors*, 22(2), 627.
- [4] Rouast, P. V., Adam, M. P., Dörner, V., & Lux, E. (2016). Remote photoplethysmography: Evaluation of contactless heart rate measurement in an information systems setting. *Applied Informatics and Technology Innovation*.
- [5] Easy Access Rules for Unmanned Aircraft Systems | EASA. 2022.
- [6] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement. 2018.
- [7] Wojke N, Bewley A, Paulus D. Simple online and realtime tracking with a deep association metric. 2017 IEEE International Conference on Image Processing (ICIP): IEEE; 2017.
- [8] Cao Z, Hidalgo G, Simon T, Wei S-E, Sheikh Y. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2021;43:172-86.
- [9] Rouast PV AM, Dörner V, Lux E. Remote Photoplethysmography: Evaluation of Contactless Heart Rate Measurement in an Information Systems Setting. *Applied Informatics and Technology Innovation Conference*; 2016. p. 1-17.

Análisis del Color, Textura y Temperatura para el Diagnóstico y Clasificación de Hemangiomas Infantiles

J. A. Pérez-Carrasco^{1,*}, B. Acha¹, J. A. Leñero², J. Bernabéu-Wittel³, C. Serrano¹

¹ Dpto de Teoría de la Señal y Comunicaciones, Universidad de Sevilla, Sevilla España, {jperez2, bacha, cserrano}@us.es

² Departamento de Electrónica y Electromagnetismo, Universidad de Sevilla, Sevilla, España, jlenero@us.es.

³ Unidad de Dermatología. Hosp. Univ. Virgen del Rocío, Sevilla, España, jbernabeuw@gmail.com

Resumen

Los hemangiomas infantiles se encuentran entre los tumores benignos más frecuentes en la infancia. Para predecir el curso clínico y aconsejar sobre el tratamiento, es fundamental determinar su ubicación, extensión, volumen y profundidad. Sin embargo, esto puede ser bastante complicado, ya que los hemangiomas pueden estar cubiertos por cabello y/o presentar patrones irregulares. El diagnóstico habitual de dichas lesiones se basa en la estimación visual de área, textura y color. En casos complicados puede ser necesario el uso de técnicas complejas como ultrasonido Doppler o resonancia magnética. En este trabajo presentamos una técnica alternativa de diagnóstico basada en el uso de la temperatura (mediante imagen termográfica) y de parámetros obtenidos de color y textura (a partir de una cámara en el espectro visible) para ofrecer un examen preciso de los hemangiomas que facilite al equipo clínico medidas de localización, extensión, color y profundidad.

En trabajos anteriores ya demostramos que debido a la alteración local de temperatura que producen los hemangiomas se puede estimar la localización y extensión del hemangioma. En este trabajo además incorporamos parámetros de color y textura obtenidos a partir de la imagen visible que permitirán dar información sobre la apariencia y profundidad de la lesión, parámetros que son muy importantes para la clasificación y el correcto tratamiento.

En este trabajo se han analizado los hemangiomas correspondientes a 53 pacientes y se han obtenido parámetros de color y textura para la clasificación en las tres categorías principales: superficial, mixto o profundo. Los resultados muestran una exactitud de 0.79.

1. Motivación

Los hemangiomas son los tumores benignos más frecuentes en la infancia. Representan neoplasias vasculares benignas que crecen mediante una rápida proliferación celular a partir de células endoteliales. La incidencia de hemangioma infantil (HI) se estima en 1-3% de todos los recién nacidos, mientras que llega al 10% en bebés menores de 1 año [1].

Para predecir el curso clínico y aconsejar sobre el tratamiento, es fundamental determinar con precisión la extensión, el volumen, la ubicación y la profundidad de los hemangiomas. Sin embargo, esto puede representar un desafío porque los hemangiomas pueden presentar patrones irregulares o estar cubiertos de pelo, o su profundidad puede ser difícil de estimar. El diagnóstico se realiza habitualmente mediante inspección clínica y

palpación, y los médicos basan su diagnóstico en características visuales como el área, la textura y el color.

Según su profundidad, los hemangiomas pueden ser superficiales, profundos o mixtos (con componentes superficiales y profundas). La gran mayoría de los hemangiomas infantiles son diagnosticados clínicamente [1-3].

En este sentido, resulta interesante desarrollar una herramienta de Diagnóstico Asistido por Computadora (CAD) para ayudar a los médicos a evaluar el diagnóstico y la evolución de los HI. Si este sistema CAD emula la evaluación de los médicos y extrae información de color y textura para estimar el color y la profundidad, se debería alcanzar un diagnóstico automático explicable, lo que haría que el sistema fuera más atractivo para los médicos [4].

Son pocos los trabajos que se han dedicado a analizar la extensión y profundidad de los hemangiomas. Entre ellos se pueden destacar los trabajos de Abage [5], donde sólo se analiza la evolución del color del HI y el tamaño, pero no la profundidad, textura o volumen; el trabajo de O'Brien [6], que utiliza características colorimétricas para distinguir entre HI y marcas de nacimiento en vino de Oporto en bebés menores de 3 meses utilizando colorímetros electrónicos, que no son fáciles de conseguir, y, finalmente, el trabajo de Sultana [7], para la evolución del hemangioma basado en color y características de área.

En este trabajo se pretende mostrar una técnica alternativa de diagnóstico basada en la termografía (como herramienta útil e inmediata para una localización precisa de los hemangiomas) y en el análisis de características de color y textura para obtener un diagnóstico y clasificación de la lesión. En anteriores trabajos hemos demostrado cómo la termografía sirve para localizar y analizar la evolución de un hemangioma analizando los valores y variaciones locales de temperatura en las lesiones infantiles y cómo la herramienta puede convertirse en técnica habitual de ayuda al diagnóstico debido a su bajo coste y fácil manejo [8-10]. Los componentes del dispositivo implementado por el equipo de este trabajo, así como el procedimiento para estimar la temperatura interior y exterior de las lesiones se detallan con mayor profundidad en los trabajos [8,10]. Para avanzar en el diagnóstico y clasificación de dichas lesiones en este trabajo se amplía la herramienta y se incorpora un análisis de color y textura de las lesiones a partir de la imagen clínica para clasificar en los principales tipos de hemangiomas infantiles: superficial, mixto y profundo.

2. Método

Para este trabajo, 53 pacientes con al menos un hemangioma infantil fueron examinados en su primera consulta de dermatología proporcionando un total de 74 imágenes. El seguimiento de estos pacientes ha sido de dos años. Para todos ellos, se tomaron dos imágenes diferentes (imágenes clínicas e infrarrojas) que fueron adquiridas simultáneamente con un protocolo específico: a una distancia fija de 30 cm y en las mismas condiciones de iluminación y temperatura ambiente.

El sistema portátil diseñado por el equipo de autores de este trabajo dispone principalmente de dos sensores de imagen, una pantalla táctil, un procesador Raspberry Pi IV y un termómetro [8-10]. Más detalles sobre los componentes del sistema portátil y del procedimiento para obtener la temperatura interior y exterior de las lesiones se encuentran en los trabajos previos [8,10].

El dispositivo se muestra en la Figura 1. Uno de los sensores es una cámara termográfica que detecta radiación en la banda de infrarrojos de onda larga (LWIR). El otro sensor es una cámara digital estándar que opera en el espectro de luz visible. El primer sensor, que opera en la banda infrarroja, será útil para la localización y el seguimiento de los hemangiomas en los pacientes infantiles, y el otro sensor, que opera en el espectro visible, será útil para la obtención de parámetros de color y textura, parámetros clave para la estimación del color y profundidad de la lesión.

La Figura 2 muestra la imagen térmica y visible de un caso de hemangioma utilizando el sistema de diagnóstico descrito. La imagen térmica muestra con mayor temperatura la zona afectada.



Figura 1. Sistema portátil utilizado. Implementación detallada del sistema. (a) Vista frontal. (b) Vista posterior. Tiene dos cámaras que funcionan simultáneamente. Uno opera en el espectro visible y la otra en la banda infrarroja.



Figura 2. Estudio de la temperatura de hemangioma infantil en dos casos diferentes. La imagen superior muestra una instantánea del hemangioma adquirida con la cámara correspondiente al espectro visible. La imagen inferior corresponde a una imagen termográfica del hemangioma que muestran valores absolutos de temperatura.

El método propuesto en este trabajo consta de dos etapas. En la primera etapa, se utiliza la imagen térmica, que proporciona la variación de temperatura entre el centro de la lesión y su piel sana circundante. Esta variación de temperatura sirve para la clasificación de la lesión como hemangioma y para conocer la localización y extensión de la misma [8-10].

En la segunda etapa del método, y una vez localizada la lesión y su extensión mediante la cámara térmica, se procede a analizar la lesión en la imagen visible para proceder a su clasificación en tres categorías: hemangioma infantil superficial, mixto o profundo.

En la Figura 3 se muestran imágenes correspondientes a los tres tipos de lesiones.

En esta segunda etapa, se hace un recorte de la zona lesionada en la imagen visible y se extraen diferentes características matemáticas de color y textura. El conjunto empleado de descriptores está formado por parámetros de textura de primer orden y parámetros extraídos de las tres coordenadas de los espacios de color RGB y el espacio de color $L^*a^*b^*$, así como mediciones de tono y croma derivadas de ellos [11]. Más específicamente, los descriptores elegidos son: media de cada plano de color R , G y B , media de luminosidad (μ_L), media de tono (μ_H), media de croma (μ_C), desviación estándar de luminosidad

(σ_L), desviación estándar de tono (σ_h), desviación estándar de croma (σ_c), media de a^* (μ_a), media de b^* (μ_b), desviación estándar de a^* (σ_a), desviación estándar de b^* (σ_b), torsión de luminosidad (s_L), curtosis de luminosidad (k_L), torsión de a^* (s_a), curtosis de a^* (k_a), torsión de b^* (s_b), y curtosis de b^* (k_b), En total el número de descriptores es 19.

Para reducir la dimensionalidad del vector de características se han empleado técnicas de selección de características hacia delante (*SFS*) y hacia detrás (*SBS*) de modo que el número final de características quedó reducido a 8 y son las siguientes:

Características elegidas: [$\mu_h \mu_c \sigma_L \mu_a \mu_b \sigma_a \sigma_b s_L$]

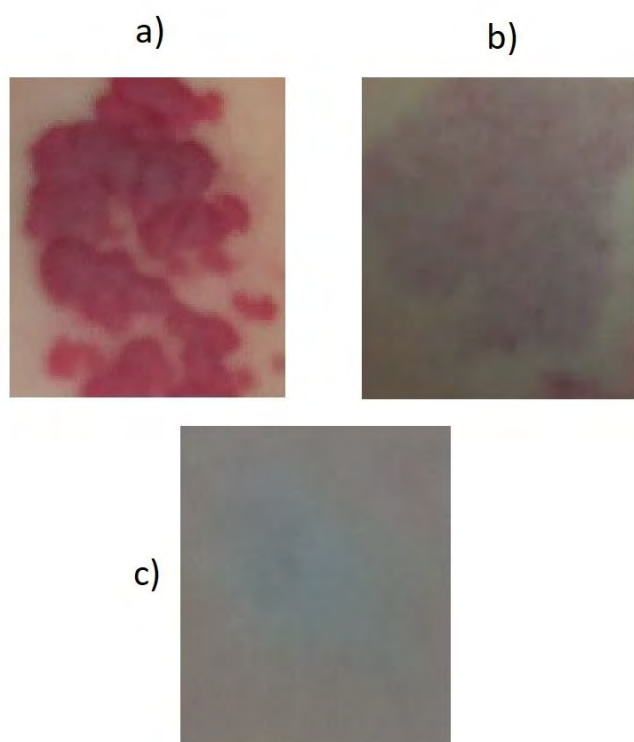


Figura 3. Tipos de hemangiomas infantiles analizados en este trabajo. a) Superficial, b) Mixto, c) Profundo. La categoría mixta incluye componentes superficiales y profundas.

3. Resultados Experimentales

El número total de hemangiomas infantiles utilizados en este trabajo ha sido 74. De ellos, 40 corresponden a hemangiomas superficiales, 24 a hemangiomas mixtos y 10 a hemangiomas profundos.

En anteriores trabajos [8-10] se analizó la temperatura y la extensión de dichos hemangiomas infantiles en los diferentes pacientes en las distintas sesiones de tratamiento mediante el uso de la cámara termográfica. En dichos pacientes se observó que los tratamientos que habían sido prescritos conllevaban una reducción del área de la región afectada. Este hecho quedó corroborado mediante la cámara termográfica, que mostraba un área menor y una reducción de la temperatura entre la región afectada y no afectada.

El poder de clasificación de las características de color y textura indicadas anteriormente ha sido evaluado mediante dos técnicas diferentes: Clasificación mediante máquinas de vector soporte (*SVM*) con núcleo gaussiano, utilizando una validación cruzada con $k=10$. y clasificación utilizando un perceptrón con una capa oculta (*MLP*) con 32 nodos, con función de activación sigmoide para la última etapa. Además, se ha utilizado el optimizador *Adam* y la entropía cruzada categórica como función de error. Para esta clasificación se ha utilizado el 70% de las imágenes para entrenamiento, 15% para validación y 15% para test. Para la clasificación mediante SVM se realizó una selección de características mediante las técnicas de selección hacia delante (*SFS*) y hacia detrás (*SBS*) para reducir el número de características finales a 8 como se indicó anteriormente.

En la Tabla 1 se muestran los parámetros de exactitud, sensibilidad y especificidad obtenidos para la clasificación mediante SVM y perceptrón en 3 clases de hemangiomas: superficial, mixto y profundo y, también en 2 clases: superficial y mixto. Se ha optado por realizar también esta clasificación en dos clases debido al reducido número de imágenes disponibles actualmente de hemangioma profundo. Se puede observar que cuando la clasificación se realiza en 3 clases la especificidad es bastante baja. Ello se debe a que las clases *HI mixto* y *profundo* se confunden debido principalmente, al escaso número de imágenes de tipo *profundo* disponibles para el entrenamiento y clasificación, y a que el tipo *mixto* contiene componentes de tipo profundo. Para la realización de la Tabla 1 y la clasificación en 3 clases se han calculado las distintas métricas (exactitud, sensibilidad y especificidad) por cada clase y se han promediado.

Las métricas elegidas han sido obtenidas a partir de los valores de verdaderos positivos (*TP*), verdaderos negativos (*TN*), falsos positivos (*FP*) y falsos negativos (*FN*) del siguiente modo:

$$Exactitud = (TN+TP)/(TN+TP+FN+FP) \tag{1}$$

$$Sensibilidad = TP/(TP+FN) \tag{2}$$

$$Especificidad = TN/(TN+FP) \tag{3}$$

Clasificador	Nº clases	3 clases	2 clases
SVM	Exactitud	0.79	0.81
	Sensibilidad	0.75	0.78
	Especificidad	0.44	0.85
MLP	Exactitud	0.75	0.78
	Sensibilidad	0.7	0.75
	Especificidad	0.43	0.82

Tabla 1. Clasificación mediante SVM y perceptrón

4. Conclusiones y líneas futuras

En este trabajo se ha mostrado cómo es posible utilizar un dispositivo de bajo costo compuesto de 2 cámaras (una infrarroja y otra para el espectro visible) para localizar, conocer la extensión y clasificar una lesión de hemangioma infantil. El objetivo de este trabajo es doble, por un lado ayudar al equipo clínico en el diagnóstico, localización y extensión del hemangioma (gracias a la imagen térmica) y, por otro lado, ayudar en la clasificación del hemangioma en uno de los tres subtipos (utilizando la imagen visible) a través de las características de color y textura que ayudan a determinar el aspecto y la profundidad de la lesión, parámetros clave para el equipo clínico.

Los resultados experimentales realizados con 53 pacientes confirman los diagnósticos clínicos en cuanto a localización y extensión del hemangioma [8-10] y se acercan bastante (0.8 de exactitud aproximadamente) a la clasificación de la lesión por el equipo clínico. Hay que hacer constar que el número disponible de imágenes es reducido, aunque se espera poder aumentarlo a medida que se siga haciendo seguimiento a nuevos pacientes. Esto permitirá obtener mejores valores en las métricas calculadas.

El dispositivo descrito en este trabajo, basado en termografía, es de bajo coste y sencillo de usar, y puede coexistir con otras técnicas auxiliares de diagnóstico, como las basadas en análisis de ultrasonido Doppler o resonancia magnética. El análisis realizado es inmediato, no invasivo y puede proporcionar resultados en tiempo real. Todas estas características lo hacen un candidato viable para ser usado por equipos clínicos en centros públicos de salud.

Como punto muy importante, hay que destacar la novedad de este trabajo, ya que este problema no ha sido abordado de manera exhaustiva por la comunidad científica y que, por tanto, hace que no haya bases de datos públicas disponibles para poder incluirlas en nuestro estudio. De hecho, el uso de técnicas de clasificación más complejas como las basadas en redes neuronales profundas (DNN) no pueden ser abordado aún debido al reducido número de imágenes.

En el futuro cercano se pretende incluir el código de clasificación en el dispositivo (entorno Python) así como obtener imágenes de mayor resolución que permitan obtener además de la clasificación medidas de área de superficie afectada y análisis más cuantitativos utilizando parámetros obtenidos del análisis del histograma. De este modo el personal clínico podrá obtener en tiempo real estimaciones reales de la superficie afectada y un diagnóstico y clasificación de la lesión sin necesidad de utilizar otras técnicas invasivas o de elevado coste, y que retrasan de manera considerable la duración del proceso y suponen un coste elevado para el sistema de salud.

Agradecimientos

Este trabajo ha sido realizado gracias a la financiación obtenida mediante los siguientes proyectos e instituciones: Proyectos de Generación de Conocimiento 2021 del

Ministerio de Ciencia e Innovación (PID2021-127871OB-I00) y Proyecto de Excelencia de la Junta de Andalucía en el ámbito del Plan Andaluz de Investigación, Desarrollo e Innovación PAIDI 2021 (PROYEXCEL_00889).

Referencias

- [1] Puig-Sanz L. Lesiones vasculares: angiomas. Disponible en: https://www.nasajpg.com/2015/01/22/angiomas/#google_vignette (Consultada: Septiembre 2023).
- [2] Moholkar S, Sawhney JS, Bhatt R. Imaging benign soft tissue lesions of the foot. *Appl Radiol*, vol 38, sup 10, 2009, pp 10–21.
- [3] Waldt S, Rechl H, Rummeny EJ, Woertler K. Imaging of benign and malignant soft tissue masses of the foot. *Eur Radiol*, vol 13, sup 5, 2003, pp 1125–36.
- [4] Van der Velden BHM, Kuijff-Kenneth HJ, Gilhuijs GA, Viergever MA. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Medical Image Analysis*, vol 70, 2022. DOI: <https://doi.org/10.1016/j.media.2022.102470>.
- [5] Abage KT, Sandri CO, Sakai LSM, Mussato LP, Pettele RR, Oliveira de Carvalho VO. Colorimetric analysis of images in the follow-up of infantile hemangiomas. *Pediatr Dermatol*, vol 37, sup 4, 2020, pp 622-625. doi:10.1111/pde.14153.
- [6] O'Brien KF, Frieden IJ, Zeymo A, Vasic J, Silverman R, Goldberg G, Carver DeKlotz CM. Analysis of lesional color to differentiate infantile hemangiomas from port-wine birthmarks in infants less than 3 months old: A pilot study. *Pediatr Dermatol*, vol 38, sup 3, 2021, pp 585-590. doi:10.1111/pde.14554.
- [7] Sultana A., Oprisescu S., Ciuc M. Automatic Evaluation of Hemangiomas for Follow-up Monitoring. *IEEE International Conference on E-Health and Bioengineering (EHB 2015)*, Romania, 2015, pp 1-4. doi:10.1109/EHB.2015.7391579.
- [8] Leñero-Bardallo JA, Acha B, Serrano C, Pérez-Carrasco JA, Ortiz-Álvarez J, Bernabéu-Wittel J. Thermography as a Method for Bedside Monitoring of Infantile Hemangiomas. *Cancers (Basel)*, vol 14, sup 21, 2022, Nov 1;14(21):5392. doi:10.3390/cancers14215392.
- [9] Leñero-Bardallo JA, Pérez-Carrasco JA, Serrano, C, Acha B, Ortiz-Álvarez J, Bernabéu-Wittel J. Uso de la Termografía para el Análisis y Seguimiento de los Hemangiomas Infantiles, *Actas del XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2022)*, Valladolid, 2022, pp 18-21. (ISBN: 978-84-09-45972-8).
- [10] Leñero-Bardallo JA, Serrano C, Acha B, Pérez-Carrasco, JA, Bernabeu-Wittel J. Thermography for the differential diagnosis of vascular malformations. *Clinical and Experimental Dermatology*. vol 1, sup 1, 2020, pp 1-5. <https://doi.org/10.1111/ced.14346>
- [11] Acha B, Serrano C, Acha JI, Roa LM. CAD tool for burn diagnosis. *Inf Process Med Imaging*, vol 18, 2003, pp 294-305. doi:10.1007/978-3-540-45087-0_25.

Concordancia entre especialistas. ¿Cómo podemos saber si tenemos una GT representativa para entrenar una red neuronal? Un estudio centrado en el Carcinoma Basocelular (BCC)

I. Matas González¹, B. Acha Piñero¹, F. Silva², A. Serrano Gotarredona², T. Toledo Pastrana³,
C. Serrano Gotarredona¹

¹ Universidad de Sevilla. Sevilla. {imatas, bacha, cserrano}@us.es.

² Servicio de Dermatología. Hospital Universitario Virgen Macarena. Sevilla

³ Servicio de Dermatología. Hospitales Quironsalud Infanta Luisa y Sagrado Corazón. Sevilla

Resumen

El cáncer de piel no melanoma más común es el carcinoma basocelular, BCC, del inglés basal cell carcinoma, con una incidencia entorno al 70% con unos criterios muy objetivos. Tener una verdad de referencia (GT, del inglés Ground Truth) adecuada es algo primordial a la hora de entrenar un modelo de Inteligencia Artificial (IA).

En este estudio se pretende analizar la fiabilidad de la verdad de referencia de los patrones BCC presentes en una lesión ya que estos patrones dermatoscópicos no pueden ser validados por biopsia.

Se recogieron los diagnósticos de los criterios BCC de 204 lesiones de cuatro dermatólogos distintos. Con ello se realizó un estudio de concordancia entre evaluadores haciendo uso de la Distancia de Hamming y de los coeficientes Cohen Kappa y Fleiss Kappa.

De acuerdo con los valores de kappa obtenidos, se obtuvo un acuerdo moderado entre los dermatólogos 1 a 3 para los patrones Pigment Network, Ulceration y Arborizing Telangiectasia. Pero para los patrones Maple Leaf, Multiple Blue-gray Globules y Blue-gray Ovoid Nest se obtuvo un acuerdo en torno a 0.5 El acuerdo entre el dermatólogo 4 y los demás fue menor que 0.4 para todos los patrones.

Esto sugiere la necesidad de emplear técnicas para inferir la verdad de referencia a partir de varios observadores o de entrenar a la red con probabilidades en lugar de con verdades absolutas.

1. Introducción

El cáncer de piel es el cáncer más común en todo el mundo [1] Principalmente se encuentran dos tipos, melanoma y no melanoma. Los tumores no melanoma más comunes son los carcinomas basocelulares (BCC) con una incidencia entorno al 70% [2], presentan los mejores criterios para el diagnóstico y tienen una alta variabilidad en cada uno de los criterios clínicos [3].

En los últimos años, la teledermatología se ha implantado rápidamente. Esta herramienta, aunque es muy útil, supone una alta carga de trabajo para los servicios de Dermatología en los hospitales públicos.

Es por ello, que una herramienta de triaje sería muy útil, para establecer una priorización entre los diferentes casos enviados desde atención primaria, de tal forma que los

casos con mayor malignidad sean atendidos urgentemente por los especialistas.

El número de artículos en la literatura científica sobre técnicas de inteligencia artificial para ayudar a la diagnosis del cáncer de piel ha crecido considerablemente debido a la existencia de bases de datos públicas [4][5][6][7]. Dichas bases de datos son muy accesibles y completas, pero por el contrario el criterio tomado para el diagnóstico, como los patrones BCC, o el número de especialistas que han participado no está disponible.

Para desarrollar una herramienta verdaderamente útil para el médico especialista es importante no sólo que obtenga buenos resultados de clasificación, sino también, poder explicar la clasificación. Una forma de explicar el diagnóstico es proporcionar al médico información sobre las características clínicas detectadas por la herramienta.

Con este fin, Serrano et al. [8] desarrollaron una herramienta de clasificación de lesiones de piel mediante la clasificación de criterios clínicos BCC. El problema de esta herramienta es que, para su entrenamiento, la verdad de referencia se extrae del diagnóstico de uno o varios dermatólogos y no puede corroborarse mediante biopsia. Por ello, es necesario analizar el consenso existente entre especialistas en la determinación de estos patrones dermatoscópicos.

Existen varios trabajos en la literatura que intentan abordar esta cuestión. Longo et al. desarrollan una herramienta para la cual, obtienen la verdad de referencia (GT: Ground Truth) de patrones BCC formada con dos observadores. Cuando no están de acuerdo, un tercero desempata el diagnóstico [9]. Lupu et al. estiman la frecuencia de aparición de los diferentes patrones de BCC a partir del diagnóstico de tres dermatólogos, pero no informan cuál es la concordancia ni cómo llegan a un acuerdo [10].

2. Metodología

2.1. Base de datos

Para este estudio, la base de datos empleada consta de 204 imágenes dermatológicas (256x256) del Hospital Universitario Virgen de Macarena de Sevilla. Cada imagen presenta uno o más patrones BCC, es decir, se tiene un problema de clasificación multietiqueta. Cada imagen fue etiquetada independientemente por cada uno de los especialistas.

En la Tabla 1 se puede apreciar el número de imágenes que presenta cada uno de los criterios dermatológicos BCC según el Dermatólogo 1 (D1), donde PN es red pigmentada, del inglés Pigment Network (criterio negativo); UL es ulceración, del inglés Ulceration; BO es nido ovoide gris azulado, del inglés Blue-gray Ovoid Nest; MG es múltiples glóbulos azul-grisáceos, del inglés Multiple Blue-gray Globules; ML es Hoja de arce, del inglés Maple Leaf, SW es rueda de carro, del inglés Spoke-Wheel; y AT es telangiectasia arborizante, del inglés Arborizing Telangiectasia. Ejemplos de esos criterios se muestran en la Figura 1.

	PN	UL	BO	MG	ML	SW	AT
Número de imágenes	16	72	63	33	28	12	102

Tabla 1. Número de imágenes con un patrón BCC en particular.

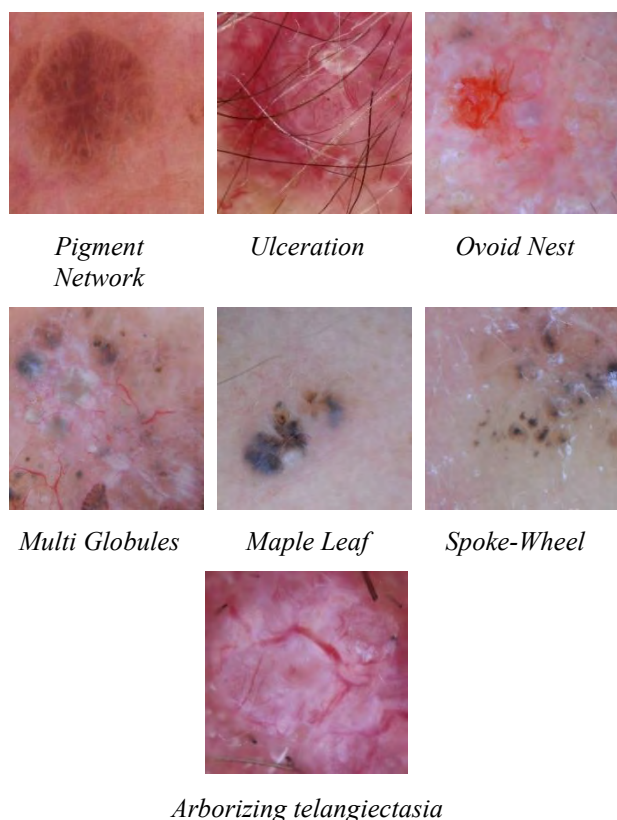


Figura 1. Ejemplos de criterios dermatoscópico BCC, mientras que Pigment Network es un criterio negativo

2.2. Acuerdo entre evaluadores

Las etiquetas fueron codificadas de forma binaria, donde 1/0 representa la presencia/ausencia de un patrón dermatológico concreto. Dado que se tienen 7 posibles patrones cada una de las imágenes será codificada con un vector [1x7] con la siguiente estructura [PN, UL, BO, MG, ML, SW, AT].

Para evaluar el acuerdo entre evaluadores se hizo uso de:

1. La distancia de Hamming [11] se calculó para cada par de evaluadores para así medir la discrepancia entre ellos. Esta medida se utiliza en la teoría de la codificación para calcular el número de posiciones donde los símbolos son diferentes. Para cada imagen y dermatólogo obtenemos una palabra de código. Cada palabra de código tiene siete bits, uno para cada criterio dermatoscópico de BCC. Una distancia igual a cero indica que las palabras de código son exactamente iguales o lo que es lo mismo, ambos especialistas están completamente de acuerdo. Por el contrario, cuando la distancia de Hamming sea igual a X significará que ambas palabras de código difieren en X posiciones, es decir, X patrones.
2. El coeficiente de Cohen Kappa [12] es un coeficiente el cual indica la concordancia entre dos evaluadores. Un valor de Cohen Kappa cercano a 1 indica una excelente concordancia entre ambos evaluadores; un valor cercano a 0 indica una concordancia aleatoria; y un valor negativo, indica total discrepancia entre evaluadores.
3. El coeficiente de Fleiss Kappa [13] es una evolución del anteriormente descrito coeficiente Cohen Kappa. Este calcula la concordancia entre tres o más evaluadores. Para este caso, el coeficiente Fleiss Kappa es calculado para cada patrón BCC.

Kappa	Nivel de acuerdo
$\kappa \geq 0.75$	Excelente
$0.4 \leq \kappa < 0.75$	Regular
$\kappa < 0.4$	Malo

Tabla 2. Relación entre los coeficientes Kappa y un valor cualitativo [15].

2.3. Herramienta de Inteligencia Artificial para detección de patrones BCC

En Serrano et al., se presentó una herramienta de detección de patrones BCC en lesiones cutáneas. El algoritmo se puede visualizar en la Figura 2. Está basado en redes neuronales profundas y en el diseño y extracción de características del color y texturas [8].

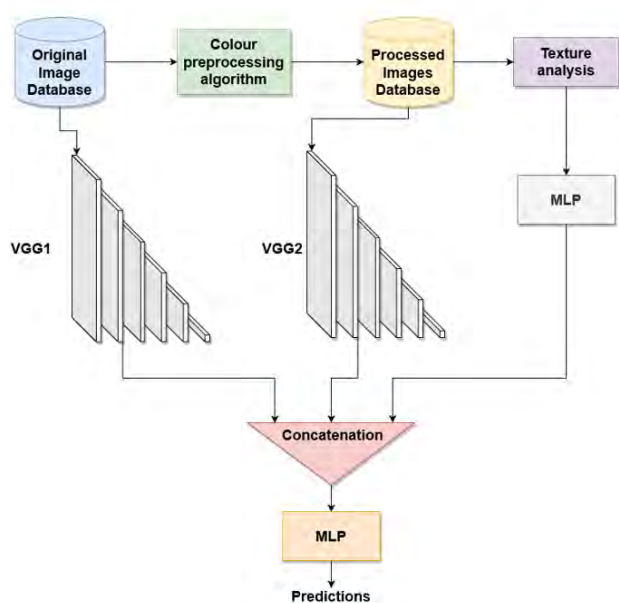


Figura 2. Arquitectura de Inteligencia Artificial [8].

En este trabajo se ha empleado esta misma arquitectura. El modelo fue entrenado con las etiquetas (GT) dadas por el Dermatólogo 1 (D1). El entrenamiento estuvo formado por 1371 imágenes, que después de aplicar la técnica *data augmentation* (DA) fueron aumentadas a 5484 imágenes. De este conjunto se extrajeron 204 imágenes (sin DA) mediante la técnica *Leave-one-out*. Estas imágenes son que se han comentado en la Sección 2.1. De esta forma, se ha completado el experimento con las 204 imágenes con una GT dada por el Dermatólogo 1.

3. Resultados

3.1. Concordancia entre especialistas

En la Figura 3 se pueden ver las distancias de Hamming entre las palabras de código de los diferentes especialistas. Se puede apreciar que la concordancia entre los dermatólogos 1, 2 y 3 (especialmente entre 1 y 2) es adecuada, con distancias 0 o 1 para la mayoría de las lesiones. Sin embargo, las distancias de Hamming cuando se comparan con el Dermatólogo 4 son más altas.

En la Tabla 3 se presentan los resultados correspondientes al coeficiente Cohen Kappa por pares de especialistas. En cada columna se muestran los valores de concordancia para cada patrón para cada uno de los pares de especialistas (filas). En esta tabla, ningún acuerdo (0,0 hasta 0,20) o un acuerdo mínimo (0,21 hasta 0,39) es resaltado en negrita. El patrón en el cual se encuentra un mayor acuerdo entre los dermatólogos es en *Arborizing Telangiectasia*. Sin embargo, también se puede apreciar un alto nivel de acuerdo entre el Dermatólogo 1 y 2 o 2 y 3 para los patrones *Pigment Network*, *Ulceration* y *Arborizing Telangiectasia*. Se puede apreciar que existe un bajo acuerdo en la determinación del patrón *Spoke Wheel*. Asimismo, existe un bajo acuerdo entre el Dermatólogo 4 y el resto, especialmente en la detección del patrón *Pigment Network*, donde el acuerdo es nulo.

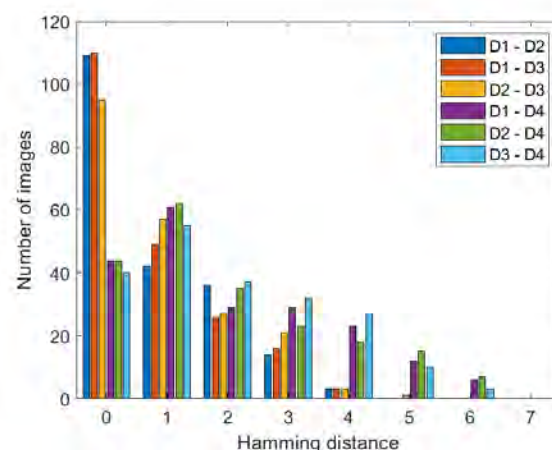


Figura 3. Comparación de la distancia de Hamming calculada por pares con cada dermatólogo.

Pares de evaluadores	PN	UL	BO	MG	ML	SW	AT
D1-D2	0.66	0.71	0.51	0.52	0.49	0.20	0.79
D1-D3	0.75	0.69	0.63	0.60	0.40	0.21	0.78
D1-D4	0.10	0.48	0.57	0.25	0.25	0.19	0.54
D2-D3	0.59	0.63	0.44	0.52	0.34	0.42	0.79
D2-D4	0.10	0.47	0.49	0.30	0.26	0.23	0.55
D3-D4	0.07	0.38	0.68	0.23	0.21	0.34	0.55

Tabla 3. Coeficiente de Cohen Kappa calculado por pares de especialistas.

En la Tabla 4 se puede apreciar los valores del coeficiente *Fleiss Kappa* obtenidos para cada criterio dermatoscópico. En dicha tabla el valor *Pa* representa el acuerdo observado y *Po* representan el acuerdo esperado si los evaluadores respondieran de forma aleatoria. Por último, κ representa el coeficiente *Fleiss Kappa*. Según este coeficiente, la concordancia entre dermatólogos es "ninguna concordancia" (<0) o "concordancia leve" ($0 < \kappa < 0,2$), excepto para el criterio de *Arborizing Telangiectasia*. Una posible razón es que, aunque la probabilidad de acuerdo observado no es baja, probabilidad de acierto aleatorio es alta porque la probabilidad de ocurrencia de los diferentes patrones dermatoscópicos es muy baja. Por el contrario, el patrón *Arborizing Telangiectasia* tiene una probabilidad de aparición aproximadamente del 0,5.

	PN	UL	BO	MG	ML	SW	AT
<i>Pa</i>	0.74	0.79	0.80	0.78	0.78	0.85	0.83
<i>Po</i>	0.85	0.77	0.77	0.8	0.85	0.90	0.75
κ	-0.71	0.12	0.11	-0.25	-0.43	-0.48	0.34

Tabla 4. Coeficiente *Fleiss Kappa* por patrón dermatoscópico.

Como se ha comentado anteriormente, este estudio se centra sobre los patrones dermatoscópicos BCC, en el cual se ha podido comprobar que hay una alta subjetividad a la hora de realizar el diagnóstico. Sin embargo, esto no es

indicativo de que el diagnóstico BCC vs No BCC sea dispar entre especialistas. Por ello, en la Tabla 6 se presentan los valores de la distancia de *Hamming* (*DH*) que al tratarse de una clasificación binaria habrá una distancia igual a 0 si coinciden en diagnóstico o una distancia igual a 2 si discrepan y el valor del coeficiente *Cohen Kappa* (*CK*) con 100 imágenes diferentes a las empleadas en el estudio anterior.

	<i>DH</i> == 0	<i>DH</i> == 2	<i>CK</i>
D2 – D3	98	2	0.94
D2 – D4	88	12	0.69
D3 – D4	90	10	0.75

Tabla 5. Valores de la Distancia de Hamming y el coeficiente de Cohen Kappa entre los dermatólogos 2, 3 y 4

3.2. Rendimiento del modelo de Inteligencia Artificial

A continuación, en la Tabla 5 se muestra el resultado del modelo visto en la Sección 2.3 con la GT dada por el Dermatólogo 1. Para ello, se presenta la sensibilidad, la especificidad y la tasa de verdaderos positivos.

	SENS	SPEC	PPV
Dermatólogo1 – AI	0.74	0.52	0.31

Tabla 6. Sensibilidad (*Sens*), especificidad (*SPEC*) y tasa de verdaderos positivos (*PPV*) de la herramienta de IA en comparación con la GT del D1

4. Discusiones y conclusiones

Existen gran cantidad de trabajos en la literatura que realizan clasificación haciendo uso de las bases de datos públicas, pero muy pocos de estos presentan un trabajo de explicación de dichas clasificaciones, ya que, estas bases de datos no contienen información adicional a la lesión, como pueden ser los patrones BCC. En [8] se presenta un trabajo de clasificación explicado entre BCC y no BCC. La Verdad de referencia empleada se basaba en el diagnóstico de un dermatólogo, ya que los patrones dermatoscópicos BCC no pueden determinarse mediante biopsia.

Este estudio llega a la conclusión de que existe una concordancia pobre entre dermatólogos en alguno de los criterios dermatoscópicos, aunque este hecho puede ser debido en parte a la baja ocurrencia de algunos criterios. Pero pone en relieve la necesidad de llegar a GT inferidas mucho más consensuadas y estables para poder así entrenar herramientas IA que puedan dar diagnósticos fiables. El siguiente paso es inferir una GT a partir de los cuatro dermatólogos y volver a entrenar al modelo en base a estas etiquetas.

5. Agradecimientos

Este trabajo ha sido realizado gracias a la financiación obtenida mediante los siguientes proyectos e instituciones: Proyectos de Generación de Conocimiento 2021 del Ministerio de Ciencia e Innovación (PID2021-127871OB-I00) y Proyecto de Excelencia de la Junta de Andalucía en

el ámbito del Plan Andaluz de Investigación, Desarrollo e Innovación PAIDI 2021 (PROYEXCEL_00889).

Referencias

- [1] Skin cancer information (jan 2023), <https://www.skincancer.org/skin-cancer-information/skin-cancer-facts>
- [2] Ciałyńska, M., Narbutt, J., Woźniacka, A., Lesiak, A.: Trends in basal cell carcinoma incidence rates: a 16-year retrospective study of a population in central poland. *Postepy dermatologii i alergologii* 35(1), 47–52 (Feb 2018). <https://doi.org/10.5114/ada.2018.73164>
- [3] Peris, K., Fargnoli, M.C., Garbe, C., Kaufmann, R., Bastholt, L., Seguin, N.B., Bataille, et al.: Diagnosis and treatment of basal cell carcinoma: European consensus- based interdisciplinary guidelines. *European journal of cancer* 118, 10–34 (Sep 2019). <https://doi.org/10.1016/j.ejca.2019.06.003>
- [4] Codella, N.C.F., Gutman, D.A., Celebi, M.E., Helba, B., Marchetti, M.A., Dusza, S.W., Kalloo, A., Liopyris, K., Mishra, N.K., Kittler, H., Halpern, A.: Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (ISIC). *CoRR abs/1710.05006* (2017), <http://arxiv.org/abs/1710.05006>
- [5] Combalia, M., Codella, N.C.F., Rotemberg, V., Helba, B., Vilaplana, V., Reiter, O., Carrera, C., Barreiro, A., Halpern, A.C., Puig, S., Malvehy, J.: Bcn20000: Dermoscopic lesions in the wild (2019)
- [6] Tschandl, P., Rosendahl, C., Kittler, H.: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data* 5 (1) (aug2018). <https://doi.org/10.1038/sdata.2018.161>
- [7] The international skin imaging , <https://www.isic-archive.com>, last access 2023/06/23
- [8] Serrano, C., Lazo, M., Serrano, A., Toledo-Pastrana, T., Barros-Tornay, R., & Acha, B. (2022). Clinically Inspired Skin Lesion Classification through the Detection of Dermoscopic Criteria for Basal Cell Carcinoma. *Journal of Imaging*, 8(7), 197. <https://www.mdpi.com/2313-433X/8/7/197>
- [9] Longo, C., Lallas, A., Kyrgidis, A., Rabinovitz, H., Moscarella, E., Ciardo, S., Zalaudek, I., Oliviero, M., Losi, A., Gonzalez, S., Guitera, P., Piana, S., Argenziano, G., Pellacani, G.: Classifying distinct basal cell carcinoma sub- type by means of dermatoscopy and reflectance confocal microscopy. *Journal of the American Academy of Dermatology* 71(4), 716–724.e1 (Oct 2014). <https://doi.org/10.1016/j.jaad.2014.04.067>
- [10] Lupu, M., Caruntu, C., Popa, M.I., Voiculescu, V.M., Zurac, S.: Vascular patterns in basal cell carcinoma: Dermoscopic, confocal and histopathological perspectives. *Oncology Letters* 17(5), 4112–4125 (May 2019). <https://doi.org/10.3892/ol.2019.10070>, <https://doi.org/10.3892/ol.2019.10070>
- [11] Wegner, P. (1960). A Technique for Counting Ones in a Binary Computer. *Commun. ACM*, 3(5), 322. doi: 10.1145/367236.367286
- [12] Carletta, J. (1996). Assessing Agreement on Classification Tasks: The Kappa Statistic. *Computational Linguistics*, 22(2), 249-254. Cambridge, MA: MIT Press. Retrieved from <https://aclanthology.org/J96-2004>
- [13] Fleiss, J. L., & Cohen, J. (1973). The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability. *Educational and Psychological Measurement*, 33(3), 613-619. doi: 10.1177/001316447303300309
- [14] Fleiss, J. L., Levin B. & Paik M. C. (2003). The Measurement of Interrater Agreement. En *Statistical Methods for Rates and Proportions* (pp. 598-626). <https://doi.org/10.1002/0471445428.ch18>

Influencia de los sistemas Brain-Computer Interface basados en Neurofeedback en las características de la red cerebral

D. Marcos-Martínez^{1,2}, V. Rodríguez-González^{1,2}, S. Pérez-Velasco^{1,2}, E. Santamaría-Vázquez^{1,2},
V. Martínez-Cagigal^{1,2}, R. Hornero^{1,2}

¹ Grupo de Ingeniería Biomédica (GIB), E.T.S Ingenieros de Telecomunicación, Universidad de Valladolid, Valladolid, España, diego.marcos.martinez@uva.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

Resumen

Las técnicas de Neurofeedback (NF) permiten a sus usuarios, mediante el uso de sistemas Brain-Computer Interfaces (BCI), la modulación voluntaria de determinados ritmos de la actividad cerebral. En este trabajo se presenta un análisis exploratorio de los efectos del NF desde una perspectiva novedosa en este campo de investigación: la teoría de redes. Para ello, se empleó las señales de electroencefalograma (EEG) registradas durante un estudio que constó de 6 sesiones de NF. Los sujetos fueron divididos en un grupo de entrenamiento (GE), que entrenó su actividad theta (4 – 8 Hz), y un grupo de placebo (GP). Las redes cerebrales se construyeron a partir de un análisis de conectividad funcional empleando Phase Lag Index (PLV) en las bandas theta (4 – 8 Hz) y alpha (8 – 13 Hz). Para analizar la estructura de las redes durante el NF se estudió el Clustering Coefficient (CLC) y el Characteristic Path Length (PL). La comparación de estas características entre GE y GP mostró unos mayores valores de CLC en el GE en theta y alpha, siendo significativa la diferencia (p -valor corregido < 0.001) en esta última banda. Por otro lado, el PL fue igual para ambos grupos y ambas bandas de frecuencia. Estos resultados podrían sugerir la utilidad del CLC como indicador del proceso de aprendizaje de neuro-modulación durante el entrenamiento NF, así como la utilidad de la metodología de teoría de redes para caracterizar los efectos del NF en la actividad cerebral.

1. Introducción

Los sistemas *Brain-Computer Interface* (BCI) permiten el control de dispositivos externos a través de la actividad cerebral del usuario [1]. Para ello, se establece un ciclo cerrado en el cual se registra, procesa y analiza dicha actividad en tiempo real. El método más utilizado en el campo de BCI para adquirir la señal de actividad cerebral es el electroencefalograma (EEG), debido a su naturaleza no invasiva, portabilidad y bajo coste. Una de las aplicaciones de los BCI son las técnicas de neuro-modulación o *Neurofeedback* (NF) [2]. Estas técnicas se basan en el condicionamiento operante, y buscan que los usuarios consigan adquirir un control voluntario sobre determinados patrones de su actividad cerebral (p.ej. la potencia de una determinada banda de frecuencia) [2]. Para lograr este control, el sistema BCI presenta en tiempo real una realimentación del patrón de actividad cerebral que se desee entrenar. Esta realimentación ayuda al usuario a identificar estrategias neurocognitivas que, a través de su

realización, le permitan modular de manera efectiva dicho patrón. Entre las estrategias típicamente se encuentran la imaginación motora, los cálculos aritméticos mentales o tratar de aumentar el nivel de relajación. Se ha demostrado que el NF tiene el potencial de inducir plasticidad cerebral [2]. Debido a esto, se plantea la hipótesis de que el NF podría normalizar la actividad cerebral, y por tanto, aliviar los síntomas en aquellos desórdenes neuropsicológicos asociados a patrones anormales de activación neuronal [3]. En consecuencia, el NF ha suscitado un gran interés debido a las posibilidades que ofrece como terapia no farmacológica y no invasiva de desórdenes como el trastorno por déficit de atención e hiperactividad o el trastorno depresivo persistente [2]. Asimismo, también se ha planteado su utilidad en el proceso de neuro rehabilitación de pacientes de ictus o en la prevención del deterioro cognitivo asociado a la edad [3].

A pesar de los resultados que señalan la efectividad del NF, aún se desconocen aspectos fundamentales de esta técnica, como el efecto que tiene sobre las distintas escalas de la actividad cerebral [3]. En este sentido, se ha señalado la necesidad de ampliar la variedad de métodos empleados en el estudio de los efectos del NF, ya que la mayoría se centran en el análisis espectral y podría no ser suficiente para determinar el impacto general del NF [4], [5]. En este contexto, el análisis basado en la teoría de redes ofrece una visión completamente diferente a los métodos de activación local [4]. Esta teoría permite estudiar los distintos elementos que componen un sistema complejo y cómo se relacionan. Para ello, se construyen redes en las que los nodos representan dichos elementos y las conexiones representan sus interacciones. Aplicada a la actividad cerebral, esta metodología permite estudiar sus propiedades de red a partir de la información de conectividad funcional registrada en las distintas regiones del cerebro [4].

La hipótesis en la que se basa este trabajo es que el análisis basado en la teoría de redes podría ser útil para estudiar los posibles cambios en la actividad cerebral inducidos por el NF. Para evaluar dicha hipótesis, se han analizado las redes construidas a partir de la conectividad funcional calculada durante un estudio de NF con sujetos jóvenes sanos. El objetivo de este trabajo es, por tanto, explorar los posibles cambios en la red cerebral derivados de un entrenamiento de NF en un grupo experimental y un grupo placebo.

2. Materiales y métodos

2.1. Estudio de NF

Las redes analizadas se obtuvieron a partir de los valores de conectividad funcional calculados sobre una base de datos registrada durante un estudio de NF [6]. En este estudio participaron 20 sujetos jóvenes sanos (media de edad de 25.5 ± 4.4 años, 12 mujeres). El estudio tuvo como objetivo el incremento de la actividad cerebral en la banda theta (θ , 4-8 Hz) en la región frontal-medial (fm- θ). Se ha observado que esta actividad está relacionada con la realización de tareas que involucran la memoria de trabajo, y se ha sugerido que su entrenamiento puede tener un impacto positivo en dicha función cognitiva [7].

El protocolo experimental se compuso de 6 sesiones de entrenamiento NF de una hora de duración distribuidas a lo largo de dos semanas. Durante el experimento se emplearon tres escenarios de entrenamiento diferentes. Estos escenarios proporcionaban realimentación cada 250 ms de la potencia fm- θ registrada en el electrodo Fz. Cada sesión de NF se compuso de 6 bloques de 5 minutos cada uno. El desarrollo de este estudio se realizó íntegramente empleando el entorno de NF ITACA [6] desarrollado en MEDUSA® [8] por el Grupo de Ingeniería Biomédica de la Universidad de Valladolid.

Con el fin de investigar posibles efectos no relacionados con el NF, se formaron dos grupos de sujetos. El grupo de entrenamiento (GE), compuesto por 10 sujetos (7 mujeres), los cuales recibieron realimentación de su actividad fm- θ durante las sesiones de NF. Por otro lado, el grupo de placebo (GP), compuesto por 10 sujetos (5 mujeres), los cuales recibieron como realimentación la potencia de su actividad cerebral en distintos rangos de frecuencia, variando en cada sesión. Las bandas empleadas para la realimentación del GP en cada sesión fueron: 16-18 Hz, 14-16 Hz, 18-20 Hz, 20-22 Hz, 12-14 Hz y 22-24 Hz. Esto permitió que los sujetos de GP pudieran experimentar cierta sensación de control, pero que no pudieran modular voluntariamente la actividad fm- θ .

Para el registro y análisis del EEG en tiempo real durante el NF se utilizó un amplificador g.USBamp (Guger Technologies OG, Graz, Austria) con 16 electrodos activos (F7, F3, Fz, F4, F8, FCz, C3, Cz, C4, CPz, P3, Pz, P4, PO7, POz y PO8) colocados en un gorro elástico siguiendo el estándar 10-10. El canal AFz se utilizó como tierra y se colocó una referencia común en el lóbulo de la oreja derecha. La señal se adquirió a una frecuencia de muestreo de 256 Hz. Para el procesamiento en tiempo real se aplicó a la señal un filtro espacial tipo Laplaciano y un filtro paso banda tipo Butterworth centrado en la banda de frecuencia de entrenamiento.

2.2. Preprocesado del EEG

Previo al cálculo de las matrices de conectividad funcional, se siguió el siguiente preprocesado del EEG registrado durante el NF. En primer lugar, la señal fue referenciada con respecto al promedio de la actividad registrada en todos los electrodos activos. A continuación, se realizó un filtrado paso banda entre 1 y 40 Hz. A la señal resultante se le aplicó el análisis de componentes independientes

(ICA) y se eliminaron aquellas componentes de origen no neuronal (artefactos oculares, musculares o cardiacos), que contaminasen la señal. Tras esto, la señal fue dividida en épocas de 3 s de duración y se aplicó un método de rechazo automático de las épocas ruidosas o que presentaran artefactos.

2.3. Análisis de redes

Las redes estudiadas se obtuvieron a partir de las matrices de conectividad funcional calculadas mediante *Phase Locking Value (PLV)* [9]. Los nodos de la red fueron representados por los distintos electrodos, y las conexiones por los valores de conectividad entre los mismos. *PLV* cuantifica el grado de sincronización del EEG entre un par de electrodos a partir de la estabilidad de su diferencia de fases instantáneas [9]:

$$PLV = \left| \frac{1}{T} \sum_{t=1}^T e^{j\Delta\phi(t)} \right|, \quad (1)$$

siendo $\Delta\phi(t)$ la diferencia de fases instantáneas de dos señales en el instante t , y T la longitud de la ventana temporal analizada. Para el uso de *PLV* es necesario que el ancho de banda de la señal analizada sea estrecho. Por ello, el análisis se llevó a cabo filtrando las épocas resultantes del preprocesado anterior en las bandas theta (θ , 4-8 Hz) y alpha (α , 8-13 Hz). La extracción de la fase instantánea de la señal se llevó a cabo mediante la transformada de Hilbert.

Se escogieron dos características para analizar la estructura global de las redes calculadas: *Clustering Coefficient (CLC)* y *Characteristic Path Length (PL)* [4]. *CLC* proporciona información sobre la presencia de comunidades o subgrupos densamente interconectados dentro de la red [4]. Esta es una propiedad necesaria para establecer que una red presenta segregación y, por tanto, tiende a formar subredes especializadas en distintas tareas. Se define de la siguiente forma:

$$CLC = \left(\frac{N}{3} \right) \sum_{i \in n} \sum_{j, h \in n} (w_{ij}^b w_{ih}^b w_{jh}^b)^{\frac{1}{3}}, \quad (2)$$

donde w_{ij} denota el valor de la conexión entre los electrodos i y j . Por su parte, *PL* proporciona información sobre el grado de integración de una red, es decir, sobre la capacidad para transmitir información de manera eficiente [4]. Se define de la siguiente forma:

$$PL = \frac{1}{N} \sum_{i \in n} \frac{\sum_{j \in n, j \neq i} d_{ij}}{n-1}, \quad (3)$$

donde d_{ij} indica la distancia mínima entre los nodos i y j . Con el objetivo de independizar estas medidas de otras características de la red, como el tamaño o la densidad, las medidas de *CLC* y *PL* de cada red fueron normalizadas con respecto al promedio de los valores de *CLC* y *PL* calculados en 50 redes aleatorias [10]. Estas se obtuvieron reordenando de manera aleatoria el valor de las conexiones de cada red original. Finalmente, se realizó un promedio de los parámetros de red calculados, obteniendo un único valor de *CLC* y *PL* por sesión para cada sujeto.

2.4. Análisis estadístico

La posible influencia del NF en las características de red se analizó comparando los valores de *CLC* y *PL* de cada grupo obtenidos en las distintas sesiones. Para evaluar la significancia estadística de los resultados se empleó la prueba U de Mann-Whitney ($\alpha=0.05$) debido a que los datos no satisfacían la condición de normalidad. Los *p*-valores obtenidos fueron corregidos mediante FDR-BH.

3. Resultados

En la Figura 1 se presentan, mediante diagrama de cajas, los valores de *CLC* obtenidos durante las sesiones de NF para los sujetos de GE y GP en las bandas theta y alpha. Se observa que los valores de *CLC* calculados en los sujetos de GE son mayores que los de GP, tanto para theta como para alpha. Sin embargo, tras realizar la corrección FDR-BH, únicamente se encontró diferencia significativa (*p*-valor < 0.001) en la comparación de los valores *CLC* entre grupos calculado en la banda alpha. En la banda theta se obtuvo un *p*-valor próximo a la significancia estadística (*p*-valor = 0.053).

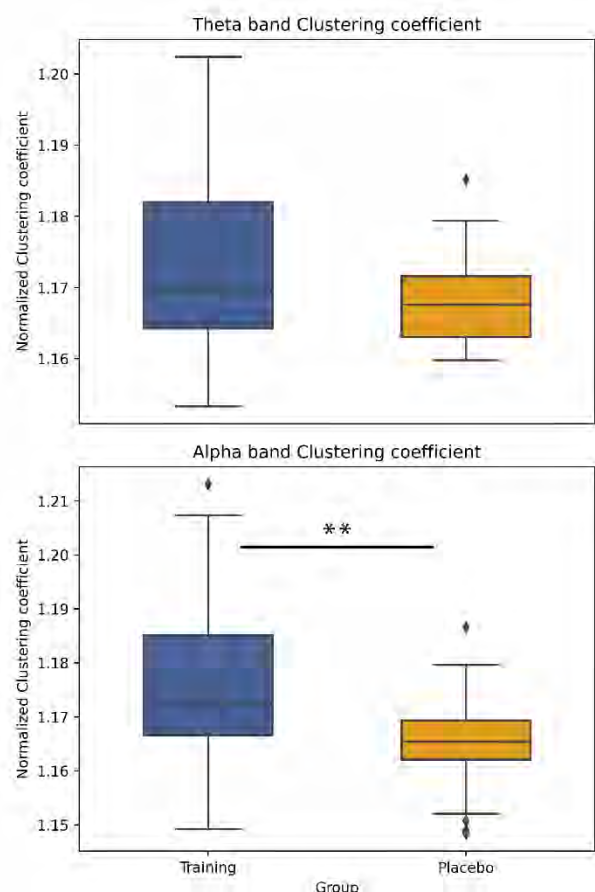


Figura 1. Diagrama de cajas de los valores de *CLC* para cada grupo de sujetos calculado en las bandas theta (superior) y alpha (inferior). Los dos asteriscos (**) indican un *p*-valor < 0.001.

En la Figura 2. se presentan, mediante diagrama de cajas, los valores de *PL* obtenidos durante las sesiones de NF para los sujetos de GE y GP en las bandas theta y alpha. Los valores de *PL* de GE en ambas bandas de frecuencia son ligeramente inferiores a los de GP, sin embargo, esta

diferencia no es significativa en ninguno de los casos (*p*-valores > 0.05).

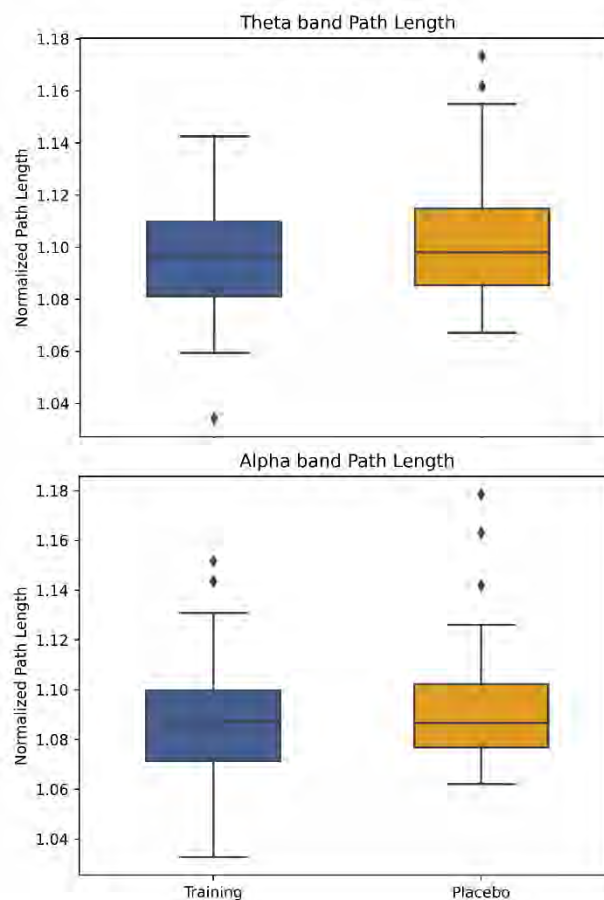


Figura 2. Diagrama de cajas de los valores de *PL* para cada grupo de sujetos calculado en las bandas theta (superior) y alpha (inferior).

4. Discusión

El objetivo de este trabajo ha sido explorar los cambios en la actividad cerebral inducidos por el NF desde la perspectiva de la teoría de redes. Para ello, se han analizado dos de las principales características globales de las redes consideradas: la segregación y la integración.

Múltiples estudios han demostrado que las características de la red cerebral no permanecen inmutables, sino que estas varían en función de las demandas de la tarea cognitiva que se lleva a cabo [11][12]. Además, no todas las características de red se ven influidas de la misma manera. Por ejemplo, un estudio reciente demostró que la automatización de tareas cognitivamente exigentes a través de la repetición daba lugar a redes más segregadas [12]. Sin embargo, la integración de estas redes no se vio afectada. Esto se encuentra alineado con los resultados obtenidos en el presente trabajo. Basándonos en los resultados de la literatura expuestos, consideramos que la diferencia en los valores de *CLC* observados entre los sujetos GE y GP podrían deberse a la automatización del proceso de neuro modulación, el cual únicamente se pudo dar en los sujetos de GE. Este proceso de automatización en NF ya ha sido mencionado previamente en la literatura [2][3]. En estos trabajos se señala cómo en un primer contacto con el NF,

el usuario debe explorar distintas tareas cognitivas que le permitan alcanzar un estado neuronal que se corresponda con el objetivo del entrenamiento. Sin embargo, a medida que este proceso se repite, el usuario va alcanzando de manera más sencilla dicho estado neuronal sin la necesidad de utilizar una tarea cognitiva como intermediario. Esto podría traducirse en una mayor especialización de las distintas partes implicadas, y por tanto un aumento de la segregación de la red cerebral. Estos resultados sugieren que el *CLC* podría ser un indicador útil para evaluar el nivel de aprendizaje de la neuro modulación en sesiones de NF. Por otro lado, la ausencia de diferencias en el *PL* entre ambos grupos sugiere que el proceso de aprendizaje durante el NF no parece tener un impacto en la característica de integración de la red cerebral del usuario.

A pesar de los resultados obtenidos, nuestro trabajo no se encuentra exento de limitaciones. En primer lugar, sería deseable incrementar la población de estudio, tanto en el GE como en el GP. Esto aportaría una mayor potencia estadística a los resultados obtenidos. En segundo lugar, ampliar las características de red estudiadas permitiría tener una mejor perspectiva de cómo varía la estructura de la red cerebral con el NF. Por último, sería interesante estudiar los efectos del NF a nivel de red una vez finalizado el entrenamiento. Esto podría aportar más información sobre la capacidad del NF de inducir cambios en la estructura de la red cerebral a largo plazo.

5. Conclusiones

En este estudio se investigaron los efectos del NF en la actividad cerebral utilizando la teoría de redes como enfoque metodológico. Esta aproximación, innovadora en el campo del NF, ha demostrado ser efectiva en la caracterización de los cambios en el EEG inducidos por el NF. Específicamente, al analizar el *CLC*, se observó una mayor segregación en las redes cerebrales de los sujetos del GE en comparación con los del GP, tanto en la banda theta como en la banda alpha. Estos resultados sugieren que el *CLC* puede ser un indicador útil del aprendizaje de la neuro modulación durante el NF. Además, destacan la importancia de emplear esta metodología para obtener un mayor entendimiento de la influencia del NF en la dinámica cerebral.

Agradecimientos

Este estudio ha sido financiado por los proyectos TED2021-129915B-I00, PID2020-115468RB-I00 y RTC2019-007350-1 financiadas por el Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación/10.13039/501100011033/, FEDER Una forma de hacer Europa; y por ‘Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)’ a través de ‘Instituto de Salud Carlos III’. D. Marcos-Martínez y S. Pérez-Velasco son beneficiarios de una ayuda PIF de la Consejería de Educación de la Junta de Castilla y León.

Referencias

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, “Brain-computer interfaces for communication and control,” *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002, doi: 10.1145/1941487.1941506.
- [2] R. Sitaram *et al.*, “Closed-loop brain training: The science of neurofeedback,” *Nat. Rev. Neurosci.*, vol. 18, no. 2, pp. 86–100, 2017, doi: 10.1038/nrn.2016.164.
- [3] T. Ros, B. J. Baars, R. A. Lanius, and P. Vuilleumier, “Tuning pathological brain oscillations with neurofeedback: A systems neuroscience framework,” *Front. Hum. Neurosci.*, vol. 8, no. DEC, pp. 1–22, 2014, doi: 10.3389/fnhum.2014.01008.
- [4] J. Gonzalez-astudillo, T. Cattai, G. Bassignana, and M. Corsi, “Network-based brain – computer interfaces: principles and applications Network-based brain – computer interfaces: principles and applications,” *J. Neural Eng.*, vol. 18, 2021.
- [5] D. Marcos-Martínez, V. Martínez-Cagigal, E. Santamaría-Vázquez, S. Pérez-Velasco, and R. Hornero, “Neurofeedback training based on motor imagery strategies increases EEG complexity in elderly population,” *Entropy*, vol. 23, no. 12, pp. 1–19, 2021, doi: 10.3390/e23121574.
- [6] D. Marcos-Martínez *et al.*, “ITACA: An open-source framework for Neurofeedback based on Brain-Computer Interfaces,” *Comput. Biol. Med.*, vol. 160, 2023, doi: 10.1016/j.compbiomed.2023.107011.
- [7] S. Enriquez-Geppert, R. J. Huster, R. Scharfenort, Z. N. Mokok, J. Zimmermann, and C. S. Herrmann, “Modulation of frontal-midline theta by neurofeedback,” *Biol. Psychol.*, vol. 95, no. 1, pp. 59–69, 2014, doi: 10.1016/j.biopsycho.2013.02.019.
- [8] E. Santamaría-vázquez *et al.*, “MEDUSA ©: A novel Python-based software ecosystem to accelerate brain-computer interface and cognitive neuroscience research,” vol. 230, 2023, doi: 10.1016/j.cmpb.2023.107357.
- [9] J. P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela, “Measuring phase synchrony in brain signals,” *Hum. Brain Mapp.*, vol. 8, no. 4, pp. 194–208, 1999, doi: 10.1002/(SICI)1097-0193(1999)8:4<194::AID-HBM4>3.0.CO;2-C.
- [10] C. J. Stam *et al.*, “Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer’s disease,” *Brain*, vol. 132, no. 1, pp. 213–224, 2009, doi: 10.1093/brain/awn262.
- [11] R. Wang, M. Liu, X. Cheng, Y. Wu, A. Hildebrandt, and C. Zhou, “Segregation, integration, and balance of large-scale resting Brain Networks Configure Different Cognitive Abilities,” *Proc. Natl. Acad. Sci.*, vol. 36, no. 48, pp. 12083–12094, 2021, doi: 10.1073/pnas.2022288118/-/DCSupplemental.y.
- [12] K. Finc *et al.*, “Dynamic reconfiguration of functional brain networks during working memory training,” *Nat. Commun.*, vol. 11, no. 1, pp. 1–15, 2020, doi: 10.1038/s41467-020-15631-z.

Validación de un protocolo de adquisición de señales electroespinoográficas en posición decúbito prono

D. I. Gracia^{1,2}, P. Soriano-Segura^{1,2}, M. Ortiz^{1,2}, E. Iáñez^{1,2}, J. M. Azorín^{1,2,3}

¹ Brain-Machine Interface System Lab, Universidad Miguel Hernández de Elche, Elche, España
{dgracia, p.soriano, mortiz, eianez, jm.azorin}@umh.es

² Instituto de Investigación en Ingeniería de Elche - I3E, Universidad Miguel Hernández de Elche, Elche, España

³ Valencian Graduate School and Research Network of Artificial Intelligence – valgrAI, Valencia, España

Resumen

Las lesiones medulares provocan graves trastornos motores y suponen un importante reto para la sociedad. Para ayudar a las personas con limitaciones motoras, los exoesqueletos robóticos de extremidades inferiores integrados con interfaces neuro-máquina han surgido como una solución prometedora. Se investiga el potencial de combinar datos de la médula espinal con señales electroencefalográficas para proporcionar un control más preciso de los exoesqueletos. En este estudio se valida un protocolo para el registro de la actividad eléctrica de la médula espinal. Este se realiza de forma no invasiva utilizando matrices de electromiografía de superficie de alta densidad posicionadas sobre el plexo braquial y lumbar. El participante realiza registros tanto en reposo como realizando movimientos isotónicos e isométricos de las extremidades inferiores y superiores y de la sección derecha e izquierda. Las señales adquiridas requieren de un proceso de filtrado para eliminar la componente cardíaca presente, realizado a partir del algoritmo de Adaptive Template Substraction. Los resultados obtenidos revelaron una clara diferenciación entre los registros realizados en estado de reposo y aquellos durante la ejecución de movimientos, así como entre los movimientos isotónicos e isométricos. Sin embargo, no se observó una variabilidad distintiva en función de la región registrada por la matriz de electrodos, ni el miembro estudiado.

1. Introducción

La lesión de la médula espinal (LME) es uno de los trastornos motores más significativos causados por daños en el sistema nervioso humano, lo cual resulta en un deterioro físico en la sociedad occidental. Este tipo de alteraciones neurológicas interrumpen las vías sensoriales y motoras, lo cual conduce a una marcha patológica permanente y un deterioro en la capacidad deambulatoria independiente. La rehabilitación física manual a cargo del personal técnico requiere un esfuerzo y tiempo considerables por parte de los terapeutas, siendo frecuentemente la restauración de la función motora variable e incompleta. Con los avances recientes en robótica, los exoesqueletos robóticos de miembro inferior han surgido como una ayuda para la marcha de personas con limitaciones motoras [1].

Aunque la electroencefalografía (EEG) es una opción interesante que se utiliza para incorporar al paciente en el control de exoesqueletos en rehabilitación, su combinación con electromiografía (EMG) también es empleada para mejorar el rendimiento [2]. Siguiendo esta tendencia, este estudio se centra en el potencial de la información transmitida por la médula espinal conocida como electroespinografía (ESG), que no solo transmite señales

del cerebro, sino que también puede procesar información de forma independiente [3]. El estudio del procesamiento de señales eferentes por parte de la médula espinal plantea una herramienta novedosa para las interfaces neuro-máquina (NMI), pudiéndose alcanzar una mayor precisión al disponer de más información e interpretaciones sobre la intención del usuario de controlar el exoesqueleto al combinarse con las señales EEG.

La actividad de la médula espinal puede ser estudiada tanto directamente, registrando la actividad neural, o indirectamente, analizando las variaciones del flujo vascular. Tradicionalmente las técnicas de registro eran invasivas y costosas, aunque actualmente se cuentan con opciones no invasivas y de menor coste, como la espectroscopia funcional de infrarrojo cercano (fNIRS) [4] o la electromiografía de superficie de alta densidad (HD-sEMG) [5]. Sin embargo, estos estudios se han enfocado en la caracterización de potenciales evocados.

Sin embargo, huelga recalcar que el registro de las señales ESG es susceptible a la presencia de artefactos causados por la señal electrocardiográfica (ECG) al estar realizándose el registro sobre el tronco del paciente. La componente ECG debe ser eliminada de forma previa al procesamiento de las señales EMG para no comprometer las características extraídas, puesto que ambas señales presentan componentes frecuenciales superpuestas. Generalmente las señales EMG se encuentran dentro de las bandas de 10 a 500 Hz, mientras que las señales ECG alcanzan frecuencias de hasta 100 Hz, con valores de amplitud superiores a los de EMG [6]. Esta problemática ha llevado al desarrollo de algoritmos basados en diferentes opciones como los filtros de paso alto, la sustracción de patrones cardíacos, la transformada de Wavelet, los filtros adaptativos o la separación ciega de fuentes.

En el presente estudio, se realizó una validación del protocolo propuesto en [7]. En el cual se propone el registro de las señales ESG mediante EMG de alta densidad (HD-sEMG) de forma no invasiva, mientras el participante realiza registros tanto en reposo como en movimiento con el objetivo de caracterizar los potenciales motores relacionados con la inervación de los miembros inferiores. El protocolo permite la comparación de las señales obtenidas correspondientes a movimientos isotónicos e isométricos de las extremidades inferiores y superiores y de la sección derecha e izquierda.

2. Materiales y métodos

2.1. Participantes

Las pruebas fueron realizadas por una voluntaria sana, de 26 años, sin limitaciones motoras ni problemas de salud conocidos. Se le informó sobre el procedimiento experimental, firmando seguidamente el consentimiento informado. El estudio se realizó de acuerdo con la Declaración de Helsinki de la Asociación Médica Mundial. Los procedimientos fueron aprobados por la Oficina de Investigación Responsable de la Universidad Miguel Hernández de Elche (DIS.JAP.09.21).

2.2. Equipos

La señal ESG fue adquirida utilizando un equipo de HD-sEMG, Sessantaquattro+, y registrada mediante el software de adquisición OTBioLab+ (OT Bioelettronica, Italia) a una frecuencia de muestreo de 2000 Hz. Se emplearon dos matrices compuestas por 32 electrodos de 1 mm de diámetro, dispuestos en una configuración de 4x8, con una distancia de 1 cm entre electrodos. Estas matrices se colocaron estratégicamente para el registro del plexo braquial, tomando como punto de referencia la vértebra T1, y del plexo lumbar, siendo el punto de referencia la vértebra T12. Además, se posicionó un electrodo de referencia sobre la muñeca derecha (Figura 1a y 1b). Antes de su colocación, la zona de aplicación de los electrodos fue preparada mediante la limpieza con gel dermoabrasivo y alcohol etílico.

Asimismo, se utilizaron galgas extensiométricas para registrar el momento inicial de los movimientos isométricos. Las galgas fueron posicionadas dentro del alcance del miembro que se solicitaba mover (Figura 1c). La galga se conectó a una placa ESP32 Dev Module (Espressif Systems, China) que transmitió la información directamente a Matlab con una frecuencia de muestreo de 2000 Hz.

En el caso de los movimientos isométricos, se emplearon IMUs (WIT Motion, China) para caracterizar temporalmente los movimientos realizados. Se colocó una IMU en cada extremidad, a nivel de la palma de la mano y



Figura 1. Posicionamiento del material empleado durante el registro del procedimiento experimental: a) 2 matrices de 32 electrodos para el registro de señales ESG a nivel braquial y lumbar, b) un electrodo de referencia y una IMU fijados en la mano derecha y c) una galga extensiométrica al alcance del pie y una IMU colocada sobre el antepié derecho con las que caracterizar el instante inicial durante la realización de movimientos.

el antepié (Figura 1b y 1c). Cada equipo registró, con una frecuencia de muestreo de 100 Hz, 13 variables: aceleración lineal, velocidad angular, ángulos de Euler y cuaternios por los tres ejes y el campo magnético.

Con el fin de reducir el ruido y las interferencias de la red eléctrica se empleó el dispositivo REOMED I 600 (REO, Alemania).

2.3. Procedimiento experimental

El experimento se estructuró en cuatro etapas distintas. Tanto en la primera etapa como en la última, la participante permaneció en una posición decúbito prono con los ojos abiertos durante dos minutos, con el fin de obtener información sobre el estado basal de las señales registradas.

En la segunda etapa, se solicitó a la participante que realizara movimientos isotónicos, específicamente flexión y extensión de la muñeca o el tobillo, de forma individual para cada uno de los cuatro registros: pierna derecha, pierna izquierda, brazo derecho y brazo izquierdo. Cada registro se dividió en una fase inicial de 15 segundos, durante la cual la persona debía permanecer relajada de manera similar al registro en reposo, seguida de cinco bloques de repeticiones. Cada bloque comenzaba con una señal auditiva de 1 segundo que indicaba al sujeto el inicio de la ventana temporal en la cual debía realizar cinco repeticiones del movimiento requerido a un ritmo libre, teniendo un tiempo máximo de 15 segundos para realizarlas. Al finalizar cada bloque, se otorgaban 10 segundos de descanso antes de iniciar el siguiente bloque de repeticiones. Al comienzo de cada registro, se proporcionaba a la participante una indicación auditiva sobre la actividad a realizar.

En la tercera etapa, se llevó a cabo el registro de movimientos isométricos. Para ello, se instruyó a la participante a ejercer presión con la punta de los dedos de la mano o del pie. Los registros en esta etapa se realizaron de forma equivalente a la etapa anterior, con la única diferencia de que la participante debía realizar dos repeticiones del movimiento en cada bloque, debiendo de mantener la fuerza durante al menos 5 segundos en cada repetición.

2.4. Análisis

En primer lugar, se seleccionaron los 5 segundos centrales de cada bloque de movimiento utilizando para ello las señales registradas en IMUs y galgas. Del mismo modo, se eligieron los 5 segundos centrales de cada etapa de reposo entre cada bloque de movimientos. Para los registros de reposo total, se calculó el valor inicial de las ventanas temporales como el promedio de los valores iniciales en los registros de las series de movimientos isotónicos e isométricos. A pesar de no reflejar la realidad de la actividad realizada, se mantiene la clasificación de instantes en movimiento y en reposo durante los registros en reposo total para poder realizar una comparación con el resto de los registros.

A continuación, a cada señal de 5 segundos se aplicó el siguiente procesamiento. En primer lugar, se aplicó un

filtro Butterworth Notch de segundo orden a 50 Hz para mitigar el ruido generado por la fuente de corriente. Posteriormente, se utilizó un filtro Butterworth paso bajo de sexto orden con una frecuencia de corte de 520 Hz para reducir el ruido de altas frecuencias. La señal resultante se promedió cada 5 muestras y se filtró nuevamente utilizando un filtro Butterworth paso alto de sexto orden con una frecuencia de corte de 10 Hz. Luego se aplicó un algoritmo de detección de picos R. La señal obtenida fue filtrada una vez más con un filtro Butterworth paso alto de sexto orden, esta vez con una frecuencia de corte de 20 Hz. Este filtro no solo eliminó la componente continua y los artefactos de movimiento, sino que también redujo parte de la contribución de la señal ECG. Finalmente, se aplicó el algoritmo de Adaptive Template Substraction [8]. Se seleccionó este algoritmo entre las diferentes opciones por su sencillez y baja frecuencia de corte, alterando mínimamente las propiedades de la señal original. Finalmente, se calcula el espectro frecuencial de la señal filtrada para cada una de las ventanas temporales definidas, promediándose por grupos.

3. Resultados y discusión

El artefacto causado por la componente ECG muestra una amplitud mayor en la matriz lumbar, posiblemente debido a su proximidad al corazón. No obstante, el proceso de desartefactado logró recuperar la señal registrada sin la presencia de la componente ECG, pero sin alterar el resto de las componentes en ambas matrices (Figura 2).

Los espectros calculados para cada ventana temporal se promediaron para las repeticiones de movimiento y relax para cada registro.

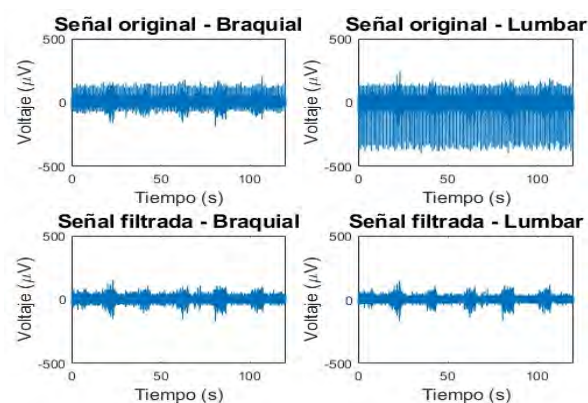


Figura 2. Señales adquiridas en el electrodo 17 de la matriz braquial y lumbar durante el registro donde la participante realizó movimientos isotónicos de pie izquierdo. Se muestra tanto en crudo (arriba) como tras haber sido filtrada la componente ECG (abajo).

Tanto en la matriz braquial (primeras dos filas de la Figura 3) como en la lumbar (últimas dos filas de la Figura 3), destacan las señales provenientes del movimiento de la mano derecha. Estas señales presentan un considerable artefacto de movimiento debido, muy probablemente, al movimiento del electrodo de referencia. Es posible que este movimiento haya causado un contacto inadecuado entre el electrodo y la muñeca, lo que habría resultado en una incorrecta referenciación.

Además, también en ambas matrices se observa que la amplitud de las señales isométricas es mayor que la de las señales isotónicas. El tener que mantener la fuerza durante un período prolongado podría provocar una transmisión de información más intensa o, al menos, más frecuente a través de la médula espinal.

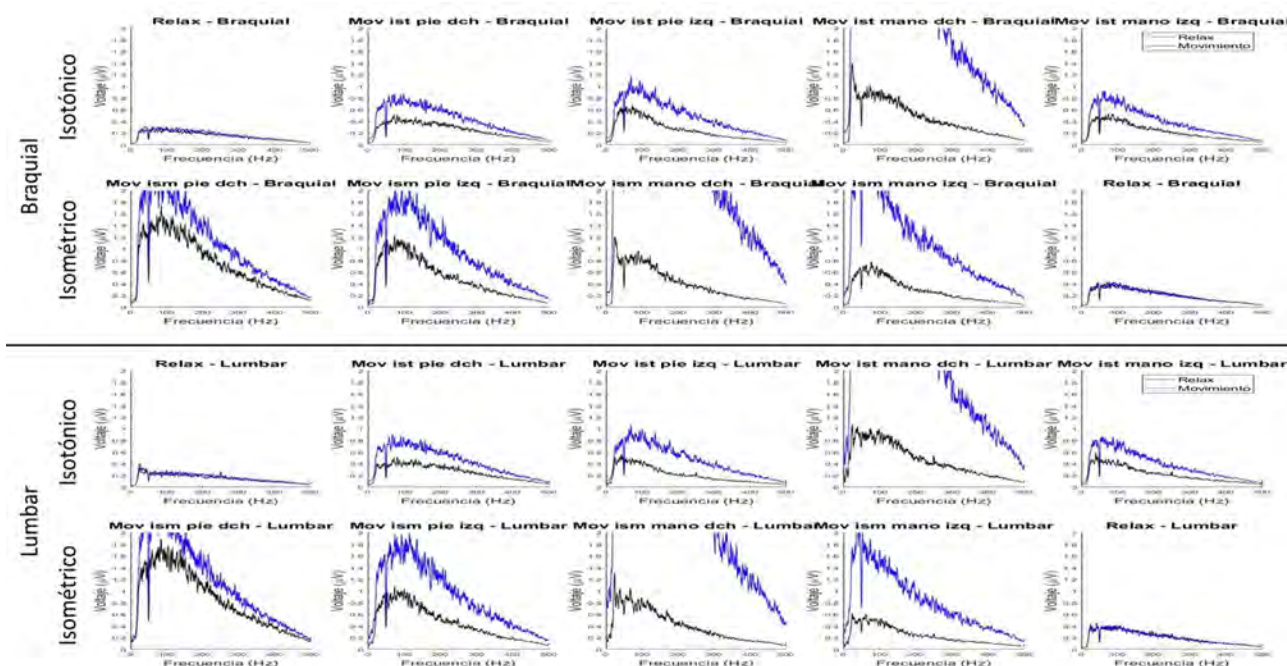


Figura 3. Espectro frecuencial del electrodo 17 de la matriz braquial y lumbar calculado como el promedio de las cinco repeticiones realizadas en cada registro en los estados de descanso y movimiento, teniendo en cuenta los cinco segundos centrales de esa tarea. Las tareas realizadas fueron: descanso mantenido por 2 minutos, tanda de movimientos isotónicos (orden seguido: pie derecho, pie izquierdo, mano derecha y mano izquierda), tanda de movimientos isométricos (orden seguido: pie derecho, pie izquierdo, mano derecha y mano izquierda) y descanso mantenido por 2 minutos.

En cuanto a la comparación entre los registros, se observa que las señales obtenidas durante los dos minutos de reposo muestran una potencia considerablemente menor en comparación con las señales durante las tareas de movimiento o los períodos de descanso entre repeticiones en los demás registros. La mayor potencia en las señales durante los períodos de descanso entre repeticiones puede deberse a que el tiempo de separación entre la realización de los movimientos fue insuficiente, lo que impidió que la participante se relajase por completo. Sin embargo, es importante destacar que, en todos los casos, las señales durante los períodos de descanso entre repeticiones muestran una amplitud menor que las señales durante los movimientos, siendo esta diferenciación más evidente en ciertos casos, como en el registro de movimientos isométricos de la mano izquierda.

Al diseñar el protocolo propuesto, se planteó la hipótesis de que la información transmitida por la médula espinal sería distinguible según si el movimiento era realizado por las extremidades superiores o inferiores, así como según si se realizaba con las extremidades derechas o izquierdas. Sin embargo, esta hipótesis ha sido descartada a raíz de los resultados obtenidos. Resultados similares se obtuvieron en trabajos anteriores [4] donde se emplearon otras técnicas de registro (fNIRS). En ellos se argumenta la conectividad intra y supraespinal, así como las conexiones interespinales del asta dorsal mediante las fibras arcuatas como explicación de los resultados obtenidos [9].

4. Conclusiones

Las pruebas realizadas con el protocolo propuesto han revelado varias problemáticas que deberán abordarse en futuras versiones. Es necesario modificar la posición del electrodo de referencia para evitar la presencia de artefactos de movimiento. Además, se deberá tener en cuenta si la referencia se posiciona en una región cercana a la zona de registro, este electrodo también registrará la componente ECG, por lo que la presencia de este artefacto se verá atenuada en los electrodos de registro. Por lo tanto, en futuros registros se plantea ubicar la referencia sobre la articulación del hombro. Por otra parte, se incrementará el tiempo dedicado al descanso entre repeticiones con el objetivo de lograr que los participantes alcancen condiciones más cercanas al estado basal durante estos períodos. Asimismo, dado que no se han observado diferencias significativas entre la señal adquirida por las matrices empleadas, en futuras pruebas se reemplazará el material utilizado por una única matriz de 64 electrodos con una distancia de 4 mm entre ellos sobre la región lumbar. Se espera que, con una mayor resolución espacial, se pueda realizar una caracterización espacial del flujo de información transmitido por la médula espinal.

No obstante, el protocolo planteado ha permitido obtener señales que son distinguibles entre los registros en reposo y los casos de movimientos isotónicos e isométricos. Sin embargo, sería interesante realizar una caracterización más completa en términos temporales y frecuenciales para asegurar la clasificación precisa de los movimientos realizados.

Por último, debido a que las diferentes tareas realizadas por los miembros superiores e inferiores durante el procedimiento experimental no han generado alteraciones notables en la señal adquirida, se plantea reducir el estudio únicamente al pie derecho (o dominante del participante). Esta reducción permitirá aumentar el número de registros para los nuevos casos (movimientos isotónicos e isométricos), posibilitando incluir diferentes condiciones sin afectar significativamente la duración total de la prueba. De esta manera, se buscará establecer diferentes condiciones más cercanas al objetivo final (neurorrehabilitación de la marcha), pero manteniendo un control de las condiciones para minimizar la presencia de artefactos en la medida de lo posible.

Agradecimientos

Esta publicación es parte del proyecto de I+D+i PID2021-124111OB-C31, financiada por MCIN/AEI/10.13039/501100011033 y por FEDER Una manera de hacer Europa. Este trabajo también cuenta con el apoyo de la Escuela Valenciana de Posgrado e Investigación en Red de Inteligencia Artificial (ValgrAI), Generalitat Valenciana y Unión Europea.

Referencias

- [1] J. L. Pons, *Wearable Robots: Biomechatronic Exoskeletons*. 2008. doi: 10.1002/9780470987667.
- [2] S. Y. Gordleeva et al., “Real-Time EEG-EMG human-machine interface-based control system for a lower-limb exoskeleton,” *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2991812
- [3] D. W. Cadotte et al., “Plasticity of the Injured Human Spinal Cord: Insights Revealed by Spinal Cord Functional MRI,” *PLoS ONE*, vol. 7, no. 9, 2012, doi: 10.1371/journal.pone.0045560.
- [4] F. Valenzuela et al., “Non-Invasive Functional Evaluation of the Human Spinal Cord by Assessing the Peri-Spinal Neurovascular Network with near Infrared Spectroscopy,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, 2021, doi: 10.1109/TNSRE.2021.3123587.
- [5] R. Koster, “Registration of somatosensory evoked potentials in the lumbar and lower thoracic spine using high-density surface electromyography,” B.S. thesis, Dpto. Ortopedia, THAUS, La Haya, Países Bajos, 2017.
- [6] L. Xu, et al., “Comparative review of the algorithms for removal of electrocardiographic interference from trunk electromyography,” *Sensors (Switzerland)*, vol. 20, no. 17, 2020, doi: 10.3390/s20174890.
- [7] D. I. Gracia et al., “Protocolo de adquisición de señales electroespino-gráficas en posición decúbito prono,” *Jornadas Nacionales de Robótica y Bioingeniería 2023*, 2023, pp. 291–296. doi: 10.20868/UPM.book.74896.
- [8] E. Petersen. Cardiac artifact removal toolbox. <https://github.com/e-pet/ecgremoval/releases/tag/1.01>, 2023. Online; recuperado el 26 de junio de 2023.
- [9] W. D. Willis and R. E. Coggeshall, *Sensory Mechanisms of the Spinal Cord*. 2004. doi: 10.1007/978-1-4615-0035-3

Conceptual design of a functional electrical stimulation cycling platform as a rehabilitation therapy for spinal cord injury or stroke

T. Olmo-Fajardo¹, D. Herrera-Valenzuela², S. González-Expósito¹, A. de los Reyes-Guzmán², I. Sinovas-Alonso², M. Salas-Monedero², F. Oliveira Barroso¹, A. Gil-Agudo², D. Torricelli¹, J. C. Moreno¹

¹ Neural Rehabilitation Group, Cajal Institute, CSIC - Spanish National Research Council, Madrid, Spain.

tania.olmo@cajal.csic.es

² Biomechanics and Technical Aids Unit, National Hospital for Paraplegics, Toledo, Spain. Associated Unit to CSIC by Cajal Institute, DP.

Abstract

Stroke and spinal cord injuries (SCI) are leading causes of disability worldwide, involving problems in mobility, balance and coordination, among others. While aerobic exercise is associated with a greater plasticity in the motor cortex of healthy individuals, its effect on neuroplasticity after suffering a stroke or a SCI is still unknown. Besides, there is no methodology to promote both cardiovascular and neuroplastic recovery in patients suffering from neurological injuries. Up to now, it has been demonstrated that physical exercise is a therapeutic intervention in many rehabilitation programs that, apart from providing clear benefits related to the physical conditioning, functionality, mood and cardiovascular health, it could also promote neuroplasticity. The literature suggests that stronger neuroplastic responses are elicited in mid-to-high intensity training programs, but the lack of homogeneity in the dose-response and the non-uniform evaluations of the neuroplasticity seem to be a limitation to generalise the obtained results. The combination of functional electrical stimulation (FES) with the benefits of cardiovascular exercise makes cycling assisted by FES a promising approach to target both the aerobic capacity and the neuromotor function. The objective of this project is to prove that aerobic exercise during personalized FES-cycling could contribute to enhance the process of neuroplasticity, and to maintain locomotor and cardiovascular function in patients with stroke or spinal cord injury.

1. Introduction

Stroke and spinal cord injuries are among the leading causes of disability, and they severely impair the quality of life of patients [1, 2]. Regardless the etiology of the spinal cord injury, i.e., traumatic or non-traumatic, the clinical consequences of SCI depend on the severity and location of the injury, usually involving the loss of sensory and/or motor functions below the level of the injury. On the other hand, stroke occurs when there is a disruption in the perfusion of oxygenated blood to the brain. It can be ischemic (due to a deficient blood flow) or hemorrhagic (due to bleeding of the blood vessels), and the lack of oxygen causes the death of some brain cells, hence involving the loss of their associated neuronal function (either motor and/or cognitive). In both conditions, rehabilitation plays a key role in the prognosis of the patients. Thanks to brain plasticity, undamaged areas can reorganize their connectivity and

restore or substitute some of the lost functions.

There is a clear socioeconomic repercussion for all patients suffering from each type of neural injury, their families and caregivers, as well as the healthcare system. This underscores the need for developing approaches to promote the recovery of the patients, in such a way they could address the changes caused by the injury and achieve a certain level of quality of life and independence. In this sense, conventional therapies are fundamental in the earliest stages of the injury to promote neuroplasticity as much as possible in these patients. Nevertheless, they present some limitations that could be overcome by means of new technologies. On the one hand, the amount of time that the professionals can spend with each patient is limited, and helping the patients to perform certain movements could involve high physical effort for the therapists. On the other hand, rehabilitation technologies offer the possibility to supervise several patients at the same time, reducing the workload of the therapists; and they could provide objective and accurate data to assess patients' evolution through time, allowing the treatments to be adapted to them based on objective information, thereby increasing therapy intensity. These technologies are likely to decrease the overall assistance and derived costs that both patients and families need to face after a neurological injury. In view of this, efforts are being made to develop innovative rehabilitation therapies involving robotics, electrical stimulation or virtual reality, among other technologies. In this context, the implementation of a rehabilitation platform consisting of the application of functional electrical stimulation during cycling seems to have the potential to have an impact in both health and socioeconomic aspects.

Electrical stimulation is a widely used technique for the treatment of motor disorders, rehabilitation of traumatic injuries and sport training since the 1960's [3]. It consists of the application of electric pulse trains through transcutaneous (adhesive, adhered to the skin), percutaneous (inserted under the skin, at a shallow depth) or implantable electrodes (placed directly into the muscle or around the nerve of interest). These trains of pulses are rarely applied directly to the muscles (which are likely to be paralysed or weakened),

but they are often transmitted to the intact motor neurons, producing action potentials that travel until the muscles, triggering a coordinated contraction aimed at producing a given movement. The use of this technique for the generation of functional movements is known as Functional Electrical Stimulation (FES). Some of the applications of FES could be the recovery or assistance for walking, cycling, paddling, grasping objects, swallowing, breathing or swimming [2]. Performing repetitive task-oriented movements plays an important role when it comes to cortical reorganization and motor re-learning, based on neuroplastic mechanisms. In this sense, several studies claim that the application of FES combined with the voluntary intention of performing a specific motion could improve the outcomes of the therapy, according to the Hebbian learning theory, promoting motor cortex modulation and excitability by means of an improved afferent feedback [2, 4].

The main focus of this project is to combine functional electrical stimulation with cycling, using transcutaneous electrodes to selectively stimulate the muscles of the lower limbs during the physical activity. The combination of the electrical muscle stimulation with the benefits of cardiovascular exercise, allows for a greater participation and activation of the muscular groups involved in the cycling task, being FES-cycling an innovative approach with a huge potential to improve cardiovascular health and neuromotor function [5]. The objective of this paper is to present a FES-cycling platform which has been proposed as an innovative rehabilitation therapy for patients with SCI or stroke, and to analyse the state of the art on the use of this technology in patients with neurological injuries. With respect to the latter, focus has been placed on the demonstrated and potential clinical benefits, as well as the tools to assess the improvement.

The paper is organized as follows: section 2 reviews the existing literature related to the use of this technology to improve the cardiovascular health, neurological condition and motor function of patients who have suffered a SCI or a stroke. Section 3 describes the FES-cycling platform, explaining its architecture and presenting the main objectives of the project, as well as the expected results. Finally, section 4 summarizes the conclusions of the paper.

2. State of the art

For a better understanding of the potential of the proposed platform, previous studies that shed light on the evidence of the prospective benefits of FES-cycling in several fields have been reviewed.

2.1. Cardiovascular and metabolic benefits

From a cardiovascular and metabolic point of view, neurological injuries are likely to induce a sensorimotor neurological impairment, as well as to limit the autonomous nervous system activity and to alter the blood flow. This has a negative impact on the cardiorespiratory and vascular responses during physical exercise, thus affecting the physiological regulation mechanisms and leading to

some complications as deep vein thrombosis, pulmonary embolism, pressure sores, atelectasis, hypotension or autonomic disreflexia, among others. Cardiovascular complications are one of the leading causes of mortality and morbidity in people with neurological injuries (such as SCI or stroke). In order to deal with that associated risk, it is important to promote physical activity and sport, and to treat some of the risk factors that could help to recover physiological responses such as haemodynamic response or vasodilation/vasoconstriction mechanisms [6].

According to previous studies, FES-cycling has been shown to improve the cardiorespiratory fitness, since the heart and the lungs increase their activity to pump more oxygenated blood to the muscles that are working. Also, the muscle contractions favour the venous return and the metabolic cost is reduced thanks to metabolic efficiency [7]. As FES causes some fatigue, the blood flow is increased in the muscles and spasms are less likely to occur. Furthermore, the insulin resistance and the glucose tolerance are improved, hence reducing the risk of diabetes associated to physical inactivity [8]. Overall, the quality of life improves, as well as the physical health, well-being and social relationships. All these improvements can be assessed by means of some parameters as the aerobic capacity, the heart rate, the blood pressure, the cardiorespiratory resistance, the stroke volume, the cardiac output, the total peripheral resistance, the cholesterol level or the glucose tolerance; and there is already evidence of significant improvement in patients suffering from SCI or stroke [9].

2.2. Neurological benefits

After suffering a SCI, there is a disruption of descending pathways, which interrupts the communication between the brain and the region of the spinal cord in charge of producing movement. In a similar way, most of the patients experience some degree of paralysis after suffering a stroke, due to a disruption in the transmission of neural impulses between the brain and the muscles. Fortunately, the brain presents a high level of plasticity, which is the ability to change and adapt in response to new experiences, by creating new neural pathways and modifying existing ones. In this sense, it also undergoes a functional recovery, by transferring some functions from a damaged area to an undamaged one. Thanks to this, neurorehabilitation programs play a key role and are likely to achieve promising results in most of the cases.

When it comes to the neurological outcomes, FES-cycling could contribute to neural plasticity and brain connectivity improvement, since some adaptive mechanisms take place in the central nervous system as a result of selective stimulation of the muscles and repetitive functional movements. The fact of coupling a FES stimulus with a voluntary intention to perform a given movement promotes the recovery according to the Hebbian learning theory [4]. It could also play a role in the recovery of the sensory perception, since the activation threshold of the neurons are reduced and the signals coming from touch or proprioception are most likely to

reach the brain [2]. For the purpose of evaluating these neuroplastic changes, there are some non-invasive methods such as the Hoffman reflex, the F-wave or the motor evoked potentials, which can be analyzed using transcranial magnetic stimulation (TMS). These techniques could help to assess if there are changes in corticospinal excitability resulting from motor learning after a given treatment, to evaluate if the changes are related to the treatment, as well as to see if they are long-lasting.

2.3. Motor benefits

From a motor point of view, neurological injuries are responsible for a high level of physical disabilities, which affects the patients' quality of life. From this perspective, FES-cycling has the potential to prevent muscle atrophy, hence reducing weakness. It also improves several gait parameters and motor scores [10, 11, 12], as well as the muscle strength [13], the trunk stability and the range of motion [14, 15]; helping to recover some functionality and voluntary control over some paralyzed muscles. Besides, it prevents bone deterioration because of inactivity [16]. The assessment of gait improvements with walking ability outcome measures and clinical motor scores is essential to define the functional recovery of patients and the effectiveness of the FES-cycling therapy.

Cycling shares some similarities with walking, from a kinematic point of view, since they are both cyclic movements which require coordinated flexion and extension motions in the joints of the lower limbs. The range of motion of cycling is higher than the one needed for gait, thus being cycling a good option to be used as therapy to maintain the functional range of motion required for walking, particularly in the ankle and knee joints. It is also important to mention that cycling is done while sitting down, reducing the risk of falls because of unsteadiness and loss of balance, hence overcoming one of the limitations present in other therapies [1]. Furthermore, it is an accessible and relatively affordable technology, compared to other approaches, such as exoskeletons, for example.

3. Rehabilitation platform

The aim is to develop an accessible and affordable FES-cycling platform, which could be used in combination with conventional therapies performed either at the rehabilitation centers or as a home-based therapy. The platform consists of an ergometer, an electrical stimulator, a PC, electrodes and sensors to control the delivered stimulation and the exercise load and intensity. Besides, the platform will incorporate a user interface to allow the clinical staff to configure the therapy in an easy and fast manner, being able to offer a flexible and personalised therapy to each patient. For the purpose of maximizing the therapeutic benefits, speeding up the recovery of patients with neurological injuries, and enhancing motivation and therapy adherence, the platform has the potential to be combined with other technologies that could provide visual and auditory feedback to the user. For the development of this training platform, it

is important to follow a user-centered design approach, which considers several application scenarios and incorporates tools and evaluation protocols to assess the efficacy of the system.

Regarding the functioning of the platform (see Figure 1), a MOTomed viva2 ergometer with a magnetic encoder along with a microprocessor reads the pedalling angle and sends it to an Arduino microcontroller. The pedalling angles from the Arduino are received in the PC in real-time and used to calculate the corresponding stimulation strategy that must be applied in each moment, in order to achieve a cycling movement. This stimulation strategy is based on muscle activation patterns of healthy subjects during cycling, measured with surface electromyography (EMG). Once calculated, the stimulation commands are sent to the stimulator (RehaStim, Hasomed), which delivers the electrical pulses to the target muscles of the patient (quadriceps, hamstrings, gastrocnemius and tibialis anterior) through transcutaneous electrodes. The stimulation strategy is configured by the therapist using an intuitive and user-friendly interface displayed on the PC, which enables the adaptation of both the electrical stimulation parameters and the cycling pattern to meet the needs of each patient. As for the visual and/or auditory feedback, it is thought to be useful to have a system capable of providing information about the patient's performance (cadence, time, etc.) to serve as motivation and to promote adherence to the therapy [17, 18].

The definition of a high-intensity training program is crucial for optimizing recovery and functional gains among patients. This program can be tailored to variables such as heart rate, systolic blood pressure and perceived exertion, ensuring personalized and efficient exercise routines. A structured 4-8 week program, consisting of 4-6 weekly 30-minute sessions, offers a balance between effectiveness and safety.



Figure 1. Overview of the rehabilitation platform components.

Concerning the different assessment tools that could be used to quantify the impact of the therapy, it has been decided to focus on the heart-rate, the oxygen consumption, the cortical excitability, as well as the improvement in motor scores (such as the Lower Extremity Motor Score, LEMS) and gait-related measures, including the 6-Minute Walking Test (6MWT).

4. Discussion and conclusions

After a detailed literature review of the potential benefits of a FES-cycling rehabilitation program in people with SCI or stroke, promising improvements could be expected from a cardiovascular, neurological and motor points of view. However, there is a lack of a methodology to address all those potential benefits at the same time. Moreover, the existing studies seem to be quite heterogeneous regarding patient selection (type and level of the injury, time since injury, age, etc.), training dosage and outcomes evaluation, making it challenging to generalize the results. Therefore, the development of protocols tailored to different application scenarios is necessary to discern the therapy's resulting improvements.

The proposed FES-cycling rehabilitation platform is a good candidate for simultaneously targeting the recovery of cardiovascular fitness, neurological and motor functions, while being able to quantify the improvements in a reliable manner.

Acknowledgments

This work is partially developed within the Research Network FUSION “Red Temática Sobre Fusión de Tecnologías Robóticas y Estimulación Eléctrica Neuromuscular para Neurorehabilitación de Trastornos del Movimiento”, grant by Agencia Estatal de Investigación (RED2022-134319-T).

References

- [1] David Barbosa, Cristina P. Santos, et al. The application of cycling and cycling combined with feedback in the rehabilitation of stroke patients: A review. *Journal of Stroke and Cerebrovascular Diseases*, 24(2):253–273, 2015.
- [2] Shiyu Luo, Haonan Xu, et al. A review of functional electrical stimulation treatment in spinal cord injury. *NeuroMolecular Medicine*, 22(4):447–463, 2020.
- [3] W. T. Liberson, H. J. Holmquest, et al. Functional electrotherapy: stimulation of the peroneal nerve synchronized with the swing phase of the gait of hemiplegic patients. *Archives of Physical Medicine and Rehabilitation*, 42:101–105, Feb 1961.
- [4] Lynsey D. Duffell, Sue Paddison, et al. The effects of fes cycling combined with virtual reality racing biofeedback on voluntary function after incomplete sci: a pilot study. *Journal of NeuroEngineering and Rehabilitation*, 16(1):149, 2019.
- [5] Diego Torricelli, Cristiano De Marchis, et al. Reorganization of muscle coordination underlying motor learning in cycling tasks. *Frontiers in Bioengineering and Biotechnology*, 8, 2020.
- [6] Elizabeth Partida, Eugene Mironets, et al. Cardiovascular dysfunction following spinal cord injury. *Neural regeneration research*, 11(2):189–194, Feb 2016.
- [7] N. Hasnan, J. P. Engkasan, et al. High-intensity virtual-reality arm plus fes-leg interval training in individuals with spinal cord injury. *Biomedizinische Technik. Biomedical engineering*, 58 Suppl 1:10.1515/bmt-4028. Epub 2013 Sep 7, Aug 2013.
- [8] E. Yaşar, B. Yılmaz, et al. The effect of functional electrical stimulation cycling on late functional improvement in patients with chronic incomplete spinal cord injury. *Spinal Cord*, 53(12):866–869, 2015.
- [9] Jan W. van der Scheer, Victoria Goosey-Tolfrey, et al. Functional electrical stimulation cycling exercise after spinal cord injury: a systematic review of health and fitness-related outcomes. *Journal of NeuroEngineering and Rehabilitation*, 18(1):99, 2021.
- [10] S. Mazzoleni, E. Battini, et al. An integrated gait rehabilitation training based on functional electrical stimulation cycling and overground robotic exoskeleton in complete spinal cord injury patients: Preliminary results. *IEEE International Conference on Rehabilitation Robotics*, 2017:289–293, Jul 2017.
- [11] Chia-Ying Fang, Angela Shin-Yu Lien, et al. The effect and dose-response of functional electrical stimulation cycling training on spasticity in individuals with spinal cord injury: A systematic review with meta-analysis. *Frontiers in Physiology*, 12, 2021.
- [12] Simona Ferrante, Emilia Ambrosini, et al. A biofeedback cycling training to improve locomotion: a case series study based on gait pattern classification of 153 chronic stroke patients. *Journal of NeuroEngineering and Rehabilitation*, 8(1):47, 2011.
- [13] Emilia Ambrosini, Simona Ferrante, et al. Cycling induced by electrical stimulation improves muscle activation and symmetry during pedaling in hemiparetic patients. *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, 20:320–30, 2012.
- [14] Hsin-Chang Lo, Yung-Chun Hsu, et al. Cycling exercise with functional electrical stimulation improves postural control in stroke patients. *Gait posture*, 35(3):506–510, 2012.
- [15] Chul Song, Jong Kim, et al. A new postural balance control system for rehabilitation training based on virtual cycling. *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, 8:200–7, 2004.
- [16] Angela Frotzler, Sylvie Coupaud, et al. High-volume fes-cycling partially reverses bone loss in people with chronic spinal cord injury. *Bone*, 43(1):169–176, Jul 2008.
- [17] Emilia Ambrosini, Elisabetta Peri, et al. A multimodal training with visual biofeedback in subacute stroke survivors: a randomized controlled trial. *European journal of physical and rehabilitation medicine*, 56, 2019.
- [18] Richard G. Ranky, Mark L. Sivak, et al. Modular mechatronic system for stationary bicycles interfaced with virtual environment for rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, 11(1):93, 2014.

An inexpensive arbitrary waveform neurostimulator for the selective activation of neurons in retinal prosthesis

P. Jimenez-Fernandez¹, H. Guzman-Miranda¹, A. Barriga-Rivera^{2,3}

¹ Department of Electronic Engineering, Universidad de Sevilla, Sevilla, Spain, hguzman@us.es

² Department of Applied Physics III, Universidad de Sevilla, Sevilla, Spain, abriviera@us.es

³ School of Biomedical Engineering, University of Sydney, Sydney, Australia, alejandro.barriga-rivera@sydney.edu.au

Abstract

This contribution presents a two-channel inexpensive arbitrary waveform neurostimulator based on a Raspberry Pi microcomputer platform and a Howland voltage-to-current converter. The system has been designed to enable the delivery of common stimulation strategies used in visual prosthesis research.

1. Introduction

Retinal electrostimulation has enabled the restoration of visual sensations in patients suffering from retinal degeneration diseases [1, 2]. By delivering current pulses, it is possible to bypass the photoreceptors and activate the surviving neurons in the retina. These electrical pulses will be eventually converted into artificially-generated visual information in the form of sequences of action potentials.

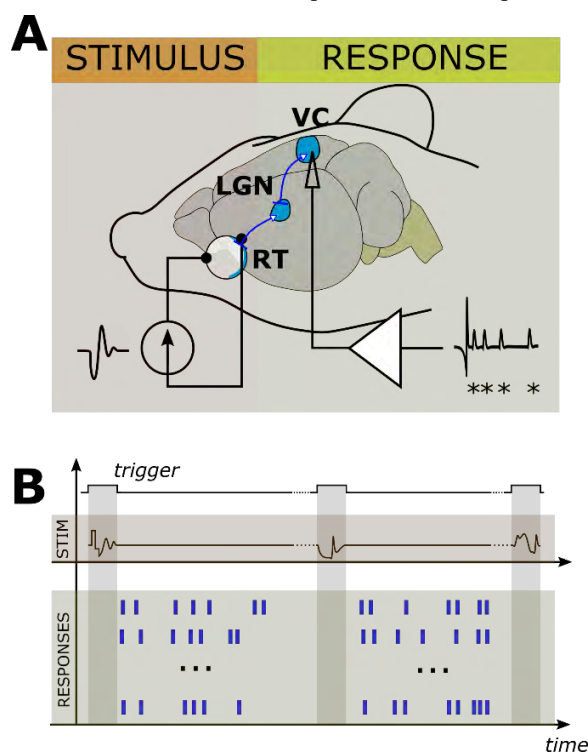


Figure 1. Schematic showing the typical experimental setup used in *in vivo* visual prosthesis research. Panel A illustrates the delivery of an arbitrary current waveform to the retina (RT) of an experimental subject. Neural responses thus generated travel to the lateral geniculate nucleus (LGN) and, subsequently, to the visual cortex (VC). An array of recording electrodes is interfaced to a neural amplifier used to obtain the neural responses. Panel B shows an schematic of the occurrence of stimuli and responses governed by a trigger signal.

This information, referred to here as ‘neural message’, propagates downstream to higher visual centres to ultimately elicit a visual sensation. There are, however, two important limitations in this process [3]: (1) the neural messages thus generated differ from those in physiological vision, and (2) the brain has undergone an important neural remodelling that may alter the way visual information is processed.

The visual prosthesis research community continues to investigate new ways of producing more physiological neural messages [4]. While classical constant-current pulses and current-steering techniques can allow for fine spatial [5, 6] and in-depth [7] recruitment of retinal neurons, kilohertz pulse trains have demonstrated, *in vitro* and *in silico*, the potential of non-conventional current waveforms to produce the selective activation of different retinal pathways [8, 9]. To date, these stimulation strategies have relied on the use of amplitude-modulated high-frequency pulse trains, as most of the neural stimulators available rely on the generation of constant-current pulses. Here we report the design of an inexpensive arbitrary waveform neural stimulator for research applications in visual prosthesis.

2. Description of the system.

Visual prosthesis research requires complex and expensive setups [10]. These experimental setups typically include a multi-channel neural recording system and a neurostimulator, both connected to deliver the targeted stimulation strategy and to record the neural activity thus elicited. While communication and synchronisation of these subsystems can be achieved in many ways, trigger signals are normally employed to indicate the delivery of a given stimulus, as shown in figure 1. This signal can be generated by the recording subsystem, by the stimulator, or by an external trigger source.

The design of the neural stimulator presented here was conceived to deliver arbitrary current waveforms with frequency components up to 30 kHz to ensure the kilohertz frequencies previously described in preferential activation experiments was attainable [9]. Secondly, given the current levels previously reported to drive ganged [11] or large electrodes [12, 13], the stimulator was design to allow for the delivery of current levels of up to ± 3 mA. Note the expected impedance value ranges from 100 Ω to approximately 10 k Ω . And thirdly, a neural stimulator designed for research must allow for different experimental

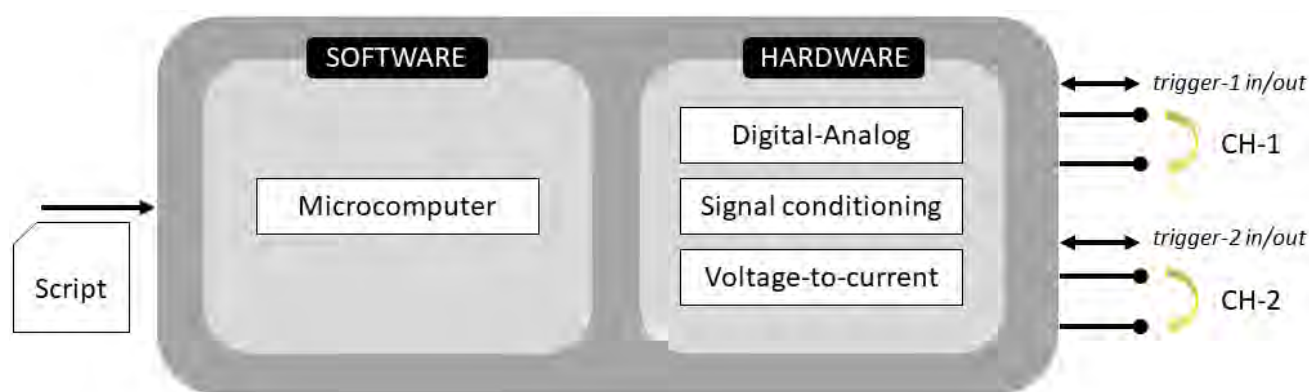


Figure 2. Conceptual diagram illustrating the subsystems of the two-channel arbitrary neural stimulator. It comprises a software subsystem running on a microcomputer, and a hardware subsystem for the generation of the current waveforms (digital-to-analog conversion, signal conditioning, and voltage-to-current conversion). A script file contains a set of instructions for the delivery of a series of stimuli and the control of the trigger signals of each channel (CH-1 and CH-2).

configurations. For that purpose, we used a microcomputer system to easily program (1) the current waveforms to be delivered, (2) the trigger regimes to be used, and (3) the workflow of the experiments.

3. Logical view

The logical view, adapted from the 4+1 view model reported by Krutchen [14], provides a picture of the different constituents of the system. The two-channel arbitrary neural stimulator described herein includes two main subsystems: a hardware subsystem and a software subsystem. The neurostimulator software runs on a Raspberry Pi (RPI) 3 Model B (Raspberry Pi Foundation, Cambridge, UK) microcomputer running Raspbian operating system. This software receives a script file containing a set of instructions via USB serial communications. These instructions codify the samples of each arbitrary waveform to be delivered as well as the inter-stimulus delays and the trigger configurations thus established. Said microcomputer is interfaced to a custom hardware subsystem designed for the generation of analog signals, and for the management of the trigger signals required to deliver the target stimulation regimen, as shown in figure 2. The output of the stimulator is to be connected to the stimulation electrodes. Shorting of the electrodes is controlled by the trigger signal. This trigger signal can be either internally generated by the neurostimulator, or received from an external device.

4. Physical view

The physical view describes the hardware components of the neurostimulator. The RPi 3 microcomputer was chosen in this neurostimulator as it is broadly used in academic applications due to the availability and the maturity of the system. On the one hand, the microcomputer provides the physical substrate for the software subsystem. On the other hand, the hardware subsystem referred to in the previous section required further description from a physical viewpoint. These subsystems were implemented as electronic circuits on a custom printed circuit board (PCB).

4.1. Power supply

The voltage compliance limits in this system have to allow for driving 10-k Ω impedances with currents of ± 3 mA. To ensure this requirement is met, we connected three TMA0515D (Traco Power, Sihlbruggstrasse, Switzerland) isolated DC-DC converters in series to provide ± 45 V. Additionally, we also included a L7805CV (STMicroelectronics, Geneva, Switzerland) voltage regulator to provide a 5-V supply to the signal conditioning circuits, as shown in figure 3.

4.2. Digital-analog conversion (DAC)

The RPi microcomputer generates digital data to codify the samples of the current waveforms at approximately 166 kHz. These data are transmitted *via* I²C protocol to a MCP4725 (Microchip Technologies, Chandle, USA) digital-to-analog module using 12 bits per sample.

4.3. Signal conditioning

Signal conditioning here includes a low-pass filtering stage, a voltage level adjustment and signal amplification. An RC low-pass filter was implemented to provide a 3-dB cut-off frequency of 160 kHz. Subsequently, voltage signal was linearly scale to adjust the 0-to-5 voltage range provided by the DAC to a symmetrical range between -3 and 3 V. For that purpose, a TL082CP (Texas Instruments, Dallas, USA) operational amplifier was topologically configured as a summing amplifier to implement equation (1), where V_{DAC} represents the voltage value provided after low pass filtering, and $V_{adjusted}$ is the adjusted voltage output.

$$V_{adjusted} = \frac{5}{6}V_{DAC} - 3 \quad (1)$$

4.4. Signal isolation.

The voltage signal thus adjusted needs to be electrically isolated from the common ground. To provide this isolated grounding, an isolation amplifier ISO124-P (Texas Instruments, Dallas, USA) was employed. The gain of this stage was 1, and consequently, it is referred to as an isolation buffer in figure 2.

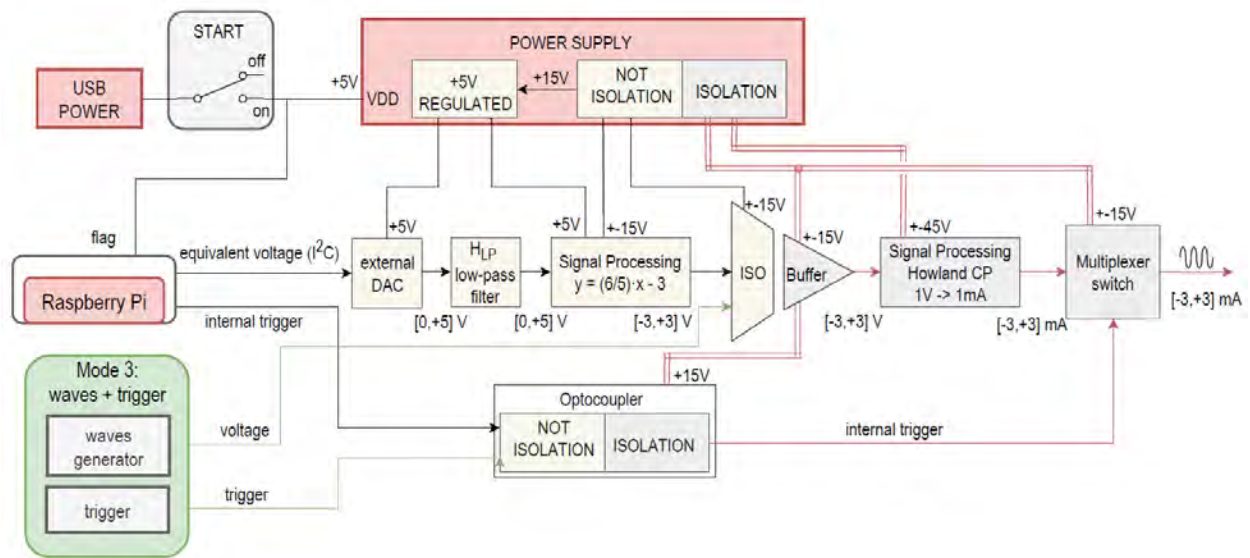


Figure 2. Block diagram of the neurostimulator circuitry. The electronics are powered using the 5-V supply provided by the Raspberry Pi board. The power supply unit provides all required voltages to the other subsystems. The analog-to-digital (DAC) converter is driven by the Raspberry Pi using the I²C communication protocol. The voltage waveform thus generated is low-pass filtered and linearly scaled between +3 and -3 V, and subsequently isolated from the common ground using an isolation amplifier (ISO). It is then converted to a current waveform using a Howland voltage-to-current converter. Shorting of the electrodes is achieved by an analog multiplexer digitally controlled by an electrically isolated trigger signal. Additionally, an external arbitrary function generator can be wired to the single processing subsystems to generate other types of current waveforms (functioning mode 3).

4.5. Voltage-to-current conversion.

Neural stimulation is normally achieved using current-controlled devices, as it allows for controlling the amount of charge delivered and recovered during the stimulus. Thus, the arbitrary voltage waveforms obtained are to be converted into current waveforms. For that purpose, an improved Howland current source [15] was implemented using a high-voltage precision operational amplifier ADA4700 (Analog Devices, Wilmington, USA) power at ± 45 V to provide a wide voltage compliance range as shown in figure 3. Note the conversion rate achieved with this circuit was 1 mA·V⁻¹.

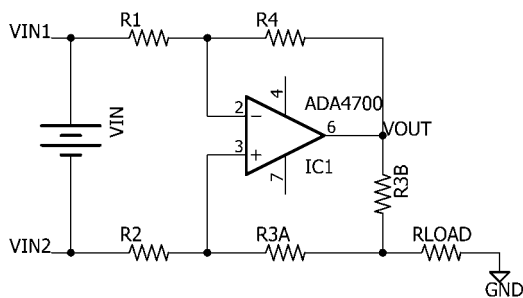


Figure 3. Circuit diagram of the Howland current source implemented in the arbitrary neurostimulator.

4.6. Trigger and electrode shorting.

The trigger signal, as described previously, is used to provide a time mark for the stimulus delivery. The trigger signal is also used to connect and disconnect the neurostimulator to the electrodes. On the one hand, with a high-level input voltage the output of the current source is connected to the target electrodes. On the other hand, with a low-level input voltage the terminals of the electrodes are shorted, and the output terminals of the current source are

also shorted. This is achieved using a CD4053BE (Texas Instruments, Dallas, USA) digitally-controlled analog multiplexer mounted as illustrated in figure 4. The trigger signal can be internally generated (functioning mode 1), or received from an external device (functioning mode 2).

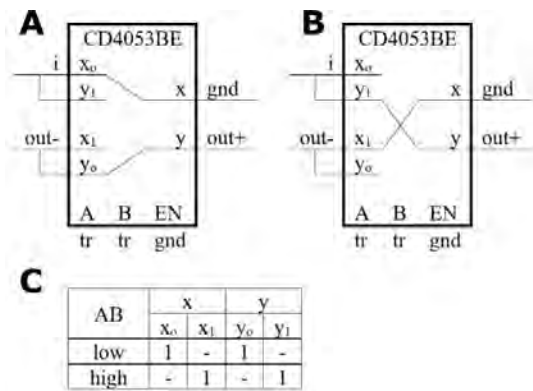


Figure 4. Connection diagram of the digitally-controlled analog multiplexer used in the arbitrary neurostimulator to allow for the shorting of the electrodes. Panel A and B illustrate the internal switching for high- and low- logic level of the trigger signal respectively. Panel C shows the electrical configuration required in terminals A and B of the CD4053BE chip.

5. Development and process views

We have described the hardware components of the arbitrary wave for neurostimulator herein presented. However, the system needs a custom software to implement a given stimulation strategy. For that purpose, the software subsystem has been developed by combining two software elements that interact to facilitate the user experience. Hence, the development and the process view referred to in the 4+1 architecture model have been combined in this section.

5.1. Matlab scripting.

The definition of the stimulus waveforms is carried out using a custom Matlab script. Briefly, this script generates all waveform samples required to drive the DAC and interstimulus delay times to be applied. The script, executed on a personal computer, generates a file containing said information. This file is to be transferred onto the RPi system using a USB memory stick.

5.2. Script execution.

A custom software developed in C language has been programmed to read the data file containing stimulus parameters and to drive the DAC of the neurostimulator using the I²C communication protocol.

5.3. Communication and interaction.

As described previously, the user defines the stimulation strategy. This typically consists of a series of stimuli, defined by the waveform samples, the trigger values (ON/OFF) and the inter-stimulus time, as illustrated in figure 1. Once the user runs the Rpi software, the neurostimulator will proceed with the delivery of the stimulation regimen programmed in the input file.

6. Experimental validation and discussion.

The neurostimulator has been tested using a 100 k Ω to ensure the system complies with strict requirements. While the neurostimulator was able to meet the current and bandwidth requirements established, a substantial ringing was observed for low-current levels, as shown in figure 5. This appears due to the low-voltage input at the entrance of the isolation buffer. This limitation can be improved by either re-scaling the gains of the signal conditioning and the voltage-to-current stages, or by using higher-end isolation buffers. Further testing is required *in vitro* and *in vivo*. In addition, while this is an experimental prototype, the channel count will need to be increased when driving multi-electrode retinal prostheses.

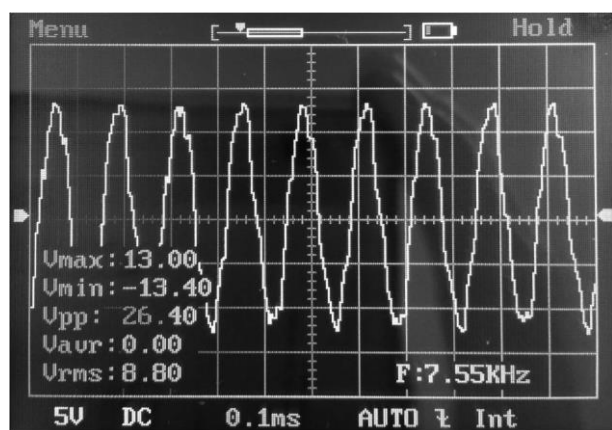


Figure 5. Voltage waveform recorded using an isolated oscilloscope from a 100-k Ω driven by a 200- μ A sinusoidal current waveform at 7.5 kHz.

Acknowledgements

This work has been funded by grant CNS2022-135870 funded by MCIN/AEI/ 10.13039/501100011033 and by “European Union NextGenerationEU/PRTR”, and

supported by ‘FUNDALUCE’ (Foundation for Fighting Blindness). Part of this work was developed by the first author as part of final year project, submitted to the University of Seville in partial fulfilment for the requirements of his undergraduate degree.

References

- [1] P. J. Allen *et al.*, "Ten year safety and stability results of a suprachoroidal retinal prosthesis," *Investigative Ophthalmology & Visual Science*, vol. 64, no. 8, pp. 3827-3827, 2023.
- [2] L. N. Ayton *et al.*, "An update on retinal prostheses," *Clinical Neurophysiology*, vol. 131, no. 6, pp. 1383-1398, 2020.
- [3] D. Caravaca-Rodriguez, S. P. Gaytan, G. J. Suaning, and A. Barriga-Rivera, "Implications of neural plasticity in retinal prosthesis," *Investigative Ophthalmology & Visual Science*, vol. 63, no. 11, pp. 11-11, 2022.
- [4] W. Tong, H. Meffin, D. J. Garrett, and M. R. Ibbotson, "Stimulation strategies for improving the resolution of retinal prostheses," *Frontiers in neuroscience*, vol. 14, p. 262, 2020.
- [5] A. Barriga-Rivera, J. W. Morley, N. H. Lovell, and G. J. Suaning, "Cortical responses following simultaneous and sequential retinal neurostimulation with different return configurations," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016: IEEE, pp. 5435-5438.
- [6] G. Dumm, J. B. Fallon, C. E. Williams, and M. N. Shivasani, "Virtual electrodes by current steering in retinal prostheses," *Investigative ophthalmology & visual science*, vol. 55, no. 12, pp. 8077-8085, 2014.
- [7] Z. C. Chen, B.-Y. Wang, and D. Palanker, "Real-time optimization of the current steering for visual prosthesis," in *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2021: IEEE, pp. 592-596.
- [8] P. Twyford, C. Cai, and S. Fried, "Differential responses to high-frequency electrical stimulation in ON and OFF retinal ganglion cells," *Journal of neural engineering*, vol. 11, no. 2, p. 025001, 2014.
- [9] T. Guo *et al.*, "Closed-loop efficient searching of optimal electrical stimulation parameters for preferential excitation of retinal ganglion cells," *Frontiers in neuroscience*, vol. 12, p. 168, 2018.
- [10] A. Barriga-Rivera *et al.*, "A 4+ 1 architecture for *in vivo* electrophysiology visual prosthesis," *Journal of Accessibility and Design for All*, pp. 81-101, 2016.
- [11] A. Barriga-Rivera *et al.*, "High-amplitude electrical stimulation can reduce elicited neuronal activity in visual prosthesis," *Scientific reports*, vol. 7, no. 1, p. 42682, 2017.
- [12] A. Barriga-Rivera, T. Guo, J. W. Morley, N. H. Lovell, and G. J. Suaning, "Retinal electrostimulation in rats: activation thresholds from superior colliculus and visual cortex recordings," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017: IEEE, pp. 1166-1169.
- [13] A. Barriga-Rivera, V. Tatarinoff, N. H. Lovell, J. W. Morley, and G. J. Suaning, "Long-term anesthetic protocol in rats: feasibility in electrophysiology studies in visual prosthesis," *Veterinary ophthalmology*, vol. 21, no. 3, pp. 290-297, 2018.
- [14] P. B. Kruchten, "The 4+ 1 view model of architecture," *IEEE software*, vol. 12, no. 6, pp. 42-50, 1995.
- [15] A. Mahnam, H. Yazdani, and M. M. Samani, "Comprehensive study of Howland circuit with non-ideal components to design high performance current pumps," *Measurement*, vol. 82, pp. 94-104, 2016.

