CASEIB 2020

LIBRO DE ACTAS

XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica

ORGANIZADORES Y COLABORADORES:

GRUPO DE INGENIERÍA BIOMÉDICA
UNIVERSIDAD DE VALLADOLID

BioRes
BIOENGINEERING AND RESUSCITATION

Sociedad Española de Ingeniería Biomédica

CASEIB
2020
XXXVIII
Congreso Anual de la Sociedad Española de Ingeniería Biomédica

LIBRO DE ACTAS

25 - 27 de noviembre de 2020

XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica CASEIB 2020

Los contenidos de 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/2020/*

Presidente del Comité Organizador:
Roberto Hornero Sánchez

Secretario del Comité Organizador:
Jesús Poza Crespo

Editores:
Roberto Hornero Sánchez
Jesús Poza Crespo
Carlos Gómez Peña
María García Gadañón

Maquetación:
Carlos Gómez Peña
Unai Irusta Zarandona
Verónica Barroso García
Aarón Maturana Candelas

Diseño de portada:
Pablo Núñez Novo

ISBN: 978-84-09-25491-0



Libro de Actas del XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica
Se distribuye bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional.

ORGANIZADORES:

GRUPO DE INGENIERÍA BIOMÉDICA
UNIVERSIDAD DE VALLADOLID

COLABORADORES:

BioRes
BIOENGINEERING
AND RESUSCITATION

SociedadEspañola de
INGENIERIABIOMEDICA

PATROCINADOR:

Bionic®
FUNDADA EN 1983

EMPRESAS COLABORADORAS CON LA SEIB:

vicomtech
MEMBER OF BASQUE RESEARCH & TECHNOLOGY ALLIANCE

THEMEL

MERCÉ V. ELECTROMEDICINA

Surgicen
Surgical Engineering

quibim

tecnalia Inspiring Business

DAWAKO®

TEQUIR
I+D+i+u

Medtronic

AVALADO POR:

International Federation for Medical and Biological Engineering
IFMBE

European Alliance for
EAMBES
Medical and Biological Engineering & Science

BIENVENIDA

Estimados congresistas,

En primer lugar, deseamos que todos estéis bien en esta situación excepcional causada por la pandemia de COVID-19. En nombre del Comité Organizador, es un honor para nosotros daros la bienvenida al XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (**CASEIB 2020**). Debido a las circunstancias extraordinarias que estamos viviendo, esta nueva edición de **CASEIB 2020** se celebra del 25 al 27 de noviembre de 2020 de manera virtual. Esta edición del congreso está organizada por el Grupo de Ingeniería Biomédica de la Universidad de Valladolid, con la colaboración del grupo de investigación BioRes (*Bioengineering and Resuscitation*) de la Universidad del País Vasco y de la Sociedad Española de Ingeniería Biomédica (SEIB).

CASEIB 2020 será la trigésimo octava edición de este congreso, lo que significa que es una reunión científica completamente consolidada en los ámbitos científico, académico y profesional de esta disciplina. Este congreso es un foro de encuentro de científicos, profesionales de la industria, ingenieros, clínicos y profesionales biomédicos interesados en conocer, analizar y debatir los temas más actuales en investigación, educación, aplicación industrial y clínica de la Ingeniería Biomédica. La organización de este congreso 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 2020 pretende ser un escaparate del estado actual de la Ingeniería Biomédica desde el punto de vista científico, al tiempo que un foro propicio para profundizar en las relaciones entre los distintos grupos de investigación de esta disciplina y para fortalecer la cohesión entre los diversos colectivos de personas que se dedican a la Ingeniería Biomédica. Somos conscientes de la importancia de la transferencia de los resultados de investigación, por lo que se acercará a la industria y al sector sanitario como actores fundamentales para poner en valor los avances en investigación básica. De manera adicional, este año hemos organizado una mesa redonda relacionada con la COVID-19, que tan presente está ahora en nuestras vidas. Asimismo, se han seguido promocionando los diversos concursos destinados a alumnos de Grado, Máster y Doctorado para fomentar la participación de jóvenes investigadores. Es nuestra intención que esta edición virtual de **CASEIB 2020**, a pesar de la distancia, siga siendo nuestro lugar de encuentro y el congreso de referencia de la Ingeniería Biomédica en España.