Análisis del estrés durante procedimientos quirúrgicos mediante laparoscopia convencional y robótica

M. J. Pérez Salazar¹, J. A. Sánchez Margallo¹, F. M. González Nuño¹, M. Kappel^{1,2}, F. M. Sánchez Margallo¹

¹ Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España, {mjperetz, jasanchez, fgonzalez, msanchez@ccmijesususon.com}

² Escuela de Ingeniería Superior, Universidad Sigma-Clermont, Aubiére, Francia, {mathilde.kappel@sigma-clermont.fr}

Resumen

Actualmente la cirugía moderna se encuentra en un punto de inflexión entre el uso de las técnicas quirúrgicas convencionales y los sistemas robóticos a la hora de decidir cuál usar durante una cirugía. Los robots quirúrgicos cada vez son más comunes en los centros hospitalarios siendo controlados por cirujanos de distintos ámbitos y niveles de experiencia. La aparición de este cambio modifica no solo la manera en la que se va a intervenir al paciente y el resultado final, sino también la mecánica del procedimiento y su impacto desde el punto de vista del cirujano. En este estudio se procede a evaluar el nivel de estrés que sufre el cirujano durante procedimientos quirúrgicos y tareas básicas en simulador. Estas intervenciones se llevaron a cabo por un grupo de 7 sujetos (4 en simulador y 3 en modelo experimental) que repetirían cada uno de los procedimientos tanto en cirugía laparoscópica convencional como en robótica. Durante el desarrollo de las tareas se registraron datos sobre el nivel de estrés, usando un sensor de la actividad electrodermal (EDA), junto con un cuestionario subjetivo sobre la carga de trabajo percibida (SURG-TLX), incluyendo el estrés, al finalizar la intervención. En este estudio se plantea calcular de forma objetiva el nivel de estrés del cirujano durante la práctica quirúrgica y comparar los valores mostrados entre la técnica laparoscópica convencional y robótica, así como su correlación con los valores subjetivos de estrés registrados.

1. Material y métodos

1.1. Destreza quirúrgica (errores/tiempo)

Cada tarea en simulador fue registrada en vídeo para su posterior evaluación por un evaluador externo experto, el cual analizó los errores cometidos en cada tarea teniendo en cuenta la dinámica de ejecución de cada una de ellas. Por norma general, se daba a entender como error no cumplir con el protocolo de actuación para la tarea, como dejar caer el objeto a la base del simulador o no cumplir con el mínimo de repeticiones u objetivos en el tiempo máximo acordado.

El tiempo total que fue necesario para el cumplimiento de cada una de las tareas se registró en la propia capturadora de vídeo utilizada para la evaluación de errores.

1.2. Análisis del estrés (EDA/SURG-TLX)

Para determinar el nivel de estrés al que se enfrentaba cada cirujano durante la actividad quirúrgica se hizo uso de dos procedimientos distintos: Un sensor para el análisis de la actividad eléctrica de la piel (EDA) y un

cuestionario subjetivo que entre otras cuestiones se les preguntaba por la demanda mental y estrés que sufrieron durante el desarrollo de la tarea o procedimiento quirúrgico.



Figura 1. Sensor para análisis de actividad electrodermica (EDA)

En el caso del sensor para el análisis de la actividad electrodermica (EDA), se empleó el dispositivo *EdaMove 4* (Movisens, Karlsruhe, Alemania) (Figura 1) para la adquisición de la señal mediante el uso de unos electrodos adheridos a la piel de la cara interior del tobillo, con el objetivo de registrar la conductancia de la piel. Este parámetro se ve afectado por variaciones en la superficie de la piel, como es el caso de la secreción de sudor debido al estrés provocado. Los resultados son procesados a través de un software específico (*Unisens Viewer*; *Movisens*) que genera, entre otros datos, gráficas que comparan la variación que se produce entre la impedancia de la piel (μS) y el tiempo transcurrido.

Al finalizar la actividad se les daba acceso a los sujetos para poder rellenar la encuesta del SURG-TLX [1]. En este punto, y de forma subjetiva, el cirujano debía puntuar, en un rango comprendido entre 1 y 10, aspectos como la demanda mental y física, temporal, complejidad y estrés experimentado entre otros, siendo este último el más significativo para nuestro caso de estudio.

1.3. Tareas y procedimientos

Cada tarea y procedimiento quirúrgico fue realizado de forma aleatoria mediante laparoscopia convencional y mediante el uso de la plataforma robótica Versius Surgical System (CMR Surgical; Cambridge, Reino Unido) (Figura 2).



Figura 2. Uso del sistema robótico VERSIUS con el simulador de corte

El protocolo de actuación fue diferente dependiendo del grupo de tareas en simulador o procedimientos en modelo experimental: Una serie de 7 tareas en simulador, con una anastomosis para los expertos, y una intervención in vivo por cada especialidad evaluada (cirugía general, urología y ginecología).

1.3.1 Tareas en simulador

En este grupo de tareas participaron 4 cirujanos con diferentes niveles de experiencia en cirugía laparoscópica (2 expertos y 2 noveles, con menos de 10 procedimientos en cirugía laparoscópica). Cada uno llevó a cabo un total de 7 tareas en un simulador, tanto para laparoscopia convencional como usando la plataforma VERSIUS, buscando evaluar una serie de habilidades quirúrgicas básicas como son la coordinación ojo/mano, precisión y destreza, entre otros. Las tareas realizadas fueron las siguientes:

- **Peg Transfer:** Consistía en trasladar 3 piezas de goma con forma de toroides alargados de un poste a otro y vuelta, cambiando la pieza de mano durante el proceso y con un tiempo máximo de 10 minutos para su realización.
- **Legumbres:** Añadiendo dificultad a la tarea anterior, se colocaron unas legumbres con distinta rugosidad en cada extremo de una pletina cuadrada, teniendo el sujeto que trasladarlas a la esquina opuesta y vuelta con cambio de mano. El tiempo máximo fueron de 10 minutos.
- **Laberinto:** Un camino de 10 aros, en los que se indicaba el orden a tomar y el sentido, donde los sujetos buscaban “enhebrar” una aguja a través de los aros introduciéndola con la mano dominante y extrayéndola con la contraria. Se disponían de 10 minutos.
- **Simulador de corte:** Esta tarea se debía realizar 2 veces, una con la mano dominante y otra con la opuesta, contando así como dos pruebas distintas. El cirujano debía cortar una pletina según el patrón que tuviera dibujado (Figura 2) teniendo 5 minutos máximos para cada una de las 2 pruebas.
- **Disección:** En esta tarea se hizo uso de un modelo ex vivo de estómago porcino. El procedimiento consistía en una incisión de 3 cm de longitud y luego separar la capa serosa del resto sin perforar la muscular. Se disponía de 5 minutos como máximo.

- **Sutura:** Aprovechando el modelo ex vivo diseccionado en la tarea anterior, se procede a realizar una sutura que consiste en el paso de la aguja y posteriormente un nudo doble y dos simples (en sentidos opuestos), con un máximo de 10 minutos y asegurándose que la apertura queda bien cerrada.

Además, los cirujanos expertos realizaron una última tarea. Esta consistía en una anastomosis uretrovesical en modelo ex vivo porcino, con hasta 8 puntos de sutura y 30 min como límite de tiempo para el caso convencional, siendo sin límite en el caso robótico debido a su uso novedoso.

1.3.2 Procedimiento quirúrgico

Los procedimientos quirúrgicos en modelo experimental se llevaron a cabo por 3 cirujanos, cada uno experto en un campo específico. Todos los procedimientos fueron aprobados por el comité de ética en experimentación animal pertinente (EXP-20230313, EXP-20230320, EXP-20230325). Cada cirujano llevó a cabo un conjunto de procedimientos quirúrgicos, tanto en cirugía laparoscópica convencional como en cirugía robótica, dependiendo de su especialidad:

- **Cirugía General:** Entre otros procedimientos realizados (colecistectomía, funduplicatura de Nissen y colectomía), se evaluó la gastrotomía.
- **Urología:** Entre otros procedimientos llevados a cabo (prostatectomía y nefrectomía parcial), se evaluó la nefrectomía total.
- **Ginecología:** Entre otros procedimientos realizados para esta especialidad (ovariectomía, miomectomía e histerectomía), se escogió la ligadura de trompas por coagulación, una trompa por cada técnica quirúrgica.

2. Resultados

Los datos extraídos de las pruebas en quirófano y su posterior análisis se muestran en este apartado, distinguiendo por un lado los resultados de las pruebas llevadas a cabo en simulador y, por otro lado, los datos obtenidos de las intervenciones quirúrgicas en modelo experimental, con el objetivo de separar ambos bloques.

Para el análisis de EDA, se evalúan 3 tipos distintos en los que las gráficas se pueden mostrar en función del estrés al que ha sido sometido el sujeto, siguiendo la metodología presentada por [2]. Por orden de mayor a menor estrés: Tipo 1 si existe un estrés inicial que no decrece y persiste en el tiempo, tipo 2 si existen fluctuaciones durante el desarrollo de las pruebas y tipo 3 si no representa ningún valor significativo durante la intervención, o en caso de aparecer al inicio este decrece exponencialmente hasta disiparse.

2.1. Pruebas en simulador

Durante las pruebas en simulador no se finalizó la ejecución del estudio mediante cirugía laparoscópica convencional del sujeto 4 por razones externas. Por tanto,

sus resultados, al no ser concluyentes, fueron descartados del estudio.

Teniendo en cuenta los datos recogidos durante las intervenciones en las que se hacía uso de la plataforma robótica y el método tradicional, los resultados de errores y tiempo transcurrido para cada tarea en simulador son mostrados en las Tablas 1 y 2.

	S1	S2	S3	S4
Peg transfer	1/5'5	4/6	0/3	0/2
Legumbres	2/4'5	2/4	3/4	2/2'5
Laberinto	5/10	9/10	7/9	6/6
Corte 1	0/3	0/5	0/5	0/2'5
Corte 2	0/2'5	0/3	0/3	0/2'5
Disección	0/5	0/5	0/5	0/1'5
Sutura	7/8'5	3/4'5	0/5	0/4'5
Anastomosis (Exp)	-	8/60	-	0/45

Tabla 1. Número de errores y tiempo (min) de realización respectivamente en simulador con sistema robótico

	S1	S2	S3
Peg transfer	1/2	1/1'5	2/2
Legumbres	5/3'5	0/0'5	4/3
Laberinto	1/9	0/2	10/8
Corte 1	0/2	0/1'5	0/2
Corte 2	1/2	0/2	0/2'5
Disección	0/5	0/3	0/5
Sutura	1/5	0/1'5	0/7
Anastomosis (Exp)	-	0/20	-

Tabla 2. Número de errores y tiempo (min) de realización respectivamente en simulador con laparoscopia convencional

Se puede observar que, si comparamos los errores cometidos durante las tareas en laparoscopia robótica y convencional (Figura 3), por norma general durante el desarrollo de las tareas robóticas los cirujanos noveles cometen menos errores en diversas tareas, mostrando la capacidad para adaptarse mejor a nuevas herramientas de asistencia quirúrgica.

El tiempo de realización, por otro lado, disminuye significativamente utilizando la técnica laparoscópica convencional, siendo especialmente pronunciada esta diferencia en el caso de expertos en laparoscopia debido a su familiarización con esta técnica frente al uso de la plataforma robótica.

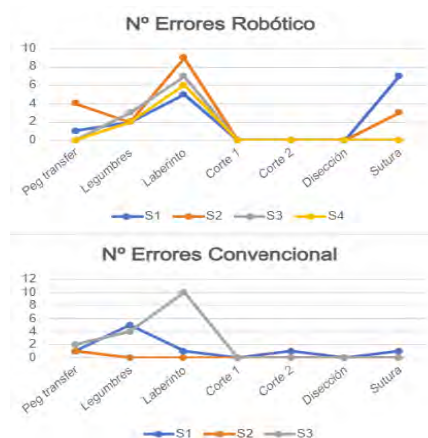


Figura 3. Representación gráfica comparativa del número de errores durante las intervenciones

Conociendo estos resultados, es de esperar que en los sujetos noveles los niveles de estrés representados por EDA sean del tipo 2, teniendo en cuenta los posibles picos en la señal generados por la falta de experiencia. Para el caso de los expertos, los resultados cambian, mostrándose del tipo 3 principalmente al no hallar apenas durante las intervenciones situaciones que varíen significativamente la demanda mental.

Atendiendo a la comparación entre los valores medios totales obtenidos por el sensor para cada categoría de sujetos (Figura 4), y considerando ambas metodologías quirúrgicas analizadas, se puede concluir con que siendo similar el estrés total experimentado por noveles y expertos, en el caso de los principiantes se demuestra claramente la preferencia hacia la asistencia robótica.

En los resultados obtenidos a partir de las encuestas SURG-TLX (Figura 5) observamos que la demanda mental incrementó para el caso de la plataforma robótica en cirujanos expertos, exceptuando el grupo de cirujanos noveles que experimentaron el sentimiento contrario o simplemente no apreciaron diferencias.

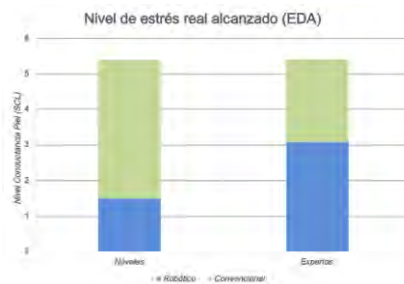


Figura 4. Comparativa del nivel de estrés medio total alcanzado por noveles y expertos.



Figura 5. Representación gráfica de la encuesta subjetiva SURG-TLX.

2.2. Procedimientos quirúrgicos

En el caso de los procedimientos quirúrgicos de cirugía general, urología y ginecología, se registró el tiempo transcurrido con la plataforma robótica y con cirugía laparoscópica convencional, junto con su valor de EDA.

Utilizando los datos comentados, se obtienen los siguientes valores medios: 28 minutos de cirugía general robótica frente a 20 min de laparoscopia; 9 min para ginecología robótica frente a 3 min en convencional; y 42 min empleados en urología robótica frente a los 26 del procedimiento convencional. En todos los sujetos se puede apreciar, como ocurrió en el caso de los sujetos expertos en el estudio con simulador, que el tiempo aumenta drásticamente al emplear la plataforma de cirugía robótica.

Si analizamos los valores de EDA, los resultados obtenidos varían significativamente entre los sujetos: Durante la cirugía general, en ambos procedimientos, se registró señales de tipo 3 siendo en robótico decreciente; en urología se obtuvo un tipo 3 en convencional frente al tipo 2 con picos de hasta $9\mu S$ en robótico; por último, en ginecología se observa un estrés de tipo 1 en convencional persistente y elevado, mientras que con el robot se obtuvo una señal de tipo 2 con una amplitud mucho menor. Todos estos resultados demuestran la diferencia de estrés que pueden sufrir los sujetos, ya no solo por la propia intervención o experiencia, sino también por la predisposición de cada sujeto y situación personal.

Atendiendo a los resultados de SURG-TLX, en estos 3 sujetos, en general, se identificó como menos compleja los procedimientos llevados a cabo mediante laparoscopia convencional, indicando una mayor demanda mental con el uso de la plataforma robótica, aunque mostrando un mayor nivel de estrés durante los procedimientos laparoscópicos convencionales.

3. Conclusiones

Teniendo en cuenta los resultados obtenidos para cada estudio se puede observar que no existe necesariamente una relación directa entre la dificultad de la tarea y el estrés al que se somete el cirujano, o la correcta ejecución de la intervención. En gran medida, el estrés subjetivo y la respuesta galvánica de la piel coinciden.

Durante las pruebas en simulador se produjo la misma discordancia, pero en sentido contrario, donde la tarea

más “estricta” fue la convencional, pero en cambio el estrés se percibió con un mayor nivel empleando la plataforma de cirugía robótica. Todo ello no implica un peor desempeño quirúrgico con asistencia robótica ya que, por ejemplo, en el simulador el cirujano que más estrés sufrió, y subjetivamente percibió, mostró mejores resultados en las habilidades quirúrgicas.

Un dato a destacar es que los sujetos que identificaron las tareas laparoscópicas más estresantes coincidieron a su vez con los procedimientos quirúrgicos con mayor estrés, los cuales se vieron en la obligación de adoptar posturas poco ergonómicas, así como esfuerzos mecánicos excesivos para el desarrollo de la actividad quirúrgica. Se entiende de esta forma a la plataforma robótica como una asistencia para fomentar la buena manera de proceder en la cirugía en beneficio del cirujano [3], resultando finalmente en una reducción del estrés subjetivo de manera general. Sin embargo, este resultado no fue similar en al caso de las tareas en simulador.

Por consiguiente, aunque las discordancias entre los sujetos apunten a que las dos modalidades son igual de válidas para el desempeño quirúrgico, el estrés sometido al cirujano durante la práctica convencional es superior a la robótica, fomentado por una mayor exigencia física. De la misma manera, haciendo uso de un simulador para evaluación de capacidades quirúrgicas se obtienen resultados con mejor técnica en la laparoscopia tradicional por cirujanos expertos, contemplándose pocas diferencias en los cirujanos noveles en proceso de formación. Por último, se observa que el aumento de errores con asistencia robótica se debe principalmente a la escasa familiarización con esta tecnología.

Agradecimientos

Este estudio ha sido cofinanciado por la Junta de Extremadura (TA18023), el Plan Complementario Biotecnología Aplicada a la Salud, cofinanciado por el Ministerio de Ciencia e Innovación con fondos de la Unión Europea NextGenerationEU, el Plan de Recuperación, Transformación y Resiliencia (PRTR-C17.I1) y el Programa Operativo FEDER Extremadura 2021-2027, y el Fondo Europeo de Desarrollo Regional (FEDER) y el Ministerio de Ciencia e Innovación de España (CPI-2019-2033-1-TRE -14).

Referencias

- [1] Wilson MR, Poolton JM, Malhotra N, Ngo K, Bright E, Masters RS. Development and validation of a surgical workload measure: the surgery task load index (SURG-TLX). *World J Surg.* 2011 Sep;35(9):1961-9. doi: 10.1007/s00268-011-1141-4. PMID: 21597890; PMCID: PMC3152702.
- [2] Schuetz M, Gockel I, Beardi J, et al. Three different types of surgeon-specific stress reactions identified by laparoscopic simulation in a virtual scenario. *Surg Endosc.* 2008;22(5):1263-1267. doi:10.1007/s00464-007-9605-1
- [3] Hurley AM, Kennedy PJ, O'Connor L et al. SOS save our surgeons: Stress levels reduced by robotic surgery. *Gynecol Surg* 12, 197–206 (2015). (<https://doi.org/10.1007/s10397-015-0891-7>)

Diseño e implementación de un modelo basado en inteligencia artificial para estimar el consumo máximo de oxígeno en adultos de mediana edad

P. Chausa¹, J. Pájaro¹, J. Solana-Sánchez^{2,3}, G. Cattaneo^{2,3}, G. España-Irla⁴, J.M. Tormos^{2,3}, P. Sánchez-González^{1,5}, D. Bartrés-Faz^{2,6,7}, A. Pascual-Leone^{2,8,9}, E.J. Gómez^{1,5}

¹ Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España, paloma.chausa/p.sanchez/enriquejavier.gomez@upm.es

² Institut Guttmann, Institut Universitari de Neurorehabilitació adscrit a la UAB, Badalona, Spain, jsolana/gcattaneo/jmtormos@guttmann.com; dbartres@ub.edu; apleone@hsl.harvard.edu

³Fundació Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Badalona, Spain

⁴Department of Physical Therapy, Movement & Rehabilitation Sciences, Northeastern University, Boston, MA, USA

⁵CIBER de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III

⁶Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain

⁷Departament de Medicina, Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona, Barcelona, Spain

⁸Hinda and Arthur Marcus Institute for Aging Research and Deanna and Sidney Wolk Center for Memory Health, Hebrew SeniorLife, Harvard Medical School, Boston, MA, USA

⁹Department of Neurology, Harvard Medical School, Boston, MA, USA

Resumen

El cálculo preciso del consumo máximo de oxígeno (VO₂max) es esencial para evaluar la aptitud cardiorrespiratoria y la salud general de un individuo. Tradicionalmente, la medición del VO₂max requiere que los pacientes se sometan a pruebas de ejercicio extenuantes en entornos clínicos, lo que limita significativamente la accesibilidad y la conveniencia. En este trabajo de investigación se presenta un modelo de estimación del valor del VO₂max basado en inteligencia artificial, que permite evaluar a las personas de forma remota mediante cuestionarios autoadministrados o midiendo variables antropométricas, evitando así la necesidad de someterse a pruebas físicas. Para realizar esta investigación, se han utilizado datos recogidos por la Barcelona Brain Health Initiative (BBHI) que alimentan varios algoritmos de inteligencia artificial con el fin de comparar y seleccionar el que ofrece mejor rendimiento. BBHI es un estudio longitudinal prospectivo en curso patrocinado por el Institut Guttmann centrado en identificar los determinantes de la salud cerebral. Este trabajo demuestra que es posible definir e implementar algoritmos basados en inteligencia artificial para la estimación del VO₂max, logrando resultados prometedores y superando el estado del arte.

1. Introducción

La capacidad cardiorrespiratoria juega un papel fundamental en el mantenimiento y la promoción de la salud y el bienestar general siendo ampliamente reconocida como una componente clave de la aptitud física y un indicador crítico de la salud cardiovascular de un individuo. Además de ser un factor fundamental en el mantenimiento de la salud del corazón y del sistema vascular, estudios recientes [1,-3] muestran su implicación en el posible desarrollo de diferentes tipos de enfermedades como enfermedades mentales o inmunológicas, entre otras.

Una de las pruebas más comunes para determinar objetivamente la capacidad cardiorrespiratoria es la

ergometría o prueba de esfuerzo, que consiste en una evaluación fisiológica que mide la cantidad máxima de volumen de oxígeno que una persona puede utilizar en un tiempo determinado durante la práctica de ejercicio aeróbico intenso. Esta medida se conoce como VO₂max y se suele expresar en relación con el peso corporal en kilogramos (ml/kg/min). Los valores de VO₂max pueden oscilar entre los 20 y los 90 ml/kg/min, donde valores mayores se asocian con mayor capacidad cardiorrespiratoria. A pesar de su utilidad, las pruebas de esfuerzo tienen algunas limitaciones, especialmente para las personas mayores. Es una prueba de ejercicio extenuante que puede provocar agotamiento, dificultad para respirar y malestar de los músculos durante y después del examen. Requiere cierto tiempo y equipo especializado, y debe ser realizada por profesionales capacitados que tengan experiencia en fisiología del ejercicio. Algunas personas pueden tener limitaciones físicas (afecciones musculoesqueléticas, trastornos respiratorios o enfermedades cardiovasculares) que les impidan realizar la prueba con su máximo esfuerzo. Además, la alta intensidad de la prueba puede provocar ansiedad o estrés en algunas personas afectando su rendimiento y capacidad para alcanzar su verdadero esfuerzo máximo. Finalmente, si bien la ergometría es generalmente segura, existe un pequeño riesgo de complicaciones asociadas con el ejercicio intenso. Es importante realizar una evaluación exhaustiva previa a la prueba para identificar cualquier riesgo potencial y garantizar una supervisión adecuada durante la prueba.

Para superar estos inconvenientes, diferentes investigadores han propuesto ecuaciones que permiten estimar el VO₂máx sin necesidad de realizar una prueba de esfuerzo. Si bien no son tan precisas como la medición directa, estas ecuaciones proporcionan una estimación

cómoda y coste-efectiva del VO₂max para estudios de población a gran escala o para personas que no pueden realizar una prueba de ejercicio máxima. A pesar de ello, aún no se ha logrado una ecuación que sirva de referencia y ofrezca resultados suficientemente buenos en todos los contextos. Por ejemplo, S.W. Lee et al [4], N. Rafique et al. [5] y I. Cieśliński et al [6] aplican estas ecuaciones a cohortes con personas de diferentes razas y características físicas, concluyendo que dichas ecuaciones no son generalizables. R. Lima et al. [7] proponen métodos alternativos para la estimación del VO₂max. La literatura científica nos muestra que existe la necesidad de seguir investigando en nuevos métodos que estimen el VO₂max de forma precisa y generalizable.

Este trabajo de investigación tiene como objetivo desarrollar un algoritmo basado en inteligencia artificial (IA) para predecir con precisión los valores máximos de consumo de oxígeno, evitando a las personas la necesidad de someterse a pruebas de esfuerzo y permitiendo realizar la medición de forma remota.

2. Materiales y métodos

Para la realización de este trabajo de investigación, se han seleccionado un total de 12 ecuaciones [8-19] que utilizan las siguientes variables: sexo, edad, altura, peso, índice de masa corporal (IMC), circunferencia de la cintura (WC), tabaquismo (CSS), frecuencia cardíaca en reposo (RHR) e información sobre actividad física (PA). La Tabla 1 describe las variables involucradas en cada una de las ecuaciones de estimación del VO₂max.

Study	Sex	Age	Height	Weight	BMI	WC	CSS	RHR	PA
Riddle	X	X	X						
Wasserman	X	X	X	X					
Jang	X	X			X		X		
Baynard	X	X			X				
Myers	X	X		X					
de Souza	X	X	X	X					
Jackson	X	X			X				X
Whaley	X	X			X		X	X	X
Matthews	X	X	X	X					X
Wier	X	X			X				X
Nes	X	X				X		X	X
Cáceres	X	X	X	X			X	X	X

Tabla 1. Variables involucradas en cada una de las ecuaciones de estimación del VO₂máx.

Los datos utilizados en esta investigación han sido proporcionados por la BBHI [20], un estudio longitudinal y prospectivo de cohortes de base poblacional en el que participan más de 6000 voluntarios entre 40 y 70 años. Esta iniciativa recoge información demográfica, socioeconómica, medidas de salud y estilos de vida, así como imágenes médicas y evaluaciones clínicas, con el objetivo de comprender y caracterizar marcadores relacionados con el mantenimiento de la salud cerebral. Parte de la cohorte (un subgrupo de 1.000 personas) ha participado en una serie de evaluaciones clínicas presenciales para determinar un fenotipado profundo de la función cerebral, incluida una evaluación de la capacidad cardiorrespiratoria mediante una prueba de esfuerzo. Para la realización de esta prueba se aplicó el protocolo de

Wasserman modificado [21] en un ciclómetro (modelo Ergoselect 4; Ergoline, Bitz, Alemania) con un sistema de análisis de gases respiratorios (Ergostik; Geratherm Respiratory). El protocolo consistió en una fase de calentamiento de 7 minutos (sin carga), una fase de carga de trabajo progresiva y una fase de recuperación de 5 minutos (carga mínima).

2.1. Flujo de trabajo

La metodología utilizada para desarrollar los modelos de IA sigue los siguientes pasos:

- **Adquisición y preprocesamiento de datos.** Se recopiló y fusionó datos de diferentes fuentes para garantizar su coherencia e integridad. Estos datos fueron preprocesados (identificación y selección de características, transformación de variables y gestión de datos faltantes) para obtener el conjunto de datos final que consta de 943 muestras.
- **Entrenamiento de los modelos.** Una vez preparados los datos, sirvieron para entrenar diferentes algoritmos de IA. Cada modelo presenta características únicas y se ajusta mejor a una naturaleza específica de los datos. Por eso, es fundamental probar variedad de modelos para encontrar cuál se ajusta mejor a los datos disponibles. En esta investigación se han utilizado varios algoritmos: Regresión Lineal, Árboles de Decisión, Random Forest, Gradient Boosting, Support Vector Machines (SVM) y dos arquitecturas diferentes de Perceptrón Multicapa (MLP) que se diferencian en el número de capas de la capa oculta (2 capas con 64 neuronas cada una versus 4 capas con 128 neuronas cada una), la función de activación (ReLU versus tanh), la función de pérdida (MAE versus MSE) y la optimización (Adam versus RMSprop).
- **Optimización.** Los modelos de IA pueden ofrecer resultados diferentes según los hiperparámetros. Los resultados se pueden optimizar ajustando sus valores y guiando al modelo para aprender mejor las características.
- **Validación.** Cuando se selecciona el mejor algoritmo, es importante aplicarlo a un conjunto de prueba (muestras que el modelo nunca ha visto) o probar con casos reales, para verificar completamente su correcto funcionamiento.

La codificación se realizó en Visual Studio Code, utilizando lenguaje Python, versión 3.11.1 a través de una computadora personal, HP Laptop modelo 15s-eq2xxx, con 12 GB de RAM y procesador AMD Ryzen 7.

2.2. Métricas de evaluación

Las métricas propuestas para evaluar el rendimiento de los modelos de regresión son:

Raíz cuadrada del error cuadrático medio. RMSE es la raíz cuadrada del promedio de las diferencias al cuadrado entre los valores previstos y los valores reales y proporciona una métrica interpretable ya que está en las mismas unidades que la variable objetivo (ml/kg/min). Representa la magnitud promedio de los residuos (errores de predicción) y penaliza los errores mayores con más fuerza. Se obtiene mediante la siguiente expresión:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}$$

donde n es el número de muestras, \hat{y}_i son los valores predichos e y_i los reales.

R-squared score. También conocido como coeficiente de determinación, o simplemente R², mide la proporción de la varianza en la variable objetivo que es predecible a partir de las características de entrada. Varía de 0 a 1, donde 1 indica un ajuste perfecto y 0 indica la incapacidad del modelo para explicar la varianza. La puntuación R² de valores negativos indica un rendimiento deficiente del modelo. Se expresa de la siguiente manera:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

donde n, \hat{y}_i y y_i representan lo mismo que en la ecuación RMSE, e \bar{y} corresponde a la media de la población observada.

3. Resultados

La Tabla 2 resume el resultado de aplicar las 12 ecuaciones de estimación del VO₂max seleccionadas a la BBHI.

Modelo	RMSE	R ²
Riddle	5,3	0,44
Wasserman	4,9	0,52
Jang	5,5	0,32
Baynard	5,1	0,51
Myers	5,2	0,51
de Souza	5,3	0,52
Jackson	6,3	0,59
Whaley	5,4	0,60
Matthews	5,3	0,61
Wier	5,2	0,63
Nes	4,5	0,65
Cáceres	5,1	0,42

Tabla 2. Resultados de las ecuaciones para la estimación del VO₂max en la cohorte BBHI

Como muestra la Tabla 2, hay una mejora significativa en las predicciones cuando se introduce la actividad física en la fórmula (a partir de la ecuación de Jackson). La Tabla 3 recoge el resultado de las métricas de los modelos implementados.

Modelo	RMSE	R ²
Linear Regression	5,3	0,49
Decision Trees	5,0	0,62
Random Forest	5,7	0,58
Gradient Boosting	4,8	0,66

SVM	5,4	0,69
MLP 1	2,1	0,71
MLP 2	2,6	0,72

Tabla 3. Resultados de modelos de aprendizaje automático en la estimación del VO₂max para la cohorte BBHI

Entre los modelos de aprendizaje automático, SVM ha ofrecido los mejores resultados con el mejor valor de R² y un excelente valor de RMSE. Los árboles de decisión, Random Forest y Gradient Boosting también realizan estimaciones aceptables, mejorando ligeramente el rendimiento de las ecuaciones. Sin embargo, la regresión lineal no logra el objetivo.

La comparativa destaca la superioridad de los modelos MLP. Además, revela que la arquitectura MLP 2 funciona ligeramente mejor que la arquitectura 1. Esto puede deberse al aumento del número de capas y neuronas ocultas. La segunda configuración de MLP ofrece una puntuación R² de 0,72, que es la más alta obtenida en este trabajo, y no parece estar sobreajustada.

Una vez seleccionado el mejor modelo de IA, se ha evaluado el impacto de introducir nuevas variables, concretamente la presión arterial (BP por su siglas en inglés – Blood Pressure) y la circunferencia del cuello (CN). En la Figura 1 se muestra una comparativa de las métricas de evaluación para todos los casos, incluyendo la combinación de ambas variables. El mejor rendimiento obtenido alcanza un valor de R² de casi 0,74 y un RMSE de 2,64.

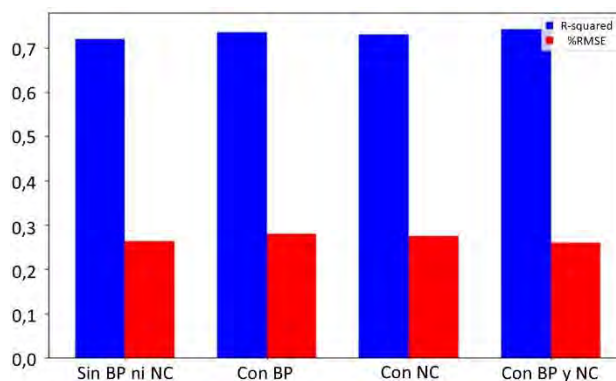


Figura 1. Comparación de métricas de rendimiento del modelo MLP en función de las variables utilizadas

Se ve que entrenar el modelo con estas dos nuevas variables agrega información a la red neuronal, lo que conduce a un mejor aprendizaje y mejores resultados.

4. Conclusiones y trabajos futuros

Esta investigación ha demostrado que es posible definir e implementar métodos basados en IA para la estimación del VO₂max, logrando resultados prometedores y superando el estado del arte.

La complejidad del cálculo del VO₂max, que involucra múltiples variables y sus variaciones individuales, se ha abordado adecuadamente a través de los modelos de IA propuestos. Los resultados de este estudio resaltan la superioridad de MLP sobre otros enfoques de aprendizaje

automático para este problema. Aprovecha su capacidad para capturar patrones complejos y relaciones no lineales dentro de las funciones de entrada, lo que conduce a un mejor rendimiento predictivo. Al explotar las representaciones jerárquicas aprendidas de los datos, las redes neuronales pueden extraer de manera efectiva características relevantes para la estimación del VO₂max.

Se ha observado que la inclusión de variables adicionales, como la presión arterial media o la circunferencia del cuello, puede proporcionar ligeras mejoras en la precisión de la estimación. Sin embargo, la practicidad de medir la presión arterial puede plantear desafíos. El acceso a dispositivos de medición adecuados y la necesidad de realizar pruebas frecuentes podrían no compensar las mejoras marginales obtenidas.

Es importante señalar que, si bien los modelos propuestos muestran resultados prometedores, existen ciertas limitaciones y áreas para seguir mejorando. Aún se necesita más investigación para desarrollar un modelo que pueda implementarse en la práctica clínica y sea generalizable. Una limitación para ello es la necesidad de una gran cantidad de datos para entrenar los modelos de forma eficaz. Recopilar y seleccionar conjuntos de datos grandes y diversos con valores de VO₂max medidos con precisión puede ser una tarea compleja. Otra vía para futuras investigaciones es la aplicación de técnicas de interpretabilidad para los modelos desarrollados. Comprender las representaciones aprendidas y las contribuciones específicas de las características involucradas en la estimación del VO₂max puede proporcionar información valiosa y mejorar la transparencia del modelo.

Agradecimientos



Esta investigación ha recibido financiación del Ministerio de Ciencia e Innovación, la Agencia Estatal de Investigación y el Fondo Europeo de Desarrollo Regional (PID2022-139298OB-C21, MCIN / AEI / 10.13039 / 501100011033 / FEDER, UE), de “La Caixa” Foundation (grant number LCF/PR/PR16/11110004), del Institut Guttmann y de la Fundació Abertis.

Referencias

- [1] B. Stubbs et al. "The association between cardiorespiratory fitness and the incidence of common mental health disorders: A systematic review and meta-analysis", *J. Affect. Disord.*, vol. 257, pp. 748–757, Oct 2019, doi: 10.1016/j.jad.2019.07.088.
- [2] M.L. Wentz and D.C. Nieman. "The compelling link between physical activity and the body's defense system", *J. Sport Health Sci.*, vol. 8, no. 3, pp. 201–217, May 2019, doi: 10.1016/j.jshs.2018.09.009.
- [3] R. Leano et al. "Cardiorespiratory fitness and cardiovascular burden in chronic kidney disease", *J. Sci. Med. Sport.*, vol. 18, no. 4, pp. 7–492, Jul 2015, doi:10.1016/j.jsams.2014.07.005.
- [4] S.W. Lee et al. "Comparison of predicted exercise capacity equations in adult Korean subjects", *J. Korean. Med. Sci.*, vol. 37, no. 14, pp. 15–117, Apr. 2022, doi: 10.3346/jkms.2022.37.e113.
- [5] N. Rafique et al. "Validity of maximal oxygen consumption prediction equations in young Saudi females", *Saudi Med. J.*, vol. 40, pp. 789–796, Aug. 2019, doi: 10.15537/smj.2019.8.24332.
- [6] I. Cieśliński et al. "External validation of VO₂max prediction models based on recreational and elite endurance athletes", *Saudi Med. J.*, vol. 25, no. 18, pp. 280–897, Jan. 2023, doi: 10.1371/journal.pone.0280897.
- [7] R. Lima et al. "A reference equation for normal standards for VO₂ Max: Analysis from the fitness registry and the importance of exercise national database (FRIEND registry)", *Prog. Cardiovasc. Dis.*, vol. 60, no. 1, pp. 21–29, Jun. 2017, doi: 10.1016/j.pcad.2017.03.002.
- [8] Riddle W, et al.. Graphical analysis of patient performance in the pulmonary function laboratory. *Proc Annu Symp Comput Appl Med Care.* 1980;1:283–290.
- [9] Wasserman K, et al. *Principles of Exercise Testing and Interpretation.* 2nd ed Malvern, PA: Lea & Febiger; 1994.
- [10] Jang T-W, et al. Estimation of maximal oxygen uptake without exercise testing in Korean healthy adult workers. *Tohoku J Exp Med.* 2012;227:313–319.
- [11] Baynard T, et al. The role of body habitus in predicting cardiorespiratory fitness: the FRIEND registry. *Int J Sports Med.* 2016;37:863–869.
- [12] Myers J, et al. A reference equation for normal standards for VO₂ max: analysis from the Fitness Registry and the Importance of Exercise National Database (FRIEND Registry). *Prog Cardiovasc Dis.* 2017;60:21–29.
- [13] de Souza ESCG, et al. A reference equation for maximal aerobic power for treadmill and cycle ergometer exercise testing: analysis from the FRIEND registry. *Eur J Prev Cardiol.* 2018;25:742–750.
- [14] Jackson AS, et al. Prediction of functional aerobic capacity without exercise testing. *Med Sci Sports Exerc.* 1990;22:863–870.
- [15] Whaley MH, et al. Failure of predicted VO₂peak to discriminate physical fitness in epidemiological studies. *Med Sci Sports Exerc.* 1995;27:85–91
- [16] Matthews CE, et al. Classification of cardiorespiratory fitness without exercise testing. *Med Sci Sports Exerc.* 1999;31:486–493.
- [17] Wier LT, et al. Nonexercise models for estimating VO₂max with waist girth, percent fat, or BMI. *Med Sci Sports Exerc.* 2006;38:555–561.
- [18] Nes BM, et al. Estimating V.O 2peak from a nonexercise prediction model: the HUNT Study, Norway. *Med Sci Sports Exerc.* 2011;43:2024–2030.
- [19] Cáceres JM, et al. Equações de predição da aptidão cardiorrespiratória de adultos sem teste de exercícios físicos. *Rev Bras Cineantropom Desempenho Hum.*2012;14
- [20] Cattaneo G., et al. (2018). The Barcelona Brain Health Initiative: A Cohort Study to Define and Promote Determinants of Brain Health. *Frontiers in Aging Neuroscience*, 10(October), 321
- [21] Wasserman, Karlman, et al. "Principles of exercise testing and interpretation." *Journal of Cardiopulmonary Rehabilitation and Prevention* 7.4 (1987): 189.

Clasificación de patologías mediante análisis acústico de la VOZ

Y. Sabuco García ¹, D. Ruiz Fernández ^{1,2}

¹ Escuela Politécnica Superior, Universidad de Alicante, Alicante, España, ysg4@alu.ua.es

² Escuela Politécnica Superior, Universidad de Alicante, Alicante, España, druiz@gcloud.ua.es

Resumen

Son múltiples los factores que intervienen en el acto de la fonación, por tanto, la voz puede evidenciar la presencia de enfermedades de diversa índole. El desarrollo de nuevos sistemas diagnósticos basados en el análisis de la voz puede ofrecer ventajas como la sencillez y la inocuidad. En este trabajo, se presenta un método de detección de sujetos patológicos a partir de la señal acústica de la voz mediante el conjunto de características ComParE de openSMILE y el modelo de ensemble Bagging.

1. Introducción

La voz es energía acústica que ha sido convertida por las cuerdas vocales a partir de energía aerodinámica que genera el aparato respiratorio [1]. Dada la gran cantidad de actores involucrados en la generación de la voz, directa o indirectamente, ésta puede ser el signo de enfermedades a diferentes niveles del organismo. Técnicas diagnósticas basadas en el análisis de voz podrían ofrecer ventajas como la inocuidad y la sencillez. Como consecuencia a esta cuestión, existen en la literatura múltiples estudios que analizan características mediante el análisis de la señal acústica a través modelos de inteligencia artificial con el fin de detectar e identificar patologías. En este sentido, se pueden encontrar referencias relacionadas con la detección de voces patológicas [2, 3] o la identificación de enfermedades como el cáncer de laringe [4], nódulos en laringe, edema de Reinke y parálisis de la laringe [2, 5], COVID-19 [6] o esclerosis lateral amiotrófica [7].

Uno de los factores limitantes en esta línea de investigación es la disponibilidad de las bases de datos y las diferencias que hay entre ellas. En ocasiones, las tareas vocales registradas o las frecuencias de muestreo son diferentes en función de la base de datos y la categorización de las patologías es muy dispar [8, 9]. Entre las bases de datos más utilizadas se encuentran la base de datos Massachusetts Eye and Ear Infirmary (MEEI) [10, 8, 2, 3, 4], la base de datos Saarbruecken Voice Database (SVD) [11, 8, 3, 12] y la base de datos Arabic Voice Pathology Database Samples (AVPD) [13, 14, 5].

En cuanto a los modelos elegidos para la detección e identificación de patologías por la voz, son numerosas las técnicas de inteligencia artificial utilizadas para la clasificación de enfermedades, se recurre tanto a modelos de aprendizaje profundo [12, 3, 4], como a otras técnicas de machine learning [7, 3, 2, 6].

Algo similar a la elección de los modelos ocurre con

respecto al análisis de la señal acústica, son diversos los algoritmos y parámetros calculados en el dominio temporal, frecuencial y cepstral que se pueden extraer y, a día de hoy, no existe consenso con respecto a qué características son las más propicias para el objetivo que se presenta. Entre los más utilizados en el dominio temporal se encuentran parámetros de frecuencia como la frecuencia fundamental y el jitter, parámetros de intensidad como el shimmer y parámetros de ruido como el Harmonic to Noise Ratio (HNR), Normalized Noise Energy (NNE) o Noise to Harmonic Ratio (NHR) [7, 4]. Mientras que en el dominio cepstral, los coeficientes basados en la percepción del habla Mel Frequency Cepstral Coefficients (MFCC) son utilizados habitualmente [8, 12, 2].

Recientemente, han aparecido diversos estudios que han hecho uso de openSMILE [3, 6, 9]. Concretamente, el conjunto de características ComParE [15] es capaz de extraer un total de 6373 de modo automático. Son pocos los estudios que hayan utilizado este conjunto para la detección e identificación de patologías a través de la voz pese a las ventajas que puede suponer hacer uso de esta herramienta, que obtiene de una manera rápida y de un modo estandarizado tantas características. El único estudio encontrado que hace uso de este sistema para detección de voces patológicas, construye un conjunto de características con las tareas vocales /a/, /i/ y /u/ en diferentes todos y utiliza el modelo de aprendizaje supervisado Support Vector Machine (SVM) [16].

En este trabajo, el objetivo será detectar voces patológicas mediante la combinación de un método de extracción de características y un método de inteligencia artificial que no se ha utilizado antes para este propósito. Se utilizará la base de datos SVD, prestando especial atención en las limitaciones que se conocen acerca de ésta. La extracción de características será mediante el conjunto de datos ComParE de openSMILE y el método de ensemble Bagging.

2. Materiales y métodos

En el siguiente apartado se procederá a explicar la metodología llevada a cabo para preparar el conjunto de datos, extraer y pre-procesar las características, entrenar y validar el modelo de clasificación. En la siguiente figura se muestra un resumen de la metodología:

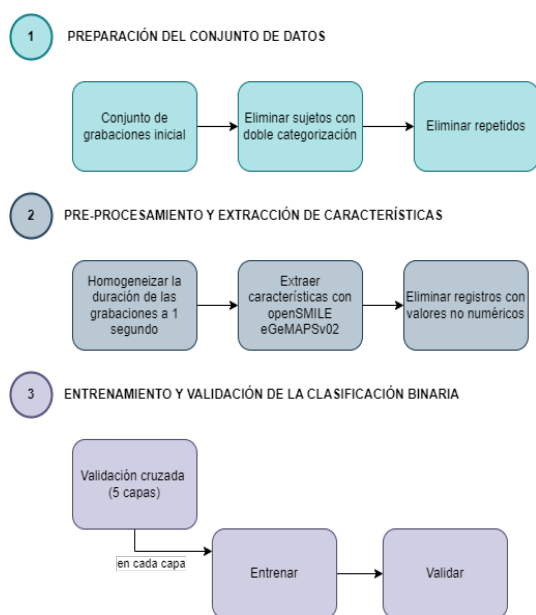


Figura 1. Metodología

2.1. Base de datos

Saarbruecken Voice Database es una base de datos abierta, distribuida por el Instituto de Fonética de la Universidad del Sarre, que posee en torno a 2000 sesiones donde se recopilaron grabaciones de las tareas vocales de los fonemas /a/, /i/ y /u/ producidas en tono normal, alto, bajo y con tono ascendente. Además, posee grabaciones de una frase en alemán y de la señal electroglotográfica (EGG). No obstante, y pese a la gran cantidad de registros que contiene, no todas las sesiones tienen todas las diferentes grabaciones recopiladas, contiene sujetos repetidos y existen individuos categorizados como sanos y patológicos al mismo tiempo[9]. Este factor no siempre es tenido en cuenta en los estudios que utilizan esta base de datos.

2.2. Preparación del conjunto de datos

Se utilizará la grabación de la tarea vocal /a/ en tono neutral de SVD cuya duración oscila de 1 a 3 segundos dependiendo del sujeto. Esta elección está basada en que es la única tarea que se recoge en las tres bases de datos más utilizadas en el estado del arte, y permitirá hacer comparativas más precisas con trabajos anteriores. Estos registros poseen una frecuencia de muestreo de 50 KHz y una resolución de 16 bits.

Se parte con el número de 2041 grabaciones de este tipo (886 hombres y 1155 mujeres). Sin embargo, como se ha dicho en los apartados anteriores, puede ocurrir que los sujetos estén repetidos, y de estos sujetos repetidos, hay algunos categorizados al mismo tiempo como sanos y enfermos. Ante esta dificultad, se decide suprimir aquellos individuos con doble categorización, así como, eliminar los individuos repetidos. Además, en ocasiones los registros no disponen de la calidad suficiente y al realizar el análisis acústico se generan valores no numéricos que invalidan el conjunto de datos, estos registros también

son suprimidos.

2.3. Pre-procesamiento y extracción de características

Para no obtener resultados dependientes de la duración se considerará sólo el primer segundo de cada grabación. Para la extracción de características se utilizará la herramienta open-source openSMILE, más concretamente, el conjunto de características denominado ComParE. Esta herramienta extrae automáticamente 6373 características. Su mecanismo se basa en extraer múltiples características de la señal acústica, lo que sus autores llaman descriptores de bajo nivel. A las cuales aplican descriptores funcionales, es decir, cálculos estadísticos como la media, la desviación estándar, cuartiles, etc.

Para la correcta validación de sistema se utilizará un método de validación cruzada de cinco capas. En este método, los datos se dividirán en cinco partes. Cada parte tendrá un conjunto de entrenamiento y uno de validación diferente que no se solapan de una iteración a otra. Habrá cinco iteraciones de entrenamiento independientes y el análisis del rendimiento se realizará con los datos de validación. La media de los datos de rendimiento de las cinco iteraciones corresponderán a los datos finales.

En la Figura 2, observamos el conjunto de datos resultante tras la preparación del conjunto de datos:

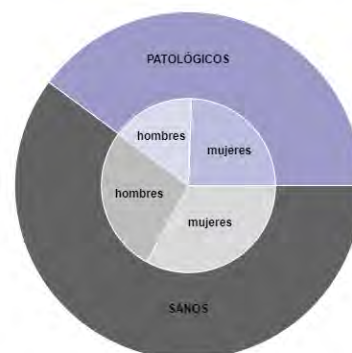


Figura 2. Distribución del conjunto de datos final

Conjunto de datos	Sanos	Patológicos	Total
Hombres	255	450	705
Mujeres	405	546	951
Total	660	996	1656

Tabla 1. Distribución del conjunto de datos final

2.4. Clasificación binaria

En un problema de clasificación, el objetivo del modelo es determinar a qué clase pertenece la entrada. Cuando se tienen dos clases objetivo se denomina, clasificación binaria. Para el problema de clasificación binaria propuesto en este estudio, discriminar entre sujetos sanos y patológicos, se utilizará un modelo de aprendizaje supervisado. Un modelo supervisado es aquel que resuelve

una tarea como la clasificación estableciendo una relación entre la entrada y la salida del modelo.

El método de aprendizaje supervisado utilizado en este estudio es el clasificador Bagging, este clasificador pertenece a los métodos ensemble o combinados. Estos métodos se basan en la combinación lineal de modelos que hacen una predicción simple, éstas se combinan para resolver un problema más complejo. En función de la combinación de los modelos simples que componen un método de ensemble podemos distinguir:

- Votación por mayoría: en éste método cada modelo simple posee un voto y la predicción se realiza por mayoría de votos.
- Bagging: cada modelo simple se entrena con un subconjunto del conjunto de entrenamiento, luego sus predicciones se agrupan por promedio o votación.
- Boosting: los modelos simples están colocados de modo secuencial de modo que un modelo intenta corregir el error de sus predecesor.

Como se ha dicho, en el caso del clasificador Bagging, éste entrena sus modelos simples con subconjuntos del conjunto de entrenamiento. Existen diferentes modos de elegir dichos subconjuntos, en este caso se hace aleatoriamente a partir del conjunto de entrenamiento extrayendo las muestras con repetición [17].

2.5. Análisis estadístico

Para evaluar la efectividad del modelo, se recurrirá a la matriz de confusión y los cálculos derivados de esta. Como se puede observar en la siguiente figura, en la diagonal principal de la matriz se encontrarán los sujetos patológicos, tanto aquellos que han sido predichos correctamente (verdaderos positivos), como aquellos que no lo han sido (verdaderos negativos).

		PREDICCIÓN	
		Verdaderos positivos (VP)	Falsos negativos (FN)
REALIDAD	Falsos positivos (FP)		
	Verdaderos negativos (VN)		

Figura 3. Matriz de confusión

Las métricas que se calcularán a partir de la matriz de confusión serán las siguientes:

- Validez (V): en inglés accuracy, representa la tasa de aciertos con respecto al total. Es una métrica que no se utiliza cuando la sensibilidad y la precisión no tienen la misma importancia en el problema planteado.

$$V = \frac{(VP + VN)}{(VP + FN + FP + VN)}$$

- Sensibilidad (S): en inglés recall, proporción de casos positivos acertados.

$$S = \frac{VP}{(VP + FN)}$$

- Especificidad (E): en inglés specificity, proporción de casos negativos acertados.

$$E = \frac{VN}{(FP + VN)}$$

- Valor predictivo positivo (VPP): en inglés precision, hace alusión a los casos positivos correctamente clasificados entre los que el modelo predice como positivos.

$$VPP = \frac{VP}{(VP + FP)}$$

- F1: es un valor que permite evaluar la precisión y la sensibilidad al mismo tiempo.

$$F1 = 2 * \frac{S * VPP}{(S + VPP)}$$

3. Resultados y discusión

En la siguiente figura, se encuentra la matriz de confusión obtenida tras el entrenamiento y validación del conjunto de datos por validación cruzada:

		PREDICCIÓN	
		845	151
REALIDAD	179		
	481		

Figura 4. Matriz de confusión resultante en el experimento

En la siguiente tabla se presentan las métricas obtenidas a partir de la matriz de confusión:

V	S	E	VPP	F1
80 %	83 %	75 %	84 %	84 %

Tabla 2. Análisis estadístico resultante

Como se puede observar la validez del clasificador Bagging que se ha utilizado en este experimento 80 %, es parecida a la obtenida en [16], donde se obtiene un 82 % de validez. Sin embargo, haría falta más información para poder realizar una comparación válida, puesto que los valores de sensibilidad, especificidad y valor predictivo positivo no están especificados en éste. Además, el conjunto de datos entre estudios no es comparable, en este estudio se está utilizando una sola tarea vocal en tono neutral, mientras que en [16] se utilizan las tareas vocales /a/, /i/ y /u/ en diferentes tonos.

4. Conclusiones

En este trabajo se ha realizado la evaluación del método de ensemble Bagging para detectar individuos patológicos frente a individuos sanos utilizando el conjunto de extracción de características automático ComParE. Los resultados de sensibilidad y especificidad son de, 83 % y 75 %, respectivamente. Se obtiene un valor de validez de 80 %, valor similar al obtenido en otros estudios. Para crear el conjunto de datos se han tenido en cuenta las limitaciones de la base de datos Saarbruecken Voice Database y se ha utilizado la tarea vocal /a/ en tono neutral para facilitar la comparativa entre experimentos.

En próximos trabajos, se deberá utilizar otras tareas vocales para ver si existe variabilidad en el rendimiento en función de la tarea y aplicar métodos de selección de características u otros mecanismos para permitir la explicabilidad de los resultados obtenidos que permitiría evaluar qué factores de la voz son los más importantes en voces patológicas, importante para la aplicación clínica final de estos modelos. Así mismo, se realizarán trabajos de identificación de patologías específicas.

Referencias

- [1] Cobeta I, Nunez F, Fernandez S. *Patología de la voz*. 2013.
- [2] Ziqi F, Wu Y, Zhou C, Zhang X, Tao Z. Class-Imbalanced Voice Pathology Detection and Classification Using Fuzzy Cluster Oversampling Method. *Applied Sciences*, 11(8), 2021.
- [3] Wu Y, Zhou CF, Ziqi F, Di W, Zhang X, Tao Z. Investigation and Evaluation of Glottal Flow Waveform for Voice Pathology Detection. *IEEE Access*, 9:30–44, 2021.
- [4] Kim H, Jeon J, Han YJ, Joo Y, Lee J, Lee S, Im S. Convolutional neural network classifies pathological voice change in laryngeal cancer with high accuracy. *Journal of Clinical Medicine*, 9(11):1–15, 2020.
- [5] Muhammad G, Alsulaiman M, Ali Z, Mesallam TA, Farahat M, Malki KH, Al-nasheri A, Bencherif MA. Voice pathology detection using interlaced derivative pattern on glottal source excitation. *Biomedical Signal Processing and Control*, 31:156–164, 2017.
- [6] Pellet P, Phillip E, Mitra S, Holland, TC. Machine Learning-based Voice Assessment for the Detection of Positive and Recovered COVID-19 Patients. *Handbook of Clinical Neurology*, 123(January):45–66, 2020.
- [7] Tena A, Claria F, Solsona F, Meister E, Povedano M. Detection of bulbar involvement in patients with amyotrophic lateral sclerosis by machine learning voice analysis: diagnostic decision support development study. *JMIR Medical Informatics*, 9(3), mar 2021.
- [8] Smekal Z, Harar P, Galaz Z, Alonso-Hernandez JB, Mekyska J, Burget R. Towards robust voice pathology detection: Investigation of supervised deep learning, gradient boosting, and anomaly detection approaches across four databases. *Neural Computing and Applications*, 32(20):15747–15757, oct 2020.
- [9] Huckvale M, Buciuileac C. Automated detection of voice disorder in the Saarbrücken voice database: Effects of pathology subset and audio materials. In *Proceedings of the Annual Conference of the International Speech Communication Association*, volume 6, pages 4850–4854. International Speech Communication Association (ISCA), aug 2021.
- [10] Massachusetts Eye and Ear Infirmary Voice Disorders Database, 1994.
- [11] Barry WJ, Putzer M. Saarbruecken Voice Database, 2008.
- [12] Zakariah M, Mohammed R, Alotaibi B, Ajmi-Guo Y, Tran-Trung Y and Elahi K, Mohammad M. An Analytical Study of Speech Pathology Detection Based on MFCC and Deep Neural Networks. *Computational and Mathematical Methods in Medicine*, 2022.
- [13] Página web de Linguistic Data Consortium - King Saud University Arabic Speech Database. Accessed on 03.07.2023.
- [14] Al-nasheri A, Muhammad G, Alsulaiman M, Ali Z, Mesallam TA, Farahat M, Malki KH, Bencherif MA. An Investigation of Multidimensional Voice Program Parameters in Three Different Databases for Voice Pathology Detection and Classification. *Journal of Voice*, 31(1), 2017.
- [15] Eyben F. *Real-time speech and music classification by large audio feature space extraction*. Number 1975. 2014.
- [16] Barche P, Gurugubelli K, Kumar-Vuppala A. Towards automatic assessment of voice disorders: A clinical approach. In *Proceedings of the Annual Conference of the International Speech Communication Association*, 2020.
- [17] Pagina web documentación Scikit-learn 1.3.0 - Ensemble Methods. Accessed on 03.07.2023.

Deformación de la presión de las vías respiratorias durante la ventilación mecánica invasiva detectada mediante técnicas de aprendizaje automático supervisado

V. Santos-Pulpón¹, C. de Haro^{1,2,3}, A. Xifra-Porxas^{1,2}, R. Fernández^{1,2,4}, G. Murias⁵, G. Muñiz-Albaiceta⁶, S. Fernández-Gonzalo¹, M. Godoy-González¹, G. Navarra-Ventura¹, G. Gomà^{1,2,3}, O. Roca^{1,3}, J. López-Aguilar^{1,2}, L. Blanch^{1,2,3}, L. Sarlabous^{1,2}

¹ Institut d'Investigació i Innovació Parc Taulí (I3PT-CERCA), Sabadell, Spain {vsantos, cdeharo, axifra, msfernandez, mrgodoy, gnavarra, ggoma, orooca, jlopeza, lblanch, lsarlabous}@tauli.cat

² CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III

³ Servei de Medicina Intensiva, Hospital Universitari Parc Taulí, Sabadell, Spain

⁴ Critical Care Department, Althaia Xarxa Assistencial Universtaria de Manresa, Spain, rfernandezf@althaia.cat

⁵ Critical Care Department, Hospital Británico, Buenos Aires, Argentina, gmurias@hbritanico.com.ar

⁶ Acute Cardiac Care Department, Hospital Universitario Central de Asturias, Oviedo, Asturias, Spain, gma@cri-lab.org

Resumen

Es esencial una interacción paciente-ventilador adecuada en aquellos pacientes que reciben Ventilación Mecánica Invasiva (VMI). Los clínicos suelen analizar las ondas generadas por el ventilador para detectar interacciones inadecuadas, pero a veces esto lleva a un diagnóstico insuficiente. La Inteligencia Artificial (IA) podría ser una herramienta poderosa en la detección de forma automática de episodios de mala interacción paciente-ventilador. El objetivo del trabajo fue desarrollar un algoritmo de IA supervisado para identificar la deformación de la presión en la vía aérea durante la VMI. Se realizó un estudio multicéntrico y retrospectivo en pacientes adultos ingresados en la Unidad de Cuidados Intensivos (UCI) sometidos a VMI. Expertos clasificaron la gravedad de la deformación de la onda de presión en vía aérea. Se entrenaron modelos de redes neuronales convolucionales y recurrentes, y se evaluaron mediante métricas de rendimiento. Se analizaron 6.428 respiraciones de 28 pacientes, y se encontró que el 42% se clasificaron como normales-leves, el 23% como moderadas y el 34% como graves en términos de deformación de la presión. Los modelos de IA mostraron una precisión del 87,7% en la red recurrente y 87% en la convolucional para la detección de la deformación de la presión. Nuestro estudio sugiere que la IA es una herramienta prometedora para identificar la deformación de la presión en vía aérea durante la VMI de forma continua y automática. Estos modelos podrían utilizarse en tiempo real para mejorar la monitorización y minimizar los periodos de interacción inadecuada entre el paciente y el ventilador.

1. Introducción

En medicina, el uso de algoritmos informáticos para la toma de decisiones está experimentando un gran avance. En entornos altamente monitorizados como las Unidades de Cuidados Intensivos (UCI), la extracción de información clínica de los datos recopilados se ha vuelto factible. El uso de estas tecnologías es relevante y prometedor, pudiendo tener un impacto positivo en la toma de decisiones médicas [1].

La Ventilación Mecánica Invasiva (VMI) es un procedimiento de soporte vital habitual en pacientes UCI no exento de efectos adversos pudiendo asociarse con un

aumento en la mortalidad. En un estudio realizado en Estados Unidos, la VMI representa 2.7 episodios por 1000 habitantes, con un coste estimado de \$27 mil millones [2]. La mortalidad global en pacientes con VMI es del 30-35%, y aumenta notablemente con la edad [3].

La monitorización de las señales de VMI se vuelve crucial en las UCIs. La interacción paciente-ventilador puede tener efecto sobre diferentes desenlaces de los pacientes en UCI con VMI. Actualmente la identificación de una correcta interacción paciente-ventilador se hace mediante una inspección visual por parte de los clínicos de las ondas del ventilador. Esto requiere de conocimientos especializados y solo se puede realizar en periodos cortos de tiempo, lo que resulta en un diagnóstico limitado. En pacientes críticos bajo VMI en modo volumen asistido-controlado con flujo constante, la deformación de la onda de la presión en la vía aérea refleja un flujo de aire insuficiente o un esfuerzo inspiratorio elevado, que puede causar lesiones pulmonares autoinfligidas por el paciente y del diafragma. Por lo tanto, los métodos automáticos que permitan identificar de manera continua estas deformaciones, pueden alertar a los clínicos y prevenir complicaciones respiratorias y el desarrollo de interacciones paciente-ventilador inadecuadas y perjudiciales en pacientes con VMI [4].

Así, el objetivo de este estudio fue desarrollar un algoritmo de *Machine Learning* (ML) supervisado para la identificación y clasificación continua de patrones de deformación de la forma de onda de presión en vía aérea en respiraciones activadas por el paciente (*trigger* paciente), en la modalidad de volumen asistido-controlado con flujo constante.

2. Material y métodos

2.1. Diseño

Estudio retrospectivo de pacientes adultos ingresados en la UCI que recibieron VMI. Se analizaron trazados de pacientes ingresados en dos UCIs de dos centros hospitalarios, que fueron sometidos a VMI en la modalidad de volumen asistido-controlado con flujo constante durante

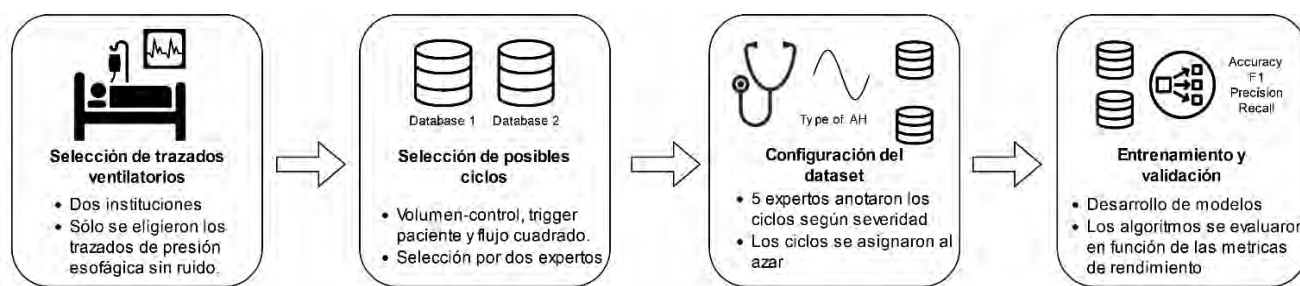


Figura 1. Pasos del estudio. Paso 1: Determinación de la población estudio y recogida de la VMI de los pacientes. Paso 2: Selección de los ciclos ventilatorios. Paso 3: Anotación de los ciclos por parte de los expertos. Paso 4: Entreno y validación de algoritmos.

más de 24 horas. En uno de los centros se obtuvo el consentimiento informado para participar en el estudio, mientras que en el otro centro el comité de revisión institucional renunció al consentimiento informado debido a la naturaleza retrospectiva del estudio. En la figura 1 se recogen las partes del proyecto.

2.2. Recogida de datos

Los datos del primer centro incluían formas de onda del ventilador (presión en la vía aérea y flujo) y trazados de presión esofágica (P_{es}) de los primeros 7 días de VMI, mientras que los datos del otro centro incluían formas de onda del ventilador (presión en la vía aérea y flujo) registradas continuamente en todo el periodo, desde el inicio de la VMI hasta la liberación. Las señales se preprocesaron con MATLAB (The MathWorks, Inc., vR2018b, Natick, MA, EE.UU.), aplicando un filtro de paso bajo *Butterworth* a 15 Hz y reduciendo la frecuencia de muestreo a 40 Hz.

Los ciclos a analizar fueron seleccionados por dos investigadores con experiencia en el procesamiento de señales de formas de onda del ventilador, que eligieron respiraciones activadas por el paciente en la modalidad de volumen asistido-controlado con flujo constante y respiraciones con y sin deformación evidente en la forma de onda de presión. El número máximo de ciclos por paciente en cada categoría se limitó a 350.

2.3. Patrones de deformación onda de presión

Cinco expertos clasificaron las deformaciones en tres patrones: normal-leve, moderada y grave. El acuerdo entre

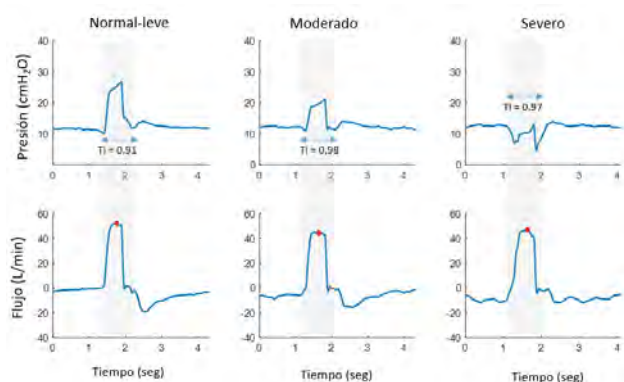


Figura 2. Ejemplo de patrones de deformación

los investigadores se determinó mediante votación por mayoría (3/5 de acuerdo [5]). En caso de desacuerdo, las respiraciones fueron reevaluadas por un experto senior para decidir su inclusión en el análisis final. La figura 2 muestra ejemplos representativo de los tres patrones.

2.4. Anotación de la severidad de deformación

Mediante la inspección visual de los trazados del ventilador (presión en la vía aérea y flujo) por parte de cinco médicos sénior intensivistas con amplia experiencia clínica en VMI interacción paciente-ventilador, tal y como sería en la práctica clínica habitual. Se clasificaron todas las respiraciones mediante una aplicación *Shiny* (ver Figura 3) desarrollada específicamente para identificar la cantidad de patrones de deformación.

2.5. Distribución del dataset

Se analizaron un total de 6.428 respiraciones. La contribución de cada paciente al conjunto de datos en cada categoría estaba condicionada a la disponibilidad de respiraciones con alto potencial para ser clasificadas por los expertos en una de las tres categorías. El número máximo de ciclos que un paciente puede aportar a cada categoría se limitó a 350 ciclos. Por lo tanto, no todos los pacientes contribuyeron por igual al conjunto de datos. De los 28 pacientes que participaron en el estudio, el 67,9%

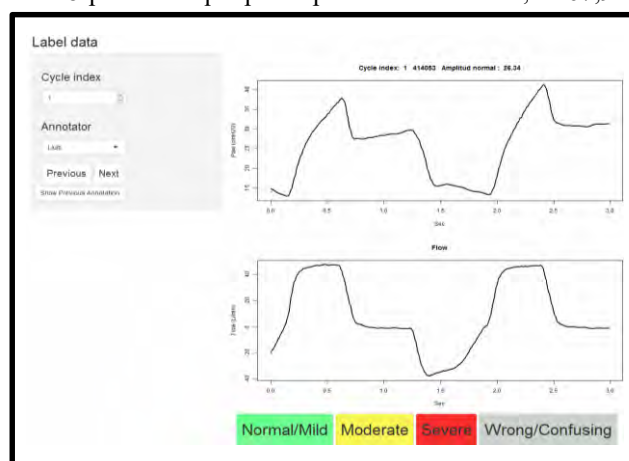


Figura 3 Vista de la aplicación *Shiny* desarrollada para la anotación de ciclos.

(19 pacientes) contribuyeron con respiraciones a las tres categorías, el 25% (7 pacientes) contribuyeron en dos categorías (4 normal-leve y moderada, 2 moderada y grave:

2 y 1 normal-leve y grave) y el 7,1% (2 pacientes) contribuyeron en una sola categoría (normal-leve o grave, respectivamente).

2.6. Algoritmos de IA

La anotación de los ciclos se utilizó para entrenar dos modelos de ML que los clasificaban en los patrones de deformación de presión: una Red Neuronal Recurrente (*Recurrent Neural Network* - RNN) y una Red Neuronal Convolutiva (*Convolutional Neural Network* - CNN). El objetivo era detectar la deformación de la presión durante la fase inspiratoria de las respiraciones. Los datos de entrada se diezmaron a 80 muestras.

Se utilizó el algoritmo RNN, apropiado para aplicaciones de secuencias largas, cuya arquitectura está diseñada para predecir una salida para cada elemento [6]. En particular utilizamos la memoria a corto plazo de larga duración (*Long Short Term Memory* - LSTM), que es el tipo más utilizado para series temporales [7] ya que aprende de las dependencias a largo plazo. En este estudio, se utilizaron dos capas ocultas de 128 neuronas y se añadió una capa totalmente conectada al final de la LSTM para clasificar en una de las tres categorías.

El algoritmo CNN utilizado realiza una convolución de una dimensión (1D-CNN) que contiene núcleos/filtros que pueden interpretarse como una aplicación de series temporales. Estos núcleos se mueven en una única dirección temporal desde el principio de la serie hacia el final, realizando la convolución. Utiliza *Batch Normalization* y *ReLU* para mejorar la convergencia y no linealidad, respectivamente. Finalmente, se aplica *Max Pooling* para reducir el tamaño del resultado obtenido [8].

Los modelos se implementaron utilizando Python (v 3.9.7) con el paquete PyTorch (v. 1.11.0) y se ejecutaron en un ordenador de sobremesa (Windows 10 Pro de 64 bits, CPU Intel(R) Core(TM) i7-6700 a 3,40 GHz y 16 GB de RAM).

2.7. Análisis estadístico

El rendimiento de los modelos RNN y CNN se evaluaron dividiendo el conjunto de datos en una división 80-20% de entrenamiento-validación, respectivamente. Este esquema se repitió 15 veces para obtener una estimación menos sesgada del rendimiento del modelo. Se calcularon las métricas de precisión, *recall*, *accuracy* y *F1-score* para evaluar su rendimiento. Además, se realizó la prueba de rangos con signo de *Wilcoxon* para investigar la relación entre los patrones de deformación de la presión y el Ti y el flujo aéreo máximo inspiratorio, corrigiendo los resultados utilizando la corrección de *Bonferroni* ($\alpha = 0,05/6 = 0,0083$). Los datos se expresaron como mediana [intervalo intercuartílico].

3. Resultados

Los expertos clasificaron un total de 6.488 respiraciones de 28 pacientes con VMI. De éstas, 302 respiraciones (4,6%) no llegaron a un acuerdo y se volvieron a analizar, de las que finalmente se excluyeron 60 respiraciones. Por lo tanto, el conjunto de datos final incluía 6.428 respiraciones clasificadas de la siguiente manera: 2.708 normales-leves (42,1%), 1.535 moderadas (23,8%) y 2.185 con

deformación grave (33,9%). La concordancia entre expertos fue del 95,4%.

El conjunto de datos de validación estaba formado por 1.287 respiraciones, 536 de ellas normales-leves (41,7%), 309 moderadas (24,0%) y 442 severas (34,4%). El algoritmo RNN clasificó con precisión el 92% de los casos normal-leve (493/536), el 80,6% de los moderados (249/309) y el 90,5% de los severos (400/442), mientras que 145 respiraciones del conjunto total de datos de validación (11,3%) se clasificaron erróneamente. El algoritmo RNN funcionó muy bien en los extremos (severo frente a normal-leve), ya que sólo etiquetó una respiración severa como normal-leve y dos respiraciones normales como severas. En conjunto, el algoritmo RNN obtuvo un *accuracy* del 87,9% [87,6 - 88,3], precisión del 87,7% [87,5 - 88,2], *recall* del 87,9% [87,6 - 88,3] y *F1-score* del 87,7% [87,4 - 88,1].

Por otra parte, el algoritmo CNN clasificó con precisión el 92% de las deformaciones de la presión normales-leves (493/536), el 74,4% de las moderadas (230/309) y el 89,6% de las severas (396/442), y 168 respiraciones del conjunto total de datos de validación (13,1%) fueron clasificadas erróneamente. Una vez más, los errores entre los extremos (severo frente a normal-suave) fueron insignificantes: 2 respiraciones normales-leves se clasificaron como severas, y 17 respiraciones severas se clasificaron como normales-leves. En general, el rendimiento de la CNN fue ligeramente inferior en comparación con el algoritmo RNN, alcanzando valores del 86,8% [86,6 - 87,4] de *accuracy*, 87% [86,7 - 87,3] de precisión, 86,8% [86,6 - 87,4] de *recall* y 86,9% [86,6 - 87,3] de *F1-score*.

4. Discusión

La principal conclusión de este estudio es que los modelos de ML pueden detectar y clasificar los patrones de deformación de presión de la vía aérea, respiración a respiración, con gran precisión.

Durante la VMI, los pacientes pueden realizar esfuerzos inspiratorios elevados en situaciones de alta demanda o flujo de aire insuficiente, lo que puede derivar en complicaciones y afectar a los desenlaces. El análisis continuo de las formas de onda de la presión mediante modelos de ML puede servir como alerta para detectar desajustes entre las demandas del paciente y la asistencia del ventilador. Esto permite un cribado no invasivo de la interacción paciente-ventilador, brindando una herramienta de diagnóstico más precisa y evitando el infradiagnóstico asociado a la inspección visual [4].

Se han desarrollado diferentes algoritmos para la detección de varios tipos de interacción paciente-ventilador. Las redes con mejores resultados han sido CNN y RNN. Los algoritmos de redes neuronales convolucionales detectaron distintos tipos de interacción paciente-ventilador con una precisión que osciló entre el 97% y el 99% [9], [10], mientras que las RNN, en particular las LSTM, obtuvieron resultados ligeramente peores, entre el 91% y el 98,3% [6]. En el presente estudio se han implementado dos redes neuronales diferentes, una LSTM y una 1D-CNN. Las CNN se consideran actualmente los modelos más

avanzados con mejores resultados en la detección de interacción paciente-ventilador, pero en nuestro estudio, el modelo RNN mostró una precisión ligeramente superior. Una explicación puede ser que las RNN son más adecuadas para manejar secuencias o datos dependientes del tiempo [6]. Las principales diferencias entre estos modelos y el nuestro son los hiperparámetros internos de las redes, que son sintonizables con los tipos de señales y la información que el algoritmo utiliza para entrenarse. El modelo de ML podría proporcionar una clasificación precisa de las respiraciones con deformación severa, basándose en el análisis de la forma de onda de presión en la vía aérea. Por lo tanto, la evaluación continua de la deformación de presión en la vía aérea mediante el uso de tecnologías de ML podría alertar a los clínicos sobre la presencia de esfuerzos inspiratorios excesivamente altos o flujo de aire insuficiente.

Por otro lado, es crucial contar con conjuntos de datos equilibrados al entrenar modelos de ML para evitar resultados inexactos o sesgados debido al desequilibrio en la distribución de etiquetas. Los modelos de ML tienden a favorecer las clases más frecuentes, lo que puede conducir a una clasificación incorrecta o insuficiente de las clases menos comunes [11]. Esto es especialmente problemático en aplicaciones críticas como la detección de enfermedades o la identificación de fraudes. Para abordar este problema, es importante equilibrar las clases en el conjunto de datos mediante la recopilación de más datos para las clases menos comunes, el submuestreo de las clases más frecuentes o la generación de datos sintéticos. Un conjunto de datos equilibrado permite a los modelos de ML aprender patrones más precisos y equilibrados, lo que mejora el rendimiento en todas las clases. Este estudio tiene varias limitaciones. En primer lugar, el modelo solo se aplicó a un modo de VMI específico, por lo que sería necesario desarrollar modelos para otros modos de VMI. En segundo lugar, los modelos RNN y CNN deben entrenarse con datos suficientes, y aunque nuestra muestra pueda parecer pequeña, ha dado buenos resultados. Por otro lado, el etiquetado a gran escala lleva mucho tiempo, y requiere de un amplio conocimiento para implementar una tarea clínica particular. Además, estos métodos predicen patologías etiquetadas, lo que restringe su aplicabilidad.

5. Conclusión

Nuestro estudio demuestra que el ML, en particular las RNN, podrían ser una herramienta excelente para identificar la deformación de la presión de las vías aéreas durante la VMI controlada por volumen y flujo de aire cuadrado, lo que permitiría tomar medidas para minimizar los periodos no reconocidos de interacción inadecuada entre el paciente y el ventilador.

Agradecimientos

Este proyecto cuenta con el apoyo del programa Pla Estratègic de Recerca i Innovació en Salut (PERIS: SLT017/20/000153) del Departamento de Salud de la Generalitat de Catalunya, España. Y del proyecto 202118 (413/C/2021) Fundació La Marató de TV3, Programa CERCA/Generalitat de Catalunya y Fundació Institut d'Investigació i Innovació Parc Taulí-I3PT y CIBER -

Consorcio Centro de Investigación Biomédica en Red-(CB06/06/1097), Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación.

Referencias

- [1] J. Gallifant *et al.*, “Artificial intelligence for mechanical ventilation: systematic review of design, reporting standards, and bias,” *Br. J. Anaesth.*, vol. 128, no. 2, pp. 343–351, 2022
- [2] H. Wunsch *et al.*, “The epidemiology of mechanical ventilation use in the United States,” *Crit. Care Med.*, vol. 38, no. 10, pp. 1947–1953, 2010
- [3] A. Esteban *et al.*, “Characteristics and outcomes in adult patients receiving mechanical ventilation: A 28-day international study,” *Jama*, vol. 287, no. 3, pp. 345–355, 2002
- [4] D. Colombo *et al.*, “Efficacy of ventilator waveforms observation in detecting patient-ventilator asynchrony,” *Crit. Care Med.*, vol. 39, no. 11, pp. 2452–2457, 2011,
- [5] V. S. Sheng *et al.*, “Majority Voting and Pairing with Multiple Noisy Labeling,” *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 7, pp. 1355–1368, 2019
- [6] M. Hüsken *et al.*, “Recurrent neural networks for time series classification,” *Neurocomputing*, vol. 50, pp. 223–235, 2003
- [7] Q. Du *et al.*, “Attention-based LSTM-CNNs For Time-series Classification,” *Proc. 16th ACM Conf. Embed. Networked Sens. Syst.*, 2018.
- [8] H. Ismail Fawaz *et al.*, “Deep learning for time series classification: a review,” *Data Min. Knowl. Discov.*, vol. 33, no. 4, pp. 917–963, 2019
- [9] Q. Pan *et al.*, “An interpretable 1D convolutional neural network for detecting patient-ventilator asynchrony in mechanical ventilation,” *Comput. Methods Programs Biomed.*, vol. 204, no. 2021, p. 106057, 2021
- [10] L. Zhang *et al.*, “Detection of patient-ventilator asynchrony from mechanical ventilation waveforms using a two-layer long short-term memory neural network,” *Comput. Biol. Med.*, vol. 120, no. March 2020, p. 103721, 2020
- [11] Q. Wei *et al.*, “The Role of Balanced Training and Testing Data Sets for Binary Classifiers in Bioinformatics,” *PLoS One*, vol. 8, no. 7, p. e67863, Jul. 2013, [Online].

Caracterización de las interacciones complejas en pacientes que pasan con éxito la prueba SBT

L. Sarlabous Uranga^{1,2}, V. Santos Pulpón¹, C. de Haro López^{1,2,3}, A. Xifra Porxas^{1,2}, C. Subirà Cuyàs^{3,4}, M. Batlle Solà^{3,4}, J. López Aguilar^{1,2}, R. Fernández Fernández^{1,2}, L. Blanch Torra^{1,2,3}

¹Institut d'Investigació i Innovació Parc Taulí (I3PT-CERCA), Sabadell, Spain

[lsarlabous,vsantos,cdeharo,axifra,jlopeza,rfernandez,lblanch}@tauli.cat](mailto:{lsarlabous,vsantos,cdeharo,axifra,jlopeza,rfernandez,lblanch}@tauli.cat)

²CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Spain

³Servei de Medicina Intensiva, Hospital Universitari Parc Taulí, Sabadell, Spain

⁴Critical Care Department, Althaia Xarxa Assistencial Universitaria de Manresa, Spain

[csubira,batmontse}@althaia.cat](mailto:{csubira,batmontse}@althaia.cat)

Resumen

La prueba de respiración espontánea (SBT) es la técnica más utilizada para identificar qué pacientes están preparados para abandonar con éxito la ventilación mecánica invasiva (VMI). Sin embargo, las tasas de fracaso, con reintubación y reconexión a la VMI son elevadas. El análisis de las interacciones complejas paciente-ventilador (CP-VI) podría ayudar a identificar aquellos pacientes que están preparados para realizar con éxito el SBT. En este trabajo se investigan dos características de las CP-VI: (1) el valor medio del incremento de irregularidad (en %) y (2) el conteo de eventos complejos que ocurren en las 6 horas previas al SBT y su asociación con el éxito o fracaso de la extubación. Las CP-VI se calculan con la entropía muestral (SE) sobre las señales respiratorias de flujo aéreo y presión en vía aérea. Estudiamos 66 pacientes que pasaron con éxito la prueba SBT y fueron extubados, de los que 10 precisaron reintubación y reconexión a la VMI en las 72 horas siguientes. El conteo promedio de eventos solo mostró diferencias en la señal de presión, mientras el valor promedio del incremento de irregularidad lo hacía en ambas señales respiratorias. El grupo no reintubados alcanzó en promedio un mayor número de CP-VI (11 vs. 7, $p < 0.05$) y de mayor magnitud (100% vs. 64%, $p < 0.05$) que el grupo reintubados, para la señal de presión de aire. Concluimos que las CP-VI podrían identificar qué pacientes podrán ser extubados con éxito. Sin embargo, son necesarios estudios con un mayor número de pacientes para confirmar los resultados obtenidos.

1. Introducción

Las Unidades de Cuidados Intensivos (UCI) tratan pacientes complejos, muchos de los cuales requieren ventilación mecánica invasiva (VMI) como soporte vital. Son esenciales las estrategias dirigidas a reducir el tiempo de VMI para disminuir las complicaciones y mejorar la calidad asistencial. La retirada controlada de la VMI (destete) es uno de los mayores desafíos en UCI [1]. El fracaso del destete en las 48-72h posteriores a la extubación programada ocurre en un 6-23% [2] y requiere reintubación de emergencia, [3].

La literatura médica recomienda el uso de varios índices fisiológicos para identificar qué pacientes están preparados para ser destetados con éxito [1][2]. La técnica más utilizada es la prueba de respiración espontánea (*Spontaneous Breathing Test*, SBT). Sin embargo, las tasas

de fracaso, con reintubación y reconexión a la VMI en las 48-72h siguientes, oscilan entre el 15% y el 20% [3] y alargan la VMI y la estancia en UCI, empeoran el pronóstico y aumentan la mortalidad en un 25-50%.

Algunos estudios han propuesto la entropía de la variabilidad respiratoria como predictor del éxito del destete de la VMI [4]. Estos estudios sugieren que el aumento de la irregularidad del patrón respiratorio puede ser un marcador de salud pulmonar [5], abriendo la posibilidad de que cierto grado de irregularidad sea normal [6]. No obstante, se basan en series temporales de las señales de presión y flujo del respirador mecánico, con un solo valor por ciclo respiratorio, no se detectan otras interacciones que ocurren en el ciclo y se pierde información crucial. Por este motivo el cálculo de la entropía debería realizarse a partir de la señal continua de presión y flujo del respirador mecánico. Nuestro grupo de investigación ha desarrollado un algoritmo automático para la detección de interacciones complejas entre el paciente y el respirador (CP-VI) calculando la entropía muestral sobre las señales de presión en la vía aérea y de flujo aéreo registradas directamente del ventilador [7].

Nuestra hipótesis es que los pacientes que pasan con éxito el SBT y se extuban, pero precisan reintubación y reconexión a la VMI en las 72 horas siguientes, presentarán menor irregularidad en las horas previas a la prueba, y esto puede ser caracterizado mediante el conteo de eventos con CP-VI, y el valor promedio del incremento de irregularidad.

2. Materiales y métodos

2.1. Población del estudio

Este estudio retrospectivo se realiza en 66 pacientes críticos adultos (>18 años), ingresados en la UCI del Hospital Althaia (Manresa) entre el 2011 y el 2016, sometidos a VMI por al menos 24 horas. Los datos clínicos y demográficos se obtuvieron de la historia clínica (Tabla 1). El Comité de Ética de Investigación con medicamentos de la Fundación Catalana de Hospitales aprobó el estudio permitiendo su realización sin la petición explícita del consentimiento informado de cada participante dado que se trata de un estudio con datos retrospectivos anonimizados.

2.2. Adquisición y análisis de datos

Las señales respiratorias de flujo aéreo y de presión en vía aérea se registraron directamente del ventilador mecánico, desde la intubación hasta la liberación de la VMI utilizando el sistema Better Care (Better Care, Barcelona, España. Patente estadounidense nº 12/538.940) con una frecuencia de muestreo de 200 Hz. Utilizamos MATLAB (The MathWorks, Inc., vR2018b, Natick, MA, USA) para el procesamiento de señales, el análisis de datos y la evaluación visual. Las señales se diezmaron a una frecuencia de muestreo de 40 Hz antes del cálculo de la entropía muestral.

Datos demográficos y clínicos	
Sexo	
Mujer, n (%)	19 (28.8 %)
Edad, años	70 (63 - 80)
APACHE II al ingreso	12 (8 - 16.25)
SOFA al ingreso	4 (3 - 6.25)
Estancia en UCI, días	8.5 (5 - 17.3)
Estancia Hospital, días	21 (13.8 - 34.3)
Días en ventilación mecánica	4.8 (2.6 - 7.6)

Tabla 1. Características de los pacientes representados como mediana (rango intercuartílico). APACHE II: Acute Physiology and Chronic Health Evaluation; SOFA: Sequential Organ Failure Assessment Score; UCI: unidad de cuidados intensivos.

2.3. Detección automática de eventos CP-VI

Definimos las CP-VI como un cambio de >50% en la frecuencia respiratoria y/o >30% de respiraciones asincrónicas de cualquier tipo (esfuerzos inefectivos, doble disparo, ciclos prolongados o disparo reverso) durante un período de 3 minutos [7]. La Figura 1 muestra un ejemplo representativo de las señales respiratorias de flujo aéreo y

de presión en vía aérea, registradas en las 6 horas previas al SBT.

La entropía muestral (*SE*) es una técnica no lineal que mide la aleatoriedad de una serie de datos [8]. *SE* se calculó en las señales de flujo (*SE-Flow* ($m = 2, r = 0,2$)) y de presión (*SE-Paw* ($m = 4, r = 0,2$)), utilizando una ventana deslizante de 30 segundos con un solapamiento del 50% [7]. En la Figura 1 se muestran series de *SE* calculadas sobre los trazados de flujo aéreo y de presión en vía aérea.

El algoritmo automático para la detección de CP-VI [7] calcula el valor máximo de *SE* alcanzado en segmentos de 15 minutos. El primer valor calculado representa el valor basal (Basal0, Figura 1). Para cada nuevo segmento analizado se calcula el porcentaje de cambio con respecto al valor basal. Si el incremento supera un cierto umbral, se considera que el segmento en cuestión contiene un evento CP-VI. Los umbrales que maximizan la detección de eventos CP-VI, se corresponden con el 25% y el 30% para el flujo aéreo y la presión en vía aérea, respectivamente. Finalmente, el valor basal se actualiza cada vez que se obtiene un valor de *SE* más bajo que el del valor basal actual (Basal1, Figura 1). En la Figura 1, se observa para *SE-Flow*, 17 eventos CP-VI con un valor promedio de incremento de irregularidad del 125.4 %.

2.4. Análisis estadístico

Para evaluar el papel de las CP-VI en la clasificación de pacientes que pasan con éxito la prueba SBT y se extuban de forma programada pero son reintubados en las 72 horas siguientes, se entrenaron 2 modelos de clasificación Bagging (BC) [9] y SVM (SVM) [10]. Las variables de entrada a los modelos se corresponden con las características derivadas en las 6 horas previas al SBT: el conteo de eventos con CP-VI, y el valor promedio del incremento de irregularidad. Estas características se compararon entre los grupos reintubados vs. no

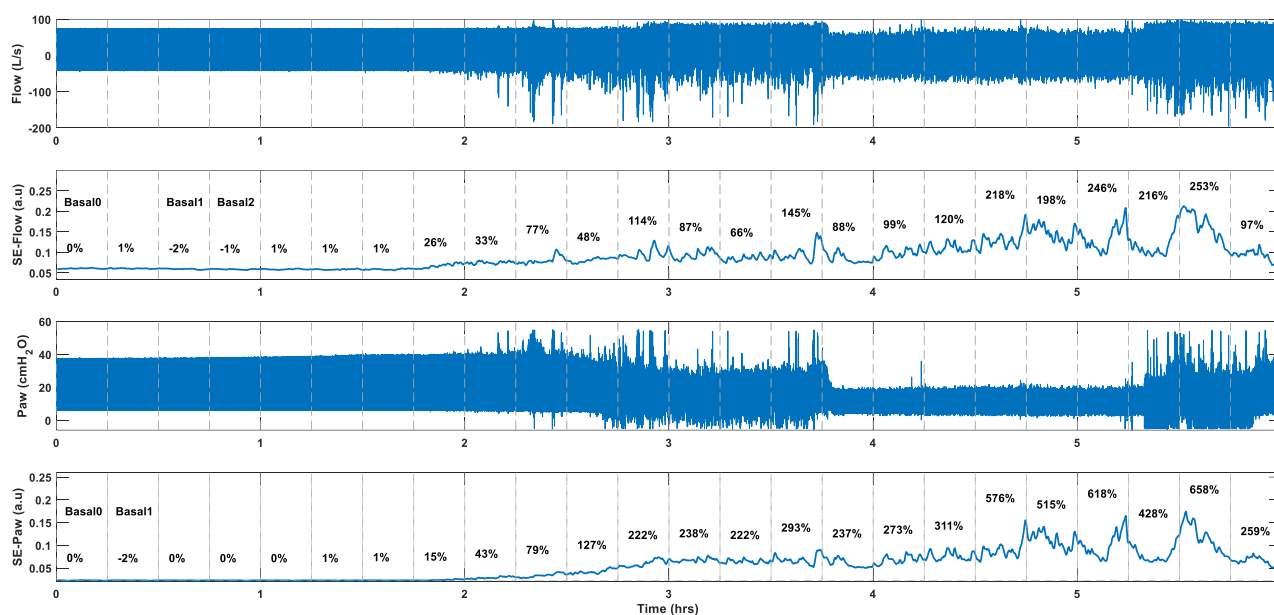


Figura 1. Representación de trazados de flujo aéreo (*Flow*), presión en vía aérea (*Paw*) y entropía muestral (*SE*) derivadas de *Flow* (*SE-Flow*) y *Paw* (*SE-Paw*). *SE-Flow* y *SE-Paw* se calcularon con $r = 0,2 \times SD$ de cada ventana deslizante superpuesta de 30 segundos, utilizando valores de m iguales a 2 y 4, respectivamente. Se representan 6 horas de ventilación mecánica previas a la prueba SBT. Los valores en % representan el incremento de irregularidad experimentado en cada segmento de 15min respecto al valor basal.

reintubados utilizando Wilcoxon rank sum test. Los valores $p < 0,05$ se consideraron significativos.

El algoritmo BC se entrenó con 100 árboles de decisión, y para mejorar el performance se implementó la técnica *k-fold cross-validation*. Mediante un análisis de sensibilidad se estimó el valor óptimo de $k\text{-fold} = 5$. Para el algoritmo SVM se investigaron diferentes *kernels* (lineal, polinómico de grado 2, 3 y 4). El *dataset* se dividió en dos subconjuntos de datos: 80% para entrenamiento y 20 % para validación. El rendimiento de los algoritmos se evaluó mediante la matriz de confusión, la sensibilidad y la especificidad. Los algoritmos se implementaron en Python3 (Anaconda3).

3. Resultados

De los 66 pacientes del estudio que pasaron con éxito la prueba SBT y se extubaron, el 85% (56) no fueron reintubados pasadas 72 horas, frente al 15% (10) que sí lo fueron. El 30% (3) de los pacientes reintubados murió en UCI, frente al 9% (5) que no fueron reintubados.

En total se analizaron 396 horas de ventilación mecánica (1584 segmentos de 15 minutos). En la Figura 2 se muestra para cada grupo de pacientes, la distribución en forma de *boxplots* de las características derivadas de las CP-VI. Por otra parte, en la tabla 2 se muestra el valor promedio de las características en cada grupo de análisis. En la tabla 3 se muestran la matriz de confusión y los valores de sensibilidad y especificidad de los diferentes modelos implementados, tanto en entrenamiento como en validación.

4. Discusión

Este trabajo representa un estudio preliminar en 66 pacientes críticos de UCI en VMI, qué tras pasar la prueba

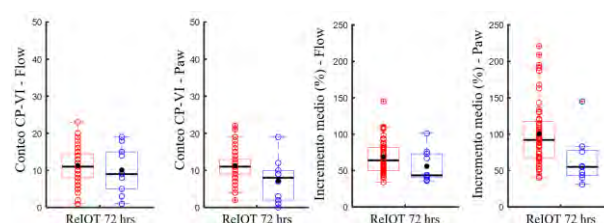


Figura 2. Conteo de CP-VI y valor promedio del incremento de irregularidad, para cada señal respiratoria, y para cada grupo: reintubados (círculos azules) y no reintubados (círculos rojos). El círculo de color negro representa el valor medio. ReIOT: reintubación.

SBT y haber sido extubados satisfactoriamente, el 15.15% de ellos requirió reintubación pasadas 72 horas, aumentando en un 21% la mortalidad. Aunque estos resultados están en línea con otros estudios [11], [12], resulta necesario investigar nuevos indicadores que ayuden a reducir el fracaso del destete. Por este motivo, nos enfocamos en caracterizar las CP-VI en las horas previas a la realización de la prueba SBT, mediante el conteo de eventos y el valor promedio del incremento de irregularidad. Además, esta caracterización podría resultar de gran utilidad al equipo clínico para identificar qué pacientes están preparados para pasar con éxito el SBT.

El conteo de eventos CP-VI y el valor promedio del incremento de irregularidad mostraron una mejor asociación sobre la señal de presión que sobre la señal de flujo (Tabla 2). Esto podría estar estrechamente relacionado con el modo ventilatorio en el que se encontraban los pacientes en las horas previas al SBT. Es conocido que, en la mayoría de los casos, los pacientes se encuentran en la modalidad presión de soporte, siendo la señal de presión la que más variación experimenta.

Características CP-VI	Flujo aéreo		<i>p</i>	Presión vía aérea		<i>p</i>
	ReIOT			ReIOT		
grupos	si	no		si	no	
Conteo eventos	10	11	0.5423	7	11	0.0183
Incremento irregularidad (%)	56	69	0.0471	64	100	0.0066

Tabla 2. Valor medio de las características: conteo de eventos y valor promedio del incremento de irregularidad desarrollados en las 6 horas previas al SBT. ReIOT: reintubación.

Algoritmo	Variables de entrada		FP	FN	TN	TP	ESP	SEN
BC	Conteo (Flow + Paw)	Ent	0	2	44	6	1	0.75
BC	Valor medio (Flow + Paw)	Ent	0	0	44	8	1	1
BC	Conteo (Flow + Paw)	Val	0	1	12	0	1	0
BC	Valor medio (Flow + Paw)	Val	1	1	11	0	0.916	0
SVM	Conteo (Flow + Paw)	Ent	0	7	45	0	1	0
SVM	Valor medio (Flow + Paw)	Ent	0	6	45	1	1	0.14
SVM	Conteo (Flow + Paw)	Val	0	2	11	0	1	0
SVM	Valor medio (Flow + Paw)	Val	0	2	11	0	1	0

Tabla 3. Rendimiento algoritmos BC y SVM en entrenamiento (Ent) y en validación (Val). Matriz de confusión: verdadero positivo (TP), verdadero negativo (TN), falso positivo (FP), falso negativo (FN), y la sensibilidad (SEN) y la especificidad (ESP). Características de entrada: conteo de eventos de CP-VI y valor promedio del incremento de irregularidad en las 6 horas previas al SBT.

Por otra parte, al analizar las características diferenciales entre los grupos reintubados y no reintubados, se encontró que el conteo promedio de eventos CP-VI sólo muestra diferencias para la señal de presión, mientras el valor promedio del incremento de irregularidad lo hacía en ambas señales. Analizando la señal de presión, los pacientes no reintubados alcanzaron en promedio un mayor número de interacciones complejas que los pacientes reintubados (11 vs.7, $p=0.0183$), siendo, también éstas, de mayor magnitud (100% vs. 64%, $p=0.0066$).

Entre las principales limitaciones del estudio se encuentran: poca separabilidad entre las clases (Figura 2), reducido conjunto de datos (66 observaciones), y un evidente desbalance en las clases (56 vs. 10) que limitan la generalización de los resultados. Todo esto conlleva a un débil comportamiento a la hora de clasificar los pacientes que precisan reintubación. En la fase de validación, por ejemplo, la sensibilidad cayó a “0” y clasificó erróneamente los pacientes en la clase reintubados. Otra limitación del estudio radica en no evaluar el efecto que tiene la modalidad ventilatoria en la que se encuentra el paciente.

En futuros estudios se debería investigar y cuantificar la cantidad de información que aportan los incrementos de frecuencia respiratoria y las asincronías en los eventos de interacciones complejas. En un análisis preliminar, no presentado en este trabajo, se ha encontrado que en algunos eventos de CP-VI la *SE* responde a cambios en la morfología de las señales de flujo aéreo y presión en la vía aérea que podrían derivarse de otros eventos no relacionados con la frecuencia respiratoria y las asincronías. Estos eventos podrían incluir las secreciones, agua en las tubuladuras, sensación de falta de aire (*air hunger*), cambios en la presión positiva al final de la espiración (PEEP), ruido de alta frecuencia y las desconexiones del ventilador.

Finalmente, este estudio abre la posibilidad de investigar una ventana de tiempo de 24–48 horas previas al SBT, con el objetivo de evaluar si las CP-VI tienen la capacidad de avanzar la realización del SBT en aquellos pacientes que presentan un mayor desarrollo de CP-VI y con mayor incremento de irregularidad. Esto podría ser relevante, ya que en muchos casos se prolonga el destete y muchos pacientes son mantenidos en VMI de forma más prolongada. Por lo tanto, el avance del SBT y destete exitoso podrían contribuir a reducir gastos en las UCIs, mejorando la gestión de los recursos y del personal asistencial.

5. Conclusiones

Las interacciones complejas entre el paciente y ventilador ocurren en todos los pacientes durante las horas previas al SBT. Los pacientes no reintubados alcanzan en promedio un mayor número de interacciones complejas y con mayor incremento de irregularidad. Sin embargo, el reducido número de observaciones unido al desbalance de las clases, limitan la generalización de los resultados, por lo tanto, resulta necesario realizar nuevos estudios en un mayor número de pacientes.

Agradecimientos

Este proyecto cuenta con el apoyo del programa Pla Estratègic de Recerca i Innovació en Salut (PERIS: SLT017/20/000153) del Departamento de Salud de la Generalitat de Catalunya, España. Y del proyecto 202118 (413/C/2021) Fundació La Marató de TV3, Programa CERCA/Generalitat de Catalunya y Fundació Institut d'Investigació i Innovació Parc Taulí-I3PT y CIBER - Consorcio Centro de Investigación Biomédica en Red- (CB06/06/1097), Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación.

Referencias

- [1] A. W. Thille, *et al.*, “The decision to extubate in the intensive care unit,” *Am. J. Respir. Crit. Care Med.*, vol. 187, no. 12, pp. 1294–1302, 2013.
- [2] F. Frutos-Vivar *et al.*, “Outcome of reintubated patients after scheduled extubation,” *J. Crit. Care*, vol. 26, no. 5, pp. 502–509, 2011.
- [3] A. W. Thille, *et al.*, “Outcomes of extubation failure in medical intensive care unit patients,” *Crit. Care Med.*, vol. 39, no. 12, pp. 2612–2618, 2011.
- [4] V. E. Papaioannou, *et al.*, “Changes of heart and respiratory rate dynamics during weaning from mechanical ventilation: A study of physiologic complexity in surgical critically ill patients,” *J. Crit. Care*, vol. 26, no. 3, pp. 262–272, 2011.
- [5] C. E. White *et al.*, “Lower interbreath interval complexity is associated with extubation failure in mechanically ventilated patients during spontaneous breathing trials” *J. Trauma - Inj. Infect. Crit. Care*, vol. 68, no. 6, pp. 1310–1316, 2010.
- [6] L. Brochard, “Breathing: Does regular mean normal?,” *Crit. Care Med.*, vol. 26, no. 11, pp. 1773–1774, 1998.
- [7] L. Sarlabous *et al.*, “Development and validation of a sample entropy-based method to identify complex patient-ventilator interactions during mechanical ventilation,” *Sci. Rep.*, vol. 10, p. 13911, 2020.
- [8] J. S. Richman *et al.*, “Physiological time-series analysis using approximate entropy and sample entropy,” *Am J Physiol Hear. Circ Physiol*, vol. 278, pp. H2039–H2049, 2000.
- [9] H. Jafarzadeh, *et al.*, “Bagging and Boosting Ensemble Classifiers for Classification of Comparative Evaluation,” *Remote sensing, MDPI*, vol. 13, p. 4405, 2021.
- [10] S. Tong, *et al.*, “Support Vector Machine Active Learning with Applications to Text Classification,” *J. Mach. Learn. Res.*, vol. 2, pp. 45–66, 2001.
- [11] A. W. Thille, *et al.*, “Weaning from the ventilator and extubation in ICU,” *Curr Opin Crit Care*, vol. 19, no. 1, pp. 57–64, 2013.
- [12] S. Jaber *et al.*, “Risk factors and outcomes for airway failure versus non-airway failure in the intensive care unit: A multicenter observational study of 1514 extubation procedures,” *Critical Care*, vol. 22, no. 1, 2018.

Aplicación de técnicas de Inteligencia Artificial Explicable para la identificación de factores relacionados con la calidad de sueño en adultos sanos

P. Chausa¹, M. Gutiérrez¹, J. Solana-Sánchez^{2,3}, G. Cattaneo^{2,3}, J.M. Tormos^{2,3}, P. Sánchez-González^{1,4}, D. Bartrés-Faz^{2,5,6}, A. Pascual-Leone^{2,7,8}, E.J. Gómez^{1,4}

¹ Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España, paloma.chausa/p.sanchez/enriquejavier.gomez@upm.es

² Institut Guttmann, Institut Universitari de Neurorehabilitació adscrit a la UAB, Badalona, Spain, jsolana/gcattaneo/jmtormos@guttmann.com; dbartres@ub.edu; apleone@hsl.harvard.edu

³ Fundació Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Badalona, Spain

⁴ CIBER de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III

⁵ Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain

⁶ Departament de Medicina, Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona, Barcelona, Spain

⁷ Hinda and Arthur Marcus Institute for Aging Research and Deanna and Sidney Wolk Center for Memory Health, Hebrew SeniorLife, Harvard Medical School, Boston, MA, USA

⁸ Department of Neurology, Harvard Medical School, Boston, MA, USA

Resumen

En los últimos años, la comunidad científica ha investigado sobre la influencia de los hábitos de vida, conductas y características de cada persona en el bienestar, mantenimiento de la salud y en la posibilidad de desarrollar enfermedades neurológicas y psiquiátricas. Los estudios realizados han demostrado que factores modificables relacionados con el estilo de vida influyen de forma determinante, siendo la calidad y los hábitos de sueño uno de esos factores. El sueño impacta en la salud, bienestar y productividad de las personas. Así, la cantidad de tiempo que dormimos es un factor predictivo de la función cognitiva influyendo incluso en la estructura del cerebro. Con el fin de extraer información sobre qué variables afectan más a la calidad de sueño, y por extensión a la salud y el bienestar de las personas, este trabajo de investigación describe la aplicación de técnicas de machine learning e Inteligencia Artificial Explicable (XAI) sobre el conjunto de datos de la Barcelona Brain Health Initiative, un estudio longitudinal y prospectivo de cohortes de base poblacional en el que participan más de 6000 voluntarios entre 40 y 70 años. Como resultado, la relevancia percibida de la rutina diaria, la percepción de llevar una vida plena y la preocupación por las opiniones ajenas, entre otros, podrían ser factores relacionados con la calidad de sueño.

1. Introducción

El sueño, y más en concreto la calidad y los hábitos de sueño son factores esenciales para el mantenimiento y mejora de la salud y el bienestar [1]. Estudios como el conducido por Tai, X.Y. et al. [2], que tratan con datos cognitivos y de sueño de 479.420 personas entre 40 y 69 años de la cohorte de la UK Biobank, nos muestran que la cantidad de tiempo que dormimos es un factor predictivo de la función cognitiva en individuos sanos de mediana y tercera edad, alcanzando el máximo rendimiento cognitivo en diferentes tests aquellos que dormían más de siete horas. Técnicas de neuroimagen empleadas en el mismo estudio mostraron que dormir entre 6 y 8 horas durante la noche se

asocia con una mayor presencia de materia gris en un patrón anatómico concreto que alcanza las regiones frontal, parietal, temporal y cerebelar. Otros estudios realizados sobre la misma cohorte han podido demostrar que dormir menos de 7 horas o más de 9 se relaciona con un mayor deterioro cognitivo [3].

A pesar de su importancia, la falta de sueño es un problema generalizado y prominente en la sociedad moderna [4]. El conocimiento de factores relacionados con la calidad de sueño puede ser de gran valor para establecer recomendaciones y definir tratamientos y medidas en aquellas personas con trastornos de sueño o dificultad para conseguir un buen descanso. En esta línea, existen estudios que aportan información sobre variables asociadas la calidad de sueño. Park K. et al [5] analizan la influencia de la actividad física y la exposición a la luz durante los dos días previos. Los factores señalados como los más influyentes en este estudio son la actividad durante el tiempo de sueño del día anterior, exposición a luz la por la tarde-noche del día anterior (de las 18:00 horas hasta que se concilia el sueño), ser mujer, la actividad física la tarde de antes (de 12 horas a 18 horas) y la exposición a luz esa misma mañana (desde despertar hasta 12:00). Alghiri Al. et al [6], analizaron factores predictores de una mala calidad de sueño en estudiantes universitarios, que, por orden de relevancia, fueron: número de horas usando dispositivos electrónicos, edad, dolores de cabeza, número de copas de té al día, índice de masa corporal (BMI), presencia de enfermedades sistémicas y dolor de cuello. También observaron que la probabilidad de tener una mala calidad de sueño era dos veces mayor en aquellos estudiantes que tuviesen o bien dolores de cabeza y/o cuello o enfermedades sistémicas. El estudio realizado por Mencar C. et al [7], se centra en factores predictivos de mal sueño para personas con apnea obstructiva, extrayendo, entre otros, ser fumador y la presencia de posibles

enfermedades sistémicas, como diabetes o, hipertensión. Concluye que el factor más influyente según todas las métricas observadas, es el índice de masa corporal (BMI). En un estudio reciente, Park S. et al [8] analizan el sueño de 30.082 personas en 11 países utilizando 52 millones de registros de actividad de dispositivos portátiles centrándose en los factores sociales y culturales que afectan el sueño de las personas. Este estudio concluye que el 55% de la variación en la calidad del sueño y el 63% en la cantidad de sueño se explican por factores sociales. A nivel individual, destacan que el sueño se vio modificado por factores como el ejercicio. Un mayor ejercicio o pasos diarios se asociaron con una mejor calidad del sueño.

Con el objetivo de avanzar en el conocimiento sobre qué factores ejercen una mayor influencia en la calidad de sueño, este trabajo propone la aplicación de técnicas de Inteligencia Artificial Explicable (XAI) a datos recogidos por la Barcelona Brain Health Initiative (BBHI) [9]. La BBHI, iniciada en 2017, está llevando a cabo un estudio longitudinal y prospectivo de cohortes de base poblacional en el que participan más de 6000 voluntarios entre 40 y 70 años, libres de diagnóstico facultativo de cualquier enfermedad neurológica o psiquiátrica en el momento de ingresar al estudio. Esta iniciativa recoge información demográfica, socioeconómica, medidas de salud y estilos de vida, así como imágenes médicas y evaluaciones clínicas, con el objetivo de comprender y caracterizar marcadores biológicos, conductuales y ambientales y sus interacciones, relacionados con el mantenimiento de la salud cerebral en la mediana edad.

2. Materiales y métodos

2.1. Fuentes de datos

Los datos analizados son las respuestas a cuestionarios online autoadministrados, los cuales están accesibles de forma segura y a demanda en un data warehouse creado por la BBHI. Para este trabajo se han utilizado todos los cuestionarios del primer año del estudio, en los que se recoge información demográfica, socioeconómica, antropométrica, presencia de diagnóstico médico y otros factores de riesgo para la salud (p. ej., tabaquismo, consumo de alcohol), así como información sobre los hábitos de vida. De forma específica, las variables utilizadas son los ítems individuales de los siguientes cuestionarios: PROMIS® (Patient-Reported Outcomes Measurement Information System; 10), PROMIS® Cognitive Abilities and Cognitive Concerns [11], DASS (Escala de Depresión, Ansiedad y Estrés; 12), MeDAS (Escala de Adherencia a la Dieta Mediterránea; 13), Cuestionario de actividad física en el tiempo libre de Godin-Shepard [14], Cuestionario de evaluación del sueño de Jenkins [15], Escala Ryff de bienestar psicológico [16], Cuestionario de reserva cognitiva [17], ELS (Escala de Vida Comprometida; 16), SOC (Escala de Sentido de Coherencia; 18), SHH-Q (Self Health Horizon Questionnaire; 19), PSQI (Índice de Calidad del Sueño de Pittsburg; 21) y LSNS (Escala de la Red Social de Lubben; [22]. El conjunto de datos inicial consta de 441 variables de 5933 adultos sanos de entre 40 y 70 años.

2.2. Metodología de análisis de datos

2.2.1. Fase de selección y preprocesamiento de datos

La selección de variables se ha realizado mediante la identificación de factores relacionados con la calidad de sueño ya reportados en la literatura científica a los que se le han añadido variables adicionales de diferentes dominios. El conjunto de datos de entrada resultante está compuesto por 20 variables de las que 13 provienen del estudio del estado del arte y 7 han sido añadidas como resultado de la exploración en diferentes dominios. La Tabla 1 enumera las 20 variables del conjunto de entrada.

Variable	Rango de Valores
Sexo	Hombre/Mujer
Hipertensión	Sí/No
Diabetes	Sí/No
Dolores de cabeza	Sí/No
Ejercicio aeróbico (días/semana)	Nunca/1-2/3-4/5-6/7
Ejercicio de baja intensidad (horas/día de ejercicio)	0/<1/1-5/6-10/>10
Gimnasio/pesas (horas/días de ejercicio)	0/<1/1-5/6-10/>10
Consumo alcohol (días/semana)	0/1-2/3-4/5-6/7
Cantidad alcohol (bebidas/día de consumo)	1/2-3/4-5/6-8/>8
Uso medicinas para dormir (último mes)	0/<1 por semana/1-2 por semana/3 o más por semana
BMI	Valor numérico
Edad	Valor numérico
Tabaquismo	Valor numérico años fumando
Relevancia de su rutina diaria (valoración subjetiva)	Mucho/Bastante/Un poco/Muy poco/Nada
Vida plena (valoración subjetiva)	Muy en desacuerdo/En desacuerdo/Ni de acuerdo ni en desacuerdo/De acuerdo/Muy de acuerdo
Edad máxima para aprender cosas nuevas (años)	40/45/50/55/60/65/70/75/ 80/85/90/95/100/105/110
Preocupación por opiniones ajenas	Muy en desacuerdo/En desacuerdo/Ni de acuerdo ni en desacuerdo/De acuerdo/Muy de acuerdo
Número de amigos cercanos	0/1/2/3-4/5-8/9 o más
Nota media académica	Número del 0 al 10

Tabla 1. Descripción de las variables de entrada

El conjunto de datos de salida se ha obtenido a partir del resultado de la escala “Jenkins Sleep Scale”. El conjunto de muestras se ha dividido en dos clases, donde la clase 0 corresponde a personas con buena calidad de sueño y está formada por las muestras cuyo resultado está por debajo de la media resultante en la población estudiada y la clase 1 a personas con mala calidad de sueño, compuesta por muestras cuyo resultado es mayor que la media.

Una vez establecidos los conjuntos de entrada y salida se realizó un preprocesamiento para transformar y normalizar las variables así como para gestionar los valores nulos. El conjunto final resultante está formado por 3216 instancias, de las que 1979 (61,54%) pertenecen a la clase 0 (buena calidad de sueño) y 1237 (38,46%) pertenecen a la clase 1 (mala calidad de sueño). Este conjunto se ha dividido en dos subconjuntos para entrenar y validar los modelos (80% y 20% respectivamente).

2.2.2. Fase de análisis y evaluación de los modelos

En este trabajo, guiados por la revisión bibliográfica, se han aplicado los siguientes algoritmos: Extreme Gradient Booster (XGBoost), Support Vector Machine (SVM) y Random Forest (RF). El desarrollo y optimización de los modelos se ha realizado haciendo uso de la librería Scikit-learn en Python, mediante la plataforma Anaconda.

La evaluación de los modelos se ha realizado en términos de precisión, sensibilidad, exactitud y F-score.

2.2.2. Explicabilidad

Para analizar la influencia de las distintas variables en la toma de decisión de los modelos, se ha utilizado Shapley Additive exPlanations (SHAP). Esta técnica ha sido elegida por su capacidad para trabajar con un gran número de modelos distintos al ser un modelo de XAI de tipo agnóstico, por su capacidad de proveer interpretaciones globales coherentes con sus explicaciones locales, y por su sencillez a la hora de ser implementado y de extraer explicaciones y representaciones visuales de las mismas.

3. Resultados

La Tabla 2 muestra los resultados de las métricas obtenidos al aplicar cada uno de los tres algoritmos propuestos.

Valores %	XGBoost	SVM	RF
Exactitud	69,41	67,70	68,48
Sensibilidad	34,45	24,79	33,61
Especificidad	89,90	92,86	88,92
F1 Score	45,43	36,20	44,08
AUC	62,18	58,82	61,26

Tabla 2. Resultados de las métricas de evaluación

Las Figuras 1 y 2 muestran los resultados de la aplicación de SHAP a los modelos XGBoost y RF. Se aplica a ambos modelos porque, además de tener métricas similares, nos permite comparar los resultados y analizar posibles diferencias. En el caso del modelo RF se han unificado las últimas 11 variables ya que no había ninguna discrepancia

con respecto a los resultados de XGBoost y por tanto no aportan información nueva para el análisis de resultados.

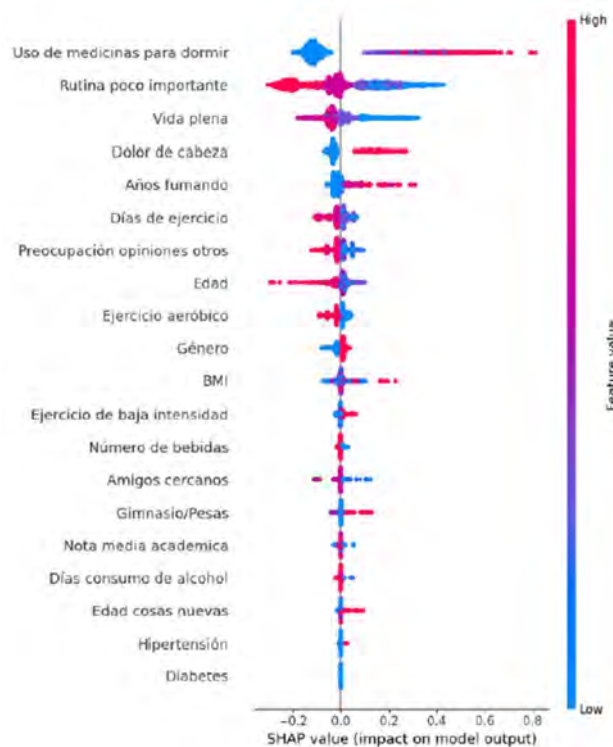


Figura 1. Análisis de influencia de variables para el modelo XGBoost

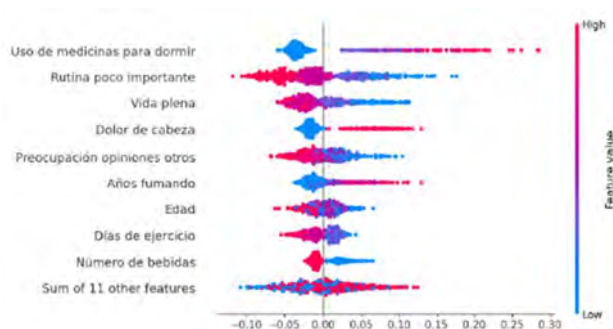


Figura 2. Análisis de influencia de variables para el modelo RF

4. Discusión y conclusiones

Como se puede observar en la Tabla 2, el modelo XGBoost ofrece, para nuestro caso de estudio, las mejores prestaciones de los 3 modelos en todas las categorías, excepto en especificidad si la comparamos con el modelo SVM. La exactitud de los tres modelos ronda el 70%, aunque no llega a alcanzarlo en ninguno de los casos. Estos niveles de exactitud coinciden con los de la bibliografía estudiada. Observando los valores de especificidad, sensibilidad y F1 score, se puede concluir que los modelos implementados tienen cierto sobreajuste respecto a la clase 0. En este sentido, se propone aplicar métodos de balanceo de datos en futuros análisis.

Las Figuras 1 y 2 muestran cómo los dos modelos coinciden en las primeras cuatro variables más influyentes en la toma de decisiones, siendo estas el uso de medicinas para dormir, la relevancia percibida por el participante de

su rutina diaria, si el participante considera que lleva una vida plena y por último la presencia o no de dolores de cabeza. El uso de medicinas para dormir y la presencia de dolores de cabeza son variables ya reportadas en la literatura científica como relacionadas con la calidad de sueño. Del análisis de datos realizado, se extrae la hipótesis de que la relevancia que la persona imprime a su rutina diaria y la sensación de tener una vida plena pueden estar asociadas a la calidad de sueño. Así, aquellos que creen tener una rutina diaria relevante y que afirman llevar una vida plena gozan de un mejor descanso y calidad de sueño.

Otras variables relevantes en las que ambos modelos coinciden, aunque variando el orden de influencia, son la edad, los años que lleva fumando el participante y la importancia que le da a las opiniones de los demás. De nuevo, la edad y el tabaquismo son factores ya conocidos mientras que la preocupación por las opiniones de otros surge como nuevo parámetro a tener en cuenta.

Cabe destacar que en este trabajo el BMI y la diabetes no aparecen como factores muy relevantes, en contra de los estudios analizados en la bibliografía. En el caso del BMI, sólo valores muy altos de este índice parecen colaborar a una mala calidad de sueño. Esta aparente contradicción con la literatura se puede explicar por el hecho de que algunos estudios se han realizado en personas con apnea obstructiva y por tanto los resultados podrían no ser generalizables a nuestra población de estudio. En el caso de la diabetes, la baja prevalencia de esta condición en el conjunto de datos, podría explicar esta diferencia.

Agradecimientos



Esta investigación ha recibido financiación del Ministerio de Ciencia e Innovación, la Agencia Estatal de Investigación y el Fondo Europeo de Desarrollo Regional (PID2022-139298OB-C21, MCIN / AEI / 10.13039 / 501100011033 / FEDER, UE), de “La Caixa” Foundation (grant number LCF/PR/PR16/11110004), del Institut Guttmann y de la Fundació Abertis.

Referencias

- [1] Grandner MA. Sleep, Health, and Society. *Sleep Med Clin.* 2017 Mar;12(1):1-22. doi: 10.1016/j.jsmc.2016.10.012.
- [2] Tai, X.Y. et al. (2022). Impact of sleep duration on executive function and brain structure. *Communications biology*, 5(1), 201. doi: 10.1038/s42003-022-03123-3
- [3] Kyle, S. D., et al. (2017). Sleep and cognitive performance: cross-sectional associations in the UK Biobank. *Sleep medicine*, 38, 85–91. doi: 10.1016/j.sleep.2017.07.001
- [4] Chattu VK, et al. The Global Problem of Insufficient Sleep and Its Serious Public Health Implications. *Healthcare (Basel)*. 2018 Dec 20;7(1):1. doi: 10.3390/healthcare7010001.
- [5] Park, K. M., et al. (2022). Prediction of good sleep with physical activity and light exposure: a preliminary study. *Journal of clinical sleep medicine*. JCSM : official publication of the American Academy of Sleep Medicine, 18(5), 1375–1383. doi: 10.5664/jcsm.9872
- [6] Alghwiri, A. A., et al. (2021). Predictors of sleep quality among university students: the use of advanced machine learning techniques. *Sleep & breathing = Schlaf & Atmung*, 25(2), 1119–1126. doi: 10.1007/s11325-020-02150-w
- [7] Mencar, C., et al. (2020). Application of machine learning to predict obstructive sleep apnea syndrome severity. *Health informatics*
- [8] Park, S., et al. Social dimensions impact individual sleep quantity and quality. *Sci Rep* 13, 9681 (2023). doi: 10.1038/s41598-023-36762-5
- [9] Cattaneo G., et al. (2018). The Barcelona Brain Health Initiative: A Cohort Study to Define and Promote Determinants of Brain Health. *Frontiers in Aging Neuroscience*, 10(October), 321
- [10] Ader, D. (2007). Developing the patient-reported outcomes measurement information system (PROMIS). *Med. Care*
- [11] Fieo, et al. (2016). Measurement Equivalence of the Patient Reported Outcomes Measurement Information System (PROMIS) Applied Cognition – General Concerns, Short Forms in Ethnically Diverse Groups. *Psychological test and assessment modeling*. 58. 255-307. 45(5 Suppl. 1), S3–S11. doi: 10.1097/01.mlr.0000260537.45076.74
- [12] Brown, T. A., et al. (1997). Psychometric properties of the depression anxiety stress scale (DASS) in clinical samples. *Behav. Res. Ther.* 35, 79–89. doi: 10.1016/S0005-7967(96)00068-X
- [13] Schröder, H., et al. (2010). A short screener is valid for assessing Mediterranean diet adherence among older. *J. Nutr.* 141, 1140–1145. doi: 10.3945/jn.110.135566
- [14] Godin, G., and Shephard, R. J. (1997). Godin leisure-time exercise questionnaire. *Med. Sci. Sports Exerc.* 29, 14–15
- [15] Jenkins, C. D., et al. (1988). A scale for the estimation of sleep problems in clinical research. *J. Clin. Epidemiol.* 41, 313–321. doi: 10.1016/0895-4356(88)90138-2
- [16] Díaz, D., et al. (2006). Adaptación española de las escalas de bienestar psicológico de Ryff. *Psicothema*, 18-3, 572-577
- [17] Rami, L., et al. (2011). Cognitive reserve questionnaire. Scores obtained in a healthy elderly population and in one with Alzheimer’s disease. *Rev. Neurol.* 52, 195–201
- [18] Trompetter, H. R., et al. (2013). Measuring values and committed action with the engaged living scale (ELS): psychometric evaluation in a nonclinical sample and a chronic pain sample. *Psychol. Assess.* 25, 1235–1246. doi: 10.1037/a0033813
- [19] Eriksson M, Lindström B. Antonovsky’s sense of coherence scale and the relation with health: a systematic review. *J Epidemiol Community Health.* 2006 May;60(5):376-81. doi: 10.1136/jech.2005.041616.
- [20] Düzel, S., et al. (2015). The subjective health horizon questionnaire (SHH-Q): assessing future time perspectives for facets of an active lifestyle. *Gerontology* 62, 345–353. doi: 10.1159/000441493
- [21] Buysse, D. J., et al. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res.* 28, 193–213. doi: 10.1016/0165-1781(89)90047-4
- [22] Lubben, J. E. (1998). Assessing social networks among elderly populations. *Fam. Commun. Health* 11, 42–52. doi: 10.1097/00003727-198811000-00008

Calibración automática para el pegado de camas en TAC

D.Sanderson ^{1,3,4}, G.Castro ^{1,3}, J.Garcia-Blas ², M.Desco ^{1,3,4,5}, M.Abella ^{1,3,4,5},

¹ Departamento de Bioingeniería, Universidad Carlos III de Madrid. Madrid, España. dsanders@ing.uc3m.es

² Computer Science and Engineering Dept., Universidad Carlos III de Madrid, Madrid, Spain ³ Instituto de Investigación Sanitaria Gregorio Marañón. Madrid, España

⁴ CIBER de Salud Mental, Instituto de Salud Carlos III. Madrid, España

⁵ Centro de Investigaciones Cardiovasculares (CNIC); Melchor Fernández Almagro, 28029, Madrid, España

Resumen

En sistemas TAC de haz Cónico (CBCT) cada adquisición se realiza con la cama inmóvil por lo que el campo de visión está limitado por el tamaño del detector. Para superar esta limitación, es común realizar adquisiciones sucesivas para distintas posiciones de la cama que posteriormente se combinan para aumentar el campo de visión en la dirección longitudinal. Para evitar la aparición de dobles bordes en el volumen resultante, es necesario conocer el desplazamiento exacto de la cama, el cual tradicionalmente se ha obtenido mediante calibraciones geométricas previas con maniqués de calibración. Esto implica que la calibración debe repetirse de forma periódica para adaptarse a los cambios que pueda sufrir el equipo. Como alternativa, en este trabajo proponemos el uso de un algoritmo de calibración automático capaz de obtener los parámetros de desalineamiento en tiempo real, permitiendo la obtención de imágenes multicama en sistemas CBCT de pequeño animal sin necesidad de una calibración previa.

1. Introducción

En sistemas TAC de haz Cónico (CBCT) el paciente se posiciona sobre una cama que permanece inmóvil durante la adquisición, por lo que el campo de visión está limitado por el tamaño del detector. Para obtener imágenes que abarquen regiones de interés (ROI) de mayor tamaño, es necesario realizar diversas adquisiciones desplazando la cama y combinarlas posteriormente en lo que se conoce como adquisición multicama. Debido a la geometría cónica del haz de Rayos X, los volúmenes reconstruidos poseen un sombreado característico en los extremos de la dirección longitudinal debido a la falta de datos, conocido como artefacto de cono. Para que no aparezca en el volumen resultante de la adquisición multicama, las adquisiciones deben realizarse con un solape espacial entre ellas.

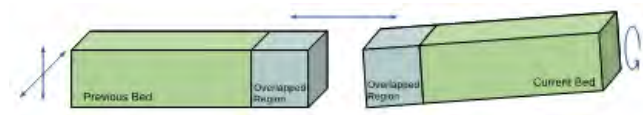


Figura 1: Representación de dos camas con solape

Debido a las tolerancias mecánicas del equipo, el desplazamiento de la cama no es ideal (Figura 1), por lo que resulta necesario calibrar este movimiento para poder combinar los volúmenes sin que aparezcan dobles bordes en la zona de solape.

Tradicionalmente, la calibración se ha realizado de forma previa a la adquisición con maniqués de dimensiones conocidas [1]. Esto podría evitarse mediante el uso

de un algoritmo de calibración automático basado en la combinación de las reconstrucciones individuales de la adquisición de cada cama para hallar los desalineamientos mediante métodos de registro. Dicho algoritmo debería cumplir los siguientes requisitos: en primer lugar debe funcionar en tiempo real para permitir su implementación en la práctica clínica y en segundo lugar debe presentar un error máximo inferior a la tolerancia máxima del equipo utilizado.