En esta nueva edición de CASEIB hemos programado una serie de actividades científicas que esperamos cumplan con vuestras expectativas. Con el objetivo de que el evento sea un lugar de enriquecimiento profesional, **CASEIB 2020** presenta los siguientes actos de interés:

- ✓ Conferencia Inaugural del Congreso a cargo de Prof. David Gozal, *Professor at the University of Missouri, Marie M. and Harry L. Smith Endowed Chair of Child Health*, y *Physician-in-Chief at the University of Missouri Women's and Children's Hospital* (EE.UU.), con el título "*Will Bioengineers Lead the Path to Solution(s) in Sleep Disorders?*".

- ✓ Conferencia de Clausura del Congreso a cargo del Prof. Álvaro Pascual-Leone, *Professor of Neurology at Harvard Medical School, Director of the Berenson-Allen Center for Noninvasive Brain Stimulation at Beth Israel Deaconess Medical Center and Harvard Medical School*, y

Program Director of the Harvard-Thorndike Clinical Research Center, con el título "*A lifelong Vital Sign of Brain Health*".

- ✓ Mesa Redonda titulada "Ingeniería Biomédica y COVID-19", en la que se analizará el papel de la Ingeniería Biomédica y la Tecnología Sanitaria en la crisis de la COVID-19.
- ✓ Mesa Redonda titulada "Desafíos y oportunidades de la Ingeniería Biomédica frente a los nuevos retos de la salud", que abordará desde varios enfoques (gestión, investigación, industria y educación) el momento actual de la Ingeniería Biomédica y sus oportunidades.
- ✓ Asamblea General de la Sociedad Española de Ingeniería Biomédica, donde se debatirá sobre las futuras actividades y objetivos de la Sociedad. Además, este año se elegirá a una nueva Junta Directiva.
- ✓ Premio José María Ferrero Corral, donde se presentarán los seis mejores trabajos realizados por estudiantes de Máster y Doctorado.
- ✓ Competición de estudiantes de Grado en Ingeniería Biomédica, donde se presentarán los mejores nueve trabajos de estudiantes de estos Grados en diferentes Universidades españolas.

El proyecto científico se completa con una amplia muestra de las líneas de trabajo desarrolladas por los investigadores en Ingeniería Biomédica. Se han aceptado 131 ponencias, revisadas por el Comité Científico y repartidas entre las diferentes áreas temáticas del congreso. De ellas, 84 se presentan como comunicaciones orales y 47 en formato de póster.

- ✓ Comunicaciones orales. Se han planificado 15 sesiones orales organizadas de forma paralela en 5 bloques de 3 sesiones cada uno. En concreto, cuatro sesiones se dedican a Señales Biomédicas, dos a Imágenes Biomédicas, dos a Sistemas de Ayuda a la Decisión Clínica, una a Bioinstrumentación, una a Biomateriales, una a Ingeniería de Rehabilitación, una a Informática Biomédica, una a Modelado de Sistemas Biomédicos, una a Simulación y Planificación Quirúrgica, y una a Telemedicina y e-Salud.
- ✓ Comunicaciones tipo póster. Se ha organizado una sesión de pósteres en la que se presentarán un total de 47 comunicaciones mediante una presentación pregrabada en vídeo; durante esta sesión se podrá interaccionar con los autores de manera directa.

Al igual que en anteriores ediciones de CASEIB, los mejores trabajos del congreso se publicarán en un número especial de la revista de acceso abierto *Entropy*. Gracias a ello, los trabajos seleccionados serán promocionados a una audiencia más amplia y actuarán como representantes del trabajo que se está realizando en España dentro de la Ingeniería Biomédica.

Finalmente, queremos agradecer a todas las personas, instituciones y empresas que han facilitado y propiciado la realización de **CASEIB 2020**. En particular queremos destacar el esfuerzo y dedicación de los autores en la preparación de sus ponencias, reconocer a los miembros del Comité Científico su buen hacer en la revisión de los trabajos científicos, y agradecer a los Conferenciantes invitados y Ponentes de las Mesas Redondas por su generosidad a la hora de compartir con nosotros su experiencia. Asimismo, quisiéramos dar las gracias a la Sociedad Española de Ingeniería Biomédica por su confianza al encargarnos la organización de esta edición de CASEIB. Queremos expresar un agradecimiento especial a los compañeros del Comité Organizador, miembros del Grupo de Ingeniería

Biomédica de la Universidad de Valladolid y del Grupo de Investigación BioRes (*Bioengineering and Resuscitation*) de la Universidad del País Vasco. Sin vosotros no hubiera sido posible **CASEIB 2020**.