Los métodos de registro se pueden clasificar en tres grupos principales según estén basados en: extracción de características, funciones de coste o correlación de fase [2]. Los métodos basados en la extracción de características constituyen la mayoría de métodos actualmente utilizados para el registro de imágenes [2]; sin embargo, dependen de algoritmos iterativos de eliminación de valores atípicos que ralentizan los tiempos de ejecución [3] y dependen de parámetros que deben ser ajustados de forma heurística. Finalmente, los métodos basados en funciones de coste como ICP [4] reducen de forma iterativa una función de coste que determina la similitud entre las imágenes, por lo que también pueden sufrir de largos tiempos de ejecución. Los métodos de correlación de fase permiten obtener de forma analítica los parámetros de transformación, por lo que son robustos y computacionalmente rápidos [5], y no requieren calibración heurística de parámetros.

No hay muchos trabajos en la literatura que propongan la aplicación de los métodos de registro al problema del pegado de camas debido al limitado uso que los equipos CBCT tienen actualmente en la clínica. Sin embargo, la creciente demanda de estos equipos hace esperar que este problema adquiera gran interés. En [6] se evalúa un algoritmo de pegado de KODAK para el registro de imágenes de CBCT dental, y se muestra que presenta limitaciones en el cálculo de la rotación. En [7] los autores utilizan un registro iterativo del punto más cercano (ICP) no fiduciario para registrar tres secciones de un cráneo, y aunque no se discuten los tiempos de ejecución, es de esperar que al utilizar un algoritmo iterativo el registro no sea especialmente rápido. En [8] se utiliza el software de Planmed Oy para el registro semiautomático de imágenes clínicas del talón, que requiere de intervención manual por parte del usuario y por tanto disminuye su comodidad de uso.

En este trabajo proponemos el uso de un algoritmo de calibración automática basado en correlación de fase para el pegado de camas en equipos CBCT, con el objetivo de obtener un registro robusto en tiempo real.

2. Materiales y Métodos

El método propuesto se basa en el algoritmo de Fourier-Miller [9], que está basado en la correlación de fase y permite identificar traslaciones, rotaciones, y escalados en imágenes 2D. Dicho algoritmo se ha adaptado al registro de volúmenes 3D, incluyendo traslaciones en los 3 ejes y rotaciones alrededor del eje z , excluyéndose el cálculo del escalado.

Mientras que es posible utilizar las reconstrucciones de ambas camas al completo, el algoritmo solo requiere de una ROI dentro de la zona de solape entre camas y que incluya suficientes detalles para poder encontrar todos los parámetros de calibración. En el caso de un escáner de pequeño animal, en el eje longitudinal la presencia de la columna vertebral proporciona una alta densidad de detalles por lo que no se requiere una ROI de gran tamaño en esa dirección. En el corte axial, que es donde se evalúa la rotación, es necesaria una ROI de mayor tamaño tanto debido a que la rotación tiene mayor impacto sobre la periferia de la imagen, como a que existe una amplia presencia de tejido blando con escasez de detalles.

2.1. Algoritmo de calibración

A continuación se describe el algoritmo adaptado. Definimos la ROI de una cama como $f_1(x', y', z')$, la cual está trasladada y rotada respecto a la ROI de la otra cama $f_2(x, y, z)$, de tal forma que:

$$f_1(x') = f_2(x \cos \theta + y \sin \theta - x_0, -x \sin \theta + y \cos \theta - y_0, z - z_0) \quad (1)$$

En el primer paso, separamos la rotación de la traslación mediante la propiedad de la equivalencia entre traslación espacial y diferencia de fase en el dominio de Fourier:

$$|F_1(\omega)| = |e^{-i2\pi\omega\mathbf{A}} F'(\omega)| = |F'(\omega)| = M_1(\omega) \quad (2)$$

donde $F(u, v, w) = F(\omega)$ es la Transformada de Fourier (FT) de una función $f(x, y, z)$, $\mathbf{A} = (x_0, y_0, z_0)$ y $F'(\omega) = F(f_2(x \cos \theta + y \sin \theta - x_0, -x \sin \theta + y \cos \theta - y_0, z - z_0))$. $M_1(\omega)$ que representa la rotación respecto a $|F_2(\omega)|$, será considerada a partir de este punto como una imagen real. Aparte de utilizar la FT 3D, dado que la corrección de la rotación solo requiere la identificación de una coordenada en el espacio de Fourier, existe la posibilidad de definir $F(\omega)$ como la media aritmética de las FT 2D de los planos axiales del volumen.

Mediante un cambio a coordenadas polares se transforma la rotación en una traslación.

$$M_1(u \cos \theta + v \sin \theta, -u \sin \theta + v \cos \theta, w) \longrightarrow M_1(\delta, \rho, \theta + \theta_0) \quad (3)$$

Utilizando la Correlación en Fase (CF), es posible cal-

cular la traslación θ_0 de forma analítica:

$$CF(F_1, F_2) = \frac{F_1^*(\omega) \times F_2(\omega)}{|F_1^*(\omega)| \times |F_2(\omega)|} \quad (4)$$

$$= \frac{e^{i\omega\theta_0} F_2^*(\omega) \times F_2(\omega)}{|e^{i\omega\theta_0} F_2^*(\omega)| \times |F_2(\omega)|} \quad (5)$$

$$= e^{i\omega\theta_0} \frac{F_2^*(\omega) \times F_2(\omega)}{|F_2^*(\omega)| \times |F_2(\omega)|} = e^{i\omega\theta_0} \quad (6)$$

dado que la fase de un número complejo multiplicado por su conjugada traspuesta siempre es 0. Por tanto:

$$F^{-1}(CF) = \delta(\theta_0) \quad (7)$$

donde $\delta(\theta_0)$ idealmente representa una delta de Dirac en las coordenadas correspondientes a θ_0 , con el cual se encuentra el error angular en el espacio Cartesiano. Para obtener una resolución inferior a la tolerancia máxima del equipo considerado en este trabajo, se duplicó el tamaño del volumen con un relleno de ceros para reducir el tamaño de pixel de la FT, dando lugar a una resolución de $0,25^\circ$.

En el segundo paso se calibra la traslación. Tras haber corregido la rotación, $f_1(x', y', z')$ y $f_2(x, y, z)$ se relacionan de la siguiente forma:

$$f_1(x', y', z') = f_2(x - x_0, y - y_0, z - z_0) = f_2(\mathbf{x} - \mathbf{A}) \quad (8)$$

Por lo tanto, utilizando la ecuación 4 aplicada a \mathbf{A} , se obtienen los valores de traslación para los tres ejes del espacio. Dado que la tolerancia del equipo es mayor que la resolución del algoritmo sin aplicar un relleno de ceros, la máxima resolución espacial en traslación se mantuvo en $\pm 0,5$ vóxeles. Para evitar posibles errores debidos al ruido de las imágenes, se exploró la aplicación de un preprocesado de imagen mediante un filtro de mediana 3×3 como en [9].

2.3. Evaluación

La evaluación en el sistema de pequeño animal PET/CT Argus (SEDECAL, España) [10]. Para este sistema, los valores de desalineamiento máximos que resultan imperceptibles a la vista son $0,125$ mm y $0,25^\circ$ en traslación y rotación respectivamente, los cuales se tomaron como la tolerancia máxima del algoritmo. Para la simulación se utilizó el atlas de ratón completo DIGIMOUSE [11], con un tamaño de matriz de $1269 \times 695 \times 3175$ y un tamaño de vóxel de $0,06\text{mm}^3$. Se simuló proyecciones de dos camas mediante el programa de simulación FUXIM [12], utilizando la geometría del sistema PET/CT Argus y un tamaño de píxel de $0,2$ mm. Se simuló un desplazamiento de camas de 18 mm, para valores de traslación y rotación pequeños (0 mm, $0,125$ mm y $0,5$ mm en x e y , 0 mm y $0,5$ mm en z , y 0° y $0,5^\circ$) y grandes ($1,0$ mm en x e y , $2,0$ mm en z , y $2,0^\circ$). La reconstrucción se realizó con el programa RAPTOR [13], obteniéndose dos volúmenes con un tamaño de matriz de $542 \times 542 \times 575$ y un tamaño de vóxel de $0,125\text{mm}^3$.

Para la evaluación en datos reales se utilizó una adquisición de dos camas de ratón y otra de rata, obtenidas a

40 kV y 400 mAs con un tamaño de píxel en el detector de 0.2 mm, y un tamaño de vóxel de reconstrucción de $0,122\text{mm}^3$. La calibración de la adquisición de ratón era $x=-0.122$ mm, $y=0$ mm, $z=0.244$ mm, $\theta=1^\circ$, y la de rata $x=0.244$ mm, $y=0.244$ mm, $z=0.732$ mm y $\theta=-1^\circ$.

Para determinar el desempeño del algoritmo y su coste computacional se han evaluado las siguientes opciones: con y sin filtro, FT 2D y 3D para el cálculo de la rotación, y las 4 ROIs con 2 tamaños en x e y (38.3 mm \times 38.3 mm y 28.3 mm \times 28.3 mm) y 3 tamaños en z (16.1 mm, 8.05 mm y 4.02 mm) que se muestran en la Figura 2. Los tiempos de ejecución se determinaron para una CPU AMD Ryzen 7 5800X y una RAM de 32 GB.

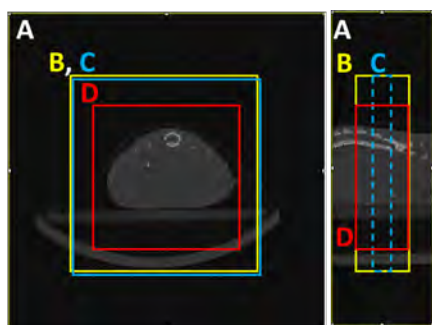


Figura 2: ROIs mostradas sobre cortes sagital y axial del volumen de ratón.

4. Resultados

En la evaluación en datos simulados se obtuvieron los valores de calibración correctos para el 83% de las 36 combinaciones de valores de desalineamiento explorados (Tabla 1). En el 17% de los casos restantes el algoritmo solo falló en la traslación en el eje y , pero con errores iguales o inferiores a la máxima tolerancia del equipo (0.125 mm y $0,25^\circ$), demostrándose así la robustez del algoritmo.

Valores reales				Valores obtenidos			
x	y	z	θ	x	y	z	θ
[mm]	[mm]	[mm]	($^\circ$)	[mm]	[mm]	[mm]	($^\circ$)
0.125	0.125	0	0.5	0.125	0.25	0	0.5
1.0	1.0	0	2	1.0	0.875	0	2
0.125	0.125	0.5	0.5	0.125	0.25	0.5	0.5
1.0	1.0	0.5	2	1.0	0.875	0.5	2
0.125	0.125	2	0.5	0.125	0.25	2	0.5
1.0	1.0	2	2	1.0	0.875	2	2

Tabla 1. Valores de traslación y rotación obtenidos para datos simulados. En negrita se indican los errores, siendo en todos los casos inferiores a la tolerancia máxima.

La evaluación en datos reales muestra una clara ventaja en el uso de la FT 3D que permitió obtener los valores de rotación en el 75% de los casos explorados (Tabla 2). Se vio que la reducción de ruido de la FT con el filtro de mediana permite una mejor aproximación de $\delta(\mathbf{A})$, como puede apreciarse en la altura, anchura y posición

del pico del panel superior de la Figura 3, obteniéndose valores correctos de traslación en el 75% de los casos (Tabla 3).

ROI	$\delta\theta[^\circ]$	$\delta\theta[^\circ]$	Tiempo	Tiempo
	3D	2D	3D [s]	2D [s]
A	1	1	14.86	11.36
B	1	0	1.74	1.96
C	1	0	0.98	1.06
D	0	0	0.9	1.45

Tabla 2. Valores de rotación obtenidos para datos reales, con filtro de mediana y FT 3D o 2D. En negrita se muestran los errores, que en todos los casos son superiores a la tolerancia máxima.

Filt	ROI	x [mm]	y[mm]	z [mm]	Tiempo[s]
SI	A	-0.122	0	0.244	2.47
	B	-0.122	0	0.244	0.36
	C	-0.244	0	0	0.24
	D	-0.122	0	0.244	0.19
NO	A	-0.122	0.122	0.244	2.01
	B	-0.122	0	0	0.30
	C	-0.122	0	0	0.16
	D	-0.122	0.122	0.244	0.19

Tabla 3. Valores de traslación obtenidos para una adquisición real de ratón, con y sin filtro de mediana. En negrita se muestran los errores, siendo en todos los casos inferiores a la tolerancia máxima.

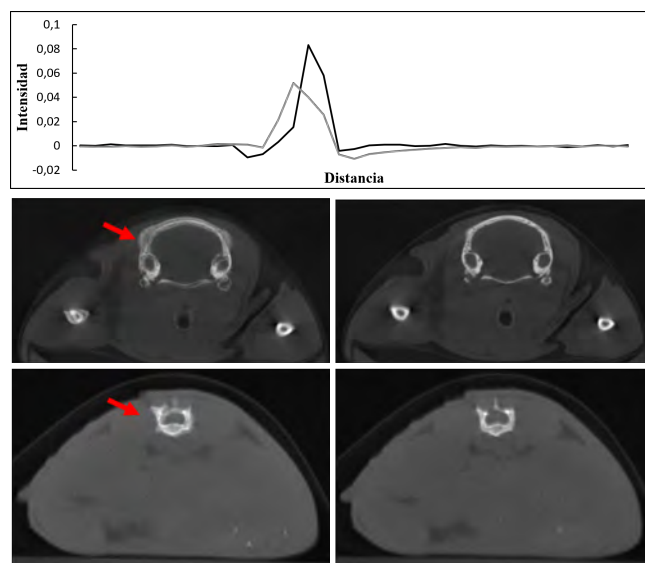


Figura 3: Fila superior: Corte 2D de la función δ del espacio de Fourier sin filtrar (gris) y filtrado (negro). Fila intermedia e inferior: Corte axial de un volumen de rata y de ratón en la zona de solape, sin calibración (izq) y calibrada (drcha). La flecha roja muestra dobles bordes.

Como se puede ver en las Tablas 2 y 3, el algoritmo falló en el cálculo de la rotación y de la traslación para las ROIs de un tamaño del 25% en el corte axial (ROI D) y del 25% en el corte sagital (ROI C), respectivamente. Para la ROI de un tamaño del 50% (ROI B) se obtuvieron los valores correctos de desalineamiento, lográndose

una reducción del 86 % en los tiempos de ejecución respecto al uso de la ROI completa (ROI A).

Por último, la Figura 3 muestra el corte axial con y sin calibración de los dos estudios evaluados donde se puede ver que el método propuesto permitió una correcta corrección de dobles bordes.

4. Discusión

Este trabajo propone un método de calibración automática para el pegado de camas en equipos CBCT basado en una adaptación del algoritmo de Fourier-Miller.

La evaluación en datos simulados muestra un acierto del 83 % de los casos. La aparición de errores en la calibración de la traslación en el eje y puede deberse a la distribución específica del contenido espacial de DIGI-MOUSE.

La evaluación en datos reales mostró errores de traslación por debajo de la tolerancia máxima. En el 75 % de los casos los errores en rotación se debieron al uso de una FT 2D, probablemente debido al desplazamiento relativo en z entre los cortes axiales de las camas que introduce una diferencia entre las imágenes que no puede explicarse solamente mediante una rotación. Por otra parte, la aplicación de un filtro de mediana tuvo un impacto significativo en los resultados gracias a permitir una mejor aproximación de $\delta(\mathbf{A})$. El uso de una ROI de un tamaño del 50 % respecto al corte axial y sagital, que permitió reducir los tiempos de ejecución en un 86 %, demostró disponer de suficiente información espacial para detectar tanto las traslaciones en los tres ejes del espacio como la rotación axial. Como esto depende del contenido espacial de cada volumen, es de esperar que al tratarse de un equipo preclínico todos los sujetos posean un fenotipo similar. No obstante, para asegurar la validez del método se evaluará en nuevos estudios de rata y ratón de distinto peso.

La evaluación preliminar muestra la viabilidad del método propuesto para la calibración automática de adquisiciones multicama en tiempo real. A pesar de que solo se ha evaluado en un equipo preclínico, la traslación a sistemas de geometría de arco en C para aplicaciones clínicas y veterinarias es inmediata.

Agradecimientos

Este trabajo ha sido apoyado por el Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación: PDC2021-121656-I00 (MULTIRAD), financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea ‘NextGenerationEU’/PRTR; PID2019-110369RB-I00/AEI/10.13039/501100011033 (RADHOR). También ha sido financiada por la Comunidad de Madrid: y en el contexto del V PRICIT (Programa Regional de Investigación e Innovación Tecnológica”); S2017/BMD-3867 RENIM-CM, cofinanciado por el Fondo Estructural y de Inversión Europeo. . El CNIC cuenta con el apoyo del Instituto de Salud Carlos III, el Ministerio de Ciencia e Innovación y la Fundación Pro CNIC.

Referencias

- [1] Mónica Abella, Juan José Vaquero, Alejandro Sisniega, Javier Pascau, A Udías, V García, Irina Vidal, and Manuel Desco. Software architecture for multi-bed fdk-based reconstruction in x-ray ct scanners. *Computer methods and programs in biomedicine*, 107(2):218–232, 2012.
- [2] Junyuan Liu, Xi Li, Siwan Shen, Xiaoming Jiang, Wang Chen, and Zhangyong Li. Research on panoramic stitching algorithm of lateral cranial sequence images in dental multifunctional cone beam computed tomography. *Sensors 2021, Vol. 21, Page 2200*, 21:2200, 3 2021.
- [3] Jiaqi Yang, Zhiqiang Huang, Siwen Quan, Qian Zhang, Yanning Zhang, and Zhiguo Cao. Toward efficient and robust metrics for ransac hypotheses and 3d rigid registration. *IEEE Transactions on Circuits and Systems for Video Technology*, 32:893–906, 2 2022.
- [4] Paul J Besl and Neil D McKay. Method for registration of 3-d shapes. In *Sensor fusion IV: control paradigms and data structures*, volume 1611, pages 586–606. Spie, 1992.
- [5] William Scott Hoge. A subspace identification extension to the phase correlation method. *IEEE Transactions on Medical Imaging*, 22:277–280, 2 2003.
- [6] S. Kopp and P. Ottl. Dimensional stability in composite cone beam computed tomography. *Dentomaxillofacial Radiology*, 39:512–516, 12 2010.
- [7] Moon Key Kim, Sang Hoon Kang, Eun Ha Lee, Sang Hwy Lee, and Wonse Park. Accuracy and validity of stitching sectional cone beam computed tomographic images. *Journal of Craniofacial Surgery*, 23:1071–1076, 7 2012.
- [8] Ruud H.H. Wellenberg, Johannes G.G. Dobbe, Jukka Erkkilä, Mario Maas, and Geert J. Streekstra. Markerless assessment of the geometric error of fused cone-beam computed tomography images of the foot constructed using stitching software. *Acta Radiologica*, 62:1341–1348, 10 2020.
- [9] Yan Song, Bo He, Lunjuan Zhang, and Tianhong Yan. Side-scan sonar image registration based on modified phase correlation for auv navigation. *OCEANS 2016 - Shanghai*, 6 2016.
- [10] J. J. Vaquero, S. Redondo, E. Lage, M. Abella, A. Sisniega, G. Tapias, M. L. Soto Montenegro, and M. Deseo. Assessment of a new high-performance a small-animal x-ray tomograph. *IEEE Transactions on Nuclear Science*, 55:898–905, 2008.
- [11] Belma Dogdas, David Stout, Arion F. Chatzioannou, and Richard M. Leahy. Digimouse: a 3d whole body mouse atlas from ct and cryosection data. *Physics in Medicine & Biology*, 52:577, 1 2007.
- [12] Monica Abella, Estefania Serrano, Javier Garcia-Blas, Ines García, Claudia De Molina, Jesus Carretero, and Manuel Desco. Fux-sim: Implementation of a fast universal simulation/reconstruction framework for x-ray systems. *PLOS ONE*, 12:e0180363, 7 2017.
- [13] Javier Garcia Blas, Monica Abella, Florin Isaila, Jesus Carretero, and Manuel Desco. Surfing the optimization space of a multiple-gpu parallel implementation of a x-ray tomography reconstruction algorithm. *Journal of Systems and Software*, 95:166–175, 2014.

Señales Biomédicas (III)

Viernes 24 8:30 - 10:00

Salón de Grados

Caracterización del Acoplo Cardiopulmonar en Pacientes Pediátricos con Apnea Obstruktiva del Sueño

P. Armañac-Julían^{1,2,*}, A. Martín-Montero^{1,3}, S. Hietakoste⁴, J. Lázaro^{1,2}, S. Kainulainen⁴, D. Gozal⁵, R. Hornero^{1,3}, P. Laguna^{1,2}, G. Gutiérrez-Tobal^{1,3}, E. Gil^{1,2}, R. Bailón^{1,2}

¹ CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Madrid, España, parmanac@unizar.es

² BSICoS group, Instituto de Investigación en Ingeniería de Aragón, Zaragoza, España

³ Grupo de Ingeniería Biomédica (GIB), Universidad de Valladolid, España

⁴ Department of Technical Physics, University of Eastern Finland, Kuopio, Finlandia

⁵ Office of the Dean, Joan C. Edwards School of Medicine, Marshall University, Huntington, WV, United States

Resumen

Este estudio tiene como objetivo investigar el uso del acoplo cardiopulmonar (CPC) como biomarcador para caracterizar la severidad de la apnea obstructiva del sueño (OSA) en niños. El análisis de CPC se basa en la coherencia tiempo-frecuencia (TFC) entre la señal de esfuerzo respiratorio y la variabilidad de ritmo cardíaco. Analizamos un total de 255 niños sin OSA, y con OSA leve, moderada y grave durante la vigilia, el sueño en fase REM y el sueño en fase no REM (NREM). Los resultados mostraron que la TFC en la banda de baja frecuencia (LF) aumentó significativamente con la severidad de la OSA tanto en NREM ($p < 0.001$) como en sueño REM ($p < 0.001$). Por el contrario, el acoplo cardiopulmonar, CPC, caracterizado mediante la TFC en la banda HF, es significativamente menor conforme la severidad de OSA aumenta, durante NREM ($p = 0.02$) y REM ($p = 0.03$). Los hallazgos sugieren que la TFC podría ser un biomarcador útil con el que evaluar la severidad de la OSA en niños y podría proporcionar información adicional sobre los mecanismos patológicos subyacentes.

1. Introducción

La apnea obstructiva del sueño (OSA) es un trastorno común del sueño caracterizado por episodios repetitivos de obstrucción de las vías respiratorias superiores durante el sueño. El estándar para el diagnóstico de OSA es la polisomnografía nocturna (PSG), con la cual se miden diversas señales fisiológicas, entre las que se incluyen electroencefalografía, electrocardiografía (ECG), varias señales respiratorias y la saturación de oxígeno (SpO₂). Sin embargo, el diagnóstico de OSA es laborioso, costoso y requiere equipo especializado y personal capacitado. Por lo tanto, existe la necesidad de herramientas alternativas que sean más convenientes y fácilmente accesibles, con el objetivo de ayudar y mejorar el proceso de diagnóstico.

El acoplo cardiopulmonar (CPC) es un fenómeno fisiológico que refleja la interacción entre la regulación cardíaca y respiratoria. Al analizar la variabilidad del ritmo cardíaco (HRV) y las señales de esfuerzo respiratorio, ambos registrados durante la PSG, se puede evaluar el CPC. La arritmia sinusal respiratoria (RSA) es la principal expresión del CPC, que es la aceleración y desaceleración

del corazón en respuesta a la inspiración y expiración, respectivamente. Esta RSA mejora el intercambio de gases pulmonares y facilita la eficiencia cardíaca, sincronizando la perfusión de sangre y la ventilación durante el ciclo respiratorio. Los estudios sugieren que el CPC puede servir como biomarcador ambulatorio eficaz para la calidad del sueño [1,2], y el CPC basado en la coherencia tiempo-frecuencia (TFC) también muestra potencial para predecir la preparación para la extubación en unidades de cuidados intensivos [3].

Aunque algunos estudios han explorado el potencial del análisis de CPC para la OSA, la investigación en esta área sigue siendo limitada. Estudios anteriores encontraron un CPC disminuido en adultos con OSA en comparación con controles sanos, lo que sugiere que podría ser una herramienta útil para el diagnóstico y monitorización de esta condición. También encontraron una preponderancia de potencia en la banda de baja frecuencia (LF) en adultos con OSA, que podría estar asociada con comportamientos anormales durante la respiración alterada del sueño, como la respiración periódica, mientras que los altos valores de CPC en la banda de alta frecuencia (HF) están asociados con RSA saludable y el sueño profundo [1,4]. Otros estudios han mostrado que el envejecimiento puede llevar a una modulación autonómica reducida en adultos mayores con OSA durante la vigilia, la fase N2 y el sueño REM, en comparación con individuos más jóvenes [5].

En este estudio, buscamos caracterizar el CPC pacientes pediátricos con OSA y explorar su potencial como herramienta de diagnóstico para esta condición. Tenemos la hipótesis de que el aumento de la severidad de la OSA también está relacionado con una disminución del CPC y la observada regulación desequilibrada del sistema nervioso autónomo (SNA) en adultos. El sueño estable, estimado por el CPC en la banda HF, mostraría reducciones en pacientes pediátricos con OSA, en comparación con pacientes que se recuperaron de OSA.

2. Materiales

El *Childhood Adenotonsillectomy Trial* (CHAT) fue un ensayo aleatorizado prospectivo diseñado para evaluar la eficacia de varios tratamientos para la OSA en niños [6]. Todos los pacientes pediátricos se sometieron a una PSG

nocturna de seguimiento en un laboratorio clínico para evaluar el estado de su enfermedad, siete meses después de que se les aplicó el correspondiente tratamiento. Nuestro estudio incluye a 255 pacientes, entre las edades de 5-10 años, tanto niños como niñas. La severidad de la OSA se estableció utilizando el índice de apnea-hipopnea (AHI) de acuerdo con las directrices establecidas [6], con categorías de severidad que varían desde no OSA ($AHI < 1$, $N=63$), hasta leve ($1 \leq AHI < 5$, $N=135$), moderada ($5 \leq AHI < 10$, $N=30$) y OSA severa ($AHI \geq 10$, $N=27$). Recopilamos datos de las grabaciones de PSG, incluyendo el ECG, las señales de esfuerzo respiratorio torácico y abdominal, e información sobre las etapas del sueño, identificadas y etiquetadas en segmentos de 30 segundos por clínicos especialistas en sueño.

3. Métodos

3.1. Preprocesado

Primero, las dos señales de esfuerzo respiratorio, $r(t)$, se utilizan para obtener la señal de frecuencia, $F_r(t)$, que se deriva usando un método de promediado espectral condicionado por la picudez [7]. Luego, la señal del ECG se sobremuestra a 1000 Hz con interpolación mediante spline cúbica y las ondas R se detectan usando un método basado en wavelets. Con el fin de explicar la regulación cardíaca por el SNA, la señal HRV se estima con el modelo de Modulación de Frecuencia de Pulso Integral Variante en el Tiempo. Dada una serie particular de latidos cardíacos después de corrección ectópica, la HR instantánea puede expresarse como $d_{HR}(t) = (1 + m(t))/T(t)$. El término $m(t)$ representa la señal moduladora, que se supone contiene la modulación del SNA, y la media instantánea-HR se obtiene mediante un filtrado de paso bajo de $d_{HR}(t)$ a 0.03 Hz. Las versiones uniformemente muestreadas se consiguen remuestreando a 4Hz. Para el lector interesado, los métodos se describen extensamente en [3].

Dado que la respiración afecta a la HRV a través de cambios en la frecuencia respiratoria y el patrón respiratorio, el análisis de HRV es guiado por la respiración. Para ello, la banda HF se establece para estar centrada en la frecuencia respiratoria, $F_r(t)$, se denota HF_C, y es variable en el tiempo: $\Omega_{HF_C}(t) = [F_r(t)-0.125, F_r(t)+0.125]$ Hz. Esto viene motivado por el hecho de que la tasa de respiración promedio en niños no es estacionaria y que, comúnmente, supera las 24 respiraciones por minuto (0.4Hz), lo cual se encuentra en el límite superior de la banda HF clásica, ya que fue establecida en población adulta. La banda LF se define usando los límites clásicos: $\Omega_{LF}(t) = [0.04, 0.15]$ Hz.

3.2. Coherencia tiempo-frecuencia

La influencia de la respiración en la HRV, es decir, el acoplo cardiopulmonar, CPC, se puede estimar usando la coherencia espectral [8]:

$$\hat{\gamma}(t, f) = \frac{|\hat{S}_{r,m}(t, f)|}{\sqrt{\hat{S}_r(t, f) \cdot \hat{S}_m(t, f)}}$$

donde $\hat{\gamma}(t, f) \in [0, 1]$. $\hat{S}_r(t, f)$ y $\hat{S}_m(t, f)$ son las densidades espectrales de potencia variable en el tiempo calculadas mediante la Distribución de Wigner Ville de la Clase Cohen de la respiración y HRV, $r(t)$ y $m(t)$, respectivamente. $\hat{S}_{r,m}(t, f)$ se corresponde con la densidad espectral de potencia cruzada. Se elige una resolución de tiempo y frecuencia de 11.25 s y 0.039 Hz, respectivamente [8]. La Figura 1 proporciona un ejemplo ilustrativo de la coherencia espectral en el dominio tiempo-frecuencia.

Se debe establecer un nivel de coherencia significativo entre HRV y respiración mediante un umbral, $\gamma_{TH}(t, f; \alpha)$. Este umbral de coherencia significativa se establece calculando la coherencia espectral, $\hat{\gamma}(t, f)$, de dos señales de ruido blanco gaussiano de 5 minutos de duración, con $\alpha = 1\%$ de probabilidad de que esas dos señales estén acopladas cuando no debería haber acoplamiento $\gamma_{TH}(t, f; 0.01) = \gamma_0$. Esto se repite iterativamente 1000 veces, y el percentil 99 de $\hat{\gamma}(t, f)$ puede establecerse como el umbral de coherencia espectral significativa, $\gamma_0 = 0.8860$.

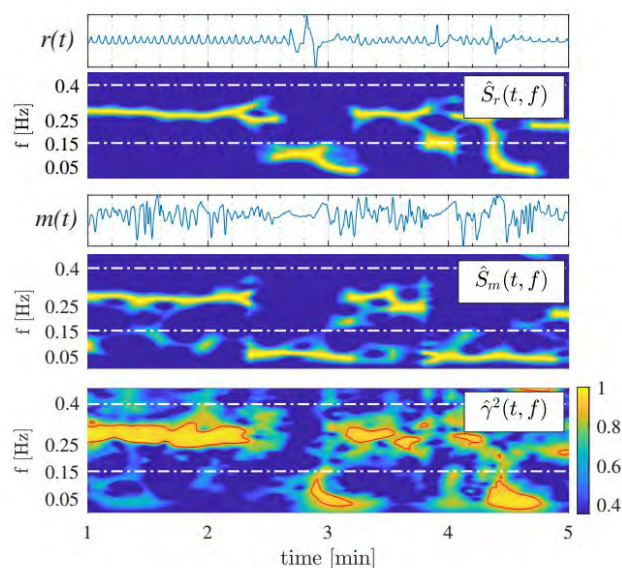


Figura 1. Coherencia Tiempo-Frecuencia. A partir de la señal respiratoria, $r(t)$ y de la señal HRV, $m(t)$, se calculan los espectros TF $\hat{S}_r(t, f)$ y $\hat{S}_m(t, f)$, respectivamente. La coherencia espectral cuadrática, $\hat{\gamma}^2(t, f)$, finalmente se obtiene de los espectros cruzados entre respiración y HRV. La coherencia espectral significativa, $\hat{\gamma}(t, f) > \gamma_0$, está delineada en rojo.

El biomarcador de CPC, basado en la TFC entre HRV y respiración en la banda HF_C [3], denotado como $\mathbb{C}_{HF_C}^T$, se compone considerando la coherencia espectral promedio por encima de γ_0 , es decir, \mathbb{C}_{HF_C} , promediado con el porcentaje de tiempo en coherencia significativa en el segmento de estudio, τ_{HF_C} (Fig. 1):

$$\mathbb{C}_{HF_C}^T = \mathbb{C}_{HF_C} \cdot \tau_{HF_C}$$

Para el CPC, por definición, C_{HFc}^T se debe calcular en la banda HFc. Sin embargo, aunque no sea una medida de acoplo cardiopulmonar, la coherencia espectral también puede calcularse en diferentes bandas espectrales como la banda LF, con la lectura TFC-LF como C_{LF}^T , la cual informa de otros fenómenos fisiológicos distintos del CPC.

3.3. Análisis estadístico

Los biomarcadores de CPC se derivan usando segmentos de 5 minutos. Realizamos un análisis separado de los resultados de CPC durante las tres etapas del sueño: vigilia (W), sueño con movimientos rápidos de los ojos (REM) y sueño no REM (NREM). Para que un segmento se considere en el análisis, debe tener al menos el 90% de su tiempo en la misma etapa de sueño. Para cada paciente, se calcula el CPC promedio en las épocas en la misma etapa de sueño a lo largo de las grabaciones nocturnas.

Las características de TFC consideradas no se ajustan ni a las pruebas de normalidad ni de homocedasticidad, por lo que se realiza una prueba de Kruskal-Wallis (KW) para comparar las diferencias en los biomarcadores entre los cuatro grupos de severidad (sin OSA, OSA leve, OSA moderada y OSA grave). Un valor $p < 0.05$ para la prueba KW puede considerarse de significación estadística.

Posteriormente, se empleó una prueba de rangos con signo emparejada para comparar las diferencias en los valores de TFC de cada paciente entre las etapas del sueño. Se considera un valor $p < 0.01$ para significación estadística, y de esta manera corregir por haber hecho múltiples comparaciones.

4. Resultados y Discusión

La Figura 2 muestra los diagramas de caja (boxplots) de la TFC en las bandas LF y HFc (C_{LF}^T y C_{HFc}^T , respectivamente), comparando los valores de los 4 grupos de severidad de OSA para las tres etapas del sueño. Los niveles de CPC en cada etapa del sueño, medidos por TFC en la banda HFc, son significativamente más bajos para las categorías crecientes de severidad de OSA durante el sueño NREM (prueba KW, $p = 0.02$) y el sueño REM (prueba KW, $p = 0.03$). Por el contrario, la TFC en la banda LF es significativamente más alta para las categorías crecientes de severidad de OSA, tanto durante NREM (prueba KW, $p < 0.001$) como durante el sueño REM (prueba KW, $p = 0.03$), lo cual es consistente con resultados encontrados en adultos [1].

La Tabla 1 muestra los p-valores de la prueba de rangos con signo, comparando los valores de TFC para las tres etapas del sueño en las bandas LF y HFc, de los diferentes niveles de severidad de OSA. El análisis estadístico muestra que existen diferencias en el CPC (C_{HFc}^T) en todas las etapas, excepto para los niños con OSA severa, estableciendo el hecho de que es necesario un análisis por separado en las etapas del sueño para la caracterización de la OSA. No se encontraron diferencias significativas en la banda LF en pacientes con OSA grave comparando valores de TFC en REM y NREM, mientras que estas diferencias son claras para pacientes sin OSA y con OSA leve. Tal y como se hipotetizó, el CPC también es significativamente

más bajo durante la vigilia en comparación con NREM y REM en todas las categorías de OSA (Tab. 1b). Según investigaciones anteriores, procesos como la apnea del sueño y la fibromialgia, que conducen a la fragmentación del sueño, han demostrado reducir la cantidad de CPC (TFC en la banda HFc) [1]. Además, se ha asociado que valores más altos de TFC en la banda LF tienen una mayor prevalencia de hipertensión y accidente cerebrovascular en adultos [9].

(a) C_{LF}^T	W-NREM	W-REM	NREM-REM
No OSA	$\ll 0.01$	$\ll 0.01$	$\ll 0.01$
OSA leve	0.03	$\ll 0.01$	$\ll 0.01$
OSA moderada	0.54	0.09	0.03
OSA severa	0.004	0.05	0.02
(b) C_{HFc}^T	W-NREM	W-REM	NREM-REM
No OSA	$\ll 0.01$	$\ll 0.01$	$\ll 0.01$
OSA leve	$\ll 0.01$	$\ll 0.01$	$\ll 0.01$
OSA moderada	$\ll 0.01$	$\ll 0.01$	$\ll 0.01$
OSA severa	$\ll 0.01$	$\ll 0.01$	0.04

Tabla 1. P-valores obtenidos del test de rango con signo pareado, comparando los valores de TFC de cada paciente entre las tres etapas del sueño. El análisis se realiza en las bandas LF (a) y HFc (b), para los diferentes niveles de severidad de la OSA. Se considera significación estadística para p-valores < 0.01 , para corregir por múltiples comparaciones.

En general, la cantidad de eventos de apnea/hipopnea es comparable entre el sueño REM y NREM [10]: aproximadamente el 88% de los episodios de 10 minutos en el sueño REM tuvieron menos de 5 eventos por episodio, y el 97% de los episodios en el sueño NREM tuvieron menos de 5 eventos de apnea/hipopnea por episodio. El aumento observado del acoplamiento en la banda LF en pacientes con OSA severa podría atribuirse a la mayor prevalencia de respiración periódica durante el sueño REM, según se informa en [1], así como a las pronunciadas variaciones cíclicas en HRV en respuesta a episodios repetidos de apnea. Sin embargo, los valores de CPC en la banda HFc fueron más altos en el sueño REM en comparación con NREM, o al menos similares para los pacientes con OSA severa, lo que explica que la disminución significativa de CPC con la creciente severidad de la OSA puede no deberse a efectos relacionados con eventos de apnea, sino más bien a otros factores fisiológicos como alteraciones en la actividad simpática.

Cabe señalar que es necesario el uso de un umbral de coherencia significativo. Muchos trabajos existentes que estudian el CPC basado en la coherencia espectral, como biomarcador de la calidad del sueño, no consideran el hecho de que dos señales de ruido blanco tendrán un nivel base de coherencia γ_0 , donde se debería reportar nula por definición. Nuestros resultados muestran que esta metodología podría proporcionar información fenotípica adicional para clasificar mejor entre las etapas del sueño, ya que la vigilia y el sueño REM a veces son indistinguibles [1].

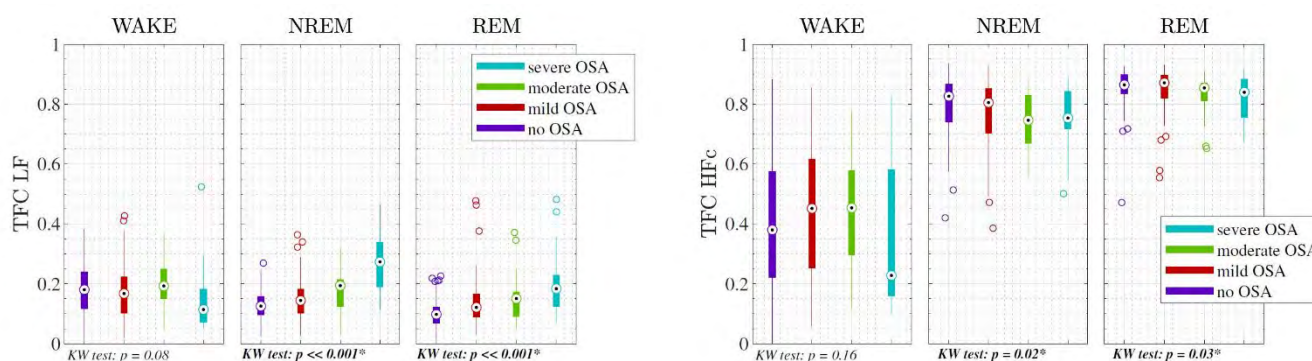


Figura 2. Boxplots de la TFC en la banda LF, C_{LF}^I , (izquierda) y en la banda HF, C_{HF}^I , (derecha), comparando los valores de los 4 grupos de severidad de la OSA (sin OSA, leve, moderada y severa), para las tres etapas del sueño (Vigilia, NREM y REM). Los resultados del test de Kruskal Wallis (KW) se muestran en cada gráfico en la parte de abajo, entre los valores de TFC de los grupos de severidad de la OSA para cada etapa del sueño.

Este estudio tiene la limitación de que se utilizaron señales respiratorias distintas a la señal de presión nasal. Estudios anteriores también han utilizado señales respiratorias alternativas [1,4], como la respiración derivada del ECG (EDR). De hecho, Varon et al. reportaron que los errores en CPC eran significativamente mayores durante los eventos de apnea que durante la actividad normal al usar señales EDR como sustituto [11]. Este hecho puede ser debido a que los movimientos del pecho capturados por la EDR, pueden no estar relacionados con la respiración real durante los eventos de apnea, lo que causa una sobrestimación del CPC, al igual que puede estar ocurriendo con las bandas de esfuerzo respiratorio usadas en este trabajo. Por ello, hemos demostrado la utilidad del CPC utilizando grabaciones de bandas de esfuerzo respiratorio, pero los trabajos futuros deberían considerar el uso de la señal de la cánula de presión nasal, lo que podría llevar a una potencial reducción de los valores de CPC en presencia de eventos respiratorios obstructivos.

5. Conclusión

En general, podemos concluir que la TFC en la banda LF podría ser un biomarcador útil para evaluar la severidad de la OSA, mientras que el CPC, medido por TFC en la banda HF, podría proporcionar información adicional sobre los mecanismos patológicos subyacentes a la OSA. Sin embargo, se necesitan estudios adicionales con muestras más grandes para confirmar estos hallazgos e investigar el uso de señales respiratorias en conjunto con el análisis de HRV.

Agradecimientos

Este trabajo fue apoyado por el CIBER-BBN a través del Instituto de Salud Carlos III, proyectos de valorización SleepyHeart y TinyHeart; por el MINECO (PID2021-126734OB-C21); y por el Gobierno de Aragón (Grupo de

Referencia BSICoS T39-23R) cofinanciado por FEDER 2014-2020.

Referencias

- [1] Thomas RJ, et al. Cardiopulmonary coupling spectrogram as an ambulatory clinical biomarker of sleep stability and quality in health, sleep apnea, and insomnia. *Sleep* 2018;41(2).
- [2] Choi JH, et al. Sleep quality change after upper airway surgery in obstructive sleep apnea: electrocardiogram-based cardiopulmonary coupling analysis. *Laryngoscope* 2015;125(7).
- [3] Armañac-Julián P, et al. Cardiopulmonary coupling indices to assess weaning readiness from mechanical ventilation. *Sci Rep* 2021;11(1).
- [4] Thomas RJ, et al. An electrocardiogram-based technique to assess cardiopulmonary coupling during sleep. *Sleep* 2005;28(9).
- [5] Trimer R, et al. Heart rate variability and cardiorespiratory coupling in obstructive sleep apnea: elderly compared with young. *Sleep Medicine* 2014;15(11).
- [6] Marcus CL, et al. A randomized trial of adenotonsillectomy for childhood sleep apnea. *N Engl J Med* 2013;368.
- [7] Kontaxis S, et al. ECG-derived respiratory rate in atrial fibrillation. *IEEE TBME* 2019;67(3).
- [8] Orini M, et al. Characterization of dynamic interactions between cardiovascular signals by time-frequency coherence. *IEEE TBME* 2011;59(3).
- [9] Thomas RJ, et al. Prevalent hypertension and stroke in the sleep heart health study: association with an ECG-derived spectrographic marker of cardiopulmonary coupling. *Sleep* 2009;32(7).
- [10] Martín-Montero A, et al. Pediatric sleep apnea: Characterization of apneic events and sleep stages using heart rate variability. *Comput. Biol. Med.* 2023.
- [11] Varon C, et al. A comparative study of ECG-derived respiration in ambulatory monitoring using the single-lead ECG. *Sci Rep* 2020;10(1).

DetECCIÓN AUTOMÁTICA DE LA APNEA DEL SUEÑO INFANTIL UTILIZANDO TÉCNICAS DE *DEEP LEARNING* Y *EXPLAINABLE ARTIFICIAL INTELLIGENCE* EN SEÑALES DE FLUJO AÉREO

V. Barroso García^{1,2}, F. Vaquerizo Villar^{1,2}, G.C. Gutiérrez Tobal^{1,2}, C. García Vicente¹, D. Álvarez González^{1,2}, D. Gozal³, R. Hornero Sánchez^{1,2}

¹ Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, veronica.barroso@gib.tel.uva.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

³ Office of The Dean, Joan C. Edwards School of Medicine, Marshall University, Huntington, Virginia Occidental, EEUU, gozal@marshall.edu

Resumen

La alta prevalencia de la apnea obstructiva del sueño (AOS) pediátrica y las limitaciones de la prueba diagnóstica estándar han fomentado el estudio de estrategias alternativas que ayuden a su diagnóstico automático. Los métodos propuestos suelen basarse en técnicas de *feature engineering*, lo que implica una complejidad y subjetividad inherente. Otros utilizan técnicas de *deep learning*, que mejoran el rendimiento diagnóstico pero carecen de transparencia e interpretabilidad. En este trabajo proponemos evaluar un modelo explicable basado en redes neuronales convolucionales (CNN) para estimar la severidad de la AOS infantil utilizando la señal de flujo (FA). Para ello, se analizaron 1638 registros de FA, que fueron divididos en segmentos de 10 minutos. El modelo CNN estimó el número de eventos apneicos por segmento. Después, se aplicó el algoritmo Grad-CAM para identificar las regiones de FA en las que se fija la CNN al hacer su predicción. El modelo propuesto mostró una alta concordancia entre el índice de apnea-hipopnea estimado y el real (coeficiente de correlación intraclass = 0.87 en el grupo de test), así como un alto rendimiento diagnóstico (κ de 4 clases = 0.38 y precisiones del 81.05%, 85.62% y 92.81% para 1, 5 y 10 eventos/h en el grupo de test). Grad-CAM reveló que la CNN se centra en el comienzo y el final del evento apneico, es decir, donde la señal FA cambia bruscamente de amplitud. Así, nuestra propuesta sería muy útil para identificar automáticamente los patrones respiratorios asociados con la AOS infantil y ayudar en su diagnóstico.

1. Introducción

La apnea obstructiva del sueño (AOS) es un trastorno respiratorio muy común y altamente prevalente. Se estima que entre el 1% y el 5.7% de los niños sufren algún grado de la enfermedad [1], aunque el porcentaje exacto puede variar según otros trastornos coexistentes [2], [3]. Los niños con AOS experimentan pausas recurrentes (apneas) y/o reducciones del flujo aéreo (hipopneas) mientras duermen, que pueden durar desde 6 segundos (2 ciclos respiratorios) hasta varios minutos cada una [4]. La presencia de estos eventos conduce a hipoxemia, hipercapnia, micro despertares y, en general, a un sueño fragmentado y no reparador que predispone a los sujetos pediátricos afectados a sufrir serias comorbilidades, como disfunciones cardiovasculares, metabólicas y

neurodegenerativas [1], [5]. Por eso, es de suma importancia que estos niños sean diagnosticados y tratados a tiempo.

El método de referencia para su diagnóstico es la polisomnografía nocturna (PSG), una prueba en la que se registra la actividad cerebral, la respiración, la frecuencia cardíaca y otros signos vitales durante el sueño [6]. Si bien su eficacia y fiabilidad son ampliamente conocidas, la PSG también es compleja, costosa, incómoda para los niños y de disponibilidad muy limitada [7]. Este hecho, unido a la alta prevalencia de la enfermedad, ha fomentado el estudio de otras estrategias que ayuden a diagnosticar la AOS automáticamente.

El procesamiento de señales cardiorrespiratorias, como la señal de flujo aéreo (FA), ha demostrado ser especialmente útil para este propósito [8]–[12]. Esta señal es capaz de reflejar la actividad respiratoria del niño mientras duerme y, por ende, los ceses respiratorios asociados con la AOS [4]. Es por ello que los métodos de *feature engineering* han resultado ser tan eficaces a la hora de caracterizar los cambios que experimenta esta señal en el dominio del tiempo y de la frecuencia cuando se producen los eventos apneicos [8]–[12]. No obstante, es importante destacar que estos métodos requieren un conocimiento previo y exhaustivo de las características a extraer [13]. Además, al depender de características diseñadas manualmente, podría haber dificultades para identificar las relaciones complejas y no lineales presentes en los datos, perdiendo así información relevante [13]. Por el contrario, los enfoques basados en *deep learning* (DL) podrían superar estas limitaciones y mejorar la capacidad diagnóstica del FA. Así, las redes neuronales convolucionales (CNN) han mostrado un alto rendimiento diagnosticando la AOS infantil a partir de señales cardiorrespiratorias [14], [15]. Si bien estos modelos son capaces de identificar automáticamente las características útiles de la señal en crudo, hay que tener en cuenta que pierden interpretabilidad y transparencia con respecto a aquellos basados en *feature engineering* [16]. Para lidiar con esta limitación, en este estudio proponemos aplicar un modelo CNN a la señal de FA junto con técnicas de *explainable artificial intelligence* (XAI) que permitan

comprender los resultados reportados y mejorar la fiabilidad de dicho modelo. Concretamente, aplicamos el algoritmo *Gradient-weighted Class Activation Mapping* (Grad-CAM) dado su éxito dotando de interpretabilidad a este tipo de redes [16]. Así, los objetivos de nuestro estudio fueron: (i) evaluar la utilidad de un modelo CNN para determinar la severidad de AOS pediátrica a partir del FA e (ii) identificar las características de esta señal en las que se fija la CNN para hacer su predicción.

2. Sujetos y señales

Nuestra investigación involucró 1638 registros polisomnográficos procedentes de la base de datos pública *Childhood Adenotonsillectomy Trial* (CHAT, #NCT00560859), la cual está disponible bajo petición en <https://sleepdata.org/datasets/chat>. El Comité Ético aprobó el protocolo de estudio y todos los tutores legales de los niños dieron su consentimiento informado.

Los sujetos fueron diagnosticados siguiendo las reglas de la *American Academy of Sleep Medicine* (AASM) [4]. A partir del índice de apnea-hipopnea (IAH) obtenido, los especialistas médicos determinaron si estos niños estaban libres de AOS (1 evento/hora (e/h) > IAH), o si padecían AOS leve (1 e/h ≤ IAH < 5 e/h), moderada (5 e/h ≤ IAH < 10 e/h) o severa (IAH ≥ 10 e/h). La Tabla 1 muestra sus datos demográficos y clínicos. La población bajo estudio fue distribuida aleatoriamente en tres grupos: entrenamiento (61.4%), validación (19.9%) y test (18.7%). No se encontraron diferencias estadísticamente significativas (p -valor > 0.01 al aplicar el test de Mann-Whitney U con corrección de Bonferroni) en edad, sexo, índice de masa corporal (IMC) ni IAH entre los grupos establecidos.

Los registros de FA utilizados se adquirieron con un termistor a frecuencias de muestreo de entre 20 y 512 Hz durante la PSG. Después, éstos fueron remuestreados a 4 Hz y estandarizados como en estudios previos [14]. Finalmente, se segmentaron en tramas de 10 minutos para alimentar la CNN propuesta en este estudio.

3. Métodos

3.1. Arquitectura del modelo CNN

La arquitectura propuesta en nuestro estudio está inspirada en el modelo CNN desarrollado en trabajos previos para estimar la severidad de AOS infantil a partir de la señal de oximetría y/o FA [14], [15]. Así, la CNN recibe como entrada los segmentos de 10 minutos de la señal de FA previamente estandarizados [14], [15]. Después, los datos son procesados mediante N_c bloques convolucionales, cada uno de ellos formado por las siguientes capas consecutivas: convolución 1D, *batch normalization*, *Rectified Linear Unit* (ReLU), *max pooling* y *dropout* [14], [15]. Así, en cada bloque se aplican N_f filtros a los datos de entrada para extraer los mapas de características (convolución), seguido de la normalización de estos (*batch normalization*) y de una activación no lineal para identificar las relaciones complejas existentes en ellos (ReLU) [13]. Después, se

	Entrenamiento	Validación	Test
Sujetos (n)	1006 (61.4%)	326 (19.9%)	306 (18.7%)
Edad (años)	7.0 [2.0]	7.0 [2.0]	6.9 [2.0]
Varones (n)	471 (46.8%)	156 (47.9%)	134 (43.8%)
IMC (kg/m²)	17.4 [6.1]	17.1 [6.4]	17.6 [6.0]
IAH (e/h)	2.6 [4.8]	2.4 [4.6]	2.3 [5.1]
No AOS (n)	219 (21.8%)	69 (21.2%)	67 (21.9%)
AOS Leve (n)	496 (49.3%)	168 (51.5%)	148 (48.4%)
AOS Moderado (n)	160 (15.9%)	44 (13.5%)	49 (16.0%)
AOS Severo (n)	131 (13.0%)	45 (13.8%)	42 (13.7%)

IMC: índice de masa corporal; IAH: índice de apnea-hipopnea. Datos presentados como número de sujetos (%) o mediana [rango intercuartil].

Tabla 1. Datos demográficos y clínicos de los sujetos.

reduce la dimensionalidad del mapa de características extrayendo aquellas que son dominantes (*max pooling*) y se reduce la interdependencia entre neuronas para prevenir el sobreajuste eliminando aleatoriamente una proporción p_{drop} de activaciones (*dropout*) [13].

Tras los N_c bloques de convolución, una capa de *flattening* concatena los mapas de características en un único vector unidimensional y, finalmente, una capa de activación lineal realiza la combinación lineal de sus elementos para estimar el número de eventos apneicos contenidos en cada segmento de 10 minutos [14], [15]. El IAH de cada sujeto se obtuvo como el promedio de los valores estimados por la CNN para todos los segmentos contenidos en su señal de FA, seguido de un proceso de regresión lineal para corregir posibles subestimaciones.

Para obtener la configuración óptima de los hiperparámetros N_c , N_f y p_{drop} , se utilizó el algoritmo de estimación del momento adaptativo (método Adam) con una tasa de aprendizaje inicial de 0.001 [15]. Al igual que en estudios previos, se empleó la función *Huber loss* con el parámetro $\delta = 1$, dada su idoneidad en distribuciones aleatorias [14].

3.2. Evaluación del modelo CNN

Para evaluar la concordancia entre los IAH estimados por la CNN y los valores reales obtenidos mediante la PSG, se empleó el coeficiente de correlación intraclase (CCI) y el error cuadrático medio (*mean-squared error*, MSE). El rendimiento diagnóstico se evaluó empleando el κ de 4 clases, la matriz de confusión y la precisión de 4 clases. También se obtuvo la sensibilidad (S), especificidad (E), precisión (P), valor predictivo positivo y negativo (VPP y VPN) y razón de verosimilitud positiva y negativa (LR+ y LR-) para los puntos de corte del IAH 1, 5 y 10 e/h.

3.3. Interpretabilidad del modelo CNN

Se aplicó Grad-CAM para dotar de interpretabilidad a nuestro modelo CNN y comprender cómo éste realiza sus predicciones [16]. Esta técnica de XAI proporciona una visualización intuitiva y contextual de las regiones de la señal de FA que son más influyentes en la estimación de la CNN. Para ello, calcula los gradientes de la predicción con respecto a los mapas de características presentes en cada capa convolucional de la CNN [16]. A partir de estos gradientes, se obtiene un mapa de activación por cada

capa convolucional, es decir, un mapa de calor que resalta con valores más elevados (colores más cálidos) las regiones del FA que más contribuyen a la predicción final de la red. El mapa de calor final se calcula promediando los mapas de activación obtenidos en cada capa.

4. Resultados

4.1. Hiperparámetros óptimos del modelo CNN

Los valores de N_c , N_f y p_{drop} variaron entre [6:1:10], [32, 64, 128, 256] y [0.0, 0.1, 0.2], respectivamente. Cada uno de los modelos configurados con estos valores fueron entrenados en el grupo de entrenamiento. La configuración que obtuvo mayor índice $kappa$ en el grupo de validación fue establecida como óptima: $N_c = 9$, $N_f = 256$ y $p_{drop} = 0.01$, con un $kappa = 0.42$.

4.2. Evaluación del modelo CNN en test

Las Tablas 2 y 3 muestran la matriz de confusión obtenida por la CNN en test, así como las métricas estadísticas derivadas de ella, respectivamente. La arquitectura propuesta alcanzó un CCI de 0.87 y un MSE de 7.90 e/h en test. Además, obtuvo una precisión global del 61.76% (189/306 sujetos correctamente clasificados) y un $kappa$ de 0.38 en test con cuatro clases. Aplicando los umbrales diagnósticos convencionales, nuestro modelo logró precisiones del 81.05%, 85.62% y 92.81% para 1, 5 y 10 e/h en test, respectivamente. También destaca la elevada LR+ alcanzada para 10 e/h (31.43).

		CNN			
		No AOS	Leve	Moderado	Severo
PSC	No AOS	19	48	0	0
	Leve	10	127	11	0
	Moderado	0	26	18	5
	Severo	0	7	10	25

Tabla 2. Matriz de confusión obtenida por el modelo CNN en el grupo de test.

IAH	S (%)	E (%)	P (%)	VPP (%)	VPN (%)	LR+	LR-
1 e/h	95.82	28.36	81.05	82.67	65.52	1.34	0.15
5 e/h	63.74	94.88	85.62	84.06	86.08	12.46	0.38
10 e/h	59.52	98.11	92.81	83.33	93.84	31.43	0.41

Tabla 3. Rendimiento diagnóstico obtenido para 1, 5 y 10 e/h por el modelo CNN en el grupo de test.

4.3. Interpretabilidad del modelo CNN en test

La Figura 1 muestra cuatro ejemplos de los resultados obtenidos con Grad-CAM en segmentos donde la predicción del modelo CNN fue correcta: (a) respiración normal y (b) eventos apneicos, y en aquellos donde la predicción fue errónea: (c) sobrestimación y (d) infraestimación. Como puede observarse en (b), para identificar el evento de apnea el modelo se centra principalmente en el comienzo y el final del mismo.

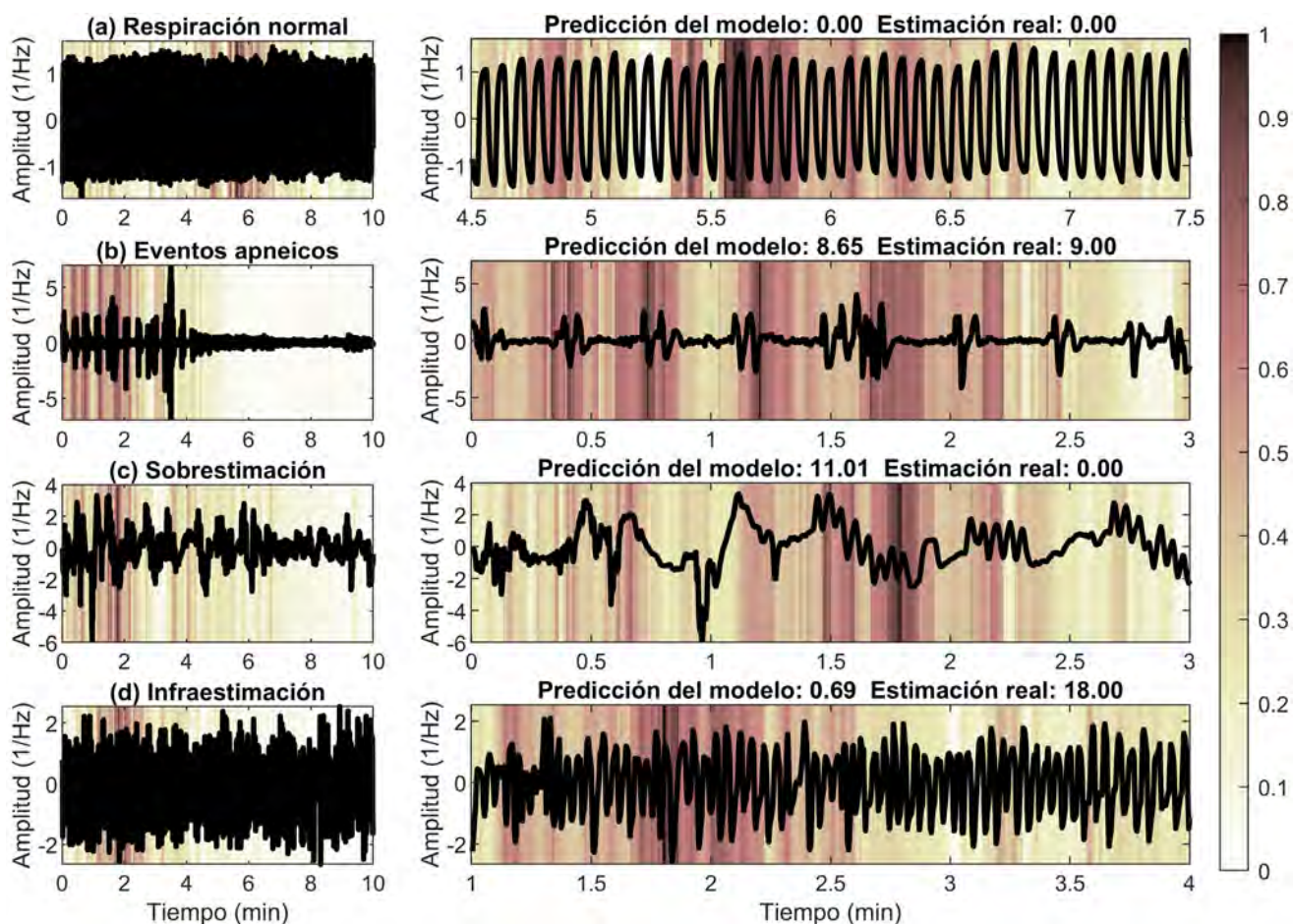


Figura 1. Resultados de Grad-CAM en (a) respiración normal, (b) eventos apneicos, (c) sobrestimación y (d) infraestimación.

5. Discusión y conclusiones

En este estudio se ha desarrollado y testado un modelo interpretable basado en CNN para estimar la severidad de la AOS infantil a partir de la señal de FA. Así, el modelo propuesto alcanzó una alta precisión diagnóstica (81.05%, 85.62% y 92.81% para 1, 5 y 10 e/h, respectivamente), especialmente en los casos de AOS moderada y severa, donde obtuvimos un robusto marcador para determinar su presencia ($LR+ \geq 10$) [11]. Además, Grad-CAM nos permitió identificar los patrones respiratorios en los que se fija la CNN y comprender los resultados que reporta. Tal y como muestra la Figura 1(a), cuando la respiración es normal el mapa de calor alcanza unos valores inferiores, por lo que la CNN se estaría centrando menos en estas regiones. No obstante, aparece un foco de atención en torno al minuto 5.6. La ausencia de cambios de amplitud en esta parte de la señal indicaría que la CNN se fija en otras propiedades del FA, como la distancia entre ciclos respiratorios. En el caso de los eventos apneicos (Figura 1(b)), podemos observar que los valores del mapa de calor se incrementan. Este aumento se mantiene a lo largo de todo el evento, aunque es particularmente más elevado al comienzo y al final. Este hecho revela que, para identificar el patrón respiratorio asociado al evento de apnea, el modelo se fija principalmente en aquellas regiones donde la señal de FA cambia bruscamente de amplitud. Sin embargo, los cambios repentinos de amplitud también pueden ser ocasionados por la presencia de ciertos artefactos, como el ruido fisiológico producido por el movimiento del niño. Como se muestra en la Figura 1(c), estos artefactos podrían conducir a una sobreestimación de los eventos apneicos. Finalmente, Grad-CAM nos permitió comprender los casos de infraestimación del modelo (Figura 1(d)). Así, los resultados obtenidos indicaron que, sin tener información sobre los *arousals* o las desaturaciones, las reducciones parciales del FA que se producen por la ocurrencia de hipopneas podrían pasar desapercibidas para la CNN.

Estudios previos de nuestro grupo analizaron la señal de FA para estimar automáticamente la severidad de AOS en niños [8]–[11]. La mayoría de ellos utilizaron métodos convencionales que alcanzaron precisiones de 73.61%–81.10% para 1 e/h, 57.14%–62.82% para 5 e/h y 70.47%–80.60% para 10 e/h. En este trabajo, se han alcanzado precisiones dentro de este rango para 1 e/h, así como notablemente superiores para 5 y 10 e/h. Otro estudio analizó conjuntamente FA con la señal de oximetría mediante CNN [14], obteniendo 84.10%–84.64%, 84.10%–93.46% y 90.26%–94.44% de precisión en 1, 5 y 10 e/h, respectivamente. Con sólo un canal, en el presente estudio se ha alcanzado un rendimiento sólo ligeramente inferior usando los mismos registros de CHAT.

Este estudio presenta algunas limitaciones que deben ser consideradas. El modelo CNN fue testado únicamente con CHAT, por lo que podría ser validado en otras bases de datos, e incluso con registros domiciliarios de FA, para aumentar la solidez de nuestros resultados. También se podría estimar el IAH utilizando otros métodos de DL, como redes neuronales recurrentes o *transformers*, y otras

técnicas de XAI, como LIME o SHAP, lo que constituye una interesante línea de investigación futura.

En conclusión, el modelo propuesto mostró un alto rendimiento estimando la severidad de la AOS infantil. La mejora obtenida respecto a las técnicas convencionales, así como la identificación de las regiones del FA que más contribuyen a la predicción, pone de manifiesto la eficacia y fiabilidad de usar conjuntamente CNN y FA para simplificar el diagnóstico de AOS en niños. Así, nuestra propuesta sería una herramienta útil para ello.

Agradecimientos

Este estudio ha sido financiado por el Ministerio de Ciencia e Innovación - AEI y ERDF (PID2020-115468RB-I00 y PDC2021-120775-I00) y por el CIBER-BBN (CB19/01/00012). G.C. Gutiérrez Tobal ha recibido una ayuda postdoctoral de la Universidad de Valladolid. D. Álvarez es beneficiario de una ayuda Ramón y Cajal (RYC2019-028566-I) del Ministerio de Ciencia e Innovación - AEI, cofinanciada por el FSE. D. Gozal ha recibido financiación del NIH (AG061824 y HL166617).

Referencias

- [1] C. L. Marcus *et al.*, “Diagnosis and management of childhood obstructive sleep apnea syndrome,” *Pediatrics*, vol. 130, 2012.
- [2] A. Moffa *et al.*, “Childhood Obstructive Sleep Apnea: from Diagnosis to Therapy—an Update,” *Current Sleep Medicine Reports*, vol. 6, no. 3, pp. 157–162, 2020.
- [3] A. C. Bitners *et al.*, “Evaluation and Management of Children with Obstructive Sleep Apnea Syndrome,” *Lung*, vol. 198, no. 2, pp. 257–270, 2020.
- [4] R. B. Berry *et al.*, “Rules for Scoring Respiratory Events in Sleep: Update of the 2007 AASM Manual for the Scoring of Sleep and Associated Events,” *J. Clin. Sleep Med.*, vol. 08, no. 05, pp. 597–619, 2012.
- [5] M. Luz Alonso-Álvarez *et al.*, “Consensus document on sleep apnea-hypopnea syndrome in children,” *Arch. Bronconeumol.*, vol. 47, no. SUPPL. 5, pp. 2–18, 2011.
- [6] C. Jon, “Polysomnography in Children,” in *Pediatric Otolaryngology for the Clinician*, Totowa, NJ: Humana Press, 2009, pp. 35–47.
- [7] P. J. Ryan *et al.*, “Validation of British Thoracic Society guidelines for the diagnosis of the sleep apnoea/hypopnoea syndrome: can polysomnography be avoided?,” *Thorax*, vol. 50, no. 9, pp. 972–5, 1995.
- [8] J. Jiménez-García *et al.*, “Assessment of Airflow and Oximetry Signals to Detect Pediatric Sleep Apnea-Hypopnea Syndrome Using AdaBoost,” *Entropy*, vol. 22, no. 6, pp. 670, 2020.
- [9] V. Barroso-García *et al.*, “Wavelet Analysis of Overnight Airflow to Detect Obstructive Sleep Apnea in Children,” *Sensors*, vol. 21, no. 4, pp. 1491, 2021.
- [10] V. Barroso-García *et al.*, “Bispectral analysis of overnight airflow to improve the pediatric sleep apnea diagnosis,” *Comput. Biol. Med.*, vol. 129, 2021.
- [11] V. Barroso-García *et al.*, “Usefulness of recurrence plots from airflow recordings to aid in paediatric sleep apnoea diagnosis,” *Comput. Meth. Prog. Bio.*, vol. 183, pp. 105083, 2020.
- [12] G. C. Gutiérrez-Tobal *et al.*, “Diagnosis of pediatric obstructive sleep apnea: Preliminary findings using automatic analysis of airflow and oximetry recordings obtained at patients’ home,” *Biomed. Signal Process. Control*, vol. 18, pp. 401–7, 2015.
- [13] I. Goodfellow *et al.*, *Deep learning*. MIT press, 2016.
- [14] J. Jiménez-García *et al.*, “A 2D convolutional neural network to detect sleep apnea in children using airflow and oximetry,” *Comput. Biol. Med.*, vol. 147, pp. 105784, 2022.
- [15] F. Vaquerizo-Villar *et al.*, “A Convolutional Neural Network Architecture to Enhance Oximetry Ability to Diagnose Pediatric Obstructive Sleep Apnea,” *IEEE J. Biomed. Heal. Inform.*, vol. 25, no. 8, pp. 2906–2916, 2021.
- [16] R. R. Selvaraju *et al.*, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization,” *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2017-Oct., pp. 618–626, 2017.

Interpretabilidad de Redes Convolucionales para el Análisis de Episodios de Tos en Enfermedades Respiratorias

P. Amado-Caballero¹, J. R. Garmendia-Leiza², M.D. Aguilar-García², C. Martínez-Fernández-de-Septiem³, A. García-Ruano⁴, M. D. Garmendia-García⁵, A. Crespo-Sedano⁶, C. Alberola-López¹, P. Casaseca-de-la-Higuera¹

¹ Laboratorio de Procesado de Imagen, ETSI Telecomunicación. Universidad de Valladolid, {pamacab@lpi.tel.uva.es}.

² Centro de Salud Los Jardinillos, SACYL, Palencia; ³ Hospital Universitario de Burgos, SACYL, Burgos; ⁴ Centro de Salud de Alaejos, SACYL, Alaejos (Salamanca); ⁵ Fundación Jiménez Díaz, Grupo Quirón, Madrid; ⁶ Complejo Asistencial Universitario de Palencia, SACYL, Palencia.

Resumen

Objetivo: Estudiar la viabilidad de utilizar metodologías de interpretabilidad de redes neuronales convolucionales para el análisis de patrones de tos en la enfermedad respiratoria.

Materiales: Se estudió un grupo de 20 pacientes adultos afectados de enfermedad respiratoria que presentaban la tos como síntoma persistente. Los pacientes se monitorizaron durante 24 horas utilizando una grabadora conectada a un micrófono de solapa. **Métodos:** La señal de audio se transformó al dominio de la frecuencia obteniéndose espectrogramas de 1 segundo de duración. Dichos espectrogramas se procesaron mediante una red neuronal convolucional para identificar los eventos de tos. Los mapas de oclusión que reflejaban las regiones de los espectrogramas que contribuyen activamente a la detección de la tos se analizaron cuantitativamente para identificar diferencias entre grupos de pacientes. **Resultados:** Se observaron diferencias significativas entre el grupo de pacientes con Enfermedad Pulmonar Obstructiva Crónica (EPOC) y grupos que comprendían otras enfermedades. **Conclusiones:** Los métodos de análisis de interpretabilidad de redes neuronales permiten explicar diferencias asociadas a la tos entre pacientes crónicos con EPOC y pacientes con otras patologías, tanto crónicas como no crónicas.

1. Introducción

Las enfermedades respiratorias constituyen una de las mayores causas de muerte y discapacidad en el mundo [1]. Enfermedades crónicas como EPOC (enfermedad pulmonar obstructiva crónica) y cáncer degeneran frecuentemente en situaciones de discapacidad y dependencia [2]. En el caso de la COVID-19, estudios recientes [3] han evidenciado necesidades de atención y monitorización domiciliaria meses después de haber pasado la enfermedad, incluso no habiendo sido hospitalizados. En el caso de las enfermedades respiratorias crónicas como EPOC o cáncer, está demostrado que el daño que produce el contagio de la COVID-19, aumenta considerablemente el riesgo de hospitalización crítica y de muerte [4]. Por esto es doblemente importante monitorizar de forma continua a estos pacientes crónicos y ser capaces de identificar exacerbaciones de su enfermedad.

El estudio encargado por la Comisión Europea sobre telemedicina publicada en [5] resaltó no sólo el alto potencial de la telemedicina en la gestión de enfermedades respiratorias y dependencias asociadas, sino también la

necesidad creciente de incrementar la investigación en esta área específica. Sin embargo, aunque la telemedicina ha mostrado potencial para la mejora del seguimiento y tratamiento de enfermedades respiratorias de una forma asequible [6], ha fracasado en alcanzar dichas expectativas [7], en parte por la inexistencia de medidas objetivas fiables para los síntomas (falta de aire, tos, jadeos). Además, las consultas remotas se beneficiarían enormemente si el médico contara con mediciones objetivas del número de veces que tose un paciente, la naturaleza de esa tos, y si el síntoma es indicativo de alguna enfermedad concreta. Esto llevaría a una menor necesidad de visitar a los pacientes en su domicilio, a un diagnóstico más temprano y a un seguimiento en tiempo real de las personas con enfermedad respiratoria en situación de dependencia.

En los últimos años, ha habido una explosión del uso de métodos basados en aprendizaje profundo o *deep learning* para resolver el problema de la detección de tos y su análisis con fines diagnósticos [8, 9, 10]. Los métodos propuestos hasta la fecha presentan en general buenos resultados, tanto a nivel de detección [8, 9], como de diagnóstico de patologías [9, 10]. Sin embargo, pese a lograr mejores resultados que otros métodos basados en aprendizaje máquina convencional, estos métodos actúan como cajas negras y no permiten interpretar si la tos como síntoma presenta alguna característica especial que se asocie a una determinada enfermedad o estadio de la misma.

En este trabajo se pretende estudiar la viabilidad de la inteligencia artificial (IA) explicable para interpretar patrones específicos en síntomas (tos) de la enfermedad respiratoria crónica. Para ello se llevará a cabo un modelado paramétrico de los mapas de oclusión que identifican regiones de activación de los patrones tiempo-frecuencia en eventos de tos al ser procesados por una red convolucional entrenada para detectar dichos eventos en señales de audio. Los parámetros del modelo se analizarán para identificar diferencias significativas entre enfermedades y su interpretación. Hasta donde sabemos, este tipo de análisis basado en IA explicable no se ha aplicado al estudio de la enfermedad respiratoria.

2. Materiales

Se llevó a cabo un estudio observacional de la evolución de la tos durante 24 horas de la vida normal de un paciente. Se grabaron prospectivamente 24 horas de audio de pacientes ambulatorios en el Área de Salud de Palencia. La base de datos generada consta de aproximadamente 15.000 eventos de tos correspondientes a 20 pacientes entre 23 y 87 años (9 mujeres, 11 hombres) con las siguientes patologías respiratorias: Enfermedad respiratoria aguda (general, 3), neumonía (4), EPOC (6), cáncer de pulmón (3), y otras como asma, bronquiectasias o sarcoidosis (resto de pacientes). Se utilizó un teléfono inteligente Android Sony Xperia Z2 para recopilar los datos, almacenando los archivos en formato WAV de 16 bits a 44,1 kHz.

Sobre la base de datos se definieron los siguientes grupos de pacientes:

- G1: Pacientes crónicos frente no crónicos.
- G2: Pacientes con EPOC frente resto de enfermedades.
- G3: Pacientes con EPOC frente resto de enfermedades excluyendo cáncer de pulmón.
- G4: Pacientes con EPOC frente enfermedad respiratoria aguda y neumonía.
- G5: Pacientes con EPOC frente enfermedades crónicas que no son EPOC.
- G6: Pacientes con EPOC frente pacientes con cáncer de pulmón.

3. Métodos

3.1. Preprocesamiento de las señales de audio

Tras submuestrear la señal de 44,1 KHz con un factor de diezmado de 5 para una frecuencia final de 8,820 KHz se definen ventanas en tiempo corto de 10 ms (89 muestras) cuya densidad espectral de potencia (DEP) se calcula utilizando una ventana de Hanning sin solapamiento con el mismo número de muestras. Dichas DEP se concatenan a lo largo de ventanas de tiempo largo de 1 s para dar lugar a espectrogramas de tamaño 45×100 . Estas representaciones tiempo-frecuencia se utilizarán como imágenes de entrada a las redes neuronales convolucionales tras calcularse su logaritmo y normalizarlas.

3.2. Identificación de las ventanas de tos

Para identificar los espectrogramas de audio correspondientes a eventos de tos se utilizó una red neuronal convolucional (CNN) [11] entrenada sobre la base de datos.

Para la CNN se optó por una arquitectura simple consistente en una capa convolucional de 32 filtros, *kernels*

2×2 y función de activación ReLU. Seguidamente se incorpora una capa de *Max-Pool*, con un tamaño de *pool* de 2×2 . Esta capa reduce la dimensionalidad de los datos para no tener una excesiva carga computacional. A continuación, se conectó una capa de dropout, con el objetivo de reducir el sobreajuste. Para ello, durante el entrenamiento, cierto número de salidas de esta capa se ignoran aleatoriamente o se “eliminan”. En nuestro caso ignoramos aleatoriamente el 10% de las salidas. Esta arquitectura se repite una vez más, tal y como se muestra en la figura 1, doblando en este caso el número de filtros (64). A continuación se intercalan una capa convolucional con 128 filtros, otra de *dropout*, y otra convolucional con 256 filtros, seguida de una *MaxPooling2D*. La salida de ésta se redimensiona (capa *flatten*) para alimentar dos capas *fully connected*, una de 512 neuronas con activación ReLU, y otra con dos neuronas con activación *softmax* que da lugar a la clasificación final.

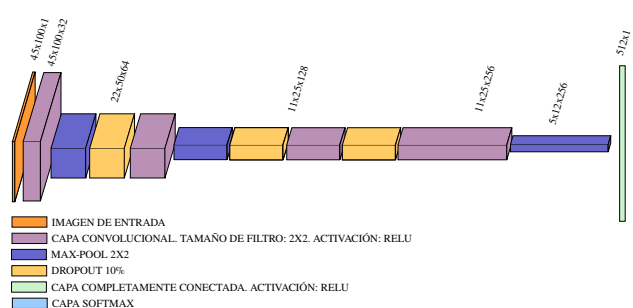


Figura 1. Arquitectura de la red convolucional utilizada.

El optimizador utilizado para el entrenamiento fue AdaMax con parámetros $\alpha = 0,002$, $\beta_1 = 0,9$, y $\beta_2 = 0,999$ [12]. El número de *epochs* utilizado fue 50 y el tamaño de *batch* 128. Para la validación de hiperparámetros y decisión sobre la finalización del entrenamiento se empleó el 20% del conjunto de entrenamiento, formado por el 70% de los clips de audio.

3.3. Interpretabilidad de la CNN. Análisis cuantitativo de mapas de oclusión

El análisis de interpretabilidad está basado en la reciente propuesta de IA explicable para el diagnóstico de TDAH [13]. Para cada paciente, se tomaron aquellos espectrogramas para los que la salida *softmax* de la red proporcionara una clasificación como Tos con un nivel de confianza superior al 90%. Con estos espectrogramas se obtuvo un mapa de calor que permitía visualizar las zonas más relevantes del espectrograma para la identificación como tos. La técnica empleada para la obtención de este mapa se conoce como mapa de oclusión [14]. Esta técnica permite obtener el mapa definiendo una máscara que se coloca sobre la imagen de entrada y, en consecuencia, oculta parte de la información de la imagen a la CNN. La probabilidad de clase de la CNN alimentada por esta entrada distorsionada será una estimación de la importancia del área oculta para la clasificación (mayor

probabilidad significa menor importancia). La máscara se desliza, se repite el proceso y los valores de salida se almacenan en una matriz que se redimensiona para que tenga las mismas dimensiones que la imagen original.

La obtención de los mapas de oclusión de cada paciente y su posterior análisis cuantitativo se realizó como sigue:

Mapas de oclusión Para cada uno de los espectrogramas de tos seleccionados, se obtuvo su mapa de oclusión. Posteriormente, se normalizaron todos los mapas para cada paciente y se promediaron por píxeles de forma que cada sujeto tiene asociado un único mapa de oclusión que caracterice su tos.

Modelado paramétrico mediante mezcla de gaussianas y análisis por contraste de hipótesis

Para cada mapa de paciente se ajustó un modelo de mezcla de gaussianas (GMM) para el análisis cuantitativo. Definimos $m(x, y)$ como el mapa de oclusión en el punto (x, y) del dominio $\chi \subset \mathbb{Z}^2$ de la rejilla 2D y $f(x, y; \theta_i)$ la función de densidad gaussiana de dos dimensiones con los parámetros θ_i . Dado que los mapas presentaban dos regiones de activación claramente diferenciadas, se optó por un modelo de 2 gaussianas. El ajuste se obtiene al aplicar la siguiente expresión:

$$\min_{p_1, p_2, \theta_1, \theta_2} \sum_{(x,y) \in \chi} \left(m(x, y) - \sum_{i=1}^2 p_i f(x, y; \theta_i) \right)^2 \quad (1)$$

siendo

$$\theta_i = [\eta_{x_i} \quad \sigma_{x_i} \quad \eta_{y_i} \quad \sigma_{y_i} \quad \rho_{x_i y_i}], \quad (2)$$

el vector de parámetros de las gaussianas, donde η se refiere a las medias, σ a las desviaciones estándar y ρ el coeficiente de correlación de Pearson de las dos componentes. Los subíndices indican la dirección de los parámetros, siendo x el eje horizontal e y el eje vertical.

El objetivo de este ajuste es evaluar la presencia de diferencias estadísticamente significativas entre los grupos de enfermedades definidos en la sección 2. Para ello, se siguió el siguiente procedimiento:

- Primero, se ordenaron las gaussianas de la mezcla de cada sujeto en función del valor del parámetro p_i en la ecuación 1. La gaussiana con menor peso será definida como no dominante.
- Se compararon las muestras grupales de parámetros entre sí para encontrar diferencias estadísticas entre ellas. Concretamente, primero se realizó una prueba de Gaussianidad en cada muestra; si no pudo descartarse la gaussianidad en ambas, se utilizó una prueba t de Student no apareada. En caso contrario, se empleó la prueba U de *Mann Whitney*.
- Una vez realizada la prueba correspondiente para todos los parámetros de la ecuación 2, se obtuvieron los *boxplot* correspondientes para los casos en que se detectaron diferencias significativas

(p – valor inferior a 0,05).

4. Resultados y discusión

Las tablas 1 y 2 muestran los p-valores obtenidos para cada uno de los parámetros de la mezcla de gaussianas. Los resultados se muestran para cada uno de los grupos de estudio descritos en la sección 2. Los parámetros analizados corresponden respectivamente a las gaussianas dominante (tabla 1) y no dominante (tabla 2), entendiéndose por dominante la de mayor p – valor.

	G1	G2	G3	G4	G5	G6
p	0.141	0.298	0.450	0.228	0.881	0.297
η_x	0.743	0.717	0.767	0.919	0.640	0.716
η_y	0.568	0.436	0.659	0.483	0.917	0.219
σ_x	0.784	0.691	0.865	0.847	0.974	0.206
σ_y	0.505	0.090	0.094	0.177	0.450	0.663
ρ_{xy}	0.345	0.030	0.030	0.014	0.386	0.151

Tabla 1. Resultados obtenidos para los p-valores de la Gaussiana dominante en todos los grupos de estudio.

	G1	G2	G3	G4	G5	G6
p	0.141	0.298	0.450	0.228	0.881	0.297
η_x	0.338	0.467	0.361	0.280	0.919	0.572
η_y	0.771	0.012	0.026	0.079	0.107	0.148
σ_x	0.526	0.522	0.565	0.856	0.969	0.360
σ_y	0.303	0.039	0.038	0.143	0.218	0.705
ρ_{xy}	0.950	0.248	0.213	0.380	0.165	0.855

Tabla 2. Resultados obtenidos para los p-valores de la Gaussiana no dominante en todos los grupos de estudio.

En el caso de la gaussiana no dominante existen diferencias significativas en la media y desviación típica para el grupo de pacientes EPOC cuando se compara con el resto de enfermedades (G2) y el resto sin incluir cáncer de pulmón (G3). Si observamos los boxplots de la figura 2 (parte derecha) se observa en efecto que la gaussiana no dominante tiene mayores valores de η_y y menores de σ_y en el grupo de pacientes con EPOC. Esto se refleja en los mapas de oclusión medianos para cada grupo que se presentan en la parte izquierda de la figura 2. En el caso de los pacientes con EPOC, la gaussiana no dominante, esto es, aquella con más intensidad en el mapa, es la que se sitúa en frecuencias más altas. Por otra parte, para los pacientes con otras enfermedades, dicha gaussiana está situada en frecuencias más bajas. En el caso de la dispersión, al tener diferentes alturas las gaussianas no dominantes en los distintos grupos, no es posible apreciar visualmente diferencias.

Para la gaussiana dominante, el coeficiente de correlación resulta significativamente diferente en las comparativas G2 (EPOC vs. resto de pacientes), G3 (EPOC vs. resto de pacientes salvo cáncer de pulmón) y G4 (EPOC vs. enfermedades no crónicas, respiratoria aguda y neumonías). En todos los casos, el coeficiente de correlación para el EPOC resulta superior (ver boxplots en figura 2, parte derecha), llegando incluso a valores negativos para los otros grupos. Esto se aprecia en las orientaciones diferentes de la huella de la gaussiana dominante para los otros grupos.

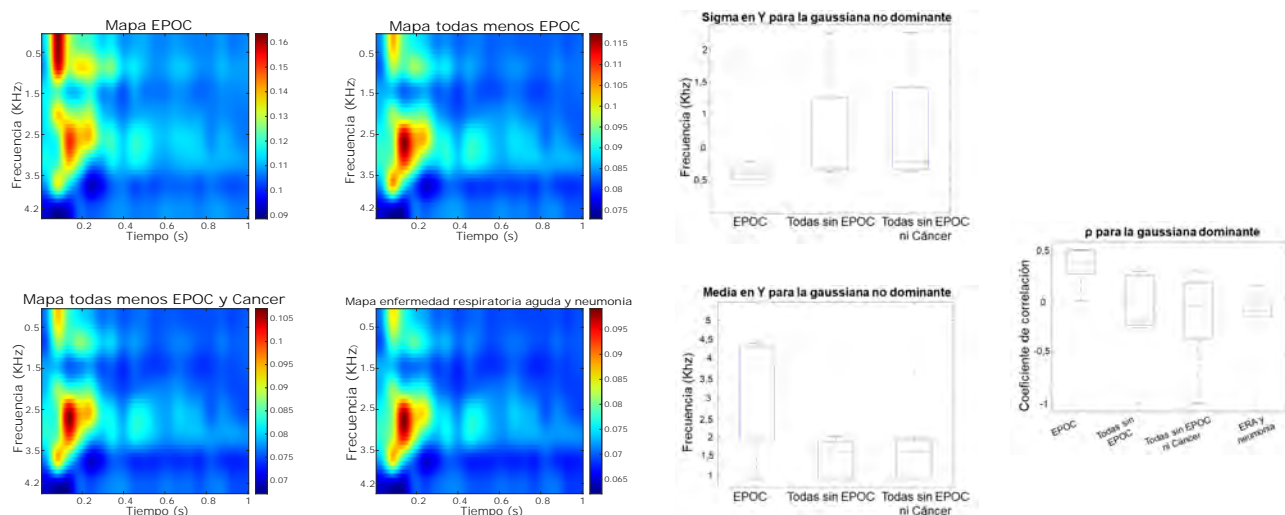


Figura 2. Mapas de oclusión medianos y boxplots de los parámetros significativos para los grupos que conforman G_2 , G_3 y G_4 .

5. Conclusiones

Este trabajo ha demostrado la viabilidad del uso de métodos de análisis de interpretabilidad de redes neuronales para explicar diferencias asociadas a la tos en diferentes grupos de pacientes. En concreto, se han encontrado diferencias entre pacientes crónicos con EPOC y pacientes con otras patologías, tanto crónicas como no crónicas. Los métodos utilizados permiten identificar qué zonas del patrón tiempo-frecuencia de un clip de audio de tos resultan de mayor relevancia para la red neuronal en cada una de las enfermedades, y pueden servir para identificar patrones específicos en las diferentes patologías.

El estudio llevado a cabo es de naturaleza preliminar, dado que los grupos de pacientes estudiados tienen un tamaño reducido para algunas de las enfermedades. En todo caso, las técnicas aplicadas resultan prometedoras y se espera que tras la ampliación de los grupos de pacientes, puedan dar lugar a más resultados significativos que permitan explicar mejor la sintomatología de las diferentes patologías respiratorias.

Agradecimientos

Este trabajo es parte del proyecto TED2021-131536B-I00, financiado por MCIN/AEI/10.13039/501100011033 y por la Unión Europea “NextGenerationEU”/PRTR.

Referencias

- [1] World Health Organisation. *The Global Impact of Respiratory Disease*. European Respiratory Society, 2017.
- [2] American Thoracic Society. Statement on home care for patients with respiratory disorders. *American Journal of Respiratory and Critical Care Medicine*, 171(12):1443–1464, 2005.
- [3] S. Belli *et al.* Low physical functioning and impaired performance of activities of daily life in COVID-19 patients who survived hospitalisation. *European Respiratory Journal*, 56(4), 2020.
- [4] World Health Organisation. WHO Coronavirus (COVID-19) Dashboard. <https://covid19.who.int/>, 2021.
- [5] European Commission. Market study on telemedicine. Technical report, European Union, 2018.
- [6] Audit Scotland. A review of Telehealth in Scotland. Technical report, Audit Scotland, 2011.
- [7] H. Pinnock *et al.* Effectiveness of telemonitoring integrated into existing clinical services on hospital admission for exacerbation of chronic obstructive pulmonary disease: researcher blind, multicentre, randomised controlled trial. *BMJ*, 347, 2013.
- [8] M. You *et al.* Automatic cough detection from realistic audio recordings using C-BiLSTM with boundary regression. *Biomedical Signal Processing and Control*, 72:103304, 2022.
- [9] A. Tena, F. Clarià, and F. Solsona. Automated detection of COVID-19 cough. *Biomedical Signal Processing and Control*, 71:103175, 2022.
- [10] J. Laguarda, F. Huetto, and B. Subirana. COVID-19 artificial intelligence diagnosis using only cough recordings. *IEEE Open Journal of Engineering in Medicine and Biology*, 1:275–281, 2020.
- [11] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.
- [12] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In Y. Bengio and Y. LeCun, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.
- [13] P. Amado-Caballero *et al.* Insight into ADHD diagnosis with deep learning on actimetry: Quantitative interpretation of occlusion maps in age and gender subgroups. *Artificial Intelligence in Medicine*, 143:102630, 2023.
- [14] M. D. Zeiler and R. Fergus. Visualizing and understanding convolutional networks. In *European conference on computer vision*, pages 818–833. Springer, 2014.

Uso de la actividad electrodermal (EDA) para la clasificación de situaciones de estrés en cirugía robótica de mínima invasión

D. Caballero¹, J.A. Sánchez-Margallo¹, M.J. Pérez-Salazar¹, M. Kappel^{1,2}, F.M. Sánchez-Margallo¹

¹ Centro de cirugía de mínima invasión Jesús Usón, Cáceres, España {dcaballero, jasanchezm, mjperez, msanchez}@ccmijesususon.com

² Escuela de Ingeniería Superior, Universidad Sigma-Clermont, Aubière, Francia, mathilde.kappel@sigma-clermont.fr

Resumen

La detección de altos niveles de estrés anticipadamente puede ayudar a prevenir problemas de salud. Además, la carencia de nuevas tecnologías para evaluar el nivel de estrés del cirujano en actividades quirúrgicas mínimamente invasivas hace necesario desarrollar nuevos métodos para solucionar esta problemática. Por consiguiente, el objetivo de este estudio es clasificar diferentes situaciones producidas durante varias sesiones quirúrgicas para determinar si producen altos niveles de estrés en los cirujanos aplicando para ello técnicas de inteligencia artificial. En este estudio, se recopilaron datos durante 26 sesiones de cirugía robótica extrayendo datos de 11 cirujanos, obtenidos de 4 sensores diferentes: actividad electrodermal (EDA), temperatura, movimiento corporal y presión arterial. Una vez generado el conjunto de datos, se transformó en dos subconjuntos con diferentes tipos de preprocesado, escalado y normalizado. Sobre ambos conjuntos de datos se aplicaron diferentes configuraciones del algoritmo de clustering, para agrupar los datos en 2, 3 y 5 grupos en función de los diferentes niveles de estrés del cirujano. Entre los resultados obtenidos, cabe destacar que los niveles más altos de estrés del cirujano están relacionados con valores altos de EDA, temperatura y presión arterial; y los niveles más bajos con valores bajos de estas mismas características. Se puede extraer también que dividir en un mayor número de grupos los niveles de estrés del cirujano define de manera más precisa los mismos. Estos resultados fueron validados utilizando la grabación de video de las sesiones quirúrgicas y evaluando si las situaciones fueron anómalas, generando estrés al cirujano.

1. Introducción

La detección de altos niveles de estrés de forma anticipada puede ayudar a prevenir numerosos problemas de salud. Por tanto, se hace urgente encontrar la aplicación de nuevas tecnologías que permita controlar y monitorizar la salud física y mental. Esto se hace especialmente esencial si extrapolamos estos niveles de estrés a las actividades quirúrgicas asistidas por plataformas robótica, debido a sus implicaciones en la calidad asistencial de los pacientes. Afortunadamente, se han llevado a cabo algunos avances a la hora de medir y monitorizar los niveles de estrés tanto en actividades cotidianas [1] como en actividades de la cirugía robótica mínimamente invasiva [2].

Las técnicas quirúrgicas mínimamente invasivas se han convertido en las técnicas quirúrgicas estándar para numerosos procedimientos. Las ventajas de la cirugía de mínima invasión han sido claramente expuestas en la literatura científica y son sobradamente conocidas. Entre estos beneficios se encuentran la disminución del dolor

postoperatorio, la reducción del consumo de analgésicos o las mejoras en los tiempos de recuperación. No obstante, también tiene ciertas desventajas para los cirujanos como son las deficiencias ergonómicas o los niveles de estrés, siendo potencialmente peligrosos para la salud del cirujano y la calidad de su desempeño quirúrgico [3].

Es por ello que la carencia de nuevas tecnologías aplicadas con el objetivo de interpretar los estados emocionales durante las actividades relacionadas con la cirugía robótica de mínima invasión, hace necesario desarrollar nuevos métodos con el objetivo de comprender, gestionar, monitorizar y dar respuesta a la relación entre el nivel de estrés del cirujano y las situaciones que se plantean en una cirugía robótica.

En relación con la medida y monitorización del estrés del cirujano durante la cirugía robótica, se ha analizado en la literatura científica un amplio número de posibles factores que podrían ser utilizadas como índices [4]. Entre los dispositivos para medir estos factores se encuentran los sensores de actividad electrodermal (EDA). Así, la utilización de EDA es óptima para evaluar el nivel de estrés del cirujano, ya que puede cuantificar cambios en el sistema nervioso del sujeto. Para medir la señal generada con los sistemas EDA, existen sensores portátiles más ligeros y con conexión inalámbrica que permiten realizar la medición de la señal de manera continua en situaciones reales de movilidad, proporcionando información detallada del cirujano que esté realizando la cirugía.

Adicionalmente, en los últimos años las aplicaciones de inteligencia artificial han visto aumentado exponencialmente su uso, permitiendo el desarrollo de diversos algoritmos lo suficientemente entrenados para ayudar en la toma de decisiones [5]. Entre las diferentes técnicas de inteligencia artificial, las técnicas de clasificación no supervisada permiten evaluar y clasificar las múltiples tuplas de datos y, por lo tanto, su aplicación a la hora de evaluar las posibles situaciones de estrés. Es por ello, que esta técnica podría ser la más adecuada para su aplicación en actividades de cirugía robótica.

Por consiguiente, el objetivo del presente estudio es clasificar diferentes situaciones producidas durante cirugías robóticas mínimamente invasivas para determinar los niveles de estrés producidos al cirujano, utilizando para ello técnicas de inteligencia artificial. De manera que, se puedan determinar de forma automática situaciones de estrés del cirujano e indicar posibles advertencias o descansos durante las cirugías, con el fin de mejorar la salud de los cirujanos.

2. Material y métodos

2.1 Diseño experimental

Se han recopilado datos durante 26 sesiones de cirugía robótica completadas por 11 cirujanos, siendo 6 cirujanos con un nivel de experiencia experto y 5 cirujanos con un nivel de experiencia novel, siendo 2 mujeres y 9 hombres de entre 25 y 50 años de edad, las sesiones de cirugía se han grabado con el objetivo de poder validar los resultados generados por el estudio. Todos los protocolos experimentales de este estudio han sido aprobados por el correspondiente comité de Experimentación Animal local (Ref. EXP-20230325, EXP-20230320 y 20230313). Los datos recopilados son los referentes a un sensor EDA, un sensor de temperatura, un sensor de presión arterial y un sensor de movimiento corporal del cirujano. Estos datos se recopilaron en un único dispositivo y se extrajeron usando el software *Unisens Viewer*, de *Movisens* (Karlsruhe, Alemania). A partir de estos datos, se generó un conjunto de datos y sobre este, se aplicaron diferentes técnicas de preprocesado, transformándolo en dos conjuntos de datos. Sobre estos datos se aplicaron técnicas de clustering, con el objetivo de clasificarlos entre situaciones con diferentes niveles de estrés del cirujano. Finalmente, se ha validado las situaciones clasificadas como de alto nivel de estrés del cirujano con los vídeos registrados de los procedimientos quirúrgicos robóticos. La Figura 1 muestra el diseño experimental seguido en el estudio.

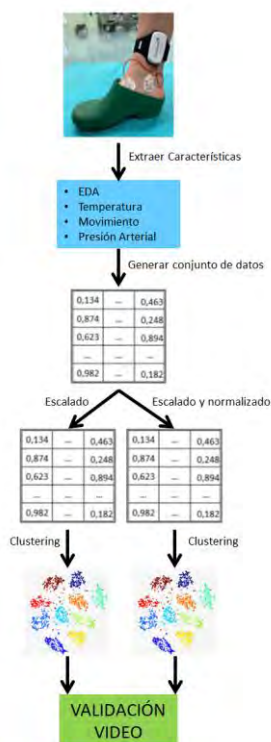


Figura 1. Diseño experimental

2.2. Robótica y Sensorización

Para la realización de las 26 sesiones de cirugía se utilizaron 2 plataformas robóticas para cirugía de mínima invasión: La plataforma *SYMANI* (*MMI*, Pisa, Italia) (<http://mmimicro.com/our-technology/how-it-works/>)

para cirugía robótica microquirúrgica y la plataforma *VERSIUS* (*CMR Surgical Ltd*, Cambridge, Reino Unido) (<http://es.cmrsurgical.com>) para cirugía robótica laparoscópica. Se utilizó en 14 sesiones la plataforma *SYMANI* y en 12 sesiones la plataforma *VERSIUS*.

Los sensores utilizados para la recopilación de los datos durante las cirugías robóticas fueron el sensor EDA, el de temperatura, el de movimiento corporal y el de presión arterial. Todos estos sensores están ubicados en un único dispositivo que es una pulsera que se ajusta en el tobillo (Figura 1), dicho dispositivo es el *EdaMove 4* (*Movisens*, Karlsruhe, Alemania) (<http://movisens.com/en/products/eda-and-activity-sensor/>).

2.3. Conjuntos de datos

A partir de todos los datos recopilados por el dispositivo *EdaMove4* se generó un conjunto de datos constituido por 2770 tuplas, que contienen los valores de EDA, la temperatura corporal, el movimiento y la presión arterial del cirujano. Estas tuplas se utilizaron para realizar la clasificación no supervisada. Para el desarrollo del presente estudio, se utilizaron dos conjuntos de datos, transformaciones del conjunto de datos original, generados después de aplicar diferentes técnicas de preprocesado, el de los datos escalados y el de los datos escalados y normalizados.

2.4. Técnicas de inteligencia artificial

Respecto a las técnicas de inteligencia artificial, como técnica de clasificación no supervisada, se aplicó en el presente estudio el clustering o agrupamiento. El clustering consiste en agrupar las tuplas de datos en varios clusters o grupos no solapados, en el cual las tuplas de datos se agrupan a la menor distancia euclídea posible de los centroides de las diferentes características [6].

En este estudio se ha aplicado como algoritmo de clustering la implementación de expectación máxima simple [7], usando para ello el algoritmo disponible en la herramienta suite de inteligencia artificial *Weka 3.8.6* (*Universidad de Waikato*, Hamilton, Nueva Zelanda). Este algoritmo se aplicó sobre los datos de EDA, temperatura corporal, movimiento y presión arterial del cirujano. Se aplicaron tres configuraciones diferentes, una de ellas clasificando solo en 2 grupos (nivel alto y bajo de estrés del cirujano), otra en 3 grupos (con nivel alto, medio y bajo de estrés del cirujano) y otro en 5 grupos (con diferentes niveles de estrés del cirujano). Una vez se obtuvieron los resultados, estos fueron validados comprobando la grabación del vídeo de la intervención.

Previo a la aplicación del clustering, se han aplicado dos tipos diferentes de preprocesado como son el escalado y la normalización, dando lugar a dos conjuntos de datos distintos. El escalado permite que todos los datos para cada una de las características se escalen a valores que se encuentran entre 0 y 1 (ecuación 1). La normalización permite que los datos previamente escalados puedan ser transformados a otro conjunto de datos con mayor integridad y una menor redundancia de los datos (ecuación 2).

$$Dato_{escalado} = \frac{Dato_{inicial} - valor_{mínimo}}{valor_{máximo} - valor_{mínimo}} \quad (1)$$

$$Dato_{normalizado} = \frac{Dato_{escalado} - valor_{medio}}{valor_{desviación\ típica}} \quad (2)$$

3. Resultados y discusión

Con el objetivo de evaluar y clasificar las posibles situaciones de estrés del cirujano durante las sesiones de cirugía robótica, se comenzó el estudio evaluando las diferentes características y se consiguió generar los dos conjuntos de datos a partir del conjunto de datos original, uno de ellos con los datos escalados para cada uno de las 4 características evaluadas (EDA, temperatura, movimiento y presión arterial) y otro conjunto de datos con los datos escalados y normalizados para las mismas características.

Una vez se generaron ambos conjuntos de datos, se procedió a aplicar el algoritmo de clustering. Así, para la primera configuración, clasificando en dos grupos (nivel alto y bajo de estrés del cirujano) con ambos tipos de preprocesado, escalado (ESC) y escalado y normalizado (ESC NORM), se obtuvieron los resultados mostrados en la Tabla 1.

	Grupo 0 ESC	Grupo 1 ESC	Grupo 0 ESC NORM	Grupo 1 ESC NORM
Nivel estrés	Bajo	Alto	Alto	Bajo
Centro EDA	0,1309	0,3127	1,0272	-0,9880
Centro Tª	0,2554	0,4966	1,0506	-1,0251
Centro Mov.	0,0023	0,2825	1,5205	-0,3378
Centro Presión Arterial	0,3822	0,9015	0,9584	-0,0801
% Tuplas en el grupo	67%	33%	33%	67%

Tabla 1. Resultados obtenidos para la configuración de dos grupos y para el conjunto de datos ESC (con escalado) y ESC NORM (con escalado y normalizado).

De la Tabla 1 se puede destacar que el grupo con nivel bajo de estrés del cirujano está en el 67% y que el grupo con nivel alto de estrés del cirujano en el 33%, independientemente de si se aplica una técnica de preprocesado u otra. También, cabe destacar que para cada una de las características los centroides son más bajos para el grupo con nivel bajo de estrés del cirujano que con nivel alto de estrés del cirujano, lo cual muestra un buen funcionamiento de la inteligencia artificial, ya que los valores bajos de EDA, temperatura y presión arterial estarían relacionados con bajos niveles de estrés del cirujano y los valores más altos con niveles más altos de estrés de los cirujanos [8,9].

En el caso de aplicar la configuración para clasificar en 3 grupos (nivel de estrés alto, medio y bajo) se obtuvieron los resultados mostrados en la Tabla 2 para el conjunto de datos escalado (ESC) y en la Tabla 3 para el conjunto de datos escalado y normalizado (ESC NORM).

Al igual que ocurrió en el caso de la configuración de dos grupos, los niveles más altos de estrés de los cirujanos están caracterizados por un valor alto de EDA, de presión arterial y de temperatura [8,9], independientemente del tipo de preprocesado aplicado.

	Grupo 0 ESC	Grupo 1 ESC	Grupo 2 ESC
Nivel estrés	Alto	Bajo	Medio
Centro EDA	0,3679	0,0003	0,1531
Centro Tª	0,4381	0,2629	0,3916
Centro Mov.	0,2829	0,0016	0,0766
Centro Presión Arterial	0,9836	0,3495	0,5861
% Tuplas en el grupo	18%	53%	29%

Tabla 2. Resultados obtenidos para la configuración de tres grupos y para el conjunto de datos ESC (con escalado).

Sin embargo, destacaremos que tanto los porcentajes en los niveles altos y bajos de estrés de los cirujanos han descendido y se han incluido en el grupo medio de estrés de los cirujanos, lo cual demuestra que es necesario un mayor número de grupos para una mejor definición de los niveles de estrés de los cirujanos en situaciones de cirugía robótica.

	Grupo 0 ESC NORM	Grupo 1 ESC NORM	Grupo 2 ESC NORM
Nivel estrés	Medio	Bajo	Alto
Centro EDA	0,1144	-0,7885	1,0525
Centro Tª	0,2206	-0,0128	1,2042
Centro Mov.	0,3337	-0,3444	1,7237
Centro Presión Arterial	0,6357	-0,9179	1,1748
% Tuplas en el grupo	27%	53%	20%

Tabla 3. Resultados obtenidos para la configuración de tres grupos y para el conjunto de datos ESC NORM (con escalado y normalizado).

En el caso de aplicar la configuración para clasificar en 5 grupos (nivel de estrés del cirujano muy alto, alto, medio, bajo y muy bajo), se obtuvieron los resultados mostrados en la Tabla 4 para el conjunto de datos escalado (ESC) y en la Tabla 5 para el conjunto de datos escalado y normalizado (ESC NOR). Al igual que ocurrió en los resultados anteriores, los niveles más altos de estrés del cirujano están caracterizados por un valor alto de EDA, de presión arterial y de temperatura [8,9]. Sin embargo, cabe destacar que tanto los porcentajes en los niveles altos medios y bajos de estrés de los cirujanos han descendido y se han distribuido a lo largo de los diferentes grupos, lo cual demuestra que con un mayor número de grupos se obtiene una clasificación más precisa de los diferentes niveles de estrés analizados.

	Grupo 0 ESC	Grupo 1 ESC	Grupo 2 ESC	Grupo 3 ESC	Grupo 4 ESC
Nivel estrés	Medio	Muy bajo	Alto	Bajo	Muy alto
Centro EDA	0,1568	0,0003	0,3771	0,0785	0,7755
Centro T ^a	0,3937	0,2593	0,4506	0,2831	0,6221
Centro Mov.	0,0215	0,0001	0,2671	0,0093	0,4001
Centro Presión Arterial	0,6451	0,2367	0,7810	0,3485	0,9919
% Tuplas en el grupo	23%	32%	17%	24%	4%

Tabla 4. Resultados obtenidos para la configuración de cinco grupos y para el conjunto de datos ESC (con escalado).

Todos los resultados obtenidos se han validado y verificado con las grabaciones de los vídeos de las diferentes cirugías robóticas llevadas a cabo, confirmando que efectivamente existe una relación entre un nivel alto o muy alto de estrés del cirujano y situaciones complejas o anómalas durante las sesiones quirúrgicas.

	Grupo 0 ESC NOR	Grupo 1 ESC NOR	Grupo 2 ESC NOR	Grupo 3 ESC NOR	Grupo 4 ESC NOR
Nivel estrés	Alto	Medio	Muy bajo	Bajo	Muy alto
Centro EDA	2,4404	0,0991	-0,7885	-0,3178	0,471
Centro T ^a	0,2958	0,0212	-1,0559	-0,6703	1,1661
Centro Mov.	0,1615	-0,2047	-0,6021	-0,2817	1,3863
Centro Presión Arterial	0,7447	0,1522	-1,7106	-0,6511	0,8908
% Tuplas en el grupo	17%	25%	32%	22%	4%

Tabla 5. Resultados obtenidos para la configuración de cinco grupos y para el conjunto de datos ESC NOR (con escalado y normalizado).

4. Conclusión

Del presente estudio se puede concluir que se ha conseguido clasificar en diferentes niveles de estrés de cirujanos, situaciones producidas durante varias sesiones de cirugía robótica en función de diferentes parámetros fisiológicos como son EDA, temperatura, movimiento corporal y presión arterial del cirujano. También se puede concluir que dividir en un mayor número de grupos los

niveles de estrés de los cirujanos dota de mayor precisión a la definición de los mismos en cada situación. Por último, cabe destacar que valores más altos de EDA, temperatura y presión arterial generan un nivel más elevado de estrés del cirujano.

Agradecimientos

Este estudio ha sido cofinanciado por la Junta de Extremadura (TA18023) y el Plan Complementario Biotecnología Aplicada a la Salud, cofinanciado por el Ministerio de Ciencia e Innovación con fondos de la Unión Europea NextGenerationEU, el Plan de Recuperación, Transformación y Resiliencia (PRTR-C17.I1) y el Programa Operativo FEDER Extremadura 2021-2027. Sin embargo, los puntos de vista y las opiniones expresadas son únicamente los de los autores. Agradecer también la realización del presente estudio a los compañeros que colaboraron en las sesiones de cirugía robótica (Elena Abellán, Laura Pires, Javier Vela, José Luis Campos, Carlos Sánchez, David Durán, José Ramón Torres, Rubén Baltá y Pedro Olivares).

Referencias

- [1] Rebolledo-Mendez G, Reyes A, Paszkowicz S, Domingo MC, Skrypchuk L. Developing a body sensor network to detect emotions during driving. *IEEE transactions on Intelligent Transportation Systems*, vol 15, 2014, pp. 1850-4. (ISSN: 1524-9050).
- [2] Hurley AM, Kennedy PJ, O'Connor L, Dinan TG, Cryan JF, Boylan, G. O'Reilly, BA. SOS save our surgeons: Stress level reduced by robotic surgery. *Gynecological surgery*, vol. 12, 2015, pp. 197-206. (ISSN: 1613-2084).
- [3] Ballantyne GH. The pitfalls of laparoscopic surgery, telepresence, and telemonitoring. Review of early clinical results. *Surgical endoscopy*, vol 16, 2002, pp. 1389-402. (ISSN: 0930-2794).
- [4] Healey JA, Picard RW. Detecting stress during real-world driving tasks using physiological sensors. *IEEE transactions on Intelligent Transportation Systems*, vol 6, 2005, pp. 156-66. (ISSN: 1524-9050).
- [5] Ávila-Tomás JF, Mayer-Pujadas MA, Quesada-Varela VJ. La inteligencia artificial y sus aplicaciones en medicina I: introducción antecedentes a la IA y robótica. *Atención primaria*, vol 52, 2020, pp. 778-84. (ISSN: 0212-6567).
- [6] Forgy EW, Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics*, vol. 21(3), 1965, pp. 768-69. (ISSN: 0006-341X).
- [7] Einicke GA, Falco G, Malos JT. EM algorithm state matrix estimation for navigation. *IEEE Signal Processing Letters* vol 17(5), 2012, pp. 437-40. (ISSN: 1070-9908).
- [8] Kim HG, Cheon EJ, Bai DS, Lee YH, Koo BH. Stress and heart rate variability: A meta-analysis and review of the literature. *Psychiatry Investigarion*, vol 15(3), 2018, pp.235-45. (ISSN: 1738-3684).
- [9] Zangróniz R, Martínez-Rodrigo A, Pastor JM, López MT, Fernández-Caballero A. Electrodermal Activity Sensor for classification of calm/distress condition. *Sensors*, vol 17(2324), 2017, pp. 1-14. (ISSN: 1424-8220).

Biomateriales

Viernes 24 8:30 - 10:00

Paraninfo

Axonal guidance using biofunctionalized silk fibroin fibers manufactured using the SFS technique

C. Castro Domínguez¹, F. Panetsos¹, G.V. Guinea², M. Ramos Gómez², J. Pérez Rigueiro²

¹ Neuro-Computing and Neuro-Robotics Research Group, Universidad Complutense de Madrid, Madrid, Spain, ccastro04@ucm.es; fivos@ucm.es

² Center for Biomedical Technology (CTB), Universidad Politécnica de Madrid, 28223 Madrid, Spain, gustavovictor.guinea@ctb.upm.es; milagros.ramos@ctb.upm.es; jose.perez@ctb.upm.es

Abstract

After an injury, the central nervous system's limited regenerative capacity severely hampers the reconnection and functional recovery of affected nervous tissue, making it an arduous task. To address this critical issue, biomaterials have emerged as a promising solution for designing scaffolds that facilitate and guide the regenerative process. Leveraging prior research on regenerated silk fibroin fibers produced via the straining flow spinning (SFS) technique, this study aims to demonstrate that biofunctionalized SFS fibers offer superior guidance capabilities compared to non-functionalized fibers. The study reveals that neurons' axons exhibit a remarkable tendency to align with the fibers' paths, in contrast to the isotropic growth observed on conventional culture plates. Additionally, the guidance ability of these fibers can be further enhanced through the biofunctionalization of the material with adhesion peptides. Proving the exceptional guidance potential of these fibers opens up exciting possibilities for their application as implants in spinal cord injuries. They could serve as a core component of a therapeutic approach that facilitates the reconnection of injured spinal cord ends, holding promise for significantly improving treatment outcomes in such cases.

1. Introduction

Spinal cord injury (SCI) leads to the loss of sensory, motor, and autonomic functions and affects about half a million patients annually, primarily due to road accidents, violence, sports, or falls. The severity of the injury depends on its location in the spinal cord, causing paraplegia or tetraplegia. SCI poses physical, psychological, and economic challenges, costing up to 1-5 million USD per patient [1,2].

Following a spinal cord injury (SCI), a series of complex processes unfold at both the cellular and molecular levels in the surrounding tissue. As time progresses, the astroglial scar undergoes maturation in the chronic phase [3], leading to the creation of a physical barrier. This barrier is characterized by the presence of molecules like chondroitin sulfate proteoglycans (CSPG) or tenascin-C [4], which significantly hinders axonal growth.

The combination of this barrier formation, along with the excitotoxic environment and the central nervous system's

(CNS) inherent limited regenerative capacity, renders the functional recovery of nerve tissue following injury nearly impossible [5]. These interrelated factors collectively limit the natural healing process, making it extremely challenging to achieve significant functional restoration after an SCI.

To date, despite active research and clinical studies, no standard, effective, and accepted treatment for restoring lost tissue function after SCI has yet been found [6]. In this context, anisotropic scaffolds appear to be a promising strategy to induce directional growth of axons [7]. Among the materials considered to create the anisotropic scaffold, silk fibroin (SF) stands out as an exceptional candidate to promote the organized growth of neuronal axons, since, in addition to being highly biocompatible and possessing excellent mechanical properties and degradation capacity, it can be processed to create scaffolds in various shapes and formats.

In this study, we demonstrate the remarkable capabilities of high-performance fibroin fibers produced through the straining flow spinning (SFS) process. These fibers not only exhibit the capacity to direct axonal growth *in vitro* [8,9], but they also show enhanced guiding abilities when functionalized with adhesion peptides. Moreover, the regenerated silkworm silk fibers show excellent biocompatibility.

The promising outcomes from this research suggest a potential application for biofunctionalized SFS fibers as implants in cases of spinal cord injuries, offering a means to reconnect the severed ends of damaged nerves. This groundbreaking approach holds significant potential for facilitating nerve regeneration and improving the outcomes of spinal cord injury treatments.

2. Materials and Methods

1. Extraction and purification of silk fibroin

SF used for fiber production is derived from silkworm cocoons (*Bombyx mori*). The process involves degumming the cocoons with a 2% Na₂CO₃ (Sigma-Aldrich) solution, dissolving the fibroin in a 9.3M LiBr (Acros Organics) solution at a concentration of 20% (w/v) at 60 °C for 4 h, and dialyzing it against water to remove salts and impurities. To create a spinnable dope, the fibroin concentration is increased by reverse dialyzing it against a solution of 10% (w/v) polyethylene glycol (PEG) (Thermo

Fisher Scientific) with a molecular weight of 8000 Da in 50 mM Tris (Fisher Scientific) base (pH 8.0) and 1 M CaCl_2 (Acros Organics) for 24h at 4°C. This results in a 16% (w/v) SF solution suitable for fiber manufacturing.

2.1. Fabrication of High-Performance Regenerated Silk Fibers through Straining Flow Spinning (SFS)

The straining flow spinning process involves injecting a dope through a silica capillary using a coaxial nozzle, controlled by a focusing fluid. The dope jet then enters a coagulating bath to complete fiber solidification. The focusing fluid has the same composition as the coagulating bath and consists of a solution of acetic acid (1 M) (Thermo Fisher Scientific) in water/ethanol (20/80 (v/v)) (Avantor VWR). The flow rates are $Q_d = 5 \mu\text{L}/\text{min}$ for the dope and $Q_f = 0.5 \text{ mL}/\text{min}$ for the focusing fluid. The fibers are collected on glass coverslips and glued using Dow Corning 3140 RTV Coating bio-compatible glue for cell cultures [10].

2.2. Mechanical characterization of fibers

For the mechanical characterization of the fibers, tensile tests were carried out in both air and water. To be able to mount the fibers in the grips with which the tests were performed, they were individually placed in aluminium foil frames, which were fixed to the grips. Prior to the test, pictures of the fibers were taken with an optical microscope and the diameters were measured using ImageJ software. The grips were clamped into a mechanical testing setup and placed on a Precisa XT 2020A balance. Tensile tests were performed at a constant rate of 1 mm/min with an Instron 4411 mechanical tester. In the case of the tests in water, the test set-up was carried out in a similar way to the tests in air, but in these cases, the frames with the fiber already assembled were introduced into a test tube with deionized water to perform the test.

Precise weight and time data were collected using custom LabVIEW software, subsequently processed to generate true stress-true strain curves. Archimedes' principle and the upward buoyant force exerted on an object immersed in a fluid were considered in the analysis of water test results.

2.3. Biofunctionalization of SF fibers

During the fiber biofunctionalization process, crosslinkers N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC) and N-hydroxy succinimide (NHS) were utilized for peptide labelling and immobilizing the peptides onto the fibers. The RGD (GRGDSP, Sigma-Aldrich) and IKVAV (H-Ile-Lys-Val-Ala-Val-OH, Bachem) peptides were labelled with fluorescein 5(6)-isothiocyanate (FITC, Sigma-Aldrich) to visualize the immobilization on the fibers. For cell culture purposes, non-fluorescent RGD and IKVAV peptides were used. The fibers were incubated with the peptides in a 4-morpholino-ethanesulphonic acid (MES, Sigma-Aldrich) buffer for 1 hour. Afterward, EDC and NHS crosslinkers were added to the solution and incubated for an additional 4 hours at room temperature [11].

2.4. Cell cultures

To study axonal guidance, human neuroblastoma cells (SH-SY5Y) differentiated with retinoic acid and primary cultured mouse cortical neurons were used. All cell cultures were maintained in a humidified atmosphere at 37 °C and 5% CO_2 . In preparation for the guidance study, SH-SY5Y cells were induced to differentiate using retinoic acid (RA) and brain-derived growth factor (BDNF) [12].

2.5. Determination of axonal guidance

In axonal guidance studies with SH-SY5Y cells, the progress was documented using an inverted light microscope. On the other hand, for primary neuronal cultures, immunofluorescence staining was performed to distinguish the cell types. The "guidance factor" (ϕ_G) was used to measure the degree of alignment of growing axons. ϕ_G is calculated by counting the intersections of axons with lines on a grid superimposed to the images taken during the cell culture at defined times. The value is calculated as the number of intersections with the perpendicular lines divided by the number of intersections with the parallel lines. A $\phi_G = 1$ indicates isotropic growth, while $\phi_G > 1$ indicates preferential orientation with the fibers.

2.6. Immunocytochemistry

In summary, the cells were washed with PBS and fixed in a 4% paraformaldehyde solution for 10 minutes. After they were permeabilized with a 0.1% Triton X-100 solution for 10 minutes and blocking was carried out using PBS with 2% BSA. The cells were then incubated overnight at 4°C with primary antibodies against GFAP (1:5000; Ab7260) and β -III tubulin (1:1000; Ab7751) diluted in PBS with Triton X-100 and BSA. After washing, secondary antibodies, anti-mouse FITC (1:200) and anti-rabbit Cy3 (1:800; Jackson ImmunoResearch) were applied for 3 hours at room temperature in the dark. Finally, neurons were washed with PBS, mounted with Fluoroshield with DAPI, and observed under a fluorescence microscope. Five images were taken at random points on the well using a Z-pattern for each case.

2.7. Statistics

The analysis was conducted using GraphPad Prism 8.0.1, and a two-way ANOVA test was employed with a significance level of $p < 0.05$.

3. Results

3.1. Mechanical characterization

After performing tensile tests on the SF fibers (Figure 1A-B), wherein weight and time data were recorded to derive true stress-strain curves (Fig. 1C-D), the fiber diameter was measured using the ImageJ software. The average diameter of the fibers was $16.8 \pm 0.6 \mu\text{m}$.

Key parameters, such as the elastic modulus, maximum stress, maximum strain, and fracture work, were derived from the curves (Table 1). It is noteworthy that while the fibers show modest mechanical properties in air, they demonstrate improved characteristics in water. This is of significant interest since these fibers will be employed in aqueous environments.

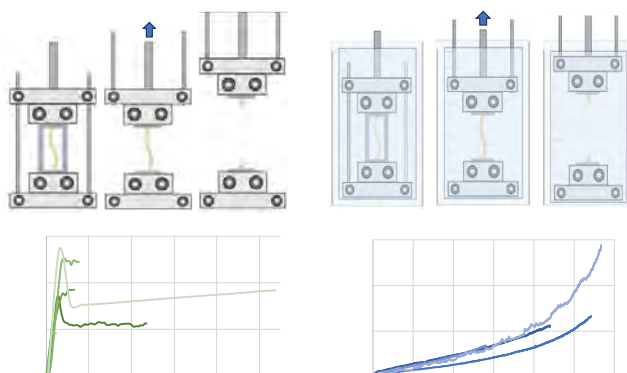


Figure 1. Mechanical properties by tensile test of SF regenerated fibers tested in air and water. Graphical representation of the tensile test in air (A) and in water (B). True stress-true strain curves of fibres test in air (C) and in water (D).

Type of assay	Elastic modulus (GPa)	Breaking stress	Breaking strain	Yield stress (MPa)	Work of fracture (MJ/m ³)
Air	4.7±0.1	29.3± 5.9	0.07± 0.07	35.6± 7.2	2.1± 1.3
Water	-	59.8±19.2	0.9± 0.1	-	20.8± 6.2

Table 1. Summary of mechanical properties of tested fibers. Table values appear as the mean ±SEM.

3.2. Effectiveness of functionalization

To evaluate the effectiveness of biofunctionalizing SFS fibers using EDC/NHS crosslinkers, RGD and IKVAV peptides bound to FITC were employed. Additionally, the unspecific adsorption of FITC to the fibers was tested by incubating them with the peptides but without the crosslinkers. Nonfunctionalized fibers were used as a control. It is observed that the fluorescence intensity was higher in the RGD- and IKVAV-functionalised samples in the presence of EDC/NHS compared to the control (p-value <0.05 and < 0.001) respectively (Fig. 2). The fluorescence of the fibers has been computed in terms of corrected total fiber fluorescence (CTFF), which is calculated by subtracting from the Integrated Density the area of selected cell by Mean fluorescence of background readings.

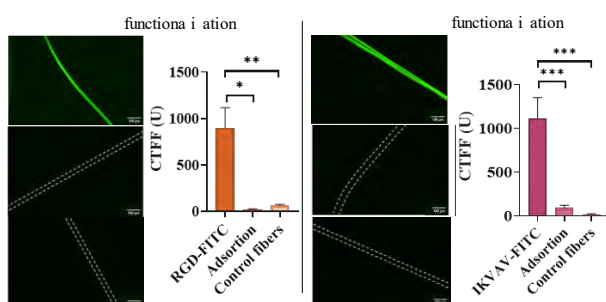


Fig. 2. Analysis of the functionalization of fibers with RGD (A) and IKVAV (B) peptides using EDC/NHS crosslinkers. (A.a) and (B.a) show the functionalized fibers in the presence of the crosslinkers. (A.b) and (B. b) show the adsorption of the peptides without the presence of the crosslinkers. (A.c) and (B.c) show the control fibers. (A.d) and (B.d) The one-way ANOVA test was performed considering the means ± SEM of te (**= p-value < 0.01; ****= p-value < 0.0001).

3.3. Determination of the axonal guidance factor (ϕ_G)

To assess the effectiveness of the guidance factor (ϕ_G) as a quantitative measure of axonal guidance, the SH-SY5Y cells were seeded on both control polystyrene plates and SFS fibers (Fig. 2).

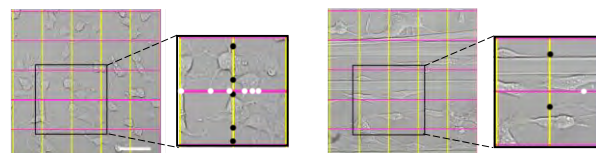


Fig.2. Axonal guidance factor (ϕ_G): Grids divided into 5 equidistant bands both horizontally and vertically were superimposed on bright field microscopy images of SH-SY5Y differentiated for 6d on control polystyrene (A) and SH-SY5Y on SFS fibers (the grids are oriented parallel and perpendicular to the fibers)

Over the course of 6, 11, and 18 days, five photos were taken of each well for analysis. Using the Two-way ANOVA test (Table 2), the results demonstrated that when cells were seeded on polystyrene, the ϕ_G value was close to 1, indicating a random and non-directed growth pattern, forming networks without any specific orientation. Conversely, when the cells were seeded on SFS fibers, the ϕ_G value exceeded 1, indicating a significant level of directionality in the axonal growth. This observation was statistically significant at day 6 (p-value < 0.0001), day 11 (< 0.01), and day 18 (< 0.0001). Thus, the ϕ_G measurement effectively quantifies and confirms the guided and directional axonal growth on the SFS fibers.

ϕ_G	Day 6	Day 11	Day 18
Control	0,8±0,04	0,9 ±0,04	0,9±0,04
Fibers	2,9±0,3	2,2±0,3	2,7±0,4

Table 2. Calculation of the ϕ_G guidance factor. SH-SY5Y cells seed onto polystyrene plates (control) and SF fibers were analyse at day 6, 11 and 18.

3.4. Axonal guidance in SH-SY5Y

The ϕ_G values obtained were summarized in Table 3. At day 6, differentiated SH-SY5Y cells showed axons growing in alignment with the fiber orientation, in contrast to the isotropic growth observed in the control group on polystyrene. At 18 days, the differences in ϕ_G remained consistent for both non-functionalized and functionalized fibers compared to the polystyrene control.

ϕ_G	Control	Nonfunctionalized fibers	Fibers - RGD	Fibers - IKVAV
Day 6	0,9±0,04	2,9 ±0,3	2 ± 0,3	2 ± 0,3
Day 11	0,9±0,04	2,2 ± 0,7	1,7 ± 0,2	1,8 ± 0,2
Day 18	0,9±0,04	2,6 ±0,4	2,4±0,3	2,1 ±0,1

Table 3. Quantification of ϕ_G in SH-SY5Y cell culture. Guidance factor measured in polystyrene plates, nonfunctionalized fibers, RGD biofunctionalized fibers and IKVAV biofunctionalized fibers of differentiated SH-SY5Y at different times 6, 11 and 18 days (Mean ±SEM).

3.5. Axonal guidance in primary culture neurons

In this study, mouse cortical cells were cultured on nonfunctionalized and biofunctionalized SFS fibers to investigate the orientation of cell processes. The cells were differentiated using specific markers for neurons and astrocytes after 7 days of culture (Fig. 5).