Os deseamos que disfrutéis del Congreso. Os aseguramos que hemos puesto lo mejor de nosotros para que **CASEIB 2020** se desarrolle de la mejor manera a pesar de tener que realizarlo de forma virtual. Estaremos a vuestra disposición para solucionar los problemas técnicos que surjan y para resolver vuestras dudas.

Un abrazo en nombre de todo el Comité Organizador,

Dr. Roberto Hornero Sánchez
Presidente del Comité Organizador del CASEIB 2020

Dr. Jesús Poza Crespo
Secretario del Comité Organizador del CASEIB 2020

COMITÉ ORGANIZADOR

Presidente:

Roberto Hornero Sánchez

Secretario:

Jesús Poza Crespo

Vocales:

María García Gadañón

Carlos Gómez Peña

Daniel Álvarez González

Gonzalo C. Gutiérrez Tobal

Javier Gómez Pilar

Colaboradores:

Elisabete Aramendi Ecenarro

Unai Irusta Zarandona

Erik Alonso González

Víctor Martínez Cagigal

Fernando Vaquerizo Villar

Pablo Núñez Novo

Verónica Barroso García

Roberto Romero Oraá

Saúl José Ruiz Gómez

Eduardo Santamaría Vázquez

Jorge Jiménez García

Adrián Martín Montero

Víctor Rodríguez González

Aarón Maturana Candelas

Javier Oraá Pérez

Víctor Gutiérrez de Pablo

Marcos Revilla Vallejo

COMITÉ CIENTÍFICO

Presidente:

Raimon Jané Campos
Universidad Politécnica de Cataluña

Secretarios:

Jesús Poza Crespo
Universidad de Valladolid

Unai Irusta Zarandona
Universidad del País Vasco

Vocales:

Begoña Acha Piñero
Universidad de Sevilla

Raúl Alcaraz Martínez
Universidad de Castilla-La Mancha

Erik Alonso González
Universidad del País Vasco

Elisabete Aramendi Ecenarro
Universidad del País Vasco

Sergio Arana Alonso
CEIT-TECNUN, San Sebastián

José Luis Arce Diego
Universidad de Cantabria

Mª Gloria Bueno García
Universidad de Castilla-La Mancha

César Cáceres Taladriz
Universidad Rey Juan Carlos

Félix del Campo Matías
Hospital Universitario Río Hortega

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

Enrique J. Gómez Aguilera
Universidad Politécnica de Madrid

Carlos Gómez Peña
Universidad de Valladolid

Javier Gómez Pilar
Universidad de Valladolid

Gonzalo C. Gutiérrez Tobal
Universidad de Valladolid

María Elena Hernando Pérez
Universidad Politécnica de Madrid

Roberto Hornero Sánchez
Universidad de Valladolid

Pablo Laguna Lasaosa
Universidad de Zaragoza

María Isabel López Gálvez
Hospital Clínico Universitario de Valladolid

Maite Mujika Garmendia
CEIT-TECNUN, San Sebastián

Gema Prats Boluda
Universidad Politécnica de Valencia

Javier Reina Tosina
Universidad de Sevilla

Ana de los Reyes Guzmán
Hospital Nacional de Parapléjicos,
Servicio de Salud de Castilla-La Mancha