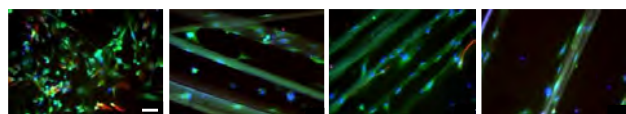


Fig. 5. Immunofluorescence images of mouse cortical cells at 7 days. Cells seeded on polystyrene (A), nonfunctionalized fibers (B), RGD-functionalized fibers (C) and IKVAV-functionalized fibers (D) at 7 days (Scale bar= 50µm). Neurons are labelled in green (β -III tubulin), astrocytes in red (GFAP) and nuclei in blue (DAPI).

Fluorescence microscopy images revealed that the cortical cultures primarily consisted of neurons in both cases. The orientation of cell development was quantified using the ϕ_G factor. For cells cultured on polystyrene (control), ϕ_G was close to 1, indicating isotropic axonal growth. However, when cells were cultured in the presence of fibers, the values of ϕ_G were greater than 1, indicating that the axons preferred alignment along the direction of the fibers. Furthermore, there were significant differences between nonfunctionalized and functionalized fibers, suggesting that the biofunctionalization influenced the axonal alignment (as confirmed by Two-way ANOVA analysis, p-value < 0.01) (Table 4).

ϕ_G	Control	Nonfunctionalized fibers	Fibers - RGD	Fibers- IKVAV
Day 7	1,1 ± 0,1	2,6 ± 0,3	5,07 ± 1	4,7 ± 0,5

Table 4. Quantification of ϕ_G in neuron primary culture. Guiding factors of axonal orientation of neurons of primary culture on polystyrene and on nonfunctionalized fibers, functionalized with RGD and IKVAV (Mean \pm SEM).

4. Discussion and Conclusion

This study focuses on the development of biomaterial implants to promote tissue regeneration in spinal cord injuries. Current clinical solutions have limited success, and the central nervous system's regenerative capacity is limited. The goal is to create scaffolds using silk fibroin, which has several advantages, including excellent mechanical properties and neurobiocompatibility.

To enhance the biointegration of SF, the fibers are biofunctionalized with peptides derived from fibronectin (RGD) and laminin (IKVAV) using crosslinking chemistry. The results show that functionalized fibers have an effect on the guide factor in both types of cells evaluated. However, RGD and IKVAV peptides had a more prominent effect on genuine neurons than on SH-SY5Y cells. This suggests that functionalized SFS fibers could be useful as implants for spinal cord injuries, facilitating the connection between the ends of the spinal cord. The broad therapeutic potential of this approach is highlighted due to the possibility of using various peptides

compatible with biofunctionalization and the use of additional soluble factors.

Funding

This study was partially funded by the Ministerio de Ciencia e Innovación (PID2020-116403RB-I00; MCIN/AEI/10.13039/501100011033), Comunidad de Madrid (MINA-CM P2022-BMD-7236) and by the agreement between the Comunidad de Madrid (Spain) and the UPM through the REACT-UE funds of the European Regional Development Fund (ERDF), as a part of the response of the European Union to the SARS-CoV2 and COVID-19 pandemic. It also has been financed by the UCM together with Banco Santander Scholarships (BOUC 11-23-2021; ref.: CT58/21-CT59/21).

References

- [1] Página web de la Organización Mundial de la salud. <https://www.who.int/news-room/fact-sheets/detail/spinal-cord-injury> (consultada en 18 Julio 2023).
- [2] Ashammakhi N, Kim HJ, Ehsanipour A, Bierman RD, Kaarela O, Xue C, Khademhosseini A, Seidlits SK. Regenerative Therapies for Spinal Cord Injury. *Tissue Engineering Part B: Reviews*, vol 25, 2019, pp 471-491.
- [3] Alizadeh A, Dyck SM, Karimi-Abdolrezaee S. Traumatic Spinal Cord Injury: An Overview of Pathophysiology, Models, and Acute Injury Mechanisms. *Frontiers in Neurology*, vol 10, 2019, pp 282.
- [4] Wiese S, Karus M, Faissner A. Astrocytes as a Source for Extracellular Matrix Molecules and Cytokines. *Frontiers in pharmacology*, vol 3, 2012, pp 120.
- [5] Tedeschi A, Omura T, Costigan M. CNS Repair and Axon Regeneration: Using Genetic Variation to Determine Mechanisms. *Experimental neurology*, vol 287, 2017, pp 409-422.
- [6] Ashammakhi N, Kim HJ, Ehsanipour A, Bierman RD, Kaarela O, Xue C, Khademhosseini A, Seidlits SK. *Tissue Engineering Part B: Reviews*, vol 25, 2019, no 6, pp 471-491.
- [7] Xue W, Shi W, Kong Y, Kuss M, Duan B. Anisotropic Scaffolds for Peripheral Nerve and Spinal Cord Regeneration. *Bioactive materials*, vol 6, 2021, pp 4141-4160.
- [8] Pérez-Rigueiro J, Madurga R, Gañán-Calvo AM, Plaza GR, Elices M, López PA, Daza R, González-Nieto D, Guinea GV. Straining Flow Spinning of Artificial Silk Fibers: A Review. *Biomimetics*, vol. 3, 2018, pp 29.
- [9] Madurga R, Gañán-Calvo AM, Mariscal T, Plaza GR, Guinea GV, Elices M, Pérez-Rigueiro J. Production of Regenerated Silkworm Silk Fibers from Aqueous Dopes through Straining Flow Spinning. *Textile Research Journal*, vol 89, 2019, vol 89, pp 4554-4567.
- [10] Lozano-Picazo P, Castro-Domínguez C, Bruno AL, Baeza A, Martínez AS, López PA, et al. Strategies for the Biofunctionalization of Straining Flow Spinning Regenerated Bombyx mori Fibers. *Molecules*, vol 27, pp 4146.
- [11] Álvarez-López A, Colchero L, Elices M, Guinea GV, Pérez-Rigueiro J, González-Nieto D. Improved cell adhesion to activated vapor silanization-biofunctionalized Ti-6Al-4V surfaces with ECM-derived oligopeptides. *Biomaterials Advances*, vol 133, 2022, pp 112614.
- [12] Castro-Domínguez C, Lozano-Picazo P, Álvarez-López A, Garrote-Junco J, Panetsos F, Guinea GV, Pérez-Rigueiro J. Axonal Guidance Using Biofunctionalized Straining Flow Spinning Regenerated Silk Fibroin Fibers as Scaffold. *Biomimetics*, vol 8, 2023, pp 65.

Microfluidic Model of the Alternative Vasculature in Neuroblastoma

Aránzazu Villasante^{1,2*}, Maria José López-Martínez^{1,2,3}, Andrea García-Lizarribar^{1,3}, Xiaofeng Peng¹,
Josep Samitier^{1,2,3}

¹Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology (BIST), Spain.

²Department of Electronic and Biomedical Engineering, University of Barcelona, Spain.

³Biomedical Research Networking Center in Bioengineering, Biomaterials, and Nanomedicine (CIBER-BBN), Spain.

Abstract

Neuroblastoma (NB) is a highly vascularized pediatric tumor arising from undifferentiated neural crest cells early in life, exhibiting both traditional endothelial-cell-driven vasculature and an intriguing alternative vasculature. The alternative vasculature can arise from cancer cells undergoing transdifferentiation into tumor-derived endothelial cells (TEC), a trait associated with drug resistance and tumor relapse. The lack of effective treatments targeting NB vasculature primarily arises from the challenge of establishing predictive in vitro models that faithfully replicate the alternative vasculature phenomenon.

In this study, we aim to recreate the intricate vascular system of NB in an in vitro context, encompassing both types of vascularization, by developing a novel neuroblastoma-on-a-chip model. We designed a collagen I/fibrin-based hydrogel with an Emod = 0.87 ± 0.28 kPa, closely mirroring NB's physiological composition and tumor stiffness. This biomaterial created a supportive environment for the viability of NB and endothelial cells. By introducing a physiological shear stress value of 36 dynes/cm², matching the range observed in arteries and capillaries, to the microfluidic chip device, we successfully induced the formation of vessel-like structures and triggered transdifferentiation of NB cells into TECs.

The vascularized neuroblastoma-on-a-chip model introduced here presents a promising and complementary strategy to animal-based research with a significant capacity for delving into NB tumor biology and vascular targeting therapy.

seamlessly integrate with high-throughput drug screening platforms^{2,3,4}.

Neuroblastoma (NB) is one type of solid tumor that could be miniaturized by using TOAC technology, and for which we need better preclinical models and a cure. NB is a malignant tumor of the neural crest cells that give rise to the sympathetic nervous system. It is a rare cancer and the most common extracranial solid tumor of childhood diagnosed in the first year of life. It frequently arises in the adrenal medulla, although the tumor's location can be anywhere in the sympathetic nervous system (abdominal origin in 80% of cases). MYCN oncogene is frequently amplified in high-risk NB, linked to an undifferentiated phenotype and poor prognosis. NB is also a highly angiogenic tumor^{5,6,7}.

Tumor vascularization is crucial for regulating growth, survival, and metastasis in solid tumors. Tumor vessels can be formed through physiological mechanisms like angiogenesis (sprouting and intussusceptive) and vasculogenesis. Initially, sprouting angiogenesis generates new blood vessels from preexisting ones, while intussusceptive angiogenesis involves splitting vessels through intra-luminal tissue pillars. Tumors also utilize embryonic vasculogenesis, where vascular precursors differentiate into endothelial cells^{8,9}. Tumor vascularization strategies also encompass cancer-specific methods like vessel co-option, integrating healthy organ vessels into the tumor, as well as alternative mechanisms such as vasculogenic mimicry and endothelial transdifferentiation. In vasculogenic mimicry, tumor cells align and form CD31-negative, periodic acid-Schiff (PAS)-positive vessel-like structures. Endothelial transdifferentiation enables cancer cells to acquire endothelial traits, including CD31 expression, resulting in tumor-derived endothelial cells (TEC)^{8,9}. Cancer-specific vessel formation strategies are likely implicated in chemoresistance across various tumors^{10,11}.

In the past two decades, numerous studies have explored the correlation between tumor advancement and angiogenesis in NB using both *in vitro* and *in vivo* experimental models. Consequently, targeting vasculature and suppressing vasculature-related factors like VEGF has been postulated as a potential approach for treating NB^{12,13}. VEGF plays a crucial role in neuroblastoma by promoting endothelial cell growth, angiogenesis, and migration. Numerous experimental anti-angiogenic treatments, including VEGF inhibitors and antibodies against VEGF receptors, strive to block the VEGF signaling pathway,

1. Introduction

In the last years, engineered human tissues designed for regenerative therapies have taken on a miniature form known as organs-on-a-chip (OOAC). These OOACs are compact plastic devices, no larger than a thumb, featuring microchannels for cultivating diverse human cell types and small chambers filled with culture medium. Within this setup, human cells are cultured in a controlled microenvironment closely mirroring organ functions, enabling the exploration of biomedical phenomena¹.

An impactful application of this technology, reshaping cancer research and the pharmaceutical sector, is the creation of tumors-on-a-chip (TOAC). The key advantage of this technology lies in the intricate design of each TOAC as a comprehensive microsystem, replicating the cellular composition and microarchitecture of native tumors while emulating specific functions in a scaled-down 3D format. This innovation expedites the translation of therapeutic discoveries to patients in need. Beyond their predictive capabilities favoring precision medicine, TOACs offer a remarkably adaptable platform utilizing minimal cell quantities and microlitre volumes of reagents, all in a cost-effective and time-efficient manner. Furthermore, TOACs

thus hindering the formation of tumor blood vessels. Despite the effort, antiangiogenic strategies effective in preclinical models have not translated to improved patient survival in clinical trials¹².

NB also employs one of the alternative vascularization mechanisms previously mentioned, where NB cells transform into TEC-expressing CD31 and MYCN amplification markers¹⁴. Current antiangiogenic approaches primarily focus on angiogenesis-generated vasculature, overlooking alternative mechanisms that could cause the clinical setbacks observed⁹. Hence, comprehending TEC biology is crucial for designing effective therapies in NB. However, the lack of effective treatments for targeting TEC and the alternative vasculature in neuroblastoma is primarily due to the challenge of creating accurate predictive *in vitro* models replicating this phenomenon. Establishing representative research models of TEC behavior is essential for advancing understanding and developing targeted therapies for neuroblastoma.

Thus, we proposed creating an NB-on-a-chip model encompassing both the “traditional” vascularization driven by angiogenesis and the alternative vasculature generated by transdifferentiation observed in patients. To achieve this, we replicated the physiological stiffness of NB by fabricating a soft collagen-based biomaterial with Young’s modulus of approximately 1kPa. This biomaterial, designed to mimic the natural composition, supported the proliferation and viability of NB and endothelial cells. Various combinations of microenvironmental factors were employed to culture NB cells, yet a static environment failed to trigger vasculature formation or transdifferentiation. However, the introduction of shear stress through the microfluidic device proved pivotal. This induced the expression of CD31 in NB cells, leading to their transformation into tumor-derived endothelial cells and the consequent generation of alternative vasculature. Our innovative NB-on-a-chip model is the first tool for evaluating different vasculature responses to drugs in a high-throughput, cost-effective, and ethically sound platform, serving as a preliminary step before assessing treatment effectiveness in more intricate models such as murine models and actual patients.

2. Materials and Methods

2.1. Cell culture

Neuroblastoma cells. SK-N-BE (2) (from the American Type Culture Collection, ATCC) was cultured according to the manufacturer’s specifications in RPMI medium supplemented with 10% (v/v) FBS and 1% penicillin/streptomycin. Cells were cultured at 37°C and 5% CO₂ in a humidified incubator.

Human Umbilical Vein Endothelial Cells (HUVEC) were purchased from Lonza and cultured according to the manufacturer’s specifications. HUVEC cells were cultured in HUVEC expansion medium EBM-2 Basal Medium (Lonza CC-3156) supplemented with EGM-2 BulletKit (Lonza CC-3156 & CC-4176) at 37°C in a humidified incubator at 5% CO₂.

Live-Dead assay. Samples were incubated in RPMI medium containing 2μM Calcein and 4μM of ethidium homodimer-1 for 30 min at 37°C, 5% CO₂, as indicated by the manufacturer’s protocol (LIVE/DEAD® Viability/Cytotoxicity Kit, Molecular Probes). Samples were imaged with a fluorescence microscope (Olympus IX81 light microscope, Center Valley, PA).

2.2. Hydrogel preparation and characterization

FX3 and FX5 formulations are both collagen I/fibrin-based hydrogels. They were fabricated using bovine collagen I (2μg/μL; ThermoFisher Scientific, #A1064401), fibrinogen from bovine blood plasma (10μg/μL; Sigma

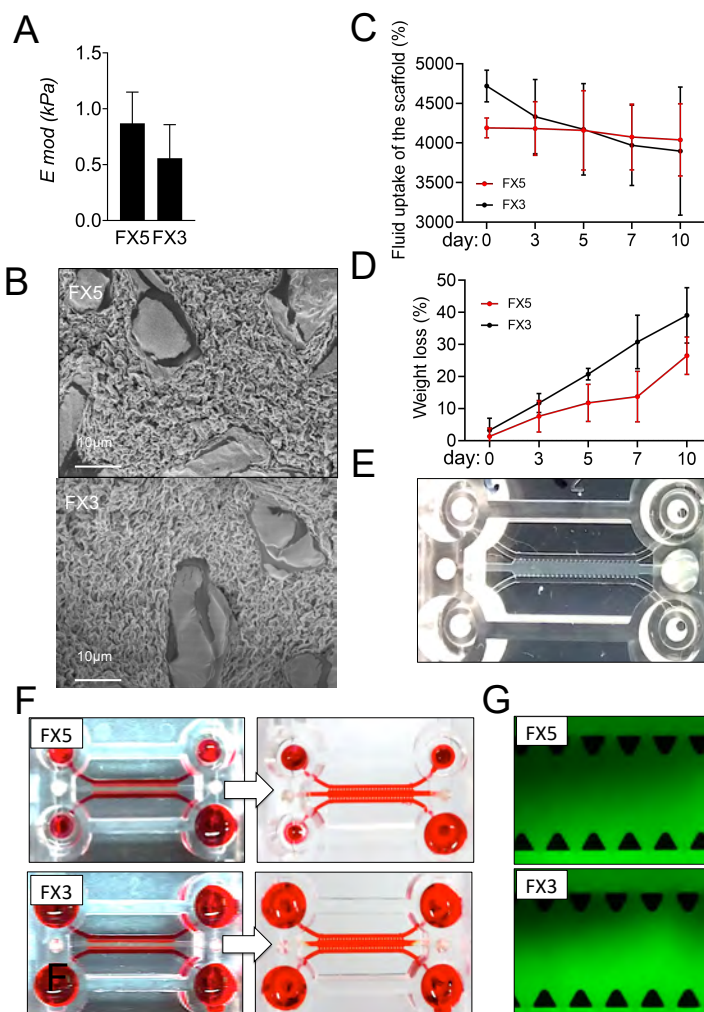


Figure 1. Characterization of collagen/fibrin-based hydrogels. (A) Equilibrium modulus (E_{mod}) of the indicated formulations FX5 ($n=3$) and FX3 ($n=3$) recapitulating native stiffness of neuroblastoma tumors. (B) Scanning Electron Microscopy (SEM) images of the FX5 ($n=3$) and FX3 ($n=3$) formulations. (C) Fluid uptake of the hydrogels in water ($n=3$ per condition and time point). (D) Degradation of the FX5 and FX3 hydrogels ($n=3$ per condition and time point). (E) Image of a butterfly-shaped microfluidic chip with a central chamber and two lateral channels for perfusion. The central chamber is filled with the FX5 hydrogel that crosslinks inside the chamber ($n=15$). (F) Images showing homogeneous diffusion of red ink through the FX5 and FX3 hydrogels at the initial (left) and the final (right) time points ($n=3$). (G) Fluorescence microscopy images of FITC-conjugated dextran diffusing into the indicated biomaterials in the gel region of the chip.

#F8630), were cross-linked with Thrombin (Sigma # T4648).

2.3. Mechanical properties

Uniaxial compression tests of materials were performed using a Zwick Z0.5 TN instrument (Zwick-Roell, Germany) with a 5 N load cell. Samples were tested at room temperature up to 30% final strain (deformation), using the following parameters: 0.1 mN preload force and 20% min⁻¹ strain rate. Stress-strain graphs were obtained from load-deformation measurements. Values for the compressive modulus were calculated from the slope of the linear region corresponding to 10–20% strain.

2.4. Microfluidic device and perfusion setup

idenTx 3 Chips (AIM Biotech) were assembled in an idenTx 9 Plate for the indicated experiments. LUC-1 luer (AIM Biotech) and male luer integral 1/16 connectors (Cole-Parmer #45518-07) were used to connect Tygon tubes (Tygon AAD04133; inner diameter = 1.27mm, outer diameter = 2.28 mm) to chip inlets and outlets. Tygon tubes were connected to the peristaltic pump tubing (Tygon LMT-55 #070534-04-ND) through a small piece of Tygon tube of 2mm inner diameter (Tygon # ACSF1S1502-C). The recirculating circuit was achieved by attaching the other extreme of Tygon tubes to a reservoir fabricated with a 50 mL Falcon. The falcon tube was also used as a bubble trap system. The culture medium was pumped using a high-precision IPC-N 8 peristaltic pump (Ismatec) and constantly perfused for 6 days to get a wall shear stress of 36 dynes/cm². The idenTx 9 plates containing the devices were placed inside a humidified incubator at 37°C and 5% CO₂.

2.5. Quantitative Real-Time PCR (qRT-PCR)

Total RNA from cells was obtained using Trizol (Life Technologies), following the manufacturer's instructions. RNA preparations were treated with "Ready-to-go you-prime first-strand beads" (GE Healthcare) to obtain cDNA. Quantitative real-time PCR was performed using PowerUp SYBR Green Master Mix (Applied Biosystems #A25742). mRNA expression levels were quantified applying the Δ Ct method, Δ Ct = (Ct of the gene of interest - Ct of Actin). PECAM primers were obtained from the PrimerBank database.

3. Results

We first aimed to engineer a biomaterial that emulates the intrinsic stiffness observed in tumors from patients to replicate the physiological microenvironment inherent to NB more accurately. As a soft tumor, neuroblastoma exhibits a characteristic low Young's modulus (E_{mod}) of $E_{mod} = 0.79 \pm 0.20$ kPa¹⁵. We finetune a collagen I/fibrin-based biomaterial by modulating the concentration of the thrombin enzyme. This adjustment yielded two distinct and consistent formulations falling within the range of natural NB stiffness, named FX5 ($E_{mod} = 0.87 \pm 0.28$ kPa) and FX3 ($E_{mod} = 0.56 \pm 0.30$ kPa) (Figure 1A, B).

The swelling percentage of the scaffolds over time in both formulations was measured to investigate their fluid absorption behavior (Figure 1C). Both materials did not

show significant differences regarding swelling behavior over time. The study of the percentage of weight loss over time demonstrated that the FX5 material is more stable in liquids than the FX3; the FX5 can remain almost intact until day 7 while the FX3 is partially degraded (Figure 1D). Based on previous experience, at least seven days of cell culturing are necessary to obtain biomimetic responses in a bioengineered cancer model¹⁷. Thus, based on this first set of experiments, FX5 was the most suitable for further studies.

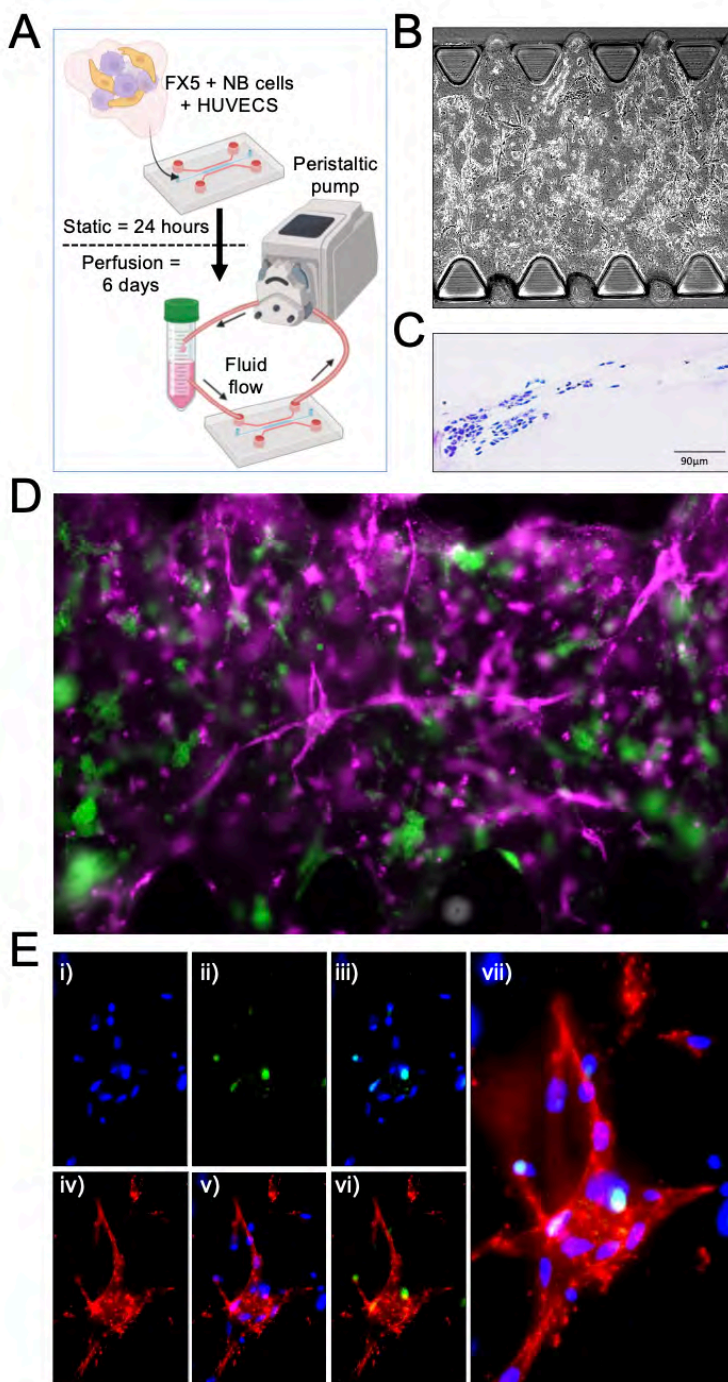
Gelation inside the chip was also analyzed, and time-dependent diffusion of ink and FITC-dextran into the hydrogel matrices was monitored. Both biomaterials were perfectly gelled within the microfluidic device's central chamber without interfering with culture media perfusion through the lateral microchannels (Figure 1E). We also observed homogenous ink diffusion and FITC-dextran through the entire hydrogel within the microfluidic chip for both formulations, which anticipates that an *in vitro* tumor model may enable the supply of nutrients through diffusion to the cells (Figure 1F-G).

Before introducing the cells inside the chip device, we explored the biocompatibility and vasculoinductive properties of the formulations. We set the culture conditions by culturing NB cells on top of the FX5 and FX3 biomaterials using RPMI medium. Live/dead assays showed both materials' biocompatibility and null cytotoxicity for at least 7 days. We also assessed the capability of the materials to support endothelial cell survival and the formation of vessel-like structures. We co-cultured NB cells and HUVECs using an RPMI and endothelial medium (EGM) mix (1:1) for 7 days. The FX5 biomaterial showed a better capacity to sustain the co-culture alive than the FX3. Still, it did not significantly affect PECAM, CD34, and MYCN expression levels but supported the formation of grouped proto-tissue-like vessels expressing CD31. Based on the results, we selected the FX5 material, the co-culture of NB/HUVEC cells, and the RPMI/EGM medium as the most suitable conditions for further studies in the chip. Consequently, a co-culture of NB cells and HUVEC were seeded within the FX5 hydrogel and cultured in RPMI/EGM mix medium for 24h, and then perfusion was applied for 6 days (Figure 2A). Vessel-like tubular structures were identified when 36 dynes/cm² were applied to the system (Figure 2B). We confirmed tissue formation by histological analysis (Figure 2C) and vasculature generation by CD31 endothelial marker staining (Figure 2D). Finally, we also demonstrated the formation of the alternative vasculature in the tissue by the identification of tumor-derived endothelial cells, which expressed both endothelial (CD31) and neuroblastoma (MYCN) markers at the same time (Figure 2E).

Conclusions

This research addresses the challenge of mimicking the neuroblastoma microenvironment by developing biomaterials with appropriate stiffness. FX5 proved to be the most suitable formulation, supporting viable co-cultures and vessel-like structures. Introducing shear stress led to enhanced vascularization, culminating in identifying

Figure 2. Analysis of endothelial cell transdifferentiation in a perfused neuroblastoma-on-a-chip. (A) Schematic depiction of the experimental protocol to build the vascularized neuroblastoma-on-a-chip. Created with BioRender.com. (B) Bright-field image of the chip seeded with NB cells and HUVEC within the FX5 hydrogel in the central chamber and cultured in RPMI/endothelial (EGM) mix medium under flow perfusion conditions for 6 days (n=9)—representative image showing a tubular network's formation under flow perfusion conditions at day 7 after cell seeding. (C) Representative image of paraffin-embedded tissue sections stained with Hematoxylin and Eosin (H&E) obtained on day 7 after cell seeding and under perfusion conditions (n=3). (D) Fluorescence image of the vascularized neuroblastoma-on-a-chip double stained for MYCN (neuroblastoma marker; green) and CD31 (endothelial marker; magenta) (n=6). (E) Image showing tumor-derived endothelial cells (TEC) in the vascularized tissue generated in the microfluidic device. (i) DAPI staining (blue, nuclei), (ii) MYCN staining (neuroblastoma marker, green); (iii) image showing co-localization of MYCN in the nuclei of neuroblastoma cells; (iv) CD31 (endothelial marker; red); (v) DAPI/CD31 double staining; (vi) MYC/CD31 double staining showing neuroblastoma cells expressing the endothelial marker indicating the presence of TEC in the tissue; (vii) Merge (DAPI-blue, MYCN-green, CD31-red).



tumor-derived endothelial cells expressing both CD31 and MYCN markers. These findings provide valuable insights into recreating neuroblastoma's alternative vasculature and contribute to developing more accurate and predictive *in vitro* models for studying tumor behavior and exploring potential therapeutic interventions.

Acknowledgments

The Nanobioengineering Lab is a member of the Biomedical Research Networking (CIBER), Spain, funded by the VI National R&D&i Plan 2008-2011 and the Instituto de Salud Carlos III (RD16/0006/0012; RD16/0011/0022) with the support of the European Regional Development Fund (ERDF), and Generalitat de Catalunya (2023SGR 015452). AV is supported by AECC, the NEN Association, and the R+D+i plan PID2020-117977RA-I00.

References

- Bhatia, S. N. & Ingber, D. E. Microfluidic organs-on-chips. *Nat Biotechnol* 32, 760–72 (2014).
- Liu, X. *et al.* Tumor-on-a-chip: from bioinspired design to biomedical application. *Microsyst. Nanoeng.* 7, 50 (2021).
- Wang, R., Zhang, C., Li, D. & Yao, Y. Tumor-on-a-chip: Perfusable vascular incorporation brings new approach to tumor metastasis research and drug development. *Front. Bioeng. Biotechnol.* 10, 1057913 (2022).
- Hakanson, M., Cukierman, E. & Charnley, M. Miniaturized pre-clinical cancer models as research and diagnostic tools. *Adv Drug Deliv Rev* 69–70, 52–66 (2014).
- Matthay, K. K. *et al.* Neuroblastoma. *Nat Rev Primer* 2, 16078 (2016).
- Otte, J., Dyberg, C., Pepich, A. & Johnsen, J. I. MYCN Function in Neuroblastoma Development. *Front. Oncol.* 10, (2021).
- Chlenski, A., Liu, S. & Cohn, S. L. The regulation of angiogenesis in neuroblastoma. *Cancer Lett.* 197, 47–52 (2003).
- Krishna Priya, S. *et al.* Tumour angiogenesis-Origin of blood vessels. *Int J Cancer* 139, 729–35 (2016).
- Belotti, D., Pinessi, D. & Taraboletti, G. Alternative Vascularization Mechanisms in Tumor Resistance to Therapy. *Cancers* 13, 1912 (2021).
- Soda, Y. *et al.* Transdifferentiation of glioblastoma cells into vascular endothelial cells. *Proc Natl Acad Sci U S A* 108, 4274–80 (2011).
- Pezzella, F. & Ribatti, D. Vascular co-option and vasculogenic mimicry mediate resistance to antiangiogenic strategies. *Cancer Rep.* 5, (2022).
- Roy Choudhury, S., Karmakar, S., Banik, N. L. & Ray, S. K. Targeting angiogenesis for controlling neuroblastoma. *J. Oncol.* 2012, 782020 (2012).
- Ribatti, D. Targeting Angiogenesis in Neuroblastoma. in *Neuroblastoma* 79–93 (Elsevier, 2019). doi:10.1016/B978-0-12-812005-7.00005-9.
- Pezzolo, A. *et al.* Tumor origin of endothelial cells in human neuroblastoma. *J Clin Oncol* 25, 376–83 (2007).
- Monferrer, E. *et al.* A three-dimensional bioprinted model to evaluate the effect of stiffness on neuroblastoma cell cluster dynamics and behavior. *Sci. Rep.* 10, (2020)

A microphysiological setup to mimic, test and analyze myocardial tissue for cardiovascular diseases

E. Yanac Huertas^{1,2}, M. Roquette¹, A. Vilche Mariscal^{1,2}, B. Lloveras Borràs¹, A. Noguera Monteagudo^{1,2}, S. Jin^{1,2}, R. Rodríguez^{1,2}, J.A. Del Río², O. Castaño^{1,2}

¹ Electronics and Biomedical Engineering, Universitat de Barcelona (UB), 08028 Barcelona, Spain

² Institute for Bioengineering of Catalonia (IBEC - BIST), Spain

Abstract

Creating cardiac tissue models for drug testing before clinical trials is still a major challenge in drug discovery. This is mainly because it is difficult to replicate the complex nature of cardiac tissue in a lab setting. One of the biggest challenges is accurately mimicking the functional features of the heart muscle. This is largely due to the immaturity of the cells used in these models, which makes it difficult to get reliable results that can be translated to human patients. Currently, in vivo models are the gold standard for evaluating new treatments. However, it is widely accepted that these animal models cannot fully reproduce human physiological responses. As a result, this mismatch often leads to failures in subsequent clinical trials. In this work, we designed two cardiac culture experiments, using GCamp6-infected cardiomyocytes, on a microfluidic chip. One of them is an aligned 2D model, and the other is a 3D tissue grown in a hydrogel of methacrylated gelatin (GelMa) and methacrylated hyaluronic acid (HaMa), the calcium traces of both were analysed at the network level, being able to obtain parameters that determine the degree of maturity and connectivity of the culture. In this way, we can ensure that we have a reproducible environment for future drug testing.

1. Introduction

Cardiovascular diseases (CVD) rank as the foremost cause of death on a global scale. Available therapeutic options consist of lifestyle and dietary modifications, medication, or surgical interventions. Since none of these interventions will completely regenerate the myocardium, it is vitally important to develop regenerative therapies to mitigate damage. Currently, the construction of *in vitro* models using innovative microfluidic techniques is gaining significance for testing these therapies during the preclinical stage. Among these diseases, our work focuses on those affecting the myocardium, which is a complex tissue with specialized cells (fibroblasts, smooth muscle cells, and cardiomyocytes) and an extracellular matrix (ECM) of collagen, elastin, fibronectin, laminin, and proteoglycans.

In this study, we designed a microfluidic chip and built two tissue culture models using mouse cardiomyocyte cell line infected for us with A calcium indicator (GCaMP6). The first model is a 2D aligned tissue with PLA fibers, and the other is a 3D model that uses an interpenetrated

hydrogel based on GelMa and HaMa to mimic the ECM. Both cultures were recorded to extract maturation information from the calcium traces.

2. Materials and methods

2.1. Electrospinning

The polymer chosen for electrospinning was polylactic acid (PLA) 70/30 due to its good adhesion and proliferation properties [1]. An 8% solution of PLA 70/30 in 2,2,2-trifluoroethanol was used, which was left overnight on a three-dimensional orbital shaker at 50 rpm.

It was chosen to electrospin the coverslips before the plasma process to attach them to the PDMS for the manufacturing of PLA fiber-based devices. The coverslips were first cleaned with ethanol before being taped on sheets of aluminum foil measuring 20 x 32 cm. Three coverslips were centered on each of these sheets, with three centimeters between each one. Finally, using adhesive tape, the aluminum foil was applied to the collector. The PLA solution was loaded into a 10 ml syringe and attached to a pump (*NE-1000, New Era Instruments*) which delivered a flow rate of 0.5 ml/h. A high-voltage power supply (*NanoNC-403010, NanoNC*) was used to provide a potential of 18 kV between the syringe needle and the grounded collector. To obtain aligned fibres, the collector used was a 90 cm diameter *NanoNC DC90*, placed 20 cm from the syringe needle, rotating at 1000 rpm for 10 minutes.

2.2. Microfluidic platform

The microfluidic platform is based on a system created by *Funamoto et al. 2012* that allows us to study cell survival, proliferation, migration, morphogenesis, and differentiation under controlled conditions [2]. The device is composed of 3 parts: a main cell culture chamber (900 μm wide and, 8300 μm long, and 150 μm high), flanked by two side channels (500 μm wide and 150 μm high), and two sets of optimized posts (200 μm wide and 150 μm high) each located between the cell culture chamber and the side channels. The chip design was made in AutoCAD and sent to CAD Art (Output City), where the mask was fabricated in acetate. The mask design was then transferred using a 4" silicon wafer as substrate and SU-8 2100 resin.

The fabrication process of the chip involves the utilization of soft lithography, a technique that employs stamps, molds, and flexible photomasks to create or replicate structures. Initially, a mixture of polydimethylsiloxane (PDMS, Sylgard 184, Dow Corning) was prepared by combining PDMS base and curing agent in a weight ratio of 10:1. This PDMS mixture was then poured onto a Petri dish containing the wafer with the desired microfluidic chip design. To eliminate air bubbles, the mixture underwent a vacuum-assisted degassing process inside a desiccator for 45 minutes. Subsequently, it was cured by placing it at 65°C for a minimum of 4 hours, allowing the polymer to solidify. The cured PDMS was carefully separated from the mold, and the desired chip designs were cut using a scalpel. Biopsy punches were used to create 1 mm holes for the cell chamber and 6 mm holes for the cell culture medium reservoirs.

Following the fabrication steps, the PDMS chips underwent a thorough cleaning process with soap, water, and compressed air to remove any contaminants. Then, in a controlled clean room environment, the chips were attached to glass covers that were previously patterned with electrospun fibers, as described earlier. Oxygen plasma was employed to activate both the PDMS chips and the glass covers for 30 seconds. By applying gentle pressure with fingers, the activated parts were carefully joined together. Subsequently, the assembled chips were placed on a hotplate at 85°C for 2 hours to allow water evaporation and enhance the bonding strength. Finally, the chips underwent sterilization by exposure to UV light for a minimum of 15 minutes to ensure a sterile working environment.

2.3. Hydrogel synthesis

The pre-crosslinked polymer ECM slurry (HAMA and GelMA) was mixed in Claycomb medium overnight at 41°C to obtain a solution with a final concentration of 4,5% GelMA and 3,5% HAMA. 0,1% w/v LAP photoinitiator solution was prepared and mixed with the polymer solution using a vortex to ensure homogeneity. To cross-link this solution, it uses UV light with a wavelength of 365 nm at 4cm, 15s and 10% lamp power giving an irradiance of 90mW/cm². To study the compression modulus of the hydrogel, uniaxial compression tests were carried out with a preload of 0.1mN and a deformation rate of 20%/min. The equipment used was an Zwick Z0.5 TN equipment (Zwick-Roell, Germany) with a 5N cell.

2.4. Cell culture

The cardiac cells used in this work are HL-1. The HL-1 cell line is a type of cardiac muscle cell line that was derived from the atrial tissue of mouse hearts. The HL-1 cell line has become a widely used tool in cardiac research because it closely mimics the properties of primary cardiac myocytes and is capable of spontaneous rhythmic contractions [3].

Cells were maintained according to the protocol provided by the Claycomb laboratory. Claycomb medium (Sigma) was supplemented with 10% FBS, 100 µg/ml penicillin/

streptomycin, 0.1 mM norepinephrine and 2 mM L-glutamine. Norepinephrine supplementation is key to the maintenance of the contractile phenotype. Cells from passages 7 to 14 were seeded at a density of 1,3x10⁵/cm² in 25 ml flasks (T25) coated with fibronectin/gelatin (5µg/ml). 1 ml of medium per 5cm² of culture area was supplied daily and divided 1:3 every 4 days. HL-1 monolayers were contractile at day 3 after passaging.

Cells were lifted at 80-90% confluence, using 0.05% EDTA trypsin, then 0.25% soybean inhibitor solution to stop the effect of trypsin and centrifuged at 500G for 5 minutes.

Furthermore, HL-1 cardiomyocytes were infected with a calcium indicator (GCaMP6s). GCaMP6s is a genetically encoded calcium indicator used in neuroscience research to monitor neuronal activity [4]. GCaMP6s and other genetically encoded calcium indicators can also be applied to cardiomyocytes to monitor calcium dynamics, important for understanding the electrical and mechanical activity of the heart and studying cardiac function. The genetic construct encoding GCaMP6s can be introduced into cardiomyocytes using techniques such as viral vectors allowing for the expression of the calcium indicator in these cells [5].

2.5. Chip seeding

For cell culture both in 2D and 3D, all the chips used had a layer of electrospinning fibers to achieve better adhesive and proliferative properties. The chips were previously irradiated with UV for 15 minutes to sterilize them. For the lifting of the cells, we waited for 80-90% confluence. 0.05% EDTA trypsin was used, then 0.25% soybean inhibitor solution was used to stop the effect of trypsin and centrifuged at 500G for 5 minutes.

For 2D assays, HL-1 cardiac cells are centrifuged and resuspended in a medium to have a concentration of 12x10⁶ cells/ml. Then 10 µg/ml of the medium with the resuspended cells is introduced into the culture chamber of the chip. The chip is left in the incubator for 1 hour to finally hydrate the channels. For hydration of the channels, one reservoir is filled with medium and aspirated from the other. medium and aspirate from the other end with a 1 ml tip cut to fit the diameter of the reservoir. diameter of the reservoir. The same process is repeated for the other parallel channel. Finally, all reservoirs are filled with 100 µl and the medium is changed daily.

For 3D assays, 960,000 cells were centrifuged in an Eppendorf and resuspended in 72µl of hydrogel mixed with 8µl of LAP at a final concentration of 3x10⁷ cells/ml. Finally, 10 µl of the solution loaded with cells was injected through the chamber inlet and UV was irradiated at a 4cm distance, during 10s, with a power of 9mW/cm². For hydration of the channels, was performed the same way as explained above in the 2D cultures, with 100 µl of medium in each reservoir and completely changed every day.

2.6. Immunofluorescence analysis of tissues

Immunofluorescence was performed using Phalloidin (Cytoskeleton Inc.), a fluorescent marker that selectively binds to actin filaments in cells, specifically targeting F-actin. To label the cell nucleus, DAPI (Sigma Aldrich), a marker that binds to DNA's adenine-thymine-rich regions, was used.

The experimental procedure began by rinsing the cells with sterile PBS for 5 minutes. Two reservoirs on one side of the chip were filled with 60 μ l of sterile PBS, allowing the liquid to flow across the chamber for the distribution of reagents. Subsequently, the cells were fixed with 4% paraformaldehyde for 15 minutes. A second wash was performed using PBS for 5 minutes. To enable the entry of staining agents, the cells were permeabilized with 0.1% Triton X in PBS for 10 minutes. Afterward, the cells were rinsed again with PBS for 5 minutes.

For staining, a solution containing DAPI at a 1:1000 dilution and Alexa Fluor 488-conjugated phalloidin at a 1:150 dilution in PBS was applied to the cells and incubated at room temperature for 1 hour. Finally, the cells were imaged using a confocal microscope, specifically the Nikon eclipse inverted model.

2.7. Alignment analysis

Images obtained by confocal microscopy and SEM were used for alignment analysis. The images were processed with ImageJ software (NIH, Maryland, USA, <http://rsb.info.nih.gov/ij>). Fast Fourier transform (FFT) analysis was used as it allows morphological quantification of the microscopic structural alignment [6]. Subsequently, using the Oval Profile Plot plugin an oval projection was placed on the FFT image and a radial summation of the pixel intensities for each angle between 0° and 360° was performed in increments of 1 [7]. Since the FFT data are symmetric, a pixel sum was performed between 0 and 180°. The degree of alignment that was present in the original image was reflected in the height and overall shape of the peak. The position of the peak in the graph reported the principal angle of the orientation. The FFT data were normalized to a reference value of 0 and plotted in arbitrary units ranging from 0 to 1 [8].

2.8. Calcium analysis

All data were preprocessed using ImageJ and self-developed scripts in Matlab and Python. The analysis was applied at the individual case and network level, both using adapted and optimized code from C. Südhof [9].

3. Results

3.1. Electrospinning and fabrication of microfluidic device

The creation of an aligned PLA mesh was successful achieving a density of 1160 fibres per mm, a fibre thickness of 712 ± 93 nm and an alignment at 94°. The mesh was correctly integrated into the base of the central culture channel of the chip. The main problems of culture

medium leakage from the sides of the chip were solved fitting the mesh as close as possible to the size of the central channel.

3.2. Hydrogel synthesis

The results of the mechanical tests show that the hydrogel reached a stiffness of $7.4k \pm 0.96kPa$, optimal for use as a cardiac scaffold. Experimentally, the loading of the pre-crosslinked polymer onto the chip was successful, preventing the hydrogel from leaking through the side channels of the chip, thus remaining crosslinked and confined in the central chamber after UV exposure.

3.3. 2D anisotropic cardiac tissue

The culture was grown directly on PLA70/30 without any coating. As a result, HL-1 adhered to the fibres without any problems, and an alignment in the direction of the fibres was evident, as can be seen in Figure 1(A), where the F-actin is represented in green (marked by the phalloidin) and in blue the nucleus of the HL-1. To quantify the alignment of the cells, the same process was used as for the fibres. The result is reflected in Figure 1(B), where the intensity of the cells is mostly at 87°.

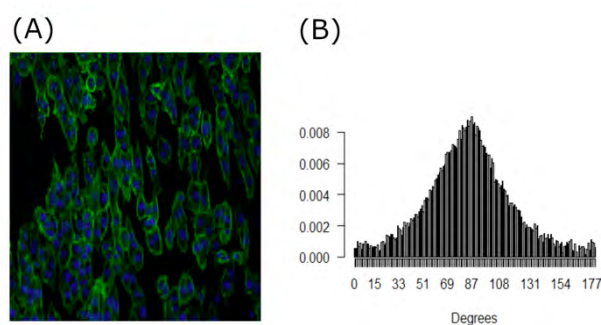


Figura 1. A) DAPI (blue) and phalloidin (green) staining at 20x; B) Radial sum of phalloidin intensities.

In another experimental case, with GCamp6-infected cells, a 30-second video was recorded, Figure 2 (A), from which network-level information could be extracted. It detected 15 synchronous network peaks, then 28.9 peaks/min and an amplitude variation coefficient of 0.05, meaning that the tissue beats with the same intensity at each synchronous contraction, Figure 2(B).

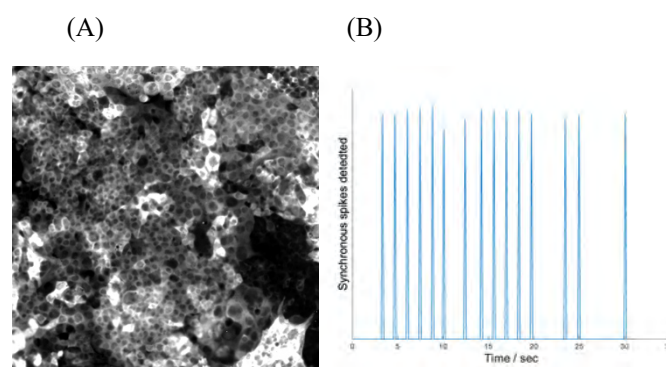


Figura 2. A) 2D HL-1 GCamp6 culture, maximum intensity projection; B) Synchronous spikes detected

3.4. 3D cardiac tissue

A 3D cardiac culture was obtained, where HL-1 cells formed spheroids in most cases, but also showed areas where cells were able to form larger aggregates, Figure 3 (A), thus enabling an interconnectivity study to be performed. The study was applied at the network level, and as a result, parameters such as a synchronous firing rate of 8.14 spikes/min, an amplitude variation coefficient of 0.01, and an inter-spike time of 7.85ms were obtained, Figure 3 (B).

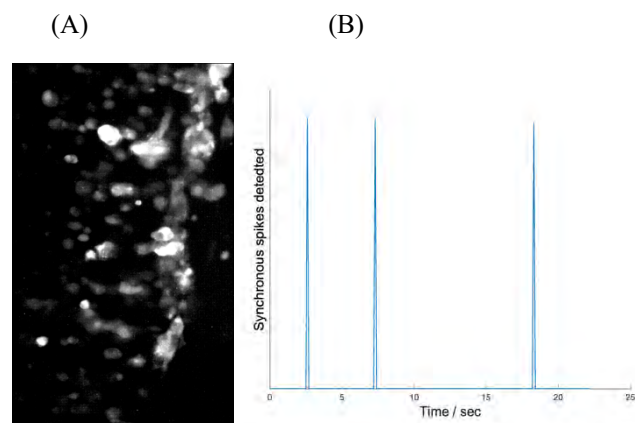


Figura 3. A) 3D HL-1 GCamp6 culture, maximum intensity projection; B) Synchronous spikes detected

The degree of cell viability of this culture was also measured using a life/death test, and the result was 80.9% for the first day.

4. Discussion and Conclusions

We achieved our two objectives, on the one hand, a cell culture reproducing a 2D anisotropic tissue, using PLA fibers aligned by electrospinning, which implies a more mature tissue, [10] where we were able to apply calcium analysis and obtain important parameters for a cellular network synchronization. On the other hand, the results of the hydrogel synthesis show that we have achieved a stiffness of $7.4k \pm 0.96kPa$, which is optimal for use as a cardiac scaffold, subsequently, a 3D myocardial culture was also achieved, this model is currently being optimized to obtain the correct cell density for good growth. In the first 3D assays, cell viability of 80.9% was obtained on the first day, which then decreased; however, in areas with high cell density on the third day, high Ca^{2+} transient connectivity was observed. Based on our results, we are working on inducing a 3D topographic signal to generate 3D-aligned tissue.

The improved results would benefit cardiac research by providing a robust and reproducible microphysiological model for testing regenerative treatments and analyzing them by calcium tracings.

Acknowledgments

We gratefully acknowledge the financial support from the Spanish Ministry of Science and Innovation (MICINN)

and the Spanish State Research Agency (AEI) through grants PID2021-124575OB-I00 and PDC2022-133918-C22.

References

- [1] Z. Álvarez et al., "The effect of the composition of PLA films and lactate release on glial and neuronal maturation and the maintenance of the neuronal progenitor niche," *Biomaterials*, vol. 34, no. 9, pp. 2221–2233, 2013.
- [2] Kenichi Funamoto, Ioannis K Zervantonakis, Yuchun Liu, Christopher J Ochs, Choong Kim, and Roger D Kamm. A novel microfluidic platform for high-resolution imaging of a three-dimensional cell culture under a controlled hypoxic environment. *Lab Chip*, 12(22):4855–4863, November 2012.
- [3] WCClaycomb, NALanson, Jr, BSStallworth, DBEgeland, JBD elcarpio, ABahin-ski, and N J Izzo, Jr. HL-1 cells: a cardiac muscle cell line that contracts and retains phenotypic characteristics of the adult cardiomyocyte. *Proc. Natl. Acad. Sci. U. S. A.*, 95(6):2979–2984, March 1998.
- [4] Tsai-Wen Chen, Trevor J Wardill, Yi Sun, Stefan R Pulver, Sabine L Renninger, Amy Baohan, Eric R Schreiter, Rex A Kerr, Michael B Orger, Vivek Jayaraman, Loren L Looger, Karel Svoboda, and Douglas S Kim. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*, 499(7458):295–300, July 2013.
- [5] YuqianJiang, YuxiaoZhou, XiaopingBao, ChuanxinChen, LaurenNRandolph, Jing Du, and Xiaojun Lance Lian. An ultrasensitive calcium reporter system via CRISPR-Cas9-mediated genome editing in human pluripotent stem cells. *iScience*, 9:27–35, November 2018.
- [6] C. Ayres et al., "Modulation of anisotropy in electrospun tissue-engineering scaffolds: Analysis of fiber alignment by the fast Fourier transform," vol. 27, no. 32, pp. 5524–5534, 2010.
- [7] S. E. Taylor, T. Cao, P. M. Talauliker, and J. Lifshitz, "Objective morphological quantification of microscopic images using a Fast Fourier Transform (FFT) analysis," *Curr. Protoc. Essent. Lab. Tech.*, vol. 2013, no. SUPPL.7, pp. 1–12, 2013.
- [8] Y. Orlova, N. Magome, L. Liu, Y. Chen, and K. Agladze, "Electrospun nanofibers as a tool for architecture control in engineered cardiac tissue," *Biomaterials*, vol. 32, no. 24, pp. 5615–5624, 2011.
- [9] Zijun Sun, and Thomas C. Südhof, "A simple Ca^{2+} -imaging approach to neural network analyses in cultured neurons" doi: 10.1016/j.jneumeth.2020.109041
- [10] López-Canosa, Adrián, et al. "A microphysiological system combining electrospun fibers and electrical stimulation for the maturation of highly anisotropic cardiac tissue." *Biofabrication* 13.3 (2021): 035047.

Lab-on-a-Chip Analysis Using Benchtop NMR Technology

M. Azagra,¹ A. Portela,¹ H. Patel,² D. Weerakonda,² J. Yeste,¹ A. Herrero,¹ M. Fallon,² Megdouda Benamara,³ M. Dubois,³ T. Antonakakis,³ J. Ramon,¹ I. Marco-Rius^{1*}

¹Institute for Bioengineering of Catalonia, Barcelona, Spain; *imarco@ibecbarcelona.eu

²Oxford Instruments, Abingdon, UK

³Multiwave Imaging, Marseille, France

Abstract

We present the design and optimization of a benchtop NMR spectrometer for real-time metabolic monitoring of 3D tissue on microfluidic platforms, utilizing hyperpolarization via dynamic nuclear polarisation. We show the modifications made to a commercial benchtop NMR spectrometer, the design and fabrication of a microfluidic platform ensuring consistent injection of hyperpolarized substrates and ongoing cell media renewal, and its integration with a radio frequency (RF) coil for data transmission and reception (Tx/Rx). Additionally, the construction of a sample carrier is presented. Preliminary NMR results from this system are also provided.

1. Introduction

Modern developments in tissue engineering present innovative methods for metabolic disease modelling and drug research by delivering in vitro models of organs and tissues combined with sensing mechanisms. Such technology could transform the pharmaceutical sector by ensuring trustworthy and predictive in vitro evaluations of chemical entities. The ability to scale down microfluidic systems and refine tissue crafting techniques has allowed scientists to produce numerous 3D in vitro setups with significant control over experiment parameters, tailored for comprehensive screening tasks. In this global and cross-sector collaboration, we have developed a tabletop NMR spectrometer compatible with microfluidic platforms housing 3D in vitro configurations. This allows for real-time, on-site, non-invasive assessment of cellular metabolism using hyperpolarization-enhanced NMR (Figure 1). Integrating this technology with tissue engineering will help track disease progression and gauge reactions to various triggers.

2. Methods

A 1.4 T commercial benchtop NMR spectrometer (Pulsar, Oxford Instruments) has been modified to allow for planar microfluidic chips to be inserted into the scanner. For this purpose, a carrier was 3D printed from a photopolymer resin to accommodate the chip outside the scanner prior to the NMR experiment, improving usability. A microfluidic chip was designed and fabricated specially for this NMR spectrometer with the following requirements: i) science chamber in the homogenous B_0 region of the spectrometer, ii) reliable injection of a hyperpolarized substrate, iii) a constant delivery and renewal of the cell media, iv) and

integrated with a RF coil for transmit/receive. The microfluidic chip was composed by a glass substrate and a single layer of polydimethylsiloxane (PDMS) and fabricated following a standard replica molding protocol¹. The science chamber was designed to be sealed on the top by a removable round cover glass to enable accessing the cell chamber and then sealing the fluidic system once the cells or engineered tissue have been inserted.

Due to the particularities of the geometry of the NMR spectrometer and the direction of its B_0 with respect to the chip plane, a Tx/Rx saddle RF coil was designed to maximise signal-to-noise-ratio and B_1 homogeneity in the science chamber. The design was optimized numerically with CST microwave studio (Finite Element Solver). Several parameters such as the number of turns, the height and the diameter of the coils were optimized to reach the highest average B_1 amplitude in the science chamber. The copper traces were 35 μm thick and placed on 200- μm -thick polyimide substrate to represent accurately the final prototype that would be produced on flexible substrate. For each case the coil was tuned and matched to 15 MHz and B_1 field normalized to 1 W of input power.

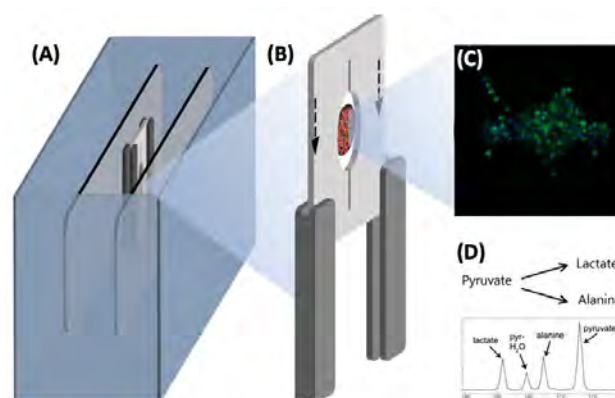


Figure 1. Overview of the work presented, including a benchtop NMR spectrometer for the detection of hyperpolarization-enhanced NMR metabolism in 3D tissue engineered cell models: (A) NMR spectrometer, (B) microfluidics chip with embedded RF coil, (C) 3D liver spheroids, (D) example of metabolic pathway observed upon injected hyperpolarised $[1-^{13}\text{C}]$ pyruvate.

A 3D cell model containing liver spheroids was developed by seeding an immortalised mouse cell line of hepatocytes (AML12) into a carboxymethyl cellulose scaffold (Figure 1C)^{2,3}. Magnetic susceptibility of both PDMS and

carboxymethyl cellulose were measured by waveguide coaxial cell (EpsiMu, Multiwave Technologies, France).

3. Results and Discussion

A PDMS-based microfluidic chip was constructed, featuring channels measuring $200\ \mu\text{m} \times 200\ \mu\text{m}$ that connect to a cylindrical chamber ($5\ \text{mm} \times 5\ \text{mm}$) intended for NMR detection. This setup incorporates a passive membrane pump, designed to transport hyperpolarized solutions. This pump, with a $20\ \text{mm}$ diameter cavity, is topped by an $800\ \mu\text{m}$ PDMS membrane, serving as a damper to diminish pressure surges during hyperpolarized solution injections. The chosen materials ensured no susceptibility artifacts. A saddle RF coil was designed and developed, with a 6-turn structure measuring $5.4\ \text{mm}$ in height and $5.5\ \text{mm}$ in diameter. A scaffold (relative permittivity 80 and conductivity $1.2\ \text{S}\cdot\text{m}^{-1}$) was integrated within the coil to facilitate probe loading. B_1 field amplitude maps depicted the coil's efficiency in the scaffold's vicinity. The scaffold's average amplitude is $1.15\ \text{mT}$ at $15\ \text{MHz}$ for a $1\ \text{W}$ power input, closely matching 75% efficiency of the typical solenoid probe in a standard Pulsar configuration. Within the scaffold volume, the homogeneity displays a relative standard deviation of 20%. Different shimming methods, such as a 3D spatial field mapping analogous to gradshim, were evaluated. These techniques yielded a linewidth under $0.1\ \text{ppm}$.

4. Conclusions

In this study, a cost-effective microfluidic platform has been developed, optimized for NMR applications spanning areas such as biomedical tissue engineering, environmental regulation, and chemical industry operations.

Acknowledgements

This work has received funding from: The European Union's Horizon 2020 research and innovation program (GA-863037); the Spanish grants with reference PID2020-117859RA-I00 funded by MCIN/AEI/10.13039/501100011033 (NARMYD), RYC2020-029099-I funded by MCIN/AEI/10.13039/501100011033 and by "ESF Investing in your future", PLEC2022-009256 funded by MCIN/AEI/10.13039/501100011033 and by the "European Union NextGenerationEU/PRTR"; The BIST (Barcelona Institute of Science and Technology)-"la Caixa" Banking Foundation Chemical Biology programme.

Referencias

- [1] Jo BH et al. Three-dimensional micro-channel fabrication in polydimethylsiloxane (PDMS) elastomer. *J. Microelectromech. Syst.* 2000; 9(1), 76
- [2] Velasco-Mallorquí F, Rodríguez-Comas J, Ramón-Azcón J. Cellulose-based scaffolds enhance pseudoislets formation and functionality. *Biofabrication.* 2021;13, 035044
- [3] Herrero-Gómez A, Azagra M, Marco-Rius I. A cryopreservation method for bioengineered 3D cell culture models. *Biomed Mater.* 2022;17(4).

Versatility of Activated Vapor Silanization (AVS) functionalization for biomedical applications

A. Álvarez-López¹, R. Tabraue Rubio¹, L. Colchero¹, G.V. Guinea¹, J. Pérez-Rigueiro¹, D. González-Nieto¹

¹Center for Biomedical Technology, Universidad Politécnica de Madrid, Pozuelo de Alarcón (Madrid), Spain
{a.aa.alvarez, raquel.tabraue, luis.colchero, gustavovictor.guinea, jose.perez, daniel.gonzalez}@ctb.upm.es

Abstract

The demand for new engineered materials encourages the development of surface modification techniques that allow the creation of tailored cell-material interphases. These interfaces should emulate the natural properties of the extracellular matrix, supporting cellular functions such as adhesion, migration, proliferation, and differentiation, among others. Activated Vapor Silanization (AVS) is a functionalization procedure that allows the immobilization of either complete proteins or peptides on a wide range of material surfaces. In this work we explore the versatility of this technique, by immobilizing two peptides (IKVAV and PHSRN), with different properties and effects, on borosilicate glass and Ti-6Al-4V surfaces. Results show that the immobilization is stable and resistant to chemical denaturing and that the IKVAV peptide enhances cell adhesion. Collectively, this functionalization strategy can accelerate the development of different peptide-decorated metallic and ceramic surfaces for biomedical applications.

1. Introduction

Current developments on surface modification techniques have allowed the engineering of cell-material interphases that emulate the natural properties of the extracellular matrix (ECM). These functionalized surfaces support essential cellular functions by presenting specific cellular binding motifs, creating more biologically favorable environments [1]. The demand for these new engineered materials arises from the low performance of traditional biomaterials in biomedical applications. For example, titanium and its alloys are commonly used in the orthopedic field because their bulk properties are similar to those of bone than other metals [2]. But despite being well established, they still face the problem of the formation of a connective tissue layer (or capsule) between the bone and prosthesis. This interface prevents the interaction between both elements leading to a loosening of the implant [3], and to the need for a new surgical intervention. Surface functionalization tries to avoid this outcome by providing anchorage for the cells and, subsequently, by controlling their spreading, migration, tissue assembly and differentiation [1]. Among all the functionalization methods, silanization has been widely used for immobilizing several types of molecules (proteins, peptides, or polymers) [4]. The traditional procedure of this technique is carried out by immersing the material in a liquid solution of an organofunctional alkoxyane [very frequently, 3-aminopropyltrietho-

xysilane (APTES)], resulting in an amino-coated layer that can bind covalently biomolecules through crosslinking agents [5]. Despite its efficacy, this procedure lacks reproducibility as it depends greatly on the detailed environmental conditions (such as humidity) [5]. In contrast to the solution phase silanization, Activated Vapor Silanization (AVS) allows the functionalization of a wide range of materials with accurate control of the deposition parameters. This technique, inspired in a combination of silanization and chemical vapor deposition, creates a thin layer of amino groups with high reproducibility. In the past, it has been demonstrated the ability of this technique to covalently immobilize complete proteins, as well as two specific peptides (RGD and fibronectin CS-1), in combination with the crosslinking chemistry EDC/NHS (EDC: N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride / NHS: N-hydroxysuccinimide). In this work, the robustness and versatility of this procedure is further explored by immobilizing two new peptides, IKVAV and PHSRN, onto two different surfaces: borosilicate glass (ceramic model material) and Ti-6Al-4V alloy (metallic model material). The effect of the functionalized surfaces is assessed through the cell culture of mesenchymal stem cells (MSC), a cell lineage with strong repercussion in the field of regenerative medicine.

2. Materials and Methods

2.1. Samples preparation

Ti-6Al-4V samples were processed as described elsewhere [6] and round borosilicate glasses were cleaned using the RCA (Radio Corporation of America) cleaning protocol as described before [7].

2.2. Activated vapor silanization functionalization

The functionalization step was carried out as described in [8]. The parameters were set as: $T_{\text{evap}} = 150$ °C, $T_{\text{act}} = 750$ °C, $P_{\text{Arg}} = 2$ mbar and $t = 20$ min. After functionalization, samples were cleaned as described in [6], but glass samples were not subjected to the sonication step.

2.3. Covalent immobilization of FITC-labeled peptides for stability assays

The peptides, IKVAV (Bachem) and PHSRN (Abyntek), were covalently labeled with fluorescein 5(6)-isothiocyanate (FITC, Sigma-Aldrich) and immobilized

on the substrates following a previously described procedure [9]. To study the stability of the immobilization, samples were incubated with 10 % sodium dodecyl sulfate (SDS) for 24 h to clean the non-covalently bound peptide. The fluorescence of the surface was observed using an inverted microscope (Leica DMI 3000B) before and after the treatment. To quantify the intensity of the signal, the images were analyzed using the ImageJ software. This procedure was performed on the first day (short-term stability), and 2 and 5 months after peptide immobilization (long-term stability).

2.4. Covalent immobilization of peptides for in vitro culture assays

The immobilization step of RGD (Sigma-Aldrich), CS-1 (Bachem), IKVAV and PHSRN was carried out as described before in section (2.3) followed by a thorough cleaning procedure as described in [8].

2.5. Cell lines

The *in vitro* effects of the peptides were assessed using mouse bone marrow mesenchymal stem cells (MSC), in passages 6-15, isolated and expanded as described in a previous work [10].

2.6. Single cells force spectroscopy (SCFS) assays

Triangular tipless cantilevers (PNP-DB, NanoWorld) were functionalized with concanavalin A following an already published procedure [11]. The data sampling and post-processing was performed as described before [12]. In this case only the glass substrates and a contact time of 15 s were used.

2.7. Cell adhesion and inverse centrifugation assays

We studied the progression of adhesion during the first hours after seeding. The intensity of cell attachment with the surface (inverse centrifugation) was determined as described before [12]. The inverse centrifugation assays were only performed on glass substrates.

2.8. Focal points staining

For the detection of the cell-surface adhesion molecule vinculin, MSC were fixed in 4% paraformaldehyde after 24h of seeding. Cells were permeabilized and blocked using a standard immunofluorescence procedure. Afterwards, cells were incubated overnight at 4 °C with the antibody anti-vinculin Alexa Fluor® 488 (sc-25336, Santa Cruz Biotechnology) diluted 1:200. Subsequently, to label filamentous actin, cells were incubated with a solution of 2 µg/mL tetramethyl-rhodamine B isothiocyanate (Phalloidin-TRITC, Sigma-Aldrich) in PBS for 1 h. The nucleus was counterstained using 0.5 µg/mL Hoechst 33258 (Sigma-Aldrich) in PBS for 5 min.

2.9. Cell survival assays

Cell survival was analyzed up to 30 days after seeding. 3×10^3 cells were seeded on each substrate previously placed on 24 well plates. When cells reached full confluence, to avoid dead and detachment, the growth

medium (10 % FBS) was replaced by 1 % FBS medium (low starvation conditions). After 3 and 30 days of incubation, cells were dyed with 1 µl/mL of calcein AM [0.5 µg/µl in dimethyl sulfoxide (DMSO)] and 1 µl/mL of propidium iodide (750 µM in PBS) for 15 min at 37 °C.

2.10. Statistics

Statistical analyses were performed using the software SigmaPlot (Systat, Germany). Data are displayed as the mean ± standard error of the mean (SEM). For stability assays, comparisons were performed using Student's t-test for paired samples. The evaluation between functionalized and bare samples in the SCFS assays was studied using Student's t-test for unpaired samples. In cell studies, a two-way analysis of variance (ANOVA) was performed to assess the differences between groups. In all cases, differences were considered statistically significant for p-values below 0.05.

3. Results and Discussion

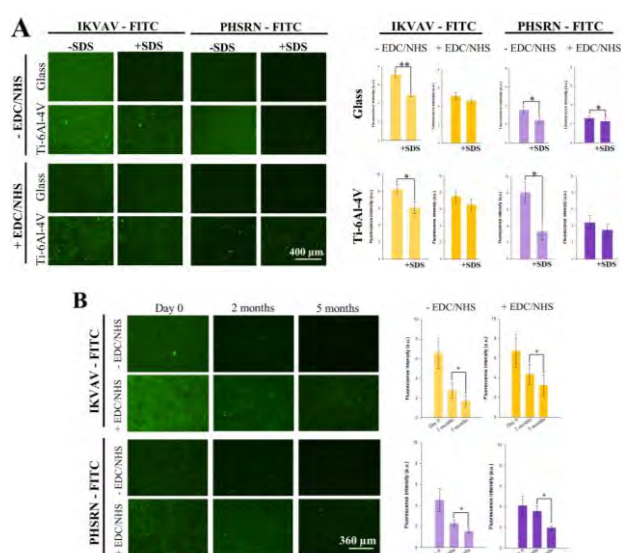


Figure 1. (A) Analysis of the fluorescence drop after SDS treatment at day 1 on glass and Ti-6Al4V (B) Analysis of the fluorescent drop after 2 and 5 months in Ti-6Al-4V (*, $p < 0.05$).

Figure 1.A shows that regardless of the material, the fluorescence drop (indirect measure of peptides attachment to the surface) after SDS treatment is lower on the immobilization process mediated by crosslinkers (+ EDC/NHS). Another shared behavior is the higher ratio and stability on the surface of IKVAV compared to PHSRN, being even higher in Ti-6Al-4V than glass. **Figure 1.B** illustrates the long-term stability of the peptides immobilization performed in this case in Ti-6Al-4V. The pattern observed on day 0 is maintained, but after 5 months we have a drop even in the samples with EDC/NHS. However, even after this period, the presence of the peptide is still observed on the surfaces, with a higher stability in the case of IKVAV.

Figure 2 shows the results from the SCFS assay performed only with the glass substrates. The detachment force necessary to pull the cell off both peptides is clearly

higher compared to the bare glass. This means that the cell interacts more strongly with both peptide-coated substrates than with the non-coated glass (control).

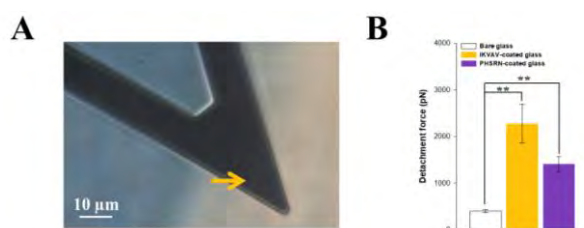


Figure 2. (A) Tipless AFM cantilever with an MSC attached near its edge (arrow). (B) Detachment force values obtained in SCFS (**, $p < 0.01$).

Between both peptides we also found differences in the time course and intensity of cell adhesion.

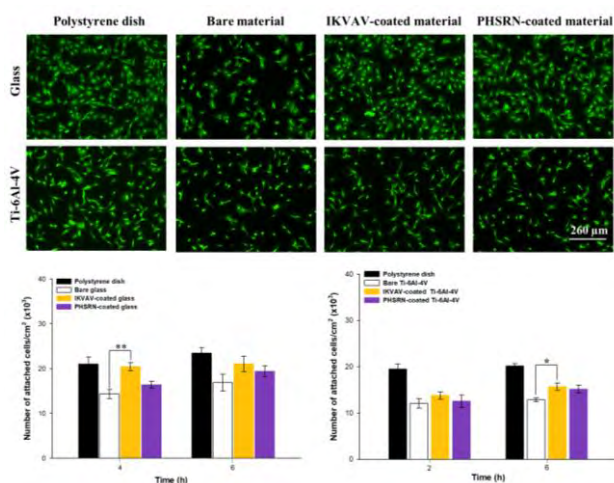


Figure 3. Progression of cell adhesion over the first hours after seeding on glass and Ti-6Al-4V (*, $p < 0.05$; **, $p < 0.01$).

On the adhesion progression assay (Figure 3) we can see a small effect on IKVAV-coated surfaces with respect to the bare substrates. On the other hand, we cannot observe significant differences between PHSRN-coated surfaces and (bare) controls.

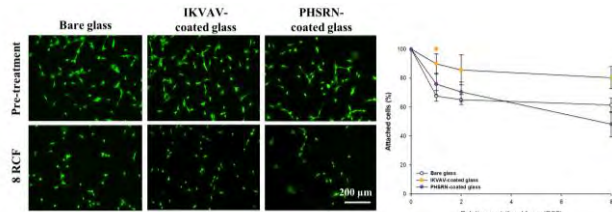


Figure 4. Intensity of cell adhesion on glass substrates. (*, $p < 0.05$, IKVAV vs bare samples).

In the intensity of cell adhesion experiment (Figure 4) we can see differences on the percentage of attached cells after centrifugation between the IKVAV-coated and bare glass samples, although this difference is only statistically significant at 1.0 x g RCF. The PHSRN-coated surface does not modify the intensity of cells attachment. We then assessed the formation of focal adhesion points of the cells to the different surfaces (Figure 5). This assay was performed on four peptides, as we did not have previous information about RGD and CS-1 decorated surfaces,

despite their characterization in previous work. In this study, there are no visible differences between the peptide-decorated or control (bare) samples for each substrate. Instead, there are contrasts depending on the material type. Glass substrates appear to promote a better formation of focal points than Ti-6Al-4V as the vinculin marking is well accumulated on the edge of the cytoskeleton.

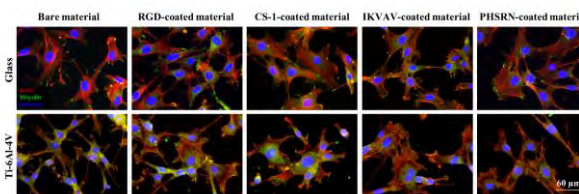


Figure 5. Focal point expression on MSC seeded on functionalized glass and Ti-6Al-4V substrates.

In Figure 6 we show a study of cell survival on peptide-coated glass surfaces. In this case, the functionalization process does not negatively influence cell survival, instead, the cells proliferate and last for 30 days on the material surface.

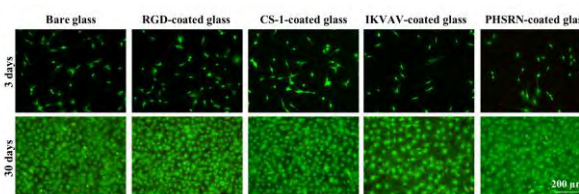


Figure 6. Survival of MSC cells seeded on functionalized glass after 3 and 30 days of culture. Living cells are marked green and dead cells are red.

4. Discussion and conclusion

The aim of this work was to show the versatility of AVS technique by assessing the stability and effect of the covalent immobilization of two new peptides with different biological activities on two different materials. This is a relevant matter as one single peptide with specific properties might not be enough to trigger all the cell signaling paths required for a defined application. In previous works, RGD and CS-1 peptides were immobilized on glass and Ti-6Al-4V using this technique. The results showed the achievement of a stable immobilization and the enhancement of the performance of the materials in *in vitro* cell cultures. The new peptides were chosen based on their described effect on cell behavior. IKVAV has been described as a promoter of adhesion on MSC cells [13] and PHSRN was chosen because of its synergy effect with RGD [14].

During the immobilization step, it has been proved that the use of crosslinkers allows the formation of covalent bonds between the peptides and the surface, providing a higher stability for the long-term. Still, we can observe a decrease in the fluorescence signal that might be due to the non-covalently bound peptides that detach over time, as happens on the samples with no crosslinkers. As this detachment seems to vary depending on the peptide, we have hypothesized that the immobilization process itself

might be affected by the properties of the molecules. We are currently studying if, for instance, the hydrophobicity of the peptides may affect the immobilization quality as more hydrophobic peptides such as IKVAV or CS-1 tend to attach in a higher ratio and be more stable over time than hydrophilic ones such as RGD or PHSRN. As for their performance in *in vitro* cell cultures, IKVAV and PHSRN seem to have less effect than RGD and CS-1. IKVAV still shows small effect on cell adhesion enhancement, but PHSRN offers no improvement compared with the bare material. In agreement with previously published data [14], PHSRN by itself does not enhance significantly MSC adhesion with respect to bare titanium, even if an initial interaction can be observed by SCFS [15]. As for the focal point development, we are unsure about why the cells cultures on Ti-6Al-4V are not able to develop strong focal adhesion points as in the glass substrates. We are currently studying if it is due to the mechanical properties of the surfaces or if it is just a slackening on their formation. This is an important matter, as focal adhesion formation is compulsory to activate migration, proliferation, or differentiation [16]. A weak formation of focal points might affect proliferation and survival of MSC on Ti-6Al-4V substrates with respect to glass substrates. These assays are also in progress. In conclusion, the AVS process proves to be a versatile procedure that allows the immobilization of a broad range of peptides in very different materials, such as titanium and glass. This fact highlights the wide range of possibilities opened by this functionalization procedure. For example, glass substrates decorated with bioactive molecules can be used in different preclinical applications to enhance cell stability during cell interaction assays, plasmid transfection or viral transduction for drug screening, while functionalized Ti-6Al-4V can be used on bone prostheses or for the design of biomimetic nanoparticles for drug delivery.

References

- [1] Rahmany MB, Van Dyke M. Biomimetic approaches to modulate cellular adhesion in biomaterials: A review. *Acta Biomaterialia*, vol 9, sup 3, 2013, pp 5431-7 (ISSN: 1742-7061).
- [2] Stich T, Alagboso F, Křenek T, Kovářik T, Alt V, Docheva D. Implant-bone-interface: Reviewing the impact of titanium surface modifications on osteogenic processes in vitro and in vivo. *Bioengineering and Translational Medicine*, vol 7, sup 1, 2022 (ISSN: 2380-6761).
- [3] Hodges NA, Sussman EM, Stegemann JP. Aseptic and septic prosthetic joint loosening: Impact of biomaterial wear on immune cell function, inflammation, and infection. *Biomaterials*, vol 278, 2021, p 121127 (ISSN: 0142-9612).
- [4] Chouirfa H, Bouloussa H, Migonney V, Falentin-Daudré C. Review of titanium surface modification techniques and coatings for antibacterial applications. *Acta Biomaterialia*, vol 83, 2019, pp 37-54 (ISSN: 1742-7061).
- [5] Sypabekova M, Hagemann A, Rho D, Kim S. Review: 3-Aminopropyltriethoxysilane (APTES) Deposition Methods on Oxide Surfaces in Solution and Vapor Phases for Biosensing Applications. *Biosensors*, vol 13, sup 1, 2023 (ISSN: 2079-6374).
- [6] Rezvanian P, Arroyo-Hernández M, Ramos M, Daza R, Elices M, Guinea GV, Pérez-Rigueiro J. Development of a versatile procedure for the biofunctionalization of Ti-6Al-4V implants. *Applied Surface Science*, vol 387, 2016, pp 652-60 (ISSN: 0169-4332).
- [7] Kern W. Evolution of silicon wafer cleaning technology. *Proceedings - The Electrochemical Society*, vol 90, sup 9, 1990, pp 3-19 (ISSN: 0161-6374).
- [8] Rezvanian P, Daza R, López PA, Ramos M, González-Nieto D, Elices M, Guinea GV, Pérez-Rigueiro J. Enhanced Biological Response of AVS-Functionalized Ti-6Al-4V Alloy through Covalent Immobilization of Collagen. *Scientific Reports*, vol 8, sup 1, 2018, pp 1-11 (ISSN: 2045-2322).
- [9] Álvarez-López A, Colchero L, Elices M, Guinea GV, Pérez-Rigueiro J, González-Nieto D. Improved cell adhesion to activated vapor silanization-biofunctionalized Ti-6Al-4V surfaces with ECM-derived oligopeptides. *Biomaterials Advances*, vol 133, 2022, p 112614 (ISSN: 2772-9508).
- [10] Martín-Martín Y, Fernández-García L, Sanchez-Rebato MH, Mari-Buyé N, Rojo FJ, Pérez-Rigueiro J, Ramos M, Guinea GV, Panetsos F, González-Nieto D. Evaluation of Neurosecretome from Mesenchymal Stem Cells Encapsulated in Silk Fibroin Hydrogels. *Scientific Reports*, vol 9, sup 1, 2019 (ISSN: 2045-2322).
- [11] Friedrichs J, Helenius J, Muller DJ. Quantifying cellular adhesion to extracellular matrix components by single-cell force spectroscopy. *Nature Protocols*, vol 5, sup 7, 2010, pp 1353-61 (ISSN: 1750-2799).
- [12] Álvarez-López A, Tabraue-Rubio R, Hernández-Escobar S, Daza R, Colchero L, Rezvanian P, Elices M, Guinea GV, González-Nieto D, Pérez-Rigueiro J. Application of single cell force spectroscopy (SCFS) to the assessment of cell adhesion to peptide-decorated surfaces. *International Journal of Biological Macromolecules*, vol 244, 2023, p 125369 (ISSN: 0141-8130).
- [13] Luo K, Gao X, Gao Y, Li Y, Deng M, Tan J, Gou J, Liu C, Dou C, Li Z, Zhang Z, Xu J, Luo F. Multiple integrin ligands provide a highly adhesive and osteoinductive surface that improves selective cell retention technology. *Acta Biomaterialia*, vol 85, 2019, pp 106-16 (ISSN: 1742-7061).
- [14] Fraioli R, Dashnyam K, Kim JH, Perez RA, Kim HW, Gil J, Ginebra MP, Manero JM, Mas-Moruno C. Surface guidance of stem cell behavior: Chemically tailored copresentation of integrin-binding peptides stimulates osteogenic differentiation in vitro and bone formation in vivo. *Acta Biomaterialia*, vol 43, 2016, pp 269-81 (ISSN: 1878-7568).
- [15] Mardilovich A, Kokkoli E. Biomimetic peptide-amphiphiles for functional biomaterials: The role of GRGDSP and PHSRN. *Biomacromolecules*, vol 5, sup 3, 2004, pp 950-7 (ISSN: 1525-7797).
- [16] Shinde A, Illath K, Gupta P, Shinde P, Lim KT, Nagai M, Santra TS. A review of single-cell adhesion force kinetics and applications. *Cells*, vol 10, sup 3, 2021, pp 1-46 (ISSN: 2073-4409).

Hydrogels with tunable characteristics to simulate the neuronal environment

A. Vilche Mariscal^{1,2}, I. Sousa Pereira⁴, A. Pascual Segura¹, MJ. López-Martínez⁴, E. Xuriguera Martín⁵, Z. Alvarez Pinto², O. Castaño Linares^{1,3,6,7}

¹ Electronics and Biomedical Engineering. University of Barcelona, Barcelona 08028, Spain

² Biomaterials for Neural Regeneration. Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology (BIST), Barcelona 08028, Spain

³ Biomaterials for Regenerative Therapies. Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology (BIST), Barcelona 08028, Spain

⁴ Nanobioengineering Group. Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology (BIST), Barcelona 08028, Spain

⁵ Department of Materials Science and Physical Chemistry, University of Barcelona, Barcelona 08028, Spain

⁶ Institute of Nanoscience and Nanotechnology (IN2UB), University of Barcelona (UB), Barcelona, Spain

⁷ CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Barcelona 08028, Spain

Abstract

The neuronal system is complex by nature, which implies a slow progression in the knowledge of the mechanisms of different diseases, and, therefore, a slow development of therapies, and drugs. However, both neurodegenerative diseases and acute injuries result in the loss of functions of the central nervous system (CNS). The use of hydrogels as bioinks in tissue engineering applied to neuronal tissue allows for mimicking the neuronal physiological environment, regulating cellular behavior, and supporting regeneration. [1] The variation of properties in tunable hydrogels allows the hydrogel to be adapted to different applications, whether for printing a scaffold or for extrusion into a microphysiological device. [2][3] The 3D scaffolds provide the physical support necessary to facilitate cellular function while the use of said hydrogels as an extracellular matrix in microphysiological devices allows long survival, adhesion, and growth facilitated by the neuroprotective nature of the hydrogel.

1. Introduction

The complexity of the central nervous system (CNS) involves a challenge in restoring damage caused by both injury and neurodegenerative diseases.[4] Neural tissue engineering uses polymers such as hydrogels as a solution to restore function in damaged neural tissues.[3][1]

Hydrogels are the main biomaterials used as extracellular matrices. These are formed by long polymer chains crosslinked by different methods that form hydrophilic 3D networks.[5] From natural or synthetic origin, these materials represent a dynamic structure that allows their remodeling by cells through different enzymes, allows differentiation, and mimics the dynamics of poroelasticity.[6][7] In recent years, the most widely used material for use as an extracellular matrix is Matrigel. This has a composition similar to that of the extracellular matrix from tumours and is easy to manipulate, but it shows a great variety of composition from batch to batch,

as is often the case with commercial polymers, and its properties cannot be modified, thus limiting the working flexibility.[8][9] For this reason, hydrogels based on methacrylated gelatin,[10][11][12] methacrylated alginate,[13][14] or methacrylated hyaluronic acid[15][16] are being used, where methacrylation control, among other parameters, allows easy modification of the properties.[11]

The chemical, physical, and biological properties of hydrogels vary depending on the polymer used, its proportion, the degree of substitution, and the degree of cross-linking. Each of its properties can be modified according to the desired application. Achieving tunable hydrogels provides better control of swelling capacity, elastic modulus, and porosity, among other properties.

The brain is an organ whose physiology, composition and mechanical properties entail a great technological challenge. In the development of biomaterials that aim to mimic said organ, it must be considered that it is anisotropic and soft. Brain parenchyma contains very little fibrous collagen I, which correlates with stiffness. Therefore, they require low stiffness, around 100Pa to 3kPa depending on the part of the brain. On the other hand, it contains different proteoglycans, glycosylated proteins that bind to water, which means a high water content in the brain, about 73-85% of the total mass.[4][17]

In addition to rigidity, brain tissue exhibits viscoelasticity under stress and dissipates stress over time.[18] The viscosity of the brain increases with age, which is directly related to the loss of neurons in aging since these present greater rigidity than glial cells and there is a decrease in the neuron-glia ratio.[17] However, not only does aging present changes in mechanical properties, but different pathologies also generate it. For example, Alzheimer's disease has decreased brain stiffness and multiple sclerosis also has decreased stiffness due to loss of

myelination, the content of which is linearly related to the stiffness. [17]

In addition to mimicking the environment in order to obtain the properties of a healthy brain and with different pathologies, it is crucial to replicate the native environment of neurons and glial cells to obtain adequate differentiation and proliferation since the intrinsic properties of the ECM guide stem cell differentiation in the absence of biochemical signals. [17][17]

This work aims to study a variety of hydrogels such as inks and bioinks that allow mimicking the native physiological environment of the central nervous system by regulating cellular behavior and supporting the regeneration of damaged nervous tissues. Furthermore, a brain co-culture system is developed in a microphysiological device that houses the different types of neuronal cells embedded in a hydrogel as an extracellular matrix.

2. Methods

2.1. Functionalization of natural polymers

GelMa: Methacrylic anhydride was added to Gelatin porcine drop by drop at 50°C for 1h. The methacrylated product GelMa was precipitated in previously (50°C) PBS followed by dialysis against water for 72h at 40°C. Finally, GelMa was lyophilized. [11]

HAMA: Methacrylic anhydride was added to HA (molecular weight 1000 kDa) (1%w/v) drop by drop at 4°C overnight. The reaction pH was maintained between 8 and 11. The methacrylated product HAMA was precipitated in ethanol followed by dialysis against water for 72h. Finally, HAMA was lyophilized. [19]

AlMa: The carboxidiimide reaction was used to activate the carboxyls of the alginate and functionalized with 2-Amino-N-methyl-acetamide (AEMA). The methacrylated product AlMa was precipitated in acetone followed by dialysis against water for 72h. Finally, AlMa was lyophilized. [20]

2.2. Synthesis of the hydrogels

The hydrogel is formed by mixing the components at 41°C overnight. To this solution, 0.05% (w/v) of the LAP photoinitiator is added and it is crosslinked using UV light.

Bioinks	%	% DS.	G'(Pa).
GelMa+AlMa+HA	7,5	40% Gelma	1000
GelMa+Hama	7,25:0,25	40% Gelma	2480
GelMa+Hama	7,0:0,5	40% Gelma	2700
GelMa+Hama	6,5;1	40% Gelma	3000
GelMa+Hama	6,25:1,25	40% Gelma	4000
GelMa	7,5	40	10000
HAMA	1	40	65

HAMA+ECM	90-10	40	48
HAMA+ECM	50-50	40	15
HAMA+ECM	10-90	40	5
ECM	6mg/ml		1,2
ECM	30mg/ml		5,6
ECM	30mg/ml*		80

Tabla 1. Bioinks composition.

2.3. Mechanical characterization of the hydrogels

Rheology: For the rheological measurements, the Discovery HR, TA instruments with the UV complement were used. 20mm PMMA discs were used, since low moduli between (106-101) Pa were expected, with an initial gap of 300µm, to ensure irradiance to the entire sample. It was measured at room temperature since the bath 10µl of the hydrogel with the crosslink was deposited by pipetting. The experimental conditions were 1 Hz and 1% strain.

Water absorption: Water absorption trends were obtained for hydrogels. The specimens were synthesized as cylindrical samples with 6mm of diameter. They were immediately weighed under dry conditions to obtain W_i and placed in 48-well plates with 1mL of 1x PBS. Measurements were collected at different times.

Degradation: Type II collagenase is an MMP8 metalloprotease capable of degrading collagen. Samples were weighed and placed in Eppendorf containers with 1x PBS until they swelled. The 1x PBS was removed and a solution of type II collagenase (Sigma-Aldrich) was added at a concentration of 1U/ml dissolved in 1x PBS. The Eppendorfs were incubated at 37°C and measured at different times.

Porosity: To study the microstructural properties of the 3D network of the hydrogel, it was stained. For this, a 650nm fluorophore (650F) was dissolved in dimethylformamide DMF at a concentration of 1mg/mL was used. This was mixed 1:100, (650F: Polymer) to obtain a concentration of 1% by weight. The solution was mixed in the vortex and left to react for 1h in the bath at 41°C, covered with aluminum foil so that the ester groups react with the amine groups of GelMa. Subsequently, it was mixed with the photoinitiator solution, introduced into the chip and crosslinked.

2.4. Microphysiological device fabrication

Microphysiological device was designed and fabricated using polydimethylsiloxane (PDMS) attached to a glass coverslip with a central channel and four reservoirs for cellular medium.

2.5. Brain Co-culture in a microphysiological device

For in vitro studies, primary astrocytes were embedded in the hydrogel and introduced into the central channel. The interconnected lateral channels facilitated medium flow and diffusion into the hydrogel via a gradient mechanism.

Subsequently, human motor neurons were introduced into the reservoirs, and wait to diffuse them to the central channel.

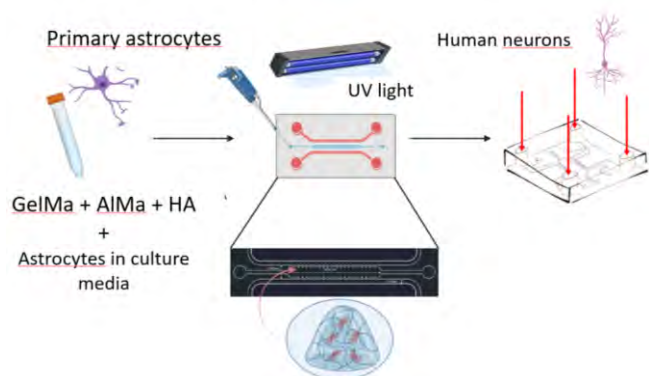


Figura 1. Conceptual sketch of microphysiological device showing the method used for cell introduction.

3. Results

Hydrogels produced from only one polymer show several disadvantages for biomedical applications.[5] Conversely, binary and ternary mixtures can result in hydrogels with better properties that resemble native tissue. In addition, they present a greater number of variables that can be modified. This is extremely important to emulate brain tissue since it does not present the same properties throughout the entire system and from pathology to pathology, it also shows significant differences, such as in the rigidity of Alzheimer's disease.

Hydrogels composed of a single component such as GelMa, HAMA have robust mechanical properties for which they exert a structural behavior. When other components such as AlMa, Hama, and ECM are added to these, their mechanical properties vary in such a way that their modulus decays since they act as plasticizers. This allows us to place hydrogels in the range of properties of the neural environment. More specifically, Table 1 shows how by increasing the proportion of Alma and HAMA in GelMA and ECM in HAMA while maintaining the same final polymer proportion, the module drops significantly.

Specifically in this system, methacrylated gelatin is used because collagen is one of the main components of the extracellular matrix. AlMa is added to the hydrogel to vary the properties, for example, to lower its stiffness and more closely resemble native tissue for the same amount of polymer and HA to improve cell migration and help maintain ECM integrity.[21][13]

The hydrogel used in this system consisted of gelatin methacrylate (GelMa) as the backbone, modified with methacrylate alginate and hyaluronic acid to increase water absorption, reduce stiffness, and improve migration and cohesion. Characterization of the hydrogel revealed desirable physical properties such as excellent water absorption capacity (Figure 2), high porosity (Figure 3), and low modulus (Table 1), which closely resemble the biophysical properties of the CNS.

For in vitro studies, the astrocytes demonstrated long-term viability, attachment, and growth within the device, while the presence of the hydrogel showed a neuroprotective effect compared to the plain device without hydrogel. Subsequently, human motor neurons were introduced into the reservoirs, resulting in the formation of homogeneously distributed neural networks within the hydrogel and in interaction with astrocytes, fostering a brain co-culture system (Figure 4).

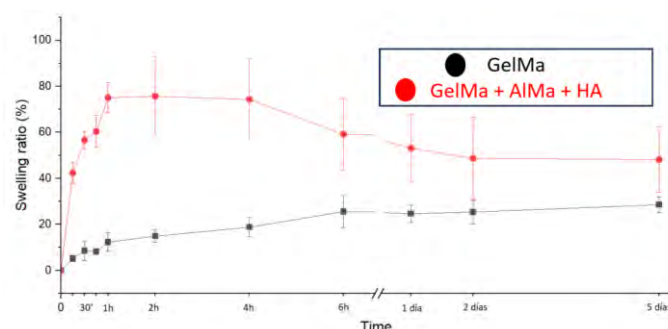


Figura 2. Results of the kinetic trend of water absorption.

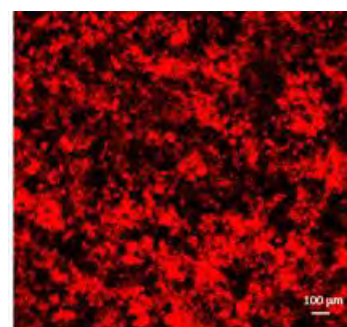


Figura 3. Micrograph of the staining of the chip. Representation of one of the x-y planes of the z-stack.

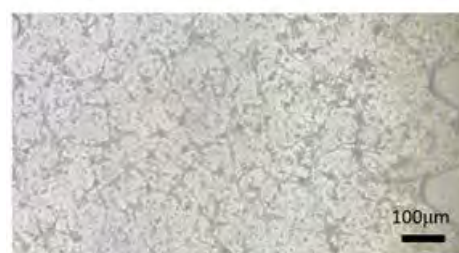


Figura 4. Co-culture astrocytes and neurons in microphysiological device.

4. Conclusions

Tunable hydrogels represent a valuable tool that allows the properties of native tissues to be effectively modelled. Controlling the degree of functionalization, proportions, and cross-linking provides the hydrogels with the desired characteristics. The incorporation of small amounts of polymers such as ALMA and HAMA in the Gelma hydrogel, and ECM in the HAMA hydrogel significantly reduces the mechanical properties by acting as

plasticizers. The hydrogel composed of GelMa AIMA HA selected for use as an extracellular matrix in the model shows high water absorption, low modulus, and high porosity mimicking the properties of the central nervous system. Overall, our results underscored the neuroprotective properties of the hydrogel within the microphysiological device over extended culture periods. This discovery holds great potential as a valuable tool for modelling neural tissue, advancing our understanding of various pathologies, and facilitating the development of innovative treatments.

Acknowledgment

We acknowledge the financial support from the European Union's Horizon Europe research & innovation program (EIC-2021-PATHFINDER-OPEN-01-01-101047099 4DBR) and Grants RTI2018-097038-B-C22, PID2021-124575OB-I00, and PID2021-124839OA-I00 funded by the Ministry of Science and Innovation of Spain (MCINAEI/10.13039/501100011033).

References

- [1] D. R. Nisbet, K. E. Crompton, M. K. Horne, D. I. Finkelstein, and J. S. Forsythe, "Neural tissue engineering of the CNS using hydrogels," *J. Biomed. Mater. Res. - Part B Appl. Biomater.*, vol. 87, no. 1, pp. 251–263, 2008, doi: 10.1002/jbm.b.31000.
- [2] S. J. Lee *et al.*, "Advances in 3D Bioprinting for Neural Tissue Engineering," *Adv. Biosyst.*, vol. 2, no. 4, pp. 1–18, 2018, doi: 10.1002/adbi.201700213.
- [3] P. Madhusudanan, G. Raju, and S. Shankarappa, "Hydrogel systems and their role in neural tissue engineering," *J. R. Soc. Interface.*, vol. 17, no. 162, 2020, doi: 10.1098/rsif.2019.0505.
- [4] E. Axpe, G. Orive, K. Franze, and E. A. Appel, "Towards brain-tissue-like biomaterials," *Nat. Commun.*, vol. 11, no. 1, pp. 10–13, 2020, doi: 10.1038/s41467-020-17245-x.
- [5] Y. S. Zhang and A. Khademhosseini, "Advances in engineering hydrogels," *Science (80-.)*, vol. 356, no. 6337, 2017, doi: 10.1126/science.aaf3627.
- [6] A. Vilalta, J. Sahuquillo, and M. A. Poca, "Matrix metalloproteinases in neurological brain lesions: A new therapeutic target?," *Rev. Neurol.*, vol. 51, no. 2, pp. 95–107, 2010, doi: 10.33588/rn.5102.2009643.
- [7] L. Su *et al.*, "Distinguishing Poroelasticity and Viscoelasticity of Brain Tissue With Time Scale," *Acta Biomater.*, vol. 155, pp. 423–435, 2022, doi: 10.1016/j.actbio.2022.11.009.
- [8] C. S. Hughes, L. M. Postovit, and G. A. Lajoie, "Matrigel: a complex protein mixture required for optimal growth of cell culture.," *Proteomics*, vol. 10, no. 9, pp. 1886–1890, 2010, doi: 10.1002/pmic.200900758.
- [9] E. A. Aisenbrey and W. L. Murphy, "Synthetic alternatives to Matrigel," *Nat. Rev. Mater.*, vol. 5, no. 7, pp. 539–551, 2020, doi: 10.1038/s41578-020-0199-8.
- [10] M. Zhu, Y. Wang, G. Ferracci, J. Zheng, N. J. Cho, and B. H. Lee, "Gelatin methacryloyl and its hydrogels with an exceptional degree of controllability and batch-to-batch consistency," *Sci. Rep.*, vol. 9, no. 1, pp. 1–13, 2019, doi: 10.1038/s41598-019-42186-x.
- [11] A. I. Van Den Bulcke, B. Bogdanov, N. De Rooze, E. H. Schacht, M. Cornelissen, and H. Berghmans, "Structural and Rheological Properties of Methacrylamide Modified Gelatin Hydrogels," *Biomacromolecules*, vol. 1, no. 1, pp. 31–38, 2000.
- [12] Y. Liu and M. B. Chan-Park, "A biomimetic hydrogel based on methacrylated dextran-graft-lysine and gelatin for 3D smooth muscle cell culture," *Biomaterials*, vol. 31, no. 6, pp. 1158–1170, 2010, doi: 10.1016/j.biomaterials.2009.10.040.
- [13] O. Jeon, K. H. Bouhadir, J. M. Mansour, and E. Alsberg, "Photocrosslinked alginate hydrogels with tunable biodegradation rates and mechanical properties," *Biomaterials*, vol. 30, no. 14, pp. 2724–2734, 2009, doi: 10.1016/j.biomaterials.2009.01.034.
- [14] J. Sun and H. Tan, "Alginate-based biomaterials for regenerative medicine applications," *Materials (Basel)*, vol. 6, no. 4, pp. 1285–1309, 2013, doi: 10.3390/ma6041285.
- [15] Y. Shmidov, M. Zhou, G. Yosefi, R. Bitton, and J. B. Matson, "Hydrogels composed of hyaluronic acid and dendritic ELPs: Hierarchical structure and physical properties," *Soft Matter*, vol. 15, no. 5, pp. 917–925, 2019, doi: 10.1039/c8sm02450b.
- [16] J. A. Burdick and G. D. Prestwich, "Hyaluronic acid hydrogels for biomedical applications," *Adv. Mater.*, vol. 23, no. 12, pp. 41–56, 2011, doi: 10.1002/adma.201003963.
- [17] H. N. Kim and N. Choi, "Consideration of the Mechanical Properties of Hydrogels for Brain Tissue Engineering and Brain-on-a-chip," *Biochip J.*, vol. 13, no. 1, pp. 8–19, 2019, doi: 10.1007/s13206-018-3101-7.
- [18] B. M. Maoz, "Brain-on-a-Chip: Characterizing the next generation of advanced in vitro platforms for modeling the central nervous system," *APL Bioeng.*, vol. 5, no. 3, pp. 1–14, 2021, doi: 10.1063/5.0055812.
- [19] D. Aycan, F. Karaca, A. Koca, and N. Alemdar, "Electro-stimulated drug release by methacrylated hyaluronic acid-based conductive hydrogel with enhanced mechanical properties," *Int. J. Biol. Macromol.*, vol. 231, no. September 2022, p. 123297, 2023, doi: 10.1016/j.ijbiomac.2023.123297.
- [20] M. Hasany *et al.*, "Synthesis, properties, and biomedical applications of alginate methacrylate (ALMA)-based hydrogels: Current advances and challenges," *Appl. Mater. Today*, vol. 24, p. 101150, 2021, doi: 10.1016/j.apmt.2021.101150.
- [21] S. J. Franco and U. Müller, "Extracellular matrix functions during neuronal migration and lamination in the mammalian central nervous system," *Dev. Neurobiol.*, vol. 71, no. 11, pp. 889–900, 2011, doi: 10.1002/dneu.20946.

Neurotecnologías (II)

Viernes 24 8:30 - 10:00
Aula Sebastián Ferignan

Evaluación del paradigma de adquisición de la actividad electroencefalográfica en estado de reposo

A. Basterra-García¹, J. Gijón-Ortego¹, V. Gutiérrez-de Pablo^{1,2}, A. Maturana-Candelas^{1,2},
R. Hornero^{1,2,3}, J. Poza^{1,2,3}, C. Gómez^{1,2}

¹ Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, alvaro.basterra@estudiantes.uva.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

³ IMUVA, Instituto de Investigación en Matemáticas, Universidad de Valladolid, Valladolid, España

Resumen

La actividad electroencefalográfica (EEG) espontánea es generada por el cerebro de forma natural en ausencia de estímulos específicos. El registro de la actividad EEG en estado de reposo (RS, *resting-state*) suele realizarse tratando de “no pensar en nada”, con lo cual se induce un estado de divagación mental (MW, *mind-wandering*). Sin embargo, existe debate sobre si los patrones neuronales inducidos son consistentes, lo cual ha motivado que diversos investigadores introduzcan la idea de registrar la actividad RS bajo un protocolo estructurado con un guiado externo (EGRS, *externally guided resting-state*). El objetivo de este trabajo es realizar un estudio comparativo entre los paradigmas RS basado en MW y EGRS, analizando los patrones de activación local. Para ello, se registró la actividad cerebral en 30 sujetos jóvenes cognitivamente sanos, con los ojos cerrados, durante una fase de MW y otra de EGRS. La actividad cerebral se caracterizó mediante diversos parámetros espectrales y no lineales. Los resultados revelaron diferencias estadísticamente significativas entre ambos paradigmas, fundamentalmente en el contenido espectral. Concretamente, se observó una aceleración de la actividad cerebral en el paradigma EGRS en comparación con el RS basado en MW. Las diferencias encontradas entre ambos paradigmas podrían deberse a que el paradigma EGRS se caracteriza por una atención focalizada en un objetivo específico, mientras que durante el RS basado en MW la atención y el nivel de alerta disminuyen.

1. Introducción

La electroencefalografía (EEG) es una técnica no invasiva que permite medir los potenciales postsinápticos generados por millones de neuronas piramidales [1]. Esta técnica es muy empleada en la práctica clínica por su bajo coste, su portabilidad y su buena resolución temporal [2]. La actividad EEG se ha analizado habitualmente usando métodos de activación local. Estos permiten caracterizar la activación sincrónica de poblaciones abundantes de neuronas. Dichas métricas se pueden agrupar en dos categorías principales: espectrales y no lineales. Las primeras evalúan diferentes características de las oscilaciones neuronales en frecuencia. De forma complementaria, las métricas no lineales se centran en propiedades fundamentales de las señales EEG, como la complejidad, la variabilidad y la irregularidad [3].

El análisis de la actividad cerebral espontánea durante el estado de reposo (RS, *resting state*) permite entender una gran cantidad de procesos neuronales subyacentes a diferentes funciones cerebrales [4]. El RS permite observar

las dinámicas cerebrales, especialmente los ritmos alfa, que es el rango de frecuencias con la mayor potencia en el espectro [5]. El RS asocia los procesos cognitivos con la dinámica cerebral intrínseca, lo que permite explicar el funcionamiento de varios mecanismos cerebrales y cómo estos se ven alterados por diversas patologías [5].

El protocolo de adquisición del EEG basado en RS tiene numerosas aplicaciones por su sencillez y comodidad con respecto a los basados en tareas específicas [6]. Desde un punto de vista cognitivo, el RS se puede concebir como un pensamiento independiente del estímulo [7]; es un estado que no está asociado a ninguna tarea específica y, por tanto, es propenso a la divagación mental o “*mind-wandering*” (MW) [8]. En los últimos años se han llevado a cabo multitud de estudios de la actividad cerebral durante el RS, que han revelado la existencia de patrones de activación temporalmente correlacionados, lo que ha dado lugar al concepto de redes en estado de reposo (RSN, *resting state networks*) [9]. El paradigma convencional de RS basado en MW se ha asociado con un nivel bajo de alerta y de vigilia [10], definiéndose como un estado mental caracterizado por una desviación de la atención. Dado que divagación mental ocupa una cantidad significativa del pensamiento despierto, diferentes estudios se han centrado en comprender los procesos neurofisiológicos que subyacen a este estado. Sin embargo, existe debate en la comunidad científica sobre la heterogeneidad y consistencia de los patrones cerebrales evocados en los registros EEG adquiridos mediante este paradigma [11, 12].

En el presente estudio proponemos evaluar la consistencia del paradigma de adquisición de señales EEG en RS basado en MW, contrastándolo con un nuevo paradigma basado en un guiado externo de los diferentes pensamientos que hay que evocar (*i.e.*, “estado de reposo guiado externamente” o *externally guided resting-state*, EGRS). Gracias a ello, pretendemos evaluar la consistencia y estabilidad de los patrones neurofisiológicos de activación local durante el estado de reposo en el paradigma convencional de RS basado en MW.

2. Materiales y métodos

2.1 Participantes

En el estudio se incluyeron 30 sujetos jóvenes cognitivamente sanos (edad $21,3 \pm 0,7$ años, media \pm desviación estándar; 15 mujeres). Padecer cualquier

enfermedad neuronal o el consumo de medicamentos neuropsiquiátricos fueron motivo de exclusión. Todos los sujetos fueron informados sobre el protocolo y propósito del estudio, y dieron su consentimiento informado para participar. El estudio fue aprobado por el Comité Ético del Hospital Clínico Universitario de Valladolid (PI-23-3254).

2.2 Diseño del experimento

Los registros se obtuvieron con los ojos cerrados mediante un protocolo con dos fases: RS basado en MW y EGRS. Previo a los registros, se reprodujo un fragmento de ruido rosa para que los sujetos partieran de una misma situación basal. La fase de RS con MW duró 3 minutos, mientras que la fase de EGRS se estructuró en cuatro etapas para guiar externamente la evocación de pensamientos: (i) pensar en amigos o familia, (ii) imaginarse un paisaje, (iii) pensar en animales, y (iv) sumar mentalmente de dos en dos. Estas instrucciones fueron aleatorizadas entre los sujetos.

2.3 Adquisición y pre-procesado de EEG

Las señales se registraron con un equipo actiCHAMP-PLUS® (BrainProducts) de 32 electrodos situados en base al sistema internacional 10-10 en las posiciones C3, C4, CP1, CP2, CP5, CP6, F3, F4, F7, F8, FC1, FC2, FC5, FC6, Fp1, Fp2, FT9, FT10, O1, O2, P3, P4, P7, P8, T7, T8, TP10, TP9, Fz, Cz, Oz, y Pz. Como referencia se usó el canal Cz. La frecuencia de muestreo fue de 500 Hz. Las señales se preprocesaron para minimizar la presencia de ruido en base a las siguientes etapas: (i) re-referenciado promedio para recuperar el canal Cz; (ii) eliminación de la media de cada canal EEG para eliminar la deriva en el nivel basal de la señal; (iii) filtrado paso banda entre 1 y 70 Hz con un filtro FIR (*Finite Impulse Response*) de orden 2000 con ventana de Hamming y filtrado de ranura para eliminar la interferencia de red a 50 Hz; (iv) segmentación de las señales en cada una de las fases de los registros; (v) análisis de componentes independientes (ICA) para eliminar las componentes relacionadas con ruido; (vi) inspección visual de épocas de 5 s y eliminación de aquellas con artefactos; y (vii) localización de las fuentes neuronales mediante el algoritmo sLORETA (*standardized Low-Resolution Brain Electromagnetic Tomography*), estimando la activación de 15.000 fuentes neuronales que fueron proyectadas en las 68 regiones del atlas Desikan-Killiany [13]. Todo el pre-procesado se realizó con MATLAB® (versión R2021b, MathWorks).

2.4 Procesado de las señales

El contenido espectral de las señales EEG se caracterizó calculando la potencia relativa (RP, *Relative Power*) para cada una de las bandas de frecuencia convencionales: delta (δ , 1-4 Hz), zeta (8-13 Hz), alfa (α , 8-13 Hz), beta-1 (β_1 , 13-19 Hz), beta-2 (β_2 , 19-30 Hz) y gamma (γ , 30-70 Hz). También se calculó la frecuencia mediana (MF, *Median Frequency*), definida como la frecuencia que contiene el 50% de la potencia de la densidad espectral de potencia. Asimismo, se estimó la frecuencia alfa individual (IAF, *Individual Alpha Frequency*), que refleja la frecuencia a la que se encuentra el pico máximo de potencia en la banda alfa [14].

Por otra parte, el análisis no lineal incluyó el cálculo de la complejidad de Lempel-Ziv (LZC, *Lempel-Ziv Complexity*), la entropía muestral (SampEn, *Sample Entropy*), y la medida de la tendencia central (CTM, *Central Tendency Measure*). La LZC estima la complejidad de secuencias finitas de símbolos, asociando valores elevados a series temporales complejas [14]. Por otro lado, la SampEn es una medida de irregularidad, aplicable a series temporales cortas y ruidosas [14]. Finalmente, el CTM es un parámetro que cuantifica la variabilidad de series temporales empleando para ello diagramas de segundo orden [14].

2.5 Análisis estadístico

Para evaluar la presencia de diferencias estadísticamente significativas entre el paradigma de RS con MW y el EGRS, se empleó la prueba no paramétrica de los rangos con signo de Wilcoxon, que se utiliza para determinar si hay diferencias entre dos conjuntos de datos pareados o dependientes [16]. Se aplicó la corrección FDR (*false discovery rate*) para controlar la proporción de falsos positivos entre todas las pruebas realizadas [17].

3. Resultados y discusión

La actividad cerebral en ambos paradigmas se caracterizó, en primer lugar, empleando diferentes parámetros espectrales (ver Figura 1). Inicialmente, se analizó el contenido espectral de la actividad EEG calculando la RP en las bandas de frecuencia convencionales. Únicamente se observaron diferencias estadísticamente significativas ($p < 0.05$) entre el promedio de las fases del paradigma EGRS y el RS basado en MW en las bandas beta-1 y beta-2; en concreto, los valores de RP fueron mayores para el EGRS en comparación con el RS con MW. Estas diferencias en los valores de RP en beta-1 y beta-2 pueden deberse a que dichas bandas de frecuencias están asociadas con el pensamiento activo, la atención y el enfoque en el mundo exterior, además de ser más sensibles a las modificaciones de los estados cognitivos y procesos mentales [1]. Por ello, las diferencias en la RP de estas bandas pueden ser más pronunciadas y detectables en comparación con el resto, que están asociadas a otros procesos cognitivos.

La actividad alfa es dominante durante el RS con ojos cerrados [18], por lo que su elevado valor en ambos paradigmas probablemente sea debido a este hecho. La disminución de RP en delta y el incremento significativo en beta-1 y beta-2 provoca que también lo haga la MF. Este aumento significativo también se observa para la IAF, lo cual puede estar relacionado con un mayor nivel de atención selectiva, concentración y procesamiento cognitivo. Estos resultados sugieren que la banda alfa puede estar más fuertemente involucrada en la coordinación y regulación de las actividades cognitivas superiores durante el EGRS [15], que en este estudio correspondería a imaginar ciertas cosas de manera activa. La aceleración de la actividad neuronal durante el protocolo estructurado podría deberse a una mayor demanda cognitiva en el EGRS y, en consecuencia, a una mayor actividad neuronal en las bandas de alta frecuencia.

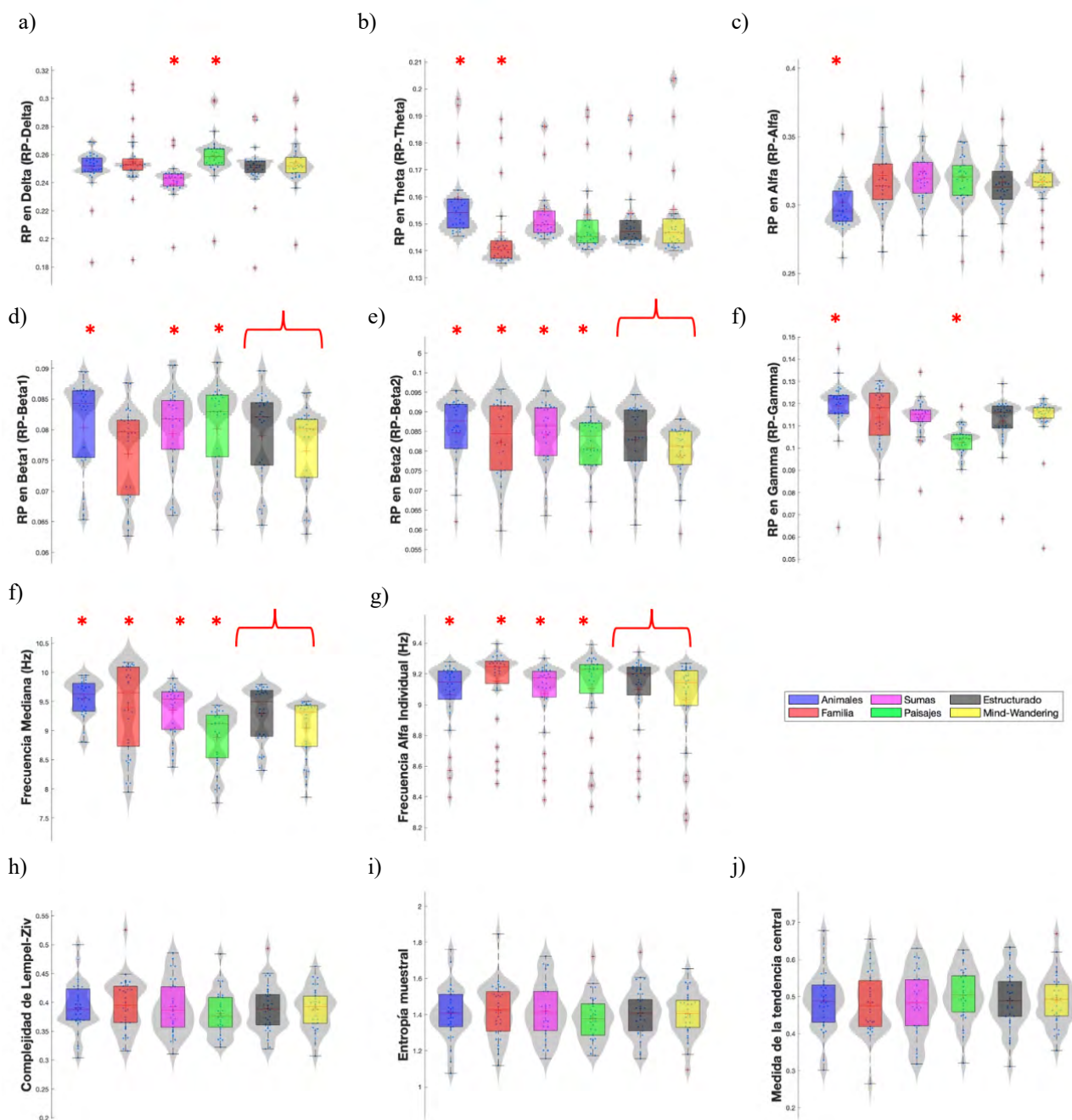


Figura 1. Distribución de las métricas calculadas, divididas en cinco grupos, que representan cada una de las tareas que componen el estado de reposo (RS) estructurado con guiado externo y el basado en “mind-wandering” (MW). Cada distribución representa una característica diferente. (a) RP en la banda delta (RP-Delta). (b) RP en la banda zeta (RP-Zeta). (c) RP en la banda alpha (RP-Alfa). (d) RP en la banda beta-1 (RP-Beta1). (e) RP en la banda beta-2 (RP-Beta2). (f) RP en la banda gamma (RP-Gamma). (g) Frecuencia mediana (MF). (h) Frecuencia alfa individual (IAF). (i) Complejidad de Lempel-Ziv (LZC). (j) Entropía muestral (SampEn). (k) Medida de la tendencia central (CTM). Los parámetros que presentan diferencias significativas entre paradigmas están marcados con un asterisco en rojo *, mientras que los parámetros que presentan diferencias entre el promedio del paradigma RS estructurado y el RS basado en MW, están marcados con una llave ($p < 0.05$, test de los rangos con signo de Wilcoxon, corrección FDR).

Por otro lado, el análisis con parámetros no lineales no reveló diferencias estadísticamente significativas en los valores de LZC, SampEn y CTM entre ambos paradigmas (ver Figura 1). Esto indica que los patrones de complejidad, irregularidad y variabilidad globales de la actividad cerebral no difieren significativamente entre el paradigma de RS basado en MW y el EGRS. No obstante, se obtuvieron valores de LZC más altos en el paradigma estructurado, lo cual sugiere una mayor complejidad del

EEG en el EGRS; esto podría indicar que este paradigma surge en presencia de un patrón heterogéneo de activación neuronal, posiblemente debido a una disminución de la sincronía cortical global. De hecho, Szostakiwskyj *et al.* [19] mostraron que la complejidad era mayor durante estados de reposo y fijación visual que durante la realización de tareas específicas. De la misma forma, la SampEn disminuyó durante el MW, lo que coincide con hallazgos encontrados en otros estudios [20]. Todos estos

trabajos sugieren que la pérdida de complejidad de las señales durante el paradigma de RS basado en MW podría deberse a un menor número de generadores corticales y/o una mayor inhibición cortical general [20].

Este trabajo tiene una serie de limitaciones. La principal es el tamaño reducido de la población muestral, que se limita a un rango de edad concreto. Además, los resultados se centran en el cálculo de parámetros de activación local a nivel global; sería conveniente analizar si los patrones de sincronización global también varían entre los paradigmas, así como analizar su distribución espacial.

4. Conclusiones

Nuestros resultados revelaron que los patrones de activación local evocados durante un paradigma EEG en estado de reposo con una estructura guiada externamente presenta diferencias significativas respecto al paradigma convencional de adquisición del EEG basado en divagación mental. Estas diferencias se observan fundamentalmente en el contenido espectral de las señales EEG, pero no cuando se analizan propiedades no lineales. Además, el aumento de la RP en las bandas de frecuencia altas en el paradigma en estado de reposo guiado externamente podría reflejar que éste se caracteriza por una mayor focalización de la atención.

5. Agradecimientos

Este estudio ha sido financiado por el Ministerio de Ciencia e Innovación / Agencia Estatal de Investigación / 10.13019/501100011033, por el Fondo Europeo de Desarrollo Regional (FEDER) ‘Una manera de hacer Europa’, bajo el proyecto PID2022-138286NB-I00, y por el CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto de Salud Carlos III. V. Gutiérrez-de Pablo es beneficiario de un contrato PIF-UVa de la Universidad de Valladolid.

6. Referencias

[1] S. Saney and J. A. Chambers, *EEG Signal Processing*. John Wiley and Sons, 2007.

[2] M. X. Cohen, “Where does EEG come from and what does it mean?,” *Trends Neurosci*, vol. 40, no. 4, pp. 208–218, 2017.

[3] V. Rodríguez-González *et al.*, “Consistency of local activation parameters at sensor- And source-level in neural signals,” *J Neural Eng*, vol. 17, no. 5, 2020.

[4] A. I. Wiesman, J. da Silva Castanheira, and S. Baillet, “Stability of spectral estimates in resting-state magnetoencephalography: Recommendations for minimal data duration with neuroanatomical specificity,” *Neuroimage*, vol. 247, p. 118823, 2022.

[5] J. Mendoza-Ruiz, C. E. Alonso-Malaver, M. Valderrama, O. A. Rosso, and J. H. Martinez, “Dynamics in cortical activity revealed by resting-state MEG rhythms,” *Chaos*, vol. 30, no. 12, 2020.

[6] R. Cassani, M. Estarellas, R. San-Martin, F. J. Fraga, and T.

H. Falk, “Systematic review on resting-state EEG for Alzheimer’s disease diagnosis and progression assessment,” *Dis Markers*, vol. 2018, 2018.

[7] B. A. Diaz *et al.*, “The Amsterdam Resting-State Questionnaire reveals multiple phenotypes of resting-state cognition,” *Front Hum Neurosci*, vol. 7, 2013.

[8] M. E. Raichle and D. A. Gusnard, “Searching for a baseline: functional imaging and the resting human brain,” *Nat Rev Neurosci*, vol. 2, no. 10, pp. 685–694, 2001.

[9] J. Cabral, M. L. Kringelbach, and G. Deco, “Exploring the network dynamics underlying brain activity during rest,” *Prog Neurobiol*, vol. 114, pp. 102–131, 2014.

[10] B. S. Oken, M. C. Salinsky, and S. M. Elsas, “Vigilance, alertness, or sustained attention: physiological basis and measurement,” *Clin Neurophysiol*, vol. 117, no. 9, pp. 1885–1901, 2006.

[11] K. Cnudde, G. Kim, W. S. Murch, T. C. Handy, A. B. Protzner, and J. W. Y. Kam, “EEG complexity during mind wandering: A multiscale entropy investigation,” *Neuropsychologia*, vol. 180, p. 108480, 2023.

[12] Koculak M, Wierzchon M. How much consciousness is there in complexity? *Frontiers in Psychology*, 2022; 13: 983315.

[13] R. S. Desikan *et al.*, “An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest,” *Neuroimage*, vol. 31, no. 3, pp. 968–980, 2006.

[14] J. Poza, C. Gómez, A. Bachiller, and R. Hornero, “Spectral and non-Linear analyses of spontaneous magnetoencephalographic activity in Alzheimer’s disease,” *J Healthc Eng*, vol. 3, pp. 299–322, 2012.

[15] S. J. Ruiz-Gómez *et al.*, “Measuring alterations of spontaneous EEG neural coupling in Alzheimer’s disease and mild cognitive impairment by means of cross-entropy metrics,” *Front Neuroinform*, vol. 12, 2018.

[16] H. B. Mann and D. R. Whitney, “On a test of whether one of two random variables is stochastically larger than the other,” *Ann. Math. Statist.* vol. 18, no. 1, pp. 50–60.

[17] Y. Benjamini and Y. Hochberg, “Controlling the False Discovery Rate: A practical and powerful approach to multiple testing,” *J R Stat Soc Ser B(Methodol.)*, vol. 57, no. 1, pp. 289–300, 1995.

[18] E. Angelakis, J. F. Lubar, S. Stathopoulou, and J. Kounios, “Peak alpha frequency: an electroencephalographic measure of cognitive preparedness,” *Clin Neurophysiol*, vol. 115, no. 4, pp. 887–897, 2004.

[19] J. M. H. Szostakiwskyj, S. E. Willatt, F. Cortese, and A. B. Protzner, “The modulation of EEG variability between internally- and externally-driven cognitive states varies with maturation and task performance,” *PLoS One*, vol. 12, no. 7, 2017.

[20] Y. Lu and J. Rodríguez-Larios, “Nonlinear EEG signatures of mind wandering during breath focus meditation,” *Curr. Resr. Neurobiol.*, vol. 3, p. 100056, 2022.

Modelo basado en redes neuronales convolucionales y altamente conexas para la ayuda al diagnóstico de pacientes en coma tras paro cardíaco

R.T. Ors Quixal¹, E. Ramírez Candela¹, S. Ruipérez Campillo¹, J.A. Pérez-Sancho¹, F. Sales Ramón¹, J. Millet^{1,2}

¹ Universitat Politècnica de València, teo.ors@gmail.com.

² Centro de Investigación Biomédica en Red Enfermedades Cardiovascular

Resumen

Muchos de los pacientes que originalmente sobreviven a una resucitación posterior a un paro cardíaco, ingresan en la unidad de cuidados intensivos en estado de coma. En este estado, el diagnóstico realizado conlleva mantener o retirar el sistema de soporte vital del paciente, por lo que un pronóstico desfavorable equivocado conlleva condenar a un paciente que podría recuperarse adecuadamente. El propósito de este estudio es desarrollar un modelo de ayuda al diagnóstico basado en aprendizaje automático que, utilizando señales de electroencefalografía de pacientes en estado de coma posterior a un paro cardíaco, pueda distinguir entre aquellos que experimentarán una recuperación favorable y aquellos que no lo harán. Para ello se propone un modelo basado en redes neuronales convolucionales 3D y en altamente conexas, con el objetivo de poder aprovechar la información derivada de la distribución de los electrodos en la superficie de la cabeza. Los resultados obtenidos reflejan la buena capacidad de este tipo de redes para poder realizar dicha clasificación.

1. Introducción

Anualmente, se producen más de seis millones de paros cardíacos en el mundo [1]. Y de entre los pacientes que inicialmente consiguen sobrevivir a la resucitación tras una parada cardíaca, la principal causa de muerte es debida a un daño cerebral severo, encontrándose gran parte de los pacientes que entran en la unidad de cuidados intensivos (UCI) en estado comatoso [2]. El protocolo establecido en estos casos consiste en mantener a los paciente artificialmente con vida al menos tres días, para permitir su recuperación neurológica parcial antes de que el profesional sanitario ejerza un diagnóstico sobre la probabilidad del paciente en coma de recuperar la consciencia [3]. Debe tenerse en cuenta, que un mal diagnóstico significa por una parte, retirar el soporte vital a una persona que en breve podría recuperar la conciencia y por otra, la inversión ineficaz de una gran cantidad de recursos, así como la ocupación de una cama en la UCI, con lo que ello conlleva.

Con el fin de mejorar la toma de decisiones, existen ensayos tipificados como la escala de coma de Glasgow (GCS), que permiten estimar dicho estado en función de parámetros como la respuesta verbal, ocular y motora. A pesar de ello, existen casos en la literatura científica, en los que, tras un pronóstico desfavorable para el paciente, se le mantuvo bajo soporte vital y este pudo recuperarse del coma satisfactoriamente.

Para reducir la subjetividad en el proceso de diagnóstico, se implementó el registro continuo de las señales

de electroencefalografía (EEG) de pacientes en coma. Además, con los avances en tecnología computacional que existen en la actualidad, tenemos la capacidad de desarrollar modelos de aprendizaje automático que pueden manejar grandes volúmenes de información tanto en los datos de entrada como en el tamaño y complejidad de la arquitectura [4, 5].

La decisión de mantener o retirar el soporte vital se toma en base a la Categoría de Rendimiento Cerebral (CRC) en la que el paciente sea clasificado. La escala CRC, se compone de cinco estados de rendimiento cerebral [3]:

- CRC 1: Buen rendimiento cerebral.
- CRC 2: Discapacidad moderada.
- CRC 3: Discapacidad severa (dependencia).
- CRC 4: Estado vegetativo persistente.
- CRC 5: Muerte cerebral.

Si el paciente se encuentra en el grupo uno o dos, de rendimiento cerebral, el soporte vital se mantiene, mientras que si se encuentra en el grupo tres, cuatro o cinco, se aconseja retirar dicho soporte. Este estudio pretende clasificar a los pacientes dentro de estos dos conjuntos: favorable y desfavorable respectivamente. Empleando únicamente sus señales de EEG registradas el tercer día o anteriores. Para ello se emplea un modelo basado en redes neuronales convolucionales tridimensionales (RNC 3D) y en redes neuronales altamente conexas (RNAC), capaz de extraer características automáticamente de las señales, comprimir la información obtenida, a partir de la cual se ofrece una respuesta de pertenencia a cada conjunto. Además, se pretende darle un enfoque ético, priorizando que hayan más falsos positivos de un diagnóstico favorable que falsos negativos [6, 7].