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

Francisco M. Sánchez Margallo
C.C.M.I. Jesús Usón

Daniel Sánchez Morillo
Universidad de Cádiz

Andrés Santos Lleó
Universidad Politécnica de Madrid

Carmen Serrano Gotarredona
Universidad de Sevilla

José María Tormos Muñoz
Instituto Guttman

Abel Torres Cebrián
Universidad Politécnica de Cataluña

Juan José Vaquero López
Universidad Carlos III de Madrid

PROGRAMA – MIÉRCOLES 25 DE NOVIEMBRE

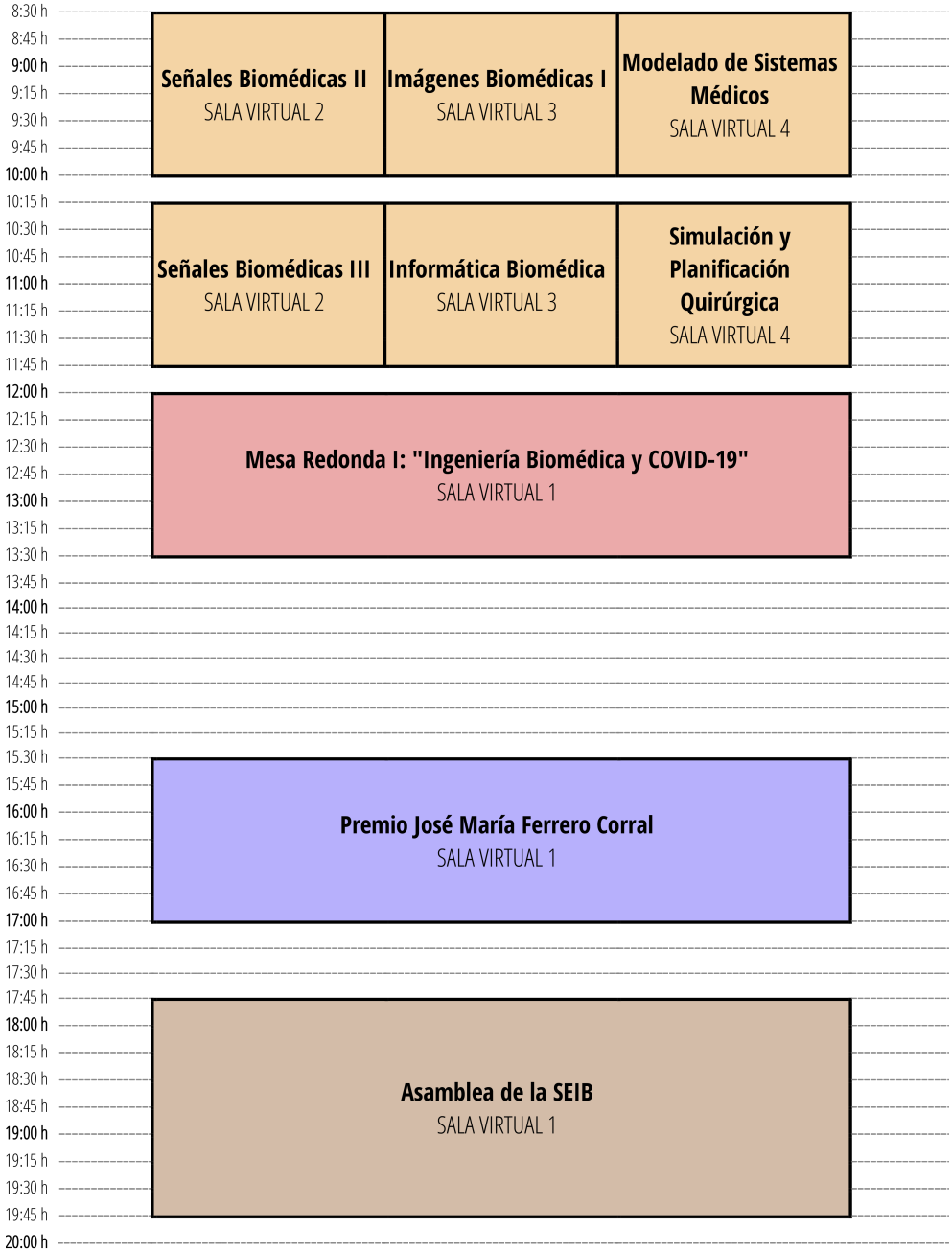


Acto Inaugural
Acto inaugural de CASEIB 2020 con la presencia de autoridades.

Competición de Estudiantes de Grado en Ingeniería Biomédica
Presentaciones de los finalistas de la competición de alumnos de los Grados en Ingeniería Biomédica.

Conferencia Plenaria Inaugural
"*Will Bioengineers Lead the Path to Solution(s) in Sleep Disorders?*", Prof. Dr. David Gozal.

PROGRAMA – JUEVES 26 DE NOVIEMBRE

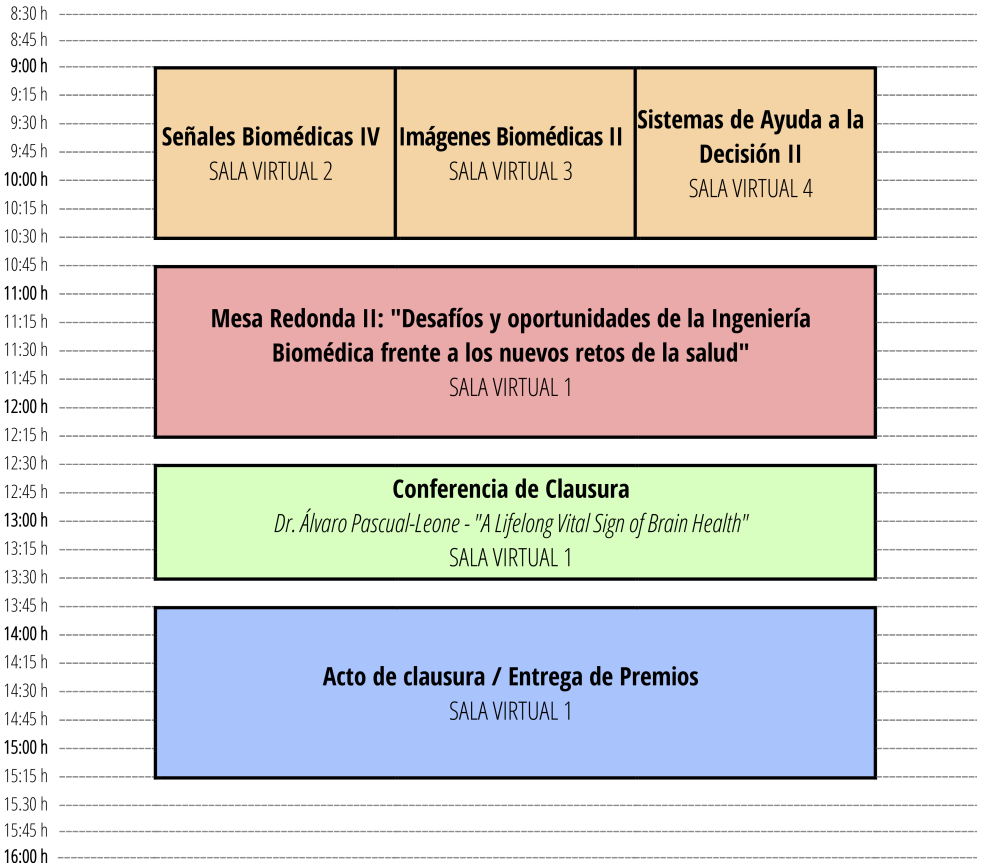


Mesa Redonda I
"*Ingeniería Biomédica y COVID-19*", moderada por el Prof. Dr. Enrique Gómez Aguilera.

Premio José María Ferrero Corral
Presentaciones de los seis finalistas del concurso de comunicaciones de estudiantes de Máster y Doctorado.

Asamblea de la SEIB
Asamblea anual de la Sociedad Española de Ingeniería Biomédica.

PROGRAMA – VIERNES 27 DE NOVIEMBRE



Mesa Redonda II
"*Desafíos y oportunidades de la Ingeniería Biomédica frente a los nuevos retos de la salud*", moderada por el Prof. Dr. Raimon Jané Campos.

Conferencia de Clausura
"*A Lifelong Vital Sign of Brain Health*", Prof. Dr. Álvaro Pascual Leone.

Acto de clausura / Entrega de Premios
Acto de clausura de CASEIB 2020 y entrega del Premio José María Ferrero Corral y de las distinciones de la Competición de Estudiantes de los Grados en Ingeniería Biomédica.

CONFERENCIA PLENARIA INAUGURAL

Will Bioengineers Lead the Path to Solution(s) in Sleep Disorders?