2. Materiales

2.1. Base de Datos

Las señales empleadas en el estudio se extraen del conjunto de datos de la base de datos de la "International Cardiac Arrest REsearch consortium" (I-CARE) construida a partir de señales de pacientes en estado de coma tras un paro cardíaco en diferentes hospitales de Estados Unidos y Europa. De entre la información disponible, se han empleado únicamente las señales de EEG registradas de forma continuada correspondientes a los tres primeros días de registro, que se corresponde con la muestra de la evolución de la señal con la que

cuentan los profesionales sanitarios a la hora de realizar el pronóstico. Todos los pacientes se encuentran etiquetados dentro de las cinco clases de CRC, manteniendo un seguimiento sobre la evolución de los pacientes de entre tres y seis meses después del alta hospitalaria. El número de pacientes perteneciente a cada clase y grupo se muestra en la Tabla 1.

3. Métodos

3.1. Preprocesado de las señales

Las señales de EEG seleccionadas de la base de datos se redujeron a un mínimo de 19 electrodos de acuerdo al sistema internacional 10-20, fueron filtradas digitalmente con filtros paso banda entre 0,5 y 30 Hz, y remuestreadas a 100 Hz. Posteriormente se re-referenciaron las señales a 18 canales bipolares a lo largo del eje longitudinal del montaje, de nasión a inion [8]. Todo esto con el fin de reducir el ruido, tanto de red como de otros artefactos, como los derivados del movimiento de los cables y/o electrodos.

Con el objetivo de que las señales de entrada al modelo mantengan su respectiva distribución en el espacio, análoga a la encontrada en la distribución de los electrodos durante la grabación de los registros, se reestructuraron los canales en tensores siguiendo la plantilla de la Figura 1.

Finalmente, como se puede apreciar en la Figura 2, se fragmentaron cada uno de los tensores de señales de cinco minutos en segmentos de cinco segundos.

Diagnóstico	CRC	Pacientes
Favorable	CRC 1	181
	CRC 2	44
	Total (1, 2)	225
Desfavorable	CRC 3	22
	CRC 4	9
	CRC 5	353
	Total (3, 4 y 5)	225

Tabla 1. Número de pacientes asociados a cada clase

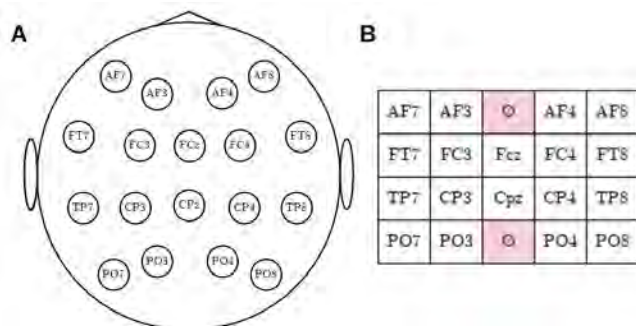


Figura 1. Conjunto de imágenes relativo a la distribución de los canales. A. Representación de los electrodos según su posición diferencial en el montaje 10-10 internacional. B. Plantilla según la cual ordenar los canales.

Todo ello con el propósito de facilitar la paralelización del modelo, reducir la demanda de recursos necesarios tanto para el entrenamiento como para la ejecución del modelo, y aumentar la cantidad de datos disponibles. Contribuyendo a mejorar su capacidad de aprendizaje durante el proceso de entrenamiento.

3.2. Diseño del modelo

El modelo consta de una capa neuronal convolucional tridimensional en la entrada (RNC 3D), una capa aplanadora (Ap) y un conjunto de cuatro capas neuronales altamente conexas (RNAC), siendo la cuarta la salida del modelo. En la Figura 3 se puede ver la arquitectura del modelo, y en la Tabla 2 consultar los valores de los parámetros asociados a su diseño.

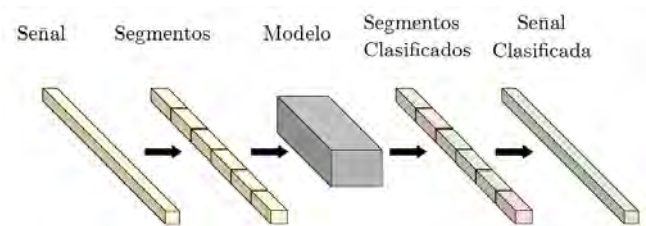


Figura 2. Fragmentación de los registros en diferentes segmentos.

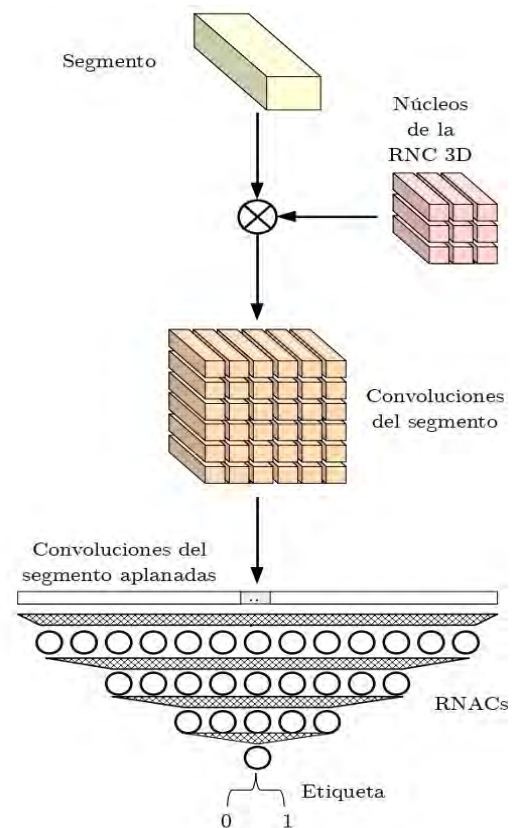


Figura 3. Arquitectura del modelo planteado.

Capa	Tipo	Kernels/Neuronas	Activación
Capa 0	RNC 3D	96*	ReLU
Capa 1	Ap	-	-
Capa 2	RNAC	256	ReLU
Capa 3	RNAC	64	ReLU
Capa 4	RNAC	16	ReLU
Capa 5	RNAC	1	Softmax

Tabla 2. Parámetros asociados a la arquitectura. *Los kernels son de tamaño $3 \times 3 \times 15$ y se desplazan $1 \times 1 \times 5$.

Esta arquitectura fue seleccionada tras una exploración de diversas arquitecturas de modelo posibles [9, 10, 11, 12]. Además, se llevó a cabo una exploración dinámica y automática de sus respectivos parámetros asociados a esta arquitectura, empleando siempre valores de potencia de dos. Para discriminar entre diferentes configuraciones, se analizaron los modelos con mayor exactitud y menor desviación típica en la clasificación.

3.3. Evaluación

Con el objetivo de disponer de más datos de las CRC 3 y 4 para una mejor representación de dicha clase, se extrajeron más horas de dichas clases que del resto en la selección de datos. Además se empleó una descompensación de los datos aportados por cada clase, aportando más (2,25:1) de la clase de CRC 1 y 5, con el fin de que el modelo pudiese ver más segmentos característicos de la clase con diagnóstico más y menos favorable.

Durante el proceso de entrenamiento del modelo, se utilizó el optimizador Adam en combinación con la función de pérdida de entropía cruzada binaria. Se implementó un sistema de monitoreo que finalizaba el entrenamiento alrededor de la época 24, comenzando desde la época 10 y con una memoria de las últimas diez épocas.

Una vez que cada segmento de un paciente fue clasificado, se interpretó el resultado como un voto mayoritario, estableciendo un umbral mínimo del 60% de los votos para un pronóstico desfavorable. Esto se hizo con el propósito de reducir el riesgo de tomar decisiones erróneas al considerar la retirada del soporte vital, especialmente en casos en los que los resultados estuvieran muy equilibrados.

Entre las métricas evaluadas para medir el rendimiento del modelo se incluyen la exactitud, la puntuación F1 y el área bajo la curva característica operativa del receptor (ABC-COR). Estas métricas proporcionan una evaluación completa y robusta de la capacidad del modelo para realizar diagnósticos precisos y respaldan su utilidad como sistemas de ayuda a la toma de decisiones en situaciones clínicas críticas.

4. Resultados

En esta sección, se presentan los resultados obtenidos a nivel de pacientes. Estos resultados se obtuvieron mediante la utilización de una técnica de validación cruzada con k-fold de cinco divisiones, lo que permitió mitigar la influencia de la variabilidad de los datos entre pacientes en los resultados. Al comparar el rendimiento

	Exactitud	Precisión	Recuperación
Segmentos	$0,82 \pm 0,02$	$0,95 \pm 0,01$	$0,87 \pm 0,01$
Paciente	$0,98 \pm 0,01$	$0,97 \pm 0,01$	$0,98 \pm 0,01$

Tabla 3. Valores promedio de las métricas y su desviación típica.

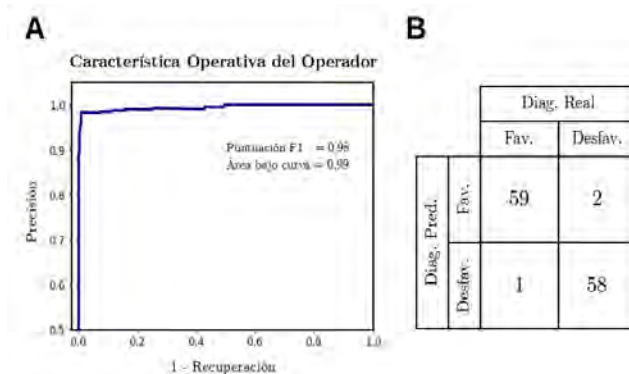


Figura 4. A. Curva ROC de la validación del modelo a nivel de paciente. B. Matriz de confusión..

del modelo a nivel de segmentos frente a nivel de pacientes, se observó una mejora significativa en la exactitud de las clasificaciones, con un aumento del 16%. Además, se logró reducir la desviación estándar entre las diferentes divisiones del k-fold en un 1%.

Como se puede ver en la Tabla 3 y Figura 4, se consiguen clasificar los pacientes con una puntuación de F1 de 0.98, una ABC-COR de 0.99 y una exactitud del 98%. Estos resultados demuestran la eficacia del modelo en la tarea de diagnóstico a nivel de paciente.

5. Conclusiones

La arquitectura de RNC 3D y RNAC ha demostrado ser eficaz para la clasificación de las señales de EEG de pacientes en coma en función de un diagnóstico favorable o desfavorable, con una exactitud del 98%. Además, la estrategia seguida para la fragmentación de la señal ha permitido su clasificación a nivel de segmentos y el voto de estos para la etiqueta global del paciente, que ha conseguido aumentar la exactitud del modelo en un 16%. También se ha cumplido con el enfoque ético propuesto, lo cual es apreciable tanto en la selección del umbral del voto como en los resultados mostrados en la matriz de confusión.

6. Limitaciones

Pese a los buenos resultados, cabe destacar que la existencia de alteraciones en la señal debidas a comorbilidades, como por ejemplo una crisis de ausencia, podría provocar que el modelo, guiase como desfavorable a un diagnóstico que realmente sería favorable. Además el modelo sólo clasifica entre pronósticos favorables o desfavorables, por lo que en un futuro se plantea el desarrollar otro modelo capaz de clasificar según CRC con el objetivo de aumentar la información en la toma de decisiones.

Agradecimientos

Este trabajo contó con el apoyo de PID2019-109547RB-I00 (Programa Nacional de Investigación, Ministerio de Ciencia e Innovación, Gobierno de España) y CIBERCVCB16/11/00486 (Instituto de Salud Carlos III).

Referencias

- [1] Rahul Mehra. Global public health problem of sudden cardiac death. *Journal of Electrocardiology*, 40(6, Supplement 1):S118–S122, 2007.
- [2] Malin Rundgren, Ingmar Rosén, and Hans Friberg. Amplitude-integrated eeg (aeeg) predicts outcome after cardiac arrest and induced hypothermia. *Intensive Care Medicine*, 32:836–842, 6 2006.
- [3] Claudio Sandroni, Alain Cariou, Fabio Cavallaro, Tobias Cronberg, Hans Friberg, Cornelia Hoedemaekers, Janneke Horn, Jerry P. Nolan, Andrea O. Rossetti, and Jasmeet Soar. Prognostication in comatose survivors of cardiac arrest: An advisory statement from the european resuscitation council and the european society of intensive care medicine. *Intensive Care Medicine*, 40:1816–1831, 12 2014.
- [4] Marleen C. Tjepkema-Cloostermans, Catarina da Silva Lourenço, Barry J. Ruijter, Selma C. Tromp, Gea Drost, Francois H. M. Kornips, Albertus Beishuizen, Frank H. Bosch, Jeannette Hofmeijer, and Michel J. A. M. van Putten. Outcome prediction in postanoxic coma with deep learning*. *Critical Care Medicine*, 47:1424–1432, 10 2019.
- [5] Zifan Xu, Jiang Wang, Ruofan Wang, Zhen Zhang, and Shuangming Yang. Deep convolutional neural network for detection of disorders of consciousness. pages 7084–7089. IEEE, 7 2020.
- [6] UNESCO. Recomendación sobre la ética de la inteligencia artificial, 2021.
- [7] Adela Cortina Vicent Botti. Ética e inteligencia artificial, 10 2020.
- [8] Wei-Long Zheng, Edilberto Amorim, Jin Jing, Wengdong Ge, Shenda Hong, Ona Wu, Mohammad Ghassemi, Jong Woo Lee, Adithya Sivaraaju, Trudy Pang, Susan T. Herman, Nicolas Gaspard, Barry J. Ruijter, Jimeng Sun, Marleen C. Tjepkema-Cloostermans, Jeannette Hofmeijer, Michel J.A.M. van Putten, and M. Brandon Westover. Predicting neurological outcome in comatose patients after cardiac arrest with multiscale deep neural networks. *Resuscitation*, 169:86–94, 12 2021.
- [9] Ashima Khosla, Padmavati Khandnor, and Trilok Chand. A comparative analysis of signal processing and classification methods for different applications based on eeg signals. *Biocybernetics and Biomedical Engineering*, 40:649–690, 4 2020.
- [10] Yun Su, Zhixuan Zhang, Xuan Li, Bingtao Zhang, and Huifang Ma. The multiscale 3d convolutional network for emotion recognition based on electroencephalogram. *Frontiers in Neuroscience*, 16, 8 2022.
- [11] Çiğdem Gülüzar Altıntop, Fatma Latifoğlu, Aynur Karayol Akın, and Bilge Çetin. A novel approach for detection of consciousness level in comatose patients from eeg signals with 1-d convolutional neural network. *Biocybernetics and Biomedical Engineering*, 42:16–26, 1 2022.
- [12] Stanley D. T. Pham, Hanneke M. Keijzer, Barry J. Ruijter, Antje A. Seeber, Erik Scholten, Gea Drost, Walter M. van den Bergh, Francois H. M. Kornips, Norbert A. Foudraïne, Albertus Beishuizen, Michiel J. Blans, Jeannette Hofmeijer, Michel J. A. M. van Putten, and Marleen C. Tjepkema-Cloostermans. Outcome prediction of postanoxic coma: A comparison of automated electroencephalography analysis methods. *Neurocritical Care*, 37:248–258, 8 2022.

Análisis de las alteraciones en la estructura neurofisiológica en la migraña

A. Rejas-Llamera^{1,*}, V. Gutiérrez-de Pablo^{1,2,*}, A.L. Guerrero³,
R. Hornero^{1,2,4}, J. Poza^{1,2,4}, C. Gómez^{1,2}

¹Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, victor.gutierrez@uva.es

²Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

³Unidad de Cefaleas, Servicio de Neurología, Hospital Clínico Universitario de Valladolid, España

⁴IMUVA, Instituto de Investigación en Matemáticas, Universidad de Valladolid, Valladolid, España

*A. Rejas-Llamera y V. Gutiérrez-de Pablo han contribuido de manera equitativa para la elaboración de este trabajo

Resumen

La migraña es un trastorno neurológico que se divide en dos subtipos dependiendo de la frecuencia de los ataques al mes: migraña crónica (MC) y episódica (ME). No existen biomarcadores que permitan realizar un diagnóstico preciso. Por ello, el objetivo de este estudio es cuantificar las posibles alteraciones producidas por los subtipos de migraña en la estructura de relaciones formada por las diferentes jerarquías de la actividad cerebral (activación local y sincronización global). Para ello, se ha analizado la actividad electroencefalográfica de 86 sujetos (18 controles, 35 pacientes con MC y 33 con ME). Se han calculado diversas medidas para caracterizar los niveles neurofisiológicos y se han evaluado las correlaciones entre ellas para generar redes de asociación que permitan realizar un análisis completo. Asimismo, se han calculado seis parámetros de red sobre las redes resultantes, analizándolos estadísticamente. Los resultados muestran que las redes de migraña presentan mayores propiedades “small world” con respecto a los controles, mostrando un mayor coeficiente de agrupamiento y una menor longitud de camino. Además, el análisis de entropía y complejidad de grafo presenta otra visión de cómo las redes presentan topologías más óptimas en MC y ME respecto a los sujetos de control. Por último, se ha observado una disminución de la centralidad y aumento de la modularidad en las redes de MC y ME, lo que indica configuraciones más robustas. Dichos cambios podrían deberse a un proceso de reestructuración cerebral mediado por fenómenos de neuroplasticidad, con el fin de mantener la integridad funcional.

1. Introducción

La migraña es una enfermedad de origen neurológico que se caracteriza por un dolor pulsátil, intenso y unilateral, que además supone la incapacitación total o parcial de aquellos que la padecen [1]. Afecta entorno al 12% de la población mundial y tiene mayor incidencia en mujeres que en hombres (18% y 6%, respectivamente), sobre todo en edad fértil [2]. Asimismo, la migraña es el trastorno neurológico que lleva asociados un mayor número de años vividos con discapacidad. Por todo ello, esta enfermedad supone un coste socioeconómico muy elevado [3]. En base al criterio de frecuencia de ataques al mes se distinguen dos tipos de migraña: crónica (15 días o más de dolor de cabeza al mes, durante al menos 3 meses) y episódica (menos de 15 días de dolor de cabeza al mes) [4]. Ambos subtipos presentan una fisiopatología diferente, tanto con respecto a síntomas como a tratamiento [2]. La migraña presenta una sintomatología muy variable, donde los síntomas oscilan entre dolores de cabeza a leves hasta la parálisis durante varios días [1]. Los ataques de migraña pueden ir

acompañados o no de un conjunto de síntomas que pueden ser de carácter visual, sensorial, motores o de lenguaje; y estos se les conoce como aura [2], [4].

La heterogeneidad del cuadro clínico dificulta su diagnóstico, a lo que hay que sumar la ausencia de biomarcadores claros y objetivos [3]. Sin embargo, se ha observado que ambos subtipos de migraña producen alteraciones diferentes a nivel neurofisiológico (*i.e.*, excitabilidad cortical alterada), las cuales son observables mediante diferentes técnicas de registro de la actividad cerebral [3], [5]. En este estudio, se ha empleado la electroencefalografía (EEG), ya que su elevada resolución temporal permite captar la dinámica eléctrica cerebral, y es una técnica no invasiva, de bajo coste y portable [6].

En el análisis de la actividad EEG se distinguen dos niveles neurofisiológicos: activación local y sincronización global [7]. La activación local permite medir características espectrales, dinámicas y no lineales presentes en la actividad EEG registrada en cada región de interés (ROI). El nivel de sincronización global engloba la conectividad funcional (CF) entre las distintas ROIs, así como las propiedades de red y la estructura de la CF multinivel entre diferentes bandas de frecuencia. A nivel de activación local, varios estudios han observado un aumento en la potencia de la banda alfa en las zonas occipital, temporal media y parietal en sujetos con migraña con respecto a controles sanos [8], [9]. Con respecto al nivel de sincronización global, se ha observado un aumento de la coherencia en el EEG en las fases previas a un ataque; sin embargo, en periodos ictal e interictal la coherencia disminuye en pacientes migrañosos con respecto a controles [10].

La hipótesis de este trabajo es que ambos subtipos de migraña alteran la jerarquía de relaciones entre los dos niveles neurofisiológicos previamente mencionados. Por tanto, el objetivo principal es explorar las diferencias en la estructura neurofisiológica entre los grupos de control, MC y ME mediante el uso de redes de asociación.

2. Materiales y métodos

2.1. Sujetos de la base de datos

Se ha registrado la actividad EEG de 86 sujetos, divididos en tres grupos: controles, MC y ME. Pacientes con otros trastornos neurológicos o psiquiátricos, bajo tratamiento preventivo, abuso de medicación o con un historial de

traumas que pudieran afectar a EEG fueron descartados. Pacientes en fase ictal y postictal fueron también excluidos. Los pacientes con migraña finalmente incluidos fueron diagnosticados en base a los criterios del *International Classification of Headache Disorders* [4]. Para los controles, el criterio de inclusión o exclusión se basó en la presencia o historial médico pasado de migraña. Los datos sociodemográficos para cada grupo se muestran en la tabla 1. El Comité Ético local del Hospital Clínico Universitario de Valladolid aprobó el estudio (PI: 17.528).

2.2. Registro y preprocesado de las señales de EEG

La actividad EEG se adquirió con un electroencefalógrafo digital BrainVision® (Brain Products GmbH) de 28 electrodos, colocados en base al sistema internacional 10-10 y con una frecuencia de muestreo de 500 Hz. Se empleó el electrodo Cz como referencia. Se realizó un registro de 10 minutos de duración de la actividad EEG en reposo, mientras los sujetos permanecían sentados, despiertos y con los ojos cerrados para evitar artefactos oculares.

Posteriormente, se preprocesaron los registros con el fin de eliminar ruido y artefactos que no se correspondieran con actividad cerebral relevante. En primer lugar, se realizó un re-referenciado promedio para recuperar la actividad en el electrodo Cz. En segundo lugar, se eliminó la media de las señales para paliar el efecto que el *offset* tiene en estas. Después, se filtraron las señales entre 0.5 y 70 Hz mediante un filtro paso banda de respuesta finita al impulso (FIR) de orden 2000 y con ventana de Hamming. Posteriormente, se realizó un filtrado en 50 Hz con un filtro FIR de ranura (ventana de Hamming y orden 2000) para eliminar la interferencia de la red. A continuación, se eliminaron artefactos mediante un análisis de componentes independientes (ICA). Después, se segmentaron las señales en épocas de 5 segundos y se rechazaron, mediante inspección visual, aquellas con artefactos o ruido de algún tipo. Por último, se localizaron las fuentes neuronales mediante el algoritmo sLORETA (*standardized Low-Resolution Brain Electromagnetic Tomography*); se estimó la activación de 15.000 fuentes neuronales, que fueron proyectadas en las 68 ROIs del atlas Desikan-Killiany.

2.3. Procesado de EEG: extracción de características

En esta etapa se llevó a cabo una extracción de características en los dos niveles neurofisiológicos que conforman la estructura jerárquica global de la actividad EEG: activación local y sincronización global. Los parámetros calculados son los siguientes:

- **Análisis espectral:** potencia relativa (RP) en las bandas de frecuencia delta (0.5-4 Hz), zeta (4-8 Hz), alfa (8-13 Hz), beta 1 (13-19 Hz), beta 2 (19-30 Hz) y gamma (30-70 Hz); frecuencia alfa individual (IAF); frecuencia mediana (MF); frecuencia al borde espectral del 95% (SEF95); frecuencia de transición (TF); entropía espectral (SE); entropía de Tsallis (TsE); entropía de Escort-Tsallis (ETsE); entropía de Rényi (RSE); varianza espectral; y asimetría y kurtosis espectral.
- **Análisis no lineal:** medida de la tendencia central (CTM); complejidad de Lempel-Ziv (LZC); auto información mutua (AMI); entropía aproximada (ApEn); entropía muestral (SampEn); entropía difusa (FuzzyEn); y

Sujetos	Controles	MC	ME
Número	18	35	33
Edad (m [std]) (años)	27.22 [9.44]	42.62 [12.05]	35.96 [12.04]
Sexo (M:H)	8:10	7:28	10:23

Tabla 1. Datos sociodemográficos de la base de datos. *m*: media; *std*: desviación estándar; *M*: mujeres; *H*: hombres.

dimensiones fractales de Higuchi (HFD) y de Katz (KFD).

- **Análisis de la dinámica:** parámetros de Hjorth (movilidad, complejidad y actividad); estadísticos de la señal en tiempo (varianza, asimetría y kurtosis).
- **Análisis de la CF:** *phase-lag index* (PLI) y *amplitude envelope correlation* (AEC).
- **Análisis de red:** grado promedio de nodo; fuerza promedio de los nodos; longitud de camino (PL); coeficiente de agrupamiento (CIC); densidad global; *closeness centrality* (CC); *betweenness centrality* (BC); entropía de grafo (GE); y *small world index* (SW).
- **Análisis de la CF multinivel:** *overlapped weighted degree* (OWD) y coeficiente de participación (P).

2.4. Redes de asociación

Con el fin de caracterizar la estructura de relaciones entre los niveles neurofisiológicos, se generaron las redes de asociación para cada grupo (controles, ME y MC). Estas son una herramienta de fusión de datos, basada en grafos, que permiten caracterizar la correlación existente entre dos o más variables (nodos), entendiendo correlación como la asociación (enlaces) entre ellas. Los sujetos de cada grupo corresponden a instancias que se van a correlar, siendo cada una de las características calculadas para cada sujeto las variables cuyas relaciones se pretende explorar.

Se empleó la correlación de Spearman para cuantificar la asociación existente entre las variables, corrigiendo por las edades de los sujetos. Para cada grupo se calcularon 50 redes de asociación mediante *bootstrap*, una técnica de remuestreo con remplazo. Las redes generadas se corresponden con grafos semiponderados: se aplicó un umbral para mantener únicamente las correlaciones mayores o iguales a 0.8 en valor absoluto, con el fin de aumentar la estabilidad de la red. A continuación, los enlaces fueron normalizados por la densidad global de cada red para eliminar el sesgo que introduce la diferencia de tamaño de las redes. Además, se eliminaron los nodos desconectados del resto de la red para evitar problemas de cálculo. Para representar gráficamente las redes, se usó el algoritmo *Force Atlas 2*, que modela los nodos como si fueran cargas que se repelen en función de su grado de nodo y los enlaces actúan como muelles [11]. En estas redes la anchura del enlace simboliza el grado de correlación, con mayor anchura cuanto mayor es su valor.

Para evaluar las propiedades globales de la red que cuantifican la estructura neurofisiológica, se calcularon varios parámetros que permiten cuantificar su modularidad, segregación (CIC), integración (PL), centralidad (*eigenvector centrality*, EVC), irregularidad

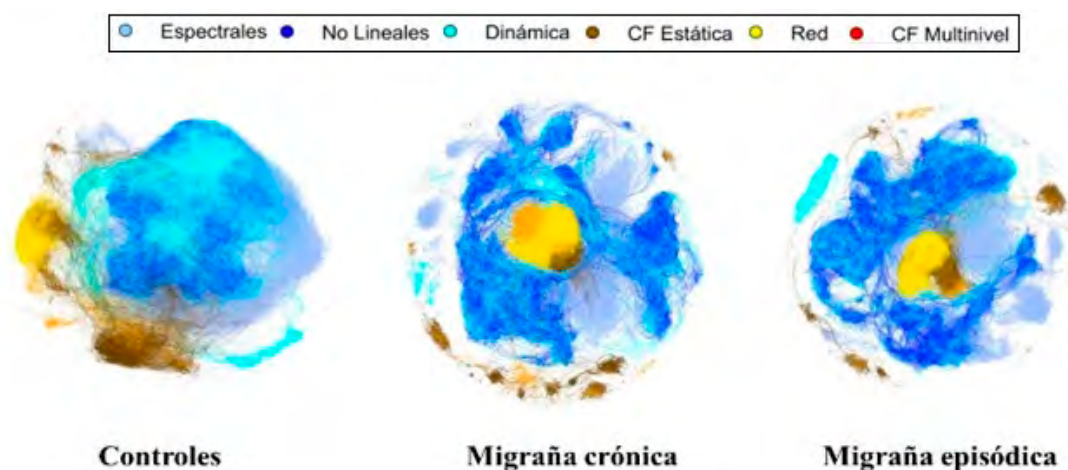


Figura 1. Redes de asociación medianas para cada grupo de sujetos.

(entropía de grafo, GE) y complejidad (complejidad de grafo, GC).

3. Resultados

La figura 1 representa las redes de asociación medianas de cada uno de los tres grupos: controles, MC y ME. Es decir, se muestra el valor mediano de los enlaces de las 50 redes generadas para cada grupo de sujetos. Visualmente, en la red de controles se pueden apreciar dos clústeres formados por medidas de sincronización global y otro de mayor tamaño formado por medidas de activación local. La presencia de ambos subtipos de migraña genera un desplazamiento y agrupamiento de los clústeres de sincronización global hacia el centro del clúster de activación local, lo cual sugiere un aumento en la asociación entre ambos niveles de análisis asociado a la migraña. Además, en ambos grupos de migraña se segregan los clústeres principales de ambos niveles neurofisiológicos formando estructuras de red más modulares.

La figura 2 resume los resultados obtenidos con los parámetros de red para cuantificar los cambios en la estructura neurofisiológica de cada grupo. Todos los parámetros de red han mostrado diferencias entre pares de grupos (p -valor < 0.05 , test U de Mann Whitney, corrección de Bonferroni). En primer lugar, se aprecian diferencias estadísticamente significativas entre todos los grupos en la modularidad, observando un valor mayor en la ME y MC. En segundo lugar, se observa un aumento estadísticamente significativo del CIC en ambos subtipos de migraña con respecto a los controles, tendencia opuesta a la que se observa para el PL y la EVC; estos dos parámetros muestran una disminución estadísticamente significativa en ambos subtipos de migraña con respecto a los controles. Asimismo, la entropía de red presenta un aumento significativo en controles y ME con respecto a MC, presentando una mayor irregularidad los dos primeros grupos. Por último, se aprecian diferencias significativas entre todos los grupos en la complejidad de red (p -valor < 0.05 , test U de Mann Whitney, corrección de Bonferroni), mostrando el grupo de ME una mayor complejidad respecto a los otros grupos.

4. Discusión

Los parámetros de red empleados permiten cuantificar diferentes propiedades de red como integración, segregación, irregularidad, complejidad, modularidad y centralidad. En primer lugar, dado que las redes de sujetos con migraña presentan mayor integración y segregación (PL bajo y CIC alto, respectivamente) que los controles, estas tienen un coeficiente de SW mayor. Esto implica que las redes de ambos subtipos de migraña presentan una configuración de red más cercana a las características SW que los sujetos de control. Las configuraciones SW presentan una topología de red óptima que se da en multitud de sistemas complejos de la naturaleza, y que se caracteriza por un equilibrio óptimo en términos de activación local y comunicación global [12]. Se ha observado que los procesos de cronificación de la migraña favorecen el desarrollo de fenómenos de neuroplasticidad en áreas corticales del procesamiento del dolor [7]. Esto sugiere que los cambios producidos en la estructura neurofisiológica en la migraña podrían ser causados por procesos de reestructuración cerebral para poder mantener la funcionalidad a pesar del desarrollo de la enfermedad.

Sin embargo, también se han observado cambios en la distribución de los pesos de la red medidos con la entropía y complejidad de grafo. La reducción que se ha observado en los valores de entropía muestra una distribución de pesos de la red que converge hacia configuraciones regulares, lo que mostraría que los pacientes, a medida que ven incrementado el número de ataques al mes, presentan configuraciones con mayor diversidad de relaciones entre nodos que los sujetos de control [13]. Por otro lado, el análisis de la complejidad de grafo aporta un matiz a mayores para poder analizar la riqueza estructural de la red [14]. A la hora de evaluar la relación entre complejidad y entropía, se ha llegado al consenso de que sistemas de alta complejidad presentan valores moderados de entropía, minimizándose la complejidad del sistema cuando la entropía es máxima o mínima [14]. Esta perspectiva permite la evaluación de la relación de estas dos medidas mediante una función parabólica [15]. Por tanto, los resultados de entropía y complejidad fortalecen la hipótesis de que la organización de las relaciones de los niveles neurofisiológicos en pacientes de migraña presenta

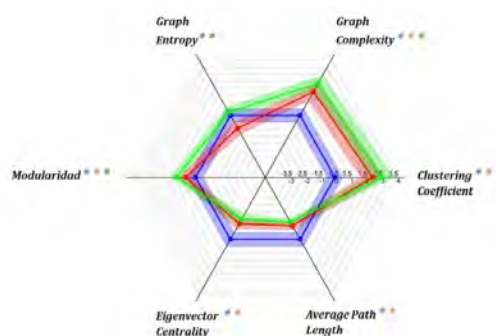


Figura 2. Parámetros de red normalizados respecto a los controles. En azul, controles. En verde, pacientes con ME. En rojo, pacientes con MC. Los asteriscos indican diferencias estadísticamente significativas entre grupos (p -valor < 0.05 , test U de Mann-Whitney, corrección de Bonferroni): en azul, entre controles y MC; en naranja, entre controles y ME; y en verde, entre MC y ME.

configuraciones más próximas a la SW, siendo más prominente esta topología en MC que en ME.

Otra rama del análisis de la organización neurofisiológica a través de sus propiedades de red es la evaluación de la especialización, en base la centralidad y la modularidad. En primer lugar, la centralidad mide cuán centralizada es la topología de la red, incrementándose la vulnerabilidad de esta cuanto mayor centralidad existe [16]. Se ha producido una disminución de la centralidad en ambos subtipos de migraña con respecto a los controles, lo cual indica una red más robusta, con un mayor número de nodos que actúan como “hubs” de la red. Además, se observa un aumento en la modularidad de las redes de migraña. Esto sugiere mayor especialización y pérdida de globalidad cerebral en la migraña, mostrando así una robustez mayor en la organización neurofisiológica en estadios más crónicos de la misma.

Existen una serie de limitaciones que se mencionan a continuación. En primer lugar, las redes han sido generadas usando el valor absoluto de las correlaciones de Spearman, impidiendo obtener información sobre la tendencia de las correlaciones. En segundo lugar, el número de iteraciones de *bootstrap* está limitado a los recursos computacionales disponibles. Por último, no se posee información de los pacientes en función de la fase migrañosa en la que se encuentren (preictal, interictal, ictal o posictal); esto permitiría apreciar cambios en la estructura neurofisiológica de los sujetos en cada uno de los estados de la enfermedad.

5. Conclusiones

Ambos subtipos de migraña producen alteraciones cuantificables en la estructura de relaciones entre los niveles neurofisiológicos. Estos cambios podrían ser producto de una reestructuración cerebral mediada por fenómenos de neuroplasticidad provocados por la aparición de la migraña.

Agradecimientos

Este estudio ha sido financiado por el Ministerio de Ciencia e Innovación / Agencia Estatal de Investigación / 10.13019/501100011033, por el Fondo Europeo de

Desarrollo Regional (FEDER) ‘Una manera de hacer Europa’, bajo el proyecto PID2022-138286NB-I00, y por el CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto de Salud Carlos III. V. Gutiérrez-de Pablo es beneficiario de un contrato PIF-UVa de la Universidad de Valladolid.

Referencias

- [1] V. Anttila *et al.*, *Genetics of migraine*, 1st ed., vol. 148. Elsevier B.V., 2018.
- [2] A. Solano Mora *et al.*, “Actualización de la Migraña,” *Rev. Medica Sinerg.*, vol. 5, no. 4, p. e447, 2020.
- [3] J. Gomez-Pilar *et al.*, “Headache-related circuits and high frequencies evaluated by EEG, MRI, PET as potential biomarkers to differentiate chronic and episodic migraine: Evidence from a systematic review,” *J. Headache Pain*, vol. 23, no. 1, pp. 1–14, 2022.
- [4] J. Olesen, “International Classification of Headache Disorders,” *Lancet Neurol.*, vol. 17, no. 5, pp. 396–397, 2018.
- [5] Y. Li *et al.*, “Abnormalities in resting-state EEG microstates are a vulnerability marker of migraine,” *J. Headache Pain*, vol. 23, no. 1, pp. 1–12, 2022.
- [6] M. Bjørk *et al.*, “What initiates a migraine attack? Conclusions from four longitudinal studies of quantitative EEG and steady-state visual-evoked potentials in migraineurs,” *Acta Neurol. Scand.*, vol. 124, no. 191, pp. 56–63, 2011.
- [7] L. L. H. Pan *et al.*, “Resting State Electrophysiological Cortical Activity: A Brain Signature Candidate for Patients with Migraine,” *Curr. Pain Headache Rep.*, vol. 26, no. 4, pp. 289–297, 2022.
- [8] B. Clemens *et al.*, “Three-dimensional localization of abnormal EEG activity in migraine: A low resolution electromagnetic tomography (LORETA) study of migraine patients in the pain-free interval,” *Brain Topogr.*, vol. 21, no. 1, pp. 36–42, 2008.
- [9] J. Gomez-Pilar *et al.*, “Exploring EEG Spectral Patterns in Episodic and Chronic Migraine During the Interictal State: Determining Frequencies of Interest in the Resting State,” *Pain Med.*, vol. 21, no. 12, pp. 3530–3538, 2020.
- [10] Z. Cao *et al.*, “Resting-state EEG power and coherence vary between migraine phases,” *J. Headache Pain*, vol. 17, no. 1, 2016.
- [11] M. Jacomy *et al.*, “ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software,” *PLoS One*, vol. 9, no. 6, pp. 1–12, 2014.
- [12] C. J. Stam & E. C. W. van Straaten, “The organization of physiological brain networks,” *Clin. Neurophysiol.*, vol. 123, no. 6, pp. 1067–1087, Jun. 2012.
- [13] X. Liao *et al.*, “Small-world human brain networks: Perspectives and challenges,” *Neuroscience and Biobehavioral Reviews*, vol. 77. Elsevier Ltd, pp. 286–300, Jun. 01, 2017.
- [14] M. Costa *et al.*, “Multiscale entropy analysis of biological signals,” *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, vol. 71, no. 2, 2005.
- [15] J. Gomez-Pilar *et al.*, “Quantification of Graph Complexity Based on the Edge Weight Distribution Balance: Application to Brain Networks,” *Int. J. Neural Syst.*, vol. 28, no. 1, pp. 1–19, 2018.
- [16] M. Rubinov & O. Sporns, “Complex network measures of brain connectivity: Uses and interpretations,” *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, 2010.

Análisis múltiple de la red funcional cerebral en sujetos con hiperintensidades de la sustancia blanca

A. Maturana-Candelas^{1,2}, R. Hornero^{1,2,3}, J. Poza^{1,2,3}, V. Jelic^{4,5}, U. Smailovic⁴, C. Gómez^{1,2}

¹ Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, aaron.maturana@uva.es.

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

³ IMUVA, Instituto de Investigación en Matemáticas de la Universidad de Valladolid, Valladolid, España

⁴ Division of Clinical Geriatrics, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Huddinge, Suecia

⁵ Clinic for Cognitive Disorders, Theme Aging, Karolinska University Hospital, Huddinge, Suecia

Resumen

Las hiperintensidades de materia blanca (white matter hyperintensities, WMHs) son indicios de lesiones en el tejido cerebral observables mediante resonancia magnética. Las WMHs son útiles para caracterizar cuadros patológicos tempranos y están asociadas a un incremento del riesgo de diversas dolencias. Sin embargo, el impacto que producen dichas anomalías en la organización de la red funcional cerebral no está plenamente explorado. Con el fin de comprender la disrupción que provocan las WMHs en la habilidad del cerebro para integrar información diversa, se han calculado métricas derivadas de análisis de redes múltiple a partir de señales de electroencefalograma (EEG). Para ello, se adquirió el EEG de 36 participantes con deterioro cognitivo subjetivo, siendo divididos en función de presencia o ausencia de WMHs atendiendo a la escala Fazekas. Posteriormente, se construyeron redes múltiple a partir de los valores de conectividad obtenidos mediante el phase lag index (PLI). El grado de nodo promedio correspondiente a la banda de frecuencia zeta mostró diferencias significativas entre los grupos. Por otro lado, no se obtuvieron diferencias en la conectividad promedio global entre ambos grupos, pero las diferencias de la contribución de cada capa a la red funcional estuvieron próximas al umbral de significancia estadística (p -valor = 0.057, test U de Mann-Whitney). Estos hallazgos sugieren que, incluso en fases preclínicas, pequeños cambios en la integridad de la materia blanca pueden desencadenar alteraciones notables en la organización del conectoma cerebral que podrían estar incluso asociados a la activación de mecanismos de compensación.

1. Introducción

Las hiperintensidades de materia blanca (WMHs) se definen como artefactos de elevado brillo hallados en imágenes de resonancia magnética (RM) cerebral. Estos elementos indican una mayor cantidad de fluido intersticial, probablemente producto de una mayor permeabilidad de la barrera hematoencefálica [1]. Las WMHs revelan lesiones del tejido cerebral vinculadas a desmielinización y la degeneración axonal [2]. A su vez, estas han sido previamente asociadas a deterioro cognitivo. Su presencia triplica el riesgo de ictus y duplica el riesgo de demencia [2]. De hecho, las WMHs están particularmente presentes en individuos de edad avanzada y pacientes con la enfermedad de Alzheimer (EA), siendo así un indicador de gran interés clínico [1]. Por ello, se desarrolló la escala Fazekas, la cual permite cuantificar la frecuencia y tamaño de las WMHs en imágenes de RM, tanto en áreas periventriculares como de sustancia

blanca profunda [3]. Esta escala ha sido validada histopatológicamente y es fácil de realizar, lo que permite un análisis cuantitativo y cualitativo de los WMHs con una buena fiabilidad [4]. La escala comprende entre valor 0 (ausencia absoluta) hasta 3 (áreas extensas y confluentes).

Al ser las WMHs indicadores que que pueden manifestar signos de daño cerebral temprano, determinar su relación con enfermedades que desembocan en demencia, como la EA, es especialmente relevante. Este tipo de trastornos se expresan en un continuo en el cual los síntomas emergen de manera paulatina, desde estadios de deterioro leves hasta una anulación absoluta de cualquier capacidad física y mental. Una fase que puede indicar estados prodrómicos de declive es el deterioro cognitivo subjetivo (DCS) [7]. Este estado preclínico, aunque no conduce necesariamente a la neurodegeneración, es considerado un marcador de riesgo de demencia [7]. La DCS es particularmente interesante, ya que se relaciona con los primeros indicios de declive cognitivo y caracterizar esta fase podría dar lugar a métodos de diagnóstico en estadios más tempranos.

Las enfermedades neurodegenerativas pueden tener una implicación crítica en las relaciones funcionales entre diversas áreas del cerebro. Por ejemplo, la EA es también conocida como un “síndrome de desconexión”, que se caracteriza por una conectividad alterada en las redes neuronales que afecta en gran medida la neurofisiología del cerebro [8]. Concretamente, está relacionada con déficit sináptico y disrupción axonal que da lugar a las alteraciones de conectividad propias de la enfermedad [8]. De hecho, estos elementos están también presentes en individuos con WMHs, sugiriéndose estos como la posible causa de las alteraciones en la conectividad funcional cerebral [9]. Una de las técnicas más empleadas para caracterizar disrupciones de la conectividad cerebral es el electroencefalograma (EEG). El EEG mide las fluctuaciones del campo eléctrico en el cuero cabelludo constituidas principalmente por potenciales post-sinápticos desencadenados en las neuronas piramidales [10]. Las señales de EEG son ampliamente utilizadas para estudiar estas dinámicas debido a múltiples ventajas como la portabilidad, el bajo coste y la no invasividad. Además, su resolución temporal excepcionalmente alta permite estudiar la naturaleza rápida y transitoria de las oscilaciones neuronales.

El estudio de la conectividad funcional cerebral es

uno de los grandes campos de investigación relacionados con la caracterización de estados neurodegenerativos, empleándose diversas métricas con el fin de distinguir funcionamientos saludables y patológicos del cerebro [11]. Uno de los planteamientos más sofisticados derivados de esta perspectiva es el análisis de redes múltiplex (ARM). El ARM permite determinar la correlación o coherencia entre fuentes de EEG en cada banda de frecuencia dispuestas en diferentes capas [12, 13]. Este enfoque lleva a la construcción de redes funcionales dependientes de la frecuencia, que pueden ser inspeccionadas con métricas de la teoría de grafos [12]. Además, el ARM permite cuantificar contribuciones de conectividad en frecuencias específicas asociadas a regiones cerebrales concretas [13].

Debido a que los WMHs pueden estar presentes previamente a signos clínicos de deterioro cognitivo, es fundamental comprender la relación entre estas lesiones y la disrupción de la conectividad funcional cerebral. Así pues, la hipótesis de este trabajo es que la red múltiplex que caracteriza la configuración funcional del cerebro se verá afectada en presencia de WMHs. Con el fin de validar esta hipótesis, se comparará la estructura global del conectoma cerebral en sujetos sin deterioro cognitivo clínico con WMHs y sin WMHs empleando diversas métricas derivadas del ARM. De este modo, será posible comprender la repercusión del daño a la materia blanca en la organización funcional del cerebro, aislando su efecto del provocado por la neurodegeneración. El uso de ARM a partir de EEG en combinación con información clínica sobre lesiones en la materia blanca en sujetos sanos es particularmente novedoso, y supone una perspectiva útil para comprender los cambios funcionales del cerebro en los primeros indicios de neurodegeneración.

2. Materiales y métodos

2.1. Sujetos

En este trabajo se empleó una base de datos compuesta por 36 sujetos con DCS, divididos en 14 sujetos con puntuación 0 de Fazekas y 22 sujetos con puntuación 1 de Fazekas. Los datos EEG de cada sujeto fueron facilitados por el equipo de investigación *Division of Clinical Geriatrics, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet*, Huddinge, Suecia. La evaluación del estado cognitivo se realizó mediante el test *Mini-Mental State Examination* (MMSE) [14]. La adquisición de datos fue aprobado por el Comité Ético del Hospital Universitario Karolinska de Huddinge. Los pacientes dieron su consentimiento por escrito para registrar, almacenar y analizar los datos clínicos. La base de datos electrónica existente que incluye a los pacientes investigados ha sido comunicada a la Junta Sueca de Inspección de Datos y entra dentro de la legislación sobre la Ley de Datos Personales y la Ley de Revisión Ética de la Investigación con Seres Humanos. Los datos sociodemográficos de los participantes se muestran en la Tabla 1.

2.2. Adquisición y pre-procesado del EEG

Los registros de EEG se adquirieron durante 10 minutos en estado de reposo en el Departamento de Neurofisiología Clínica del Hospital Universitario Karolinska-Huddinge. Se colocaron 19 electrodos según el sistema estándar 10-20 con impedancias de electrodo inferiores a 5 kΩ y una frecuencia de muestreo de 256 Hz. Los pacientes estaban sentados en una habitación con luz normal y sonido atenuado, con los ojos cerrados. Un técnico supervisó todo el registro del EEG para alertar al paciente (mediante estimulación acústica) en caso de somnolencia. Se aplicó un filtro paso bajo de 1 a 70 Hz con ventana Hamming de orden 1000. A continuación se segmentó el EEG en épocas de 2 segundos. Finalmente, el EEG fue filtrado en las bandas de frecuencia clásicas: delta (δ , 1-4 Hz), zeta (θ , 4-8 Hz), alfa (α , 8-13 Hz) y beta (β , 13-30 Hz). El procesado y pre-procesado del EEG se realizaron con MATLAB[®] (versión R2021b, Mathworks, Natick, MA).

	<i>N</i>	Edad (años) ($\bar{x} \pm \sigma$)	MMSE ($\bar{x} \pm \sigma$)	Sexo (M:F)
Faz. 0	14	56.43 ± 4.40	28.64 ± 1.50	5:9
Faz. 1	22	59.50 ± 7.01	28.38 ± 1.24	6:16

Tabla 1. Datos sociodemográficos correspondientes a la base de datos empleada. Faz.: puntuación Fazekas; MMSE: puntuación Mini-Mental State Examination; \bar{x} : media; σ : desviación típica; M: masculino; F: femenino.

2.3. Análisis de EEG

Antes de construir las redes múltiplex, se realizó un análisis de conectividad. Entre las numerosas métricas de conectividad que existen en la literatura, el *phase lag index* (PLI) tiene la capacidad de ser relativamente insensible a los efectos de la conducción de volumen [15]. PLI cuantifica la asimetría de la distribución de las diferencias de fase instantáneas entre dos series de la siguiente manera [15]:

$$PLI = |\langle \text{sign}[\sin(\Delta\phi(t_k))] \rangle|, \quad (1)$$

donde $\Delta\phi(t_k)$ indica la diferencia de fase para cada muestra entre dos series temporales y $\langle \cdot \rangle$ diseña el valor medio [15]. Los valores de PLI resultantes oscilan entre 0 y 1, donde 0 indica ausencia de acoplamiento o una diferencia de fase centrada en cero o π ; y 1 indica un acoplamiento de fase perfecto, no centrado en cero [15]. Como resultado de aplicar la métrica PLI, se obtuvo una matriz de adyacencia de tamaño 19x19 para cada época, banda de frecuencia y sujeto, representando todas las conexiones entre pares de electrodos. Se calculó la media de todas las matrices de adyacencia de todas las épocas en cada banda de frecuencia, lo que resultó en 4 matrices (cada una correspondiente a una banda de frecuencia) por cada sujeto. La integración de estas capas dio lugar a la red múltiplex.

Para cuantificar la heterogeneidad de la contribución de conectividad de un nodo a las distintas capas de una

red, se utilizó el coeficiente de participación (P). Este parámetro estima hasta qué punto un nodo asigna sus valores de conectividad a una sola capa o a múltiples capas de forma homogénea. P se ha utilizado previamente para evaluar la alteración de la red funcional en la EA [12, 13]. El primer paso para obtener este parámetro es calcular el grado del nodo i en una capa ψ [13]:

$$s_i^{[\psi]} = \sum_{j=1}^n w_{ij}^{[\psi]}, \quad \psi = \{\delta, \theta, \alpha, \beta\}, \quad (2)$$

donde w_{ij} es el valor de conectividad entre el nodo i y j para la banda ψ que, en nuestro caso, corresponde al valor calculado a partir del algoritmo PLI. Además, $n = 19$, de acuerdo con el número de canales. Cada capa ψ se obtuvo a partir del promedio por épocas de los valores PLI entre cada canal. A continuación, el grado de nodo $s_i^{[\psi]}$ de cada canal se sumó en todas las capas, lo que dio como resultado el grado ponderado superpuesto (G):

$$G_i = \sum_{\psi} s_i^{[\psi]}, \quad \psi = \{\delta, \theta, \alpha, \beta\}, \quad (3)$$

que se calculó para cada uno de los 19 canales de cada sujeto. Por último, se calculó el P para cada nodo i como se muestra a continuación [13]:

$$P_i = \frac{M}{M-1} \left[1 - \sum_{\psi} \left(\frac{s_i^{[\psi]}}{G_i} \right)^2 \right], \quad \psi = \{\delta, \theta, \alpha, \beta\}, \quad (4)$$

donde $M = 4$ corresponde al número de capas, siendo estas la representación de la matriz de conectividad de cada banda de frecuencia. Los resultados obtenidos del cálculo de P pueden oscilar entre 0 y 1. Un valor de P igual a 1 indica valores de conectividad distribuidos homogéneamente entre capas, mientras que $P = 0$ significa que todas las conexiones con el canal se concentran en una sola capa. En este estudio, los valores de G , P y grado de nodo en capas individuales fueron obtenidos promediando por sujetos y épocas, obteniendo un valor promedio por cada canal. De este modo, es posible caracterizar el modelo de configuración funcional del cerebro de cada grupo de estudio. Posteriormente, los datos representativos de cada canal fueron comparados con el fin de establecer diferencias estadísticamente significativas. Debido a que los datos no eran normales ni homocedásticos, se empleó el test no paramétrico U de Mann-Whitney para determinar dichas diferencias entre los dos grupos.

3. Resultados y discusión

Primero, se inspeccionó el grado de nodo promedio de cada banda de frecuencia de manera individual (Tabla 2). La banda zeta fue la única que mostró diferencias estadísticamente significativas (p -valor < 0.001) entre los grupos Fazekas 0 y Fazekas 1, siendo mayor el valor de conectividad en el grupo de Fazekas 1. Este resultado está alineado con las alteraciones de PLI en zeta reportadas anteriormente al comparar sujetos en estadios preclínicos y pacientes con EA [18]. A pesar de que los sujetos en este estudio son preclínicos en su totalidad, sí podría asumirse que el grupo Fazekas 1 se encuentra en un estado más cercano a la patología. Por ello, se sugiere que las redes funcionales que trabajan en la banda zeta podrían ser más sensibles a disrupciones de fase. La potencia en la banda zeta también parece verse afectada por las WMHs, mostrando valores más elevados en sujetos con presencia de las mismas e independientemente de la edad [19]. Estas alteraciones en la potencia de zeta podría tener relación con la disrupción temprana de estas oscilaciones en presencia de WMHs, aunque esta sea mínima.

Adicionalmente, se profundizó el análisis inspeccionando los valores G y P derivados de la red múltiple en su totalidad. Los valores de G no mostraron diferencias significativas entre grupos. Se obtuvo un valor promedio entre sujetos y canales de 16.582 ± 0.170 en el grupo de Fazekas 0 y 16.605 ± 0.169 en el grupo de Fazekas 1. Por otro lado, para los valores de P , se obtuvo un valor de 0.9788 ± 0.0013 para el grupo Fazekas 0; y de $0.9779 \pm 8,716 \cdot 10^{-4}$ en el grupo Fazekas 1. En este caso, el p -valor obtenido se encontró cercano al umbral de significancia estadística (0.057). Este hallazgo implica que, aunque los valores de conectividad en términos globales son similares, no es así en lo que respecta a la contribución de cada capa (cada banda de frecuencia) a la configuración de la red funcional cerebral. Posiblemente, las diferencias obtenidas en la banda zeta pudieran ser la principal contribución a la disparidad en los valores de P entre grupos. Un valor de P más cercano a 1 en el grupo de Fazekas 0 implica una contribución más homogénea por parte de todas las capas. Esto podría indicar una mayor capacidad de procesamiento cerebral integrando información de diversa índole; de hecho, valores mas reducidos de P han sido asociados a una pérdida de especialización de regiones cerebrales específicas [13]. Sin embargo, esta conclusión debe de ser considerada con cautela, ya que es sabido que las fases prodrómicas de la EA se caracterizan por exhibir dinámicas eléctricas

	Delta ($\bar{x} \pm \sigma$)	Zeta ($\bar{x} \pm \sigma$)	Alfa ($\bar{x} \pm \sigma$)	Beta ($\bar{x} \pm \sigma$)
Fazekas 0	5.133 ± 0.038	4.429 ± 0.020	4.562 ± 0.108	2.459 ± 0.080
Fazekas 1	5.150 ± 0.029	4.468 ± 0.024	4.548 ± 0.113	2.438 ± 0.058
p -valor	0.179	< 0.001	0.620	0.414

Tabla 2. Valores de grado de nodo por banda de frecuencia y resultados estadísticos de la comparación entre grupos (prueba U de Mann-Whitney). \bar{x} : media; σ : desviación típica.

con patrones opuestos al resto del continuo [16]. Dicha particularidad ha sido especialmente asociada a los mecanismos de compensación cerebral propios de estadios que preceden a la EA. Se ha reportado evidencia de la capacidad de plasticidad del cerebro para compensar la neurodegeneración precoz en pacientes con EA preclínica, incluso antes de la aparición de los primeros síntomas [16, 17].

A pesar de los hallazgos obtenidos en este estudio, es preciso mencionar ciertas limitaciones. Primero, se cuenta con un número reducido de sujetos con DCS. Un estudio con múltiples participantes de una amplia variedad de estadios permitiría mejorar la caracterización de la relación entre progresión de la EA y WMHs. Asimismo, los valores de Fazekas están muy asociados al deterioro cognitivo, y aislar totalmente el efecto de las WMHs de otros factores neurodegenerativos es inviable. Por ello, sería necesario casos infrecuentes en los que pacientes con EA sin WMHs participasen en subsecuentes estudios con el fin de estudiar su impacto en la electrofisiología cerebral.

4. Conclusiones

En este estudio se ha comprobado la relación entre la presencia de WMHs y la organización de la red funcional cerebral en sujetos preclínicos. El grado de nodo promedio muestra diferencias significativas únicamente en la banda zeta, lo cual revela una mayor sensibilidad de este tipo de oscilaciones a las lesiones asociadas a las WMHs. Ello podría estar asociado a procesos cerebrales más dependientes de la integridad de la estructura axonal. Por otro lado, la conectividad global promedio no se ve afectada entre sujetos con Fazekas 0 y Fazekas 1; sin embargo, sí hay indicios de que la contribución de cada banda de frecuencia sufre alteraciones entre dichos grupos. Esto es evidencia de la relación entre pequeñas lesiones estructurales del cerebro y alteraciones significativas en la red funcional cerebral.

Agradecimientos

Este estudio ha sido financiado por el Ministerio de Ciencia e Innovación / Agencia Estatal de Investigación /10.13019/501100011033 y por el Fondo Europeo de Desarrollo Regional (FEDER) ‘Una manera de hacer Europa’, bajo el proyecto PID2022-138286NB-I00, y por el CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto de Salud Carlos III. La base de datos empleada en este estudio está alojada en un proyecto financiado por el programa de investigación e innovación Horizonte 2020 de la Unión Europea bajo el acuerdo de subvención Marie Skłodowska-Curie número 676144 (*Synaptic Dysfunction in Alzheimer Disease*, SyDAD). A. Maturana-Candelas es beneficiario de una ayuda PIF de la “Consejería de Educación de la Junta de Castilla y León”.

Referencias

- [1] Meral A. Tubi, et al. White matter hyperintensities and their relationship to cognition: Effects of segmentation algorithm. *NeuroImage*, 2020.
- [2] Joanna M. Wardlaw, et al. What are white matter hyperintensities made of? *Journal of the American Heart Association*, 2015.
- [3] Franz Fazekas, et al. MR signal abnormalities at 1.5 T in Alzheimer’s dementia and normal aging. *American Journal of Roentgenology*, 1987.
- [4] Philip Scheltens, et al. White matter changes on CT and MRI: An overview of visual rating scales. *European Neurology*, 1998.
- [5] 2022 Alzheimer’s disease facts and figures. *Alzheimer’s & Dementia*, 2022.
- [6] D.B. Carr, et al. Current concepts in the pathogenesis of Alzheimer’s disease. *The American Journal of Medicine*, 1997.
- [7] Robert Stewart. Subjective cognitive impairment. *Current Opinion in Psychiatry*, 2012.
- [8] X. Delbeuck, et al. Alzheimer’s disease as a disconnection syndrome? *Neuropsychology review*, 2003.
- [9] E.C.W. van Straaten, et al. Disturbed phase relations in white matter hyperintensity based vascular dementia: An EEG directed connectivity study. *Clinical Neurophysiology*, 2015.
- [10] Ira J. Rampil. A primer for EEG signal processing in anesthesia. *Anesthesiology*, 1998.
- [11] Fabrizio Vecchio, et al. Resting state cortical EEG rhythms in Alzheimer’s disease. In *Supplements to Clinical Neurophysiology*, volume 62, pages 223–236. 2013.
- [12] Lihui Cai, et al. Functional integration and segregation in multiplex brain networks for Alzheimer’s disease. *Frontiers in Neuroscience*, 2020.
- [13] Meichen Yu, et al. Selective impairment of hippocampus and posterior hub areas in Alzheimer’s disease: an MEG-based multiplex network study. *Brain*, 2017.
- [14] Marshal F. Folstein, et al. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 1975.
- [15] Cornelis J. Stam, et al. Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Human Brain Mapping*, 2007.
- [16] Sinead Gaubert, et al. EEG evidence of compensatory mechanisms in preclinical Alzheimer’s disease. *Brain*, 2019.
- [17] Simona Gardini, et al. Increased functional connectivity in the default mode network in mild cognitive impairment: A maladaptive compensatory mechanism associated with poor semantic memory performance. *Journal of Alzheimer’s Disease*, 2015.
- [18] Casper T. Briels, et al. Reproducibility of EEG functional connectivity in Alzheimer’s disease. *Alzheimer’s Research & Therapy*, 2020.
- [19] Fanny Quandt, et al. Higher white matter hyperintensity lesion load is associated with reduced long-range functional connectivity. *Brain Communications*, 2020.

Implementing Emotion Detection from Speech for Psychological Assessment of Elderly People: A Comparative Study of Python-based Approaches and Existing Solutions

I. Warnants¹, N. Tsiogkas¹, J. Roca Gonzalez², F.J. Ortiz Zaragoza², I. Méndez³, JA. Vera Repullo², J.P. Serna⁴

¹ Catholic University of Leuven (KU Leuven), Leuven, Belgium, ismael.warnants@student.uhasselt.be, nikolaos.tsiogkas@kuleuven.be

² Departamento de Tecnología Electrónica, Universidad Politécnica de Cartagena, Cartagena, España, {jroca.gonzalez, francisco.ortiz, jose.vera}@upct.es

³ Department of Developmental and Educational Psychology, Universidad de Murcia, Espinardo, Murcia, 30100, España, inmamendez@um.es

Development Systems & Services, S.L. Algezares, 30157 Murcia, España. juanpedroserna@gmail.com

Abstract

In the last ten years, the number of people over 65 has increased 30% in Spain. This trend is anticipated to grow and require more healthcare personnel. To prevent this, people should live longer independently instead of in care homes. The ADDIM system will assist them in living independently. The research presented in this paper is part of the mood detection of the user in the ADDIM (Asistencia Domiciliaria Digital Integral para Mayores) system. This is a Digital platform for monitoring older people's health, safety, companionship, and emotional support at home based on robotics, artificial intelligence, and ambient assisted living.

To detect user emotions, the right speech corpus, feature extraction methods, preprocessing methods, and machine learning models have to be selected. Based on the detected emotion, the robot will interact with the user to perform predefined actions. The final mood of the user will be estimated using this output in conjunction with visual feedback and the sensors in the user's home with the ADDIM system.

Three speech corpora are selected with retraining to achieve personalized detection based on the user's previous recordings. In addition, this will ensure that the detection is improved over time, which has yet to be implemented in other research. Finally, the implementation uses dimensional emotion detection instead of discrete emotion detection. This augments the number of detectable emotions.

1. Introduction

Over the last decade, Spain witnessed a 30% increase in its population aged 65 and above. This is attributed to a rise in life expectancy from 79.1 years in 2000 to 83.2 years in 2019, making it one of the countries in the European Union and worldwide with the highest life expectancy among older adults [1]. While this is a positive outcome, it has implications for the country's healthcare system and the elderly population's mental health. The growth in elderly people living alone may lead to mental disorders impacting their quality of life. In this regard, several experts have spoken out in favor of a change in the model of dependency care in Spain, pointing out among the priorities the availability of comprehensive and integrated care in the home to relieve the pressure on healthcare personnel [2]. As an example, during COVID-19 there was a shortage of

care response capacity in nursing homes, causing a dramatic situation. Therefore, person-centered care and promoting adapted assisted living should be implemented to support elderly people living alone in Spain.

In this research, we propose the ADDIM system, a digital platform for monitoring the health, safety, companionship, and emotional support of older people at home. In this system, a Speech Emotion Recognition (SER) model will detect the user's emotions and work together with the rest of the ADDIM system's components to personalize the elderly user's care. The implementation goal is to be a proof-of-concept to check the feasibility of triggering actions in a robot based on the result provided by the SER model. Future research will explore the possible ways to improve the SER model and its real-world limitations. The aim is to detect emotions in audio fragments instead of continuous speech.

2. Framework

The ADDIM system extends mental assessment and mood monitoring with pro-active coaching capabilities, provided by a companion social robot, which suggests activities to older people to improve their mood and mental well-being.

The proposed companion robot and Ambient Assisted Living (AAL) system comprises several components, summarized below. It should be remembered that different deployments of the system are possible, from the simplest (only AAL) to the most complex model, including the social robot. Using low-cost commercial components (such as low-cost commercial sensors), open-source software, and communication standards (ROS, Node-RED, MQTT, Home Assistant) also favor the incorporation of new hardware and software for future developments.

- Assistive mobile robot: Designed to navigate autonomously around the house to attend to the users' needs or suggest an activity according to their mood.
- Home automation sensor ecosystem: Includes a set of low-cost sensors, both commercial and self-designed, to monitor the user's lifestyle.

- Speech Emotion Recognition (SER): A Python implementation of a machine learning algorithm to detect the emotion of the user from an audio file sent by the robot to the computer.
- Application for psychological data acquisition: Carries out a psychological study and thus relates the information obtained from the Empatica E4 to the user’s state of mind, such that basic questions about activity/well-being can be answered. This is another input for the machine-learning algorithm.
- Central home-assistant unit: Collects home automation data on the person’s routines at home and activates the necessary devices. This unit, an embedded PC to house the artificial intelligence algorithms for the coordination and functioning of the various well-being devices. ROS, Home Assistant, and IoT were used to make the information transparent among the system’s different elements. The SER model will run in this part of the system.



Figure 1. A scheme of the main elements.

3. State-of-the-art

Following the WHO’s [3] recommendations on ethics and governance in the field of health, artificial intelligence, as used in this work, holds great promise for public health practice, as it has great benefits, such as protecting human autonomy, promoting the well-being and safety of individuals and the public, ensuring transparency, clarity and intelligibility, and guaranteeing inclusiveness, equity and sustainability.

The field of mood recognition has gained in importance in the world of artificial intelligence. Successful studies have been conducted with wearables [4], [5], such as IoT sensors for behavioral inference [6]. However, putting the system to practical use is still the challenge we are facing with the proposed system.

Different sensor systems have been proposed to collect physiological and activity data and analyze them to automatically assess people’s moods. Smart monitoring systems use smartphones and self-reported depression reports to assess people’s moods, as in [7]. However, the latter system does not provide counseling. In the H2020 Help4mood project, data on daily activities are converted

into graphical, textual, and conceptual summaries that can be communicated to clinicians providing external assessment.

Automatic mood monitoring and assessment are important components of support therapies and daily life coaching. Ecological momentary assessment (EMA) can ease patients’ mental burdens since the traditional assessment tests are too long to be used repeatedly. In addition, the latest smartphones and wearable devices make the EMA approach much more feasible as a solution for monitoring mental illness and offer economically friendly solutions.

The above-described method of gathering information on mental well-being has been used in several works. In [8], a system is proposed to monitor potential depressive patterns in elderly people living alone. Emotional states are assessed from diverse sources, such as surveys, smartwatches, and EMA questionnaires. EMA methods for obtaining emotional information can expand this research area to practical applications. For example, in [9], a system that gathers information on a person’s activities can detect long-term stress patterns. Stress detection during daily real-world tasks through advances in intelligent systems is described in [10]. Techniques such as machine learning were also applied [11] to monitor elderly people’s moods via intelligent sensors on wristbands.

An issue to be addressed is the usability of these modern assistive technologies by the elderly. In [12], the authors present a user-centered design for a web-based multi-modal user interface tailored for elderly users of near-future multi-robot services.

3.1. Empathic vs. non-empathic Robot

In [13] they look at the effects of the robot named Ryan with or without emotional assessment in a conversation with the user. They noticed little difference in the empathic or non-empathic version of the robot in conversation with the user. However, when the word count was measured, and the exit survey of all participants was finished, they noticed that the emotional version of Ryan was more engaging and likable by the participants. And that the empathic version of the robot encourages users to have longer conversations compared to the non-empathic version. This was suggested to have the potential to decrease depression in the users.

3.2. Speech Emotion Recognition (SER)

In the state of the art [14], there are two methods to detect emotions: discrete and dimensional. The dimensional was chosen as it has two parameters to plot any emotion on the Speech Affective Space Model (SASM) [15]. Figure 2 shows the Speech Affective Space Model used in the state of the art. The SASM is not limited to the number of labels on which the model is trained and can, because of this predict more emotions compared to the available emotions in the speech corpus. The Euclidean distance is used to measure the distance of the predicted emotion to the closest discrete emotion. For each dimension (Valence and Arousal) the concordance correlation coefficient (CCC) is calculated to measure the accuracy and precision of the predicted results compared to the expected results. In the

state of the art the CCCV is 0.28 ± 0.05 and the CCCA is 0.58 ± 0.05 [14].

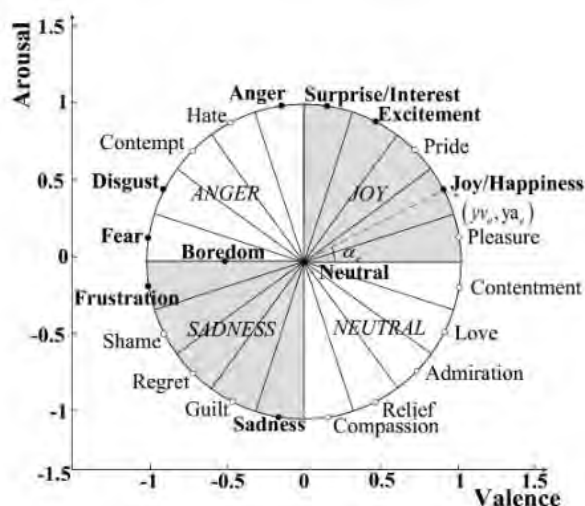


Figure 2. Representation of emotions on the Speech Affective Space Model (SASM) by using Arousal-Valence values [14]

4. Methodology

To create a SER model, the first step is to select a speech corpus, but as there are no speech corpora with dimensional values, the discrete values are mapped onto the SASM and the model output will be compared to these converted Valence-Arousal values. The selected speech corpora was EmoDB to compare our results to the state of the art. Our results were similar for other speech corpora. It has to be noted that most speech corpora are artificial and the expression of emotions can vary across a lot of factors (ethnicity, age, etc.) [16].

The second step is to extract the features from the audio files from the speech corpora. For this, a wide range of methods and features can be used. The best results were achieved using a library by Renovamen (Github) [17]. It consists of 320 features that are normalized. The ones with the highest correlation to the output are chosen with principal component analysis (PCA).

With these features, the model can be trained and optimized. The selected models are deep learning models as they improve with the more data they get. This is beneficial as the goal is to retrain later on with the final emotion from the user we get from all the inputs of the ADDIM system that will detect the user's emotion and can retrain the Emotion Recognition (ER) models like this one. A CNN, LSTM and CNN+LSTM model are used to compare to the state of the art. The CNN+LSTM model combines CNN layers as the input and LSTM layers as the output. These models are optimized by changing the amount and types of layers and other parameters.

The models will be evaluated on their accuracy, MSE, RMSE, R², CCCV, CCCA, and the number of inputs. The amount of inputs is also correlated with the amount of calculations that have to be performed and, thus also the time to train and predict.

5. Results & discussion

The models were trained and used to predict emotions on a CPU and were able to predict in a matter of seconds.

Table 1 compares our models with the State Of The Art [14]. The CNN model and CNN+LSTM model show great performance and outperform the state of the art. The LSTM results are not promising, but still in line with the state of the art. The training and validation loss of the models are shown in Figure 3-5.

The results with the CNN+LSTM model are achieved with less inputs compared to the other models. This means that there are less features required and thus also less computing time.

Model	CNN	LSTM	CNN+LSTM	SOTA
Accuracy (%)	78.78	36.36	78.78	/
MSE	0.1110	0.2501	0.1269	/
RMSE	0.3333	0.5001	0.3562	/
R ²	0.6553	0.2429	0.6413	/
CCCV	0.8531	0.5573	0.7615	0.28±0.05
CCCA	0.8331	0.5253	0.8901	0.58±0.05
Inputs	46	41	40	/

Table 1. Model training results for CNN, LSTM and CNN+LSTM with EmoDB, compared to the the State-Of-The-Art (SOTA)

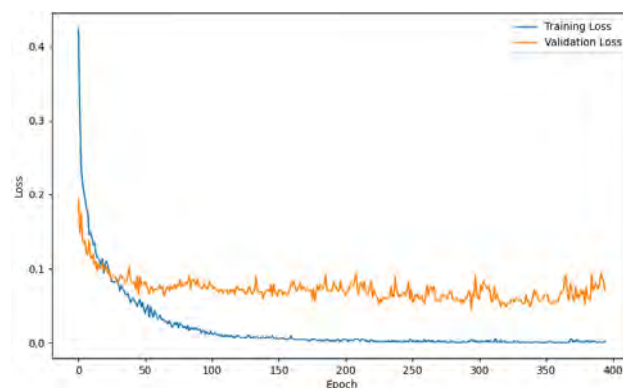


Figure 3. Training and validation loss during training of the CNN model on the EmoDB speech corpus

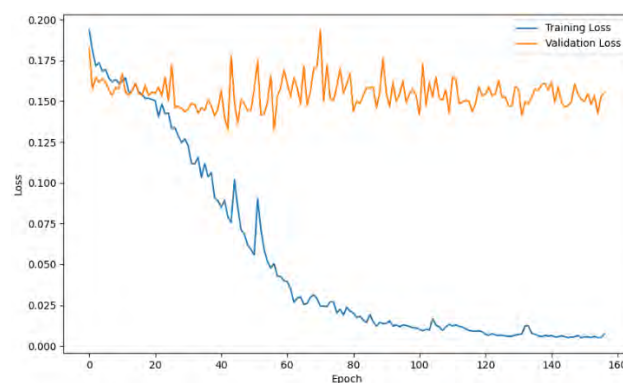


Figure 4. Training and validation loss during training of the LSTM model on the EmoDB speech corpus

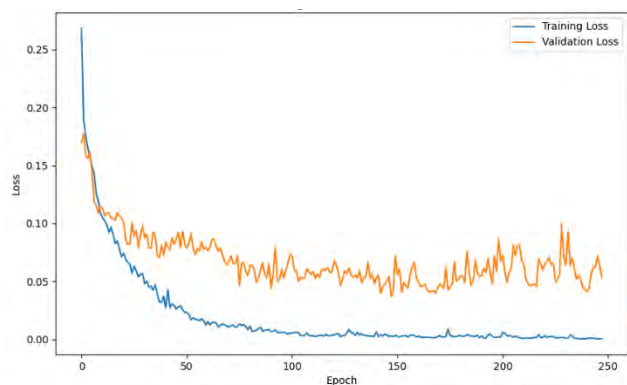


Figure 5. Training and validation loss during training of the CNN+LSTM model on the EmoDB speech corpus

The features that were selected with principal component analysis (PCA) as inputs for the models were consistent during testing while comparing with other speech corpora. This seems to show that there are features that indicate the amount of Valence and Emotion regardless of the language, recording method, and other variables.

The models only get features that can be extracted from the audio files, so there is no speech context used for these results. This means that conversations cannot be recorded by the model, which is again better for privacy.

The results of the SER will be added with the other models into the system to get an even more accurate result for the user's emotion. This emotion will be used for retraining the model to adapt to the user and become better over time. This is possible because we use deep learning models which thrive with the increased amount of data.

There are however some limitations. The EmoDB speech corpus is limited and artificial in nature, which is the same for other speech corpora. It is also difficult to find speech corpora in Spanish, especially for elderly people.

6. Conclusion

The CNN and CNN+LSTM models have proven to outperform the state of the art. This while maintaining privacy by running locally on a CPU and predicting fast.

These results should be retrained on a Spanish speech corpus, if possible, but the features that are used to train and predict seem to be consistent across speech corpora and languages.

Future work must prove the advantages of retraining in a real-world scenario. Results for audio filtering could also be explored for real-world scenarios. The model could also be adapted to work in real-time to respond even faster compared to the speed we have now.

These improvements will benefit the promotion and monitoring of the emotional state and therefore the quality of life of older people.

Acknowledgements

The HIMTAE project, Robwell subproject (reference RTI2018-095599-A-C22) has been funded by: Programa

Estatad de Investigación, Desarrollo e Innovación Orientada a los Retos de la Sociedad, en el marco del Plan Estatal de Investigación Científica y Técnica y de Innovación 2013-2016.

The pilot project ADDIM-Poncemar (Asistencia Domiciliaria Digital Integral para Mayores en Poncemar) funded by the "Fundación Integra Digital" of the Region of Murcia is currently being developed in the Day Care Centers of the Poncemar Foundation, in the Health Campus of Lorca, University of Murcia. We thank them for their collaboration and involvement in the testing of the solutions presented in this communication.

Finally, we would like to thank the previous CASEIB organising committees for their kindness in allowing us to use their style guides as a reference for this document.

References

- [1] "Spain data | World Health Organization." Accessed: Aug. 18, 2023. [Online]. Available: <https://data.who.int/countries/724>
- [2] V. Davey, *Situación en España de la evaluación de sistemas de atención a personas mayores en situación de dependencia. Informe en Red nº 28*. 2021. doi: 10.13140/RG.2.2.28439.09128.
- [3] "Ethics and governance of artificial intelligence for health." Accessed: Sep. 26, 2023. [Online]. Available: <https://www.who.int/publications-detail-redirect/9789240029200>
- [4] O. M. Mozos *et al.*, "Stress Detection Using Wearable Physiological and Sociometric Sensors," *Int. J. Neural Syst.*, vol. 27, no. 02, p. 16500416, Mar. 2017, doi: 10.1142/S0129065716500416.
- [5] M. Koutli, N. Theologou, A. Tryferidis, and D. Tzovaras, "Abnormal Behavior Detection for Elderly People Living Alone Leveraging IoT Sensors," in *2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)*, Oct. 2019, pp. 922–926. doi: 10.1109/BIBE.2019.00173.
- [6] Calatrava, F.M.; Ortiz, F.J.; Vera, J.A.; Roca, J.; Jiménez, M.; Martínez, O., "Heterogeneous system for monitoring daily activity at home and the well-being of the elderly," in *In Proceedings of the XLII Conference on Automatic: Minutes Book*, Castelló, Spain, 2021, pp. 632–639.
- [7] A. Doryab, "Detection of behavior change in people with depression," Jul. 2014.
- [8] H. Kim, S. Lee, S. Lee, S. Hong, H. Kang, and N. Kim, "Depression Prediction by Using Ecological Momentary Assessment, Actiwatch Data, and Machine Learning: Observational Study on Older Adults Living Alone," *JMIR MHealth UHealth*, vol. 7, no. 10, p. e14149, Oct. 2019, doi: 10.2196/14149.
- [9] R. Kocielnik, N. Sidorova, F. M. Maggi, M. Ouwerkerk, and J. H. D. M. Westerink, "Smart technologies for long-term stress monitoring at work," in *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, Jun. 2013, pp. 53–58. doi: 10.1109/CBMS.2013.6627764.
- [10] M. V. Gómez-Gómez, M. V. Bueno-Delgado, C. Albaladejo-Pérez, and V. Koch, "Augmented Reality, Virtual Reality and Mixed Reality as Driver Tools for Promoting Cognitive Activity and Avoid Isolation in Ageing Population," in *Smart Objects and Technologies for Social Good*, I. M. Pires, S. Spinsante, E. Zdravevski, and P. Lameski, Eds., in Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Cham: Springer International Publishing, 2021, pp. 197–212. doi: 10.1007/978-3-030-91421-9_15.
- [11] "EnrichMe - Collaborative Projects | PAL Robotics." Accessed: Sep. 26, 2023. [Online]. Available: <https://pal-robotics.com/collaborative-projects/enrichme/>
- [12] A. Di Nuovo *et al.*, "The multi-modal interface of Robot-Era multi-robot services tailored for the elderly," *Intell. Serv. Robot.*, vol. 11, no. 1, pp. 109–126, Jan. 2018, doi: 10.1007/s11370-017-0237-6.
- [13] H. Abdollahi, M. Mahoor, R. Zandie, J. Sewierski, and S. Qualls, "Artificial Emotional Intelligence in Socially Assistive Robots for Older Adults: A Pilot Study," *IEEE Trans. Affect. Comput.*, pp. 1–1, 2022, doi: 10.1109/TAFFC.2022.3143803.
- [14] S. Jing, X. Mao, and L. Chen, "Automatic speech discrete labels to dimensional emotional values conversion method," *IET Biom.*, vol. 8, no. 2, pp. 168–176, 2019, doi: 10.1049/iet-bmt.2018.5016.
- [15] N. Kamaruddin and A. Wahab, "Human behavior state profile mapping based on recalibrated speech affective space model," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2012, pp. 2021–2024. doi: 10.1109/EMBC.2012.6346354.
- [16] A. Hanjalic, "Extracting moods from pictures and sounds: towards truly personalized TV," *IEEE Signal Process. Mag.*, vol. 23, no. 2, pp. 90–100, Mar. 2006, doi: 10.1109/MSP.2006.1621452.
- [17] X. Zou, "Speech Emotion Recognition." Aug. 17, 2023. Accessed: Aug. 18, 2023. [Online]. Available: <https://github.com/Renovamen/Speech-Emotion-Recognition>

Sesión Póster III

Viernes 24 10:00 - 11:00

Claustro

Estimadores del retardo entre las series de QT y RR en registros ECG de prueba de esfuerzo: evaluación en simulación

C. Pérez¹, E. Pueyo^{1,2}, JP. Martínez^{1,2}, L. Sörnmo³, P. Laguna^{1,2}

¹Grupo BSICoS, I3A, IIS Aragón, Universidad de Zaragoza, Zaragoza, España {cperez, epueyo, jpmart, laguna}@unizar.es

²CIBER-Bioingeniería, Biomateriales y Nanomedicina (CIBER BBN), España

³Department of Biomedical Engineering, Lund University, Sweden, leif.sornmo@bme.lth.se

Resumen

La adaptación lenta del intervalo QT ante cambios abruptos de la frecuencia cardíaca (FC) se ha identificado como un marcador de riesgo de arritmias ventriculares. Sin embargo, estos cambios abruptos no son fáciles de inducir en pacientes. Recientemente, se ha propuesto un método para cuantificar este tiempo de adaptación en electrocardiogramas (ECG) grabados durante prueba de esfuerzo. Para ello se calcula el retardo entre la serie real de QT y una serie estimada del QT instantáneo y sin memoria que correspondería a la FC real en cada momento. En este trabajo se evalúa este método en un entorno controlando usando señales simuladas de ECG correspondientes a una prueba de esfuerzo. Las señales se obtienen a partir de un simulador al que se le ha incorporado la capacidad de introducir una memoria del QT y se ha añadido la contaminación con ruido muscular (RM) variante en el tiempo. Los resultados muestran que: a) la delineación del final de la onda T sobre la primera derivación transformada espacial usando “Análisis de Componentes Periódicas” ofrece el menor error para valores de SNR bajos en prueba de esfuerzo, b) los valores obtenidos con el estimador propuesto resultan mínimamente sesgados para SNR de 25 a 50 dB, siendo la estrategia de menor sesgo la que estima el QT instantáneo, mediante una corrección previa de los pares [QT, RR] en el pico la prueba de esfuerzo.

1. Introducción

El incremento de la heterogeneidad espaciotemporal de la repolarización ventricular está relacionado con la inestabilidad cardíaca, la cual podría desencadenar arritmias ventriculares y muerte súbita cardíaca (SCD) [1]. El tiempo de adaptación del intervalo QT ante cambios abruptos de la frecuencia cardíaca (FC) ha sido identificado como un marcador de riesgo arrítmico [2]. Los cambios abruptos de tipo escalón, donde habitualmente se evalúa esta adaptación, no siempre son fáciles de observar y/o generar, pudiendo incluso no estar presentes en registros ambulatorios. Por ello, se propuso en anteriores trabajos estimar la memoria del intervalo QT como el retardo temporal entre la serie QT real y una serie estimada de QT instantáneo y sin memoria que correspondería a la FC real en cada momento, permitiendo así evaluar cambios graduales de la FC en lugar de cambios abruptos [3]. Este tipo de cambios graduales son los que se observan en señales ECG registradas durante una prueba de esfuerzo (EST), en este caso en forma de rampa. En [3] se evaluó el retardo del intervalo QT en registros reales, calculándolo de forma independiente en el tramo de ejercicio y en el tramo de recuperación de la prueba. Los resultados

mostraron que el rango de valores del tiempo de adaptación del intervalo QT en respuesta a cambios graduales de FC era similar a los evaluados para cambios abruptos de la FC [3] y que permitían discernir entre sujetos con enfermedad coronaria severa y leve. Sin embargo, la robustez y la precisión de estos métodos no pueden evaluarse analizando señales reales de prueba de esfuerzo, donde el verdadero valor de la memoria del QT es desconocido, y otros tipos de pruebas electrofisiológicas en paralelo para determinarlo son difícilmente accesibles.

Alternativamente, en este estudio se han generado señales simuladas de ECG cuya dinámica de la serie RR procede de datos reales de grabaciones durante una EST, y en cuyo simulador se ha incluido la posibilidad de tener una relación entre QT y RR predeterminada. Esta está descrita por una relación no lineal y se considera adicionalmente una memoria del QT al seguir variaciones de RR [4] donde el tiempo de adaptación es un parámetro de entrada al simulador. Los ECGs simulados se contaminaron con ruido muscular (RM) sintético o real, con varias relaciones señal-ruido (SNR) predefinidas.

Los principales objetivos de este estudio son: (1) Evaluar la delineación del final de la onda T (T_e), empleando diferentes métodos de transformación espacial de derivaciones de ECG para enfatizar dicha onda; (2) estudiar la robustez y la precisión de la estimación, con el método descrito, del tiempo de adaptación del intervalo QT para diferentes niveles de ruido.

2. Métodos

2.1. ECGs en simulación

Se generaron señales ECG simuladas [4] con diferente tiempo de adaptación (τ) del intervalo QT con un patrón de RR predefinido. Este patrón es similar al observado durante una EST, pero con (y sin) una modificación en el pico de esfuerzo. En esta modificación se mantuvo el valor de RR en el pico durante 10 minutos para así facilitar la convergencia a un estado estacionario de la relación QT-RR antes de empezar el periodo de recuperación, y evaluar el impacto en la estimación de la no convergencia en la realidad. Se generaron 10 señales por cada valor de τ , cuyos valores se tomaron de 10 a 50 s en pasos de 10 s, resultando en un total de 50 ECGs simulados limpios, formados por las 8 derivaciones estándar, V1-V6, I y II, y con una frecuencia de muestreo de 1000 Hz. Para la producción de estas señales se utilizó la última versión del simulador

descrito en [4]. Los parámetros del simulador definidos por el usuario para producir estas señales se resumen en la Tabla 1. Los ECGs simulados fueron contaminados por señales de RM sintéticas o reales con un SNR definido.

Por otro lado, se crearon señales de RM de media cero siguiendo el modelo descrito en [4], que es constante en las zonas basales y linealmente creciente/decreciente en la etapa de ejercicio/recuperación. La varianza en el pico de ejercicio de esta señal sintética se definió como 4 veces la varianza en la zona basal del inicio de la prueba (*basal inicio*), que se calculó de acuerdo al valor de SNR deseado:

$$SNR_l = 20 \log_{10} \left(\frac{A_{QRS,l}}{RMS_{noise,l}} \right) \quad (1)$$

donde $A_{QRS,l}$ es la amplitud pico a pico del complejo QRS medio de la derivación l (este complejo comprende un intervalo de 100 ms centrado alrededor del pico de la onda R) [6,7]. Toda la señal de RM sintética se escaló según el valor RMS de una ventana de 60 s alrededor del pico de ejercicio, $RMS_{noise,l}$.

	Parámetro	Valor
Duración fases (min)	Basal inicio	10
	Rampa ejercicio	*
	Pico	10
	Rampa recuperación	*
	Basal final	10
Frecuencia Cardíaca (bpm)	Basal inicio	80
	Pico	165
	Basal final	95
Frecuencia respiratoria (Hz)	Basal inicio	0.25
	Pico	0.7
	Basal final	0.3

Tabla 1. Parámetros señales de ECG simuladas

*Tanto la duración de la zona de ejercicio como la de recuperación se definen de acuerdo a la información de duración y posición del pico de ejercicio de las 10 señales de ruido reales disponibles extraídas de ECGs registrados durante una prueba de esfuerzo [6].

Además, se extrajeron señales RM reales a partir de ECGs de pacientes grabados durante una EST [6] usando un filtro paso-banda, cuyas frecuencias de corte paso bajo y alto son 10 y 200 Hz, respectivamente. Fijado el valor de SNR que se deseaba, se obtuvo un factor de escala para la señal RM correspondiente al valor RMS calculado de la ecuación (1).

Las señales RM añadidas a cada derivación son independientes unas de otras, tanto si son señales sintéticas como reales. Los valores de SNR definidos para ambos tipos de señal RM son 25, 20, 25, 40 y 50 dB. Se tiene un conjunto de 500 ECGs contaminados con ruido, tanto sintético como real.

2.2. Delineación de la onda T

En este apartado se evalúa la precisión en la delineación del final de la onda T, T_e , usando una estrategia basada en delinear cada una de las 8 derivaciones disponibles y, tras aplicar una regla, seleccionar una única marca (Mleads), así como delineando solo la primera derivación

transformada tras aplicar una técnica de transformación espacial (LSR) de las derivaciones de ECG.

Las señales de ECG contaminadas se filtraron antes de ser delineadas. Primero, el ruido de alta frecuencia y los posibles artefactos fueron atenuados por un filtro Butterworth paso-bajo, de orden 6 y frecuencia de corte 50 Hz implementado de forma bidireccional para obtener fase cero. Después, la línea de base se atenuó empleando una interpolación con splines cúbicos.

Para obtener las derivaciones transformadas, se definieron diferentes métodos basados en dos técnicas de LSR: Análisis de Componentes Periódicas (π CA) [8] y Análisis de Componentes Principales (PCA) [9]. Se emplearon estas técnicas para enfatizar la onda T en la primera derivación transformada. Posteriormente, se delineó la derivación [10] para obtener el punto T_e [11]. Los diferentes métodos basados en las dos técnicas de LSR mencionadas, con una estimación única de la transformación o variable a lo largo del tiempo, son:

- π CA: o $G\pi CA_1$, la transformación se aprende en cada ventana de 150 s, recalculando la matriz de transformación en cada ventana, con periodicidad de 1 latido ($P = 1$).
- $G\pi CA_3$: La transformación se aprende en cada ventana de 150 s, recalculando la matriz de transformación y considerando periodicidades de 1, 2 y 3 latidos ($P = 3$).
- πCA_0 : $G\pi CA_{1,0}$, donde la matriz de transformación se estima una sola vez en la primera ventana de 150 s de la señal y $P = 1$. La misma matriz de transformación se aplica al resto de la señal.
- $G\pi CA_{3,0}$: La matriz de transformación se calcula una sola vez usando la información de la primera ventana de 150 s y se aplica al resto de la señal con $P = 3$.
- PCA: técnica de PCA donde la matriz de transformación se recalcula en cada ventana de 150 s.
- PCA₀: en este caso, la matriz se estima una sola vez usando la información de los primeros 150 s, y luego se aplica al resto de la señal.

Para estudiar el comportamiento de cada método en función del nivel de ruido, se tomó como valor de referencia la delineación del punto T_e en cada latido k , $T_{e,k}^r$, de una señal ECG simulada y con un nivel inapreciable de ruido, considerando para ello una SNR = 50 dB, con lo que se evita singularidades en algunos cálculos matriciales. Por tanto, el error de delineación, ϵ , se calcula como:

$$\epsilon = \sqrt{\frac{1}{K} \sum_{k=1}^K (T_{e,k}^r - T_{e,k})^2} \quad (2)$$

donde K es el número de latidos en los que $T_{e,k}$ está definido.

2.3. Comportamiento del tiempo de adaptación del intervalo QT como estimador

El retardo entre la serie de QT real, $d_{QT}(n)$, y la serie de QT sin memoria (estimada usando el valor instantáneo de

la FC), $d_{QT}^i(n)$, refleja el tiempo de adaptación del intervalo QT ante cambios graduales de la FC, siguiendo estos cambios una evolución de tipo rampa (como los producidos en las fases de ejercicio y recuperación). La serie $d_{QT}^i(n)$ se obtuvo ajustando un modelo de regresión hiperbólica a partir de los valores de la serie RR, $d_{RR}(n)$ y QT, $d_{QT}(n)$ [3], por lo que se necesita calcular tanto la serie RR como la serie QT para evaluar el retardo.

Para ello, tanto el pico de la onda R como el inicio del complejo QRS (QRS_0) se obtienen al aplicar la estrategia Mleads, es decir, delineando todas las derivadas disponibles y quedándonos finalmente con solo una marca [10]. Sin embargo, el punto T_e se obtiene al delinear la primera derivación transformada del método $G\pi CA_{3,0}$ [3].

El retardo entre $d_{QT}^i(n)$ y $d_{QT}(n)$ se estima en la rampa de ejercicio, $\hat{\tau}_e$, y en la rampa de recuperación, $\hat{\tau}_r$, de forma independiente, usando un estimador basado en mínimos cuadrados:

$$\hat{\tau} = \arg \min_{\tau} \sum_{n=n_0}^{n_e} (d_{QT}^i(n) - d_{QT}(n + \tau))^2; \tau \in \{-I, \dots, I\}, \quad (3)$$

donde I delimita el rango plausible de τ , y n_0 y n_e definen el punto inicial y final de las rampas, respectivamente.

En [3] se desarrolló un método automático para determinar los puntos n_0 y n_e . El inicio de la rampa de ejercicio se definió como el punto de cambio entre el área basal inicial, la cual se supone plana, y la subsecuente área de ejercicio que sigue una tendencia tipo rampa. Esta misma idea se empleó para definir el final de la rampa de recuperación, buscando ahora el punto que delimita la rampa creciente de recuperación y la zona basal plana del final de la prueba. Tanto el final de la rampa de ejercicio como el inicio de la rampa de recuperación se establecen al 55% del total de la rampa de $d_{QT}^i(n)$, de modo que estos puntos están alejados del pico de ejercicio (véase [3]).

3. Resultados y discusión

La Figura 1 muestra la media y la desviación estándar de los errores de delineación del final de la onda T, ϵ , para cada método definido, SNR y tipo de RM usando los ECGs simulados que mantienen el valor de RR en el pico de esfuerzo durante 10 min. Se observa que cuanto menor es el valor de SNR, los métodos derivados de alguna técnica de LSR tienen una delineación más precisa que la estrategia Mleads. Cuando las señales ECG son contaminadas por RM sintético, $G\pi CA_{3,0}$ registra errores de 5.97 ± 3.32 , 3.71 ± 2.09 , 2.67 ± 1.79 y 1.85 ± 1.00 ms para SNR = 25, 30, 35 y 40 dB, respectivamente. Los métodos basados en PCA muestran el menor error cuando el valor de SNR es alto, pero este error no está muy alejado del que se obtiene con cualquier método basado en πCA (6.35 ± 3.78 , 3.56 ± 2.07 , 2.00 ± 0.98 y 1.19 ± 1.00 ms para SNR = 25, 30, 35 and 40 dB, respectivamente, con PCA_0). La delineación obtenida con la estrategia Mleads exhibe el peor comportamiento, excepto para SNR = 25 dB (6.09 ± 2.26 , 4.74 ± 1.70 , 3.67 ± 1.39 y 2.89 ± 1.42 ms para SNR

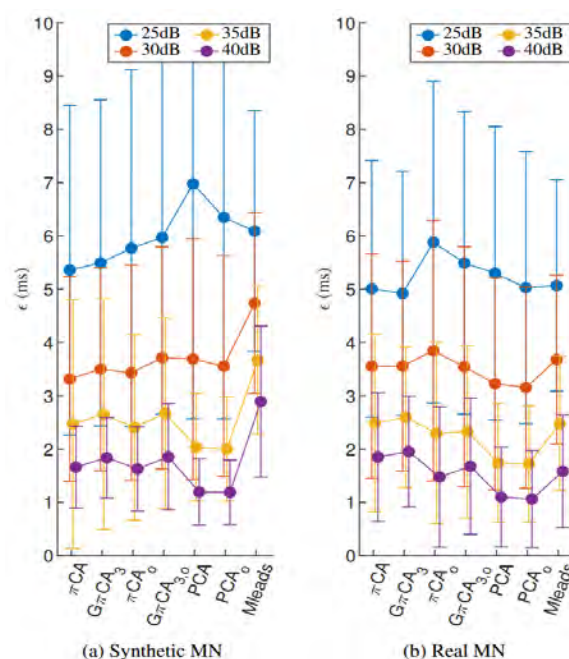


Figure 1. Media y desviación estándar del error de delineación, ϵ , para cada método de delineación, valor SNR y tipo de RM usando los ECGs simulados que mantienen el valor de RR en el pico de esfuerzo durante 10min.

= 25, 30, 35 y 40 dB, respectivamente). Una tendencia menos clara se observa en los errores obtenidos con las señales de ECG contaminadas con RM real. Cuanto menor es el valor de SNR, mejor es el comportamiento de πCA . En los casos con un valor de SNR menor de 25 dB, las mejores delineaciones se obtienen con el método de πCA para ruido real y con $G\pi CA_{3,0}$ para RM sintético. Estos resultados arrojan el mismo comportamiento estudiado en un trabajo anterior [11], donde la variabilidad latido a latido a lo largo de la serie QT extraída del ECG se usó como una medida subrogada del error en la delineación del punto T_e . Cuando el valor de SNR es igual a 25 dB, la estrategia Mleads no es capaz de delinear el punto T_e del 33% del total de latidos, mientras que los métodos basados en alguna técnica de LSR son capaces de linear dicho punto en el 99% de los latidos.

Debido a que encontramos crítico delinear el punto T_e para señales con un valor de SNR menor de 25dB, en el resto del trabajo se escogió la primera derivada obtenida con el método $G\pi CA_{3,0}$ para delinear el punto T_e , de acuerdo con los resultados que se han descrito.

Se calcularon la media y la desviación estándar de los retardos estimados, tanto en ejercicio, $\hat{\tau}_e$, como en recuperación, $\hat{\tau}_r$, para ECGs con diferente τ , diferentes valores de SNR y diferentes tipos de RM. Tal como puede observarse en la Figura 2, se obtuvieron resultados similares cuando los retardos son estimados en señales ECGs contaminadas por RM sintético o real. El tiempo medio de adaptación estimado en la rampa de ejercicio $\hat{\tau}_e$ está ligeramente sesgado, siendo estos valores estimados más altos que el tiempo de adaptación definido en las señales simuladas. En los resultados extraídos de la recuperación, se observa una menor influencia del valor de

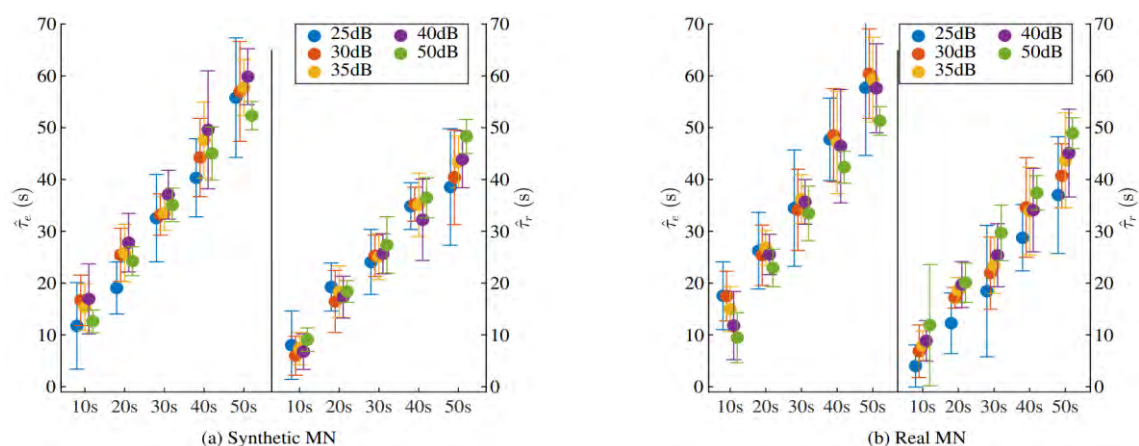


Figure 2. Media y desviación estándar de los retardos estimados en las rampas de ejercicio, $\hat{\tau}_e$, y en recuperación, $\hat{\tau}_r$, calculados en los ECGs con diferente tiempo de adaptación del intervalo QT, valor SNR y tipo de ruido muscular.

SNR cuando el tiempo de adaptación a estimar es menor. Además, también se observa que los valores de $\hat{\tau}_r$ están sesgados, pero en este caso los valores estimados son menores que los definidos. Este sesgo no excede el 20% en ningún caso. En general, la desviación estándar es mayor para valores de SNR menores.

4. Conclusión

En este estudio se han generado señales ECG simuladas que siguen un patrón similar a aquellas registradas durante una prueba de esfuerzo, donde se establece la relación temporal entre QT y RR y se añade ruido muscular con un SNR bien definido. Se confirma que delinear el punto final de la onda T en la primera derivación transformada espacial procedente de la técnica de “Análisis de Componentes Periódicas” ofrece una mayor precisión cuando el valor SNR es bajo. Además, se comprueba que es posible estimar el tiempo de adaptación del intervalo QT ante cambios tipo rampa de la FC utilizando un método que compara la serie de QT y una serie de QT estimada usando el valor instantáneo de la FC. Los resultados obtenidos están mínimamente sesgados respecto al tiempo de adaptación del intervalo QT definido en la producción de las señales ECG simuladas para un rango de SNR entre 25 dB y 50 dB.

Agradecimientos

Este trabajo se financió con proyectos PID2022-140556OB-I00 and TED2021-130459B-I00 (MICINN), y por el Gobierno de Aragón (Grupo de referencia BSICoS T39-23R and project LMP94_21. C. Pérez agradece al Gobierno de Aragón por su beca de doctorado personal, IIU/796/2019.

Referencias

[1] Artyeva NV. Dispersion of ventricular repolarization: Temporal and spatial. *World Journal of Cardiology* 2020; 12(9):437.
 [2] Pueyo E, Smetana P, Caminal P, De Luna AB, Malik M, Laguna P. Characterization of QT interval adaptation to RR

interval changes and its use as a risk-stratifier of arrhythmic mortality in amiodarone-treated survivors of acute myocardial infarction. *IEEE Transactions on Biomedical Engineering* 2004;51 (9).

- [3] Pérez C, Pueyo E, Martínez JP, Viik J, Laguna P. Characterization of Impaired Ventricular Repolarization by Quantification of QT Delayed Response to Heart Rate Changes in Stress Test. In *Computing in Cardiology*, volume XLVII. IEEE Computer Society, 2020.
- [4] Bachi L, Halvaei H, Pérez C, Martín-Yebra A, Petrénas A, Johnson L, Marozas V, Martínez JP, Pueyo E, Stridh M, Laguna P, Sörnmo L. ECG modeling for simulation of arrhythmias in time-varying conditions. *IEEE Transactions on Biomedical Engineering* 2023; In press (v).
- [5] Petrénas A, Marozas V, Solôsenko A, Kubilius R, Skibarkiene J, Oster J, Sörnmo L. Electrocardiogram modeling during paroxysmal atrial fibrillation: application to the detection of brief episodes. *Physiological Measurement* 2017;38 (11):2058.
- [6] Bailón R, Sörnmo L, Laguna P. A robust method for ECG based estimation of the respiratory frequency during stress testing. *IEEE transactions on biomedical engineering* 2006;53(7):1273–1285.
- [7] Martín-Yebra A, Sörnmo L, Laguna P. QT interval adaptation to heart rate changes in atrial fibrillation as a predictor of sudden cardiac death. *IEEE Transactions on Biomedical Engineering* 2022;69(10):3109–3118.
- [8] Sameni R, Jutten C, Shamsollahi MB. Multichannel electrocardiogram decomposition using periodic component analysis. *IEEE Transactions on biomedical engineering* 2008;55(8):1935–1940.
- [9] Castells F, Laguna P, Sörnmo L, Bollmann A, Roig JM. Principal component analysis in ECG signal processing. *EURASIP Journal on Advances in Signal Processing* 2007; 2007:1–21.
- [10] Martínez JP, Almeida R, Olmos S, Laguna P. A wavelet based ECG delineator: evaluation on standard databases. *IEEE Transactions on biomedical engineering* 2004;51(4).
- [11] Pérez C, Martín-Yebra A, Viik J, Martínez JP, Pueyo E, Laguna P. Eigenvector-based spatial ECG filtering improves QT delineation in stress test recordings. In *2021 55th Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2021; 261–264.

Aplicación multiplataforma para gestión farmacológica del servicio de Neonatología (SAFINEo)

O. de Francisco Ortiz¹, D. Vera Carrillo², C. Gómez Fernández³, J.L. Leante Castellanos⁴

¹ Departamento Ingeniería Mecánica, Materiales y Fabricación, Universidad Politécnica de Cartagena, Cartagena, España, oscar.defrancisco@upct.es

² Departamento Ingeniería Mecánica, Materiales y Fabricación, Universidad Politécnica de Cartagena, Cartagena, España, david.verac@edu.upct.es

³ Unidad de Neonatología, Hospital General Universitario Santa Lucía, Cartagena, España, cristina.gomez3@carm.es

⁴ Unidad de Neonatología, Hospital General Universitario Santa Lucía, Cartagena, España, josel.leante@carm.es

Resumen

El desarrollo de la administración de medicación en unidades de recién nacidos sigue siendo un punto crítico debido a su complejidad y la necesidad de ajustarse a los cambios en la población neonatal según edad gestacional, días de vida, peso, patologías e incompatibilidades con otros fármacos. El proceso manual actual implica múltiples cálculos matemáticos que pueden resultar en errores médicos, complicaciones clínicas y mayores tiempos y costos de preparación. Para abordar estos desafíos, se ha elaborado una aplicación móvil que facilita y agiliza la preparación de medicamentos, basándose en las recomendaciones más recientes para la población neonatal. La aplicación, diseñada para dispositivos accesibles como tabletas y teléfonos inteligentes en las principales plataformas (Android e iOS), busca paliar efectos adversos y reducir los tiempos requeridos para mejorar la seguridad y eficiencia en la administración de medicamentos en recién nacidos.

1. Introducción

El cuidado de los neonatos en unidades de neonatología es una tarea altamente especializada y delicada. La administración precisa y segura de medicamentos es esencial para garantizar un adecuado tratamiento y minimizar los riesgos asociados a dosificaciones incorrectas [1]. La falta de preparaciones comerciales diseñadas específicamente para neonatos ha sido una problemática latente en esta área [2].

En la actualidad, la preparación y administración de fármacos en periodo neonatal es una tarea difícil debido a la falta de preparaciones comerciales específicas para esta población, lo que lleva a dosis poco definidas basadas en el peso o superficie corporal. Esta indefinición puede resultar en dosis tóxicas o ineficaces, con posibles efectos adversos. Además, la necesidad de diluir los fármacos de forma precisa antes de su administración parenteral al neonato aumenta la probabilidad de errores en el cálculo, elección del diluyente, estabilidad de la mezcla o concentración final [3].

Con el fin de mejorar este aspecto, se ha desarrollado SAFINEo, una aplicación móvil multiplataforma innovadora y versátil, creada por un equipo multidisciplinario de ingenieros, médicos y enfermeros especializados en neonatología, en colaboración con la Universidad Politécnica de Cartagena (UPCT) y el servicio

de Neonatología del Hospital General Universitario Santa Lucía de Cartagena (HSL).

Esta aplicación, dirigida a pediatras y enfermeros, tiene el objetivo de agilizar, facilitar y precisar el cálculo de dosificaciones, diluciones y sueroterapias de los fármacos en neonatología, con el fin de reducir al máximo la tasa de errores descritos anteriormente.

Para su desarrollo, se utilizaron las plataformas IOS y Android, y los lenguajes de programación empleados fueron Swift [5] y Kotlin [6], respectivamente. La base de datos fue implementada y desplegada mediante Google Firebase, específicamente Firebase Firestore [7], que ofrece una variedad de utilidades para el lanzamiento de aplicaciones móviles y su monitoreo.

1.1. Motivación

La motivación detrás del desarrollo de SAFINEo radica en la necesidad de optimizar y precisar el cálculo de dosificaciones y diluciones de fármacos en neonatología. Los errores en la administración de medicamentos en esta población pueden tener consecuencias graves para la salud de los neonatos, ya que su organismo en desarrollo es particularmente sensible a cualquier variación en la dosis [4]. SAFINEo busca reducir al máximo la tasa de errores, agilizando el proceso de preparación y administración de medicamentos, lo que se traducirá en una atención más segura y eficaz para estos pequeños pacientes.

2. Desarrollo

El desarrollo de SAFINEo se ha realizado dentro de un proyecto conjunto entre de la UPCT y la unidad de Neonatología del HSL. Basado en una estrecha colaboración entre ingenieros y expertos en neonatología, permitió una comprensión profunda de las necesidades y requerimientos específicos del servicio de neonatología. La aplicación se ha diseñado para las plataformas IOS y Android, utilizando los lenguajes de programación Swift y Kotlin, respectivamente. La base de datos se implementó utilizando la plataforma Google Firebase, que proporciona una sólida infraestructura para el lanzamiento de aplicaciones móviles y el seguimiento de su uso (Figura 1).

Para la concepción de la APP se han seguido los estándares y regulaciones más relevantes como son la ley de protección de datos, normas de calidad y seguridad de

datos de salud (ISO 27001), normativas de recetas electrónicas, farmacovigilancia y para historias clínicas electrónicas (HCE) así como estándares de interoperatividad (HL7 y CDA) y de accesibilidad y usabilidad. Finalmente, también se tuvieron en cuenta las regulaciones de la AEMPS en España.

Las herramientas de programación o IDEs (Entornos de Desarrollo Integrado) utilizadas para el desarrollo de la aplicación SAFINeo han sido Android Studio y Xcode. Estas plataformas proporcionan un conjunto de herramientas en una sola interfaz gráfica, que incluye un editor de código fuente, compilaciones locales automatizadas y un depurador para encontrar errores en la aplicación.

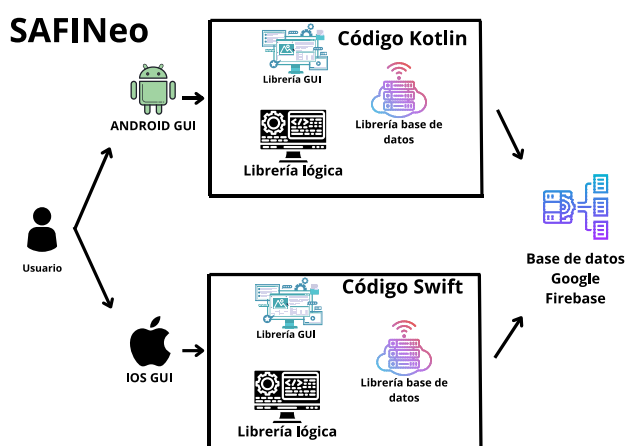


Figura 1. Arquitectura de la APP

Android Studio, el IDE oficial para el desarrollo de aplicaciones Android, está basado en IntelliJ IDEA y ofrece un potente editor de texto y herramientas necesarias para la implementación de aplicaciones. Entre sus características se encuentran un sistema de compilación flexible, un emulador rápido y eficaz, y la integración con GitHub, entre otros.

Por otro lado, Xcode es el IDE creado por Apple para el desarrollo de aplicaciones iOS y macOS. Ofrece una amplia variedad de funciones y herramientas para la creación y edición de código, así como un simulador de dispositivos iOS y macOS para probar y depurar aplicaciones sin necesidad de tener un dispositivo físico.

Los lenguajes de programación empleados en el desarrollo de SAFINeo han sido Kotlin y Swift. Kotlin, un lenguaje de código abierto creado por JetBrains, es utilizado para mejorar la productividad y legibilidad del código en aplicaciones de Android, combinando características orientadas a objetos y funcionales. Swift, desarrollado por Apple, es un lenguaje moderno y multiplataforma utilizado en aplicaciones iOS, macOS, watchOS y tvOS. Ambos lenguajes destacan por su seguridad, velocidad y facilidad de uso.

Para la implementación de la base de datos en SAFINeo, se ha utilizado Google Firebase, específicamente Firestore. Firebase es una plataforma móvil y web adquirida por Google que proporciona diversas

herramientas y servicios en la nube para el desarrollo de aplicaciones. Firestore, una base de datos NoSQL en tiempo real, permite almacenar y sincronizar datos en tiempo real, brindando una experiencia de usuario rápida y fluida. Su escalabilidad y capacidad para manejar grandes cantidades de datos lo convierten en una opción ideal para la aplicación SAFINeo.

Para el diseño de la aplicación se incluyeron en la base de datos todos los fármacos utilizados en la Unidad de Neonatología clasificándolos por clases y familias tras un profundo estudio de características similares.

Para cada fármaco se dispone de las opciones del cálculo de dosificaciones y diluciones en función de las características del paciente como edad gestacional, días de vida y peso. Adicionalmente se incluye información del fármaco relacionada con la presentación, efectos adversos, administración, indicaciones, monitorización e incompatibilidades (Figura 2).



Figura 2. Pantalla principal fármaco Aciclovir

Por otra parte, se dispone de dos opciones/pantallas adicionales:

- **Medicación de urgencia:** Esta opción calcula de manera conjunta la dosificación y la dilución necesaria de los fármacos utilizados para atender de forma rápida y efectiva a los recién nacidos que requieren cuidados intensivos y/o se encuentran en situaciones de urgencia o emergencia. Estos medicamentos se administran en casos críticos, como problemas respiratorios, infecciones y trastornos cardiovasculares, entre otras emergencias médicas que ponen en riesgo la vida del neonato (Figura 3). El objetivo de la

medicación de urgencia es estabilizar al paciente y mejorar su condición en situaciones críticas.

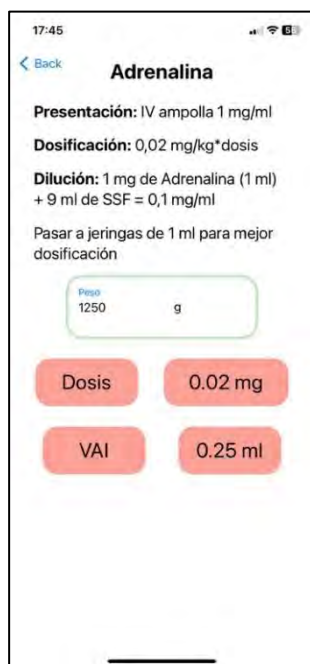


Figura 3. Ejemplo fármaco medicación de urgencia

- Sueroterapia: Esta opción realiza cálculos de las cantidades necesarias a utilizar, extraer y diluir para la preparación del suero hipertónico, heparinizado y glucosado, con el fin de preparar la mezcla prescrita por el facultativo (Figura 4). La sueroterapia consiste en administrar soluciones intravenosas para mantener el equilibrio de líquidos y electrolitos en el neonato.

Esta terapia es fundamental en el cuidado neonatal, ya que los recién nacidos son particularmente susceptibles a desequilibrios hídricos y necesitan un control preciso de sus niveles de fluidos y electrolitos para mantener su salud y bienestar.

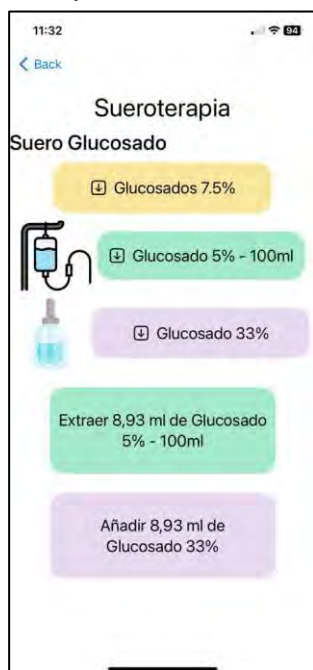


Figura 4. Ejemplo pantalla Sueroterapia: Suero Glucosado

Estas dos opciones adicionales amplían las capacidades de la aplicación, proporcionando herramientas precisas y confiables para el cálculo y administración de medicamentos en neonatología. Su implementación promete mejorar significativamente la atención médica neonatal en situaciones de urgencia y emergencia, brindando a los profesionales de la salud una herramienta esencial en el cuidado y bienestar de los recién nacidos.

3. Estructura

3.1. Patrón de programación VMMV

El patrón VMMV es una variante del patrón Modelo-Vista-Controlador (MVC), que busca desacoplar la lógica de presentación de la vista y mejorar la testabilidad de la aplicación. En VMMV, la comunicación entre la vista y el modelo se realiza a través de un componente intermediario conocido como ViewModel (Figura 5). Esta separación permite un mejor manejo de la lógica de negocio y aumenta la flexibilidad para realizar pruebas unitarias.

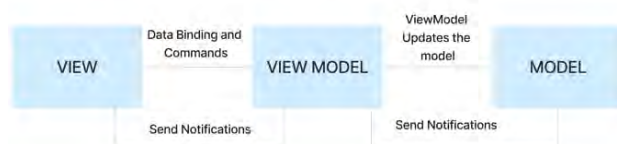


Figura 5. Patrón de programación VMMV

En el contexto de Kotlin, la implementación del patrón VMMV se facilita gracias a la compatibilidad inherente del lenguaje con programación funcional y orientada a objetos. Los ViewModel se encargan de manejar el estado de la vista y procesar las interacciones del usuario, mientras que los Modelos representan la capa de acceso a datos y lógica de negocio. Además, Kotlin ofrece bibliotecas y herramientas que favorecen la adopción del patrón VMMV, simplificando así su integración en el desarrollo de aplicaciones móviles.

En el contexto de Swift, el patrón VMMV se implementa mediante clases y estructuras que representan los modelos y ViewModels respectivamente. La estructura de tipos del lenguaje, junto con el uso de protocolos y extensiones, permite una implementación clara y eficiente del patrón VMMV. Además, la herramienta de enlace de datos nativa de Swift, SwiftUI, facilita aún más la adopción del patrón, permitiendo una interacción fluida entre los componentes de la arquitectura.

El desarrollo del software se realizó utilizando un modelo de desarrollo en V (V-Model) que, aunque basado en un modelo en cascada (Waterfall), enfatiza la realización de pruebas en cada etapa del desarrollo garantizando así la calidad final del producto.

3.2. Estructura base de datos

Firestore es una base de datos NoSQL, orientada a documentos, proporcionada por Google como parte de su plataforma Firebase. Es una solución escalable, flexible y en tiempo real, diseñada para almacenar, sincronizar y recuperar datos para aplicaciones web y móviles. La estructura de Firestore se basa en colecciones, documentos y campos, lo que la hace adecuada para una variedad de

casos de uso, desde aplicaciones simples hasta aplicaciones empresariales complejas.

En SAFINeo existen dos colecciones que engloban la base de datos, Fármacos y Medicación de Urgencia. En el primero se encuentran todos los fármacos presentes en la aplicación, cada fármaco es un documento. En cada documento (fármaco) existen diferentes campos entre los que destacan Nombre, Administración, Efectos adversos, Incompatibilidades, Indicaciones, Monitorización, Presentación... Todos estos campos son de tipo String ya que contienen la información relacionada. Los documentos están también compuestos de otros tipos de campos como son Concentración, Dosis, tipo de dosificación...(Figura 6). Estos campos son de tipo int o float dependiendo de la característica de cada uno. En Medicación de Urgencia, se encuentra los fármacos presentes en esta sección, como anteriormente cada fármaco es un documento y en él se encuentra campos más específicos de cada fármaco.

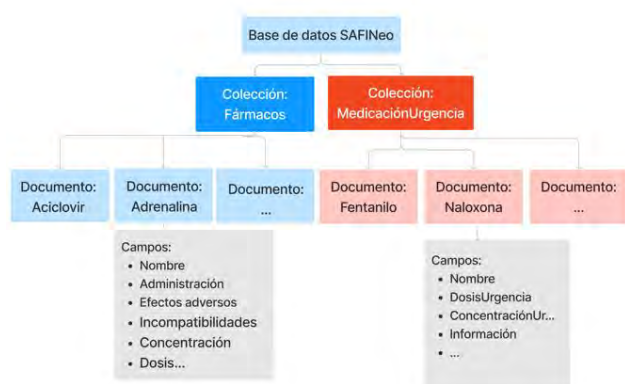


Figura 6. Estructura de la base de datos

4. Proceso de pruebas

Una etapa fundamental en el desarrollo de SAFINeo fue su evaluación en el entorno clínico real del servicio de Neonatología. La aplicación fue puesta a prueba por pediatras y enfermeros especializados, quienes proporcionaron retroalimentación valiosa sobre su funcionamiento y utilidad. Durante las pruebas, se recopilaban datos para validar la precisión de las dosificaciones y diluciones calculadas por SAFINeo, comparándolas con las prácticas convencionales. Los errores en los cálculos de dosificaciones, diluciones y sueroterapia se redujeron a un 0% durante el período de pruebas (Grupo de control de 10 profesionales sanitarios durante 15 días de uso).

5. Conclusiones

La aplicación SAFINeo ha demostrado ser una herramienta prometedora y valiosa en el servicio de neonatología. Su capacidad para calcular con precisión las dosificaciones y diluciones de fármacos ha sido confirmada durante las pruebas realizadas en un entorno clínico real. SAFINeo ofrece una interfaz intuitiva y fácil de usar, lo que facilita su adopción por parte de los profesionales de la salud. Su implementación en el servicio de Neonatología del Hospital Universitario Santa Lucía de Cartagena ha

mostrado resultados alentadores en la reducción de errores y mejora de la seguridad en la administración de medicamentos.

5.1. Trabajo Futuro

A pesar de los logros alcanzados, el desarrollo de SAFINeo sigue siendo un proceso continuo. Entre los trabajos futuros, se contempla la ampliación de la base de datos de medicamentos, incluyendo nuevos fármacos y preparaciones. Asimismo, se buscará la integración de SAFINeo con sistemas de registro de pacientes y la posibilidad de compartir datos con otros profesionales de la salud. La aplicación se ha desarrollado con el objetivo de facilitar su integración en el sistema sanitario actual, incluyendo su conexión con los sistemas de EHR (Electronic Health Record) como el CGM Selene utilizado en el hospital. Se espera que SAFINeo siga evolucionando para seguir mejorando la atención neonatal y convertirse en una herramienta indispensable para los profesionales en el cuidado de los neonatos.

Agradecimientos

Queremos agradecer sinceramente al alumno David Vera Carrillo por su destacada contribución en el desarrollo de la aplicación SAFINeo como parte del desarrollo de su TFM. Su habilidad y dedicación en la programación han sido fundamentales para el éxito de este proyecto científico. Su trabajo ha permitido crear una herramienta invaluable para mejorar la atención médica neonatal. Estamos agradecidos por su compromiso y entusiasmo en esta investigación.

Referencias

- [1] Chatterjee S. et al. Drug utilization study in a neonatology unit of a tertiary care hospital in eastern India. *Pharmacoepidemiology and drug safety*, vol. 16, no 10, 2007, p. 1141-1145 (ISSN:1099-1557).
- [2] Dotta A, Braguglia A, Salvatori G. Pharmacological research in neonatology. *The Journal of Maternal-Fetal & Neonatal Medicine*, vol. 24, no sup1, 2011, p. 44-46 (ISSN: 1476-4954).
- [3] Quesada A, López O, Jiménez E, Cabañas MJ. Guía de administración de fármacos en el período neonatal para enfermería. Elsevier Spain, 2009 (ISBN: 978-84-7592-7756)
- [4] Benitz W. Infectious Disease and Pharmacology: Neonatology Questions and Controversies. Elsevier Health Sciences, 2018 (ISBN: 032354391X).
- [5] García CG, et al. Swift vs. objective-c: A new programming language. *International Journal of Interactive Multimedia and Artificial Intelligence (IJIMAI)*, 2015, vol. 3, no 3, p. 74-81. DOI:10.9781/ijimai.2015.3310
- [6] Horton J. Android Programming with Kotlin for Beginners: Build Android apps starting from zero programming experience with the new Kotlin programming language. Packt Publishing Ltd, 2019 (ISBN: 1789615402)
- [7] Moroney L. The firebase realtime database. *The Definitive Guide to Firebase: Build Android Apps on Google's Mobile Platform*, 2017, p. 51-71 (ISBN: 1484229428).

Modelado, diseño y fabricación de un sistema inyector de aerosoles basado en modelos anatómicos

M. Abellón Caselles¹, D. Ojados González², J. F. Roca González³, M. García Fuentes², H. Flores Aparicio²

¹ Universidad Politécnica de Cartagena, Cartagena, España {marco.abellon}@edu.upct.es

² Servicio de Apoyo a la Investigación Tecnológica (SAIT), Universidad Politécnica de Cartagena, Cartagena, España {lola.ojados, hector.flores, marta.garcia}@upct.es

³ Departamento de Automática, Ingeniería Eléctrica y Tecnología Electrónica, Universidad Politécnica de Cartagena, Cartagena, España {jroca.gonzalez}@upct.es

Resumen

Este estudio trata de diseñar, modelar, y fabricar un sistema inyector de aerosoles basado en modelos anatómicos, para simulación del fenómeno de tos humana.

Para ello, se han buscado alternativas para la obtención del diseño en forma del modelo 3D de una cabeza humana. Para ello, en primer lugar, se llevó a cabo el escaneado 3D de un maniquí con forma humana, y como segunda opción se definió mediante software CAD un proceso de diseño que permite modelar en 3D dicha cabeza. Una vez conseguido el modelo se valoró su reproducción mediante técnicas de fabricación aditiva con equipos de impresión 3D, con el objetivo de acoplar el sistema inyector de aerosoles diseñado y prototipado.

Por otro lado, se diseñó un sistema formado por diferentes dispositivos que permiten generar un flujo que simula el comportamiento del fluido que expulsa una persona en forma de aerosol al toser. El fluido que se expulsa debe cumplir unas condiciones lo más similares a la tos humana, con el fin de que el alcance de este estudio sirva de referencia para un posterior estudio de la propagación de la tos tanto en espacios abiertos como cerrados. Para finalizar, se realizaron pruebas de funcionamiento del sistema inyector de aerosoles obtenido.

1. Introducción

La presencia de la pandemia debida a la aparición del virus SARS-CoV-2 con la que se convive a nivel mundial desde finales del año 2019, genera la necesidad de desarrollar proyectos de investigación en búsqueda de soluciones que afronten los problemas que presenta dicho virus.

Estudios realizados por investigadores expertos en la materia afirman que el SARS-CoV-2 se propaga desde la boca o la nariz de una persona infectada en forma de pequeñas partículas líquidas que expulsa cuando tose, estornuda, habla o respira. Estas pequeñas partículas van desde pequeños aerosoles hasta gotitas respiratorias de mayor tamaño.

Es por ello, que para frenar los contagios se debe estudiar la forma de propagación del virus, su alcance tanto en lugares cerrados como abiertos y la posibilidad de diseñar y fabricar EPIs que permitan a la población mantener unos hábitos de vida lo más cercanos posible al modo de vida que se llevaba antes de la existencia del virus y la aparición de contagios.

2. Objetivos

El presente estudio tiene como objetivo principal modelar, diseñar y fabricar un sistema inyector de aerosoles basado en modelos anatómicos. Este sistema debe recrear la respuesta fisiológica de un paciente de SARS-CoV-2, por lo que el inyector lanzará al exterior un flujo de un fluido que simule una mezcla de aire, saliva y mucosidad y que por tanto tendrá unas propiedades determinadas. Dicho inyector quedará integrado en una réplica de cabeza humana diseñada a partir de modelos reales en la medida de lo posible, que podrá ser materializada mediante técnicas de fabricación digital como es la impresión 3D.

La finalidad del proyecto es la de ayudar al estudio dentro del ámbito científico y sanitario ya que dentro de la sociedad son los campos más cualificados para afrontar la situación, desde la prevención de la población sana hasta la cura de la población ya contagiada. Por ello, desde el campo de la ingeniería se pretende aportar conocimientos que lleven a cabo una rápida y exitosa solución al problema que causa el virus que circula entre toda la población mundial.

3. Materiales y métodos

3.1. Materiales

Para la realización del proyecto se emplearon diferentes tecnologías de software como SolidWorks 2020, 3DSlicer, Meshmixer, Blender, Ansys 2022 R1, Autodesk Inventor 2018, Arduino IDE y Geomagic. Las piezas se diseñaron mediante CAD (diseño asistido por computadora) para poder exportar los archivos contenidos, en diferentes formatos compatibles con otros softwares específicos.

Además, se usaron equipos de fabricación digital, como las impresoras 3D, para la obtención de piezas, y la cortadora láser, para hacer los cortes sobre los materiales para la estructura del dispositivo de toside, instrumentación electrónica para el desarrollo de la parte de control y programación del dispositivo, incluido el material electrónico para la regulación del sistema, como es la placa de Arduino y sus componentes, además de la instalación de aire comprimido para la realización de los ensayos.

En cuanto al método de fabricación, se tuvieron en cuenta normas básicas y recomendaciones de diseño para la

correcta fabricación en impresión 3D: evitar ángulos rectos dado que son grandes concentradores de tensiones, combinar y no intersectar los cuerpos de un mismo modelo para evitar problemas al exportar a un archivo STL, el modelo debe ser un sólido cerrado y no puede contener huecos, todos los elementos del modelo tienen que tener un grosor mínimo, en este caso 0,8 mm, no reproducir detalles menores a 0,3 mm, evitar las partes en voladizo que requieran de uso de soportes, posicionar la pieza en la forma más adecuada para cada proceso de fabricación aditiva para resistir las cargas, adecuar la relación entre resolución y mallado de triángulos en el archivo STL, disponer de una base plana para tener una superficie de apoyo mayor y por tanto conseguir una mejor estabilidad y calidad de impresión.

3.2. Metodología

El procedimiento se separó en 8 fases diferenciadas: (i) desarrollo del inyector, (ii) diseño del modelo anatómico de la cabeza humana, (iii) revisión de los avances, (iv) simulación de funcionamiento, (v) impresión del inyector, (vi) determinación de las conexiones entre las piezas, (vii) regulación electrónica del simulador y (viii) ensayo.

En primer lugar, se usó SolidWorks para el diseño 2D y 3D (Figura 1). Se diseñó una tobera convergente-divergente para elevar la velocidad de manera notable, partiendo de un flujo que penetra en la tobera a una presión inicial superior a la presión atmosférica proporcionada por un sistema de aire comprimido. A la salida de la tobera el flujo presenta una presión inferior a la presión de entrada, pero superior a la presión atmosférica. Como referencia se tomó el inyector diseñado por Bo Zhang en el proyecto “*Diseño y caracterización de un simulador de tos*” desarrollado en el Instituto de Tecnología de Nueva Jersey en 2017 [1]. El inyector funciona como impulsor del fluido, además, se conectó un tubo que va desde el inyector a un reservorio, consiguiendo que a través del efecto Venturi se aspiren gotas gruesas (en comparación con las gotas finas del flujo) desde el reservorio hasta el inyector. El objetivo del reservorio es añadir gotas gruesas al aerosol que contiene las gotas finas, absorbiendo la disolución del reservorio a través de la depresión generada en la garganta. Cuando el flujo que pasa por el inyector se mezcla con la disolución absorbida, ésta es descompuesta en gotas gruesas. Además, se escogió un material hidrofóbico que minimiza el rozamiento del agua con la pared del inyector.



Figura 1. Modelo del inyector en 3D

En la segunda fase se diseñó el modelo anatómico de la cabeza humana. Para la ejecución de esta fase se plantearon

dos opciones. La primera opción consiste en escanear un maniquí en 3D con escáner portátil de marca y modelo ARTEC EVA, generando una nube de puntos, y, a partir de ella, una malla, posteriormente simplificada con el software Meshmixer. Esta malla tratada se exporta con herramientas CAD creando un modelo sólido editable, el cual se puede modificar a libre elección. Finalmente, el modelo deseado se exporta a formato STL para su impresión en 3D. En la otra opción, que fue la desarrollada y elegida, se obtuvo el modelo a partir de 3D Slicer, software de segmentación médica, y se realizó su posterior optimización (Figura 2).



Figura 2. Resultado final del busto modelado en 3D

En la tercera fase se analizó el resultado del modelado 3D del busto y se decidió no llegar a fabricarlo para optimizar recursos, por lo que se sustituyó por un soporte para ensayos de funcionamiento del inyector, modificándose el diseño de este para poder incorporar el reservorio y fabricar el conjunto en futuros estudios.

Para la cuarta fase de simulación de funcionamiento del inyector se usó una versión educacional de un software de mecánica de fluidos. Los resultados no se han incluido en este estudio ya que no se dispone de una versión completa de este software que permita publicar resultados de investigación, pero sí sirvió para realizar comprobaciones básicas. Se propone un estudio más completo de esta fase cuando se disponga de la versión de software adecuada.

En la quinta fase, se imprimió el inyector en la impresora 3D Form 3B, con tecnología de estereolitografía láser de baja fuerza LFS. Primero se importó el archivo STL del modelo a Preform, software de configuración de parámetros de impresión que gestiona la impresora para cada proceso de fabricación. Se cambió la orientación de la pieza a una donde fuera posible configurar soportes en zonas de fácil acceso para su posterior retirada (Figura 3).

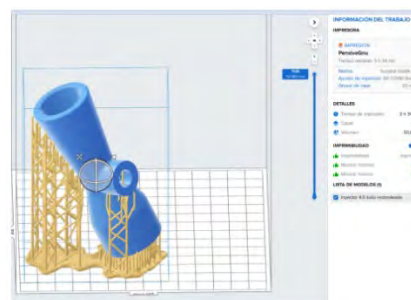


Figura 3. Orientación adecuada para la impresión 3D

Para la impresión se empleó una resina biocompatible denominada ‘Surgical Guide Resin’. Cuando el proceso de impresión finalizó, se extrajo la pieza con la base sobre la que se mantiene y se sometió a un proceso de lavado en alcohol isopropílico durante 15 minutos. Tras el lavado, se puso la pieza en la base del horno de curado, donde recibió calor a 60° mientras rotaba durante 18 minutos. Con esto se obtuvo la pieza final impresa en 3D (Figura 4).



Figura 4. Resultado de la fabricación del inyector

Para la fase seis, realización de las conexiones entre las piezas (Figura 5), se empleó un tubo de tecalán para la conexión entre los distintos elementos del dispositivo. Se usó un compresor que proporciona 5 bares de presión, desde el que se emite aire comprimido que pasa por un nebulizador que arrastra parte de líquido en partículas muy finas. El nebulizador del que se disponía estaba destinado a uso médico y no era compatible con adaptadores estándar para tubos de tecalán, por ello, se procedió a diseñar un adaptador optimizado compatible con el sistema. Para simular la tos humana se insertó una válvula antes de la salida del inyector para que el flujo saliente no fuera continuo. Dicha válvula se conectó por tomas rápidas al tubo de tecalán y estas tomas se unieron a la válvula, cuyos agujeros son de rosca 1/8” NPT. Cuando la válvula es alimentada, cambia el estado a abierto, dejando pasar el flujo de aire con gotas finas a través del tubo de tecalán hasta el inyector. Para acoplarla al inyector, se fabricó otro adaptador que encajaba en el mismo y que se puede encajar y extraer tantas veces como se desee.

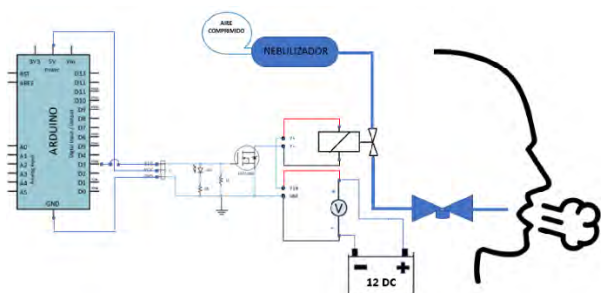


Figura 5. Esquema de las conexiones del sistema

En la fase siete, para la regulación electrónica del simulador (Figura 6), se pretendía conseguir una simulación real del tosido. Un tosido real de un humano consiste básicamente en un proceso en que la glotis se contrae hasta cerrarse, manteniendo en los pulmones y en la cavidad torácica, gran cantidad de aire [2]. Para conseguir un proceso similar, se insertó la válvula cerrada

sin dejar pasar ningún flujo, hasta recibir la señal de abrir durante el tiempo determinado, controlada por Arduino, que, al recibir una entrada dada por el usuario, permite que se abra la válvula durante el tiempo requerido. La señal de salida se programó en una de las salidas digitales siendo enviada al módulo MOSFET, el cual tiene la función de conmutar la fuente de 12V con una válvula de solenoide.

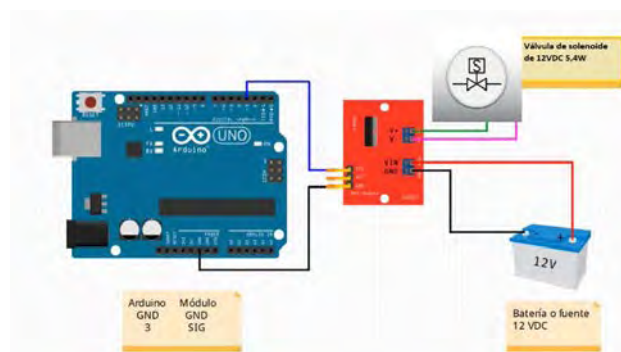


Figura 6. Esquema electrónico

En la última fase se puso en funcionamiento todo el sistema. En primer lugar, del nebulizador, lleno de agua con colorante azul, sale un flujo que se compone de aire y agua, hasta llegar a la válvula de solenoide que se encuentra cerrada. Dicha válvula, regulada por la placa Arduino, reacciona a la entrada introducida por el usuario abriendo durante el periodo de tiempo programado (en este ensayo se situaba entre 10ms - 100ms) y dejando pasar el flujo que entra rápidamente en el inyector arrastrando el líquido del depósito en forma de gotas finas. Frente a la salida del inyector se colocó un papel en blanco donde quedó proyectada la mezcla de aire y gotas de fluido líquido para observar la dispersión, cantidad y tamaño de las gotas.

Para poder rellenar el nebulizador conforme se consumiera el líquido almacenado en su interior, se utilizó una aguja que quedó incrustada en la pared lateral del reservorio. Ya que el volumen a rellenar era mínimo no fue necesaria una sección muy grande, por lo que se probó primero con una aguja de 0,3mm, se comprobó que era demasiado endeble y se aumentó el diámetro de sección de la aguja a 0,8mm. Esta sección era la adecuada ya que estaba equilibrada tanto en caudal como en rigidez para permanecer insertada en la resina, pero no era capaz de atravesar el grosor completo del depósito. Como solución, se taladró la pared del nebulizador con una taladradora Dremel 3000 y con una broca de 0,5mm, previamente a la inserción de la aguja para que quedara ajustada en el orificio de la manera más estable y el mayor tiempo posible.

Paralelamente, como accesorio, se fabricaron dos soportes en impresión 3D para colocar el inyector y se realizó el corte láser de piezas rectangulares de cartón pluma posicionar estos soportes y para sujetar el papel donde se proyectan y se pueden observar las gotas expulsadas por el simulador de tosido (Figura 7).



Figura 7. *Presentación final*

4. Resultados

El primer resultado de este estudio, es el propio simulador de tosido, pero en este punto se describen además los resultados correspondientes a los ensayos realizados con el simulador. Para el primer ensayo se llenó el depósito del inyector y se ejecutaron varios tosidos con presión a 8 bares y tiempo de apertura de la válvula a 100 ms. Se obtuvieron unos resultados bastante alejados de la realidad ya que la cantidad de líquido expulsada era demasiado elevada, llegando incluso a emparar el papel (Figura 8).

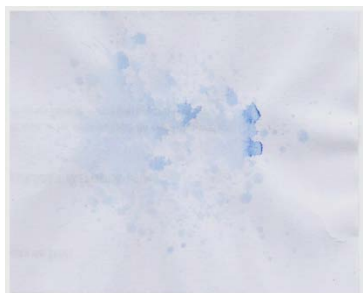


Figura 8. *Gotas expulsadas en el ensayo a 8 bar*

En el segundo ensayo, se cambió la presión a 6 bares manteniendo el mismo tiempo de apertura y con el nivel del depósito no tan lleno, siendo la cantidad y el tamaño de las gotas similar al expulsado en un tosido real (Figura 9).



Figura 9. *Gotas expulsadas en el ensayo a 6 bar*

5. Conclusiones

La principal conclusión que se obtiene tras analizar los resultados de este estudio es que se han logrado los objetivos básicos planteados en la propuesta inicial que hacen posible la funcionalidad del prototipo del simulador diseñado, modelado y fabricado. El proceso de diseño del inyector se ha realizado en dos etapas: una primera etapa donde se ha tomado de referencia el modelo de Bo Zhang [1] y una segunda etapa donde se ha adaptado el modelo a las necesidades de los ensayos planteados para la realización de este estudio. El diseño de un modelo de

cabeza ha supuesto una etapa de investigación y aprendizaje en el ámbito del diseño aplicado a la medicina.

El uso de la tecnología de impresión 3D ha sido fundamental para la fabricación de los prototipos y de los soportes. Se considera un método muy válido ya que no requiere unos conocimientos demasiado avanzados en técnicas de fabricación digital y es económica frente a otros métodos de fabricación, lo que hace que sea una tecnología que da acceso a un rango amplio de usuarios y apta para obtención de prototipos.

En cuanto a la regulación electrónica, se han empleado componentes electrónicos al alcance de usuarios con conocimientos básicos en electrónica y que están disponibles en el mercado a un coste moderado, lo que da accesibilidad al dispositivo. La programación que controla la placa Arduino y que por tanto regula el sistema tiene un código que, por medio de una entrada a través de Monitor Serie es capaz de habilitar una salida digital que abre el paso del flujo durante un tiempo programado.

Finalmente, como trabajos futuros, se propone una investigación en profundidad donde se estudie el flujo en la salida cuando se varíen otros factores, como probar con otro líquido con mayor viscosidad, aumentar la temperatura del líquido o incluso modificar el diámetro de los tubos y que de esta manera se optimice el sistema hasta obtener unos resultados que sean idénticos a la respuesta fisiológica que presenta un humano al nacer. Una vez que se obtengan los resultados buscados, sería el momento de diseñar una cabeza que cumpla con una geometría similar a la humana. Esta cabeza albergará el inyector en su interior y servirá para probar equipos de protección individual como son mascarillas, permitiendo comparar su eficacia en su variedad de diseño y composición, incluso validar nuevos modelos.

Agradecimientos

Se hace necesario agradecer el apoyo y la ayuda prestada por el personal técnico del Servicio de Diseño Industrial y Cálculo Científico (SEDIC) del Servicio de Apoyo a la Investigación Tecnológica (SAIT) y del FABLAB de la UPCT, para la elaboración del presente trabajo. Gracias también a estos servicios centrales de la UPCT y a sus responsables por poner a mi disposición los recursos técnicos necesarios para el desarrollo de este estudio, la obtención del prototipo resultante y la realización de los ensayos pertinentes.

Referencias

- [1] Zhang B, Zhu C, Ji Z, Lin CH. Design and characterization of a cough simulator. *J Breath Res.* 23 de febrero de 2017;11(1):016014.
- [2] Página web de Normon. *La tos: qué es y por qué se produce.* <https://www.normon.es/articulo-blog/la-tos-que-es-y-por-que-se-produce> (Consultada: Mayo 2022).

A case study: Improving face detection in Cartagena Vessel Traffic Service Operators to detect fatigue onset using thermographic images through the YOLO5Face model.

F. Crestelo Moreno¹, J. Roca González², J. Suardiáz Muro², G. Boil³

¹ Marine Science and Technology Department, University of Oviedo, Gijón, Spain, fcrestelo@uniovi.es

² Department of Electronic Technology, Technical University of Cartagena, Cartagena, Spain, jroca.gonzalez@upct.es

³ G. Boil, CITI University, Ulaanbaatar, Mongolia, gregory@citi.edu.mn

Abstract

Fatigue is a common physiological state that can have a significant impact on human performance and well-being, so its early detection is essential in a number of fields, such as maritime safety, which is the subject of this article. Traditional methods of fatigue assessment methods often rely on subjective self-reports or performance metrics, which can lack sensitivity and objectivity. In recent years, thermography has emerged as a promising non-invasive tool that has been validated to detect the onset of fatigue in humans. One way to detect the onset of fatigue is to monitor the temperature of the orofacial area using thermographic imaging. This article provides an overview of the principles of thermography, focusing on the analysis of orofacial thermographic images in real work situations in 16 port control operators in Cartagena. For this purpose, images of these subjects were taken during their watch and the points of interest mentioned above were identified using the YOLO5Face model, a version of YOLO5 trained to recognise faces. As a result of using this model, it was possible to identify and monitor points of interest in thermographic images in an uncontrolled environment.

1. Introduction

Fatigue has always been a concern for the maritime industry. In 1993, the International Maritime Organisation (IMO) adopted an initial Resolution A.772(18) on fatigue factors in manning and safety, which was followed by a number of subsequent recommendations and guidelines. In this Resolution, the IMO recognises that there is no universally accepted definition of fatigue, but generally considers fatigue to be a state of exhaustion, overexertion, sleeplessness or excessive tiredness. Accurate and early detection of fatigue is therefore essential to prevent accidents, enhance safety, optimise productivity and improve quality of life.

According to the European Maritime Safety Agency (EMSA), out of a total of 823 accidents analysed between 2014 and 2020, 89.5% of all occurrences were related to human actions, either at the accident event level or at the contributing factor level [1].

Infrared thermal imaging has been used in medicine since the early 1960s. Working groups within the European Association of Thermology produced the first publications on the standardisation of thermal imaging in 1978 [2].

In the context of human fatigue research, thermography is used to detect changes in skin temperature, mainly in the orofacial region, with particular interest in the area of the nose, where temperature changes may indicate physiological stress or fatigue [3]. When a person experiences fatigue or physical or mental stress, blood flow patterns and heat distribution on the surface of the skin can change. These changes can be visualised and analysed using thermographic imaging.

The main objective of this paper is to describe the methodology for detecting the occurrence of fatigue in vessel traffic service operators (VTSOs) by using model training for thermographic images.

2. Materials

The primary tool employed in this study was the FLIR E-60 thermal imaging camera (FLIR Systems, Inc., Wilsonville, OR, USA). The FLIR E-60 is a state-of-the-art thermal imaging camera designed for high-resolution thermal imaging and temperature measurement applications.

Detector Type	Uncooled microbolometer
Spectral Range	7.5 - 13.0 μm
Temperature Range	-20°C to 650°C (-4°F to 1202°F)
Thermal Sensitivity	<0.05°C at 30°C (86°F)
Resolution	320 x 240 pixels
Image Refresh Rate	60 Hz
Accuracy	$\pm 2\%$ or $\pm 2^\circ\text{C}$ ($\pm 3.6^\circ\text{F}$), whichever is greater

Table 1.- Key features of the FLIR E60 thermal imaging camera

The FLIR E-60 camera was chosen for its exceptional thermal sensitivity and high-resolution capabilities, which were crucial for capturing and analysing temperature variations in the subjects under investigation.

3. Experimental Setup and sample

The experiments were carried out in the Port of Cartagena, in the Traffic Control Room, a controlled environment to ensure accurate thermal data collection, following strict protocols such as those developed by the University of Glamorgan for the recording and analysis of thermographic images in humans [4], which take into account factors such as room size, ambient temperature, humidity, atmospheric pressure and radiation sources.

Sixteen active VTSOs from the Port Control of Cartagena (Spain), a centre attached to a cooperation agreement with the Polytechnic University of Cartagena, participated in this study. These participants have a wide range of age, work experience, or gender [N=16 mean (\pm SD) age = 41 ± 9.27 yrs.; total experience in the workplace = 5 ± 5.85 yrs.; 69 % male].

The study design and procedures were previously reviewed and approved by the Scientific and Ethical Research Committee of the Technical University of Cartagena, and all the volunteers were free to withdraw from the study at any time.

This occupational sector was chosen because it is a complex socio-technical system [6-8], where factors such as interaction with technology, long on-call times, or the high responsibility are triggers for the onset of fatigue.

4. Method

The thermographic images were taken in real work situations using the previously mentioned thermographic protocols.

The camera is positioned at a distance of no more than 1 metre from the subject and at the appropriate angle to frame the orofacial area, but not more than 45 degrees [5](Figure 1).



Figure 1.- Placing the thermal imaging camera for an orofacial study

The method we used to pinpoint the forehead / the nose is the following: we first localise the centre of the eyes and use it as the diagonal of a square whose two others vertices are the forehead and the nose. Thus, the challenge is to localise on each frame the eyes of the subject, with recording where the volunteers are constantly moving to perform their tasks, where some use glasses which are opaque in a thermal image, and with constant passage of other people, as it is a shared centre with the Port

Authority. To automate this task, we choose the AI tool: a version of the YOLO5 model (fifth version of ‘You Look Only Once’ model) was used, more specifically the YOLO5Face model, a version of YOLO5 designed for face recognition [9].

This model can run on a variety of hardware configurations. We ran it on a local CPU (1.4GHz Intel Core i5 quad-core) and it only took a few seconds to a minute per video, depending on the length of the video. The recommended configuration by the developers of YOLO is a GPU with at least 8GB of memory.

The original YOLO5 model is trained to detect faces on RGB images. However, the face recognition from thermal images suffers from the lack of large annotated thermal datasets, especially for ‘in the wild’ images, meaning images in an uncontrolled environment. This difficulty has been overcome by Kuzdeuov et al. [10]. They constituted a dataset, Thermal Face in the Wild (TFW), which is a collection of thermal images in which faces are manually annotated with bounding boxes and the centres of the eyes, nose tips and corners of the mouth. It consists of 9,982 images and 16,509 labelled faces. The novelty of the dataset is the integration of a large amounts of images / faces captured in the wild (5,112 faces in a controlled environment, 1,748 faces in a semi-controlled environment, 9,649 faces in the wild). Faces in many situations were captured, such as occlusion, head pose variations, variations of scales, locations, weather condition, and wearing of a face mask. These situations are known to be challenging. Then the YOLO5Face model was re-trained using this dataset, resulting in a highly accurate model, with average precision scores of 97% on the test dataset of outdoor images. We used their model to our orofacial region identifying purpose (Figure 2).

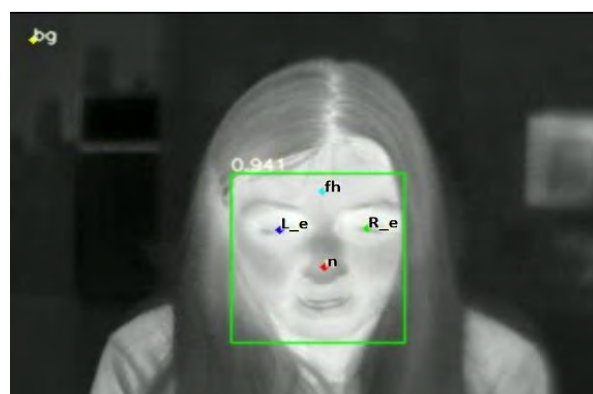


Figure 2.- Capture of the orofacial region identifying the points of interest in one of the VTSOs (bg= background; n= nose; L_e = left eye; R_e= Right Eye; fh= forehead)

To do so, we needed the FLIR thermal video specific format, the SEQ format, to be supported by the frames extracting function and then to be fed into the re-trained YOLO5Face model. It appeared that such a function wasn't provided by the FLIR company, nor available on the internet. To address this need, we rewrote the extracting function to support the SEQ format by integrating the fnv library provided by FLIR, library that process the SEQ

videos. One of the challenges of this process is that thermal frames are given in two formats: the 8-grayscale format, that is then converted in RGB images and given to the YOLO5Face model for face recognition and localisation of Point Of Interests (coordinates in the frame), and the 16-grayscale format that encapsulates the temperature at each pixel of the frame, specifically at the POIs. By combining the use of these two formats, the resulting function can retrieve orofacial coordinates and temperatures from the frames.

A feature added to the extracting function is the ratio argument. It aims to handle heavy videos, allowing only a percentage of frames to be extracted, for instance a ratio of 0.1 meaning that one frame is extracted every 10 frames. Indeed, the FLIR camera used is designed to capture videos with a fps close to 7.5, which means one frame captured each 0.13 s, and some videos lasts for 5 or more hours, that results in a substantial number of frames to process. The temperature does not need such an accurate monitoring. Then the addition of the ratio feature was decided, allowing the temperature to be monitored once per second (ratio close to 0.1) or even once per minute (ratio close to 0.002).

As a result of this improves, the data processing sequence is as follows: first, the extraction function detects the SEQ videos and extracts and processes the frames from the videos in 8 and 16 grayscales format. Then the extracted 8-grayscale frames are analysed by the trained YOLO5Face model. Once we have the coordinates of the landmarks, we read the temperatures on the 16-grayscale frames and record it.

However, during the acquisition of the points of interest, it appeared that two problematic situations occur and must be filtered out.

In the first situation, one or more faces are detected by the model, in which case only the temperature of the first detected face is recorded. We don't have control on which face is detected at first by the model, because it depends on the position of the faces on the frame. Regarding such situations, we make the assumption that any other faces shown is sporadic. Thus, the recorded temperatures might 'jump' when a change of 'face' occurs. This can be filtered out by averaging temperatures over long enough periods (moving average technique), as discussed in the results section.

In the second situation, no face is detected. First, neither the forehead nor the nose coordinates are recorded (null value in the database). For the recorded temperature, several strategies are implemented: record a temperature of 0, record a temperature equal to the reference temperature in the background, record the last recorded temperature of the nose/forehead. The user can then decide which strategy fits better to his analysis. In this study we chose the first option, recording a temperature of 0 when no faces appear.

Finally, once all the videos have been processed, the data is stored in a csv file, one file for each video. The csv file contains a database of the coordinates of the left eye, right eye, forehead, nose and background and the temperatures measured at these points.

Results

To explain the applied method for proceeding the raw collected data, we showcase one of the results obtained. The watch displayed lasts more than 6 hours; it has been recorded at a ratio close to one frame each second. As shown in the figure (Figure 3), the raw data isn't readable, for the reasons raised above.

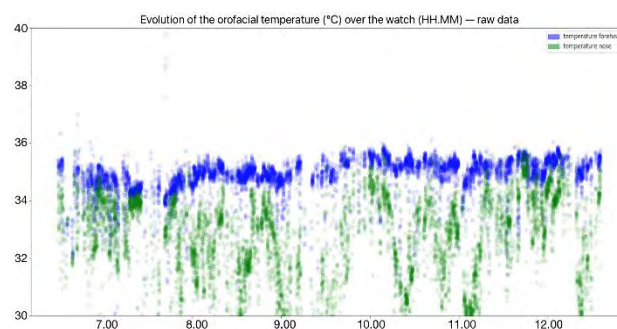


Figure 3.- Nose and forehead temperature raw curves

To make the data more understandable and instructive we will follow a three steps procedure:

First, we won't show the points where the temperature is 0. It means we cut out the points where no faces appear. And for averaging purposes we replace this temperature with the last temperature recorded for each passage where the face is out of the frame.

Second, we try to assess when another face is detected in the frame. To do so we monitor the motion of the left eye of the detected face. The assumption is that a high variation in the left eye position means another face have been detected. The limitation of this assumption is when the other face detected is close enough to the face of the subject so that it could be misinterpreted as a natural movement of the head of the subject. To find these high variations we do a statistical analysis of the movement of the eye: we calculate the variations of position between each two consecutive times 'dX' and then calculate its mean 'm' and its standard deviation 'std'. Then, each point beyond the 'm + 4std' value is considered an outlier in the variations (the number 4, or sometimes 3, is a thumb-rule in the outlier's detection). We then paint in red all the resulting points.

Third, we use a 'moving average' method: for each time, we calculate the average temperature of the temperatures recorded in the window of 5 min around this point (starting from 2.5 min before and until 2.5 min after this point). Under the assumption that each apparition of another face is sporadic (few seconds to one minute at most), the period of 5 minutes seemed reasonable filter out the little jumps observed in the raw data while still representing the temperature of the subject. It also allows to address any fluctuations of the camera recording process.

Following these three steps, we get a much smoother temperature curve for the nose and the forehead, as displayed in figure (Figure 4). We can then identify trends, for instance a decreasing of the temperature for the first half reflecting the fatigue of the subject and then high variations of the temperatures on the second half that could

be interpreted as the subject resisting the fatigue. A few red points appear, meaning either that the used method to identify other faces recording isn't effective or that the showing of other faces is indeed sporadic. But looking at the regularity of the temperature curve, it doesn't seem to impact very much the study and the 'moving average technique' is enough to smooth these singularities.

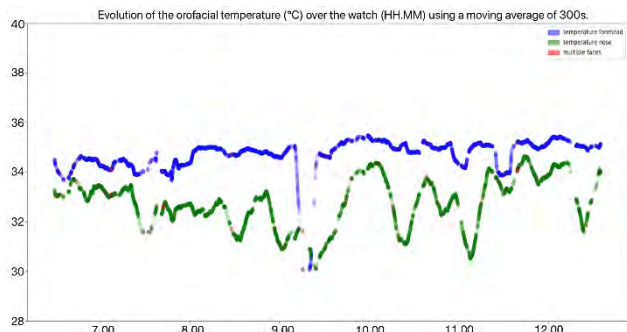


Figure 4.- Nose and forehead temperature curves processed

The results obtained were verified by manually determining the coordinates of the points of interest on the nose and forehead using geometry. For this purpose, anatomical reference structures were used, such as the lines connecting the lateral commissures of each eye to the perpendicular with respect to the nasal tip.

The temperature results obtained between the manual method and the algorithm did not differ by more than a tenth of a degree, which is within the tolerance in this field [11]. The test was performed by checking the temperature records every 15 minutes of recording for each volunteer and comparing them with those obtained every second by the algorithm.

As a result, 200 temperature points were manually extracted to represent the nasal temperature curve, and these results were crossed with the real-time temperature curve obtained by the algorithm, resulting in more than 8000 points, which is unattainable manually. For the showcased watch, the manually recorded temperatures fit well the above curves (Figure 5).

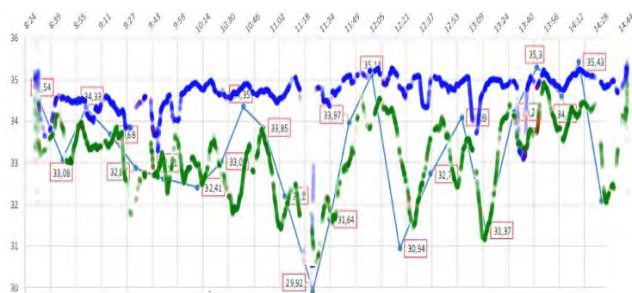


Figure 5.- Manual and automatic temperature-comparing curves

Conclusion

Despite the fact that the recordings were made in a real working situation, and the thermography protocols could not be strictly applied the results are quite consistent and the temperature trends are clear so that these temperature variations could be analysed to determine the occurrence of fatigue. The results showed that nasal temperature was

a reliable indicator of the port coordinators' mental workload factor, which is closely related to fatigue. It decreased when decision making was required and increased during periods of relaxation. However, the results could be improved by making corrections in situations where more than one face is recognised. We cannot predict which face will be recognised first, but we can measure 'gaps' in the coordinates of the reference points. This can improve tracking and eliminate the assumption of intermittent situations with multiple faces.

Acknowledgements

We would like to thank the Spanish Search and Rescue Service and the volunteers of the Port Control of Cartagena for their collaboration in this study.

References

- [1] EMSA, «Annual overview of marine casualties and incidents 2021», 2021. <http://emsa.europa.eu/newsroom/latest-news/item/4266-annual-overview-of-marine-casualties-and-incidents-2020.html>
- [2] Aarts, N.J.M. et al. Thermographic terminology. Acta Thermographica, 1978. Suppl. 2. 1-30
- [3] Marinescu, A.C., Sharples, S., Ritchie, A.C., Sánchez López, T., McDowell, M., Morvan, H.P., 2018. Physiological Parameter Response to Variation of Mental Workload. Human Factors: The Journal of the Human Factors and Ergonomics Society 60, 31–56. <https://doi.org/10.1177/0018720817733101>
- [4] Ammer, K., 2008. The Glamorgan Protocol for recording and evaluation of thermal images of the human body. Thermology International 18, 125–129.
- [5] Peterson, M.D., Joseph D. Bronzino, Donald R. (Ed.), 2013. Medical Infrared Imaging: Principles and Practices. CRC Press, Boca Raton. <https://doi.org/10.1201/b12938>
- [6] Crestelo Moreno, F., Soto-López, V., Menéndez-Teleña, D., Roca-González, J., Suardiáz Muro, J., Rocés, C., Paño, M., Fernández, I., Díaz-Secades, L.A., 2023. Fatigue as a key human factor in complex sociotechnical systems: Vessel Traffic Services. Frontiers in Public Health 11.
- [7] Nuutinen, M., Savioja, P., Sonninen, S., 2007. Challenges of developing the complex socio-technical system: Realising the present, acknowledging the past, and envisaging the future of vessel traffic services. Applied Ergonomics 38, 513–524.
- [8] Relling, T., Lützhöft, M., Hildre, H.P., Ostnes, R., 2019. How vessel traffic service operators cope with complexity – only human performance absorbs human performance. Theoretical Issues in Ergonomics Science 21, 418–441. <https://doi.org/10.1080/1463922X.2019.1682711>
- [9] YOLOv5, “Yolov5,” <https://github.com/ultralytics/yolov5>. 1, 2, 3, 5, 6, 9
- [10] TFW: Annotated Thermal Faces in the Wild Dataset Askat Kuzdeuov,, Dana Aubakirova, Darina Koishigarina, and Huseyin Atakan Varol, Senior Member, IEEE
- [11] Lahiri, B.B., Bagavathiappan, S., Jayakumar, T., Philip, J., 2012. Medical applications of infrared thermography: A review. Infrared Phys Technol 55, 221–235. <https://doi.org/10.1016/j.infrared.2012.03.007>

GB3 as a Novel Therapeutic Target in Neuroblastoma's Alternative Vasculature through Stiffness Modeling

Aránzazu Villasante^{1,2*}, Josep Corominas¹, Clara Alcón¹, Andrea García-Lizarribar^{1,3}, Jaime Mora⁴,
Mónica López-Fanarraga⁵, Josep Samitier^{1,2,3}

¹Institute for Bioengineering of Catalonia (IBEC), Barcelona, Spain.

²Department of Electronic and Biomedical Engineering, University of Barcelona, Barcelona, Spain.

³Biomedical Research Networking Center in Bioengineering, Biomaterials, and Nanomedicine, Madrid, Spain.

⁴Oncology Department, Pediatric Cancer Center Barcelona, Hospital Sant Joan de Deu, Barcelona, Spain.

⁵University of Cantabria-IDIVAL, Santander, Spain.

Abstract

Neuroblastoma (NB) is a childhood cancer from sympathetic nervous system cells. NB exhibits cellular heterogeneity, with N, S, and I-type cells displaying distinct tumorigenic potentials. NB is highly vascularized, and blood vessels can form through various mechanisms, including endothelial transdifferentiation, leading to the development of tumor-derived endothelial cells (TECs) associated with chemoresistance. We lack specific biomarkers for TECs, relying mainly on general endothelial-expressed proteins like CD31 or stemness-related markers, which are crucial for the survival of healthy cells. Therefore, identifying new TEC biomarkers is vital for effective NB therapies.

A stiffness-based platform simulating human arterial and venous stiffness was developed to study neuroblastoma TECs in vitro. N and I-type cells cultured on arterial-like stiffness transdifferentiated into TECs, while S-type cells did not.

We employed TECs derived from N-type NB cells as a model to explore new TEC markers. Our focus was on Gb3, a glycosphingolipid receptor, which is overexpressed in some cancers and correlates with angiogenesis, metastasis, and drug resistance. Remarkably, we found that TECs expressed Gb3, confirming Gb3 as a novel marker for TECs and opening possibilities for targeted therapeutic intervention.

1. Introduction

Neuroblastoma (NB) is a rare and malignant tumor originating from neural crest cells in the sympathetic nervous system, commonly diagnosed in children within the first year of life. It often occurs in the adrenal medulla but can manifest anywhere in the sympathetic nervous system. Approximately 50% of cases are already metastatic at diagnosis, with common sites of metastasis being bone marrow, bone, liver, and skin. More than 40% of children experience relapses despite surgical and chemotherapy treatments¹.

NB is a highly vascularized tumor, and vascularization plays a crucial role in tumor growth, survival, and metastasis. Tumors can create blood vessels through various mechanisms, including angiogenesis. In addition, tumors can apply alternative strategies known as vasculogenic mimicry and endothelial transdifferentiation, which do not involve the host vascular endothelium. During vasculogenic mimicry, tumor cells align to form a vessel-like structure negative for CD31 staining and positive for periodic acid-Schiff staining. Endothelial

transdifferentiation is another alternative mechanism by which cancer cells acquire endothelial properties, such as expression of CD31, and become tumor-derived endothelial cells (TECs)^{2,3}.

TECs are likely involved in chemoresistance in numerous tumors, including NB^{4,5}. However, current antiangiogenic strategies are designed to target the “classical mechanisms” but not the alternative vasculature, producing clinical failures and adverse outcomes^{2,3}. This is partly because we lack specific targetable biomarkers for TECs besides the general endothelial expressed proteins such as CD31 or the stemness-related markers Oct-4 and TNC⁶, also expressed in healthy cells. Thus, it is critical to identify new TECs biomarkers to design effective therapies.

The glycosphingolipid receptor Gb3/CD77, synthesized by the A4GALT gene, has been associated with invasiveness, metastasis, angiogenesis, and drug resistance in various cancers. Blocking Gb3 has reduced tumor angiogenesis in animal models, but its presence in TECs has not been studied^{7,8}.

In this study, we aimed to obtain TECs in a laboratory setting using a platform based on substrate stiffness. Given the stem-like properties observed in neuroblastoma cells, we hypothesized that these cells could transform into TECs when placed on polydimethylsiloxane (PDMS) substrates, similar to healthy endothelial progenitor cells (EPCs)⁹. To achieve this, we developed a biocompatible PDMS-based platform that mimicked the stiffness of both arterial and venous human blood vessels. Our findings revealed that a subset of neuroblastoma cells cultured on the arterial-like PDMS transformed into TECs. On the contrary, the less malignant NB cells did not transform any stiffness conditions. This model system effectively replicated the heterogeneity in neuroblastoma tumors, offering the potential for targeting TECs specifically to prevent resistance and tumor recurrence in pediatric patients.

Using this model system, we investigated the possibility of identifying a new marker of TECs and focused on Gb3. We confirmed the presence of Gb3 in TECs, paving the way for potential therapeutic interventions to disrupt the growth and progression of NB tumors.

2. Materials and Methods

2.1. PDMS-based platform

Polydimethylsiloxane (PDMS) Sylgard 184 (Dow Corning) was mixed in a weight ratio of 1:75 and 1:70 (cross-linker: pre-polymer) to get equilibrium modulus values of $E_{mod} = 101.38 \pm 6.22$ kPa, and $E_{mod} = 140.70 \pm 5.24$ kPa, respectively. PDMS Sylgard 527 was mixed with proportions of 1:1 and 1:2 (A component/B component), corresponding to Equilibrium modulus values of $E_{mod} = 3.39 \pm 0.81$ kPa and $E_{mod} = 7.09 \pm 1.16$ kPa, respectively.

2.2. Cell culture

Neuroblastoma cell line SK-N-BE (2) was cultured in RPMI medium supplemented with 10% FBS and 1% penicillin/streptomycin at 37°C and 5% CO₂ in a humidified incubator.

2.3. Flow Cytometry Analysis

Cells cultured on PDMS of varying stiffness levels were collected, stained with CD31 antibodies (dilution 1:100, DAKO, #M0823) in Hank's Balanced Salt Solution (HBSS) with 2% Fetal Bovine Serum (FBS) for 30 minutes on ice, and analyzed using flow cytometry. Negative controls were included, and data analysis was performed using Flowing Software.

2.4. Immunofluorescence (IF)

Cells were fixed, permeabilized, and blocked before being incubated with primary antibodies for CD31 (dilution 1:50, DAKO, #M0823 or 1:50, Abcam #ab28364) diluted in antibody diluent (Dako, S3022) in a humid chamber overnight at 4°C. Gb3/CD77 (1:50, Biolegend, clon 5B5) was used for co-stains. Afterward, they were incubated with appropriate secondary antibodies and Hoechst for fluorescence or confocal microscopy imaging.

2.5. Quantitative Real-Time PCR (qRT-PCR)

Total RNA from cells was obtained using Trizol (Life Technologies), following the manufacturer's instructions. RNA preparations were treated with “Ready-to-go you-prime first-strand beads” (GE Healthcare) to obtain cDNA. Quantitative real-time PCR was performed using PowerUp SYBR Green Master Mix (Applied Biosystems #A25742). PECAM and A4GALT primers were obtained from the PrimerBank database

2.6. Datasets for genomics analysis.

The R2 Genomics Analysis and Visualization Platform (<http://r2.amc.nl>) was utilized to investigate A4GALT mRNA levels in NB tumors. Data for NB cell lines can be accessed through the GEO ID: GSE28019. The NB primary tumor data originates from Versteeg's Laboratory study of 88 human NB samples (GEO ID: GSE16476.).

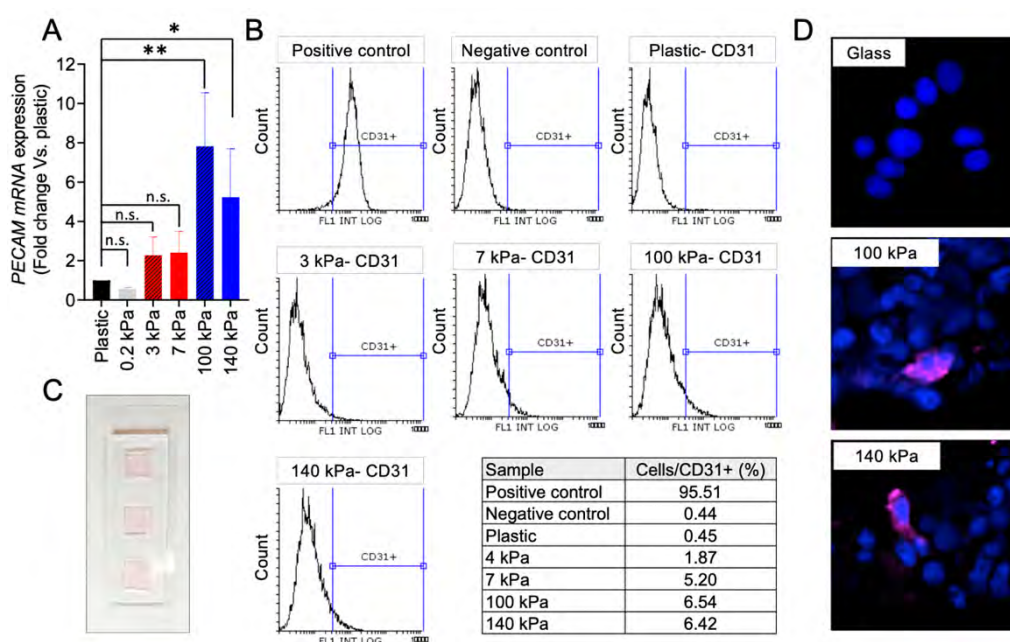


Figure 1. Endothelial cell transdifferentiation neuroblastoma cells on venous-like and artery-like PDMS materials. (A) mRNA levels of *PECAM* in SKNBE (2) cells cultured on the indicated substrates stiffness for 7 days. Fold change was calculated by first normalizing to GAPDH levels in the individual samples and then to the corresponding levels in cells cultured in plastic culture dishes. Data are shown as average \pm SD (n=3). Statistical significance was determined by the two-tailed Student's t-test: *p < 0.05; **p < 0.01; ***p < 0.001; ns, not significant. (B) FACS analysis of CD31 endothelial marker in SKNBE (2) cells cultured on the indicated PDMS substrates for 7 days (n=4). HUVEC were used as positive control. Cells incubated without primary antibody were used as negative controls. (C) Cell culture chamber slide fabricated with PDMS Sylgard 184 mixed in a weight ratio of 1:10 (cross-linker: pre-polymer) with dimensions of 20 mm x 60 mm (W x L), cell culture wells of 10 mm by 10 mm, and 0.80mm thickness. (D) Representative images of CD31 expression (magenta) in SKNBE (2) cells cultured on the indicated substrates; nuclei are stained with Hoechst 33342 (blue) (n=4). Cells were cultured in the chamber slides for 7 days.

3. Results

We cultured NB cells on venous-like and arterial-like PDMS substrates to induce transdifferentiation into TECs. We prepared two different venous-like PDMS using Sylgard 527 with proportions of 1:1 and 1:2 (A component/B component), corresponding to Equilibrium modulus values of $E_{\text{mod}} = 3.39 \pm 0.81 \text{ kPa}$ (the lowest value reported for the venous tissue), and $E_{\text{mod}} = 7.09 \pm 1.16 \text{ kPa}$ (the value published to induce venous differentiation *in vitro*), respectively. Next, we fabricate the arterial mimicking physiological stiffness substrates. We used Sylgard 184 with a composition of 1:70 (curing agent/base) and an $E_{\text{mod}} = 140.70 \pm 5.24 \text{ kPa}$ for obtaining the highest described stiffness value for mammal arteries. Sylgard 184 with a composition of 1:75 (curing agent/base) was used to obtain the optimum *in vitro* value of

substrate stiffness to differentiate EPCs into arteries, $E_{\text{mod}} = 101.38 \pm 6.22 \text{ kPa}$.

NB cells do not express PECAM/CD31 transdifferentiation marker in compliant substrates (**Figure 1A, B**), nor in highly soft materials such as $E_{\text{mod}} = 0.2 \text{ kPa}$ (**Figure 1A**). However, NB cells show significant PECAM mRNA expression levels compared to the plastic and soft controls when cultured for 7 days on arterial-like PDMS substrates (**Figure 1A**). A subpopulation of CD31+ originated in the 100kPa and 140kPa plates was also observed by flow cytometry (**Figure 1B**). For further analysis of the transdifferentiated cells, we performed immunofluorescence for CD31. To this end, we fabricated a device with chambers placed on a glass slide and coated with 1:70 and 1:75 Sylgard 184 (**Figure 1C**). As expected, NB cells did not express CD31 when cultured on the glass control, but we confirmed the presence of TECs in the 100kPa and 140kPa chamber slides (**Figure 1D**). The venous PDMS platform also induced transdifferentiation (**Figure 1B**), but the subpopulation of TECs was smaller than the one obtained with the arterial-like PDMS substrates. Interestingly, the 7kPa plates favored transdifferentiation better than the 3kPa, as described for healthy EPCs differentiation.

Once the model to transdifferentiate NB cells *in vitro* was set up and validated, we focused on the possible expression of Gb3 by TECs. Gb3 is a globoside, so gene expression analysis is unsuitable for detecting Gb3 cellular levels. However, determining mRNA levels of A4GALT, the enzyme that catalyzes the transfer of galactose to lactosylceramide to form globotriaosylceramide, can give us essential information about Gb3 biosynthesis.

Comparing cells and tumors from patients, we observed that A4GALT expression levels were higher in NB tumors with MYCN amplified compared to non-amplified cells in monolayer (**Figure 2A**). The NB cells SKNBE(2), displaying MYCN amplified, expressed significantly higher levels of A4GALT when cultured on the arterial-like stiffness substrates compared to conventional plastic culture plates, strengthening the capability of the model to mimic some patients' features (**Figure 2B**). Interestingly, A4GALT levels were higher in the 100kPa plates, which was also the most optimum condition to obtain TECs *in vitro*. So, the 100kPa condition was used as a model system to evaluate the possible co-expression of CD31 and Gb3. Immunofluorescence studies of CD31 and Gb3 confirmed that TECs co-expressed both markers, which points out Gb3 as a novel target for TECs (**Figure 2C**).

4. Discussion

Due to the TECs subpopulation is not present in traditional plastic dishes, we aimed to design and build a model to mimic endothelial transdifferentiation *in vitro* to identify a novel target molecule expressed in TECs. To develop the simplest model, we proposed to tune the stiffness of the substrate, focusing on the reported stiffness of very well-studied vasculatures, which are arteries (50–150 kPa) and venous tissue (3–50 kPa). Due to the stemness characteristics of NB tumors, we also focused on the established stiffness used to differentiate endothelial

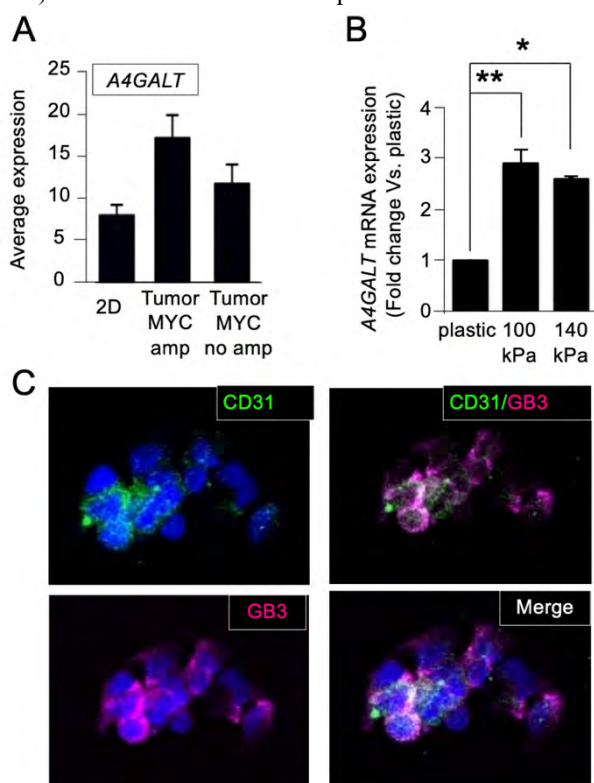


Figure 2. Expression of Gb3 in tumor-derived endothelial cells. (A) Comparison of A4GALT mRNA average expression of human neuroblastoma cells in monolayer (2D) and human neuroblastoma tumors with or without amplified MYCN using BoxPlot analysis based on publicly available expression array data. The analysis was conducted with the AMC Onco-genomics software tool (www.amc.com). (B) mRNA levels of A4GALT in SK-N-BE(2) cells cultured on the indicated substrates for 7 days. The fold change was calculated by normalizing to GAPDH levels within each sample and then relative to A4GALT levels in cells cultured on plastic culture dishes. Data are presented as the average \pm standard deviation (SD) from three independent experiments ($n=3$). Statistical significance was determined by the two-tailed Student's t-test: * $p < 0.05$; ** $p < 0.01$. (C) Representative images illustrating the presence of CD31 (green) and Gb3 (magenta) in SK-N-BE(2) cells cultured on chamber slides coated with the 100 kPa substrate for 7 days; nuclei are counterstained with Hoechst 33342 (blue)

progenitors *in vitro* (7 kPa for venous and 128 kPa for artery).

We demonstrated that NB cells mainly transdifferentiated in stiffer substrates in the range of arteries. Interestingly, the reported stiffness for the collagenous bone (frequent secondary site for NB) is about 100kPa¹⁰, similar to the value observed for the best-evaluated condition for NB transdifferentiation. Our data indicate that NB cells' aggressiveness could be related to the capacity to transdifferentiate. Low tumor vascularity in NB is associated with a favorable prognosis¹¹. Vascularization enables tumors to obtain oxygen and nutrients for growing up and favors cancer dissemination to distant organs^{2,3}. Thus, transdifferentiation into TECs is desirable for malignant progression linked to tumor aggressiveness and metastatic potential. Our results point to that conclusion.

Importantly, traditional cancer models lack cell heterogeneity observed in clinical samples in patients. Adaptation to a plastic plate eliminates subpopulations of cells involved in survival in the harsh environment that is the human body, for example, cells controlling tumor vasculature formation or the cancer stem cell niche⁵. Those subpopulations represent only a small percentage within a tumor cell mass and its microenvironment, but they are critical for tumor development. Considering this, we can argue that drugs that affect cancer cells in plastic culture models may not act over essential subpopulations of cells in native tumors because they are not represented in the *in vitro* model. Thus, it seems critical to recapitulate *in vitro* those populations of cells that have a relevant role in tumor maintenance *in vivo* to assess a specific treatment against them. Therefore, we can propose the stiffness-based PDMS platform as a simple method to recapitulate cell heterogeneity and predict NB aggressiveness and metastatic potential based on the capacity to transdifferentiate into TECs. In addition, the platform can be used as a previous drug-testing step before more complex models to discard ineffective treatments in a fast, cheap, and high-throughput way compatible with standard pharmaceutical industry equipment and protocols.

Recent research has highlighted the role of Gb3 in cancer. Studies have shown that Gb3 expression is upregulated in certain cancer types, including neuroblastoma, although its role in neuroblastoma was not as extensively studied as in other cancer types. Using the platform, we identified Gb3 as a novel marker present in TECs. The overexpression of Gb3 was not only in the TECs but also in the bulk of the tumor, which makes Gb3 a perfect candidate for targeted therapies and diagnostic purposes. Researchers have been exploring several strategies to target Gb3 as a potential therapeutic approach against cancer¹².

5. Conclusions

In conclusion, our study aimed to design and develop an *in vitro* model that mimics endothelial transdifferentiation to identify novel target molecules expressed in tumor-derived endothelial cells. By tuning the substrate stiffness to mimic arterial and venous tissues and endothelial progenitor differentiation conditions, we observed that NB cells predominantly transdifferentiated in stiffer substrates, resembling arterial stiffness. This finding suggests that NB

cells' aggressiveness may be related to their capacity to transdifferentiate into TECs, indicating a potential link between tumor aggressiveness and vascularization. Importantly, the stiffness-based PDMS platform allowed for recapitulating cell heterogeneity observed in clinical samples, often lacking in traditional cancer models. Moreover, our study identified Gb3 as a novel marker present in TECs, making it a promising candidate for targeted therapies and diagnostic purposes. While further research is required to fully understand the role of Gb3 in neuroblastoma and its clinical applications, our findings support the utility of the stiffness-based platform as a tool for predicting NB aggressiveness and evaluating targeted treatment approaches.

ACKNOWLEDGMENTS

The Nanobioengineering Lab is a member of the Biomedical Research Networking (CIBER), Spain, funded by the VI National R&D&i Plan 2008-2011 and the Instituto de Salud Carlos III (RD16/0006/0012; RD16/0011/0022) with the support of the European Regional Development Fund (ERDF), and Generalitat de Catalunya (2023SGR 015452). AV is supported by AECC, the NEN Association, and the R+D+i plan PID2020-117977RA-I00.

REFERENCES

1. Matthay, K. K. *et al.* Neuroblastoma. *Nat Rev Primer* 2, 16078 (2016).
2. Krishna Priya, S. *et al.* Tumour angiogenesis-Origin of blood vessels. *Int J Cancer* 139, 729–35 (2016).
3. Belotti, D., Pinessi, D. & Taraboletti, G. Alternative Vascularization Mechanisms in Tumor Resistance to Therapy. *Cancers* 13, 1912 (2021).
4. Pezzolo, A. *et al.* Tumor origin of endothelial cells in human neuroblastoma. *J Clin Oncol* 25, 376–83 (2007).
5. Villasante, A. *et al.* Vascularized Tissue-Engineered Model for Studying Drug Resistance in Neuroblastoma. *Theranostics* 7, 4099–4117 (2017).
6. Pezzolo, A. *et al.* Oct-4+/Tenascin C+ neuroblastoma cells serve as progenitors of tumor-derived endothelial cells. *Cell Res* 21, 1470–1486 (2011).
7. Celi, A. B., Goldstein, J., Rosato-Siri, M. V. & Pinto, A. Role of Globotriaosylceramide in Physiology and Pathology. *Front. Mol. Biosci.* 9, 813637 (2022).
8. Desselle, A. *et al.* Anti-Gb3 Monoclonal Antibody Inhibits Angiogenesis and Tumor Development. *PLoS ONE* 7, e45423 (2012).
9. Xue, C. *et al.* Substrate stiffness regulates arterial-venous differentiation of endothelial progenitor cells via the Ras/Mek pathway. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1864, 1799–1808 (2017).
10. Murphy, C. M., O'Brien, F. J., Little, D. G. & Schindeler, A. Cell-scaffold interactions in the bone tissue engineering triad. *Eur. Cell. Mater.* 26, 120–132 (2013).
11. Chlenski, A., Liu, S. & Cohn, S. L. The regulation of angiogenesis in neuroblastoma. *Cancer Lett.* 197, 47–52 (2003).
12. Birklé, S. *et al.* Inhibition of tumor angiogenesis by globotriaosylceramide immunotargeting. *Oncol Immunology* 2, e23700 (2013).

Potential role of Calcineurin isoform A-β1 in neurovascular niche stem cell dynamics as a therapeutic target

V. Llorente¹, E. Bello-Arroyo¹, M. López-Olañeta², E. Fernández-Cortés¹, E. Lara-Pezzi², M. Desco^{1,2,3,4,*}, M.V. Gómez-Gaviro^{1,3,4,*}

¹ Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), Madrid, Spain.

² Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, Spain.

³ Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid, Spain.

⁴ Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Leganés, Spain.

*e-mail: desco@hggm.es / vgomez@hggm.es

Abstract

Calcineurin A-β1 (CnAβ1) is a noncanonical isoform of calcium-based protein phosphatase Calcineurin with a unique C-terminal domain that sets it apart from the canonical forms and results in completely different functions and interactions, sometimes directly opposing the effect of standard Cn. The actual roles of CnAβ1 in different tissues have been barely studied so far, however, and any processes it could be part of in the brain in particular are currently unknown. Our preliminary data suggests CnAβ1 could play a role in neural stem cell (NSC) regulation and dynamics; as the natural capabilities for adult neurogenesis and post-injury repair of the brain and its NSCs are limited, any insight on these processes and the potential involvement of CnAβ1 in them could provide new venues for therapeutic and neuroengineering approaches, as they concern neurodegenerative pathologies and cerebrovascular lesions. Further research will be carried out with advanced 3D imaging and tissue clearing techniques in order to verify and expand these findings.

1. Introduction

Calcineurin (Cn) is a protein phosphatase that controls Ca²⁺ dependent cell signalling processes through most tissues and is highly conserved through eucaryote organisms [1]. Its most notable role is part of the adaptive immune response, through the activation of T cells through NFATs [2]. Cn is a heterodimer with several isoforms which retain a high similarity between them, where free Ca⁺ binds to the B subunit, activating the A subunit and allowing Calmodulin to bind to it as well. However, the splicing variant Calcineurin A-β1 (CnAβ1) contains a unique C-terminal domain (Figure 1) that partakes in completely different processes; its actual roles have only recently begun to be studied [3], [4], and this research has mostly focused on its role in the heart, where standard Cn takes part in cardiac hypertrophy processes but CnAβ1 has the opposite effect [5]. Its potential relevance in neural processes remains completely unexplored thus far.

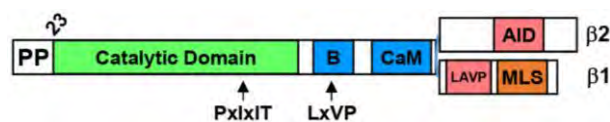


Figure 1: CnA structure, showcasing the alternative C-terminal domain on CnAβ1 compared to the canonical CnAβ2 form. B: subunit-binding domain, CaM: calmodulin-binding domain, LAVP: inhibitory sequence, AID: autoinhibitory domain, MLS: membrane localization sequence [6].

Neural stem cells (NSCs) are multipotent self-renewing cells that give rise to all the different components of the central nervous system (CNS) during embryonic development, but they are not limited to early life stages, as NSCs persist into adulthood and retain some neurogenic activity. These remain in specialized neurovascular niches, which harbour a specific environment to maintain, complement and aid the activity of NSCs [7]–[11]. NSCs also activate in response to lesions or illness, and mount and attempt to repair and regenerate after pathologies like stroke [12], although these are not enough for a complete recovery. NSCs are also negatively affected by neurodegenerative diseases, with their decline potentially being one of the causes of the cognitive symptoms [13]. Thus, studying the functions and dynamics of NSCs and neural progenitor cells (NPCs) could be key in mitigating the effects of these pathologies, finding new treatments, and unlocking the neuroregenerative potential present in the neurovascular stem cell niches.

Preliminary results from the study of a knockout model for CnAβ1 (described and validated in [5]) have shown that its expression is increased in NSCs, hinting at a possible role in regulation of stem cell processes. As the only innate source of neurogenesis and possible post-injury repair present in the brain, any factors in NSC dynamics may hold the key for future therapeutic targets and possible treatment or diagnosis methods.

2. Materials and Methods

Adult male CnA β 1^{Δi12} [5] (KO) and WT C57BL6/J mice were used for this study. A first set of animals (N=6) was injected with Bromodeoxyuridine (BrdU) daily for a week, then sacrificed two weeks later, to allow the BrdU to be lost by all actively dividing cells and only remain in slow-cycling ones. The second set (N=3) received a single BrdU injection 24h before sacrifice, to label proliferating cells during that time period.

Animals were subjected to transcardial perfusion [14] and their organs were extracted and fixed in 4% paraformaldehyde (PFA) overnight. Brains were then embedded in 2% agarose blocks and cut into 70 μ m sections using a vibratome. Sections were labelled with fluorescent antibodies using an indirect immunofluorescence protocol [15]. Primary antibodies used in this study were primed against Sox2 and BrdU.

Samples were analyzed with a Leica TCS-SPE confocal fluorescence microscope. Images of the samples were acquired and processed with the Fiji image processing software. The number of antigen-positive cells within the niche area was quantified for each image, then averaged over the available images. For non-nuclear markers, the amount of antigen-positive area was quantified and divided over the area of the niche itself in each image, then averaged over the available images.

All experimental procedures involving animals included in this project were carried out by qualified personnel in conformity with the “3Rs Principle” and with European Union Directive 2010/63/EU, under PROEX 269.1/22, and were approved by Hospital Gregorio Marañón’s Research Committee and Ethical Committee for Animal Experimentation (ES280790000087).

3. Results

Different cellular markers were studied in order to detect the effects of CnA β 1 knockout on NSC niche cell populations, as any disruptions or differences would imply a role of CnA β 1 in those processes.

The Sex Determining Region Y-box 2 (Sox2) transcription factor is key in maintenance of embryonic and neural stem cells by upholding self-renewal capabilities, and thus was used as a marker for NSCs and NPCs [16].

KO samples displayed a higher amount of Sox2+ cells compared to WT ones (**Figure 2**), thus implying larger populations of NSCs and NPCs being present in the niches of CnA β 1-deprived animals in standard conditions.

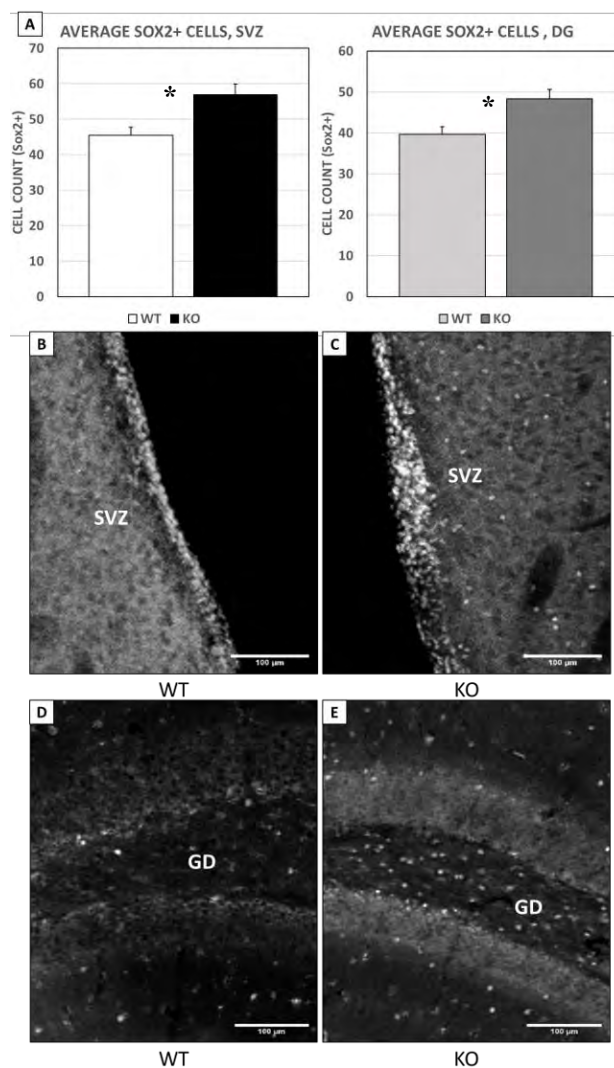


Figure 2: A) Average Sox2+ cell count in neurovascular niches (SVZ and DG), comparison between WT and KO CnA β 1. N=9. $p < 0.005$. Error bars correspond to SEM. B,C,D,E) Sox2+ cell populations, 20x magnification confocal microscopy, Gray: Sox2+, B) WT SVZ, C) KO SVZ, D) WT DG, E) KO DG.

Bromodeoxyuridine (BrdU) was used to label slow-cycling cells, corresponding to quiescent newly-generated NSCs [17], as opposed to proliferating NPCs; this was done to check the level of self-renewal displayed by NSCs in KO and WT animals. KO samples showcased a higher amount of BrdU+ slow-cycling cells in the NSC niches compared to WT (**Figure 3**).

This result could be explained by either a higher amount of NSCs being generated by symmetric division from existing NSCs compared to WT, or a higher pre-existing NSC population with roughly the same rate of self-renewal events as the WT.

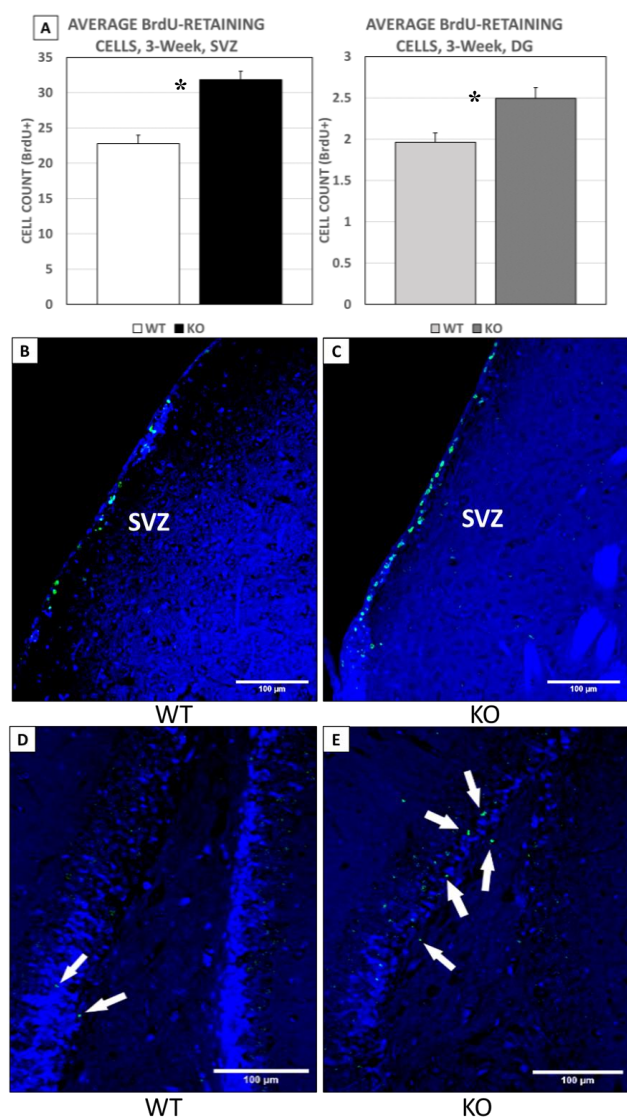


Figure 3: **A)** Average slow-cycling BrdU+ cell count (identified as quiescent NSCs) in neurovascular niches (SVZ and DG), comparison between WT and KO CnA β 1. N=6. $p < 0.005$. Error bars correspond to SEM. **B,C,D,E)** Slow-cycling BrdU+ cell populations, 20x magnification confocal microscopy. Blue: DAPI (nuclear marker), Green: BrdU, **B)** WT SVZ, **C)** KO SVZ, **D)** WT DG, **E)** KO DG.

BrdU was also used to label any and all cells dividing in the neurovascular NSC niches in a comparatively short period of time of 24 hours, to check for the overall rate of cell proliferation. In this case, no significant differences were observed between WT and KO (**Figure 4**).

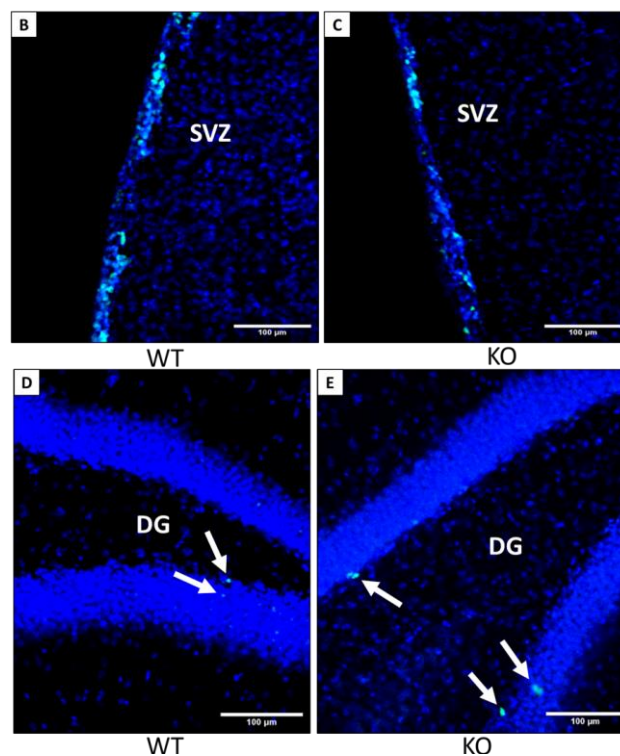
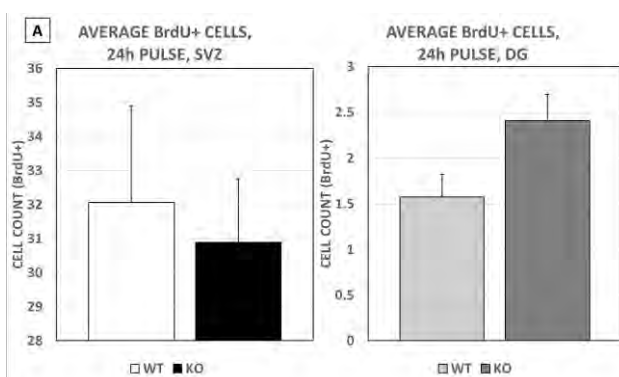


Figure 4: **A)** Average short-pulse (24h) BrdU+ cell count in neurovascular niches (SVZ and DG), comparison between WT and KO CnA β 1. N=3. Not significant. Error bars correspond to SEM. **B,C,D,E)** Short-pulse BrdU+ cell populations, 20x magnification confocal microscopy. Blue: DAPI (nuclear marker), Green: BrdU, **B)** WT SVZ, **C)** KO SVZ, **D)** WT DG, **E)** KO DG.

These similar levels of cell proliferation in the niche between KO and WT, when taking into account the results of the previous markers, would point at the altered processes being related to either the rate of NSC self-renewal through symmetric division (with CnA β 1 directing NSCs towards differentiation and thus higher NSC population depletion) or to the baseline NSC population present in the niche (with CnA β 1 regulating the amount of stem cells present in the niche); these are thus the processes where CnA β 1 is more likely to play a significant role.

4. Conclusion

Calcineurin A- β 1 may play a significant role in the neurovascular stem cell niche by regulating NSC dynamics such as proliferation and population maintenance. Any insights on this topic could open the door for the use of CnA β 1 as a target for therapeutic and neuroengineering approaches in conditions where a more robust regenerative response from NSCs and/or a higher population is desired. As such, further research will be carried out with expanded cell marker targets and techniques, including the use of advanced 3D imaging methods such as SPIM in combination with optical tissue clearing protocols.

Acknowledgements

This work has been funded by Instituto de Salud Carlos III (ISCIII) through the projects DTS22/00030 co-funded by the European Union, and PT20/00044 co-funded by European Regional Development Fund "A way to make

Europe”. It has also been supported by Ministerio de Ciencia e Innovación through the grant PLEC2022-009235 funded by MCIN/AEI /10.13039/501100011033 and by the “European Union NextGenerationEU/ PRTR”, and through the grant PID2021-127033OB C21/MCIN/AEI/10.13039/501100011033.

References

- [1] F. Rusnak and P. Mertz, “Calcineurin: form and function,” *Physiol. Rev.*, 2000.
- [2] J. Roy and M. S. Cyert, “Identifying new substrates and functions for an old enzyme: Calcineurin,” *Cold Spring Harb. Perspect. Biol.*, vol. 12, no. 3, 2020, doi: 10.1101/cshperspect.a035436.
- [3] R. Bond, N. Ly, and M. S. Cyert, “The unique C terminus of the calcineurin isoform CNA1 confers non-canonical regulation of enzyme activity by Ca²⁺ and calmodulin,” *J. Biol. Chem.*, vol. 292, no. 40, pp. 16709–16721, Oct. 2017, doi: 10.1074/jbc.M117.795146.
- [4] J. M. Gomez-Salineró, P. G. Pavía, and E. Lara-Pezzi, “CnA β 1 shifts cardiac metabolism,” *Aging*, vol. 11, no. 3. United States, pp. 839–840, Jan. 2019. doi: 10.18632/aging.101789.
- [5] L. Padrón-Barthe *et al.*, “Activation of Serine One-Carbon Metabolism by Calcineurin A β 1 Reduces Myocardial Hypertrophy and Improves Ventricular Function,” *J. Am. Coll. Cardiol.*, vol. 71, no. 6, pp. 654–667, 2018, doi: 10.1016/j.jacc.2017.11.067.
- [6] X. Li *et al.*, “Calcineurin A β -Specific Anchoring Confers Isoform-Specific Compartmentation and Function in Pathological Cardiac Myocyte Hypertrophy,” *Circulation*, vol. 142, no. 10, pp. 948–962, Sep. 2020, doi: 10.1161/CIRCULATIONAHA.119.044893.
- [7] L. C. Fuentealba, K. Obernier, and A. Alvarez-Buylla, “Adult neural stem cells bridge their niche,” *Cell Stem Cell*, vol. 10, no. 6, pp. 698–708, Jun. 2012, doi: 10.1016/j.stem.2012.05.012.
- [8] M. V. Gómez-Gaviro, R. Lovell-Badge, F. Fernández-Avilés, and E. Lara-Pezzi, “The vascular stem cell niche,” *J. Cardiovasc. Transl. Res.*, vol. 5, no. 5, pp. 618–630, Oct. 2012, doi: 10.1007/s12265-012-9371-x.
- [9] M. V. Gómez-Gaviro *et al.*, “Betacellulin promotes cell proliferation in the neural stem cell niche and stimulates neurogenesis,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 4, pp. 1317–1322, Jan. 2012, doi: 10.1073/pnas.1016199109.
- [10] C. E. Scott *et al.*, “SOX9 induces and maintains neural stem cells,” *Nat. Neurosci.*, vol. 13, no. 10, pp. 1181–1189, Sep. 2010, doi: 10.1038/nn.2646.
- [11] V. Llorente, P. Velarde, M. Desco, and M. V. Gómez-Gaviro, “Current Understanding of the Neural Stem Cell Niches,” *Cells*, vol. 11, no. 19, p. 3002, 2022, doi: 10.3390/cells11193002.
- [12] O. Lindvall and Z. Kokaia, “Neurogenesis following stroke affecting the adult brain,” *Cold Spring Harb. Perspect. Biol.*, vol. 7, no. 11, Nov. 2015, doi: 10.1101/cshperspect.a019034.
- [13] J. Terreros-Roncal *et al.*, “Impact of neurodegenerative diseases on human adult hippocampal neurogenesis,” *Science (80-.)*, vol. 374, no. 6571, pp. 1106–1113, Nov. 2021, doi: 10.1126/science.abl5163.
- [14] J. Wu, Y. Cai, X. Wu, Y. Ying, Y. Tai, and M. He, “Transcardiac perfusion of the mouse for brain tissue dissection and fixation,” *Bio-protocol*, vol. 11, no. 5, p. e3988, Mar. 2021, doi: 10.21769/BioProtoc.3988.
- [15] K. Harb, M. Bertacchi, and M. Studer, “Optimized immunostaining of embryonic and early postnatal mouse brain sections,” *Bio-protocol*, vol. 11, no. 1, p. e3868, Jan. 2021, doi: 10.21769/BioProtoc.3868.
- [16] H. Suh, A. Consiglio, J. Ray, T. Sawai, K. A. D’Amour, and F. H. H. Gage, “In Vivo Fate Analysis Reveals the Multipotent and Self-Renewal Capacities of Sox2+ Neural Stem Cells in the Adult Hippocampus,” *Cell Stem Cell*, vol. 1, no. 5, pp. 515–528, Nov. 2007, doi: 10.1016/j.stem.2007.09.002.
- [17] K. Obernier *et al.*, “Adult Neurogenesis Is Sustained by Symmetric Self-Renewal and Differentiation,” *Cell Stem Cell*, vol. 22, no. 2, pp. 221–234.e8, Feb. 2018, doi: 10.1016/j.stem.2018.01.003.

A microphysiological model designed for the research of angiogenic sprouting in an extracellular matrix

A. Noguera-Monteaudo^{1*}, J. Ferré-Torres¹, A. Lopez-Canosa², A. Vilche^{1,2}, P. Barceló¹, B. Van Durme³, Y. Dong³, E. Xuriguera¹, R. Rodríguez^{1,2}, S. Van Vlierberghe³, E. Engel², A. Hernandez-Machado¹ and O. Castaño^{1,2}

¹ University of Barcelona, 08028, Barcelona, Spain

² Institute for Bioengineering of Catalonia, Spain

³ Ghent University, Ghent, Belgium.

Abstract

Angiogenesis is an important biological process for vascular development, as well as being involved in different health problems such as cancer, inflammatory processes, infections, and some autoimmune diseases, among others. In recent years, microphysiological devices have been proposed to study angiogenesis due to their ability to accurately recreate in vivo microenvironmental conditions. This study presents a platform to study the early stages of angiogenesis together with an in-silico model, presenting an innovative approach that allows us to better analyze the dynamics of angiogenesis. The presented platform allows the formation of an angiogenic gradient through a porous hydrogel, thus provoking the angiogenic response of endothelial cells and therefore studying the process under the desired conditions. In addition, the use of the two-photon polymerization technique has allowed us to print 3D hydrogels with the desired structure within microfluidic devices. Our results show that this microphysiological device, together with the developed mathematical model, is a valuable tool for studying the complex process of angiogenesis.

1. Introduction

Angiogenesis is a complex biological process based on the formation of new blood vessels from pre-existing ones. This process is not only crucial in tissue growth and repair but also implicated in a wide variety of health issues, such as cancerous growth, implant revascularization, malignant disorders, inflammatory problems, ischemia, infections, and immune diseases, making it a highly researched process [1]. In the field of cancer, the inhibition pathways of angiogenesis have been studied to minimize cancer growth [2-4], while in regenerative medicine, various angiogenic materials and factors have been investigated to control and enhance revascularization [5-6]. Despite extensive research on angiogenesis, there is still a lack of understanding of the mechanisms involved in this process.

Angiogenesis depends on the interactions of endothelial cells (EC) with the extracellular matrix (ECM), which has a scaffolding role and guides the process, and angiogenic factors. Normally, the angiogenesis process is initiated through a hypoxic situation that generates a vascular endothelial growth factor (VEGF) gradient, which promotes EC proliferation and migration. [7] Due to the influence on cells of their environment, the fabrication and characterization of artificial ECM has become of great interest. For this, many collagen hydrogels and fibrin-based hydrogels have been used as ECM due to their

biocompatibility and the possibility of controlling their properties. [8-10]

However, recent advancements in microfluidic technology have provided a valuable tool for investigating angiogenesis in vitro [10-11]. Microfluidic systems allow replicating in the in vitro microenvironment, providing a controlled platform for studying angiogenesis. These platforms allow the control of individual parameters, which makes them a good option to study the effect of different parameters on angiogenesis.

In this work, a microphysiological device [12] has been developed, along with a mathematical model [13], to study the early stages of angiogenesis in EC through a porous hydrogel. At the cellular level, experiments are performed to study the angiogenic response and EC progression through a fibrin hydrogel using a VEGF gradient. The properties of the fibrin hydrogel are characterized both at the physical level, using rheology; and at the structural level, staining the fibrin with NHS-Rhodamine and observing the structure in the confocal microscope. In addition, based on these experiments, a pass-field mathematical model of a single equation is developed to predict this same advance, and thus study the temporal evolution more effectively. This model focused on modeling the dynamics and morphology of budding, providing a new framework to understand cell migration and the influence of biomaterial structure.

Finally, with the use of two-photon polymerization (2PP) [14], we have managed to print 3D hydrogels with structures designed within the microfluidic devices, which will allow us to study the effect of different geometries on angiogenesis and blood vessel formation.

2. Materials and methods

2.1. Fibrin hydrogel preparation

Fibrin hydrogel was prepared from commercial lyophilized fibrinogen (Sigma Aldrich, >75%) and bovine thrombin (Thermo Fisher Scientific, High purity grade). A 2.5 mg/mL fibrinogen solution was prepared by dissolving lyophilized fibrinogen in Dulbecco's phosphate-buffered saline (DPBS). Bovine thrombin was dissolved in DPBS to obtain a final concentration of 100 U/ml. The two solutions were mixed at a ratio of 100:2 v/v (fibrinogen: thrombin). The mixture was maintained for 15 min at 37 °C to allow de-polymerization.

2.2. Characterization of the 3D structure

The Leica TCS SP5 confocal microscope equipped with a 40X/1.4 N.A. (numerical aperture) lens was used to elucidate the microstructural properties of the 3D networks of the hydrogel. To enable visualization of fibrinogen fibers, NHS-rhodamine (Thermo Fisher, >90%) was reacted with fibrinogen before polymerization of the hydrogel. It was allowed to react so that the ester groups of NHS-rhodamine react with the fibrinogen amine groups, allowing visualization of the hydrogel in the confocal excited sample with a wavelength of 561 nm at a laser power of 20% and detected at 570-615 nm. The void ratio was evaluated using Fiji/ImageJ software [15-16]. For void ratio analysis, confocal images were binarized by applying a threshold, and the void ratio was calculated from the area of black pixels (pores) concerning the total area in a stack of images.

2.3. Rheology

For the rheological measurements, an Anton Paar MCR 702 Dynamic Mechanical Analyzer was used. Viscosity was measured as a function of shear rate at a constant temperature of 20 °C and 37 °C. For elastic shear modulus (G') and viscous shear modulus (G''), a 20 mm plate with a 0.5 mm gap was used, corresponding to a sample volume of 0.2 ml. These measurements were carried out at a constant temperature of 37 °C and an excitation frequency of 1 Hz, varying the amplitude of deformation from 0.001 to 100%.

2.4. Device design and fabrication

The angiogenesis-based platform design was performed using CAD software (AutoCAD 2019) (Figure 1). The design consists of a central chamber (1300 μm wide, 8800 μm long, 150 μm high) connected to two lateral channels (750 μm wide, 150 μm high) that have deposits (radius 3 mm) at the ends.

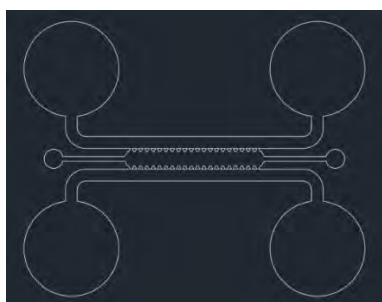


Figure 1. Horizontal section of the design and cell culture model of the microfluidic platform for the study of angiogenesis response of EC.

The master mold was fabricated on a silicon wafer using standard photolithography techniques with an SU8-3050 photoresist (MicroChem) in a cleanroom. PDMS (SYLGARDTM 184 Silicone Elastomer Kit) was used for the fabrication of the device. Elastomer and curing agent were weighed and mixed in a ratio of 10:1 w/w (elastomer: curing agent). It was then degassed and cured overnight at 65 °C, before peeling the PDMS from the master. The

devices were cut into individual devices and punched. Before adhering the PDMS to a coverslip using an air plasma chamber (Harrick Plasma PCD-002-CE) for 30 seconds, the devices were cleaned. Finally, the chips were heat treated overnight at 85 °C to stabilize the bond and sterilized for 30 minutes in ultraviolet radiation (UV 254 nm).

2.5. Cell culture and angiogenesis quantification

The mixture of fibrinogen and bovine thrombin was injected into the central chamber and left at 37 °C for 15 minutes to allow polymerization. Rat endothelial progenitor cells (rEPC) were resuspended and introduced into one of the side channels of the chip. Then, the chips were incubated for 45 minutes in an upright position to allow cell attachment to the fibrin hydrogel. After the incubation period, the other channel was hydrated by adding a cell medium. Finally, the reservoirs were filled with medium and kept at 37 °C, 5% CO₂ incubator for 3 days. The desired VEGF concentration was introduced into the channel where the cells were not found from the second day and were maintained until the end of the experiment. Endothelial cells were fixed on the fourth day with 4% paraformaldehyde (Sigma Aldrich) for 15 min after washing cells with sterile PBS. The cells were permeabilized using a 0.5% Triton X-110 solution, before incubating the cells with phalloidin TRICT and DAPI. Imaging was performed on a Leica TCS SP5 confocal microscope. The DAPI was excited with UV light at a laser power of 20% and detected at a wavelength of 420-470 nm; and TRICT phalloidin was excited with a wavelength of 561 nm at a laser power of 20% and detected at a wavelength of 570-615 nm. The images were processed using ImageJ software and the Angiogenesis Analyzer plugin [17] to quantify the angiogenic response under different conditions.

2.6. Two-photon polymerization

Scaffolds of a gelatin-based ink (HYDROTECH INX U100 (BioInx, Ghent, Belgium) were processed by 2PP at $\lambda=800$ nm at 80 mW using a NanoOne Bio printer (UpNano, Vienna, Austria). Scaffolds were previously designed using Fusion 360 (Autodesk, CA, US) and loaded in the device as STL file.

3. Results

3.1. Characterization of the 3D structure

The use of Rhodamine allows the visualization of the structure of the fibrin hydrogel network within the microfluidic device and under the same conditions as in the cell experiments (Figure 2 (a)).

Images show the fibrinogen fibers in red. In addition to an irregular structure, the images show a very fibrillar structure with a very open scaffold providing high interconnectivity between the pores.

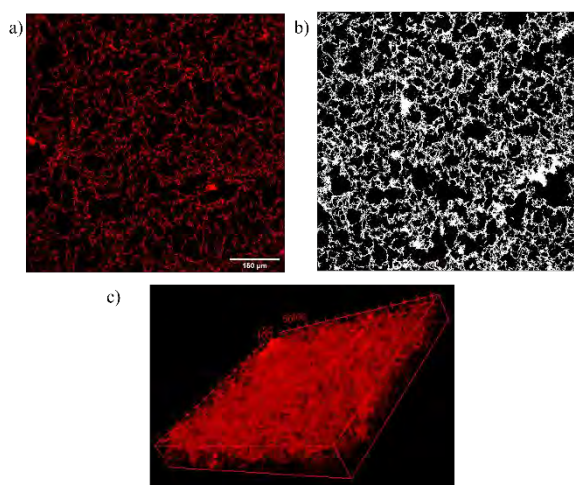


Figure 1: NHS-Rhodamine staining for 3D network analysis of fibrin hydrogel. (a) Confocal image of fibrin hydrogel networks inside the microfluidic chip with 20X objective. (b) Image with brightness adjustment and threshold applied to the calculation of voids. (c) 3D rendering from a stack of z-images of the fibrinogen hydrogel network in ImageJ.

It can be observed that the pores are not very well defined and have irregular shapes and discontinuities in the walls, making it difficult to define a pore size. From different image stacks, the porosity of the material is analyzed by applying a threshold (Figure 2 (b)). From the analysis of the images, the pore volume fraction is found to be $75.37 \pm 1.29\%$. Furthermore, it is possible to create a reconstruction from the image stacks to visualize the structure of the fibrin hydrogel in 3D (Figure 2 (c)).

3.2. Rheology

The mechanical properties of the fibrinogen hydrogel are analyzed with a rheometer. The viscosity of the fibrin hydrogel is analyzed at 20 °C (room temperature) and 37 °C (cell growth temperature) as a function of shear rate. At the two temperatures, it can be observed that the viscosity decreases with increasing shear rate, thus obtaining a shear-thinning response. It can be observed that the decrease in viscosity as a function of shear rate shows a logarithmic trend. It is observed that the viscosity decreases with increasing temperature. At 20 °C, the viscosity ranges from 7 Pa·s to 100 mPa·s; and at 37 °C, the viscosity ranges from 5 Pa·s to 10 mPa·s. Storage modulus (G') and loss modulus (G'') are also determined at 37 °C. From these results, a storage modulus (G') with a value of around 200 Pa and a loss modulus (G'') with a value of around 40 Pa are obtained. At low deformations, the storage modulus is larger than the loss modulus, indicating that the material has elastic behavior, and the hydrogel network properties are measured. As the deformation increases, we can see that there is a point where G' and G'' intersect, indicating that the material begins to have a more fluid behavior (yield point). This phenomenon is irreversible, and the gel is destroyed at the end of the rheometer measurements, so a new hydrogel is used for each experiment. Through rheometry, the different mechanical properties of fibrin hydrogel have been obtained (Table 1).

	20 °C	37 °C
Viscosity range	7 Pa·s - 100 mPa·s	5 Pa·s - 10 mPa·s
Storage Modulus (G')	-	≈ 200 Pa
Loss Modulus (G'')	-	≈ 40 Pa

Table 1. Mechanical properties of fibrin hydrogel

3.3. Angiogenesis analysis

The angiogenic response of the ECs inside the microfluidic device is compared for a control situation (without VEGF angiogenic factor), and a situation with a concentration of 0.1 µg/ml of VEGF (Figure 3).

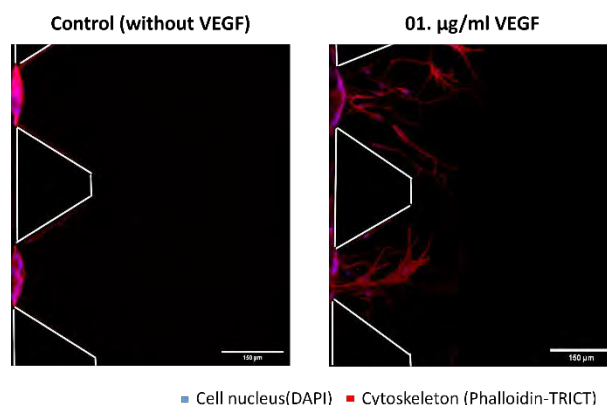


Figure 3: Analysis of the cellular response of rEPCs under different angiogenic conditions on the microfluidic chip. Comparison of cellular response of rEPCs without VEGF (Control) with a cellular response with VEGF concentration of 0.1 µg/ml.

To compare quantitatively, the migration distance parameter is used as the distance from the initial point to the farthest point of each segment. By doing this, we obtain a migration of $45.22 \pm 25.48 \mu\text{m}$ for the control and $290.20 \pm 52.93 \mu\text{m}$ with a VEGF concentration of 0.1 µg/ml. This demonstrates that the microfluidic system presented can generate a VEGF gradient that stimulates the angiogenic response of EC and, therefore, studying angiogenesis.

3.4. Two-photon polymerization in microfluidic device

Due to the autofluorescence exhibited by the photoinitiator included in the formulation of the 2PP ink, the structure within the microfluidic device can be observed using a fluorescent microscope and thus verify the printed structure (Figure 4).

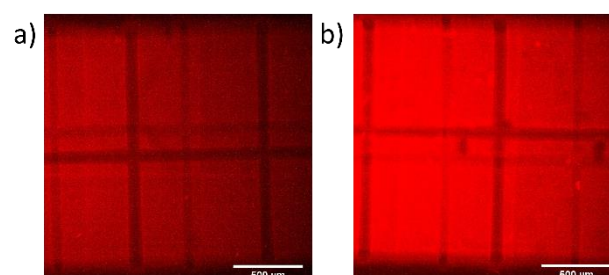


Figure 4: *Autofluorescence analysis of gelatin-based hydrogel to analyze the 3D structure of the 2PP-printed hydrogel inside the microfluidics device. Images of the structure for a concentration of (a) 12% and (b) 15% with 5X objective in a fluorescent microscope.*

In these images, we can see that the desired structure has been printed inside the microfluidic device successfully with the two concentrations of gelatin. In both cases, you can see the channels of different sizes and heights that cross from one side to the other and the two long channels that cross perpendicularly. In general, it has been seen that with the 15% concentration, there are more problems with bubbles that can affect the printed structure. This demonstrates that using the 2PP technique we can 3D print hydrogels with the desired structure inside microfluidics devices.

Conclusions

In conclusion, microfluidic technology has the potential to revolutionize our understanding of angiogenesis by providing a controlled platform to study the mechanisms involved. The presented microfluidic chip is a step towards this goal and allows the investigation of individual parameters in a controlled microenvironment. Furthermore, the possibility of printing engineered 3D structures within microfluidics devices using two-photon polymerization presents a wide variety of possibilities.

Acknowledgment

We gratefully acknowledge the financial support from the Spanish Ministry of Science and Innovation (MICINN) and the Spanish State Research Agency (AEI) through grants RTI2018-097038-B-C22, PID2021-124575OB-I00 and PDC2022-133918-C22, and the the financial support from the European Union's Horizon Europe research & innovation program (EIC-2021-PATHFINDER-OPEN-01-01-101047099 4DBR). Finally, we also appreciate the support of Research Foundation Flanders (FWO) (1SH3W24N).

References

- [1] A.M.A.O. Pollet and J.M.J. den Toonder, "Recapitulating the vasculature using Organ-on-Chip technology," *Bioengineering*, (2020).
- [2] Ansari, M.J., Bokov, D., Markov, A. et al. Cancer combination therapies by angiogenesis inhibitors; a comprehensive review. *Cell Commun Signal* 20, 49 (2022).
- [3] Rajabi M, Mousa SA. The Role of Angiogenesis in Cancer Treatment. *Biomedicines*. 2017; 5(2):34.
- [4] R. K. Jain, D. G. Duda, J. W. Clark, and J. S. Loeffler, "Lessons from phase III clinical trials on anti-VEGF therapy for cancer," *Nat. Clin. Pract. Oncol.*, vol. 3, no. 1, pp. 24–40, (2006)
- [5] Saberianpour, S., Heidarzadeh, M., Geranmayeh, M.H. et al. Tissue engineering strategies for the induction of angiogenesis using biomaterials. *J Biol Eng* 12, 36 (2018)
- [6] Hanieh Shokrani, et al. Cell-Seeded Biomaterial Scaffolds: The Urgent Need for Unanswered Accelerated Angiogenesis, *International Journal of Nanomedicine*, 17: 1035-1068, (2022).
- [7] Pitulescu, M. E., Schmidt, I., Giaimo, B. D., Antoine, T., Berkenfeld, F., Ferrante, F., et al. Dll4 and notch signalling couples sprouting angiogenesis and artery formation. *Nat. cell Biol.* 19, 915–927, (2017).
- [8] Blache, U., and Ehrbar, M. Inspired by nature: Hydrogels as versatile tools for vascular engineering. *Adv. wound care* 7, 232–246, (2018).
- [9] S. Chung, R. Sudo, P. J. MacK, C. R. Wan, V. Vickerman, and R. D. Kamm, "Cell migration into scaffolds under coculture conditions in a microfluidic platform," *Lab Chip*, vol. 9, no. 2, pp. 269–275, (2009).
- [10] E. Akbari, G. B. Szychalski, and J. W. Song, "Microfluidic approaches to the study of angiogenesis and the microcirculation," *Microcirculation*, vol. 24, no. 5, pp. 1–8, (2017).
- [11] K. Haase and D. Roger, "Advances in on-chip vascularization," (2017).
- [12] A. Lopez-Canosa et al. Microfluidic 3D platform to evaluate endothelial progenitor cell recruitment by bioactive materials. *Acta Biomater.* 151:264-277, (2022)
- [13] Ferre-Torres J., Noguera-Monteaquedo A., Lopez-Canosa A., Romero-Arias J. R., Barrio R., Castaño O., Hernandez-Machado A. Modelling of chemotactic sprouting endothelial cells through an extracellular matrix. *Frontiers in Bioengineering and Biotechnology*, 11, 2296-4185, (2023).
- [14] S. O'Halloran, A. Pandit, A. Heise, A. Kellett, Two-Photon Polymerization: Fundamentals, Materials, and Chemical Modification Strategies. *Adv. Sci.* 2022, 2204072.
- [15] Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. Fiji: An open-source platform for biological-image analysis. *Nature Methods*, 9(7), 676–682, (2012).
- [16] Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*, 9(7), 676–682. <https://doi.org/10.1038/nmeth.2019>
- [17] Carpentier G, Martinelli M, Courty J and Cascone I. Angiogenesis Analyzer for ImageJ. 4th ImageJ User and Developer Conference proceedings. Mondorf-les-Bains, Luxembourg. ISBN: 2-919941-18-6 : 198-201, (2012).

Producción de membranas de Bruch artificiales de fibroína de seda mediante *Dynamic Dope Destabilization Spinning* para la creación un Epitelio Pigmentario de la Retina funcional

S. Martínez Rodríguez^{1,2,12}, N. Jemni-Damer^{1,2,3}, N. Soldado Vilches², A. Guedán-Durán^{1,2,3}, AM. Quintana Prego^{1,2}, F. Panetsos^{1,2,3,4}, D. González-Nieto^{2,4,5}, J. Pérez-Rigueiro^{2,4,5,6,13}, F. Rojo^{2,4,6,13}, G. Guinea^{2,4,5,6,13}, A. Virtuoso⁸, G. Cirillo⁸, M. Papa⁸, C. Largo-Aramburu⁹, F. Armadá-Maresca¹⁰, G. Kumar Srivastava^{11,12}

¹ Neuro-computing & Neuro-robotics Research Group, Complutense Univ of Madrid, Spain

² Center for Biomedical Technology (CTB). Universidad Politécnica de Madrid. 28223

Pozuelo de Alarcón (Madrid), Spain

³ Neural Plasticity Research Group, Institute for Health Research San Carlos Clinical Hospital (IdISSC), Madrid, Spain

⁴ Silk Biomed SL, Madrid, Spain

⁵ Biomedical Research Networking Center in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain

⁶ Departamento de Ciencia de Materiales. ETSI Caminos, Canales y Puertos. Universidad

Politécnica de Madrid. 28040 Madrid. Spain

⁷ Dept. Photonic Technology & Bioengineering, ETSI Telecommunication, Polytechnic Univ of Madrid, Spain

⁸ Division of Human Anatomy – Neuronal Networks Morphology Lab, Dept. Mental, Physical Health and Preventive Medicine, Univ of Campania “Luigi Vanvitelli”, Naples, Italy

⁹ Experimental Surgery. University Hospital, La Paz (IdiPAZ), Madrid Spain

¹⁰ Ophthalmology Service, La Paz University Hospital, Madrid, Spain

¹¹ Departamento de Cirugía, Oftalmología, Otorrinolaringología y Fisioterapia, Facultad de Medicina, Universidad de Valladolid, Valladolid, Spain

¹² Instituto Universitario de Oftalmobiología Aplicada, Facultad de Medicina, Universidad de Valladolid, Valladolid, Spain

¹³ Biomaterials and Regenerative Medicine Group, Instituto de Investigación Sanitaria del

Hospital Clínico San Carlos (IdISSC). Calle Prof. Martín Lagos s/n. 28040 Madrid. Spain

Resumen

La retina, capa interna del ojo, tiene como función principal recibir las señales lumínicas del exterior y transmitir las, en forma de señales nerviosas, al cerebro. La capa externa de la retina, el epitelio pigmentario de la retina (EPR), alberga una serie de células polarizadas, cuya función principal es actuar como barrera hematorretiniana externa, esencial para el correcto funcionamiento de los fotorreceptores. La parte basal de las células del EPR, se encuentra en contacto con la membrana de Bruch, matriz multicapa que separa la retina de la coroides.

A día de hoy, existen múltiples patologías degenerativas oculares que generan alteraciones estructurales y funcionales, que provocan la muerte celular y desencadenan pérdida de la visión. En estadios avanzados de estas enfermedades resulta insuficiente administrar cualquier molécula, por lo que es necesario restaurar el tejido dañado. Para ello, el uso de un biomaterial con las características adecuadas es requisito indispensable. En el presente trabajo, se utiliza fibroína de seda (FS), un biopolímero biocompatible y biorreabsorbible procedente del gusano de seda doméstico *Bombyx mori*. La seda procedente de los capullos se procesa y se fabrican mallas de fibras en diferentes orientaciones, en paralelo y en entramado. Para ello se utiliza la técnica de hilado *Dynamic Dope Destabilization Spinning* (D3S), que es un sistema que permite obtener fibras de altas prestaciones mecánicas. Estas mallas,

actúan como soporte artificial para las células del EPR, que son posteriormente sembradas sobre estas para crear un biohíbrido funcional.

1. Introducción

En la actualidad, nos enfrentamos a un desafío considerable en el ámbito de la salud visual: numerosas patologías degenerativas del sistema visual carecen de un tratamiento definitivo. Un ejemplo paradigmático de esta problemática es la Degeneración Macular Asociada a la Edad (DMAE), la cual representa la causa principal de ceguera irreversible en personas mayores en países desarrollados [1]. En las etapas iniciales de esta enfermedad, se observa un deterioro principalmente en dos componentes críticos del ojo: la membrana de Bruch y el EPR. Estos elementos desempeñan un papel fundamental, ya que el EPR secreta factores esenciales para la supervivencia y el funcionamiento adecuado de las células de la neuroretina, así como para el desarrollo de los vasos sanguíneos en la coroides [2]. Además, las células del EPR actúan como barrera hematorretiniana externa debido a sus uniones estrechas, que están organizadas por proteínas como ZO-1, -2 y -3. Además, estas células desempeñan funciones clave, como

transportar nutrientes, desechos, iones y gases entre los vasos sanguíneos coroideos y los fotorreceptores. También, fagocitan los segmentos externos de los fotorreceptores, que se renuevan constantemente [3].

Sin embargo, hasta la fecha, el enfoque terapéutico predominante para abordar estas enfermedades implica la aplicación repetida de inyecciones intravítreas. A pesar de que estas inyecciones pueden ralentizar el avance de la enfermedad, no logran detenerla por completo. Además, este tratamiento se caracteriza por su invasividad, lo que afecta significativamente a sus vidas cotidianas [4]. Dado que estas afecciones son crónicas y requieren un tratamiento constante, se convierten en un problema de salud pública con un impacto socioeconómico significativo. Estas circunstancias resaltan la urgente necesidad de desarrollar una terapia innovadora que no solo detenga la progresión de la degeneración, sino que también repare y proteja el tejido dañado, permitiendo la recuperación de la visión en los pacientes.

Actualmente, existen algunas alternativas a los tratamientos moleculares. Principalmente, la terapia celular y terapias de medicina regenerativa. La primera de estas presenta la principal limitación de la falta de soporte físico y la baja supervivencia celular a largo plazo. En el caso de las membranas de Bruch artificiales que se han estudiado hasta la fecha, estas son esencialmente en formato de film y presentan como principal limitación las pobres propiedades físico-mecánicas que estas poseen [5].

Por este motivo, nuestro trabajo se enfoca en el desarrollo de una membrana de Bruch artificial que sirva como sustrato para la creación de un EPR funcional mediante mallas de FS. Este biopolímero, obtenido a partir de los capullos de los gusanos de seda, presenta propiedades como neuroprotección, capacidad antiinflamatoria y propiedades antioxidantes. Además, posee unas propiedades mecánicas excepcionales, logradas gracias al método de hilado D3S, que permite obtener fibras de altas prestaciones. Es biocompatible, lo que permite la adhesión, migración y proliferación de células en el andamio sin provocar respuestas inmunitarias adversas y, además, es biodegradable y biorreabsorbible.

2. Métodos

2.1. Fabricación de los biomateriales

Para la fabricación de nuestros biomateriales se utilizó fibroína de seda, que se extrajo a partir de los capullos del gusano *Bombyx mori*. Se trocearon en pequeños fragmentos y se eliminó la sericina. A continuación, se llevó a cabo una serie de diálisis para concentrar la solución final. Una vez obtenida, fue procesada mediante la técnica de hilado D3S. El formato fabricado para este trabajo fue mallas de fibras en paralelo y de fibras en entramado.

2.2. Estudio de las propiedades físico-mecánicas de las fibras

Se tomó una fibra individualmente y se colocó en un soporte de aluminio. La fibra se fijó cuidadosamente en el

soporte de un bastidor metálico. La máquina comenzó a tirar de la fibra, mientras que registró los datos de tiempo y masa soportada hasta que la fibra se rompió y terminó el ensayo. Los datos registrados se procesaron y se calculó la tensión, deformación y el módulo de elasticidad de cada fibra mediante las siguientes fórmulas:

$$\sigma = s(e + 1); e = v t/L; s = (mg)/\pi R_0^2; \varepsilon = \ln(e + 1)$$

Donde: σ : tensión real, s = tensión ingenieril, e : deformación ingenieril, v : velocidad (fijada en 1mm/s), t : tiempo; L_0 : longitud inicial, m : masa, g : fuerza de la gravedad, R_0 : radio, ε : deformación real.

2.3. Cultivo de la línea celular

Para el estudio se utilizó una línea celular humana del epitelio pigmentario de la retina: ARPE-19 (*Arising Retinal Pigment Epithelia*). Se cultivó a 37 °C y 5% de CO₂ en medio de cultivo rico en glucosa (*Dubelcco's Modified Eagle medium: Nutrient Mixture F12*) (1:1, 12634010, Gibco) con un 10% de suero fetal bovino (FBS, 10100, Gibco) y un 1% de penicilina/estreptomina (15140122, Gibco). Las células crecieron en placas P100 hasta alcanzar un 80% de confluencia.

2.4. Diferenciación celular

Previo a su utilización para los diferentes experimentos, se diferenciaron las células mediante medio de cultivo DMEM F12 suplementado con FBS (10%), L-glutamina (1%), penicilina/estreptomina (1%), N1 (1%), taurina (1%), nicotinamida (10mM), hidrocortisona (0,04%) y aminoácidos no esenciales (1%). Se cultivaron las células a 37 °C y 5% de CO₂ durante dos semanas. Una vez que las células alcanzaron la confluencia, se disminuyó la cantidad de FBS al 3%.

2.5. Integración de las células en el biomaterial

Las mallas fueron montadas en *CellCrown inserts* para una placa multipocillos P24. Se utilizaron *Transwell* comerciales con el mismo diámetro y tamaño de poro conocido para el grupo control (8 μm, *Corning Costar 3422, Sigma Aldrich*). Las células, una vez confluentes, se levantaron con tripsina-EDTA 0,05% (25300062, Gibco). Se centrifugaron a 125 g, a 21 °C durante 5 min. Se cultivó una densidad de 6×10^4 células por *insert* y *Transwell*. Las mallas con las células se dejaron incubando a 37 °C y 5% de CO₂ durante 30 min para evitar que las células difundieran al pocillo. Después, se añadió medio de cultivo de diferenciación a cada pocillo y se cultivaron las células durante 12 días, renovando el medio cada 2-3 días.

2.6. Ensayo de viabilidad celular XTT®

Se mezclaron 40 μL de reactivo de acoplamiento de electrones (PMS, Metil sulfato de N-metil dibenzopirazina) con 2 ml de XTT. Se preparó la mezcla en 4 ml de medio de cultivo específico. A continuación,

se añadieron 750 µL de la preparación a cada pocillo y se incubó durante 3 horas a 37 °C y 5% de CO₂. Transcurrido este tiempo, se agitó suavemente la placa y se midió la absorbancia espectrofotométrica mediante un lector de microplacas (ELx800, BioTek). La longitud de onda para medir la absorbancia del producto de formazán fue 450 nm.

2.7. Estudio de la funcionalidad de las células: formación de uniones estrechas (ZO1)

Las mallas biohíbridas fueron lavadas 3 veces con 1X PBS (*Phosphate Buffered Saline*). Se fijaron las células con una solución en gradiente de paraformaldehído (PFA) al 2% durante 10 min, y, al 4% durante el mismo tiempo a temperatura ambiente. Una vez fijadas, fueron lavadas de nuevo con 1X PBS para eliminar restos de PFA. Se utilizó el anticuerpo anti-ZO-1 (*Zónula occludens* o proteína de uniones estrechas) hecho en *rabbit* (1:500, ab221547, abcam) para marcar las uniones estrechas entre las células del EPR. Las muestras se incubaron con el anticuerpo primario hasta el día siguiente a 4 °C y en agitación suave. El anticuerpo secundario fue CyTM3 *AffiniPure Goat Anti-Rabbit IgG* (H+L) (1:500, 111-165-003, Jackson ImmunoResearch), que se incubó 4 horas a 4 °C cubriéndolo de la luz. Se lavaron las muestras y se realizó contratinción con Hoechst. Se lavaron de nuevo y se montaron en portaobjetos, que se cubrieron con *Fluoromount Aquous Mounting* (F4680, Sigma-Aldrich).

3. Resultados

3.1. Propiedades físico-mecánicas de las fibras

Se estudió la capacidad de deformación de las fibras de FS en respuesta a diferentes tensiones aplicadas. Como se observa en la figura 1, en la que se representan los datos de deformación real (ϵ) frente a la tensión real (σ), la curva consta de dos regiones diferenciadas: una región de régimen elástico y una de régimen plástico. Tras la realización del ensayo de tracción en aire de cada fibra, se calcularon los valores de tensión y deformación real, gracias a las fórmulas descritas anteriormente. Los resultados obtenidos, confirmaron que las fibras se estiraron hasta que la tensión alcanzó aproximadamente 16 MPa, donde comenzó la región de régimen plástico. En esta última, las fibras continuaron deformándose de manera irreversible, hasta alcanzar el valor de tensión máxima, que fue aproximadamente 18 MPa, tras el cual se rompieron. Se calculó la recta de la región de régimen elástico, cuya pendiente corresponde al módulo de elasticidad. Este fue del orden de 10³, mayor que la membrana de Bruch, cuyo módulo de elasticidad es de entre 7 y 19 MPa (figura 1)

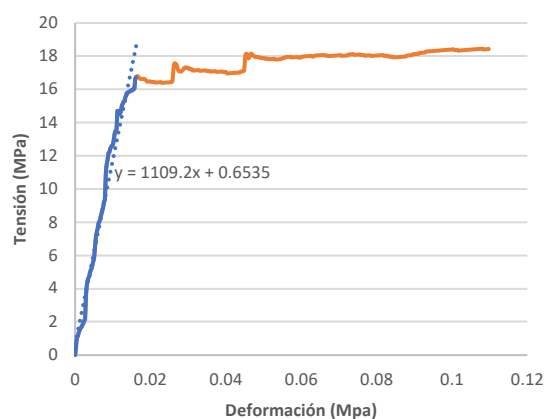


Figura 1. Curva de tensión-deformación representativa de las fibras de FS. Se diferencian la región elástica (azul), y la región plástica (naranja). La pendiente de la recta de la región elástica corresponde al módulo de elasticidad (1109,2 MPa). Se utilizó un tamaño muestral de n=6.

3.2. La viabilidad celular fue mayor en las mallas de fibras en entramado que en las mallas de fibras en paralelo

Previo a los estudios sobre la funcionalidad de las mallas biohíbridas, se estudió su biocompatibilidad mediante el ensayo de viabilidad celular XTT. En la gráfica (figura 2) se representa, mediante barras, el porcentaje de viabilidad celular obtenido en cada tipo de malla biohíbrida y en el *Transwell* (Control). Se observó una disminución estadísticamente significativa en la viabilidad celular de las mallas de fibras en paralelo con respecto al control. Por el contrario, en las mallas de fibras en entramado no se observaron diferencias estadísticamente significativas en la viabilidad celular con respecto al control.

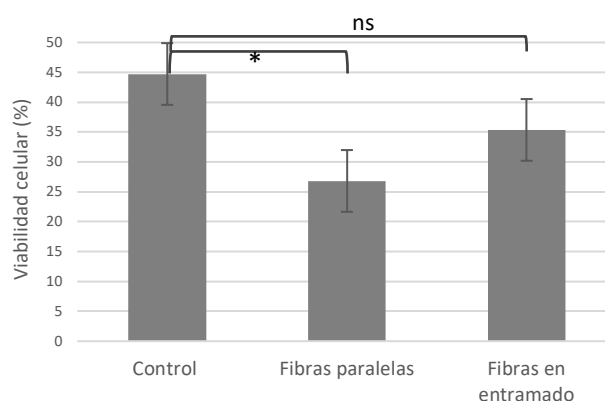


Figura 2. Viabilidad de las ARPE-19 sembradas en el *Transwell*, las mallas de fibras en paralelo y en entramado. Se observó una disminución significativa de la viabilidad celular en las mallas de fibras en paralelo con respecto al control. En las mallas de fibras en entramado no se observaron diferencias significativas en la viabilidad celular con respecto al control. (* $p < 0,001$; ANOVA de una vía con post-análisis de Tukey; $n = 6$)

3.3. Los biohíbridos presentaron evidencias de la formación de barrera hematorretiniana

La capacidad de las células del EPR de actuar como barrera hematorretiniana se debe principalmente a las uniones estrechas que estas forman entre sí, mediante la expresión de ocludinas, principalmente. Por ello, se estudió la formación de uniones estrechas en las mallas biohíbridas. Como se observa en la figura 3, se vio la expresión de ZO-1 entre las células de los tres grupos, que indicó la formación de uniones estrechas. Se observó un marcaje más difuso en algunas células de las mallas de los dos tipos con respecto al control. Por otro lado, tanto en el grupo control como en las mallas biohíbridas de fibras en entramado se observaron evidencias de la formación de un monocapa celular, no siendo así en las mallas de fibras en paralelo. En el grupo control, las células mostraron su morfología cuboidal típica en mosaico, que se observó también en las células de las mallas de fibras en entramado. En estas últimas, algunas células mostraron una morfología triangular, la cual también se observó en las mallas de fibras en paralelo.

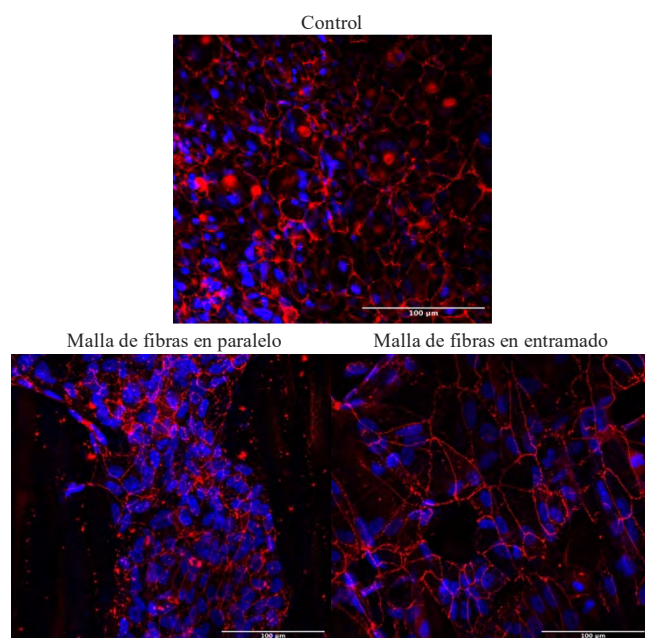


Figura 3. Expresión de ZO-1 en las células ARPE-19. En azul, núcleos marcados con Hoechts y en rojo, el marcaje de ZO1. Las células de todos los grupos expresaron de forma positiva ZO-1 indicando la formación de uniones estrechas entre células. Las imágenes fueron tomadas con el microscopio Laser Confocal Espectral para las mallas. Se utilizó un tamaño muestral de $n=3$ por cada grupo. Barras de escala: 100 μm

4. Conclusiones

- La técnica de hilado D3S permite obtener fibras de altas prestaciones mecánicas.
- Las mallas de FS permiten una mayor viabilidad de células del EPR cuando las fibras de FS se disponen en entramado que en paralelo.
- Las células del EPR forman uniones estrechas en ambos biohíbridos.

- El formato de mallas más idóneo para crear una membrana de Bruch es el de fibras en entramado.

Referencias

- [1] Mitchell, P.; Liew, G.; Gopinath, B.; Wong, T.Y. Age-Related Macular Degeneration. *Lancet* **2018**, *392*, 1147–1159, doi:10.1016/S0140-6736(18)31550-2.
- [2] Hurley, J.B. Retina Metabolism and Metabolism in the Pigmented Epithelium: A Busy Intersection. *Annu. Rev. Vis. Sci.* **2021**, *7*, 665–692, doi:10.1146/annurev-vision-100419-115156.
- [3] Galloway, C.A.; Dalvi, S.; Shadforth, A.M.A.; Suzuki, S.; Wilson, M.; Kuai, D.; Hashim, A.; Macdonald, L.A.; Gamm, D.M.; Harkin, D.G.; et al. Characterization of Human iPSC-RPE on a Prosthetic Bruch's Membrane Manufactured from Silk Fibroin. *Investig. Ophthalmol. Vis. Sci.* **2018**, *59*, 2792–2800, doi:10.1167/iovs.17-23157.
- [4] Stahl, A. The Diagnosis and Treatment of Age-Related Macular Degeneration. *Dtsch. Arztebl. Int.* **2020**, *117*, 513–520, doi:10.3238/arztebl.2020.0513.
- [5] Tan, Y.S.E.; Shi, P.J.; Choo, C.J.; Laude, A.; Yeong, W.Y. Tissue Engineering of Retina and Bruch's Membrane: A Review of Cells, Materials and Processes. *Br. J. Ophthalmol.* **2018**, *102*, 1182–1187, doi:10.1136/bjophthalmol-2017-311390.

Agradecimientos

Este estudio fue parcialmente financiado por el Ministerio de Ciencia e Innovación (PID2020-116403RB-I00 financiado por MCIN/AEI/10.13039/501100011033) y EIC-Pathfinder de la Unión Europea. Programa en el marco del proyecto THOR (Acuerdo de subvención número 101099719). Este trabajo también es parte del programa MINA-CM (número S2022/BMD-7236), financiado mediante la convocatoria para la implementación de programas de actividades de I+D entre grupos de investigación de la Comunidad de Madrid en Biomedicina 2022 (Orden 1171/2022).

Analysis of Peroxiporin-Mediated H₂O₂ Fluxes in HeLa Cellular Membrane: A Therapeutic Potential in Redox signaling disorders

PM Soria^{1,2}, E. Arévalo-Nuñez de Arenas¹, I. Sorrentino¹, A.K. Molina-Oviedo¹, I. Medraño-Fernandez^{1,2}

¹ Department of Bioengineering, University Carlos III de Madrid, 28911, Madrid, Spain ² Corresponding authors: pau.soria.sanchez@gmail.com; imedrano@ing.uc3m.es

Abstract

Once considered merely a byproduct of oxidative respiration with deleterious effects, hydrogen peroxide (H₂O₂) has now emerged as a pivotal second messenger influencing cellular processes by modulating protein function and gene expression. A particularity of the mechanism of action of this molecule resides in its compartmentalized production within cells. To reach its targets, H₂O₂ must cross cellular membranes aided by members of the family of aquaporin (AQP) channels. In HeLa cells, it has been demonstrated that AQP8 is the unique transporter mediating H₂O₂ internalization upon an exogenous challenge, despite the presence of abundant mRNA corresponding to AQP3. In this study, we have utilized the genetically encoded H₂O₂-specific ratiometric HyPer probe stably expressed in HeLa cells cytosol (HeLa HyPerCyto) to explore this paradox by measuring and characterizing H₂O₂ permeability across both channels. Live imaging analyses revealed that AQP8 is the transporter preferentially used at the highest concentrations tested, while AQP3 appeared to facilitate H₂O₂ transport at the lower concentrations. These findings deepen into the comprehension of the complex regulation that must be deployed to control a physiological signal which can be potentially harming, providing insights into peroxiporin-mediated H₂O₂ management.

1. Introduction

ROS are intermediate chemical species formed during the reduction of molecular oxygen (O₂) to water (H₂O) in aerobic organisms as part of their utilization of this vital life-sustaining molecule. These highly reactive molecules present a dualistic nature: at low concentrations, ROS -in particular H₂O₂, the most stable species- play crucial roles in cell signaling; however, an excessive accumulation of ROS causes cellular damage, affecting DNA, proteins, and lipids. The latter phenomenon is referred to as oxidative stress, which arises from the imbalance between ROS production and the capacity of cellular antioxidant systems to scavenge them [1] and has been implicated in various diseases, including neurodegenerative disorders, aging, cardiovascular diseases, and cancer.

Physiological production of H₂O₂ predominantly occurs in enclosed intracellular compartments, mainly in mitochondria, endoplasmic reticulum (ER) and the extracellular space due to membrane-bound NADPH-

oxidases (NOX) [2]. As redox targets are situated in the cytosol, regardless the source, the second messenger must traverse a lipidic barrier using facilitated transport. AQPs are integral membrane protein channels that primarily facilitate the passive transport of water and other small, uncharged solutes across lipid bilayer membranes, following their concentration gradient. The structure of AQPs is characterized by four polypeptides forming a homotetrameric configuration. In mammals, a total of thirteen distinct isoforms of AQPs have been identified until date. Among them, the isoforms endowed with H₂O₂-transporting capacity are referred as to peroxiporins. Specifically, the isoforms AQP3, AQP8, AQP9, and AQP11 lie in this category [3]. These PXP differ in their cellular location: AQP3, AQP8, AQP9 are plasma membrane resident peroxiporins, whereas AQP11 is located at the endoplasmic reticulum (ER). Interestingly, all membrane channels have been shown to influence different pathways downstream tyrosine kinase receptors. Furthermore, the isoform AQP8 can be reversibly gated depending on cellular needs, revealing that these routes can be controlled at the level of membrane passage [4].

The existence of different transporters to internalize H₂O₂ predicts either differential usage or selective regulation of passage. Interestingly, HeLa cells internalize exogenously supplied H₂O₂ using AQP8, while transcripts for AQP3 are readily detected in this cellular model [5]. Here, we provide preliminary evidence for the former, anticipating that the control of H₂O₂ permeability is more complex than what expected.

2. Materials and methods

2.1. Culture of the stable HeLa HyPerCyto cell line

HeLa cells stably expressing the H₂O₂-specific ratiometric HyPer Probe [4] were gently provided by the lab of Roberto Sitia (IRCCS San Raffaele, Milan, Italy). Routine cultures were maintained in Dulbecco's modified Eagle's medium (DMEM, Gibco) supplemented with GlutaMAXTM-I medium (Gibco, Thermofisher), 10% fetal bovine serum (FBS Good, PAN-Biotech, Germany) and 2mM Penicillin – 10mg/ml Streptomycin (P/S, PAN-

Biotech). The antibiotic G418 (Sigma) was added to a final concentration of 0.5mg/ml for selective plasmid expression. The cells were cultured at 37°C, in a 95% humidified atmosphere and with of 5% CO₂ and split when 90% of confluence was reached. To assure that no contamination with the intracellular parasite mycoplasma developed in the cultures, cells were periodically treated with 2.5µg/ml of the preventive antibiotic plasmocinTM prophylactic (InvivoGen). All steps were performed in a sterile environment within a level 2 biosafety cabinet.

2.2. Gene downregulation

For each experiment, 20x10⁴ *HyperCyto*-expressing HeLa cells were seeded in glass coverslips within six-well plates and the expression of each AQP downregulated after 24 hours using 90pmol of the specific siRNA (for AQP3, purchased from ThermoFisher, and for AQP8 from Ambion Life Technologies) in presence of RNAiMAX lipofectamin (Invitrogen). The cells were incubated for six hours before replacing the medium by a fresh one, and the cells were left to in the incubator for 72 hours prior to H₂O₂ permeability analyses.

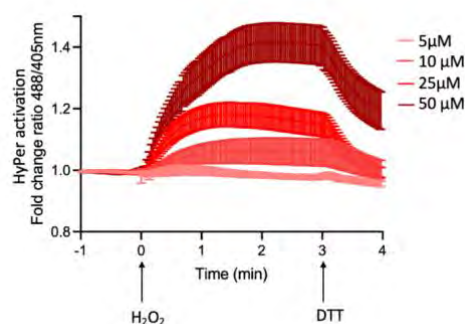
2.3. Functional-microscope assay

Cells on coverslips were incubated for 10 minutes at room temperature in Ringer buffer (RB) which contained 140 mM NaCl, 2mM CaCl₂, 1mM MgSO₄, 1.5mM K₂HPO₄, 10mM Glucose and 10mM HEPES at a pH of 7.4. RB was disposed and repared fresh after two weeks of use. Preceding the incubation of the cells with RB, H₂O₂ 30% (Sigma) was diluted in RB at different concentrations (5µM, 10µM, 25µM and 50µM). Besides, an aliquot of 5mM of the reducing agent dithiothreitol DTT (Sigma) used as probe reduction control was also prepared in RB. Finally, the glass coverslips containing the cells were placed on an Attofluor TM cell chamber (Thermo Fisher) for live-cell microscopy Live images were collected using intervals of excitation of 2 seconds with two led lights at 488 and 405nm, for a lapse of five minutes using a Leica Dmi8 microscope equipped with HC PL FLUOTAR L 20X/0.40 DRY and Hamamatsu C11440-42U Camera, and mean fluorescence intensity recorded for both channel at 526nm emission. After the first minute of acquisition, H₂O₂ was added at the indicated concentrations for each experiment. At the fourth minute, DTT was added to reduce the probe and provide evidence for oxidation. The ratio of emission of the probe (488/405 nm) of at least 25 cells was calculated using ImageJ software and averaged. Depending on the experiment conducted, results were depicted as either basal ratio or fold change of the ratio against time, in each case accompanied by the standard error of the mean (SEM). In addition, the rate of probe activation by H₂O₂ in non-silenced and silenced cells was quantified as a percentage, using the response to a challenge of 50 µM of H₂O₂ of non-silenced cells as reference value.

3. Results and discussion

3.1. H₂O₂ plasma permeability in HeLa cells

To further understand how H₂O₂ permeates the plasma membrane and the contribution in each case of the different membrane transporters expressed in the HeLa cell line, H₂O₂ transport was titrated using the genetically encoded H₂O₂-specific *HyPer* probe [6] using live imaging. This probe is a ratiometric sensor, as its derived fluorescence is the product between two fluorescence emissions, one excited at 488nm and the second at 405nm, and both collected at 526nm [6]. Peroxiporin-mediated H₂O₂ internalization has been traditionally investigated using exogenous challenges that assure high responses of the sensors used (typically, 50-100µM [2],[4]). However, living cells probably experience fluxes of much lower intensity, as these concentrations induce terminal oxidative modifications incompatible with life in a substantial protein pool (Medraño-Fernández, unpublished results). Thus, when designing a systematic analysis of cell permeability in which the capacity of different transporters is assayed, is mandatory to also include lower concentrations. Following these lines, the response of HeLa cells stably expressing the probe in their cytosol (HeLa *HyPerCyto*) to controlled boluses of either 5µM, 10µM, 25µM or 50µM was recorded. Exposing HeLa *HyPerCyto* cells to those, generated a change in the ratio of fluorescence of increasing intensity that paralleled the increase in H₂O₂ concentration used to stimulate the cells. The effect was quantified across experiments by representing the mean fold change of the



488/405nm ratio against time \pm standard error of the mean (SEM), producing a dose-response curve for each concentration as shown in Figure 1.

Figure 1. Time-course of *HyPer* activation using scaled concentrations of H₂O₂. Fold change ratio of activation of the probe is plotted against time \pm SEM.

In these kinetic representations no changes were recorded during the first minute of acquisition to allow for sample equilibration, and at time 0, the exogenous oxidant was added causing an increase in the ratio as the H₂O₂ molecules irrupt in the cytosol; H₂O₂ continues activating the probe (so increasing the ratio) until an equilibrium is reached caused either by transporter saturation, action of the antioxidant defense or a combination of both. At the end, an excess of 5mM dithiothreitol (DTT) was added to the cells leading to a reduction in the ratio as corresponding to the addition of a reducing agent,

confirming that the shift in ratio observed is indeed an oxidation of the probe.

The results obtained regarding the internalization of 50 μM H_2O_2 in HeLa *HyPerCyto* cells were consistent with the previous investigations conducted which served as the foundation for this study [2]. In accordance with the stimulus applied, the level of activation of the probe was diminished while reducing the H_2O_2 bolus concentration. Indeed, at the lowest concentration used, probe activation only slightly got over the baseline, indicating that the concentration used is vicinal to the lower resolution of the probe. If we consider that permeability across cellular membranes has been estimated to be around 300:1 (from 300 molecules outside, 1 would get in) [7], this means that in HeLa cells any process producing H_2O_2 would be only detectable with *HyPer* when generating at least 16nM of cytosolic H_2O_2 molecules. However, at all concentrations, the accelerated kinetics at time 0, can only be explained by the presence of a transporter, indicating that the selected challenges are suited to serve as bases to further analyze peroxiporin-mediated H_2O_2 fluxes: once obtained the steady state responses to each concentration we can interfere channel expression to determine their specific permeability ranges if existing.

3.2. Characterization of the differential roles of AQP3 and AQP8 on guiding H_2O_2 fluxes

To investigate whether the lack of sensitivity of AQP3 to a 50 μM challenge in HeLa cells is due to differences in the capacity of transport of this channel respect to AQP8, we selectively silenced the expression of both channels either individually or in combination using specific siRNAs. Single downregulation of each transporter will uncover the specific H_2O_2 transporting capabilities of each isoform when cells are exposed to different boluses of H_2O_2 , while the double silencing of both will reveal additive contributions. The quantification of H_2O_2 uptake was measured after the addition of exogenous H_2O_2 and the results were represented as percentage (%) of H_2O_2 transport time for clarity. As reference value due to its well-established levels of response [6], we used a 50 μM challenge of H_2O_2 in non-silenced cells (labelled as Ctrl in Figures 2 to 5).