Miércoles 25 de noviembre, 17:00 h – Sala Virtual 1

Moderador: Prof. Dr. Roberto Hornero Sánchez. *Universidad de Valladolid*.



Prof. Dr. David Gozal

Professor at the University of Missouri
Marie M. and Harry L. Smith Endowed Chair of Child Health
Physician-in-Chief at the University of Missouri Women's and Children's Hospital

El Prof. David Gozal, MD, MBA, PhD (Hon) es uno de los líderes científicos a nivel mundial en sueño pediátrico y anterior presidente de la *American Thoracic Society*.

Como neumólogo pediátrico y especialista en medicina del sueño, el Prof. Gozal ha desarrollado numerosas investigaciones para entender los mecanismos neurobiológicos implicados en el control respiratorio, poniendo especial énfasis en los desórdenes pediátricos del sueño y su relación con trastornos neuroconductuales, cardiovasculares y metabólicos. El Prof. Gozal fue el primer investigador en demostrar de manera concluyente que la apnea obstructiva del sueño durante la infancia conduce a un rendimiento académico adverso, y que dicho efecto es parcialmente reversible con el tratamiento. El Prof. Gozal también ha investigado la efectividad de nuevas tecnologías de diagnóstico, que incluye trabajos pioneros sobre biomarcadores de la apnea del sueño basados en métodos de inteligencia artificial. Asimismo, ha explorado cómo diversos tratamientos no quirúrgicos pueden reducir los síntomas conductuales y mejorar la capacidad cognitiva de niños afectados por la apnea del sueño. En definitiva, la trayectoria investigadora del Prof. Gozal ha tratado de aportar nuevos enfoques conducentes a un tratamiento personalizado de los trastornos del sueño.

Como fruto de su intensa actividad investigadora, el Prof. Gozal ha publicado más de 700 artículos originales, reseñas en revista y capítulos de libro. Su investigación ha sido financiada de forma continua por el *National Institute of Health* (NIH). Mantiene numerosas colaboraciones con instituciones y grupos de investigación de todo el mundo. Ha desempeñado y continúa desempeñando diferentes cargos en múltiples organizaciones, incluido el de Presidente de la *American Thoracic Society* (ATS), la principal sociedad profesional en el mundo de la medicina pulmonar, la atención crítica y la medicina del sueño. El impacto y la excelencia de sus trabajos en el área de los trastornos pediátricos del sueño han sido distinguidos por numerosos premios internacionales.

CONFERENCIA PLENARIA DE CLAUSURA

A Lifelong Vital Sign of Brain Health

Viernes 27 de noviembre, 12:30 h – Sala Virtual 1

Moderador: Dr. José María Tormos Muñoz. *Instituto Guttman*.



Prof. Dr. Álvaro Pascual-Leone

Professor of Neurology at Harvard Medical School
Director of the Berenson-Allen Center for Noninvasive Brain Stimulation at Beth Israel Deaconess Medical Center and Harvard Medical School
Program Director of the Harvard-Thorndike Clinical Research Center

Álvaro Pascual-Leone, MD, PhD, es Prof. de Neurología en la *Harvard Medical School*, Científico Senior en el *Arthur Marcus Institute for Aging Research* y Director del *Center for Memory Health at Hebrew SeniorLife* (HSL) en Boston. Asimismo, ostenta el cargo de Director Médico del *Guttmann Brain Health Institute* y Director Científico de la *Barcelona Brain Health Initiative*. Además, el Prof. Pascual-Leone es cofundador de Linus Health y TI Solutions.

El Prof. Pascual-Leone es un neurólogo cognitivo especializado en demencia, enfocado en la atención integral, personalizada y centrada en el paciente y el apoyo multidisciplinar y holístico de las personas con deterioro cognitivo, sus familias y cuidadores. Una de las principales líneas de investigación actuales del Prof. Pascual-Leone consiste en caracterizar y promover la salud del cerebro a lo largo de la vida, prevenir el deterioro cognitivo relacionado con la edad y minimizar la discapacidad en pacientes con trastornos neuropsiquiátricos al tiempo que se obtienen conocimientos fundamentales sobre la función del cerebro humano. Asimismo, es considerado un líder internacional en el desarrollo tecnológico, la aplicación clínica y la investigación de la estimulación cerebral no invasiva. Sus investigaciones tienen como objetivo la comprensión de los mecanismos que controlan las redes cerebrales y su plasticidad a lo largo de la vida, así como la capacidad para ser modificadas para el mejor resultado conductual del individuo, prevención de la discapacidad y restaurar la función en pacientes neuropsiquiátricos.