When assaying 50 μM of H_2O_2 uptake (Figure 2), we replicated former results [2] and the capacity of internalization of the H_2O_2 relied solely in AQP8 (si8), as expected. Indeed, the cells in which the expression of AQP8 was downregulated experienced a complete inhibition of H_2O_2 transport, while AQP3-silenced cells (si3) presented a probe activation profile that returned to control values. In accordance, H_2O_2 internalization in cells incubated simultaneously with both siRNAs (si8/si3) was also abrogated, an effect that can be attributed to accurate silencing of the AQP8 isoform.

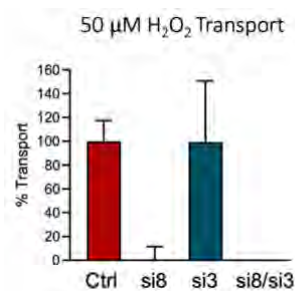


Figure 2. Quantification of exogenous 50 μM H_2O_2 transport. Results are represented as percentage of H_2O_2 transport respect to non-silenced cells \pm SEM. Ctrl, non-silenced cells; si8, AQP8-silenced cells; si3, AQP3-silenced cells; si8/si3, cells silenced both for AQP8 and AQP3.

As for 25 μM of H_2O_2 (Figure 3), inhibition of the activity of AQP8 due to intervention of its expression (si8) abrogated H_2O_2 internalization even though AQP3 transporters were present. Conversely, when silencing the expression of AQP3 (si3), full activation of the probe to control values was achieved. The double silencing of both channels (si8/si3) presented similar results respect to the single silencing of AQP8. Altogether, the data obtained suggested that, at 25 μM of H_2O_2 , AQP8 was again responsible for H_2O_2 internalization and that AQP3 was not relevant. It must be noted that reproducibility across replicates is less efficient than when assaying the 50 μM boluses, a fact that could reflect that when a reduced gradient across membranes is applied, the mechanisms of internalization are not so effectively measured with the probe used. Thus, to further support these preliminary conclusions on channel usage, more experiments would be needed.

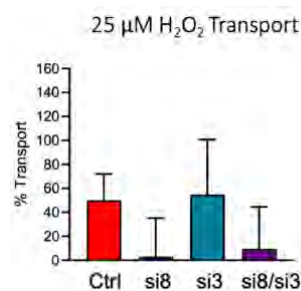


Figure 3. Quantification of exogenous 25 μM H_2O_2 uptake. Results are represented as in Figure 2. Ctrl, non-silenced cells; si8, AQP8-silenced cells; si3, AQP3-silenced cells; si8/si3, cells silenced both for AQP8 and AQP3.

By lowering the H_2O_2 concentration by challenging cells with 10 μM (Figure 4), the situation is reversed: AQP8-silenced cells (si8) exhibited similar probe activation values when compared to non-silenced cells (Ctrl), probably due to the presence of active AQP3 channels, while when AQP3 expression was silenced (si3), H_2O_2 internalization is inhibited. This observation suggested that at the concentration used, AQP8 function might not

be relevant while AQP3 takes its place as necessary transporter.

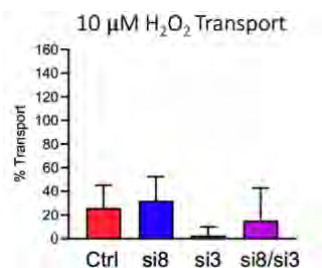


Figure 4. Quantification of exogenous 10μM H₂O₂ internalization. Results are represented as in Figure 2. Ctrl, non-silenced cells; si8, AQP8-silenced cells; si3, AQP3-silenced cells; si8/si3, cells silenced both for AQP8 and AQP3.

Similarly, even if when the H₂O₂ input was reduced to 5μM (Figure 5), transport was greatly reduced, reflecting a poor gradient for internalization, only when AQP3 is expressed (si8), H₂O₂ transport is active. Thus, the most probable scenario at this concentration levels is the one in which AQP3 takes the role of leading transporting channel, while AQP8 is dispensable.

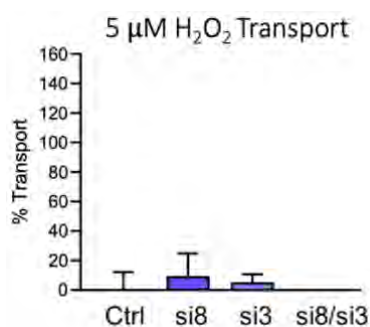


Figure 5. Quantification of exogenous 5μM H₂O₂ transport. Results are represented as percentage of H₂O₂ transport respect to non-silenced cells using as reference value the response of non-silenced cells to 50μM H₂O₂ ± SEM. Ctrl, non-silenced cells; si8, AQP8-silenced cells; si3, AQP3-silenced cells; si8/si3, cells silenced both for AQP8 and AQP3.

Taken together, all the findings presented indicate that at higher concentrations of H₂O₂ (50μM and 25μM), the main responsible for permeation is AQP8 (Figures 2 and 3) since its absence has a noticeable impact on the efficiency of the transport. On the other hand, at lower concentrations (Figures 4 and 5), AQP3 is suggested to be primarily accountable for the transport of H₂O₂ across the plasma membrane.

4. Conclusions

AQP8 was identified as the primary transporter for 50 μM and 25 μM concentrations of H₂O₂, while AQP3 exhibited a significant contribution to H₂O₂ transport at concentrations of 5μM and 10μM. These results suggest that the permeability ranges when two peroxiporins are present at the same time in the same cellular type is

subjected to complex mechanisms of regulation, as concentration distinction would only be possible by means of gating or selective plasma membrane expression of each isoform. Therefore, the data presented provide insights into distinctive transport capacities of AQP3 and AQP8 within the context of HeLa cells, shedding light on the significance of these isoforms in mediating H₂O₂ fluxes but interestingly, also open new research questions. Due to the implications of both isoforms in diverse cellular processes, further investigations will be needed to deeply understand the detailed peroxiporin activity of AQP3 and AQP8, the control of their specific permeability ranges, and the broader implications that their regulation would imply. Therefore, this report lays the foundation for future investigations that could pave the way for novel therapeutic avenues to tune signaling in redox signaling-dependent diseases.

Acknowledgements

This project has been developed in the UC3M facility core CleanRooms for Bioengineering. Financial support was provided by the Madrid Government (Comunidad de Madrid) under the Multiannual Agreement with UC3M in the line of "Research Funds for Beatriz Galindo Fellowships" (REDOXSKIN-CM-UC3M), and in the context of the V PRICIT (Regional Programme of Research and Technological Innovation), and from "Proyectos de I + D + I" (PID2020-114230 GA-I00) funded by MCIN/AEI/10.13039/501100011033/.

Referencias

- [1] H. Kimura, T. Sawada, S. Oshima, K. Kozawa, T. Ishioka, and M. Kato, "Toxicity and Roles of Reactive Oxygen Species," *Curr. Drug Targets - Inflamm. Allergy*, vol. 4, no. 4, pp. 489–495.
- [2] M. Bertolotti *et al.*, "Tyrosine kinase signal modulation: a matter of H₂O₂ membrane permeability?," *Antioxid. Redox Signal.*, vol. 19, no. 13, pp. 1447–1451, Nov. 2013, doi: 10.1089/ars.2013.5330.
- [3] D. Nordzike and I. Medraño Fernández, "The Plasma Membrane: A Platform for Intra- and Intercellular Redox Signaling," *Antioxidants*, vol. 7, p. 168, Nov. 2018, doi: 10.3390/antiox7110168.
- [4] I. Medraño-Fernandez *et al.*, "Stress Regulates Aquaporin-8 Permeability to Impact Cell Growth and Survival," *Antioxid. Redox Signal.*, vol. 24, no. 18, pp. 1031–1044, Jun. 2016, doi: 10.1089/ars.2016.6636.
- [5] G. Pellavio *et al.*, "Regulation of Aquaporin Functional Properties Mediated by the Antioxidant Effects of Natural Compounds," *Int. J. Mol. Sci.*, vol. 18, no. 12, p. 2665, Dec. 2017, doi: 10.3390/ijms18122665.
- [6] V. V. Belousov *et al.*, "Genetically encoded fluorescent indicator for intracellular hydrogen peroxide," *Nat. Methods*, vol. 3, no. 4, pp. 281–286, Apr. 2006, doi: 10.1038/nmeth866.
- [7] M. Carmona *et al.*, "Monitoring cytosolic H₂O₂ fluctuations arising from altered plasma membrane gradients or from mitochondrial activity," *Nat. Commun.*, vol. 10, no. 1, Art. no. 1, Oct. 2019, doi: 10.1038/s41467-019-12475-0.

Predicción de complicaciones postoperatorias a partir de la actividad neuronal de pacientes pediátricos sometidos a anestesia mediante modelos de *Machine Learning*

L. Gutiérrez-de Pablo¹, J. Gomez-Pilar^{1,2}, E. Gómez-Pesquera³, Eduardo Tamayo³, R. Hornero^{1,2}

¹ Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, lauragudepablo2@gmail.com.

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España.

³ Servicio de Anestesiología, Hospital Clínico Universitario de Valladolid, Valladolid, España.

Resumen

Dado que el cerebro es el órgano diana para muchos anestésicos, una monitorización precisa de la profundidad anestésica es esencial para prevenir complicaciones postoperatorias, como los Trastornos del Comportamiento (TCPO). En este trabajo se analizaron señales del electroencefalograma (EEG) de 115 pacientes pediátricos adquiridos y monitorizados mediante el sistema SedLine®. Se calcularon características espectrales y no lineales del EEG en tres estados: vigilia preoperatoria, anestesia y cirugía. Se encontraron diferencias significativas entre los estados de vigilia y anestesia, mostrando una actividad neuronal más rápida y compleja durante la anestesia en estos pacientes. Posteriormente, se realizó un análisis de asociación entre las variables electrofisiológicas, clínicas y sociodemográficas. Se estableció una relación entre variables clínicas y sociodemográficas y los TCPO, además de identificarse la importancia de la potencia relativa en Delta-1 en la predicción de TCPO a corto plazo. Por último, se desarrollaron modelos de Machine Learning (ML) basados en regresión logística y análisis de discriminante lineal, que incorporaron variables del EEG, clínicas y sociodemográficas, logrando una precisión del 92% y del 82% en la predicción de TCPO a los 7 y 28 días, respectivamente. Estos hallazgos sugieren que la monitorización perioperatoria del EEG puede mejorar la caracterización de la profundidad anestésica en pacientes pediátricos, una población particularmente vulnerable. Además, los modelos predictivos basados en ML pueden ser herramientas valiosas para prevenir y gestionar las complicaciones postoperatorias en esta población.

1. Introducción

La hospitalización, la cirugía y la anestesia tienen un impacto psicológico substancial en pacientes pediátricos, debido a la posibilidad de manifestar un posterior cambio en el comportamiento, también llamado Trastorno del Comportamiento en el Postoperatorio (TCPO) [1]. Hasta el 60% de los pacientes pediátricos sometidos a algún tipo de cirugía ambulatoria pueden presentar TCPO tras dos semanas de la intervención [2]. Los TCPO tienen gran impacto psicológico en el paciente, como indicios de ansiedad, agresividad o irritabilidad [1]. Asimismo, estos también van ligados a un peor postoperatorio y recuperación, una mayor incidencia del dolor en la incisión quirúrgica y una disminución en la calidad del sueño. Por último, estos trastornos también van asociados a consecuencias sociales, como dificultades para relacionarse con otros niños o problemas académicos [3], [4]. Aunque hay estudios que indican posibles factores de riesgo de los TCPO como la ansiedad preoperatoria, el sexo o la edad [4], la etiología de este fenómeno es todavía desconocida. Actualmente, los TCPO se evalúan a partir del cuestionario

Post Hospitalization Behaviour Questionnaire (PHBQ) una vez que el paciente ya sufre estas alteraciones. Por tanto, predecir la aparición del TCPO sería de gran ayuda para lograr un tratamiento personalizado y minimizar las consecuencias de estas alteraciones.

Como se ha comentado, estos fenómenos ocurren en pacientes sometidos a una intervención y, por tanto, que han sido anestesiados. La anestesia general se puede definir como una depresión fármaco-inducida y reversible del Sistema Nervioso Central que resulta en hipnosis, amnesia y analgesia. Los fármacos anestésicos utilizados para la anestesia general pueden ser intravenosos, como el propofol, o inhalatorios, como el sevoflurano [5]. El sevoflurano es el anestésico inhalatorio más utilizado en pacientes pediátricos por sus múltiples ventajas, como su rápida hipnosis, su poca irritación hacia las vías superiores y su buen perfil de seguridad cardiovascular, entre otros beneficios [6]. Sin embargo, no se libra de inconvenientes, como el incremento de la incidencia de aparición de delirios y vómitos postoperatorios, así como su asociación a actividad epileptiforme en el electroencefalograma (EEG), aunque con una menor incidencia [6]. Por ello, una monitorización neurológica del paciente es necesaria para observar su estado durante la cirugía e incluso prevenir complicaciones debidas a la anestesia.

Actualmente, la técnica utilizada para la monitorización neurológica es el EEG [7]. El EEG es una de las técnicas más habituales para estudiar de forma no invasiva la dinámica electrofisiológica cerebral. Concretamente, el equipo de anestesia del Hospital Clínico Universitario de Valladolid (HCUV) utiliza el monitor SedLine® para registrar la actividad eléctrica neuronal de los pacientes durante la cirugía. Este sensor consta de seis electrodos, dispuestos según el sistema internacional 10-20, colocados sobre el cuero cabelludo en la zona frontal y frontopolar. Este monitor, además de mostrar el EEG en tiempo real, ofrece distintos parámetros como el *Patient State Index* (PSI), que ofrece información sobre el grado de consciencia del paciente [7]. Este parámetro, aunque fácil de interpretar, es habitual que no se corresponda con la clínica del paciente y menos en pacientes pediátricos.

Considerando todo lo expuesto, el objetivo de este trabajo es doble. En primer lugar, determinar medidas espectrales y no lineales del EEG que consigan caracterizar la profundidad anestésica en pacientes pediátricos y observar los cambios en la dinámica electrofisiológica del paciente anestesiado, en comparación con el paciente en vigilia

preanestésica. Por otra parte, de forma innovadora, se pretende construir modelos predictivos de los TCPO a corto y medio plazo a partir de variables del EEG, clínicas y sociodemográficas con el objetivo de crear una herramienta predictora de si los pacientes van a sufrir o no TCPO y crear de forma previa, un tratamiento y atención personalizada.

2. Materiales y métodos

2.1. Bases de datos

La base de datos utilizada consta de 115 registros de EEG tomados con el equipo SedLine® a pacientes pediátricos de entre 2 y 12 años, sin patología neuropsiquiátrica que no fueran sometidos a cirugía de urgencia. Después de un análisis inicial, se descartaron 15 registros por no cumplir con los estándares mínimos de calidad. Asimismo, no fueron incluidos pacientes cuyos padres, por barreras del idioma, no pudiesen contestar a las preguntas del cuestionario PHBQ y, también, aquellos padres que no quisieron firmar el consentimiento informado. Además de los registros y los cuestionarios, el personal sanitario del HCUV aportó datos adicionales. En primer lugar, unas marcas de tiempo útiles para la segmentación posterior de la señal en fases operatorias, así como datos clínicos y sociodemográficos de todos los pacientes.

2.2. Preprocesado

Los registros de EEG tomados con el monitor SedLine®, con frecuencia de muestreo 178.154 Hz, fueron sometidos a un preprocesado exhaustivo. Este preprocesado tuvo los siguientes pasos:

- Filtrado utilizando un filtro *Finite Impulse Response* (FIR) paso banda entre 0.1 y 70 Hz y un filtro FIR elimina banda para eliminar el ruido de red eléctrica.
- Rechazo manual de épocas, en el cual se eliminaron aquellas tramas planas o muy ruidosas.
- Rechazo automático adaptativo, en el cual se eliminaron aquellas épocas que tuvieran al menos una muestra que superara el umbral impuesto. La condición para este rechazo fue que el porcentaje total de épocas eliminadas de todos los registros no superase el 20% del total.
- Segmentación en épocas de 6,5 segundos, suficiente para analizar las bandas más lentas del EEG.

2.3. Procesado

Una vez preprocesadas las señales de EEG, se procedió a llevar a cabo la extracción de características. En primer lugar, se realizó un análisis espectral en el cual se calculó la *Power Spectral Density* (PSD) mediante el método de Welch. A partir de ésta, se calcularon diferentes parámetros como la potencia relativa (RP) en Delta-1, Delta-2, Zeta, Alfa, Beta-1, Beta-2 y Gamma, calculada como la potencia en cada una de estas bandas dividida entre la potencia global. También se calculó la frecuencia mediana (MF), que aporta información sobre la velocidad de oscilación media de la señal y es calculada como la frecuencia a la cual se encuentra el 50% de la potencia en la PSD; la asimetría espectral (Sk), medida de la forma de la PSD; y, por último, la entropía espectral (SE), que aporta información sobre la irregularidad del espectro de la señal. Posteriormente, se

realizó un análisis no lineal de las señales, en el que se calculó la entropía muestral (SampEn), como medida de la irregularidad de la señal, y la complejidad de Lempel-Ziv (LZC), estimación de la complejidad de la serie temporal.

Tras la extracción de características, se dividieron en tres fases operatorias utilizando las marcas de tiempo aportadas por el personal sanitario del HCUV. Las fases operatorias fueron: vigilia preanestésica (desde la colocación de los electrodos hasta la inducción anestésica), anestesia (desde la inducción hasta la educación de esta) y cirugía (desde el inicio hasta el fin de la intervención). Una vez obtenidos los parámetros segmentados, se realizó un análisis estadístico para estudiar las posibles diferencias entre los tres estados. Concretamente, se utilizaron el test de Friedman para comparar las tres fases operatorias de una misma variable, y el test de Wilcoxon, para comparar de dos en dos las fases de una misma variable. Posteriormente, se realizó un análisis de asociación entre las variables clínicas, sociodemográficas y electrofisiológicas. Estas asociaciones fueron más adelante representadas en redes de asociación con la herramienta Gephi y el algoritmo ForceAtlas 2, lo que proporcionó un análisis exploratorio para observar el comportamiento potencial que tendrían los modelos de *machine learning* (ML). Para la creación de estos modelos predictivos se utilizaron dos clasificadores: la regresión logística (LR) y el análisis discriminante lineal (LDA).

3. Resultados y discusión

En primer lugar, se puede observar en la Fig. 1 la distribución de las variables espectrales calculadas en las tres fases operatorias. Se aprecia como la RP en Delta-1 disminuye durante la anestesia y cirugía, contrario al resto de bandas de frecuencia, en las cuales la RP aumenta, con excepción de la RP en Delta-2, en la cual no hubo diferencias en los tres estados operatorios. Aunque no existen estudios previos para contrastar estos resultados, esta tendencia es contraria a lo estudiado hasta ahora en pacientes adultos, lo que podría ser debido a la neurofisiología cerebral en la población pediátrica relacionada con procesos de maduración cerebral. Como indicio de este desarrollo neurológico, algunos estudios sugieren un aumento en la RP en Zeta [8]. En la RP en Zeta también se observan diferencias entre anestesia y cirugía. Esto es relevante ya que, durante la cirugía, el paciente es sometido a estrés y dolor, algo que no ocurre antes de la intervención. Por tanto, los resultados sugieren que la potencia en la banda Zeta podría ser una característica relacionada con el estrés y/o dolor. Por otra parte, otros estudios sugieren un aumento en la RP en Alfa como biomarcador de la edad: a mayor pico Alfa, menor edad [9]. Asimismo, los aumentos en las RP de Beta-1 y Beta-2 podrían ser causados por la presencia de oscilaciones rápidas en el EEG debido a la inducción con sevoflurano o, también, por estimulación dolorosa [7]. Finalmente, el aumento de la RP en Gamma podría darse por las mismas razones que en las bandas de frecuencia Beta-1 y Beta-2.

Por otra parte, este mismo aumento de las potencias de frecuencias más altas durante la anestesia y cirugía puede constatare gracias a la MF. Estudios previos observaron este efecto en pacientes anestesiados con sevoflurano

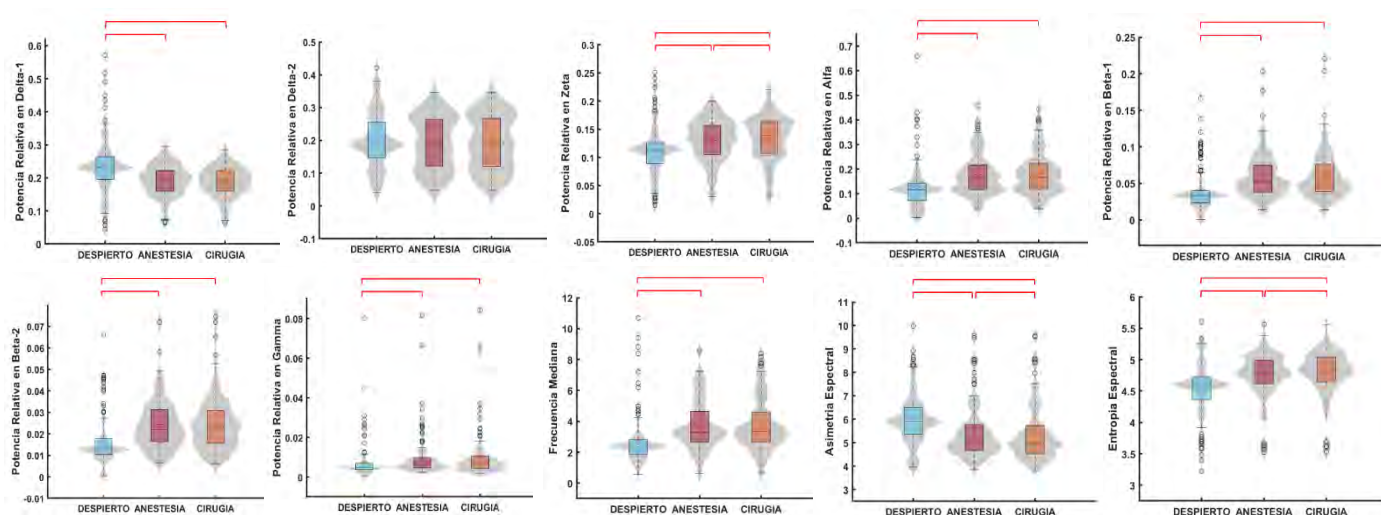


Figura 1. Variables espectrales extraídas del EEG divididas en diferentes fases operatorias.

debido a la actividad epileptiforme en el EEG o por exposición a movimientos [10], lo que podría ser el caso durante intervenciones quirúrgicas. Se puede observar una disminución en la Sk durante la anestesia y cirugía, lo que indicaría que la PSD se ha desplazado hacia frecuencias más elevadas. Por último, la SE, estudiada en algunos artículos como indicador del efecto hipnótico del sevoflurano [11], aumenta en los estados de anestesia y cirugía, lo que indica una mayor irregularidad en la PSD. De nuevo, en la Sk y la SE el EEG ha sido capaz de observar diferencias entre los estados de anestesia y cirugía.

En segundo lugar, podemos observar en la Fig. 2 la distribución de las variables no lineales calculadas en las tres fases operatorias. Se puede observar un aumento en la SampEn durante la cirugía y la anestesia, en comparación con vigilia. Algunos autores sugieren la SampEn como predictor de la profundidad anestésica en adultos debido a su relación lineal con parámetros como el PSI o el *Bispectral Index* (BIS) [7], [12]. Asimismo, se puede observar un aumento en la LZC durante la anestesia y la cirugía, lo que indicaría un incremento en la complejidad de las señales electrofisiológicas en esos períodos.

Como método exploratorio, se realizó un análisis de asociación entre todas las variables. Las redes de asociación resultantes pueden observarse en la Fig. 3. En el análisis de las variables en vigilia, se identifican dos módulos diferenciados: en azul las variables del EEG y en naranja y rojo las variables clínicas y sociodemográficas, destacando las variables objetivo en rojo. Aunque las conexiones con las variables objetivo provienen principalmente de variables dentro del mismo módulo, es relevante la relación entre la RP en Delta-1 y los TCPO a los 7 días, sugiriendo la importancia del estado preoperatorio en las consecuencias postoperatorias. Asimismo, se presentan las redes de asociación para los estados de anestesia y cirugía, donde se mantienen los módulos previos. En este caso, las variables del EEG no se relacionan directamente con las variables objetivo, pero sí con otras variables clínicas y sociodemográficas.

Finalmente, se crearon modelos de ML para predecir TCPO a corto y medio plazo. Inicialmente, se dividieron al 50% los datos en conjuntos de entrenamiento y test, pareados por

sexo y edad. Más adelante, se organizó la información para cada prueba, variando las características de entrada en los modelos. En primer lugar, se utilizaron solo variables derivadas del EEG. En segundo lugar, se incorporaron variables clínicas y sociodemográficas. Por último, se emplearon variables seleccionadas mediante el algoritmo *Fast Correlation-Based Filter* (FCBF). Para los modelos de predicción de TCPO a los 7 días se seleccionaron la agitación a los 15 y 30 minutos, la edad y la Sk en vigilia. En cuanto a los modelos de predicción de TCPO a los 28 días, se seleccionaron la agitación a los 30 minutos, el tipo de cirugía, la RP en Delta-1 durante la cirugía y la RP en Zeta en vigilia.

La eficacia de cada modelo fue evaluada con nueve parámetros: la sensibilidad, la especificidad, el valor predictivo positivo y negativo, la razón de verosimilitud positiva y negativa, la precisión, el coeficiente kappa de Cohen y el *Area Under the Receiving Operating Curve* (AUROC). Comparando los modelos de predicción de TCPO a los 7 días, se observa que los resultados con las variables de EEG tienen baja precisión en ambos casos en comparación con otras variables. Las variables clínicas y sociodemográficas dan buenos resultados en ambos modelos, como era de esperar según las redes de asociación. Destaca el modelo de LR con las variables seleccionadas, con una precisión del 92% y una AUROC de 0.968, lo que haría de este un modelo muy fiable para predecir los TCPO a corto plazo. Entre los modelos de predicción de TCPO a los 28 días, nuevamente sobresale el modelo de LR con las variables seleccionadas, en el que se obtiene una precisión del 82% y una AUROC de 0.735. Estos resultados

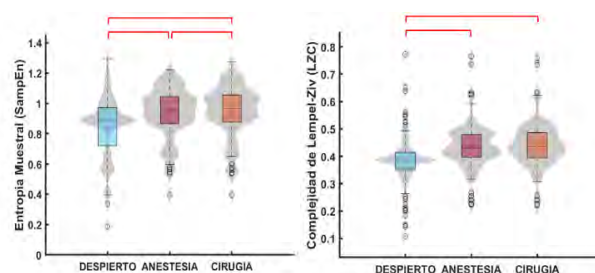


Figura 2. Variables no lineales extraídas del EEG divididas en diferentes fases operatorias.

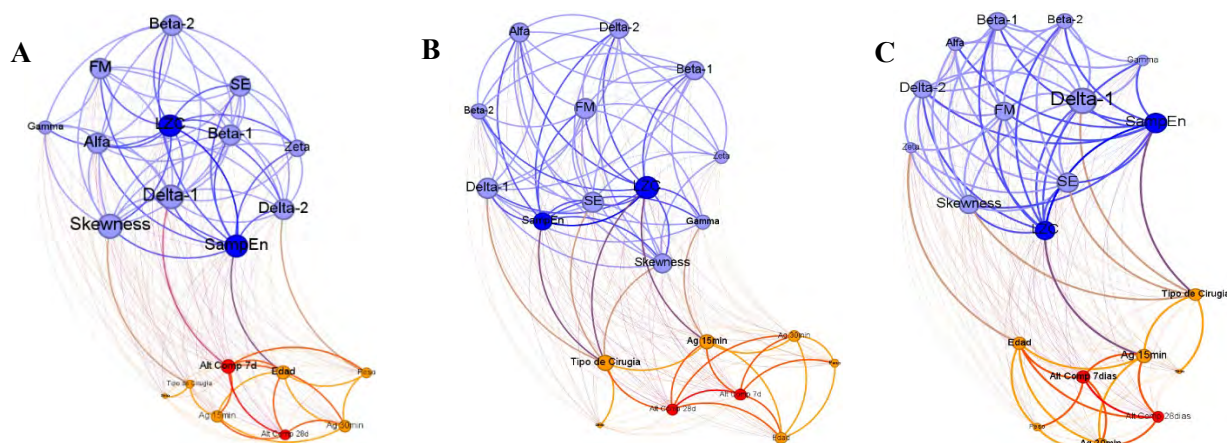


Figura 3. Redes de asociación entre las variables electrofisiológicas, clínicas y sociodemográficas: (A) vigilia, (B) anestesia y (C) cirugía.

indicarían que los parámetros extraídos del EEG aportarían información complementaria a las características clínicas y sociodemográficas, ayudando a construir modelos más robustos.

3.1. Limitaciones y líneas futuras

En primer lugar, el tamaño de la base de datos es limitada para su utilización para la construcción de modelos de ML. Además, debido a la prevalencia clínica del TCPO, existe desbalanceo entre el grupo de pacientes con y sin TCPO a los 28 días. En futuros trabajos, se podría aumentar la muestra y extender los registros de vigilia preanestésica, dado a que mostraron ser relevantes para predecir complicaciones en el postoperatorio. Además, se podrían incluir variables relevantes como la evolución de la concentración de sevoflurano o del PSI durante la cirugía, duración de la operación, ayudando a proporcionar modelos predictivos más robustos. Como futuras líneas de investigación, se pueden considerar la exploración de variables electrofisiológicas adicionales para evaluar la nocicepción del paciente y analizar el EEG durante el sueño en pacientes pediátricos hospitalizados después de la cirugía, comparándolo con los registros obtenidos durante la anestesia.

4. Conclusiones

Las variables extraídas del EEG han demostrado diferencias estadísticamente significativas entre los estados de vigilia, anestesia y cirugía. Además, se ha observado una actividad neuronal con mayor velocidad y complejidad durante la anestesia con sevoflurano en pacientes pediátricos. Asimismo, el análisis de redes de asociación reveló conexiones directas entre la mayoría de las variables clínicas y sociodemográficas con las variables objetivo. Específicamente, en el estado de vigilia, se destaca la relación directa entre la RP en Delta-1 y los TCPO a los 7 días, subrayando la importancia del estado preoperatorio en las complicaciones posteriores. La singularidad de este estudio radica en su enfoque innovador al emplear modelos de ML para predecir los TCPO. Los resultados muestran como las variables seleccionadas, que incluyen clínicas, sociodemográficas y electrofisiológicas, mejoran la capacidad predictiva de los modelos, destacando la importancia de las variables extraídas del EEG en la predicción de TCPO.

Agradecimientos

Este estudio ha sido financiado por el departamento de Bioingeniería de la Universidad Carlos III de Madrid y la Gerencia Regional de Salud de Castilla y León (GRS 2636/A/22).

Referencias

- [1] Stargatt, R., Davidson, A. J., Huang, G. H., Czarnecki, C., Gibson, M. A., Stewart, S. A., & Jansen, K. (2006). A cohort study of the incidence and risk factors for negative behavior changes in children after general anesthesia. *Pediatric Anesthesia*
- [2] Yuki, K., & Daaboul, D. G. (2011). Postoperative maladaptive behavioral changes in children. *Middle East Journal of Anaesthesiology*.
- [3] Caldas, J. C. S., Pais-Ribeiro, J. L., & Carneiro, S. R. (2004). General anesthesia, surgery and hospitalization in children and their effects upon cognitive, academic, emotional and sociobehavioral development - a review. *Pediatric Anesthesia*.
- [4] Tripi, P. A., Palermo, T. M., Thomas, S., Goldfinger, M. M., & Florentino-Pineda, I. (2004). Assessment of risk factors for emergence distress and postoperative behavioural changes in children following general anesthesia. *Pediatric Anesthesia*.
- [5] Barash, P. G., Cullen, B. F., Stoelting, B. F., Cahalan, R. K., Stock, M. C., Ortega, R., Sharar, S. R., & Holt, N. F. (2018). *Anestesia clínica* (8th ed.). Wolters Kluwer.
- [6] Jöhr, M., & Berger, T. M. (2005). Paediatric anaesthesia and inhalation agents. *Best Practice & Research Clinical Anaesthesiology*.
- [7] Rigouzzo, A., Khoy-Ear, L., Laude, D., Louvet, N., Moutard, M.-L., Sabourdin, N., & Constant, I. (2019). EEG profiles during general anesthesia in children: A comparative study between sevoflurane and propofol. *Pediatric Anesthesia*.
- [8] Cornelissen, L., Kim, S.-E., Purdon, P. L., Brown, E. N., & Berde, C. B. (2015). Age-dependent electroencephalogram (EEG) patterns during sevoflurane general anesthesia in infants.
- [9] Kh, L., Yh, K., Yj, S., & Mk, O. (2015). The Patient State Index is well balanced for propofol sedation. *Hippokratia*.
- [10] Jaaskelainen, S. K., Kaisti, K., Suni, L., Hinkka, S., & Scheinin, H. (2003). Sevoflurane is epileptogenic in healthy subjects at surgical levels of anesthesia.
- [11] Escontrela Rodríguez, B., Gago Martínez, A., Merino Julián, I., & Martínez Ruiz, A. (2016). La entropía espectral en la monitorización de la profundidad de la anestesia. *Revista Española de Anestesiología y Reanimación*.
- [12] Liu, Q., Ma, L., Fan, S.-Z., Abbod, M. F., & Shieh, J.-S. (2018). Sample entropy analysis for the estimating depth of anaesthesia through human EEG signal at different levels of unconsciousness during surgeries.

The importance of integral time length windows for the classification of activities of daily living based on Machine Learning techniques

A. Ruiz¹, E. Carbone², B. Larraga¹, E. Rocon³, A. Gutiérrez¹

¹ ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain.

² Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Turin, Italy

³ Centre for Automation and Robotics, CSIC-UPM, 28500 Madrid, Spain

Abstract

Pathological tremor, a prevalent movement disorder seen in essential tremor (ET) and Parkinson's disease (PD) patients, is the most common tremor disorder impacting the quality of life of those who suffer from it. This study proposes a method to classify daily life activities using a single wrist-worn IMU for tremor patients. The used dataset involves IMU recordings from the dominant arm during 11 tasks performed by ET and PD patients. Signal features were extracted from different sized windows and used to train Random Forest (RF) and Support Vector Machine (SVM) models, training 10 different models overall. Results shows that although larger window sizes, particularly the 10 seconds window, provided highest average F1-score, certain specific activities were better classified with shorter windows. This approach outperforms prior studies by achieving improved classification outcomes and opens a new line in continuous tremor monitoring. Future research could explore the combination of various window lengths to identify optimal window durations for further accuracy refinement.

1. Introduction

Pathological tremor is the most common movement disorder, characterized by involuntary and rhythmic oscillations of a part of the body, affecting mostly the hands [1]. It is commonly seen in general medical practice; essential tremor (ET) affects approximately 4% of the population above 65 years old [2], and Parkinson's disease (PD) has become the second most common neurodegenerative disorder after Alzheimer's disease [3].

As a direct cause of these symptoms, the quality of life of the patients affected by tremor is significantly impacted, gradually undermining their ability to perform their activities of their daily living (ADLs) from early stages and throughout the course of the disease [4].

The most effective treatment of tremor is medication, although drugs are usually prescribed on a trial-and-error basis, which coupled with a wide range of side-effects, represent the major drawback of the actual therapeutic strategies [4]. Motor symptoms caused by pathological tremor are typically assessed based on the mechanical demonstration of tremor and quantified using movement disorder clinical scales such as the UPDRSIII (*Unified Parkinson's Disease Rating Scale*) and the Fahn-Tolosa-

Marin scale [5]. However, these evaluation methods are hampered by the bias in the performance of the patients, which may be caused by placebo effects or the “white coat syndrome”. In this case patients tend to apply an extra effort due to the presence of a clinician, resulting in a biased reflection of their motor ability. Additionally, this syndrome may also constrain the evaluation of the medication's effect on the patient, which is essential to study the evolution of the symptoms with medication dosages [4,6].

Nowadays, the combination of wearable sensing technology and data mining algorithms to recognize movement disorders has shown an increasing potential, as they can be used to quantify motor symptoms. In subjects with no tremor, ADLs' classification with inertial measurement units (IMUs) has grown in the past years, approaching recognition of postural movement and activities related to motion [6].

However, the application of these methods to patients, which increases the difficulty of the task due to the tremorous component of their movements, has not yielded conclusive results yet. Previous studies have focused on detecting and classifying ADLs in patients of PD and ET using several IMUs [6-9]. Particularly, one of these studies focused on classifying a series of fine and gross movement activities in patients using four IMUs along the arm, with the objective of moving towards an every-day life application that could allow a continuous monitoring [4].

In this paper, a new method which provides an improvement in the classification of activities is proposed. This method uses only one IMU on the wrist and aims to improve the classification performed by a previous study [4], using the same dataset and reducing the number of IMUs used from four to one. Increasing the precision of the classification of activities in PD and ET patients would contribute to the objective of continuous monitoring of tremorous movement and the assessment of the medication's impact throughout the evolution of the disease.

2. Methodology

A dataset which contains records from 4 IMUs, placed over the dominant arm is used. This database contains ET and PD patients' data carrying out different tasks according to a

specific protocol [4]. Nonetheless, this work focuses on the IMU placed at the third distal of the forearm, analyzing closest kinematic movements to the wrist. The recorded and analyzed activities were:

- Combing hair (CB)
- Buttoning the buttons of a lab coat (BB)
- Cutting a fake steak (CE)
- Eating the previously cut pieces with a fork (EF)
- Simulate drinking (SD)
- Opening and closing a tupperware container (OT)
- Turning 3 pages in a book/magazine (TB)
- Printing their name/signing a document (SN)
- Simulate tooth brushing (TB)
- Turning doorknob (TD)
- Resting arms on table (RE)

These tasks encompass both fine and proximal movements, portraying two levels of precision.

2.1. Data characteristics

The dataset is composed of acceleration and angular velocity in all three axes (x, y, z) from 16 patients, whose gender and age were not required. Each participant carried out the 11 aforementioned tasks, repeating each of them between three and six times, except for the “RE” task, which was only performed once. The tremorous and voluntary movements were separated by an adaptative algorithm based on frequency separation, and the signals were resampled at 1kHz [4].

2.2 Preprocessing and filtering

As a first step, the data was exported to dataframes to visualize and analyze the IMU signals.

First of all, a data screening was implemented. The signals were then analyzed empirically, considering the time distribution of each task. The signals that were farthest from the distribution center were discarded, as they could potentially represent erroneous recordings and lead to misclassification.

2.3 Feature extraction

Considering that the duration of the signals was different and variable for different tasks and patients, the signals were divided in different sized overlapping windows. The selected windows were (see Figure 1):

- 2.5 seconds with 1 second of overlap
- 5 seconds with 2 second of overlap
- 10 second with 4 second of overlap
- 15 second with 6 second of overlap
- 20 second with 8 second of overlap

Once the signals were segmented, the following features were extracted from each window [4,10]: mean, standard deviation, median, maximum, minimum, difference between first and last value of window, variance, and RMS.

2.4 Classification approach

The features for each window were extracted and collected in five different dataframes corresponding to the different windows sizes. The dataset was divided into test and train groups with a 30-70% proportion respectively. It is worth noting that this division was done within patients, making sure that the patients within the train group were not in the same test group to avoid bias and overfitting problems. In this way, 11 patients were used for training and 5 for testing. All the samples were normalized using MaxMinScaler from sklearn’s library in Python.

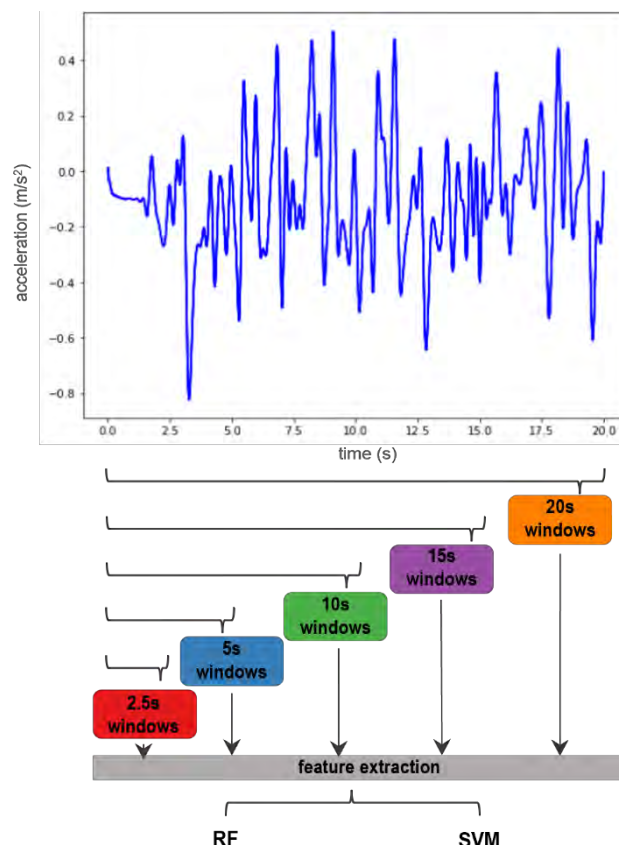


Figure 1: Methodology followed to train and test different models based on segmentation of different sized windows.

Then, two models were used: a Random Forest (RF) and a Support Vector Machine (SVM) classifier. Both models were trained with the five different window sizes proposed, hence ten different models were trained and tested.

3. Results

The F1-score was the selected metric to assess the performance of the classification models. Table 1 shows the average F1-score calculated for the ten different models. Although the SVM model provides slightly better results,

F1 score	2.5s	5s	10s	15s	20s
SVM	70.65%	77.64%	81.22%	79.76%	75.48%
RF	67.73%	74.11%	80.05%	77.93%	77.34%

Table 1: Average F1-score of different classifiers

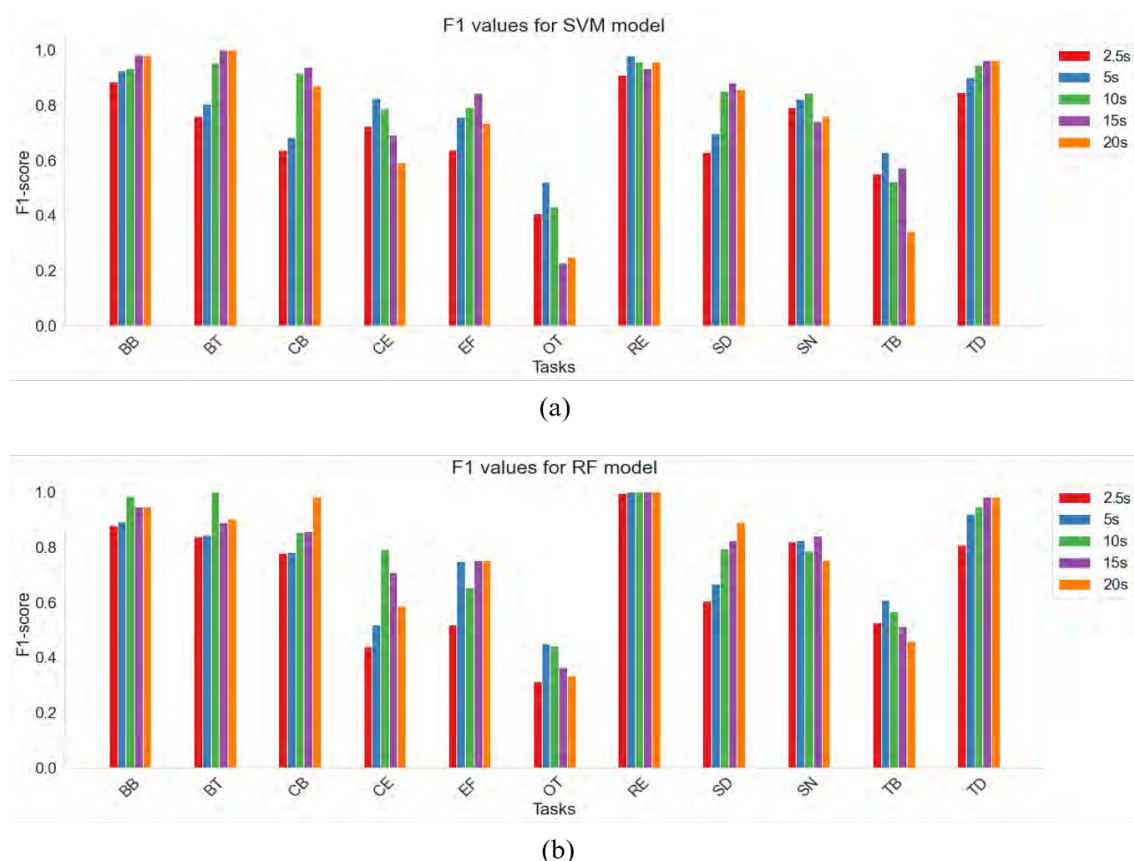


Figure 2: Bar graph showing the F1-score obtained in the classification using different time windows, in both SVM (a) and RF (b) models.

there is no significant difference, so it cannot be determined that one model is better than the other.

The trained and tested model with the 10 seconds window performed the best classification in both models, followed by the 15 seconds one. This shows that larger windows provided better results than the shorter ones.

However, to fully understand how each individual task was being classified by the different models, Figure 2 shows the F1-score obtained by different windows for both the SVM (Figure 2a) and RF (Figure 2b).

Despite the fact that the highest average F1-scores were given by larger windows, these Figures show that, depending on the task, some of them were better classified with models trained with shorter windows.

4. Discussion

The assessment and evaluation of tremor and its evolution with medication dosages remains a challenging problem. The existing clinical evaluation methods are limited by the bias in the performance of the patients and the subjectivity of the evaluator. This could be improved by implementing a continuous monitoring in an every-day life application which could allow to correlate the tremor to the activity carried out. However, the recent developed state-of-art methods in ADL classification in tremor patients have not yielded conclusive evidence yet. They use fixed size windows to train machine learning models, but there is no clear consensus on which window size should be preferably used. Most designs are based on randomly chosen values

from past successful cases, which may not necessarily be the optimal fit for the problem being addressed. In this paper, a different methodology to classify ADLs considering the segmentation of the signals in different time-sized overlapping windows is presented. Although it was noted that the best results were generally given by the 10 seconds window, the performance of the classifiers was analyzed task by task.

On the one hand, some of the tasks were found to be better classified with shorter windows, such as OT and TB tasks. For these tasks, the F1-score was higher with the 5 seconds windows trained model. None of the tasks showed to be better classified with 2.5 seconds windows.

On the other hand, most of the tasks were best classified with windows between 10 and 20 seconds. For example, for CB, SD, and TD tasks the larger windows, 15 seconds, and 20 seconds windows, provided a higher F1-score. Regarding BB, BT, and EF tasks, it was seen that windows equal to 10 seconds or above provided better results, taking into consideration both SVM and RF models. Finally, RE and SN tasks showed smaller differences between the results provided by different windows, reaching similar F1-score values for each of them.

Table 2 shows the different results obtained from the tested models for every different task. The first column shows the F1-scores achieved in [4]. It is worth noting that this previous study did not include RE's classification results, although the data was present and labeled in the dataset. For this reason, the value doesn't appear in the first column, but

was obtained for the following ones. The following two columns show the results obtained from both classifiers, SVM and RF, trained and tested with the segmentation of 10 seconds windows, as it gave the best overall results. Finally, the last two columns show the results obtained considering the optimal window which gave the highest F1-score for each task (see Figure 2).

These last two columns provided the best results, increasing notably the performance in the classification from the previous studies [4]. Therefore, a classification based on different time-sized windows seems to improve the prediction of ADLs.

Tasks	Previous results	10s SVM	10s RF	max SVM	max RF
BB	86,55%	93,15%	97,06%	98,18%	97,06%
BT	85,00%	95,24%	97,78%	100%	97,78%
CB	82,80%	91,53%	88,24%	93,88%	98,04%
CE	52,31%	78,72%	65,96%	82,54%	65,96%
EF	46,40%	79,25%	74,51%	84,44%	78,26%
OT	32,43%	43,24%	36,36%	52,00%	42,31%
RE	-	95,65%	100%	97,92%	100%
SD	83,34%	85,25%	82,14%	88,14%	92,31%
SN	71,43%	84,44%	81,32%	84,44%	81,32%
TB	33,33%	52,46%	62,50%	62,90%	64,00%
TD	92,21%	94,55%	94,74%	96,15%	98,11%

Table 2: F1-score obtained for each task and model.

5. Conclusion and future development

In this study, a different methodology to classify ADLs based on different time-sized windows was presented, which brings flexibility in the recognition of different tasks and movements carried out by the ET and PD patients. In general, it can be stated that considering the 10 seconds window models, the methodology presented gives better results than the ones shown in previous studies, considering that only one of the IMUs (the one on the wrist) was taken into consideration, instead of four in the arm.

Results show that although the segmentation in 10 seconds windows provided a better overall classification, some tasks are better classified depending on the window size, providing even better results than the ones obtained considering only the 10 seconds window model.

Consequently, this supposes a step forward in the objective of real-time activity classification for the study and evaluation of medication dosage and its effects on tremor. In the light of the results presented in this paper, a classification based on the combination of different time windows which segments the signals in different sizes should be considered as a future development. Moreover, a new algorithm to evaluate the optimal window length should be developed.

Acknowledgements

This work has been partially financed by Universidad Politécnica de Madrid under the Técnicas de Control y Análisis de Variables Cinemáticas Aplicadas a la Extracción de Patrones de Movimiento project (RP2109560107) and the National Program for Scientific, Technical and Innovation Research 2021-2023 within the context of the project NETremor (TED2021-130174B-C32).

References

- [1] A. Anouti, W. C. Koller, and K. City, “Articles Tremor Disorders Diagnosis and Management.” vol. 162, no. 6. pp. 510–513, Jun. 1995. PMID: 7618310
- [2] B. Thanvi, N. Lo, and T. Robinson, “Essential tremor - The most common movement disorder in older people,” *Age and Ageing*, vol. 35, no. 4. pp. 344–349, Jul. 2006. doi: 10.1093/ageing/afj072.
- [3] F. Demrozi, R. Bacchin, S. Tamburin, M. Cristani, and G. Pravadelli, “Toward a Wearable System for Predicting Freezing of Gait in People Affected by Parkinson’s Disease,” *IEEE J Biomed Health Inform*, vol. 24, no. 9, pp. 2444–2451, Sep. 2020, doi: 10.1109/JBHI.2019.2952618.
- [4] J. I. Serrano, S. Lambrecht, M. D. del Castillo, J. P. Romero, J. Benito-León, and E. Rocon, “Identification of activities of daily living in tremorous patients using inertial sensors,” *Expert Syst Appl*, vol. 83, pp. 40–48, Oct. 2017, doi: 10.1016/j.eswa.2017.04.032.
- [5] S. Fahn, “Classification of movement disorders,” *Movement Disorders*, vol. 26, no. 6. pp. 947–957, May 2011. doi: 10.1002/mds.23759.
- [6] K. Frank, M. Josefa, V. Nadales, P. Robertson, and M. Angermann, “Reliable Real-Time Recognition of Motion Related Human Activities Using MEMS Inertial Sensors.” in *Proceedings of the 23rd International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2010)*, pp. 2919–2932, 2010.
- [7] L. Sigcha *et al.*, “Deep learning and wearable sensors for the diagnosis and monitoring of Parkinson’s disease: A systematic review,” *Expert Syst Appl*, vol. 229, p. 120541, Nov. 2023, doi: 10.1016/j.eswa.2023.120541.
- [8] H. Nguyen, K. Lebel, S. Bogard, E. Goubault, P. Boissy, and C. Duval, “Using Inertial Sensors to Automatically Detect and Segment Activities of Daily Living in People with Parkinson’s Disease,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 1, pp. 197–204, Jan. 2018, doi: 10.1109/TNSRE.2017.2745418.
- [9] B. Jiang, J. J. Han, and J. Kim, “A Wearable In-home Tremor Assessment System via Virtual Reality Environment for the Activities in Daily Lives,” in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, Jul. 2022, pp. 1117–1120. doi: 10.1109/EMBC48229.2022.9871008.
- [10] M. G. Martín, “Contributions To Human Motion Modeling And Recognition Using Non-Intrusive Wearable Sensors”, Ph.D dissertation, Universidad Politécnica de Madrid, 2022, doi: <https://doi.org/10.20868/UPM.thesis.70493>.

Impact of integrated signals for doing HAR using Deep Learning models

M. Jaén-Vargas¹, J. García Martínez¹, K. Reyes Leiva^{1,2}, M. Trujillo-Guerrero¹, F. Fernandes³, S. Barroso Gonçalves⁴, M. Tavares Silva⁴, D. Simões Lopes^{3,5}, R. Abizeid Daou⁶, J. Serrano Olmedo^{1,7}

¹ Bioinstrumentation and Nanomedicine laboratory (LBN), Center for Biomedical Technology (CTB), Universidad Politécnica de Madrid, Madrid, Spain. lbn@ctb.upm.es

² Engineering Faculty, Universidad Tecnológica Centroamericana UNITEC, San Pedro Sula, Honduras

³INESC ID, Lisbon, Portugal

⁴IDMEC, Instituto Superior Técnico, University of Lisbon, Portugal

⁵Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

⁶Univeridad La Sagesse, Beirut, Lebanon

⁷Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, (CIBER-BBN), Madrid, Spain

Abstract

Human Activity Recognition (HAR) is having a growing impact in creating new applications and is responsible for emerging new technologies. Also, the use of wearable sensors is an important key to exploring the human body's behavior when performing activities. Hence, the use of these dispositives is less invasive and the person is more comfortable. In this study, a database that includes three activities is used. The activities were acquired from inertial measurement unit sensors (IMU) and motion capture systems (MOCAP). The main objective is differentiating the performance from four Deep Learning (DL) models: Deep Neural Network (DNN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and hybrid model Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM), when considering acceleration, velocity and position and evaluate if integrating the IMU acceleration to obtain velocity and position represent an increment in performance when it works as input to the DL models. Moreover, compared with the same type of data provided by the MOCAP system. Despite the acceleration data is cleaned when integrating, results show a minimal increase in accuracy for the integrated signals.

1. Introduction

Human Activity Recognition (HAR) is a growing technology that is having a huge impact due to its great capability to obtain data from wearable sensors in order to have a better understanding of the human body [1], opening all types of possible applications. Fields such as medicine [2] or sports [3], [4] can leverage HAR to make advances towards better treatments and preparations related to human activity and prevention of injuries. This recent interest in HAR technology is also due to the use of Artificial Intelligence, that can extract features from data and later classify and make predictions, allowing a faster processing of data. This data process is made in AI by neuronal networks that try to simulate the human brain,

The neuronal networks work as a group of interconnected neurons similar to the human brain [5], and they process the information from obtained data, recognize patterns, learn and improve. Machine Learning (ML) is the AI technique that makes use of this neural networks with several models that can be helpful in different use cases [6].

These models are formed by different layers emulating interconnected neurons, and they differ between them in terms of complexity and structure. The data is provided as a model which allows the algorithm to learn and classify in order to make predictions.

In HAR, the input data is the obtained one from wearable sensors worn by a person during a physical activity with the objective to measure the body movements [7]. Using this method is possible to give to the model information of a person doing different activities and so, the model can differentiate which data correspond to each activity.

In this study, two types of sensors are used for this purpose, inertial measurement unit sensors (IMU) and motion capture systems (MOCAP), when performing three different activities: walking, sitting and squatting. These activities can contain valuable information about the person's body. This information is passed as an input to the ML models, and it can learn to differentiate the activities and help with the study of the body's health.

The data registered using sensors correspond to tri-axial data. While Mocap primarily provides position data, it also takes into account velocity and acceleration. For IMU the sensors obtain acceleration data, which then needs to be integrated to obtain velocity and position information. Hence, the signal is cleaned so it should be better as input, this is the main case study where MOCAP and IMU will be compared, as for IMU a manual integration is made so the impact of the integration will be evaluated.

To evaluate ML models' performance only de accuracy is considered, then classifying only the accuracy when predicting and classifying is considered. In this research Deep Learning, (DL) an AI technique that is having a great impact nowadays, can bring a new point of view, which can extract the features that models focus om decision making, this way the internal process can be understood [8]. With the knowledge of the internal process, it is possible to see if it is easier for the DL models to differentiate the activities using the integrated signals.

2. Materials and methods

2.1. Preprocessing of data

The sensor's data was acquired in the Lisbon's Biomechanics Laboratory belonging to Instituto Superior Técnico (IST). It was approved by the ethics committee of IST in January 2020 (Ref. nr. 1/2020 (CE-IST)).

MOCAP markers were placed on the body of ten people as well as two IMU: one on the left ankle and the other on the left wrist [8].



Figure 1. MOCAP and IMU sensors placed in ankle and wrist.

Although a MOCAP system consisting of 26 markers was used, only the left ankle and left wrist markers, along with two IMUs located in the same position, were considered for this specific data analysis. Before using DL models, data preprocessing was made to clean the signals in order to obtain better results. The three different activities measured (sitting, walking and squatting) were labelled 0, 1 and 2 for the models to later classify. The data from the two sensors were organised so 8 people was reserved for the training and the 2 remaining to validate and to not have the models know of any previous data about those people.

Each activity was filtered with a Butterworth filter to eliminate unwanted peaks in the signal and the offset of the accelerometer was corrected using a 3D algorithm maintaining the sensor static a certain time, so the only force measured is gravity.

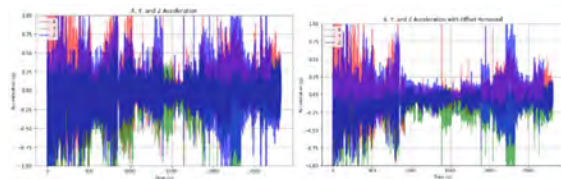


Figure 2. Comparison between raw acceleration data and filtered acceleration data

From MOCAP system acceleration, velocity and position data are obtained directly from the measurement process, but from IMU only acceleration data was obtained, for getting the velocity and position data an integration with the trapeze rule is made, which allows to determine the function f for all the points by applying the algorithm in a recursive way knowing the distance h between two adjacent points.

2.2. Training

With all the data prepared a sliding window technique was used to divide the data into subsequences that will be the input of the Deep Learning models that were defined to be compared: DNN, CNN, RNN and a CNN-LSTM hybrid. DNN is characterised for having a big number of hidden

layers between the input and output, CNN using convolution can do more complex operations, RNN is related to time sequences due to its capacity to remember previous steps, and CNN-LSTM mix the advantages of CNN and LSTM being able to after CNN interpreted subsequences give the output as sequences to LSTM.

A sliding window technique is used to give the data to the model. The data is divided into subsequences [9] and after that 80% of the data is used for training and the remaining 20% for validating if the model is training in a correct way.

2.3. Deep Learning data extraction

After training the algorithms a classification process was carried over in order to understand which models get better result and what are the parts of the data used for making decisions when classifying.

To know which part of the data the algorithm is using to choose between one activity, or another two main techniques were used: Class Activation Map (CMAP) and Features Map. By this method it is possible to see if the velocity and position data is more useful for the algorithm than the acceleration and to see if the same part of the data is used to decide or not.

The Class Activation Map (CMAP) is a heat map that allows seeing in a graphic way which part of the given data is used to predict the activities. This is done by the calculation of the gradients that inform us about the direction in which the loss function has a more prolonged ascent [10]. Being our data two-dimensional the CMAP will be represented as a line being the higher points the points where the algorithms are giving more importance to the data in order to make the classification, but it can be presented in a more visual form by using its matrix form.

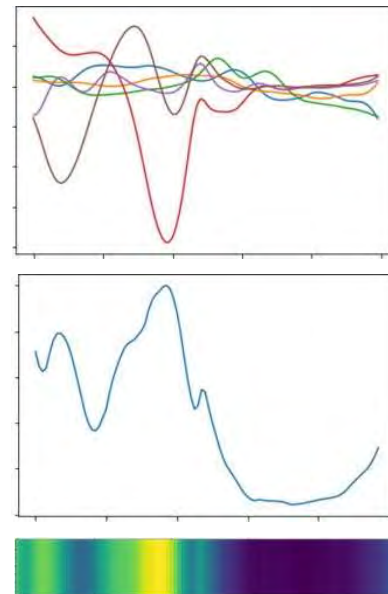


Figure 4. Walking measure of MOCAP acceleration with its CMAP in line and matrix form

On the other hand, the Features Map provide information about the process of classification made in all the different layers, being able to see the development of the

information treatment along all the layers with the output of the final layer being the label of the activity. This information is presented as graphics with different dimensions corresponding to the dimension of each layer, being the tridimensional ones represented as images and the bidimensional as graphic bars.

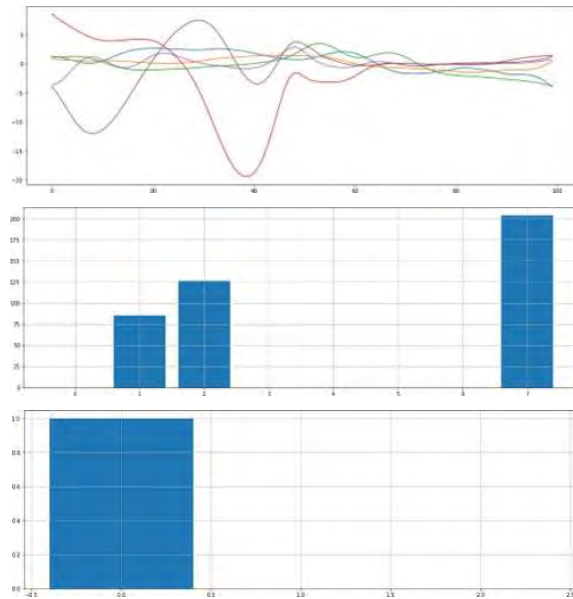


Figure 5. Output of the different layers of a CNN model for an acceleration walking sample

3. Results

As previously mentioned, the main objective is to determine if integrated signals can produce better classification when using Deep Learning.

Model	Acceleration		Velocity		Position	
	MOCAP	IMU	MOCAP	IMU	MOCAP	IMU
DNN	0,9988	0,9773	0,9988	0,9907	0,9995	0,9862
CNN	0,9881	0,9773	0,9948	0,9907	0,995	0,9862
RNN	0,9986	0,9634	0,9931	0,9943	0,993	0,9846
CNN-LSTM	0,9989	0,9995	0,9996	0,9995	0,9997	0,9983

Table 1. Accuracy of Deep Learning models

Model	Acceleration		Velocity		Position	
	MOCAP	IMU	MOCAP	IMU	MOCAP	IMU
DNN	0,0047	0,0642	0,0009	0,0302	0,002	0,0402
CNN	0,0382	0,0642	0,0171	0,0302	0,002	0,0402
RNN	0,006	0,0962	0,0216	0,0302	0,004	0,0463
CNN-LSTM	0,0032	0,0032	0,0014	0,0018	0,001	0,0042

Table 2. Losses of Deep Learning models

As shown in table I, the velocity and position, signals that were integrated, have an accuracy of 99,96% as maximum value for MOCAP velocity, 99,97% for MOCAP position,

99,95% for IMU velocity and 99,83% for IMU position, all being slightly better results than the maximum accuracy of the acceleration for both MOCAP and IMU, which are 99,89% for MOCAP and 99,95% for IMU.

Moreover, table II shows the losses of the models, with the integrated signals being its lowest values 0,09% for MOCAP velocity, 0,1% for MOCAP position, 0,18% for IMU velocity and 0,42% for IMU position, again slightly better results than the obtained with acceleration, 0,32% for both MOCAP and IMU.

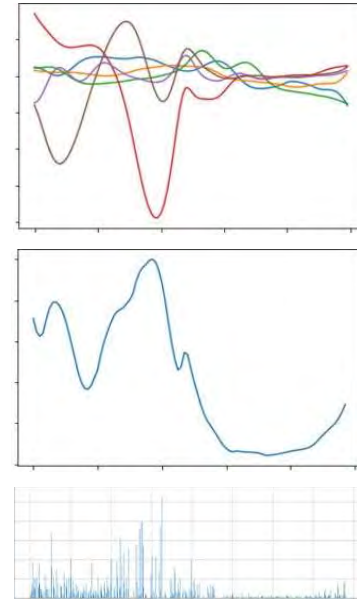


Figure 6. Comparison between input data, CMAP and output of Flatten layer for a walking sample

As shown in Fig. 6 for all the activities the DL models focused on the curves with considerable slopes (of positive or negative value) in order to differentiate the activities and the output of the Flatten layer, in charge of grouping the features obtained in previous layers, is similar to the obtained CMAP, being a good indicator of how the models classify.

To confirm all this results a new database is used to compare the obtained results, having this database a higher number of activities (12 activities from 30 people) from daily living measured with the accelerometer and gyroscope of a mobile phone [11].

Model	Accuracy	Losses
CNN	0,9537	0,1284
RNN	0,9684	0,0941
CNN-LSTM	0,9697	0,0858

Table 3. Accuracy and losses of DL models for new database

Again, the best results are given by the CNN-LSTM model in both accuracy and losses. In this case only acceleration is shown as it is reaffirmed that integrated signals obtain better results but with no significant difference.

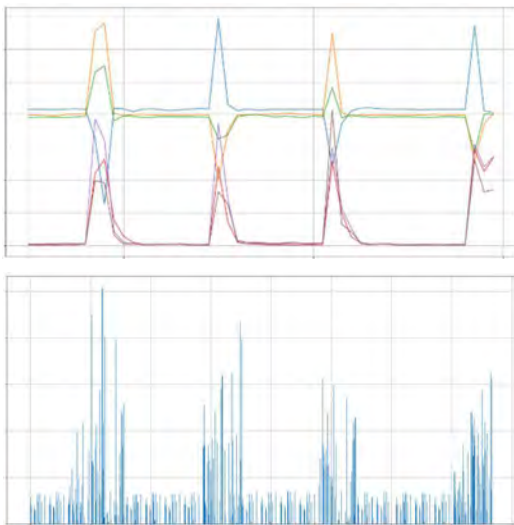


Figure 7. Comparison between input data for 12 activities and output of CNN model Flatten layer.

For feature extraction with Deep Learning the models keep focusing on the curves with considerable slopes in order to make a classification, as shown in Fig. 7.

4. Discussion

Most of the studies use the acceleration as signal to feed the AI models. In this study, the integration of acceleration to obtain velocity and position as input signals to AI algorithms is considered. As consequence, an increment in accuracy and a decrease in losses is obtained compared with acceleration results. Hence, this can be an indicator that working with integrated signals can produce a better classification in HAR.

Furthermore, for IMU the velocity results indicate a better performance than position even though it is an integration from velocity, and it should be the other way around, but this is due to an error that is propagated in integration. Also, this behaviour is confirmed in MOCAP results, where position has indeed better performance metrics. However, MOCAP does not have this problem when integrating because the position data is directly measured by the system.

Regarding the process where the model differentiates and activity from other, two methods were used: CMAP and Feature Map. Both show that Deep Learning models are focused on the curves with considerable slopes to classify the activities.

5. Conclusions

After studying all the results in detail, the values obtained by the different neural networks for acceleration, velocity, and position reveal important data. The hybrid CNN-LSTM model provides the best results in comparison with the rest of the models, being the highest accuracy 99,97% and the lowest loss of 0,1%. Moreover, when comparing the results of integrated and non-integrated signals, there is an improvement in using integrated signals, confirming the hypothesis that cleaner signals facilitate the classification of different networks, but the difference between them is

so minimal that it should not be necessary to integrate signals to obtain good enough results for DL models.

Comparing MOCAP and IMU, we also notice that although MOCAP results are generally better, the difference is minimal. Thus, despite MOCAP being a more expensive sensor system, it produces similar results to cheaper IMU sensors.

Based in Deep Learning data extraction we can also confirm that, for different activities with acceleration, velocity, or position from various sensors, the neural networks focused on curves with significant slopes (positive or negative) to perform classification.

We can assert all of this by the reproduction of the same procedure with a larger database of activities, getting to the same conclusions from the previous database.

References

- [1] T. Sztyley, H. Stuckenschmidt, and W. Petrich, "Position-aware activity recognition with wearable devices," *Pervasive Mob Comput*, vol. 38, pp. 281–295, Jul. 2017, doi: 10.1016/j.pmcj.2017.01.008.
- [2] G. Ogbuabor and R. La, "Human activity recognition for healthcare using smartphones," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Feb. 2018, pp. 41–46, doi: 10.1145/3195106.3195157.
- [3] S. Zhang *et al.*, "Deep Learning in Human Activity Recognition with Wearable Sensors: A Review on Advances," *Sensors 2022, Vol. 22, Page 1476*, vol. 22, no. 4, p. 1476, Feb. 2022, doi: 10.3390/S22041476.
- [4] Z. Zhuang and Y. Xue, "Sport-related human activity detection and recognition using a smartwatch," *Sensors (Switzerland)*, vol. 19, no. 22, Nov. 2019, doi: 10.3390/s19225001.
- [5] M. C. Nwadiugwu, "Neural Networks, Artificial Intelligence and the Computational Brain," Dec. 2020, doi: 10.48550/arXiv.2101.08635.
- [6] L. Sai Ambati and O. El-Gayar, "Human Activity Recognition: A Comparison of Machine Learning Human Activity Recognition: A Comparison of Machine Learning Approaches Approaches," *Journal of the Midwest Association for Information Systems Journal of the Midwest Association for Information Systems (JMWAIIS) (JMWAIIS)*, vol. 2021, 2021, doi: 10.17705/3jmwai.000065.
- [7] N. Y. Hammerla, S. Halloran, and T. Plötz, "Deep, convolutional, and recurrent models for human activity recognition using wearables," *IJCAI International Joint Conference on Artificial Intelligence*, vol. 2016-January, pp. 1533–1540, 2016.
- [8] M. Jaén-Vargas *et al.*, "A deep learning approach to recognize human activity using inertial sensors and motion capture systems," in *Frontiers in Artificial Intelligence and Applications*, IOS Press BV, Oct. 2021, pp. 250–256, doi: 10.3233/FAIA210196.
- [9] M. Jaén-Vargas *et al.*, "Effects of sliding window variation in the performance of acceleration-based human activity recognition using deep learning models," *PeerJ Comput Sci*, vol. 8, 2022, doi: 10.7717/PEERJ-CS.1052.
- [10] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning Deep Features for Discriminative Localization," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-December, pp. 2921–2929, Dec. 2015, doi: 10.1109/CVPR.2016.319.
- [11] J.-L. Reyes-Ortiz, L. Oneto, A. Samá, A. Ghio, X. Parra, and D. Anguita, "Transition-Aware Human Activity Recognition Using Smartphones".

Detector automático de artefactos en señales neuronales basado en técnicas de Inteligencia Artificial

A. Gil-Correa¹, S. Pérez-Velasco^{1,2}, V. Rodríguez-González^{1,2}, H. Hoshi³, Y. Shigihara³, C. Gómez^{1,2}, J. Poza^{1,2,4}

¹ Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, amalia.gil@estudiantes.uva.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

³ Precision Medicine Centre, Hokuto Hospital, Japón

⁴ IMUVA, Instituto de Investigación en Matemáticas, Universidad de Valladolid, Valladolid, España

Resumen

La magnetoencefalografía (MEG) es una técnica no invasiva que permite estudiar el funcionamiento del cerebro en diversas condiciones y patologías. Sin embargo, las señales MEG pueden estar contaminadas por señales indeseadas de origen neuronal, sesgando los resultados. Por ello, resulta necesario minimizar la presencia de estas antes de analizar los datos. Uno de los métodos más empleados para abordar este problema es el Análisis de Componentes Independientes (ICA). Sin embargo, ICA no ofrece una clasificación en función del origen de las componentes de las señales, sino que esta depende, en la mayoría de los casos, de la interpretación subjetiva de un técnico. En el presente trabajo se ha desarrollado un sistema basado en Deep Learning (DL) que, haciendo uso de una variante de la Red Neuronal Convolutiva EEG-Inception, permite diferenciar y clasificar las componentes ICA de origen neuronal y las componentes indeseadas. Para ello, se diseñaron y evaluaron dos sistemas de clasificación diferentes: binario y multiclase. La base de datos empleada estaba formada por señales MEG de 473 sujetos en estado de reposo. La precisión en la clasificación binaria fue del 96.92%, y en la clasificación multiclase se alcanzó el 86.90%. Además, por primera vez, se introdujo un mecanismo para que la propia red realizara una reevaluación de las componentes ICA, lo que reveló que algunas de ellas estaban mal identificadas en la base de datos. Por lo tanto, la herramienta propuesta no solo mejora la clasificación automática de las componentes ICA, sino que también puede facilitar y mejorar el etiquetado manual efectuado por los técnicos, reduciendo la subjetividad asociada a este proceso.

Palabras clave

Magnetoencefalografía, artefactos, Deep Learning, Redes Neuronales Convolutivas, EEG-Inception, clasificación.

1. Introducción

La electroencefalografía (EEG) y la magnetoencefalografía (MEG) son técnicas no invasivas que permiten medir los campos electromagnéticos generados por el cerebro [1]. Los registros MEG ofrecen una mejor resolución espacial y una menor distorsión de la actividad neuronal que los registros EEG, si bien su coste es notablemente mayor [2]. Ambos permiten la identificación de patrones de actividad cerebral. Así, se puede obtener información acerca del funcionamiento del cerebro y de posibles alteraciones patológicas, resultando de vital importancia para el diagnóstico y caracterización de muchas patologías neuronales [3].

Garantizar la obtención de datos fiables es un importante reto tras haber adquirido un registro EEG o MEG. La

presencia de ruido o artefactos, provenientes de fuentes ajenas a la actividad cerebral, puede interferir con las señales de interés. Esto provoca una disminución en la relación señal a ruido (SNR), empeorando en gran medida la calidad de la señal neuronal adquirida. La eliminación o mitigación de estos artefactos es de vital importancia para el estudio y caracterización de la función cerebral y de las alteraciones asociadas a enfermedades que afectan al sistema nervioso central. El Análisis de Componentes Independientes (ICA, *Independent Component Analysis*) es una técnica muy empleada para este propósito, pues permite separar los datos originales en distintas componentes independientes, propias de las diferentes fuentes que generan la señal adquirida. Sin embargo, la clasificación de dichas componentes suele ser efectuada de forma manual por expertos. Se trata de una tarea ardua y costosa, sobre todo si la cantidad de datos obtenidos tras el registro es muy elevada. Además, conlleva el riesgo de introducir errores, dado que implica un alto grado de subjetividad [4]. Parece necesario, por tanto, desarrollar un método de detección automática y objetiva de artefactos en señales neuronales que facilite la tarea y ofrezca resultados fiables en un corto período de tiempo.

La Inteligencia Artificial (IA) agrupa una serie de técnicas que tratan de imitar la inteligencia humana y realizar tareas complejas, desde proponer soluciones a problemas cotidianos a diseñar e implementar sistemas inteligentes de ayuda a la decisión. Con el objetivo de minimizar la complejidad y subjetividad inherentes a la detección de artefactos en señales neuronales a partir de las componentes ICA, en este trabajo se ha optado por hacer uso de la IA. En concreto, se ha planteado una herramienta basada en *Deep Learning*, dada su capacidad de aprendizaje con un elevado número de datos no estructurados, permitiendo analizar y clasificar la señal sin necesidad de que se extraigan características previas, y su analogía con las redes neuronales propias del ser humano [5]. Un entrenamiento adecuado de estas puede permitir la identificación de componentes independientes asociadas a ruido y actividad cerebral, lo que facilitaría el trabajo del técnico encargado del etiquetado y reduciría el tiempo necesario para esta tarea.

El objetivo principal de este trabajo consiste, por tanto, en el diseño y generación de una herramienta fiable que permita, de manera objetiva, detectar las componentes ICA ruidosas en registros MEG, facilitando y objetivando la tarea a los técnicos especialistas.

2. Materiales y métodos

2.1. Base de datos

Los datos empleados proceden de un total de 473 sujetos. Los registros fueron realizados en el Hospital Hokuto (Obihiro, Japón). Para adquirir las señales MEG se utilizó un sistema MEG Vision PQ1160C (Yokogawa Electric) con 160 canales, mediante el uso de gradiómetros axiales. Se empleó una frecuencia de muestreo de 1000 Hz y un filtro paso-bajo a 200 Hz. Los registros tuvieron una duración de cinco minutos por cada sujeto. En dicho tiempo los participantes se encontraban tumbados, con los ojos cerrados y en reposo. Para cada sujeto y a partir de la señal MEG se obtuvieron 160 componentes ICA con el algoritmo *Extended Infomax ICA*. Posteriormente, un técnico especialista etiquetó cada componente ICA según la consideraba actividad cerebral o artefacto. El etiquetado que se realizó fue el siguiente: “0” para componentes cerebrales, “1” para componentes cardíacas, “2” para componentes de red eléctrica, “3” para componentes oculares y “4” para otro tipo de componentes no neuronales. Además, se han tenido en cuenta sólo las 40 primeras componentes, dado que son las que tienen mayor potencia y, por lo tanto, recogen la mayoría de la información de la señal original. Como los artefactos se sitúan mayoritariamente en las primeras componentes, al escoger las 40 primeras se recoge la mayor parte de estos. En la Tabla 1 se puede ver el número de muestras que contiene la base de datos para dichas componentes. Para un análisis posterior en el que se realiza una reevaluación del etiquetado se han tomado las 40 primeras componentes, pero debido al desbalanceo se han añadido los artefactos de las 160 componentes totales; es decir, se han considerado todos los artefactos de la base de datos, pero sólo las 40 primeras componentes etiquetadas como no artefacto, ya que son las más relevantes para reconstruir la señal y a las cuales el técnico habrá puesto mayor atención en el etiquetado. La razón por la cual se eliminan estas componentes en ambos análisis es que el técnico encargado de etiquetarlas puede equivocarse al clasificar una componente como cerebral, mientras que es bastante improbable que una componente determinada artefacto no lo sea.

2.2. Modelos de clasificación

Con el fin de identificar las componentes ruidosas se han propuesto dos modelos de clasificación de artefactos, una clasificación binaria y una clasificación multiclase. El primer modelo se enfoca en identificar la presencia de artefactos en la señal, sin especificar el tipo de componente ruidosa. Se trata de un modelo sencillo capaz de diferenciar entre componentes de origen cerebral y componentes correspondientes a artefactos. El segundo modelo busca

Etiqueta	Número de componentes ICA
Componentes cerebrales (“0”)	17934
Componentes cardíacas (“1”)	727
Componentes de red eléctrica (“2”)	28
Componentes oculares (“3”)	43
Otras componentes (“4”)	188

Tabla 1. Número de componentes ICA de cada tipo.

detectar la presencia de artefactos de diferentes orígenes en la señal registrada, y llevar a cabo una clasificación de estos en una categoría específica. Dicha categoría es la asignada por el técnico que se encargó de etiquetar cada componente, y que se resume en la Tabla 1.

2.3. Configuración de la red EEG - Inception

En este trabajo se ha hecho uso de la red neuronal convolucional (CNN, *Convolutional Neural Network*) EEG-Inception; esta permite la clasificación de señales procedentes de registros de EEG y la detección de potenciales relacionados a eventos (ERP, *Event-Related Potentials*) [6]. Puesto que originalmente esta CNN fue diseñada para el análisis de señales EEG, ha sido necesario modificar sus bloques para adaptarla al propósito de este trabajo, según se ve en la Figura 1. En la red EEG-Inception modificada se han introducido dos entradas. La primera consiste en las características extraídas de la base de datos, también denominadas *features*. Cada *feature* es una componente ICA extraída a partir de una señal de 5 minutos de duración. La segunda entrada está formada por una matriz mezcla que sirve para reintroducir una componente ICA en la señal MEG original, es decir, permite conocer qué canales registran la componente evaluada. Con la primera entrada se ha llevado a cabo el grueso de las modificaciones de la red:

- Para la clasificación binaria se realizan dos convoluciones de 25×1 , es decir, dos ventanas temporales de 125 milisegundos, mientras que para la clasificación multiclase se realizan tres convoluciones de 100×1 , 50×1 y 25×1 , que se corresponden con tres ventanas de 500, 250 y 125 milisegundos. En todos los casos se submuestreó a una frecuencia de 200 Hz.
- A continuación, se concatenan las salidas obtenidas tras las convoluciones, y se reducen las dimensiones aplicando un *Average-Pooling*. Cabe destacar que la capa *Depthwise* ha sido eliminada dado que únicamente se introduce una componente temporal cada vez, es decir, sólo se dispone de un canal de entrada.
- El siguiente bloque lo forman otras dos convoluciones de 6×1 (ventana de 3 segundos para el caso binario y de 12, 6 y 3 segundos para el caso multiclase), dado que tras el *Average-Pooling* del primer bloque las escalas se ven reducidas, permitiendo la extracción de características adicionales.
- Para terminar con esta entrada, se efectúan dos convoluciones de 8×1 y de 4×1 (48 y 384 segundos para el caso binario), seguidas de sus *Average-Pooling* correspondientes. La última capa extrae las características de la señal al completo, es decir, es función de los cinco minutos de señal introducida. Finalmente, a la salida obtenida asociada a la primera entrada se le aplica un *Global Average-Pooling*, que calcula la media de los valores de una ventana de la matriz de salida. Esto se concatena con el único bloque aplicado a la entrada ICA, un *Global Max-Pooling*, que toma el valor máximo de la matriz de recomposición. Tras esto, se combina la información de ambas entradas en un bloque *Dense* de activación *softmax* con el que se obtiene la clasificación binaria o multiclase.

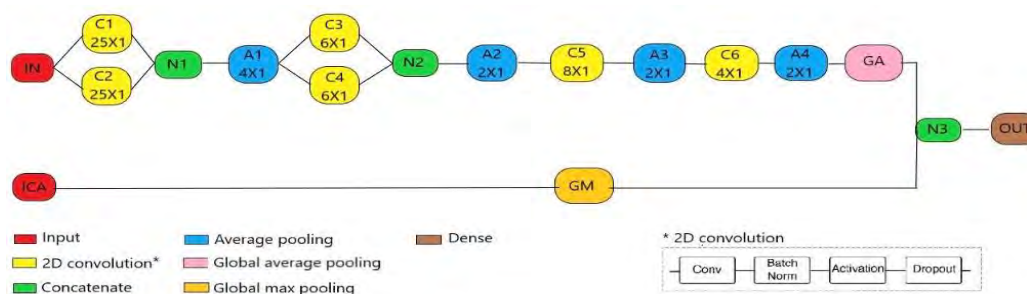


Figura 1. Arquitectura de la red EEG-Inception adaptada a este trabajo.

2.4. Parámetros de entrenamiento de la red

A continuación, se describen algunos de los hiperparámetros que definen el comportamiento de la red. El número de filtros temporales escogido ha sido de 8, ya que permite la correcta extracción de características de la señal MEG sin suponer un elevado coste computacional. Además, el *Dropout rate* que se ha elegido ha sido de 0.4; este valor se ha considerado óptimo para evitar el sobreajuste a los datos de entrenamiento que pudiera llegar a producirse.

Por otro lado, el conjunto de datos está desequilibrado, es decir, hay muchas más componentes etiquetadas como no artefactos que como artefactos. Por ello, se ha balanceado la importancia de cada clase durante el entrenamiento; en concreto, se le ha dado mayor importancia a las clases minoritarias y menor a las mayoritarias.

Finalmente, para determinar la calidad del proceso de clasificación, se han empleado diferentes métodos de evaluación del rendimiento: sensibilidad, especificidad, precisión, matriz de confusión y coeficiente *kappa*. Se ha hecho uso del método de validación cruzada *k-fold* con un valor de *k* igual a 5, dado que permite una adecuada generalización de los resultados obtenidos sin aumentar el coste computacional. Además, se ha empleado la técnica de *early-stopping*, que evita que se produzca un sobreajuste durante el entrenamiento. Por último, la función de pérdidas empleada ha sido *Categorical Cross-Entropy*, que calcula la diferencia entre la distribución de probabilidad de las etiquetas reales y la predicha por el modelo para cada una de las clases, buscando minimizar la distancia entre dichas distribuciones [6].

3. Resultados

3.1. Clasificación binaria

En la clasificación binaria se han obtenido valores de sensibilidad del 67.14%, especificidad del 98.56%, precisión del 96.92%, y coeficiente *kappa* 0.6824, es decir, un valor moderado de concordancia. La matriz de confusión puede verse en la parte izquierda de la Figura 2.

Además, se ha llevado a cabo una reevaluación de la base de datos por la propia red. Se han tomado aquellas componentes etiquetadas como no artefacto cuyo *score* (valor numérico que indica la probabilidad de que una determinada clase sea la correcta para una determinada entrada) para dicha categoría era menor que un cierto valor, en concreto 0.9, para calcular nuevamente los parámetros

de la red sin ellas y sin reentrenar. Luego se ha repetido este último proceso, pero reentrenando la red.

El proceso se ha repetido *N* veces, siendo *N* el punto en el cual la red deja de mejorar la sensibilidad ofrecida. Se ha descubierto mediante un análisis observacional que, efectivamente, se trata de componentes dudosas debido a diversos eventos que aparecen en ciertos tramos de la señal y, en total, se han eliminado de la base de datos 5296 de estas componentes. La precisión conseguida tras la última iteración ha sido del 99.26% (sensibilidad = 95.09%, especificidad = 99.96%, y coeficiente *kappa* = 0.9696), lo cual indica un aumento de ésta respecto al análisis inicial.

3.2. Clasificación multiclase

Para la clasificación multiclase, con el dataset original, la sensibilidad de la red frente a la componente cardíaca ha sido del 74.41%, frente a la componente asociada a la red eléctrica ha sido del 42.85%, frente a la componente ocular ha sido del 81.39%, y frente a otro tipo de componentes artefactuadas ha sido del 25.00%. La especificidad ha sido del 88.13%, la precisión lograda del 86.90%, y el coeficiente *kappa* 0.3450. En la parte derecha de la Figura 2 se muestra la matriz de confusión correspondiente.

4. Discusión

La red desarrollada ha permitido una evaluación objetiva de las componentes ICA, ofreciendo unos resultados de clasificación moderados. Los resultados obtenidos con la herramienta diseñada han demostrado que la clasificación manual y subjetiva realizada por los técnicos especialistas inducía a error en un número notable de ocasiones. Se ha observado que varias señales que originalmente estaban etiquetadas como componentes no artefactuadas en realidad sí que contenían partes ruidosas que limitaban el aprendizaje de la herramienta desarrollada.

Estudios previos muestran que incluir las señales electrocardiográficas (ECG) o electrooculográficas (EOG) puede ser beneficioso para la detección de artefactos [7].

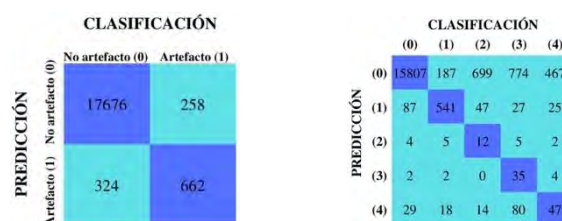


Figura 2. Matrices de confusión obtenidas tras la clasificación binaria y multiclase para las 40 primeras componentes ICA.

Así, basándose en un análisis previo mediante ICA y haciendo uso de redes neuronales, es posible aplicar diferentes estrategias de separación dependiendo del origen de la actividad que se esté detectando. Otros estudios han hecho uso de *higher order statistics* junto con medidas de entropía para discriminar de forma automática los artefactos [8]. Por otro lado, diversos estudios emplearon exclusivamente señales MEG, tal y como se ha planteado en este trabajo. Duan et al. (2013) emplearon una Máquina de Vector Soporte entrenada con cinco características extraídas de las componentes ICA, para llevar a cabo una clasificación binaria, y consiguieron la mayor especificidad de todos los estudios analizados con únicamente 10 sujetos. La precisión que obtuvieron fue del 97.41% [9]. Garg et al. llevaron a cabo dos estudios, uno en el año 2017 y otro en el año 2021. En el primero se emplearon una CNN 1D y 49 sujetos, habiéndose realizado el etiquetado previo por un experto, y la precisión obtenida fue del 95.86% [10]. Los resultados mejoraron notablemente tras el segundo estudio, donde se disponía de 217 señales MEG adquiridas durante diferentes tareas. Aquí se empleó una CNN para clasificar los artefactos oculares, de origen cardíaco, y otro tipo de artefactos, y la precisión que se obtuvo fue del 98.95% [11]. El hecho de emplear únicamente las señales MEG y la alta precisión obtenida los llevó a concluir que no era necesario la identificación manual de artefactos. Por otro lado, el modelo diseñado en este trabajo supera notablemente en el número de sujetos empleados al resto de estudios analizados. Además, se han propuesto hasta tres modelos diferentes de detección de artefactos, lo que hace que sea una herramienta versátil que garantiza la obtención de resultados fiables.

Aunque la herramienta diseñada haya mostrado su utilidad, existen una serie de limitaciones y líneas futuras que hay que mencionar. La primera limitación está asociada al número de sujetos de la base de datos, lo cual afecta la validez y la generalización de los resultados obtenidos. En segundo lugar, la calidad de los datos juega un papel clave en el rendimiento de la red; como ha sido entrenada con componentes ya etiquetadas por medio de un técnico, si dichas etiquetas son erróneas o imprecisas, la red neuronal puede aprender de forma incorrecta y comenzar a clasificar equivocadamente. Por otro lado, podría extenderse el uso de la herramienta diseñada a señales EEG, dado que la red EEG-Inception fue pensada para tratar con este tipo de señales y el análisis sería muy similar al realizado con señales MEG. Además, una posible continuación de este trabajo sería detectar aquel tramo de señal neuronal que está contaminado con un artefacto, pero que no afecta a la componente completa sino parcialmente, y tratar de mitigar su efecto para poder reconstruir posteriormente la señal cerebral. Finalmente, se podría hacer uso de técnicas de *Explainable Artificial Intelligence* (XAI), que permiten entender los resultados ofrecidos por la red entrenada.

5. Conclusiones

Los artefactos presentes en las señales provenientes de registros MEG distorsionan la señal neuronal subyacente, ocasionando una degradación en su calidad y dificultando su interpretación. El uso de la CNN EEG-Inception

adaptada a este trabajo ha demostrado ofrecer una alta capacidad de detección y clasificación de los artefactos más comunes en señales MEG, consiguiendo disminuir el sesgo asociado al etiquetado manual de componentes ICA que puedan estar contaminadas por artefactos. Además, la herramienta diseñada ha sido capaz de mejorar la calidad de la base de datos mediante un proceso iterativo de reevaluación para mejorar su desempeño.

Agradecimientos

Este estudio ha sido financiado por el ‘Ministerio de Ciencia e Innovación – Agencia Estatal de Investigación’ y el Fondo Europeo de Desarrollo Regional (FEDER), bajo el proyecto PID2022-138286NB-I00, y por el CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto de Salud Carlos III. S. Pérez-Velasco es beneficiario de una ayuda PIF de la Consejería de Educación de la Junta de Castilla y León.

Referencias

- [1] M. X. Cohen. Where Does EEG Come From and What Does It Mean?, *Trends Neurosci*, vol. 40, 2017, pp. 208–218.
- [2] V. Rodríguez-González et al. Exploring the Interactions Between Neurophysiology and Cognitive and Behavioral Changes Induced by a Non-pharmacological Treatment: A Network Approach, *Front Aging Neurosci*, vol. 13, 2021, p. 696174.
- [3] K. Sternickel et al. Biomagnetism using SQUIDS: status and perspectives. *Supercond Sci Techn*, vol. 19, 2006, p. S160.
- [4] P. Croce et al. Deep Convolutional Neural Networks for feature-less automatic classification of Independent Components in multi-channel electrophysiological brain recordings. *IEEE Trans Biomed Eng*, vol. 66, 2018, pp. 2372–2380.
- [5] K. O’Shea et al. An Introduction to Convolutional Neural Networks. arXiv preprint arXiv:1511.08458, 2015.
- [6] E. Santamaria-Vazquez et al. EEG-Inception: A Novel Deep Convolutional Neural Network for Assistive ERP-Based Brain-Computer Interfaces, *IEEE Trans Neural Syst Rehabil Eng*, vol. 28, 2020, pp. 2773–2782.
- [7] L. Breuer et al. Ocular and cardiac artifact rejection for real-time analysis in MEG, *J Neuroscience Methods*, vol. 233, 2014, pp. 105–114.
- [8] N. Kamal Al-Qazzaz et al. Automatic artifact removal in EEG of normal and demented individuals using ICA-WT during working memory tasks, *Sensors*, vol. 17, 2017, p. 1326.
- [9] F. Duan et al. Boosting specificity of MEG artifact removal by weighted support vector machine. En *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 6039–6042. 2013.
- [10] P. Garg et al. Automatic 1D convolutional neural network-based detection of artifacts in MEG acquired without electrooculography or electrocardiography, *2017 International Workshop on Pattern Recognition in Neuroimaging (PRNI)*, Toronto, ON, Canada, 2017, pp. 1–4.
- [11] A. H. Treacher et al. Megnet: automatic ICA-based artifact removal for MEG using spatiotemporal convolutional neural networks. *NeuroImage*, 241:118402, 2021.

Measuring cognitive load in a simulation game

B. Larraga-García¹, V. Ruiz Bejerano¹, Á. Gutiérrez¹

¹ ETSI de Telecomunicación, Universidad Politécnica de Madrid, Madrid. España.
{blanca.larraga, veronica.ruiz, a.gutierrez}@upm.es

Abstract

Cognitive Load (CL) measurement has gained interest across various domains as performance and efficiency can be affected if CL is not properly managed. The human cognitive structure has some limitations which are important to consider when accomplishing a task. Additionally, the study on how a primary task is impacted when a secondary task is accomplished has been studied. This affects to the working load memory and, therefore, to the CL. This work uses a tool, a multitasking game, that allows to infer cognitive load. This game considers a primary task and different types of secondary tasks: visual, memory and auditory. An experimental test has been performed using this game and objective, subjective and performance information has been gathered. The objective information was obtained thorough different sensors, the subjective information through the NASA-TLX and Paas questionnaires and the performance information through a score value provided by the game. An impact on performance on the primary task was perceived when secondary tasks were activated. Nevertheless, this impact is different depending on the type of secondary task accomplished. The memory task seems to impact the most on the performance on the primary task. Additionally, the subjective information did not correlate with any other variable gathered along the whole experimental test. With respect to the objective information, the eye measurements correlate with different physiological variables, showing that they provide reliable information with respect to measuring CL. Finally, interesting results on how the different difficulty levels impact performance is obtained.

1. Introduction

Interest in measuring Cognitive Load (CL) has increased not only in the learning environment but also in many other domains such as surgery, driving, manufacturing or emergency services [1]. This is due to the fact that, an appropriate management of the CL may support performance and efficiency [2]. The CL Theory (CLT) began with the aim to solve problems on students' learning in the late 1970s. Since then, a lot of research has been made on this field providing information about how the human cognitive structure works. Considering this cognitive structure, the new information is managed by the working memory, which has a limitations in time and capacity. Nevertheless, with time, some information goes to the long-term memory in which such capacity and time limitations do not exist. Therefore, to avoid unnecessary information while performing a task will avoid to overload the working memory capacity. This involves also to limit distractions and to deliver the information in an appropriate manner so that the working memory is not affected. As a consequence, the CL will not increase. However, if this is not properly man-

aged, problems may arise. When mental task demand increases, people become unable to maintain adequate engagement in the task performed which leads to an increase in errors that, in some areas, might have an enormous impact. Nevertheless, it is not easy to measure the CL and different approaches have been taken in the literature. These approaches rely on either subjective, objective or performance information. The most frequent approaches for assessing CL have been subjective measurements [3, 4]. These measurements gather the information from users through specific questionnaires. Nevertheless, new approaches have been developed to integrate objective information, such as physiological data from the body [5, 6]. In these studies the impact on cardiovascular, temperature, respiration and electrodermal measures are presented. They state that these variables are the most popular in CL assessment. Additionally, the eyes provide valuable information according to [7], which states that the pupil diameter increases when CL increases. This is because the autonomous nervous system is affected when the CL increases and therefore, the pupil diameter changes accordingly. Moreover, other eye measures such as fixations and blinks are also of interest [8]. As the CL changes the number of fixations as well as the duration is affected; similarly, the number of blinks and their duration change. These measures have been proposed as the most reliable physiological information [2]. Finally performance measures, which consider how a task is performed, are also considered. This is specially studied when different tasks are performed simultaneously, as the human cognitive structure is composed of multiple interdependent modules which have a great potential but, they are limited to execute one task at a certain time, which is a human cognitive structure limitation. These studies focus on measuring how performance on a task is affected when another one is performed at the same time. This phenomenon is known as multitasking, and its influence on CL has also been extensively examined in the literature[9]. Considering all this, there are studies that find some correlations between the information provided by this three sources of data (subjective, objective or performance) whereas other do not; nevertheless, all of them agree that further research is needed. The aim of this study is to analyze the impact that a multitasking game that combines a primary and a secondary task has on CL by analyzing objective, subjective and performance information. Furthermore, the investigation will encompass an examination of how various types of secondary tasks influence performance on the primary task.

2. Materials

2.1. Participants

Seven healthy participants volunteered to participate in the study: 3 women and 4 men, with an average age of 28 years ($SD \pm 6.8$). The total time of the study lasted 60 minutes. All the participants had normal or corrected-to-normal vision and were right-handed.

2.2. Software used

The software used within this study has been developed in Unity 3D [10]. The game presents a multitasking environment in which different types of tasks are implemented to be executed simultaneously. There is a primary task, which consists on a blue square jumping a red obstacle as shown in Figure 1.

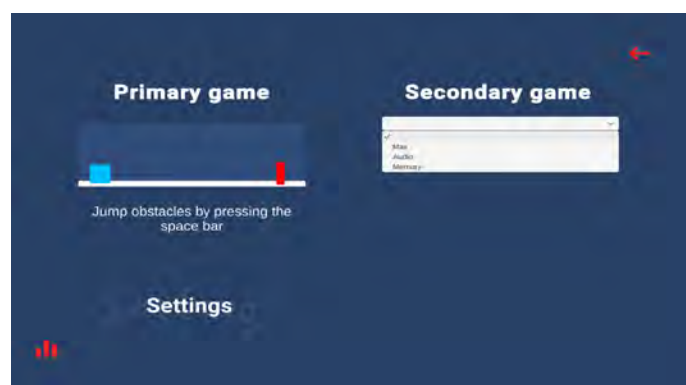


Figure 1: Multitasking game in which the primary task and the different secondary tasks are shown.

A secondary task is implemented taking into account visual, auditory and memory inputs.

- **Visual:** the secondary task that deals with visual inputs presents a list of numbers and the participant needs to find the maximum one. This list of numbers changes considering if the answer provided is correct or not. If the answer is correct the range of number presented increase and if it is not correct, the range of numbers remains the same.
- **Memory:** the secondary task that uses the memory is based on the n-back task. A white square appears with a small blue square inside. This blue square could be placed in any of the four corners. The participant must remember the position of the small square of the first figure, for the 0-back task. Therefore, a sequence of figures appears and each of them must be compared with the first one to identify if they match or not. For the 1-back task, each figure must be compared with the previous one within the sequence of figures. And finally, for the 2-back task, each figure must be compared with the one that appeared two positions before.
- **Auditory:** the secondary task that deals with auditory input consists on listening to a list of figures and colours with the goal to identify the one that is repeated. The list will increase its length if the

answer is correct and it will remain with the same length if the answer is not correct.

Additionally, both primary and secondary tasks, have different difficulty levels: low, medium and high. This means that for the primary and secondary tasks that involves visual and auditory inputs, the speed in which the bar moves or in which the sequence of numbers or figures are displayed/listened changes. For the secondary task that involves memory, the difficulty changes from low (0-back task), to medium (1-back task) and to high (2-back task).

2.3. Hardware used

Several hardware devices are used to measure the vital signs and the eye measurements. The Zephyr BioModule inserted into the Zephyr strap gathers the respiration rate [11]. The Empatica wristband is used to gather the superficial temperature and the galvanic skins response [12]. The Fitbit Versa 2 gathers the heart rate [13]. Finally, the Pupil Core eye tracking platform shown in Figure 2 processes the information with respect to the pupil size, the blink and the fixations duration [14]. The Pupil Core eye tracker has three cameras: the world camera that records data at 120 Hz and two eye cameras that gathers all the eye-related information at 200 Hz. For this pilot study, the pupil size in millimeters as well as the duration of the blinks and the fixations, in seconds, are analyzed.



Figure 2: Pupil core eye tracker[14].

Finally, the subjective information is also gathered by using the NASA-TLX questionnaire [15] and the Paas scale [16]. The NASA-TLX questionnaire measures the workload experienced when a task is performed and takes into account six dimensions: mental, physical, temporal, effort, performance and frustration. It is currently the most used one. On the other hand, the Paas scale focuses only on the mental effort experienced. Even though it considers only one aspect, according to [17] this scale is appropriate to detect Cognitive Load changes.

3. Methods

The study accomplished within this work consisted of a set of games in which the different complexities of both, primary and secondary tasks were involved. Each participant performed a session composed of 18 trials of 0.65 minutes each of them. Every trial comprised either just the primary task or a primary and a secondary task at the same time. These trials were grouped in three blocks with 6 trials each of them. After each block, a

5-minutes break was made. The combination of trials were randomized and in Table 1, one example of a block of trials is shown.

Trial	Task	Task	Task	Task
1	P1	P1 & V1	P1 & V2	P1 & V3
2	P1	P1 & M1	P1 & M3	P1 & M2
3	P1	P1 & A3	P1 & A1	P1 & A2
4	P3	P3 & M2	P3 & M1	P3 & M3
5	P3	P3 & A1	P3 & A2	P3 & A3
6	P3	P3 & V3	P3 & V1	P3 & V2

Table 1. A set of trials in which P corresponds to the primary task, V to the visual, M to the memory and A to the auditory secondary task. Additionally, the numbers 1, 2, 3 refer to the difficulty levels low, medium and high respectively.

Once a block was finished and before the 5-minutes break, all the participants fulfilled a NASA-TLX questionnaire which measures workload and a Paas scale to measure mental effort experienced.

4. Results and discussion

All the physiological information from the participants were gathered and analyzed. First of all, an analysis on the performance on primary and secondary tasks was performed. As shown in Figure 3, it is perceived that the primary task is affected when a secondary task is performed at the same time. In blue, the collisions which refer to the performance on the primary tasks increases when a secondary task is included. In fact, the secondary task that affects the most on the primary task performance is the visual task, followed by the memory task and finally by the auditory one. Nevertheless, when having a look at the performance of the secondary tasks, wrongs in Figure 3, it is seen that the auditory game seems to be the one that performed the worst. Finally, the Time-Up variable which counts the percentage of times that the answer is not provided within the given time is also shown. Nevertheless, its impact is almost imperceptible.

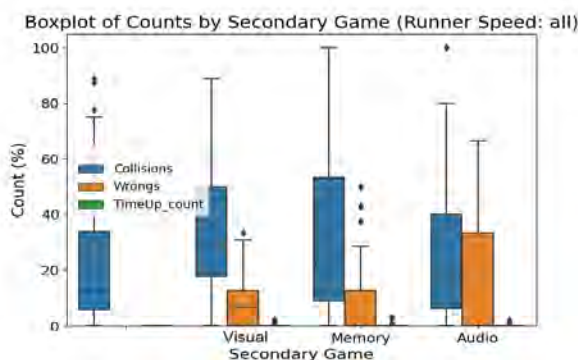


Figure 3: Primary and secondary task performance considering the percentage of collisions, errors and time-up events.

Then, a deep analysis on all the relationship between the physiological variables, the subjective information

provided by the questionnaires and the performance on the game is carried out. To accomplish this analysis, the values per physiological variable considered are: the maximum, the minimum, the average and the peak-to-peak values; and Figure 4 shows the results obtained. The subjective information is the one provided by the NASA-TLX and Paas questionnaires, and the performance information is provided by the score value obtained once a block of trials is finished. First of all, the questionnaire responses do not present a correlation higher than 0.5 with any other variable. This coincides with previous works, that highlight that the response to these questionnaires may vary a lot between individuals [4, 18] and that, additionally, the cognitive load varies along a task and these questionnaires are fulfilled once the task has finished [19]. Therefore, they do not estimate cognitive load in real time.

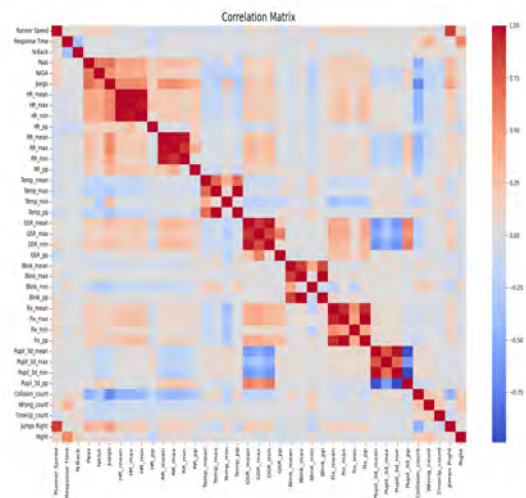


Figure 4: Correlations between all the objective and subjective information gathered from the whole session for all the participants.

The performance on the game correlates specially with heart rate and respiration rate increasing the score obtained on the game when the heart and respiration rates increase. This could be related with the motivation and excitement on the game [20]. Then, an analysis on the physiological variables is performed as they are able to detect changes along the game providing real-time information to assess cognitive load. Between these variables, some correlations are obtained. First of all, the change in the dimension of the pupil (pupil peak-to-peak variable) correlates with the galvanic skin response value as, according to [2] the pupil variation is an indicator of cognitive load as well as the galvanic skin response [21]. Additionally, the fixations duration correlates also with the heart rate and the galvanic skin response. Usually, when the cognitive load increases the fixation duration increases as the attention is focused in order to try to solve a task [22]. According to [5, 23] eye measures provide reliable information that can be measured easily by using a single device whereas to obtain similar

information from other physiological variables may require the use of more devices. On top of that, the right and wrong responses do not show any correlation with the objective, subjective and performance information. What it is found relevant is that when the response time increases there are more right answers. This makes sense as a higher response time means that there are more seconds given to the user to provide an answer. Therefore, if this time is higher, there is more time to provide an answer and the number of correct responses increases. What is interesting to see is that, the number of times that the obstacle is avoided increases with the speed of the obstacle. This was perceived during the sessions as to avoid the obstacle at lower speeds seems to be more complex as more attention is needed. With higher speeds, the jumps are more automatically performed.

5. Conclusions

This work has successfully measured objective, subjective and performance information when a multitasking activity is accomplished. Additionally, how a secondary task impacts on the performance of a primary task has been studied. Moreover, relationships between all these variables have been extracted. Of course, this is a preliminary work with some limitations such as the number of participants that should be increased to confirm the preliminary findings obtained within this work.

Acknowledgements

This research has been supported by the European Commission within the context of the project RESCUER, funded under EU H2020 Grant Agreement 101021836. Additionally, it has been partially supported by the Universidad Politécnica de Madrid under the call First Projects 2023 and under the Técnicas de Control y Análisis de Variables Cinemáticas Aplicadas a la Extracción de Patrones de Movimiento project (RP2109560107).

References

- [1] A.T. Duchowski, C. Biele, A. Niedzielska, K. Krejz, and Kiefer P. The index of Pupillary Activity. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1(1), 2018.
- [2] R. Cong, K. Tago, and J. Qun. Measurement and verification of cognitive load in multimedia presentation using an Eye Tracker. *Multimedia Tools and Applications*, 81(7), 2022.
- [3] G. Cierniak, K. Scheiter, and P. Gerjets. Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior*, 25(1), 2009.
- [4] J. Leppink, F. Paas, C. P. van der Vleuten, T. van Gog, and J. J. van Merriënboer. Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, 45(1), 2013.
- [5] F. N. Biondi, B. Saberi, F. Graf, J. Cort, P. Pillai, and B. Balasingam. Distracted worker: Using pupil size and blink rate to detect cognitive load during manufacturing tasks. *Applied Ergonomics*, 106(1), 2023.
- [6] O. Anmarkrud, A. Andresen, and I. Braten. Cognitive load and working memory in multimedia learning: Conceptual and measurement issues. *Educational Psychologist*, 54, 2019.
- [7] S. Chen, J. Epps, and F. Chen. An investigation of pupil-based cognitive load measurement with low cost infrared webcam under light reflex interference. *Proceedings of 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1(1), 2013.
- [8] J. Hyönä, Radach R., and Deubel H. *The mind's eye cognitive and Applied Aspects of Eye Movement Research*. North-Holland, 2003.
- [9] Ö. Örun and Y. Akbulut. Effect of multitasking, physical environment and electroencephalography use on cognitive load and retention. *Computers in Human Behavior*, 92:216–229, 2019.
- [10] Unity 3D. <https://unity.com/es>, 2023.
- [11] Medtronic. <https://www.zephyranywhere.com/system/components>, 2023.
- [12] Empatica. <https://www.empatica.com/research/e4/>, 2023.
- [13] Fitbit. <https://www.fitbit.com/global/es/products/smartwatches/versa>, 2023.
- [14] Pupil Labs. <https://pupil-labs.com/products/core/>, 2023.
- [15] NASA:B. Gore. <https://humansystems.arc.nasa.gov/groups/TLX/>, 2023.
- [16] F. Paas. Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84:429–434, 1992.
- [17] P. Ayres. Something old, something new from cognitive load theory. *Computers in Human Behavior*, 113:106503, 2020.
- [18] A. Skulmowski and D.G. Rey. Measuring Cognitive Load in Embodied Learning Settings. *Frontiers in Psychology*, 8(2), 2017.
- [19] M. Ahmad, I. Keller, D. Robb, and Katrin Lohan. A framework to estimate cognitive load using physiological data. *Personal and Ubiquitous Computing*, 27(1), 2020.
- [20] G. Chanel, C. Rebetez, M. Betrancourt, and T. Pun. Boredom, engagement and anxiety as indicators for adaptation to difficulty in games. *Proceedings of the 12th international conference on Entertainment and media in the ubiquitous era*, 7(1), 2008.
- [21] N. Nourbakhsh, F. Chen, Y. Wang, and RA. Calvo. Detecting users' cognitive load by galvanic skin response with affective interference. *ACM Transactions on Interactive Intelligent Systems*, 7(3), 2017.
- [22] M. Holmqvist, K. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer. *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press, 2011.
- [23] S. Chen and J. Epps. Automatic classification of eye activity for cognitive load measurement with emotion interference. *Computer Methods and Programs in Biomedicine*, 110(2), 2013.

Protocolo para el entrenamiento cognitivo de la memoria semántica mediante tDCS

T. Pascual¹, D. I. Gracia^{2,3}, V. Quiles^{2,3}, M. Ortiz^{2,3}, J.M. Azorín^{2,3,4}, E. Iáñez^{2,3}

¹ Universidad Miguel Hernández de Elche, Avda. de la Universidad de Elche, S/N, 03202, Elche, España; teresa.pascual@alu.umh.es

² Brain-Machine Interface System Lab, Universidad Miguel Hernández de Elche, Avda. de la Universidad de Elche, S/N, 03202, Elche, España; {dgracia, vquiles, mortiz, jm.azorin, eianez}@umh.es

³ Instituto de Investigación en Ingeniería de Elche - I3E, Universidad Miguel Hernández de Elche, Avda. de la Universidad de Elche, S/N, 03202, Elche, España

⁴ Valencian Graduate School and Research Network of Artificial Intelligence – valgrAI, Camino de Vera, S/N, 46022, Valencia, España

Resumen

La estimulación directa transcraneal por corriente continua (tDCS) en áreas implicadas en la memoria semántica supone beneficios contribuyendo a la recuperación del procesamiento léxico-semántico que se refleja en asociaciones inconscientes, evocando conceptos, símbolos y palabras. Por ello, se está empleando como herramienta de neurorrehabilitación para diferentes tipos de patologías como son las afasias o el alzhéimer. El siguiente protocolo es una propuesta para aplicar tDCS en el lóbulo temporal del hemisferio izquierdo y observar los efectos en la recuperación de las asociaciones semánticas mediante un test de asociación implícita (TAI) que mide los prejuicios. Se plantea realizar el test antes y después de la estimulación tanto en inglés como en castellano esperando que el tiempo de reacción ante las asociaciones presentadas sea menor en el grupo que reciba la estimulación respecto al grupo control.

1. Introducción

La tDCS es un método de estimulación no invasivo e indoloro el cual consiste en aplicar corrientes eléctricas sobre el cuero cabelludo con el objetivo de estimular áreas específicas del cerebro [1]. Esta técnica ha sido utilizada en varios campos de investigación, entre ellos en el aprendizaje y la plasticidad cerebral como demuestran los estudios de Reis et al., 2009, o en campos de neurorrehabilitación como también confirman los estudios de Lefaucheur et al., 2017 [2][3].

La tDCS se aplica mediante la inducción de una corriente continua de bajo voltaje en diferentes zonas del cuero cabelludo, generalmente en sesiones cortas que tienen una duración de 15 a 30 minutos. Esta corriente actúa modulando la excitabilidad neural en las capas más externas de la corteza cerebral, actuando con un excitador (ánodo) o inhibidor (cátodo) de la actividad neuronal en áreas específicas. Esta excitación se consigue por medio del uso de electrodos que se colocan en las zonas que se quiere trabajar para conseguir los resultados esperados. Softwares como Simnibs permiten simular las corrientes aplicadas con diferentes combinaciones de electrodos,

tamaños e intensidades para comprobar que éstas están actuando sobre las áreas deseadas [4].

Esto resulta en una técnica prometedora en investigación y tratamiento clínico como demuestra Bolognini et al., 2011, en su estudio en pacientes post-ictus [5]. En concreto, en este estudio se propone un protocolo para trabajar las zonas de la corteza de Brodmann [6] relacionadas con la memoria semántica entendida, como define Tulvin (1972) como la parte del sistema de memoria a largo plazo que almacena y recupera conocimiento general sobre el mundo, conceptos, significados de palabras y hechos abstractos no estando ligada a experiencias personales [7].

A nivel anatómico, como indica la investigación de Eickhoff et al., 2009, donde mediante técnicas de neuroimagen proporcionan una descripción detallada de las áreas cerebrales involucradas en la memoria semántica concluyendo que se encuentra en diferentes áreas distribuidas en el cerebro, concretamente abarcando zonas del lóbulo temporal y frontal [8]. Algunas de las regiones clave encontradas son:

- Área temporal lateral (ATL): involucrada en representación y procesamiento semántico de palabras y conceptos. Donde se encuentra el área de Wernicke, relacionada con la comprensión del lenguaje hablado y escrito [9].
- Área parietal inferior: relacionada con la recuperación semántica y el procesamiento de información. Donde se encuentra el giro angular [10].
- Corteza prefrontal inferior: asociada al procesamiento del lenguaje, el uso y la recuperación del conocimiento semántico.
- Corteza cingulada anterior: implicada en la supervisión de la memoria y el monitoreo del acceso a la información.

Gracias a este análisis con la herramienta “BrainMap” [11] y al estudio de Ardila et al., 2016, donde se menciona la región perisilviana del hemisferio izquierdo comprendido

como un sistema léxico semántico relacionado con el área de Wernicke, que incluye el área de Wernicke central (reconocimiento de palabra) y un área de Wernicke extendida (asociaciones lingüísticas), se proponen las zonas a estimular [12].

Para poder evaluar la mejora del rendimiento de la memoria semántica se empleará un test de asociación implícita (TAI), método de medición de actitudes y preferencias validado por los científicos Anthony G. Greenwald y Mahzarin R. Banaji en la década de los 90, resultando un instrumento de medición implícito, que está relacionados con pensamientos o acciones con objetos sociales [13].

En el presente artículo se propone un protocolo para medir la mejora de la memoria implícita al realizar el test TAI tanto en castellano como en inglés antes y después de la estimulación. Los experimentos contendrán dos grupos para poder comparar un grupo control que no recibirá estimulación (sham) frente a otro que sí recibirá estimulación (tDCS). Se medirán parámetros como el tiempo de reacción al realizar el test, así como la cantidad de fallos.

2. Material y métodos

2.1. Equipo de estimulación starstim32

Para aplicar la estimulación tDCS se propone el dispositivo Starstim32. Mediante su software NIC permite optimizar y personalizar los protocolos de tDCS seleccionando el montaje de electrodos, tiempo de estimulación, intensidad en cada electrodo y si son ánodos o cátodos. A su vez también permite aplicar los protocolos tanto de estimulación tDCS como sham (Figura 1).

Para aplicar la estimulación se utilizarán electrodos de 1 cm de radio y una intensidad máxima de 0.5 mA por electrodo, sin superar en ningún caso una densidad de corriente de 0.1583 mA/cm², que aún queda muy lejos de los valores límites de seguridad. Estos electrodos son más localizados frente a los electrodos de otros equipos que son de tipo esponja y abarcan mucha más zona de forma global, pero

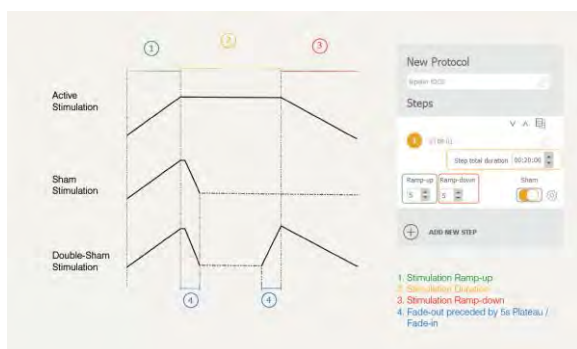


Figura 1. Captura del software del Starstim 32 con los métodos de estimulación tDCS y sham. En ambos casos, una rampa de 3 segundos al inicio permite alcanzar la intensidad seleccionada y de igual forma una rampa de 3 segundos de bajada al finalizar el tiempo de estimulación. La diferencia es que la estimulación tDCS mantendrá la intensidad durante el tiempo de estimulación mientras que el grupo sham volverá a valor 0 durante ese periodo.

no focalizada. Los ánodos serán los electrodos por los que se aplica la corriente y aumentan la excitabilidad de las neuronas; mientras que los cátodos serán aquellos electrodos por los que se inhibiría. Es necesario poner gel conductor en los electrodos empleados para asegurar una impedancia reducida y que la estimulación pueda aplicarse.

2.2. TAI/software

Para evaluar las asociaciones implícitas se utilizó el TAI de la Figura 2. Los resultados de los tests implícitos revelan qué patrones y preferencias existen subconscientemente en la mente de los consumidores que determinan su comportamiento. Para la elección de estas palabras se recurre a la base de datos de Stadthagen-Gonzalez et al., 2018, quién ha realizado una lista de palabras clasificadas de formas científica por su valencia emocional [14].

El test trata de medir la ruta rápida o implícita, por lo que las respuestas contaminadas por la mente consciente se suprimen y eliminan. Durante la prueba los participantes asocian palabras o imágenes con diferentes categorías y el tiempo de respuesta se utiliza para inferir las preferencias subyacentes.

En primer lugar, los participantes reciben una serie de instrucciones claras sobre cómo realizar la prueba, que consistirá en categorizar la información del monitor que presentará secuencialmente palabras o imágenes. Esto se hará presionando dos teclas del teclado, “e” e “i”, para producir así una categorización rápida de cada estímulo.

También habrá que tener en cuenta si las respuestas son congruentes (relacionados entre sí) o incongruentes (no relacionados) con los prejuicios de la persona. Además, se puede realizar en diferentes idiomas, en este caso se propone realizar el TAI en castellano (lengua materna) y en inglés (a modo de tarea distractora). El TAI mide tanto la rapidez como la precisión de las respuestas. Se presupone que, a respuestas más rápidas y precisas, asociaciones más fuertes y a respuestas más lentas y menos precisas, asociaciones menos fuertes.

La prueba se estructura en bloques de preguntas, donde se evalúan diferentes asociaciones. En este caso se propone un test de 5 etapas con las siguientes categorizaciones:

1. Simple de conceptos (32 ítems): mariposa vs araña
2. Simple de atributos (32 ítems): Bueno vs Malo
3. Combinada (64 ítems): mariposa o bueno vs araña o malo

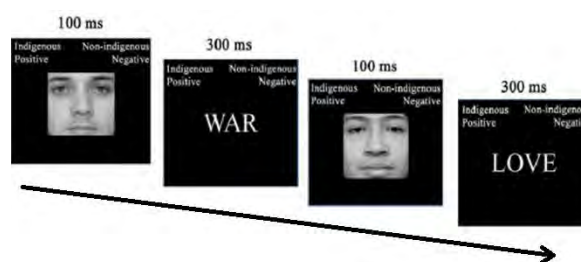


Figura 2. Esquema experimental del TAI que se refiere a todas las combinaciones posibles del bloque de asociación con ejemplo “indígena-positivo” “no indígena- negativo” [15].

4. Simple de conceptos (32 ítems): araña vs mariposa
5. Combinada (64 ítems): mariposa o malo vs araña o bueno

El emplear un test de 5 etapas sirve tanto para minimizar el efecto orden en los resultados reduciendo de esta manera los sesgos en las asociaciones, como para reducir el error causal, así si un participante comete un error en un bloque, es menos probable que ese error individual tenga un impacto significativo en la medida global si se promedian los resultados en varios bloques.

Para la obtención de los resultados, el mismo TAI genera una tabla Excel final con los diferentes tiempos de respuesta y errores para cada una de las preguntas, obteniendo 224 ítems por usuario y prueba.

2.3. Procedimiento experimental

La realización de la prueba experimental comienza con la preparación de los diferentes softwares con el fin de no ralentizar la prueba en sus diferentes etapas: el software NIC para la estimulación tDCS y el test TAI para la prueba cognitiva. El usuario será instrumentado con el equipo de estimulación y se situará frente a un ordenador para poder realizar con comodidad el test TAI.

1. La fase preestimulación contemplará: en primer lugar, el TAI en castellano y posteriormente en inglés, con una duración de unos 6 minutos cada uno (Figura 3).
2. A continuación, en la fase de tDCS o sham se aplicará el protocolo durante el mismo tiempo a ambos grupos. El grupo tDCS recibirá la estimulación indicada durante 20 minutos mientras que el grupo sham no la recibirá para actuar como grupo de control. El diseño del montaje es un punto clave que se comenta en el siguiente apartado.
3. Por último, la fase postestimulación, se volverán a realizar los TAI, primero en castellano y luego en inglés.

Por tanto, la duración total del experimento se estima 1 hora (15 minutos de instrumentación y 45 minutos de experimentación). Dicho protocolo ha sido simulado con un usuario para medir el tiempo de instrumentación, confirmar la duración de cada fase, y asegurar una presentación adecuada del TAI. En la Figura 4 puede verse como ejemplo el setup preparado para un usuario durante esta prueba experimental simulada.

2.4. Criterios para el diseño del montaje de estimulación en el área de la memoria semántica

El protocolo de estimulación se centra en el área semántica, que involucra el lóbulo temporal del cerebro. Se eligió estimular el área temporal del hemisferio izquierdo, crucial para la memoria, el lenguaje, la memoria auditiva y el reconocimiento facial. Por ello, siguiendo el sistema 10-10 de EEG se propondrán diferentes zonas en las que se puede centrar la tDCS para influir en funciones lingüísticas [16]:

- Área de Wernicke: Ubicada en el lóbulo temporal corresponde a las áreas 22 (AB22), 39 (AB39) y 40 (AB40) de Brodmann. Se encuentra con mayor frecuencia en el hemisferio izquierdo del cerebro, pero no siempre.

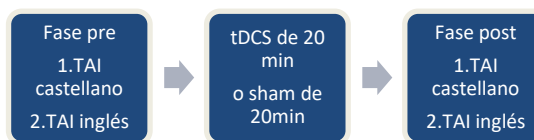


Figura 4. Secuencias de tareas realizadas en cada trial. En la fase pre se realizará el TAI en castellano y luego en inglés. Luego unos 20 minutos de estimulación o sham. Por último, fase postestimulación, de nuevo TAI en castellano y luego en inglés.

- Giro Angular: Ubicado en el lóbulo parietal (AB39). Ayuda a decodificar diferentes elementos de información sensorial para comprender el lenguaje. Área terciaria que juega un importante rol en la integración multimodal (visual, auditiva y táctil), en el procesamiento semántico, cálculo, lectura y escritura.

El software Simnibs permite evaluar diferentes configuraciones de estimulación cerebral. En lugar de utilizar solo un ánodo y un cátodo, se está trabajando en un montaje con múltiples ánodos y cátodos alrededor del área de interés. Esto ayuda a focalizar la estimulación en esa zona específica y afectar áreas cercanas. El tiempo de estimulación elegido es de 20 minutos, basado en la literatura científica.

3. Discusión

El protocolo presentado permite obtener datos en diferentes variables, que, junto con el diseño de un montaje adecuado de estimulación, permitirá validar diferentes hipótesis. Hay que tener en cuenta que, debido a las diferencias anatómicas de los usuarios, la estimulación sobre el área de la memoria semántica puede no ser todo lo exacta que sea necesario. Un objetivo a futuro consideraría realizar una resonancia magnética funcional a cada usuario

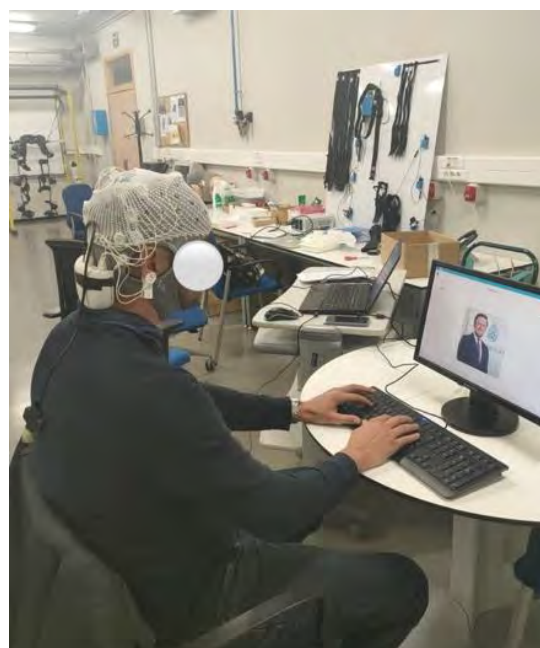


Figura 3. Setup de prueba experimental con el usuario llevando el equipo de estimulación y sentado frente a un ordenador para realizar el test TAI.

mientras realiza tareas relacionadas con la memoria semántica para determinar el área cerebral de cada usuario. Y mediante la simulación de su propia resonancia poder afinar el montaje de electrodos para enfocar la estimulación únicamente en el área de interés. No obstante, en este caso, el montaje buscado busca estimular con varios ánodos el área de interés rodeándola con varios cátodos para contenerla únicamente en el área de interés.

Todos los usuarios realizarán los mismos tests antes y después de la estimulación para permitir una comparación intrausuario. La hipótesis es que el grupo de estimulación mostrará una mejora relativa mayor que el grupo de control. Se aplicará tDCS en áreas específicas para reducir los tiempos de respuesta en asociaciones de palabras, especialmente en el grupo de estimulación. Se comparará el rendimiento entre los tests de inglés y castellano, esperando mayores tiempos de respuesta y errores en inglés, especialmente en el grupo de estimulación. Se comenzará con una muestra de 40 usuarios para conclusiones iniciales y se planea aumentarla para obtener resultados más robustos.

4. Conclusiones

En el presente estudio se ha propuesto un protocolo para trabajar el rendimiento de la memoria semántica de manera no invasiva mediante el uso de la tDCS y su evaluación mediante un TAI. Se diseñará un montaje de estimulación evaluando con Simnibs que se enfoque en el área de interés relacionado con la memoria semántica y se realizarán pruebas iniciales con al menos 40 personas.

Esto permitirá validar las diferentes hipótesis planteadas para establecer que la tDCS afecta de manera positiva el área de la memoria y que por tanto puede mejorar los resultados en los test TAI.

En una primera prueba piloto un usuario ha obtenido tiempos de respuesta similares en los test en inglés y castellano y se observa una ligera reducción tanto del tiempo de reacción como de los errores, pero con únicamente un usuario aun no es posible obtener datos concluyentes y será necesario comprobar el efecto de la estimulación tDCS con el montaje debidamente finalizado.

Referencias

[1] Horvath, J. C., Carter, O., & Forte, J. D. (2014). Transcranial direct current stimulation: five important issues we aren't discussing (but probably should be). *Frontiers in Systems Neuroscience*, 8. <https://doi.org/10.3389/fnsys.2014.00002>

[2] Reis, J., Schambra, H. M., Cohen, L. G., Buch, E. R., Fritsch, B., Zarahn, E., Celnik, P., & Krakauer, J. W. (2009). Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. *Proceedings of the National Academy of Sciences of the United States of America*, 106(5), 1590-1595. <https://doi.org/10.1073/pnas.0805413106>

[3] Lefaucheur, J. P., Antal, A., Ayache, S. S., Benninger, D., Brunelin, J., Cogiamanian, F., Cotelli, M., De Ridder, D., Ferrucci, R., Langguth, B., Marangolo, P., Mylius, V., Nitsche, M. A., Padberg, F., Palm, U., Poulet, E., Priori, A., Rossi, S., Sackellmann, M., . . . Paulus, W. (2017). Evidence-based guidelines on the therapeutic use of Transcranial Direct Current Stimulation (TDCS). *Clinical*

Neurophysiology, 128(1), 56-92. <https://doi.org/10.1016/j.clinph.2016.10.087>

[4] Thielscher, A., Antunes, A. and Saturnino, G.B. (2015), Field modeling for transcranial magnetic stimulation: a useful tool to understand the physiological effects of TMS? IEEE EMBS 2015, Milano, Italy

[5] Bolognini, N., Vallar, G., Casati, C., Latif, L. A., El-Nazer, R., Williams, J., Banco, E., Macea, D. D., Tesio, L., Chessa, C., & Fregni, F. (2011). Neurophysiological and behavioral effects of TDCS combined with Constraint-Induced Movement therapy in Poststroke patients. *Neurorehabilitation and Neural Repair*, 25(9), 819-829. <https://doi.org/10.1177/1545968311411056>

[6] Judaš, M., Ceganec, M., & Sedmak, G. (2012). Brodmann's map of the human cerebral cortex — or Brodmann's Maps? *Translational Neuroscience*, 3(1), 67-74. <https://doi.org/10.2478/s13380-012-0009-x>

[7] Tulving E. (1972). Episodic and Semantic Memory. In E. Tulving, & W. Donaldson (Eds.), *Organization of Memory* (Pp. 381-403). Cambridge, MA Academic Press. - References - Scientific Research Publishing. (s. f.). [https://www.scirp.org/\(S\(lz5mqp453edsnp55rrgjt55.\)\)reference/referencespapers.aspx?referenceid=2919588](https://www.scirp.org/(S(lz5mqp453edsnp55rrgjt55.))reference/referencespapers.aspx?referenceid=2919588)

[8] Eickhoff, S. B., Laird, A. R., Fox, P. T., Bzdok, D., & Hensel, L. (2014). Functional segregation of the human dorsomedial prefrontal cortex. *Cerebral Cortex*, 26(1), 304-321. <https://doi.org/10.1093/cercor/bhu250>

[9] Diéguez-Vide F, Peña-Casanova J. Cerebro y lenguaje: Sintomatología neurolingüística. Madrid: Editorial Medica Panamericana, 2012.

[10] Vivas, Leticia, & García García, Ricardo. (2010). Contribución desde la neuropsicología al estudio de la memoria semántica. *Avances en Psicología Latinoamericana*, 28(2), 239-250. Retrieved August 21, 2023, from http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1794-47242010000200008&lng=en&tlng=es.

[11] *BrainMap Taxonomy | Stimulus Type | Words*. (s. f.). http://www.brainmap.org/taxonomy/stimulus_type/Words.html

[12] Ardila, A., Bernal, B., & Rosselli, M. (2016). Área cerebral del lenguaje: una reconsideración funcional. *Revista De Neurología*. <https://doi.org/10.33588/rn.6203.2015286>

[13] Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4-27. <https://doi.org/10.1037/0033-295x.102.1.4>

[14] Stadthagen-Gonzalez, H., Ferré, P., Pérez-Sánchez, M. Á., Imbault, C., & Hinojosa, J. A. (2017). Norms for 10,491 Spanish words for five discrete emotions: happiness, disgust, anger, fear, and sadness. *Behavior Research Methods*, 50(5), 1943-1952. <https://doi.org/10.3758/s13428-017-0962-y>

[15] Infante, A. L., Hurtado, Martín, S., & Ibanez, A. (2009). Neurociencia social en Sudamérica: la cognición sujeta a claves emocionales, sesgos raciales, procesos. . *ResearchGate*. https://www.researchgate.net/publication/299766572_Neurociencia_Social_en_Sudamerica_La_cognicion_sujeta_a_claves_emocionales_sesgos_raciales_procesos_empaticos_efectos_contextuales_y_toma_de_decisiones

Self-Supervised Audio Spectrogram Transformer para el Reconocimiento de Emociones en Señales de Voz: Un Enfoque de Entrenamiento Semi-Supervisado

Oscar Valls ¹, Rocío del Amor ¹, Adrián Colomer ¹, Valery Naranjo ¹

¹ Instituto Universitario de Investigación en Tecnología Centrada en el Ser Humano (H-TECH)
 Universidad Politécnica de Valencia (UPV)
 {osvallo, madeam2, adcogra, vnaranjo}@htech.upv.es.

Resumen

La detección de emociones en señales de voz es un campo de las tecnologías de reconocimiento automático del discurso. Conforme los algoritmos de aprendizaje profundo han ido evolucionando, los servicios de reconocimiento de voz son cada vez más precisos. La señal de voz puede informar acerca del estado emocional del hablante. Variaciones en el tono, cadencia del habla o volumen de la voz, son algunas de las características que transmiten parte del contenido emocional del mensaje transmitido. Los sistemas de reconocimiento de emociones abren un amplio campo de aplicaciones en el ámbito de la psicología ya que pueden facilitar el reconocimiento del estado emocional del paciente, así como asistir en decisiones, diagnósticos y monitorización. Las arquitecturas basadas en transformers, gracias al entrenamiento semi-supervisado, pueden emplearse para múltiples tareas de reconocimiento del discurso, en este caso, el reconocimiento de emociones. En este artículo se presentan las técnicas empleadas en el entrenamiento supervisado del Self-Supervised Audio Spectrogram Transformer para la detección de emociones empleando la base de datos IEMOCAP. Gracias a realizar un *fine-tuning* parcial del modelo se ha llegado a obtener un porcentaje de *accuracy* (68.2%) que supera al del estado del arte (60.5%).

1. Introducción

La salud mental es uno de los grandes problemas a los que se enfrenta la sociedad, según la Organización Mundial de la Salud (OMS) [1], en 2019, 970 millones de personas sufrieron algún trastorno mental, entre los más frecuentes se encuentran, el trastorno por ansiedad, que sufrieron 300 millones de personas, y el trastorno por depresión, que sufrieron 280 millones de personas, siendo ambos trastornos especialmente frecuentes entre mujeres y gente joven. Además, el suicidio supuso la causa del 1,3% de las muertes a nivel mundial, es la cuarta causa de muerte en jóvenes de entre 15 y 29 años, produciéndose el 77% de ellos en países con bajos recursos. Todo ello pone de manifiesto la necesidad de una optimización y crecimiento del sector de la psicología clínica, dando a los terapeutas la posibilidad de prestar atención al mayor número posible de pacientes mediante la automatización y digitalización de ciertos procesos de la terapia.

La evolución de los algoritmos de aprendizaje profundo, y en especial las arquitecturas basadas en *transformers* [2], han permitido el desarrollo de tecnologías de reconocimiento del discurso, llegando a realizar con mucha eficacia tareas como la transcripción, el reconocimiento

del hablante, y en este caso el reconocimiento del contenido emocional en la señal de voz. Este tipo de sistemas han demostrado también ser eficaces en aplicaciones relacionadas con el campo de la psicología, siendo capaces de asistir en la detección de pacientes con depresión y la remisión de la propia enfermedad [3], la predicción de los marcadores fisiológicos de estrés [4], la detección de las emociones relacionadas con la ansiedad [5] y el suicidio [6].

Pese al avance que supone el uso de las redes neuronales para el reconocimiento de emociones, el desarrollo de estas técnicas viene limitado por la falta de bases de datos anotadas de un tamaño suficientemente grande para aprovechar la complejidad y dimensiones de las arquitecturas empleadas hoy en día, que son principalmente las basadas en *transformers* [2], para las tareas relacionadas con el reconocimiento del discurso. Esto da lugar a que los modelos, que funcionan bien en tareas como la transcripción, para la cual si que existen bases de datos de tamaño suficiente, tiendan a no generalizar bien en tareas de reconocimiento de emociones y cuando se emplean los modelos en condiciones de uso reales. Para abordar este problema, las técnicas de entrenamiento semi-supervisado resultan de gran utilidad, ya que en el preentrenamiento el modelo aprende a generar representaciones del audio que recibe a la entrada, empleando para ello una base de datos no anotada, por lo que puede ser de gran tamaño y de uso generalista. Posteriormente, únicamente es necesario realizar un *fine-tuning* del modelo con una base de datos de menor tamaño para entrenarlo en el reconocimiento de emociones. En este artículo se propone una mejora en el proceso de *fine-tuning* del modelo SSAST [7] para la clasificación de emociones, mediante el cual se obtienen unos resultados de *accuracy* superiores a los presentados en el artículo del SSAST y comparables a otros modelos del estado del arte.

2. Materiales

2.1. Base de datos IEMOCAP

Interactive Emotional Dyadic Motion Capture (IEMOCAP) [8] es la base de datos de emociones más extendida en el estado del arte. Contiene 7380 muestras que forman unas 9 horas y media de audio que consisten en diálogos guionizados e improvisados por 10 hablantes distintos, 5 hombres y 5 mujeres. Tal y como se observa en la figura 1, la base de datos contiene un total de 6 emo-

ciones (*neutral, frustración, enfado, tristeza, felicidad y excitación*). Aunque, como se observa en la figura 1, no todas las emociones representadas por igual y la base de datos queda desbalanceada. Para mitigar el problema, se toman únicamente las muestras de 4 emociones, que son las siguientes: enfado, neutral, tristeza y felicidad.

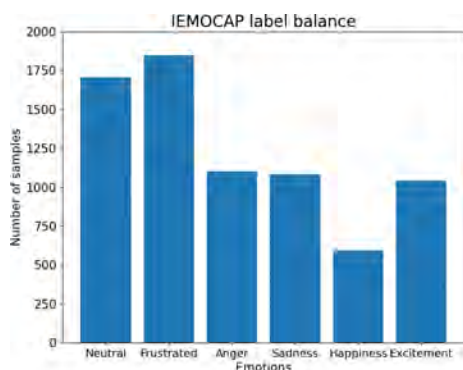


Figura 1: En el eje X se muestran las distintas emociones representadas en la base de datos. En el eje Y, el número de muestras de cada emoción.

Las características de los audios son las siguientes:

- Frecuencia de muestreo = 16 kHz
- Duración media de los audios = 4,5 segundos
- Duración del audio más corto = 0.5 segundos
- Duración del audio más largo = 34,138 segundos

3. Métodos

3.1 Self-Supervised Audio Spectrogram Transformer. SSAST

Recientemente, los *transformers* han demostrado ser más eficaces que las arquitecturas que se venían empleando tradicionalmente en una gran variedad de tareas relacionadas con el procesamiento de imágenes y de audio. En el caso del procesamiento de imágenes, el *Vision Transformer* (ViT) [9] puede obtener mejores resultados que las redes neuronales convolucionales (CNNs) [10]. En el procesamiento de audio, el *Audio Spectrogram Transformer* (AST) [11] y el *Keyword Transformer* (KWT) [12], que son modelos inspirados en la arquitectura del ViT, pero que funcionan con el espectrograma del audio y los MFCCs respectivamente, también llegaron a resultados a la altura del estado del arte en tareas como la clasificación de escenas de audio o la detección de palabras clave. Un aspecto crítico, que diferencia en gran medida a los ViT de las CNN, y que a su vez, es el principal inconveniente de los *transformers*, es que la cantidad necesaria de datos anotados para entrenarlos y que alcanzan su máximo rendimiento es mucho mayor que la necesaria para entrenar de forma óptima una CNN [9]. De esta problemática surge el modelo SSAST [7], que es un modelo AST entrenado de forma semi-supervisada, gracias a esto, se minimiza la cantidad de datos anotados necesarios para entrenarlo. El preentrenamiento se

realiza empleando la combinación de las bases de datos: *AudioSet-2M* [13], compuesta por 2 millones de muestras de eventos sonoros de 10 segundos. Esta base de datos está destinada a entrenar modelos en tareas de audio no relacionadas con el discurso y *Librispeech* [14], compuesta por 960 horas de grabaciones de audiolibros, leídos por más de 1000 hablantes distintos. Es la base de datos más extendida para entrenar modelos en tareas relacionadas con reconocimiento del discurso.

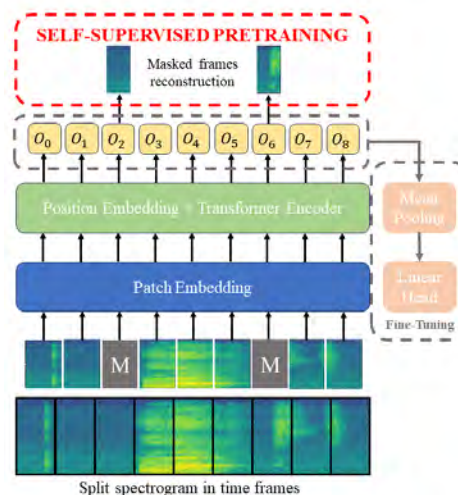


Figura 2: Estructura del modelo SSAST. A la entrada, se enmascaran (*M*) algunas ventanas para el preentrenamiento. A la salida se obtiene una secuencia de embeddings $[O_0, O_1, \dots, O_n]$, que representan el audio de entrada.

La base de datos resultante de combinar ambas está compuesta de 2,234 millones de muestras. 1,953 millones pertenecen a *AudioSet* y 281k muestras pertenecen a *Librispeech*. El modelo SSAST funciona con el espectrograma de la señal de audio, por lo que requerirá de un preprocesamiento para convertir de señal 1D a espectrograma. Dicho espectrograma se divide en parches de tiempo y frecuencia o en ventanas temporales. En el caso de las tareas relacionadas con la señal de voz, el modelo funciona mejor dividiendo el espectrograma en ventanas temporales y tratando el espectrograma como una secuencia temporal de esas ventanas [7]. Cada ventana se proyecta linealmente en un espacio latente, la secuencia de dichas proyecciones se introducen como entrada en los bloques del *transformer*. En este caso, los mecanismos de atención del *transformer* se aplican únicamente en la dimensión del tiempo, relacionando las ventanas entre ellas y obteniendo así una secuencia de *embeddings* a la salida del *transformer* como representación del audio a la entrada. En la figura 2 se observa que la tarea pretexto del preentrenamiento no supervisado del SSAST consiste en enmascarar algunos de los parches o ventanas aleatoriamente a la entrada del *transformer* y luego reconstruir los parches o ventanas enmascarados a partir de las representaciones que el *transformer* entrega de estos a su salida. Empleando la pérdida de reconstrucción de dichos parches o ventanas se consi-

que entrenar el SSAST de forma no supervisada para generar un vector de representación $[O_0, O_1, \dots, O_n]$. del audio que recibe a la entrada. Dicho vector es el que se emplea para realizar el entrenamiento supervisado para realizar una tarea concreta. Hasta ahora, el SSAST únicamente ha sido probado como extractor de características para el reconocimiento de emociones, es decir, se emplea un modelo preentrenado para obtener las representaciones de los audios a la entrada y con esas representaciones se entrena un clasificador para realizar el reconocimiento de las emociones. En este artículo se propone un *fine-tuning* parcial del SSAST, es decir, entrenando algunos de los últimos bloques del *transformer* junto con el clasificador para realizar las predicciones.

3.2. Fine-tuning parcial del SSAST

Las versiones de modelos SSAST preentrenados con las que se han realizado los experimentos son:

- *Tiny model*: 12 bloques con 3 cabezas de atención y el tamaño de los *embeddings* es de 192. Tiene 6 millones de parámetros.
- *Base model*: 12 bloques con 12 cabezas de atención y el tamaño de los *embeddings* es de 768. Tiene 89 millones de parámetros.

Existe una versión con un tamaño intermedio entre el *tiny* y el *base* pero ha sido descartada para simplificar el estudio y por no estar publicados los pesos del modelo preentrenado. En primer lugar, se ha realizado un estudio de ablación para determinar el número de capas entrenables óptimo para que el modelo sea capaz de clasificar mejor las emociones. Una vez realizado el estudio, se ha evaluado el mejor modelo mediante una *4-fold cross validation*. La métrica de evaluación en el estudio de ablación es la *accuracy* del conjunto de test y en la *cross validation* también es la *accuracy* obtenida en el conjunto de test, pero en este caso se promedian las *accuracies* de los modelos entrenados con los distintos *folds*. En los experimentos de ablación se han probado también distintas combinaciones de hiperparámetros. Para ello, se mantuvieron la partición de entrenamiento, validación y test; es decir, no se realizó validación cruzada para reducir el número de entrenamientos. Tras el primer experimento, se ha tomado el número de capas entrenables óptimo, con la combinación de hiperparámetros que haya obtenido mejor resultado de *accuracy* en el conjunto de test, para realizar la validación cruzada y así comprobar la robustez del modelo frente a variaciones en las particiones de la base de datos y se replica el marco experimental seguido para la obtención de los resultados presentados en el paper del SSAST [7].

4. Resultados y discusión

En la Tabla 1 se muestran los resultados de *accuracy* obtenidos por cada modelo con distinto número de capas congeladas. La mejor *accuracy* se obtiene para los modelos: *Tiny* con 8 capas entrenables (68.9% acc) y *Base* con 8 capas entrenables (70.3% acc). Los resultados del paper del SSAST [7] se obtienen de mediante el

benchmark SUPERB: Speech processing Universal Performance Benchmark) [15], según el cual, el *fine-tuning* se realiza con el modelo SSAST preentrenado y completamente congelado, entrenando únicamente la capa *fully-connected* de clasificación. Se emplean las mismas cuatro clases de la base de datos IEMOCAP y se realiza una validación cruzada. Con el fin de que los resultados sean comparables, se realiza la validación cruzada, cuyos resultados aparecen en la Tabla 2, donde se puede observar que el método de *fine-tuning* parcial obtiene para ambos modelos mejores resultados en clasificación que los obtenidos con SUPERB. Estos resultados ponen de manifiesto que el entrenamiento del SSAST con 8 de 12 capas congeladas, ha permitido al modelo aprender a generar las características óptimas para detectar emociones en señales de voz.

Nº de capas congeladas	<i>Tiny model</i>	<i>Base model</i>
12	57.4 %	57.7 %
11	56.5 %	60.5 %
10	58.2 %	63.2 %
9	68.1 %	69.8 %
8	68.9 %	70.3 %
7	65.8 %	68.8 %
6	66.0 %	68.5 %
5	67.8 %	69.7 %

Tabla 1: Resultados del estudio de ablación expresados en términos de *accuracy* en el conjunto de test.

Modelo	acc	f1-score
<i>Tiny (Fine-tuned)</i>	67.5 % ± 1,433	0.665
<i>Tiny (SUPERB)</i>	55.7 %	-
<i>Base (Fine-tuned)</i>	68.2 % ± 0,446	0.668
<i>Base (SUPERB)</i>	60.5 %	-
<i>Wav2Vec 2.0 [16]</i>	63.4 %	-
<i>HuBERT [17]</i>	64.9 %	-

Tabla 2: Resultados de la validación cruzada en términos de *accuracy* y *f1-score* en el conjunto de test comparados con resultados de *accuracy* de otros modelos del estado del arte.

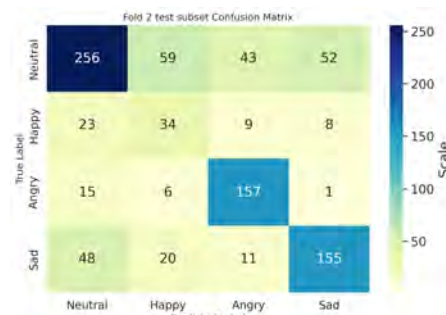


Figura 3: Matriz de confusión del conjunto de test del modelo *tiny* entrenado con el fold 2. En esta partición se ha obtenido el peor resultado de *accuracy*.

Para analizar más en detalle los resultados de *accuracy* del modelo SSAST *fine-tuned*, se realiza un estudio de la confusión. Véase Figura 3. La emoción menos representada (felicidad) es la que el modelo más confunde, y en la mayoría de casos con la clase neutral, que es la mayormente representada. El modelo reconoce mejor el enfado y la tristeza, pero aquellas muestras que no reconoce de dichas emociones también las clasifica como neutral. Por lo que se considera que el desbalanceo de los datos es un factor limitante a la hora de obtener mejores resultados en clasificación de emociones.

5. Conclusiones

En este artículo se han explorado las posibilidades del SSAST en la detección de emociones en señales de voz. El preentrenamiento ha dotado al modelo de la capacidad de generar representaciones del audio que recibe, y dichas representaciones pueden ser empleadas para múltiples tareas. Con una cantidad de datos no necesariamente grande y realizando un *fine-tuning* parcial, es decir, entrenando las últimas capas junto con el clasificador, se consigue optimizar la generación de dichas representaciones para obtener mejores resultados en el reconocimiento de emociones. No obstante, los resultados de clasificación no son del todo satisfactorios, ya que hay mucha confusión en la emoción menos representada. Esto pone de manifiesto la necesidad de bases de datos de suficiente tamaño y balance para esta tarea, puesto que si bien existen modelos de reconocimiento del discurso que rinden muy bien en otras tareas como la transcripción o la detección de palabras clave, esto es porque sí que existen bases de datos adecuadas para ello. En el reconocimiento de emociones aún hay mucho margen de mejora en este aspecto.

Agradecimientos

Este trabajo ha sido realizado gracias a los fondos recibidos del Ministerio de Ciencia e Innovación por el proyecto DIPSY (TED2021-131401B-C21) y los fondos de la Generalitat Valenciana (GVA) por el proyecto CONTACTS2 (CIPROM/2022/20)

Referencias

- [1] Mental disorders. 2022. Available at <https://www.who.int/news-room/fact-sheets/detail/mental-disorders>.
- [2] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2023.
- [3] L. Hansen, Y. P. Zhang, D. Wolf, K. Sechidis, N. Lagedgaard, and R. Fusaroli. A generalizable speech emotion recognition model reveals depression and remission. *Acta Psychiatr Scand*, 145(2):186–199, Feb 2022. Epub 2021 Dec 14.
- [4] Alice Baird, Andreas Triantafyllopoulos, Sandra Zänkert, Sandra Ottl, Lukas Christ, Lukas Stappen, Julian Konzok, Sarah Sturmbauer, Eva-Maria Meßner, Brigitte M. Kudielka, Nicolas Rohleder, Harald Baumeister, and Björn W. Schuller. An evaluation of speech-based recognition of emotional and physiological markers of stress. *Frontiers in Computer Science*, 3, 2021.
- [5] Rathi Adarshi Rammohan, Jeevan Medikonda, and Dan Issac Pothiyil. Speech signal-based modelling of basic emotions to analyse compound emotion: Anxiety. In *2020 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER)*, pages 218–223, 2020.
- [6] John Gideon, Heather Schatten, Melvin Mcinnis, and Emily Mower Provost. Emotion recognition from natural phone conversations in individuals with and without recent suicidal ideation. pages 3282–3286, 09 2019.
- [7] Yuan Gong, Cheng-I Jeff Lai, Yu-An Chung, and James Glass. Ssast: Self-supervised audio spectrogram transformer, 2022.
- [8] Carlos Busso, Murtaza Bulut, Chi-Chun Lee, Abe Kazemzadeh, Emily Mower, Samuel Kim, Jeannette N. Chang, Sungbok Lee, and Shrikanth S. Narayanan. IE-MOCAP: interactive emotional dyadic motion capture database. *Language Resources and Evaluation*, 42:335–359, 2008.
- [9] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale, 2021.
- [10] Yann Lecun and Y. Bengio. Convolutional networks for images, speech, and time-series. 01 1995.
- [11] Yuan Gong, Yu-An Chung, and James Glass. Ast: Audio spectrogram transformer, 2021.
- [12] Axel Berg, Mark O'Connor, and Miguel Tairum Cruz. Keyword transformer: A self-attention model for keyword spotting. In *Interspeech 2021*. ISCA, aug 2021.
- [13] Jort F. Gemmeke, Daniel P. W. Ellis, Dylan Freedman, Aren Jansen, Wade Lawrence, R. Channing Moore, Manoj Plakal, and Marvin Ritter. Audio set: An ontology and human-labeled dataset for audio events. In *Proc. IEEE ICASSP 2017*, New Orleans, LA, 2017.
- [14] Vassil Panayotov, Guoguo Chen, Daniel Povey, and Sanjeev Khudanpur. Librispeech: An asr corpus based on public domain audio books. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5206–5210, 2015.
- [15] Shu wen Yang, Po-Han Chi, Yung-Sung Chuang, Cheng-I Jeff Lai, Kushal Lakhotia, Yist Y. Lin, Andy T. Liu, Jiatong Shi, Xuankai Chang, Guan-Ting Lin, Tzu-Hsien Huang, Wei-Cheng Tseng, Ko tik Lee, Da-Rong Liu, Zili Huang, Shuyan Dong, Shang-Wen Li, Shinji Watanabe, Abdelrahman Mohamed, and Hung yi Lee. Superb: Speech processing universal performance benchmark, 2021.
- [16] Alexei Baevski, Henry Zhou, Abdelrahman Mohamed, and Michael Auli. wav2vec 2.0: A framework for self-supervised learning of speech representations, 2020.
- [17] Wei-Ning Hsu, Benjamin Bolte, Yao-Hung Hubert Tsai, Kushal Lakhotia, Ruslan Salakhutdinov, and Abdelrahman Mohamed. Hubert: Self-supervised speech representation learning by masked prediction of hidden units, 2021.

Mejorando el Análisis de Emociones en texto mediante Llama 2

Mario Rico Ibáñez ¹, Rocío del Amor ¹, Valery Naranjo ¹

¹ Instituto Universitario de Investigación e Innovación en Tecnología Centrada en el Ser Humano, HUMAN-tech, Universitat Politècnica de València, Valencia, España

Resumen

La salud mental es crucial en la sociedad actual, y la falta de digitalización en las consultas de psicólogos clínicos en España destaca la necesidad de soluciones innovadoras. Este trabajo presenta una perspectiva novedosa sobre el análisis de emociones, utilizando el modelo Llama 2 para clasificar emociones de manera efectiva. Este proyecto se enfoca en conjuntos de datos formales, alejándonos de fuentes informales como redes sociales y clasificaciones binarias. A través de una cuidadosa exploración y expansión de datos, destacamos cómo Llama 2 supera a los enfoques convencionales. Este artículo destaca la mejora en el análisis de emociones que ofrece Llama 2 y su potencial para revolucionar el procesamiento del lenguaje natural en este campo.

1. Introducción

La salud mental se considera un pilar fundamental para el bienestar personal en España. Sin embargo, el país enfrenta un desafío significativo, ya que solo cuenta con 5.6 psicólogos clínicos por cada 100,000 habitantes, lo que representa aproximadamente una octava parte de la media europea [1]. Esta disparidad subraya la necesidad imperante de encontrar soluciones que puedan aliviar la saturación de los profesionales de la salud mental y, al mismo tiempo, brindar una atención adecuada.

En el último año, se ha sido testigo de un notable auge en el campo de la inteligencia artificial generativa, en una gran cantidad de ámbitos relacionados con el procesamiento del lenguaje natural. Esto resalta de manera inequívoca la necesidad y la pertinencia de la digitalización del sector de la psicología, aprovechando al máximo estas tecnologías de vanguardia.

Trabajos previos en clasificación de emociones [2] se enfocan principalmente en clasificación de emociones binaria, es decir, emoción positiva o negativa y utilizan bases de datos extraídas de las redes sociales, lo cual significa un problema debido a que no se adapta completamente a la realidad de nuestra situación, la mejora de la digitalización del sector de la psicología clínica. Además, los modelos utilizados en estudios[3] que si coincidían con nuestro punto de vista, clasificación multiclase y datos extraídos de fuentes como encuestas, eran significativamente más pequeños en términos de parámetros en comparación a los modelos de vanguardia.

Por tanto, se propone un nuevo punto de vista a la hora del análisis de las emociones. La utilización de *Large Language Models* que dista de las soluciones aportadas en el estado del arte en el cual se utilizan modelos como LSTM [4] BERT[5] y similares.

2. Materiales y Métodos

2.1. Bases de Datos

En un primer paso, se realizó una revisión de todos los conjuntos de datos disponibles en la web, evaluando la naturaleza y la cantidad de clases que cada uno contenía. El estudio requería datos que se adecuasen al contexto en el que el modelo iba a implementarse, por lo que se identificaron fuentes que abarcaban una amplia gama de orígenes, como Twitter, Reddit, películas, entre otros. Sin embargo, la atención se centró principalmente en aquellos provenientes de encuestas, ya que se alineaban más con los criterios necesarios para la investigación. El resumen de todas las bases de datos disponibles se encuentra en la Tabla 1.

Datos	Nº de muestras	Clases	Naturaleza
Goemotions[6]	58009	28	Reddit
Emoint[7]	3609	4	Tweets
ISEAR[8]	7666	7	Encuesta
E-Stimulus [9]	2414	7	Artificial
E-Detection	40000	13	Tweets

Tabla 1. Resumen bases de datos disponibles

Además de la necesidad de enfocarse en bases de datos con características específicas, el conjunto de datos ISEAR presentaba las clases que expertos en el campo recomendaron considerar: felicidad, tristeza, enfado, asco, miedo, culpa y vergüenza. Para mejorar los resultados, se optó por aumentar la cantidad de datos disponible. Luego de explorar diversas técnicas de aumento de datos en el procesamiento del lenguaje natural, como la sustitución de palabras por sinónimos, la inserción de ruido en las oraciones e incluso la eliminación de palabras, se llegó a la conclusión de que estas opciones podrían comprometer la calidad de los datos. Por lo tanto, se decidió recurrir a la expansión de datos mediante el uso de chatGPT. Empleando un *script* en Python junto con un *prompt* adecuado, se lograron obtener los resultados deseados. Esta técnica permitió aumentar el conjunto de datos de 7,666 muestras a 10,700, lo que contribuyó de manera significativa a la investigación.

Datos	Nº de muestras
Anger	1096
Disgust	1096
Fear	1095
Sadness	1096
Shame	1096
Joy	1094
Guilt	1093

Tabla 1. Resumen bases de datos disponibles

Posteriormente, debido a las características de los modelos de lenguaje de gran envergadura, se procedió a

combinar ISEAR con GoEmotions, seleccionando únicamente las clases necesarias. Con el fin de intentar mejorar los resultados obtenidos con los mismos. Debido a las características de los *Large Language Models* se pudo aumentar el tamaño de la base de datos con sentencias provenientes de redes sociales o similares, ya que únicamente es necesario que aprendan a realizar la tarea solicitada, clasificación de texto según las emociones. A pesar de aumentar la base de datos de entrenamiento, se mantuvo el 20 % de los datos originales para un partición de test para comprobar el funcionamiento del modelo en el contexto deseado.

2.2. Modelos

El estado del arte de la clasificación de emociones en texto se encuentra en codificadores basados en *TRANSFORMER*[10], como serían BERT o RoBERTa [11]. Sin embargo, el crecimiento notable de modelos de lenguaje como GPT-3.5 y GPT-4, impulsados en gran parte por aplicaciones como ChatGPT, ha llevado a empresas como Meta a desarrollar modelos de código abierto. Esta tendencia ha generado un interés significativo en las ventajas que estos modelos avanzados ofrecen en comparación con los modelos tradicionalmente utilizados. Estos modelos avanzados se destacan por su capacidad para contextualizar información, generar texto coherente, adaptarse a tareas específicas y mantener conversaciones naturales. Además, su entrenamiento continuo promete mejoras constantes en su desempeño, lo que los convierte en opciones atractivas para aplicaciones de clasificación que requieren una comprensión profunda del lenguaje y el contexto en el que se utilizan.

Modelos generativos

LLaMA: Open and Efficient Foundation Language Models [12]

Este modelo, al igual que la mayoría de los modelos utilizados en tareas basadas en Procesamiento del Lenguaje Natural (NLP), se basa en *TRANSFORMER*, una arquitectura ampliamente empleada, su enfoque es similar al de modelos como GPT. Es importante destacar que este modelo ha sido entrenado utilizando un extenso corpus de datos que abarca 4.749 terabytes.

La estrategia usada en estos entrenamientos se basa en el *instruction-tuning*, una técnica similar al *fine-tuning* con el cual alinearemos los pesos del modelo para realizar ciertas tareas. A pesar de ser una herramienta muy potente, con los recursos disponibles (NVIDIA TITAN V) no se consiguió llegar a hacer funcionar este modelo de manera correcta.

Llama 2: Open Foundation and Fine-Tuned Chat Models [13]

El modelo es una mejora con respecto al anterior. Se mejoraron los datos del entrenamiento en referencia al curado. Ha sido entrenado con aproximadamente 2 trillones de tokens con un mayor cuidado en el filtrado y el tratamiento de los mismos. En cuanto al tamaño de los modelos disponemos de 7B, 13B y 70B, además de las respectivas opciones de *chat*.

La implementación se ha conseguido gracias al uso de

gráficas más potentes además de técnicas de *parameter efficient fine tuning*, entre las cuales encontramos *Low-Rank Adaptation* [14] y *Quantized Low-Rank Adaptation*[15].

El *Low-Rank Adaptation* (LORA) es una técnica que congela los pesos preentrenados y agrega matrices de descomposición de rango entrenables en cada capa de la arquitectura *TRANSFORMER*. Esto reduce significativamente el número de parámetros entrenables para tareas posteriores. Las ventajas de esta técnica de entrenamiento radican en un menor consumo de la memoria gráfica, VRAM, debido a la descomposición de las matrices. Además, podemos seleccionar a que lugares de la arquitectura *TRANSFORMER* queremos añadirle estos pesos query, key, value o output, partes esenciales de las arquitecturas sobre las que se trabaja en el NLP. Por tanto, al poder seleccionar los lugares en la arquitectura sobre los que queremos modificar y no trabajar con los pesos iniciales, además de guardar estos pesos con matrices de un rango seleccionado conseguimos realizar el entrenamiento de manera más eficiente sin tener pérdidas notables en el modelo.

Quantized Low-Rank Adaptation (QLORA) se basa en LORA, pero se distingue por una característica clave que lo hace extremadamente eficiente en el manejo de modelos de gran envergadura: la cuantización.

En este proceso, se produce una compensación entre el tiempo de entrenamiento y la utilización de memoria. Para lograrlo, los valores del modelo base se cargan en una representación de 4 bits (equivalente a 16 niveles), en contraposición a los convencionales 32 bits en punto flotante que se utilizan comúnmente en procedimientos de este tipo.

Esta transformación se logra a través de la normalización de los pesos del modelo, donde se establece su media en cero y su varianza en uno. Luego, estos pesos normalizados se cuantizan a 4 bits. En el caso de la normalización nf4 [16] empleada en el entrenamiento, los niveles de cuantización se distribuyen de manera uniforme y equidistante. Finalmente, en el proceso de des-cuantización, los pesos se restauran a su precisión completa de 32 bits. Es importante destacar que durante el *forward* y *backward pass*, los pesos base se elevan temporalmente a 32 bits para realizar los cálculos necesarios, y después se restauran a su estado original de 4 bits una vez completado.

Concluido este proceso, únicamente se modificarán en memoria los pesos inyectados mediante LoRA, los cuales se conservan en un formato de 32 bits. El resto no se modifican y permanece en su formato original de 4 bits. Esto permite que QLORA optimice de manera significativa el rendimiento sin requerir una cantidad excesiva de memoria.

La estructura resultante, basada en el diseño de los métodos de entrenamiento y los resultados obtenidos, se compone de las siguientes etapas :

En primer lugar, se realiza un ajuste de formato en la sentencia que se introducirá en el *tokenizador* y, posteriormente, en el modelo. Cada *Large Language Model*

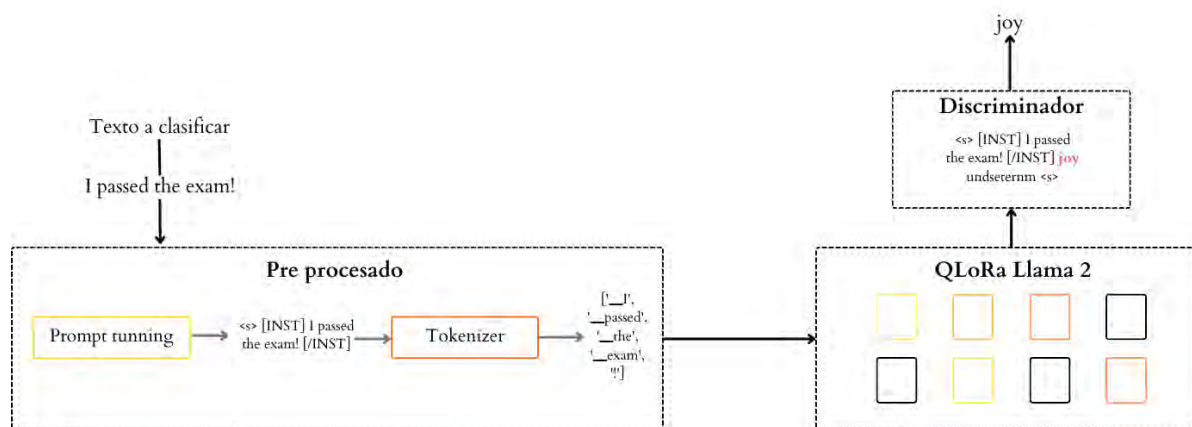


Figura 2. Estructura del sistema.

(LLM) tiene una estructura de entrada en la que funciona de manera óptima. El sistema seleccionado puede verse en la Figura 2.

El modelo que había sido entrenado para emular un chat, cuyo formato es el siguiente:

<s>[INST] Sentence [/INST] Output <s>

Donde *Sentence* y *Output* representan las entradas y salidas del sistema. Utilizando la interfaz de entrenamiento de *HuggingFace*, se puede realizar un ajuste fino supervisado.

A continuación, se encuentra el *tokenizador*, encargado de dividir y asignar cada palabra a diferentes tokens. Dado que se trata de un modelo de gran envergadura, Llama 2, posee un amplio vocabulario y es capaz de identificar la mayoría de las palabras sin dificultad.

El tokenizador precede al modelo, que será entrenado y posteriormente optimizado para su uso en la detección de sentimientos. Dado que es un modelo generativo, puede generar más tokens de los necesarios, por lo que se aplicará una limitación.

Finalmente, se emplea un módulo de discriminación en el que se selecciona únicamente la emoción detectada, descartando el resto, ya que podría generarse ruido o incluso podría explicar la razón detrás de la emoción.

3. Resultados y discusión

3.1. Entrenamiento llama 2

Para actualizar los pesos de LoRA se ha realizado de un entrenamiento basado en 4 épocas, dependiendo del tamaño del dataset consta de un mayor o menor número de pasos. Una tasa de aprendizaje de $\mu = 3e - 6$, un decaimiento de los pesos $\gamma = 0,001$, el optimizador AdamW y un tamaño de lote de 32, implementando una tasa de aprendizaje lineal.

En cuanto a los parámetros característicos de la implementación de *Quantized Low Rank Adaptation* han sido: La precisión nf4, Un rango en la descomposición de la matriz de 16, lo cual hará que la matriz de representación de los pesos esté representada por 16 filas independientes, se compromete la precisión del modelo pero se mantiene la estabilidad y reduce el tamaño de los pesos LoRA. Por último, se añade una probabilidad de

dropout de 0.1 con el fin de evitar en cierta manera el sobreajuste. El resultado comienza con un error de $3e6$, que converge rápidamente a 0,5 y se mantiene estable durante las épocas de entrenamiento.

Tras este entrenamiento se consigue un modelo generativo el cual es capaz de clasificar las emociones propuestas.

3.2. Comparación de resultados

En los estudios realizados en el estado del arte, obtenemos que las arquitecturas más empleadas son RoBERTa, DistilBERT[17], BERT entre otras.

En la siguiente tabla se tiene una comparación de los resultados obtenidos por este estudio basado en la misma base de datos sobre la que trabajamos y los obtenidos con el modelo generativo Llama 2.

Modelo	Precisión sobre test
LLAMA 2	0.7832
BERT	0.7009
RoBERTa	0.7431
DistilBERT	0.6693
XLNet[18]	0.7299

Tabla 2. Comparación de resultados

A la vista de los resultados se concluye una mejora en el análisis del estudio de las emociones gracias a estos nuevos modelos de gran envergadura. En la figura 3 se observa la matriz de confusión después de realizar la evaluación sobre el conjunto de datos de test.

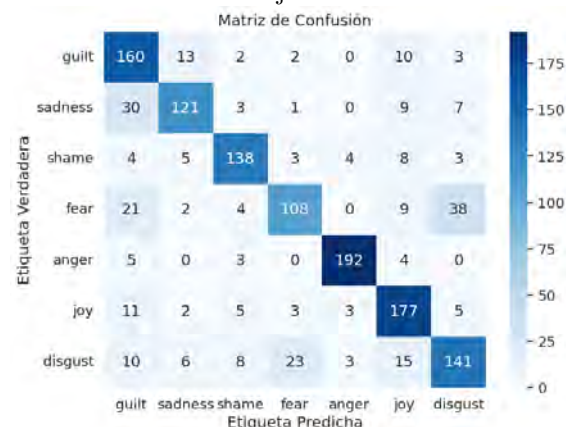


Figura 3. Matriz de confusión.

Además, tras un breve estudio de los fallos cometidos por el modelo, se observa que las emociones que no está detectando correctamente es debido a una falta de contexto en el que se podría identificar tanto la emoción etiquetada en la base de datos como la detectada por el modelo. Véase un ejemplo:

<s>[INST] In this task, you will be performing a classification exercise aimed at identifying the underlying emotion conveyed by a given sentence. The emotions to consider are as follows: Anger, Joy, Sadness, Guilt, Shame, fear or disgust
Sentence: when i for the first time failed a junior exam [/INST] shame nobody likes <s>

En el caso anterior se detecta la emoción vergüenza pero la etiquetada de la base de datos es asco. Por tanto cabría esperar que sobre una base de datos igual de extensa pero con las etiquetas mejor seleccionadas para casos como este o, aceptando varias clasificaciones por sentencia tendríamos mejores resultados.

4. Conclusiones

Tras el estudio realizado, se observa una mejora en los resultados obtenidos en la clasificación de emociones, con resultados mejores que los propuestos en el estado del arte. Es así gracias a que las nuevas arquitecturas presentadas por empresas como Meta mejoran de manera significativa la comodidad y el desarrollo de trabajos sobre NLP. Debido a la gran capacidad que tienen de comprender el comportamiento humano y ser fácilmente adaptable a diferentes tareas. Por tanto, este trabajo constituye una nueva dirección a la hora de la clasificación de texto, en este caso, emociones. A medida se comiencen a optimizar estos entrenamientos, se comenzarán a observar ventajas significativas en el campo del procesamiento del lenguaje natural, llegando incluso estos modelos a ser funcionales en tareas aplicando diferentes técnicas de modificación de *prompt*.

Referencias

[1] Juan Carlos Duro Martínez. ¿Sabemos cuántos profesionales especialistas en psicología clínica trabajan en el Sistema Nacional de Salud Español? *Papeles del Psicólogo*, 42:81 – 93, 08 2021.

[2] Armin Seyeditabari, Narges Tabari, and Wlodek Zadrozny. Emotion detection in text: a review, 2018.

[3] Acheampong Francisca Adoma, Nunoo-Mensah Henry, and Wenyu Chen. Comparative analyses of bert, roberta, distilbert, and xlnet for text-based emotion recognition. In *2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, pages 117–121, 2020.

[4] Sepp Hochreiter and Jürgen Schmidhuber. Long Short-Term Memory. *Neural Computation*, 9(8):1735–1780, 11 1997.

[5] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding, 2019.

[6] Dorottya Demszky, Dana Movshovitz-Attias, Jeongwoo Ko, Alan Cowen, Gaurav Nemade, and Sujith Ravi. Goemotions: A dataset of fine-grained emotions, 2020.

[7] Venkatesh Duppada and Sushant Hiray. Seernet at emoint-2017: Tweet emotion intensity estimator, 2017.

[8] K. R. Scherer and H. G. Wallbott. Evidence for universality and cultural variation of differential emotion response patterning. *Journal of Personality and Social Psychology*, 66:310–328, 1994.

[9] Diman Ghazi, Diana Inkpen, and Stan Szpakowicz. Detecting emotion stimuli in emotion-bearing sentences. In Alexander F. Gelbukh, editor, *Computational Linguistics and Intelligent Text Processing - 16th International Conference, CICLing 2015, Cairo, Egypt, April 14-20, 2015, Proceedings, Part II*, volume 9042 of *Lecture Notes in Computer Science*, pages 152–165. Springer, 2015.

[10] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need, 2023.

[11] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Roberta: A robustly optimized bert pretraining approach, 2019.

[12] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. Llama: Open and efficient foundation language models, 2023.

[13] Hugo Touvron, Louis Martin, and Kevin Stone et al. Llama 2: Open foundation and fine-tuned chat models, 2023.

[14] Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models, 2021.

[15] Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. Qlora: Efficient finetuning of quantized llms, 2023.

[16] Davis Yoshida. Nf4 isn't information theoretically optimal (and that's good), 2023.

[17] Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter, 2020.

[18] Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Ruslan Salakhutdinov, and Quoc V. Le. Xlnet: Generalized autoregressive pretraining for language understanding, 2020.

Agradecimientos

Este trabajo ha sido realizado gracias a los fondos recibidos del Ministerio de Ciencia e Innovación por el proyecto DIPSY (TED2021-13101B-C21) y los fondos de la Generalitat Valenciana (GVA) por el proyecto COM-TACTS2 (CIPROM/2022/20).

Aplicación de Realidad Virtual para Simulación de Recorridos de Evacuación de espacios sanitarios

K. Essamhi Saidi¹, D. Ojados González^{2,3}, G. Salcedo Eugenio³, M. García Fuentes^{2,3}, I.J. Ibarra Berrocal^{2,3}

¹Universidad Politécnica de Cartagena, Cartagena, España {kamilia.essamhi}@edu.upct.es

²Laboratorio de Investigación, Desarrollo e innovación de Tecnologías Biomédicas (LIDiTeB), Hospital General Universitario Santa Lucía - Universidad Politécnica de Cartagena, Cartagena, España, {lola.ojados, marta.garcia, isidro.ibarra}@upct.es

³Servicio de Apoyo a la Investigación Tecnológica (SAIT), Universidad Politécnica de Cartagena, Cartagena, España {gustavo.salcedo}@upct.es

Resumen

Este trabajo tiene como objetivo diseñar un prototipo de experiencia virtual a partir del modelado 3D del Hospital General Universitario Santa Lucía de Cartagena, Área 2 del Servicio Murciano de Salud y de la programación de contenidos de RV para reproducir un entorno virtual donde simular recorridos de evacuación en situaciones de emergencia. Para acotar el alcance de este trabajo y validar el resultado del prototipo obtenido, es suficiente trabajar sobre una sección del complejo hospitalario donde se encuentra el “Laboratorio de Investigación Desarrollo e innovación de Tecnologías Biomédicas (LIDiTeB)”, ubicación elegida como punto de partida del recorrido de evacuación a simular mediante técnicas de RV.

El objetivo principal es demostrar la capacidad de la técnica de Realidad Virtual aplicada al campo de la Prevención de Riesgos Laborales y en concreto a recorridos de evacuación, haciendo evidente su constitución como nueva herramienta para facilitar y dar una nueva alternativa a las vías de evacuación gestionadas de forma convencional en entornos públicos concurridos, como hospitales.

Se desarrollará un prototipo de gemelo digital para llevar a cabo una experiencia inmersiva piloto en una escena virtual del edificio. Esto permitirá ensayar actividades definidas por el Servicio de Prevención de esta institución, como simulacros de evacuación, que por los riesgos que entrañan, no se podrían experimentar en la realidad física. En resumen, se aprovechará la Realidad Virtual para mejorar la seguridad y preparación del personal en casos de evacuación en un entorno hospitalario complejo.

1. Introducción

La realización de este trabajo viene motivada por la intención de garantizar una evacuación segura del edificio elegido, el Hospital General Universitario Santa Lucía (HGUSL) de Cartagena perteneciente al área II del Servicio Murciano de Salud (SMS), y confirmar la corrección de los recorridos de evacuación establecidos. De esta manera sería posible facilitar y ayudar a la rápida y eficaz evacuación, investigar y poner en marcha el avance tecnológico al servicio y beneficio de los usuarios y sensibilizar a todos los usuarios a través de una realidad visual y experimental sin daños. El hospital es de nueva construcción y dispone de elementos constructivos, técnicos y protocolos organizativos que garantizan la

evacuación de todas las personas, por lo que se cuenta con que su diseño cumple con todas las normativas vigentes actualmente en cuanto a las condiciones básicas de seguridad en caso de incendio, y que proporciona la seguridad suficiente para una evacuación eficaz en caso de emergencia o catástrofe.

2. Objetivos

El principal objetivo es demostrar la capacidad de la tecnología de Realidad Virtual aplicada a recorridos de evacuación para mejorar la información de los usuarios y facilitar el proceso de evacuación de edificios en condiciones de seguridad y autonomía en casos de situaciones de emergencia. Para alcanzar este objetivo se busca obtener un prototipo de evacuación en situaciones de emergencia en un espacio limitado de pequeño tamaño, el Laboratorio de Investigación Desarrollo e innovación de Tecnologías Biomédicas (LIDiTeB), ubicado en la primera planta del HGUSL y las instalaciones colindantes, con la intención de ampliarlo al resto de recintos del hospital. Para lograr tal objetivo será necesario profundizar en investigaciones relacionadas con la PRL e incorporar nuevas tecnologías como la RV para la confección y programación de los escenarios. Se pretende evidenciar la importancia de la prevención en materia de evacuación de espacios y control de riesgos y ayudar a crear una cultura preventiva general más allá del cumplimiento legal, para lo que el método de implementación de técnicas como la realidad virtual es una herramienta que resulta eficaz.

3. Materiales

Se ha hecho uso de dos espacios de trabajo para la realización de este proyecto: LIDiTeB y SEDICUPCT-FABLAB que se trata del laboratorio de fabricación digital, disponible en la UPCT y que forma parte de la red mundial de laboratorios fabulosos o laboratorios de fabricación digital impulsada desde el Massachusetts Institute of Technology (MIT). En este laboratorio se dispone de gadgets de RV, como las gafas de Realidad Virtual HP Verber o la plataforma de locomoción Virtualizer VR, utilizadas en este trabajo.

Se ha utilizado el software CAD Rhinoceros para modelar el edificio a partir de los planos del HGUSL proporcionados por el personal técnico del mismo, y el software Unity para realizar el entorno interactivo. Unity es un motor gráfico o una herramienta para el desarrollo gráfico-físico de animación y sonidos de inteligencia artificial, tiene una serie de comandos y funciones de programación que permite diseñar, crear y producir el funcionamiento de un entorno interactivo de videojuegos o aplicaciones de RV / RA para cualquiera que sea la aplicación de destino [1].

4. Metodología

4.1. Levantamiento 3D del edificio

El levantamiento 3D del edificio se realizó en dos fases. En la fase de análisis se describió su arquitectura y distribución como complejo hospitalario longitudinal que sigue un eje norte-sur paralelo a la autovía (CT-33). Se presenta para su comprensión un diagrama conceptual del complejo hospitalario [2] (Figura 1).



Figura 1. Diagrama conceptual del edificio

Por otro lado, en la fase de síntesis del levantamiento 3D se generó un modelo tridimensional de las tres plantas: Planta Baja, donde se ubican las salidas principales al exterior del edificio en busca de puntos de encuentro, Planta 1 o planta técnica, donde se ubica el laboratorio LIDiTeB y por ende donde empiezan las simulaciones de los recorridos de la escena virtual y Planta Terraza, planta por encima del LIDiTeB, compuesta por una gran terraza, que tiene la función de cubierta del laboratorio.

4.2. Etapa de modelado y creación de escena con interacción

Partiendo de los planos 2D en formato AutoCAD, se obtuvo el modelado 3D con el software Rhinoceros de las tres plantas (planta baja, primera y terraza) siguiendo un proceso constructivo de cinco fases:

- Primera fase: levantamiento de elementos estructurales (pilares, forjados, cajas de escaleras de obra, huecos de ascensores y escaleras mecánicas).
- Segunda fase: definición de las particiones interiores y los muros exteriores de las plantas baja y primera. Estos elementos verticales están contruidos con materiales mixtos, acero, vidrio y muros de hormigón.
- Tercera fase: levantamiento sin detallar de los seis bloques de unidades hospitalarias del edificio, sólo se representó su geometría de forma esquemática ya que excede del alcance de este trabajo. Se simuló el entorno lo más parecido posible para que a los usuarios les sea sencillo reconocer los puntos de salida.

- Cuarta fase: se insertaron las carpinterías tanto exterior como interior (ventanas, puertas, barandillas, etc) y se detalló el mobiliario urbano.
- Quinta fase: se modeló el interior del LIDiTeB, lo más parecido posible a la realidad en función de las limitaciones de Rhinoceros.

El modelado 3D obtenido con Rhinoceros se exportó como archivo fbx ya que es el mejor formato compatible para importar modelos 3D desde Unity, se realizó la importación del modelo 3D en Unity y se procedió a definir los materiales y texturas de los elementos del edificio modelado: suelos, muros, particiones, escaleras, barandillas, galería, etc, gracias a la biblioteca de texturas *Textures.com* [3]. Para mejorar el realismo visual en el modelo y ayudar a la ambientación y gameplay de la simulación se procedió a la iluminación de todos los espacios cerrados (Figura 2).

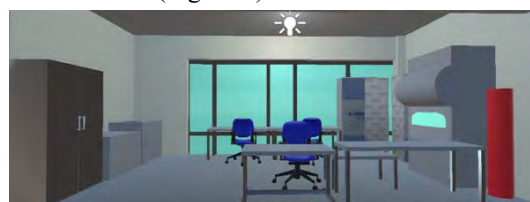


Figura 2. Texturas e iluminación dentro del LIDiTeB (Unity)

Posteriormente, se realizó el proceso de colocación de la señalización necesaria en materia de prevención, basada en la señalización real ya posicionada en el edificio y además se añadieron las señales que se ha considerado necesario para completar la señalización de los recorridos seleccionados (Figura 3).



Figura 3. Señales en pasillo de consultas externas (Unity)

Con la intención de dotar a la escena de la posibilidad de realizar un ensayo realista de evacuación se añadió una simulación de fuego junto a la cubeta ubicada en una de las poyatas de LIDiTeB. Para crear fuego y humo se importó un paquete ya preconfigurado para estos efectos de la página Unity Asset Store [4].

Además, se creó una interfaz de usuario (Figura 4) usando un sistema UI. El canvas que se configuró desde Unity como pantalla inicial del contenido RV a ensayar, incluye un panel, la imagen del logo LIDiTeB [5], botones de inicio y salida, además de información del lugar desde el que se pretende realizar la simulación del recorrido virtual de evacuación hacia el exterior del edificio donde se encuentra este laboratorio. También, desde esta pantalla de inicio, se informa de los objetivos con los que está planificada esta escena (EVACUACIÓN LIDiTeB) y las bases de funcionamiento de la escena.



Figura 4. Interfaz de usuario. Pantalla inicial

Para finalizar la configuración de la escena creada y poder simular y ensayar los recorridos de evacuación, se creó un script donde se ubicó en el modelo virtual el punto de encuentro existente en el entorno físico real (Figura 5).



Figura 5. Menú final punto de encuentro

4.3. Conexión de hardware con el entorno digital creado en Unity

Se añadió un Prefab del paquete estándar Assets para controlar la cámara de la escena RV con el teclado del ordenador desde el que se podrá gestionar el recorrido por el usuario, lo que permitirá navegar por la escena en primera persona.

En segundo lugar, se descargó e importó el paquete de herramientas XR de Interaction Toolkit de Unity para conectar las gafas y mandos RV.

Finalmente, para la conexión de la Plataforma Virtualizer VR con Unity se utilizó un software de desarrollo SDK (Software Development Kit) compatible con el SDK del andador y se importó el paquete CybSDK en Assets [6].

5. Resultado y discusión

Partiendo del interior del LIDiTeB hasta el exterior del hospital, se detalló el edificio lo máximo posible dentro de lo que permite el software Rhinoceros (Figura 6). Además, se obtuvieron las escenas en el software Unity mejorado las vistas con texturas, añadiendo luces a las escenas, mobiliario urbano y puntos de encuentro (Figura 7) y se detallaron tanto los componentes de evacuación como las señalizaciones, protección contra incendios, planos de evacuación, efecto de fuego y humo y avatares (Figura 8).



Figura 6. Resultado del modelo 3D en Rhinoceros



Figura 7. Resultado del modelo 3D en Unity



Figura 8. Resultado interior LIDiTeB

Se representaron recorridos de realidad virtual en función del tipo de inmersión, dando lugar a tres modalidades diferentes de este mismo desarrollo.

Primero se configuró una experiencia semi inmersiva o realidad virtual de escritorio, donde el usuario interactúa con una pantalla haciendo uso del ratón y teclado del propio ordenador, por lo que no se siente sumergido en el mundo virtual pero sí tiene capacidad para interactuar con la escena.

La siguiente modalidad es la obtenida mediante la conexión del desarrollo en Unity con las gafas RV HP Reverb consiguiendo una inmersión total en la escena de recorrido de evacuación creada, donde el usuario que porta las gafas RV interactúa y se siente inmerso en la misma. El espacio es ilimitado tanto en la escena virtual como en el mundo físico, por lo que, para facilitar el recorrido por la escena virtual, cuando se trata de largas distancias, se programan acciones como la de teletransporte, gestionadas por los mandos RV que complementan a las gafas RV

Por último, se ha conseguido una inmersión total en la escena de recorrido de evacuación creada usando la plataforma locomotora Virtualizer VR. Esta comunicación permite al usuario caminar, correr y moverse por la escena de manera controlada y sin desplazamiento físico del fuera de la plataforma pero siendo ilimitado el desplazamiento por la escena virtual.

En los tres niveles de inmersión la experiencia RV es monousuario, aunque puede mostrarse en monitor o pantalla de proyección a espectadores externos a la escena virtual, además esta experiencia se puede convertir en multiusuario si se lanza desde una plataforma que permita comunicar más de una unidad de gafas RV mediante software.

Para mostrar estos resultados se han grabado videos en función del *gadget* usado en los recorridos generados. El recorrido principal, que pasa por las escaleras de evacuación de la galería y el recorrido alternativo, más largo, que pasa por el pasillo de las consultas externas del hospital. Estos vídeos se han recogido en la tabla 1.







	Modalidad 1	Modalidad 2	Modalidad 3
Recorrido principal			
Recorrido alternativo			

Tabla 1. Códigos QR de los videos por modalidad y recorrido

6. Conclusiones

Se ha logrado el objetivo planteado de creación de contenidos de RV aplicados a PRL para la realización de simulacros de evacuación en un edificio público en situación de emergencia, siendo la implementación de aplicaciones de Realidad Virtual y Aumentada, una buena estrategia de formación y ensayo como elemento de apoyo a la PRL. Además, permite entender mejor y más rápido situaciones de emergencia, ya que el usuario se sitúa en un entorno lo más fiel posible a la realidad de su espacio de trabajo, actuando en el entorno virtual como lo haría en una situación real de emergencia. Con los gadgets adecuados, la frontera entre lo real y lo irreal se reduce, permitiendo al usuario vivir una experiencia realista al sentirse rodeado e inmerso en este entorno digital, siendo una forma muy efectiva para retener información y aprender. Los tres recorridos creados y testados según el nivel de inmersión obtenido en la configuración del contenido de Realidad Virtual con ordenador de sobremesa combinado con teclado y ratón, con gafas RV y con plataforma locomotora, confirman el buen funcionamiento de las dos rutas de evacuación propuestas.

Referencias

- [1] Technologies, U. (s. f.). Unity - Manual: Introducción a los componentes.
<https://docs.unity3d.com/es/530/Manual/Components.htm>
- [2] Soleundurraga. (2022). Nuevo Hospital Universitario de Santa Lucía / CASA sólo arquitectos. ArchDaily en Español.
<https://www.archdaily.cl/cl/02-77564/nuevo-hospital-universitario-de-santa-luciacartagena-murcia-espana-casa-solo-arquitectos>
- [3] Textures for 3D, graphic design and Photoshop! (s. f.). Textures for 3D, graphic design and Photoshop!
<https://www.textures.com/library>
- [4] Particle Pack | VFX Particles | Unity Asset Store. (2019, 22 marzo). Unity Asset Store.
<https://assetstore.unity.com/packages/vfx/particles/particle-pack-127325>
- [5] LIDiTeB - Laboratorio de Investigación, Desarrollo e innovación de Tecnologías Biomédicas. (2023, 30 mayo). Inicio » LIDiTeB. LIDiTeB. <http://liditeb.sait.upct.es/>
- [6] Cyberith. (2022, 16 mayo). VR Walking Platform - Cyberith Virtualizer VR Treadmills. Cyberith Virtualizer | Professional Virtual Reality Treadmills.
<https://www.cyberith.com/>

Realidad virtual aplicada a la ingeniería del factor humano. Caso de estudio “Manipulación de paciente en enfermería”

G. Salcedo Eugenio¹, A. Macián Morales¹, D. Ojados González¹, I.J. Ibarra Berrocal¹

¹Servicio de Apoyo a la Investigación Tecnológica (SAIT), Universidad Politécnica de Cartagena, Cartagena, España
{gustavo.salcedo, alvaro.macian, lola.ojados, isidro.ibarra}@upct.es

Resumen

En este proyecto se implementa la realidad virtual y sistemas de captura de movimiento aplicada a la prevención de riesgos laborales, ingeniería del factor humano, análisis de posturas y estudios ergonómicos.

Estas tecnologías permiten reproducir en un mundo virtual, las posturas de un trabajador realizando una tarea, observar detalladamente cada movimiento y analizar ángulos de las articulaciones.

El caso de estudio se centra en la manipulación de pacientes en Enfermería. Se realiza un estudio descriptivo sobre una población de enfermeros, realizando la tarea de manipulación de un paciente en una camilla.

La captura de movimiento (también Mocap) es una forma de registrar digitalmente los movimientos humanos. Los datos de captura de movimiento grabados se asignan en un modelo digital en el software 3D (Unity 3D) para que el personaje digital se mueva como la persona que realiza los movimientos.

1. Introducción

Los estudios epidemiológicos realizados en el sector socio sanitario indican que la incidencia de lumbalgias es, según la OIT, entre el doble y el triple más que en otros sectores más asociados a las cargas físicas. Las causas se deben a la movilización manual de pacientes como el factor de riesgos ergonómicos que mayor índice de absentismo laboral. Según datos del INSS, la primera causa de incapacidad temporal por enfermedad común es la lumbalgia y la principal causa de los sobreesfuerzos ha sido la movilización de enfermos, representando más de un 55% sobre el total. (Ver Figura 1).

La movilización de pacientes se caracteriza por la singularidad de los ambientes de trabajo en condiciones de carga a manipular (personas), espacio, ayudas técnicas y frecuencias. Además de ser una de las tareas más habituales, presenta las siguientes singularidades:

- Las formas y volúmenes presentan extremidades que se pueden movilizar con varios grados de libertad.
- Los pesos que se deben manipular por término medio son mucho más elevados que en los sectores industriales.
- La voluntariedad de la persona atendida, la no colaboración del sujeto o la colaboración inadecuada del sujeto.

- Los espacios son reducidos y diseñados para el confort de la persona atendida, no el trabajador.

En el sector sociosanitario se dan unas características específicas frente a las cuales las metodologías tradicionales usadas para los estudios ergonómicos presentan algunas oportunidades de mejora continua.



Figura 1. Movilización manual de paciente

Fuente: <https://auxiliarenfermeria.es/movilizacion-de-pacientes-2/>

Estas aproximaciones generales están siendo complementadas por métodos específicos para la manipulación de pacientes y el estudio ergonómico de las tareas propias del sector sanitario. En este proyecto se implementa las nuevas tecnologías como son la Realidad Virtual y Captura de Movimientos para realizar un análisis y estudio ergonómico de la Manipulación y movilización manual de pacientes en enfermería.

Se puede definir la Realidad Virtual (VR) como “un mundo virtual generado por ordenador en el que los usuarios sienten estar dentro de él”. Permite simular una experiencia sensorial, en un espacio real o imaginario, a través del cual se puede interactuar en estos entornos, pudiendo tocar y mover objetos, caminar, conducir. (Ver Figura 2).



Figura 2. Realidad Virtual

Fuente: <https://www.edsrobotics.com/blog/realidad-virtual-que-es/>

Existen tecnologías que permiten replicar el movimiento de un cuerpo y luego usarla en un ordenador para darle diversas aplicaciones, este proceso de obtener los datos de posición en cada instante del tiempo se conoce como captura de movimiento llamada también MoCap (Motion Capture) por sus siglas en inglés. (Ver Figura 3).

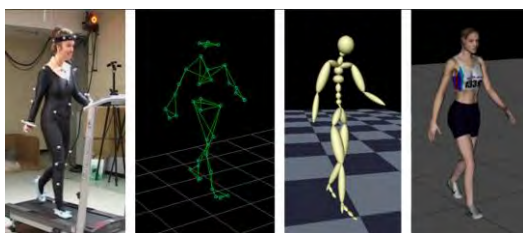


Figura 3. Mocap (Captura de movimiento)

Fuente:

<http://www.mitak.or.kr/upload/data/3%20Rotoscoping.pdf>

MoCap es un proceso que permite grabar y describir el movimiento de un cuerpo en el espacio, el cual es representado por los cambios de traslación y rotación, gracias a dichos cambios es posible calcular velocidades y aceleraciones que hoy en día son empleados como información valiosa ya sea para estudiar el comportamiento de un algún objeto en el espacio o de toda la cadena cinemática del cuerpo de una persona, un animal o alguna cosa.

Las dos anteriores tecnologías se aplican a la ingeniería del factor humano, específicamente a la evaluación ergonómica. Los métodos de evaluación ergonómica permiten identificar y valorar los factores de riesgo presentes en los puestos de trabajo para, posteriormente, en base a los resultados obtenidos, plantear opciones de rediseño que reduzcan el riesgo y lo sitúen en niveles aceptables de exposición para el trabajador. La exposición al riesgo de un trabajador en un puesto de trabajo depende de la amplitud del riesgo al que se expone, de la frecuencia del riesgo y de su duración (las posturas forzadas, los levantamientos de carga, la repetitividad de movimientos, etc). Dicha información es posible obtenerla mediante métodos de evaluación ergonómica. Existen una gran variedad de métodos de evaluación ergonómicos: RULA, REBA, Ecuación NIOSH, OWAS, OCRA, etc.

El objetivo principal de este proyecto es implementar la Realidad Virtual y el Sistema de Captura de Movimiento aplicada a la prevención de riesgos laborales concretamente a la ingeniería del factor humano en la manipulación y movilización manual de pacientes en enfermería, reproduciendo en un mundo virtual las posturas de un trabajador realizando una tarea, observar detalladamente cada movimiento y analizar los ángulos con el fin de determinar si se pueden producir lesiones musculoesqueléticas durante la ejecución de la tarea.

2. Metodología

Utilizando un sistema MOCAP (Perception Neuron 2.0), basándose en sensores inerciales de movimiento, se captura y registra digitalmente los movimientos humanos en un modelo virtual (Ver Figura 4). El

Perception Neuron incorpora un software (AXIS NEURON) que permite grabar y visualizar en tiempo real el movimiento resultante de la captura sobre un modelo biomecánico de una persona y ajustar su antropometría según nos interese (Ver Figura 5). Posteriormente los archivos se exportan al software 3D (UNITY) donde se realiza la programación de los ángulos críticos para realizar un análisis biomecánico y determinar así las posibles lesiones musculoesqueléticas (Ver Figura 6).

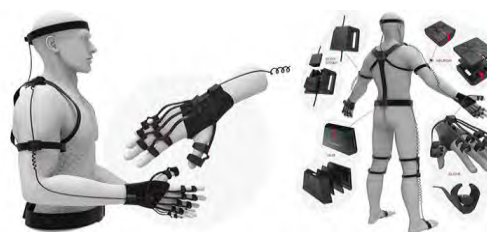


Figura 4. Mocap (Perception Neuron 2.0)

Fuente: https://xinreality.com/wiki/Perception_Neuron



Figura 5. Axis Neuron



Figura 6. Software UNITY 2017

3. Resultados

El caso de estudio se centra en la manipulación y movilización manual de pacientes en enfermería. Se genera un estudio descriptivo sobre una población de enfermeros, realizando la maniobra de movilización de un paciente desde una camilla hacia una silla de ruedas (con y sin grúa de movilización de pacientes) (Ver Figura 7).



Figura 7. Equipamiento Hardware

Las mediciones a realizar sobre las posturas adoptadas por el trabajador son fundamentalmente angulares (los ángulos que forman los principales miembros del cuerpo respecto a determinadas referencias). Cada segmento corporal del modelo virtual se visualiza en la simulación 3D la variación de parámetros durante la ejecución de una tarea (Ver Figura 8).

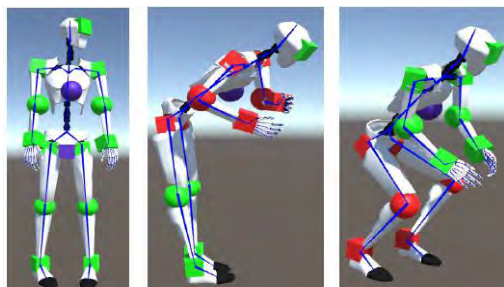


Figura 8. Posturas

Resultados del estudio descriptivo de movilización manual de paciente desde camilla hacia silla de ruedas (Ver Figura 9 y 10).



Figura 9. Movilización manual



Figura 10. Gráfica de movilización manual

Resultados del estudio descriptivo de movilización con grúa de paciente desde camilla hacia silla de ruedas. (Ver Figura 11 y 12).



Figura 11. Movilización con grúa



Figura 12. Gráfica de movilización con grúa

4. Conclusiones

Este sistema es una herramienta tecnológica para analizar y evaluar las posturas de un trabajador, genera un modelo 3D en tiempo real, fiable, fácil de calibrar, configurable para percentiles de ambos géneros y los datos obtenidos se almacenan en una base de datos y pueden ser utilizados para cualquier método de evaluación ergonómica.

En la movilización manual de pacientes la carga a levantar es relativa al peso del paciente lo que genera un factor de riesgo de lesión en la zona lumbar, la flexión de tronco es el parámetro mayor y la flexión de antebrazo es aproximadamente la mitad.

En la movilización de pacientes utilizando grúa la carga a levantar es nula y no hay riesgo de lesión en la zona lumbar, la flexión de tronco es similar a la flexión de antebrazos generando más posturas de menor riesgos por la manipulación de la grúa y aumentando el tiempo de realización de movilización del paciente.

Agradecimientos

Por último, agradecer el apoyo del SEDIC (Servicio de Diseño Industrial y Cálculo Científico) y el grupo de investigación NDSM (Nuevos dispositivos de seguridad en máquinas) de la UPCT. Al Instituto de Seguridad y Salud Laboral de la Región de Murcia por la financiación de este proyecto, agradecer expresamente su compromiso.

Referencias

- [1] Real Decreto 487/1997, de 14 de abril, sobre disposiciones mínimas de seguridad y salud relativas a la manipulación manual de cargas que entrañe riesgos, en particular dorso lumbar, para los trabajadores.
- [2] Guía Técnica para la evaluación y prevención de los riesgos relativos a la manipulación manual de cargas. INSHT (Instituto Nacional de Seguridad e Higiene en el Trabajo).
- [3] Paul Dotte. Método de movilización de los pacientes. Ergonomotricidad en el ámbito asistencial. Ed-Elsevier Massson; 2012.
- [4] Técnica de movilización en paciente. Internet. Disponible en <http://www.auxiliar-enfermeria.com/movilizaciones.htm#marcO2>
- [5] Esquema posiciones anatómicas. Internet. Disponible en http://www.auxiliar-enfermeria.com/esquemas/esquema_posiciones.htm
- [6] CAPTURACIÓN DE MOVIMIENTO.

http://en.wikipedia.org/wiki/Motion_capture

[7] PERCEPTION NEURON MOTION TECHNOLOGIES.

<https://neuronmocap.com/>

NTP 601 Evaluación de las condiciones de trabajo: carga postural. Método REBA (Rapid Entire Body Assessment). Instituto Nacional de Seguridad e Higiene en el Trabajo, España.

Holograma Piramidal 360 aplicado a la ingeniería del factor humano. Caso de Estudio Ergonómico “manipulación de paciente en Enfermería”

A. Macián Morales¹, G. Salcedo Eugenio¹, I.J. Ibarra Berrocal¹, D. Ojados González¹

¹Servicio de Diseño Industrial y Cálculo Científico (SEDIC) del Servicio de Apoyo a la Investigación Tecnológica (SAIT), Universidad Politécnica de Cartagena (UPCT), Cartagena, España
{alvaro.macian, gustavo.salcedo, isidro.ibarra, lola.ojados}@upct.es

Resumen

En este proyecto se desarrolla un dispositivo de visualización holográfica en 3D aplicada a la Prevención de Riesgos Laborales, Ergonomía, análisis de posturas, estudios ergonómicos y formación del personal de enfermería. Esta tecnología permite a trabajadores de enfermería visualizar sus posturas durante la realización de sus tareas proyectadas en un holograma 3D. Permitiendo observar detalladamente cada movimiento y postura. La proyección holográfica 3D es una innovadora herramienta de aprendizaje.

La pirámide holográfica es un sistema compuesto por un proyector formado por una pirámide invertida que es capaz de generar imágenes tridimensionales dentro de su espacio de proyección. La imagen proyectada se ve desde cualquier ángulo de observación. Los cuatro lados de la pirámide están fabricados de un material transparente para que el público pueda ver a través de ella por todas partes y puede reproducir imágenes de vídeo o gráficos generados por ordenador. Esto se produce a través de la creación de reflejo en la superficie y reflexiones. La propia pirámide se utiliza como una especie de prisma que reúne la luz de cuatro proyecciones de vídeo en una imagen sólida. Los hologramas generados son de reflexión, lo que brinda la sensación de que las imágenes son tridimensionales y que están flotando en el aire. También permite reproducir escenas de vídeos ya grabadas o mostrar un holograma en vivo y tiempo real.

1. Introducción

La pirámide holográfica [1] es un sistema compuesto por un proyector formado por una pirámide invertida que es capaz de generar imágenes tridimensionales dentro de su espacio de proyección. La imagen proyectada se ve desde cualquier ángulo de observación.

Esta tecnología permite visualizar las posturas de un trabajador proyectadas en un holograma piramidal 360°. El uso de hologramas en procedimientos de estudios de puestos de trabajo ergonómicos sirve para realizar un análisis técnico completo en 360 grados, permitiendo observar detalladamente desde todas las perspectivas cada movimiento y postura del trabajador.

2. Objetivo

En este proyecto se desarrolla un dispositivo de visualización holográfica en 360° aplicada a la Prevención de Riesgos Laborales, Ergonomía, análisis de posturas, estudios ergonómicos y formación del personal

sanitario en la correcta realización de sus tareas de manera ergonómicamente correcta.

El personal sanitario podrá observar en el holograma un caso de estudio de un enfermero realizando la tarea de extracción de sangre en un paciente siguiendo dos técnicas diferentes. La primera con el paciente tendido en una camilla, y la segunda, sentado en una silla. Se ha realizado un estudio estadístico sobre una población de 10 enfermeros, realizando las dos técnicas de extracción de sangre y posteriormente se han analizado los resultados del estudio.

3. Metodología

Inicialmente se realiza el montaje de un estudio de grabación de dimensiones (3m x 3 m) con telas de croma verde compuesto por un Kit de Estudio de grabación, Kit de Iluminación, 4 Cámaras de grabación (ubicadas en las vistas lateral derecha e izquierda, vista frontal y posterior), una camilla, una mesa y una silla. (Ver Figura 1).



Figura 1. Montaje de Estudio de grabación Croma Verde

En el estudio se realiza la grabación a 10 enfermeros(as) realizando la técnica de extracción de sangre a un paciente de dos maneras distintas. La primera técnica con el paciente sentado en una silla, simulando la extracción de sangre y la segunda con el paciente tumbado en una camilla. (Ver Figura 2 y 3).



Figura 2. Técnica de extracción de sangre paciente sentado en silla.



Figura 3. Técnica de extracción de sangre en camilla.

El objetivo de las practicas es ver en qué medida los enfermeros realizan correctamente este tipo de intervención y actividad y cuáles son los puntos críticos en los que a la larga le pueden ocasionar complicaciones ergonómicas a nivel personal debido a una mala intervención enfermera. En este caso, la intervención del procedimiento de extracción de sangre [2] es una de las técnicas enfermeras que más complicaciones de trabajo sufre. Después de realizar todas las grabaciones, los archivos generados por las cuatro cámaras, se guardan en formato de video y se realiza la edición de videos utilizando el software Sony Vegas Pro 14.

3.1. Software Edición de Video (Sony Vegas Pro 14)

La herramienta Sony Vegas Pro 14[3] es un software de edición de video de gran calidad. Incluye innovadoras herramientas: edición de material de SD, HD, 2K y 4K en 2D y 3D estereoscópicas, compatibilidad con formato amplio, control de audio sin precedentes y efectos de transformación y procesamiento impulsados por la GPU.

Es un potente motor de composición y efectos que combina hasta 32 efectos de alta calidad, a los que se les pueden añadir fotogramas clave, por cadena de efectos y aplicarlos a eventos, pistas, medios o proyectos. Automatización de efectos y compresión/extensión de tiempo, mientras aplica efectos de audio personalizables en tiempo real, como Ecuador, Reverberación, Retardo y muchos más.

Mediante el Vegas Pro 14 se editan y sincronizan los videos de las cuatro cámaras para a continuación renderizar los videos definitivos. Posteriormente se proyectan en el Holograma Piramidal.

3.2. Holograma Piramidal 360

La pirámide holográfica [1] es un sistema compuesto por un proyector formado por una pirámide invertida que es capaz de generar imágenes tridimensionales dentro de su espacio de proyección. La imagen proyectada se ve desde cualquier ángulo de observación. La proyección holográfica 3D es una innovadora herramienta de aprendizaje. Los cuatro lados de la pirámide están fabricados de un material transparente para que el público pueda ver a través de ella por todas partes, se reproducen imágenes de video o gráficos 3D generados por ordenador. Esto se produce a través de la creación de reflejo en la superficie y reflexiones. La propia pirámide se utiliza como una especie de prisma que reúne la luz de cuatro proyecciones de video en una imagen sólida. Los hologramas generados son de reflexión, lo que brinda la

sensación de que las imágenes son tridimensionales y que están flotando en el aire. También permite reproducir escenas de videos ya grabadas o mostrar un holograma en vivo y tiempo real. (Ver Figura 4 y 5).

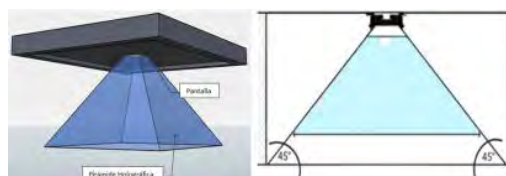


Figura 4. Holograma Piramidal



Figura 5. Holograma Piramidal 360

3.3. Software RULER

RULER [4] es una herramienta para realizar estudios ergonómicos que permite medir los ángulos formados por los segmentos corporales del trabajador sobre fotografías tomadas en el puesto de trabajo. Es necesario utilizar un número suficiente de fotografías desde diferentes puntos de vista (alzado, perfil y vistas de detalle) para lograr medir la verdadera magnitud.

El procedimiento para realizar las mediciones es el siguiente:

- Primero hay que arrastrar el archivo de la fotografía del trabajador sobre el icono “Imagen”. También puedes hacer click sobre él recuadro para escoger el archivo.

- Ajustar la imagen para enfocar los ángulos que se quieren medir mediante los botones de zoom y desplazamiento.

- Usar el mouse para colocar los cuadrados del medidor sobre los puntos convenientes. El cuadrado rojo debe colocarse sobre la articulación que se desea medir. El cuadrado amarillo y el cuadrado azul debes colocarlos sobre los ejes de los dos miembros adyacentes a la articulación. Después se visualizan los ángulos.

La siguiente figura muestra un ejemplo de medición de Flexión de Tronco en el levantamiento manual de carga. (Ver figura 6)



Figura 6. Software Ergonomía "RULER"

Para poder realizar una evaluación ergonómica aplicando el método REBA [5] se deben medir los siguientes ángulos:

- Flexión y rotación lateral del tronco.
- Flexión del cuello.
- Rotación e inclinación lateral de la cabeza.
- Flexión de las rodillas.
- Flexión/ extensión del brazo.
- Flexión del antebrazo.
- Flexión / extensión de la muñeca.

4. Resultados

Tras el estudio estadístico con una muestra de población de 10 enfermeros(as), se realizaron las grabaciones de las dos técnicas de extracción de sangre de un paciente. Se obtienen los videos desde las cuatro principales vistas laterales, frontal y posterior; generando una perspectiva detallada de los trabajos. (Ver Figura 7 y 8).



Figura 7. Grabación de Técnica de extracción de sangre paciente en silla.

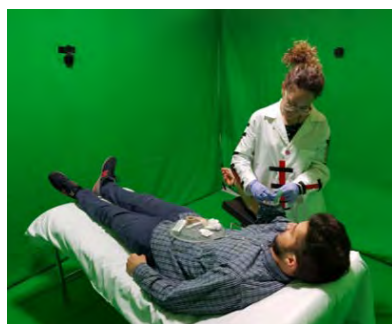


Figura 8. Grabación de Técnica de extracción de sangre paciente en camilla.

4.1. Edición de Videos

La edición de los videos se realiza utilizando el Software Sony Vegas Pro 14, se importan los cuatro videos generados por las cámaras SJCAM 5000+ en formato .mp4, después se sincronizan todos los videos. (Ver Figura 9).



Figura 9. Edición de video "Sony Vegas Pro 14"

Se utiliza la herramienta de "Clave Cromática" para eliminar el fondo verde reemplazándolo por un fondo negro. (Ver Figura 10).



Figura 10. Efecto Croma Sony Vegas

Por último, se editan simultáneamente las cuatro vistas de videos, centrando, modificando el tamaño y aplicando filtros de colores. Se renderizan los videos finales y se guardan en cuatro pendrives.

4.2. Holograma Enfermero en extracción de Sangre

Los cuatro pendrives se introducen en los cuatro proyectores que finalmente reproducen los videos sobre la pirámide de metacrilato, generando un efecto de holograma que muestra las tareas de un enfermero extrayendo sangre a un paciente. (Ver Figura 11 y 12).

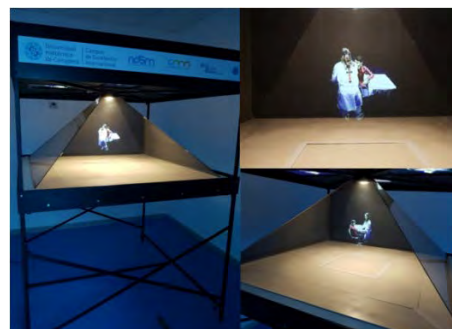


Figura 11. Holograma de enfermero extrayendo sangre a un paciente



Figura 12. Vistas global de Holograma 360°

4.3. Análisis de Posturas y Ángulos

Utilizando el software RULER y las imágenes de las cuatro vistas, el técnico de prevención puede analizar detalladamente y desde una perspectiva global, cada segmento corporal y los ángulos críticos durante la tarea aplicando los métodos REBA[5] y RULA[6](Ver Figura 13 y 14).

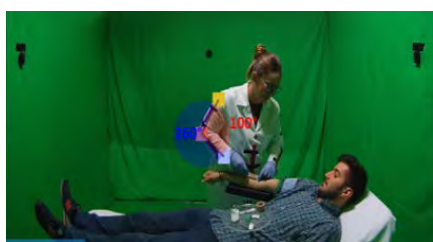


Figura 13. Vista frontal / Ángulos "Software RULER"



Figura 14. Vista Posterior / Ángulos "Software RULER"

En la tabla 1 se muestran algunas de las medidas de ángulos obtenidas.

Trabajador/a	Extracción de sangre en silla		Extracción de sangre en camilla	
	Flexión del tronco	Flexión codo	Flexión del tronco	Flexión codo
1	50°	82°	9°	100°
2	3°	103°	8°	106°
3	27°	128°	10°	133°
4	20°	85°	2°	93°
5	10°	99°	27°	91°
6	41°	100°	32°	78°
7	5°	103°	15°	94°
8	16°	90°	5°	102°
9	52°	56°	13°	96°
10	15°	96°	22°	95°

Tabla 1. Medidas de ángulos obtenidos

Según el análisis ergonómico de las muestras el 46% de los enfermeros realizaron la técnica ergonómica correctamente con el paciente sentado en una silla, frente al 61.53% que lo realizaron correctamente con el paciente tumbado en camilla.

5. Conclusiones

El método de visualización holográfica permite hacer un estudio global de la actividad realizada por un trabajador, permitiendo observar detalladamente en 360 grados cada movimiento y ángulos de las secciones corporales de una persona. El holograma Piramidal 360 es una herramienta innovadora que podrá utilizar el Técnico de Prevención para elaborar un análisis ergonómico más profundo y con puntos de vista diferente, al tener una visión completa de cada postura, ángulo y movimiento. El Enfermero podrá analizar y observar sus propios movimientos y posturas durante la realización de su tarea, generando un autoaprendizaje de las mejoras en las técnicas de extracción de sangre y corrigiendo las posturas forzadas.

Agradecimientos

El equipo de investigadores que ha llevado a cabo este trabajo, tiene presente la gran ayuda que supone contar con el respaldo y apoyo del Instituto de Seguridad y Salud Laboral (ISSL) de la Región de Murcia que por medio del grupo de Nuevos Dispositivos de Seguridad en Máquinas (NDSM) del Servicio de Apoyo a la Investigación Tecnológica (SAIT) de la Universidad Politécnica de Cartagena facilita recursos para el desarrollo de investigaciones en el campo de la prevención y el cuidado de la salud, y está muy agradecido por ello.

Referencias

- [1] Dispositivo holográfico piramidal. Disponible en web: <https://luftscreen.com.mx/piramide-holografica-360/>
- [2] Extracción de Sangre. Procedimientos Especiales. Asociación Española de Biopatología Médica.
- [3] Software Sony Vegas Pro. Disponible en web: <https://www.vegascreativesoftware.com/es/vegas-pro/>
- [4] Software RULER. Disponible en web: <https://www.ergonautas.upv.es/herramientas/ruler/ruler.php>
- [5] Método para evaluación ergonómica REBA. Disponible en web: <https://www.ergonautas.upv.es/metodos/reba/reba-ayuda.php>
- [6] Método para evaluación ergonómica RULA. Disponible en web: <https://www.ergonautas.upv.es/metodos/rula/rula-ayuda.php>

Aplicación de la Realidad Virtual para la capacitación en materia de comunicación en crisis sanitarias

M.T. Herrero Ezquerro¹, A. Garrido Rubio², A. Macián Morales³, D. Ojados González³

¹ Instituto Universitario de Investigación en Envejecimiento, Universidad de Murcia, Murcia, España {mtherrer}@um.es

² Departamento de Ciencia Política y de la Administración, Universidad de Murcia, Murcia, España {agarrido}@um.es

³ Servicio de Apoyo a la Investigación Tecnológica (SAIT), Universidad Politécnica de Cartagena, Cartagena, España {alvaro.macian, lola.ojados}@upct.es

Resumen

Este proyecto trata el desarrollo de una aplicación de Realidad Virtual para formar a los portavoces sanitarios sobre cómo realizar una correcta comunicación durante una crisis sanitaria. Este tipo de comunicación posee características particulares que demandan conocimientos específicos y generan un impacto significativo en la sociedad, lo cual las distingue de otros tipos de comunicaciones.

La aplicación consta de una parte formativa y otra parte de entrenamiento. En la parte formativa, el usuario recibe la información de un formador virtual y podrá elegir sobre qué temas quiere obtener información interactuando con su voz. Los módulos formativos aprovechan las virtudes de la realidad virtual enriqueciendo la experiencia con estímulos visuales de madera inmersiva.

Por otro lado, en la parte de entrenamiento el usuario puede practicar como realizar una comunicación en una sala de prensa virtual. Y al finalizar recibirá un feedback con resultados e indicaciones para mejorar.

1. Introducción

La realidad virtual (VR) es una tecnología que permite a los usuarios sumergirse en un mundo virtual y experimentarlo de manera similar a la realidad. Esto se logra mediante el uso de dispositivos de entrada y salida, como gafas de realidad virtual o guantes de realidad virtual, que permiten al usuario interactuar con el mundo virtual y recibir estímulos sensoriales como el sonido y las imágenes.

Gracias a las nuevas tecnologías, a lo largo de estos últimos años hemos ido viendo cómo ha evolucionado la forma en la que se han ido distribuyendo los contenidos formativos e incluso hemos podido comprobar cómo la introducción de determinadas mecánicas de juego (gamificación) ha mejorado la adherencia de profesionales y estudiantes a los contenidos, aportando un mayor grado de motivación.

Dentro del campo de la comunicación sanitaria, apenas existen aplicaciones de la RV, salvo la herramienta ACORFORMed [1], para capacitar a los médicos para dar malas noticias.

Sí que es posible encontrar aplicaciones para la comunicación en general, cómo simuladores para aprender a hablar en público que centran su capacidad

formativa en la comunicación no verbal como la app CHIARA [2].

2. Objetivo

El objetivo de este trabajo es el desarrollo de una aplicación de Realidad Virtual como complemento para la capacitación en materia de comunicación en crisis sanitarias. Esta iniciativa ha sido promovida por la Cátedra Fundación ASISA de la Universidad de Murcia en el ámbito de la Comunicación Sanitaria coordinada por la Dra. Dña. M. Trinidad Herrero Ezquerro, presidenta de Lyceum de Ciencia y miembro de la Real Academia Nacional de Medicina. El contenido y temática ha sido desarrollado por D. Antonio Garrido Vicedecano de la Facultad de Derecho de la Universidad de Murcia. La Universidad Politécnica de Cartagena ha llevado a cabo la programación de motores gráficos de RV para la configuración de la actividad formativa resultante.



Figura 1. Entidades participativas del proyecto

Este tipo de comunicación posee características particulares que demandan conocimientos específicos y suelen generar un impacto significativo en la sociedad. Se trata de una comunicación dinámica que varía sus características según la fase de la crisis en la que se encuentre, por lo que no son de gran ayuda los clásicos manuales orientados a otras crisis más convencionales. Todo esto hace que la comunicación en crisis sanitarias sea distinta de otros tipos de comunicaciones, y requiera una capacitación diferente.

La estructura del contenido de la aplicación se muestra en la figura 2.

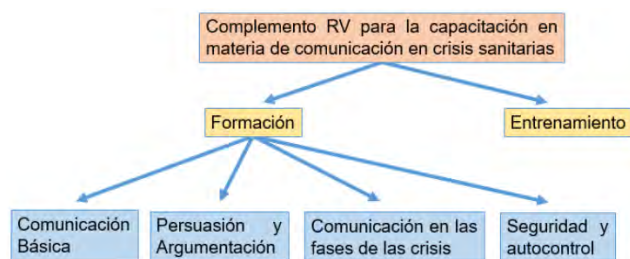


Figura 2. Contenido de la aplicación

Contiene cuatro módulos de formación sobre los aspectos importantes para realizar una correcta comunicación sanitaria. En estos módulos, el usuario, haciendo uso de gafas de realidad virtual, recibe la información de un formador virtual representado por un avatar y puede elegir sobre qué temas quiere obtener información mediante una serie de preguntas.

Por otro lado, la aplicación contiene un módulo de entrenamiento, en el que el usuario realiza una comunicación sanitaria en un entorno virtual, donde los avatares son periodistas que realizan preguntas al portavoz sanitario (el usuario). Y al finalizar genera un feedback con resultados e indicaciones para mejorar.

3. Metodología

Para el desarrollo de esta aplicación son necesarios los elementos mostrados en la figura 3.



Figura 3. Metodología para el desarrollo

3.1. Software de desarrollo VR

Para el desarrollo de la app se ha utilizado Unity3D, herramienta que permite crear contenidos para diversas plataformas digitales (PC, videoconsolas, móviles, etc.). Cuenta con una gran comunidad de usuarios y una documentación muy completa, lo que facilita enormemente su uso. Además, tiene un gran soporte para VR lo que permite conectar gadgets como gafas VR de diferentes marcas y modelos, como OCULUS o HTC VIVE.

3.2. Avatares. Representación de las personas

Un Avatar es una representación gráfica de una persona en entornos digitales.

En contenidos virtuales formativos, los avatares son un recurso didáctico de mucho valor puesto que ayudan al alumno a comprender mejor el contenido y a guiarle por los distintos temas como si de un profesor se tratara,

ayudando a humanizar y propiciando un acercamiento a los alumnos reforzando el mensaje que se pretende transmitir.

Por otro lado, los avatares pueden ser utilizados para simular situaciones reales donde se interactúa con otras personas.

En esta aplicación se utilizan avatares para representar a los formadores y alumnos en los módulos formativos y a los periodistas y cámaras en el módulo de entrenamiento.

Los avatares han sido creados con el software DAZ Studio. Este software tiene un complemento, Face Transfer, que permite transferir la cara de una persona real al avatar a partir de una imagen fotográfica.

Con esta herramienta se han creado los avatares de los formadores, periodistas y alumnos presentes en los escenarios virtuales. En la figura 4 se muestran los avatares de los formadores.

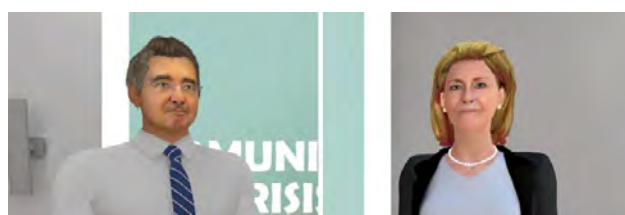


Figura 4. Avatares de los formadores

Además del aspecto visual del avatar, es necesario dotarlo de movimiento para que pueda desarrollar su tarea. Para ello se ha empleado la tecnología MOCAP, que consiste en realizar la captura de movimiento mediante un traje dotado de sensores y el traslado de dicho movimiento a un modelo digital (avatar).

En el desarrollo de esta aplicación hemos utilizado el traje MOCAP Perception Neuron, el cual consta de 32 sensores inerciales.

En la figura 5 se muestra el momento de captura de movimiento del cuerpo mediante MOCAP para posteriormente animar los avatares de los formadores.



Figura 5. Captura de movimiento del formador

Posteriormente estos movimientos son transferidos al avatar del formador de nuestra aplicación.

Aparte del movimiento del cuerpo, necesitamos dotar de vida la cara del avatar. Para ello existe software que mediante una cámara es capaz de rastrear dichos movimientos.

En este caso se ha utilizado la app FaceCap. Esta aplicación hace uso de las cámaras de profundidad real de

dispositivos iPhone para rastrear las diferentes zonas de la cara. La figura 6 muestra el uso de un iPhone con la app FaceCap para capturar los gestos de la cara.

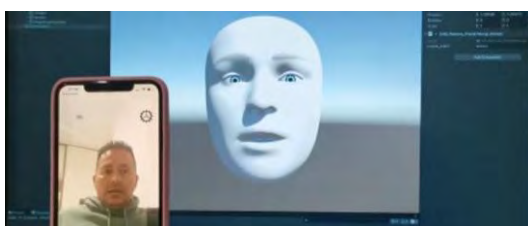


Figura 6. Captura facial con FaceCap

Empleando estas herramientas se han capturado los movimientos de los formadores de la actividad formativa.

3.3. Entorno virtual inmersivo

En una simulación virtual formativa es importante que el entorno o escenario de la simulación sea similar al entorno real donde se realiza la tarea. Se consigue así que la experiencia de la simulación sea inmersiva, el usuario siente la experiencia con todos sus sentidos, mejorando el aprendizaje al añadir estímulos sensoriales y emocionales.

En este trabajo, se ha optado por modelar en 3D los entornos adecuados, utilizando el software Fusion 360 de Autodesk. La aplicación contiene dos entornos virtuales, uno para el módulo de formación y otro para el de simulación. El entorno para los módulos formativos consiste en un aula, que ha sido diseñada desde cero. Este entorno puede verse en la figura 7.



Figura 7. Aula formativa virtual

Por otro lado, el entorno para el módulo de simulación o entrenamiento consiste en una sala de prensa (figura 8). Este escenario se ha modelado tomando como base el salón de grados de la facultad de derecho de la Universidad de Murcia (figura 9).



Figura 8. Sala de prensa virtual



Figura 9. Salón de grados real

En las fotos se puede comparar el grado de realismo alcanzado mediante modelado 3D. Además, mediante este método se puede modificar y añadir mobiliario y detalles de los que puede que no dispongamos en la sala real, convirtiendo así un salón de grados en una sala de prensa.

3.4. Reconocimiento de voz

El reconocimiento de voz es la capacidad de una máquina o programa para identificar palabras y frases en lenguaje hablado y convertirlas a un formato legible por máquina.

Esta tecnología utiliza Machine Learning y Procesamiento del Lenguaje Natural (PLN), para crear chatbots o voicebot, logrando que, en base a ejemplos, sean capaces de detectar lo que el usuario necesita a través del texto o voz y de mantener una conversación con respuestas concretas y coherentes.

La aplicación utiliza esta tecnología para que el usuario pueda interactuar mediante la voz con los bots (avatares de los formadores) y para analizar el contenido semántico de la comunicación sanitaria que realizará el usuario durante su práctica.

La herramienta utilizada es Wit.ai de Meta [3], la cual permite integrar el reconocimiento de voz en diversos software de desarrollo como Unity3D.

En la aplicación web de Wit.ai se entrena la inteligencia artificial para reconocer la intención del usuario a partir de lo que se comunique. Las intenciones del usuario son las divisiones del contenido formativo que éste puede pedir al formador que le informe.

Por ejemplo, una intención es saber la respuesta a la pregunta “¿Son las crisis sanitarias como otras crisis? ¿Se comunica en ellas igual que se hace en otras crisis?”

Se ha creado una pantalla dónde el usuario puede ver las posibles preguntas/temas que puede realizar en el módulo formativo que se encuentre. Esta pantalla está situada sobre la mesa delante del usuario y tiene unos botones interactivos (mediante el uso de los mandos VR) para navegar por las posibles preguntas como muestra la figura 10.

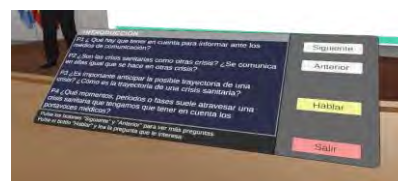


Figura 10. Pantalla con las posibles preguntas/temas

3.5. Feedback

En el módulo de entrenamiento se necesita que la aplicación le diga al usuario si lo está haciendo bien o en qué aspecto de la comunicación puede mejorar. Por lo tanto, debe ser capaz de analizar y evaluar los parámetros no verbales y paraverbales, importantes durante cualquier tipo de comunicación básica.

Respecto a la Comunicación no verbal:

- Gestos con las manos.
- Contacto visual con cada espectador.

Y en cuanto a Comunicación paraverbal:

- Silencios durante el discurso para resaltar una idea.
- Ritmo de palabras por minuto.
- Volumen.

Para analizarlos se utilizan los inputs de los gadgets de RV. Rastreo de posición de las gafas y manos, y el micrófono.

3.6. Contenido formativo

El objetivo de esta aplicación es formar. Para ello es necesario generar un contenido formativo que la aplicación pueda transmitir al usuario. La manera de transmitir esta información es mediante los formadores virtuales, los cuales informan al usuario sobre los temas que él solicita mediante su voz. El contenido formativo ha sido desarrollado por los Doctores Dña. María Trinidad Herrero y D. Antonio Garrido, expertos en comunicación sanitaria.

El contenido formativo trata los siguientes módulos:

- Módulo 1. Comunicación Básica: Este módulo incorpora formación a los portavoces sanitarios en lenguaje no verbal, paraverbal y en telegénia e imagen.
- Módulo 2. Persuasión y Argumentación: Este módulo incorpora formación a los portavoces sanitarios en discurso persuasivo y argumentación.
- Módulo 3. Comunicación en las fases de las crisis: Este módulo incorpora formación a los portavoces sanitarios en la comunicación desde una perspectiva dinámica, según la fase de la crisis en que se encuentren: fase de latencia, fase de eclosión, fase de explosión o fase de síntesis.
- Módulo 4. Seguridad y autocontrol: Este módulo incorpora formación a los portavoces sanitarios en seguridad y autocontrol mental, reducción de la ansiedad y confianza en sí mismo para hacer frente a situaciones crispadas y muy fluidas y dinámicas.

4. Resultados

Se han creado los avatares (formadores, alumnos y periodistas) y entornos virtuales necesarios para la aplicación

Además, se han desarrollado los contenidos formativos gracias a la colaboración con expertos en la materia. En

las figuras 11 y 12 se muestran algunas capturas de los módulos formativos.



Figura 11. Captura del contenido formativo P2



Figura 12. Captura del contenido formativo P5

En el módulo de entrenamiento se ha programado la parte del análisis del entrenamiento y feedback.

En la figura 13 se muestra el menú dónde el usuario puede ver los resultados de sus entrenamientos.



Figura 13. Menú de resultados

5. Conclusiones

La realidad virtual es una tecnología muy interesante para el desarrollo de herramientas formativas ya que aporta un alto grado de estímulos sensoriales que enriquece la experiencia del alumno. Además, permite simular situaciones de la vida real el número de veces que sea necesario para mejorar capacidades y aptitudes frente a ellas.

6. Referencias bibliográficas

- [1] Aplicación ACORFORMed. Disponible en web: <http://www2.lpl-aix.fr/~acorformed/index.html>
- [2] Aplicación CHIARA. Disponible en web: <https://bechiara.com/es/>
- [3] Herramienta NLP Wit.ai de Meta. Disponible en web: <https://wit.ai/>

ÍNDICE DE AUTORES

A			
Abella, Monica	266, 540	Armañac Julián, Pablo	545
Abellón Caselles, Marco	614	Ayala Fernandez, Unai	222
Abizeid Daou, Roy	650	Azagra, Marc	574
Acha, Begoña	476, 492, 496	Azkue, Laiene	222
Aguado Del Hoyo, Alejandra	39	Azorín, José	381, 385, 504
Aguado Sarrió, Eric	290		662
Aguilar García, María Dolores	553	B	
Aguilera Jiménez, Elena	402	Badosa, María Carmen	460
Agulló Mertens, Elena	214	Bailón, Raquel	18, 343, 545
Alandí Rocafull, Irene	22	Balsa, Alejandro	85
Alberola López, Carlos	274, 306, 553	Barceló Pons, Pol	630
Albusac Jiménez, Javier Alonso	157, 456	Barriga Rivera, Alejandro	488, 512
Alcaide León, César	140	Barrios Álvarez De Arcaya, Javier	111, 231, 452
Alcaraz, Jordi	190	Barro, Óscar	427
Alcaraz, Raúl	10, 235, 247	Barroso García, Verónica	549
Alcázar Laynez, Manuel Luís	488	Barroso Gonçalves, Sergio	650
Alcobendas Maestro, Mónica	456	Bartrés Faz, David	520, 536
Alcón, Clara	622	Bastarrika, Gorka	282
Alió Sanz, Jorge	368	Basterra García, Álvaro	585
Allegue, José Manuel	310	Bastir, Markus	115, 251, 254
Alonso, Erik	435	Batlle, Montserrat	532
Alonso González, Cristina	165	Bayo, Fernando	294
Alonso Linaje, Gonzalo	161	Bello Arroyo, Elisabet	626
Álvarez, Zaida	580	Beltran Carrero, Alberto	356
Álvarez Gonzalez, Daniel	549	Benamara, Megdouda	574
Álvarez López, Aroa	576	Benavent, Diego	85
Amado Caballero, Patricia	553	Berjano Zanón, Enrique	123, 131
Ambrosiani Fernández, Jesús	165	Bernabéu Wittel, José	492
André, Jean-Marc	389	Bernat, Lluís	243
Aniorte García, Alicia	218	Blanch, Lluís	528, 532
Antolí Francés, Antonio	102	Blanco Coloma, Laura	157, 456
Aramendi, Elisabete	435	Blasco García, Jesús Damián	68
Arbelo, Elena	227	Boil, Gregory	618
Arce Diego, José Luís	127, 364, 448	Borrego, Miguel	476
Arevalo Núñez De Arenas, Eduardo	136, 638	Boutinguiza, Mohamed	136
Arias González, Felipe	136, 427	Bragós Bardia, Ramón	322
Armada Maresca, Félix	634	Brooks, Claire J.	294

Bru Orts, David	214	Corcoy, Rosa	81
Burgos Olmos, Manuel Antonio	115, 251, 254	Corominas, Josép	622
C		Corpas Del Moral, Teresa	286
Caballero, Daniel	347, 557	Cotty, Rory	360
Cabello Murillo, Eduardo	35	Coupé, Perrick	43, 47
Cadena Santoyo, Manuel	356	Crespo Sedano, Andrea	553
Callejón Leblic, M ^a Amparo	165	Crestelo, Fernando	618
Calvillo Arbizu, Jorge	73	Cubero, Lucía	27, 39
Calvo De Mora, María	302	Cubero, Lucía	372
Calvo Haro, José Antonio	402	Cuenca Dacal, Luís	131
Campello, Juan	60	Cummins, Nicholas	343
Camps Herrero, Julia	290	Curila, Karol	6
Carbone, Enrico	646	D	
Carreras Vidal, Lourdes	190	De Francisco Ortiz, Óscar	610
Carrillo López, Laura	314	De Gregorio Bermejo, Carla	56, 174
Casado, Rubén	239	De Haro, Candelaria	528, 532
Casanova Batlle, Enric	178	De La Calle García, Beatriz	56, 174
Casaseca De La Higuera, Pablo	553	De La Rica Martínez, Alba	106
Castañeda, Carlos E.	262	De León Luís, Juan	27, 443
Castaño Linares, Óscar	570, 580, 630	De Los Reyes Guzmán, Ana	94, 157, 456
Castells Ramón, Francisco	239		508
Castells Ramón, Francisco	589	De Ramón Fernandez, Alberto	318
Castro Blas, Javier	540	Del Amor Del Amor, María Rocío	352, 360, 666
Castro Domínguez, Cristina	562		670
Castro Leal, Gonzalo	540	Delicado Correa, Gloria Del Rocío	282
Cattaneo, Gabriele	520, 536	Demidova, Marina	2
Cavas Martínez, Francisco	368	Desco, Manuel	266, 540, 626
Ceruelo Abajo, Silvia	157, 456	Díaz Gutierrez, Nuria	356
Chausa, Paloma	520, 536	Díaz Lantada, Andrés	198
Chiva De Agustín, Miguel	302	Díaz Pineda, Jaime	389
Cinesi Gómez, César	414, 418	Díaz Serrano, Pamela	31
Cirillo, Giovanni	634	Dobos, Agnes	630
Cisnal, Ana	161	Dobson, Richard	343
Ciurana Gay, Joaquim	178	Domingo Díez, Javier	140, 144
Cobo, Antonio	81	Domingo Díez, Javier	144
Colchero, Luís	576	Dong, Yiming	630
Collavini, Santiago	123	Dubois, Marc	574
Colomer Granero, Adrián	666		
Company Se, Georgina	322	E	
Conejero Rodilla, Andrés	22	Eijsvogels, Thijs	18
Conesa Carrillo, Miguel Angel	314	Elena, Mar	488
Conesa Celdrán, Agustín	106	Engel López, Elisabeth	630

Escribano Cano, Pilar	10	García, Esther	343
España, Goretti	520	García, Manuel	10
Essamhi Saidi, Kamilia	674	García Fernández, Ignacio	231, 452
Esteban Ortega, Francisco	115, 251, 254	García Fuentes, Marta	406, 614, 468
Estevez Labori, Francisco	119		674
		García Gadañón, María	480
F		García Gómez, Gabriel J.	60, 243
Fabregat Sanjuan, Albert	202	García Linares, José Antonio	73
Fallon, Matthew	574	García Lizarribar, Andrea	566, 622
Fanjul Vélez, Félix	127, 364, 448	García López, Gabriel	68
Farré, Ramón	190	García Martínez, Javier	650
Feliu Batlle, Jorge Juan	68	García Mesa, Pablo	356
Fernandes, Francisco	650	García Ruano, Agustín	553
Fernández, Fernando	81	García Rueda, Isabel María	115, 254
Fernández, Javier	94	García Sáez, Gema	258
Fernández, Rafael	528, 532	García Sánchez, Laura	52
Fernández Arias, Mónica	136, 427	García Torres, Fernando	360
Fernández Corazza, Mariano	123	García Vicente, Clara	14, 549
Fernandez Cortes, Elena	626	Garmendia Aguilar, María Dolores	553
Fernández Costa, Juan Manuel	194	Garmendia Leiza, José Ramón	553
Fernández Del Cerro, Carlos	266	Garnés Camarena, Óscar	331
Fernandez Del Río, José Antonio	570	Garrido, José Francisco	310
Fernández Gonzalo, Sol	528	Garrido Rubia, Antonio	686
Fernandez Martínez De Septien, Carmen	553	Gavara, Nuria	190
Fernández Rodríguez, Álvaro	389, 393	Ghosh, Subrata	360
Fernandez Sánchez, M. Teresa	464	Gijón Ortego, Jorge	585
Ferré Torres, Josép	630	Gil, Eduardo	545
Ferrer Riquelme, Alberto	290	Gil, Javier	136, 427
Figuere Memba Massoko, Ludovic	102	Gil Agudo, Ángel	508
Filgueiras Rama, David	306, 431	Gil Correa, Amalia	654
Flores, Héctor	414, 418, 422	Giraldo, Beatriz	194
	614	Godoy González, Marta	528
Fraille, Juan Carlos	161	Goicolea, José María	52
Franco Terriza, Lucía	488	Gomà, Gemma	528
Frías, Marcos	460	Gómez, Carlos	585, 593, 597
Fuster Barceló, Caterina	282		654
		Gómez, Neurys	2
G		Gómez Aguilera, Enrique J	35, 77, 85, 98
Galera, Cristina	314		149, 278, 286
Gallego Ferrer, Gloria	210		298, 302, 520
Gallo, Juan	439		536
Ganoza Quintana, José Luís	364	Gómez De Los Infantes Peña, Susana	402
Garate, Francisco	77	Gómez Fernandez, Cristina	610

Gómez García, Carmelo	368	Herrera Valenzuela, Diana	508
Gómez Gaviro, María Victoria	626	Herrero, Alba	574
Gómez Grande, Adolfo	298	Herrero Ezquerro, María Trinidad	68, 310, 686
Gómez Pesquera, Estefanía	642	Herrero Quevedo, Fernando	372
Gómez Pilar, Javier	642	Herrero Tudela, María	468, 480
Gómez Valverde, Juan José	356	Hietakoste, Salla	545
Gomis, Pedro	227	Hornero, Roberto	14, 397, 468
Gomis Tena, Julio	327		480, 500, 545
Gonzalez, Victor M.	464		549, 585, 593
González, Cintia	81		597, 642, 654
González, Iván	496	Hoshi, Hideyuki	654
González Cabrero, José	368	Hotopf, Matthew	343
Gonzalez Costea, Belen	77	Huerta Herraiz, Álvaro	235
González Domínguez, Ramón	182		
González Expósito, Sara	508	I	
González Morcillo, Carlos	157, 456	Iacucci, Marietta	360
González Nieto, Daniel	576, 634	Iañez, Eduardo	381, 385, 504
González Nuño, Francisco M.	516		662
González Suárez, Ana	131, 119	Ibanez, Mario	670
Gonzalez Vazquez, Joaquín J.	243	Ibarra Berrocal, Isidro J.	169, 674, 678
Gozal, David	14, 545, 549		682
Gracia, Desirée	504, 662	Iniesta, José	85
Guedán Durán, Atocha	634	Irastorza, Ramiro	123
Guerrero, Ángel Luís	593	Isabel Roquero, Alba	227
Guijarro Martínez, Paula	27	Isasi, Iraia	435
Guinea Tortuero, Gustavo Víctor	562, 576, 634		
Gutiérrez, Álvaro	646, 658	J	
Gutiérrez, Marco	536	Jacinto, Misael	262
Gutierrez De Pablo, Laura	642	Jaén Vargas, Milagros	650
Gutiérrez De Pablo, Víctor	585, 593	Jara, Carlos A.	60
Gutiérrez Tobal, Gonzalo César	14, 468, 545	Javierre, Miguel	18
	549	Jelic, Vesna	597
Guzman Miranda, Hipólito	512	Jemni Damer, Nahla	634
H		Jiménez Fernandez, Pablo	512
Halamek, Joséf	6	Jiménez García, Jorge	14
Haro, Josép María	343	Jiménez Mallebrera, Cecilia	460
Hayes, Brian	360	Jiménez Muñoz, Marta	68
Hernandez, Alicia	77	Jin, Shana	570
Hernández Abad, Ignacio	35	Jódar Reverte, María	314
Hernandez Machado, Aurora	630	Juan Poveda, Javier	381, 385
Hernando Pérez, María Elena	81, 258	Juárez, Sergio	94
Herrera Tirado, Vanesa	157, 456	Junquera, Isabel	327
		Jurak, Pavel	6

K		López García, Luís	102
Kainulainen, Samu	545	López García, Marcos	89
Kappel, Mathilde	347, 516, 557	López González, Verónica	443
Kerexeta Sarriegi, Jon	222	López Guerrero, José Antonio	352
Kloosterman, Manon	18	López Martínez, María José	566
Kontaxakis, Giorgos	31, 377	López Olañeta, Marina	626
Kontaxis, Spyridon	343	López Riquelme, Juan Antonio	68
Kovalyk Borodyak, Oleksandr	472, 484	López Santiago, Javier	31
Kumar Srivastava, Girish	634	Lozano Berrio, Vicente	157, 456
		Lozano Ibáñez, Miguel	231, 452
L		Ludusan, Cristina	406
Laguna, Pablo	2, 545, 606	Luque Del Toro, Ivan	186
Lamers, Femke	343	Lusquiños, Fernando	427
Lara Pezzi, Enrique	626		
Larburu Rubio, Nekane	222	M	
Largo Aramburu, Carlota	634	Machado, Isidro	352
Larraga, Blanca	646, 658	Macian Morales, Álvaro	169, 678, 682
Larrea, Andima	435		686
Larrén, Estela	402	Magalhaes, Andreia	439
Larrosa Capacés, Susana	202	Malanda Trigueros, Armando	331, 335
Lax, Ana María	314	Malpica, Norberto	270
Lázaro, Jesús	545	Manero, José María	427
Lazo Maestre, Manuel	165	Manjón Herrera, José Vicente	43, 47
Leante Castellanos, José Luís	610	Mansilla Gil, Jorge	165
Leñero Bardallo, Juan Antonio	492	Marco Rius, Irene	574
Letouzé, Théodore	389	Marcos, Diego	397, 500
Leva, Álvaro	227	Marcos Jorquera, Diego	318
Liberos, Alejandro	111, 231, 452	Marínez Rodrigo, Arturo	235
Licona Beltrán, Alisson	258	Mariscal, Cristina	335
Llinas, María Teresa	310	Martin, Alba	2, 343
Llombart Bosch, Antonio	352	Martín, Mercedes	377
Llorente, Vicente	626	Martín Fernandez, Marcos	274
Lloveras Borràs, Berta	570	Martín Montero, Adrián	14, 545
Lobato Gómez, Carlos	153	Martín Pinacho, Jesús Javier	278
López, Ana	77	Martin Sierra, Pablo	339
López, María José	580	Martínez, Juan Pablo	2, 6, 606
López, Maribel	468, 480	Martínez, Sofia	634
López Aguilar, Joséfina	528, 532	Martínez Abellan, Lucia	406
López Ales, Emilio	274	Martínez Blanco, África	190
López Canosa, Adrian	630	Martínez Cagigal, Víctor	500
López Dolado, Elisa	94	Martínez De Juan, José Luís	22, 327
López Expósito, Patricio	294	Martínez Mateu, Laura	431
López Fanarraga, Mónica	622	Martínez Muñoz, Gemma	106

Martínez Rodrigo, Arturo	247	Noguera Monteagudo, Adrià	570, 630
Martínez Sánchez De La Torre, Rubén	385	Novella Navarro, Marta	85
Martínez Sober, Marcelino	102	Novelli, Antonello	464
Martínez Zarzuela, Mario	56, 174	Núñez, Olivier	81
Masè, Michela	10		
Masias Bruns, Mireia	186	O	
Matcham, Faith	343	Oates, Martin John	106
Maturana, Aarón	585, 597	Ochandiano, Santiago	39
Mediavilla Santos, Lydia	402	Oddo, Silvia	123
Medrano Fernandez, Iria	206, 638	Ojados Gonzalez, Lola	169, 218, 406,
Menchón Lara, Rosa María	274		414, 418, 422
Méndez, Inmaculada	601		614, 674, 678
Merino Caviedes, Susana	306		682, 686
Meseguer, Pablo	352, 360	Olivas Vergara, Otto	372
Milara, Eva	298, 302	Oliveira Barroso, Filipe	508
Millet, José	239, 589	Olmo Fajardo, Tania	508
Miras Cabrera, Baltasar	106	Orozco López, Onofre	262
Molina Oviedo, Angie Katherine	206, 638	Ors Quixal, Rafael Teodoro	239, 589
Morales Loro, David	198	Ortega, José María	266
Morales Sánchez, Juan	472, 484	Ortiz, Mario	381, 385, 504
Morell, Sergio	43, 47		662
Moreno, Juan C.	508	Ortiz Abellán, Carmen	290
Moreno Blanco, Diego	98	Ortiz Zaragoza, Francisco	64
Moreno Montes, Eva María	377	Ortiz Zaragoza, Francisco José	601
Moscardó I Chafer, Isabel	102	Ortuno, Juan	31
Muñiz Albaiceta, Guillermo	528	Ozturk, Berke	464
Muñoz Barrutia, Arrate	282		
Murias, Gastón	528	P	
Muzo, David Chushig	182	Pace Bedetti, Horacio M.	22
		Pacurar, Razvan	406
N		Padovano, Daniele	247
Naranjo, Valery	352, 360	Pajares Ruiz, Virginia	322
Naranjo Hernández, David	89	Pájaro, Javier	520
Naranjo Ornedo, Valeriana	666, 670	Palacios, Saul	6
Narayan, Vaibhav	343	Palacios Gutiérrez, Tomás	198
Naredo Sánchez, Esperanza	372	Palazón Palau, David	484
Navallas Irujo, Javier	335	Palmieri, Flavio	227
Navarra Ventura, Guillem	528	Pamies Vilà, Rosa	202
Navarro, Samuel	352	Panetsos, Fivos	562, 634
Navas, Pedro	52	Panetsos Petrova, Fivos	31
Nescolarde Selva, Lexa	322	Papa, Michele	634
Nguyen, Uyen Chau	6	Pardo, Rebeca	377
Noé, Sergio	190	Pareja Jiménez, Marta	178

Parras Burgos, Dolores	64, 106	Portela, Alejandro	574
Pascau, Javier	27, 39, 372	Pose Díez De La Lastra, Alicia	443
	443	Pou, Juan	136, 427
Pascual, Ana	580	Pou Álvarez, Pablo	136
Pascual, Teresa	662	Poza, Jesús	585, 593, 597
Pascual Leone, Álvaro	520, 536		654
Pascual Rubio, Vicenç	202	Pozo, Miguel Ángel	377
Pastor, José Manuel	247	Pozo Rodríguez, Gustavo	35
Pastor García, Francisco	318	Prats Boluda, Gema	22, 327
Patel, Hetal	574	Prats Montalbán, Josémanuel	290
Pavón Pulido, Nieves	68	Prinzen, F.	6
Payo Gutiérrez, Ismael	94	Puche García, Violeta	431
Pedraja, Juan	286	Puerta, Paula	464
Pedrerá Jiménez, Miguel	258	Pueyo, Esther	6, 18, 606
Peng, Xiaofeng	566	Puig Miquel, Teresa	178
Penninx, Brenda	343	Punset, Miquel	427
Perales Gómez, Raquel	157, 456		
Pérez, Beatriz	81	Q	
Pérez, Cristina	606	Quiles, Vicente	662
Pérez, Juan José	123	Quintana, Manuel	98
Pérez, Sofía	343	Quintana Prego, Ana María	634
Pérez Benito, David	294	Quiñonero López, Gema	439
Pérez Carrasco, José-Antonio	492		
Pérez Fernández, Nicolás	282	R	
Pérez Jiménez, Sandra	98	Rafecas Codern, Albert	322
Pérez Mañanes, Rubén	402	Ramirez, Elisa	239, 589
Pérez Martínez, Juan José	131	Ramírez, Julia	2
Pérez Martos, Salvador	314	Ramón, Javier	574
Pérez Ramos, Alejandro	115, 251, 254	Ramón, José L.	60
Pérez Rigueiro, José	562, 576, 634	Ramón Azcón, Javier	194
Pérez Salazar, Manuel J.	347, 516, 557	Ramos Gómez, Milagros	140, 144, 562
Pérez Turiel, Javier	161	Ravelli, Flavia	10
Pérez Velasco, Sergio	397, 500, 654	Raya, Rafael	339
Picazo Selva, Lucia	115, 254	Reina Tosina, Javier	73, 89, 165
Piella Fenoy, Gemma	186	Rejas Llamera, Adrián	593
Plasencia Rodríguez, Chamaida	85	Reverte Ribo, Virginia	310
Platonov, Pyotr	2	Reyes Leiva, Karla	650
Plesinger, Filip	6	Rieta, José J	10, 235, 247
Pollán, Marina	81	Rigo Vidal, Agnes	202
Polo Hortiguera, Cristina	381	Rio Álvarez, Angel	464
Polonio Alcalá, Emma	178	Riu Costa, Pere Joan	322
Pons Andrés, Neus	210	Rius Costa, Vanesa	202
Porta, José María	460	Riveiro, Antonio	136

Roca, Oriol	528	Ruiz, Verónica	658
Roca González, Joaquín	64, 310, 406	Ruiz Canales, Antonio	106
	414, 418, 601	Ruiz Fernández, Daniel	318, 524
	614, 618	Ruiz Gávez, Rubén	397
Rocón, Eduardo	646	Ruiz Gutiérrez, Martín	194
Ródenas, Juan	10	Ruiz Pérez, Marina	43, 47
Rodrigo Bort, Miguel	111, 231, 452	Ruiz Vitte, Ainhoa	646
Rodríguez, Daniel	136		
Rodríguez Ballabriga, Ángel	202	S	
Rodríguez Contreras, Alejandra	427	Sabuco García, Yolanda	524
Rodríguez Falces, Javier	335	Sáez, Gema	81
Rodríguez Galván, Justino	306	Saiz, Javier	327, 431
Rodríguez Gonzalez, Blanca	270	Salaberria, Ruth	435
Rodríguez González, Víctor	500, 654	Salamanca Gonzalez, Celina	206
Rodríguez Herrero, Agustín	262	Salas Monedero, Miriam	508
Rodríguez Matesanz, Manuel	149	Salazar Carrasco, Lucía	153
Rodríguez Robles, Francisco	484	Salcedo Eugenio, Gustavo	169, 218, 422, 674, 678, 682
Rodríguez Rodríguez, Myriam	402		
Rodríguez Rosa, David	94	Samitier, Josép	439, 566, 622
Rodríguez Trujillo, Romen	570, 630	Sánchez, Patricia	35, 77, 98, 149
Rodríguez Vila, Borja	270		278, 286, 298
Rojas Lozano, Pilar	356		302
Rojo, Francisco	634	Sánchez García, Daniel	27
Roldán, Mònica	460	Sánchez Gómez, Serafín	165
Román Martínez, Isabel	73	Sánchez González, Javier	306
Romero, Cristina	243	Sánchez González, Patricia	520, 536
Romero, Lucía	431	Sánchez López De Pablo, Cristina	339
Romero, Roberto	468, 480	Sánchez Margallo, Francisco Miguel	153, 347, 516
Romero De Antonio, Pau	231, 452		557
Romero Olóriz, Carlos	488	Sánchez Margallo, Juan A.	153, 347, 516
Romitti, Giada Sira	111, 231, 452		557
Romo Marín, Ángela	56, 174	Sánchez Nanclares, Gorka	314
Ron Angevin, Ricardo	389, 393	Sancho Gómez, José Luís	472
Ropero Romero, Francisco	165	Sanderson, Daniel	540
Roquette, Matilde	570	Santacroce, Giovanni	360
Rosell Ferrer, Javier	322	Santamaría Vázquez, Eduardo	397, 500
Rozalen, Gemma	278	Santos Mayo, Carlos	39
Rubio, Ana	352	Santos Pulpón, Verónica	528, 532
Rubio Bolívar, Javier	98	Sanz, Amaia	435
Rubio Mayo, Paula	258	Sanz Martí, Concha	102
Ruipérez Campillo, Samuel	239, 589	Sanz Prieto, Daniel	115, 251, 254
Ruiz, Cristina Soguero	182	Sarandeses, Pilar	298
Ruiz, Emma	81	Sarlabous, Leonardo	528, 532

Scotlandi, Katia	352	Termenón Rivas, María	111, 231, 452
Sebastian Aguilar, Rafael	111, 231, 452	Toledo Pastrana, Tomás	496
Seiffert, Alexander P.	35, 278, 286	Tolosa Pardo, Laia	210
Sellés Navarro, Inmaculada	472	Tormos Muñoz, José María	520, 536
Serna, Juan P.	601	Tornero López, Jesús	385
Serra Almor, Dolors	231, 452	Torrado Carvajal, Angel	270
Serrano, Carmen	476, 492, 496	Torricelli, Diego	508
Serrano Gotarredona, Amalia	496	Tortosa, Luís	227
Serrano Olmedo, José Javier	140, 144, 650	Tristán Vega, Antonio	306
Serrano Pertierra, Esther	464	Trujillo Guerrero, María Fernanda	650
Setién Doderó, Fernando	127	Tsiogkas, Nikolaos	601
Shigihara, Yoshihito	654		
Siddi, Sara	343	U	
Silva, Francisca	496	Ubeda, Andrés	60, 243
Simmross Wattenberg, Federico	274	Urendes Jiménez, Eloy	339
Simoés Lopes, Daniel	650		
Simón Martínez, Cristina	56, 174	V	
Sinovas Alonso, Isabel	508	Vairavan, Srinivasan	343
Smailovic, Una	597	Valle, César	331
Smisek, Radovan	6	Valls Lozano, Óscar	666
Solana, Javier	520, 536	Van Dam, Peter	18
Soldado Vilches, Noemi	634	Van Durme, Bo	630
Somolinos Simon, Francisco Javier	258	Van Vlierberghe, Sandra	630
Soria Sánchez, Paula	638	Vaquerizo Villar, Fernando	549
Soriano Segura, Paula	504	Vaquero, Juan José	294
Sornichero Martínez, Javier	368	Veganzones, Miguel	161
Sornmo, Leif	606	Velasco Álvarez, Francisco	389, 393
Sorrentino, Ilaria	206, 638	Vera Carrillo, David	610
Sosa Santos, Rene Fernando	448	Vera Repullo, José Alfonso	64, 601
Souiade, Lilia	140, 144	Verdu Monedero, Rafael	472, 484
Sousa, Ines	580	Vilas Iglesias, Ana M.	136
Spagnolo, Emanuela	77	Vilche, Anna	580
Stashuk, Daniel	331	Vilche Mariscal, Anna	570
Suardiaz, Juan	618	Villanueva Bádenas, Estela	210
Suárez Vega, Víctor Manuel	282	Villar, José R.	464
Subirà, Carles	532	Villasante, Aranzazu	439, 566, 622
		Virtuoso, Assunta	634
T		Vizcaino Martin, Francisco Javier	393
Tabraue Rubio, Raquel	576		
Tamayo, Eduardo	642	W	
Tapia Galisteo, José	81, 258	Warnants, Ismael	601
Tavares Silva, Miguel	650	Weerakonda, Dian	574
Tejedera Villafranca, Ainoa	194		

X

Xifra Porxas, Alba 528, 532

Xuriguera Martin, Elena 580, 630

Y

Yanac, Eduardo 570

Yeste, José 574

Z

Zammarchi, Irene 360



CASEIB 2023 XLI

Congreso Anual de la Sociedad Española de

Ingeniería Biomédica

Contribuyendo a la salud basada en valor



ISBN: 978-84-17853-76-1



9 788417 853761