El Dr. Pascual-Leone ha publicado más de 800 artículos científicos revisados por pares y sus trabajos tienen un elevado impacto (índice h 178; índice i10 680). Thomson Reuters lo ha reconocido como uno de los 15 mejores neurocientíficos del mundo y una de las “mentes científicas más influyentes del mundo”. Ha sido galardonado con numerosos premios internacionales y es miembro electo de la Real Academia Española de Ciencias. La actividad investigadora del Prof. Pascual-Leone ha sido financiada de manera reiterada por el *National Institute of Health* (NIH). Su trabajo tiene gran atractivo para el público general y la difusión y divulgación de sus aportaciones a través de artículos en prensa, documentales de televisión y radio y varios libros es extensa.

MESAS REDONDAS

Mesa Redonda I: 'Ingeniería Biomédica y COVID-19'

Jueves 26 de noviembre, 12:00 h – Sala Virtual 1

Moderador

- Prof. Dr. Enrique Gómez Aguilera. *Secretario General de la Sociedad Española de Ingeniería Biomédica*.

Participantes

- Prof. Dra. Teresa Riesgo Alcaide. *Secretaria General de Innovación, Ministerio de Ciencia e Innovación*.
- D. Juan Carlos Antony García. *Regional Executive Manager, GE Healthcare*.
- Prof. Dra. Laura M. Lechuga Gómez. *Instituto Catalán de Nanociencia y Nanotecnología (ICN2). Plataforma Salud Global del CSIC. Miembro del comité de expertos científicos (Grupo de Trabajo Multidisciplinar) asesor del Ministerio de Ciencia e Innovación en COVID-19. Coordinadora del proyecto europeo CONVAT*.
- Dr. Josep M. Campistol Plana. *Director General del Hospital Clinic, Barcelona*.
- Dr. Tomás Ruiz Albi. *Médico Adjunto del Servicio de Neumología del Hospital Río Hortega de Valladolid. Miembro del Comité Asesor de Expertos frente al COVID-19 de la Junta de Castilla y León*.

Mesa Redonda II: 'Desafíos y oportunidades de la Ingeniería Biomédica frente a los nuevos retos de la salud'

Viernes 27 de noviembre, 10:45 h – Sala Virtual 1

Moderador

- Prof. Dr. Raimon Jané Campos. *Presidente de la Sociedad Española de Ingeniería Biomédica*.

Participantes

- Dra. Raquel Yotti Álvarez. *Directora del Instituto de Salud Carlos III*.
- Prof. Dr. Manuel Desco Menéndez. *Universidad Carlos III de Madrid. Jefe de Servicio de Medicina y Cirugía Experimental, Hospital Universitario Gregorio Marañón. Coordinador del Área de Ingeniería Biomédica de la Agencia Estatal de Investigación*.
- Dr. Sergio Muñoz Lezcano. *Responsable de Innovación y Emprendimiento de la Federación Española de Empresas de Tecnología Sanitaria (FENIN). Coordinador de la Plataforma Española de Innovación en Tecnología Sanitaria*.
- Prof. Dr. Javier Saiz Rodríguez. *Universidad Politécnica de Valencia. Vicepresidente de la Sociedad Española de Ingeniería Biomédica*.

COMPETICIONES

Premio José María Ferrero Corral

Jueves 26 de noviembre, 15:30 h – Sala Virtual 1

CASEIB 2020 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 seis finalistas presentarán sus trabajos durante 10 minutos, a los que le seguirán 5 minutos de debate con el tribunal.

Tribunal

- Prof. Dr. Raimon Jané Campos. *Universidad Politécnica de Cataluña*.
- Prof. Dr. Manuel Desco Menéndez. *Universidad Carlos III de Madrid*.
- Prof. Dr. Javier Saiz Rodríguez. *Universidad Politécnica de Valencia*.

Trabajos seleccionados

- María Bikuña Izaguirre, *et ál. "Characterization of PCLPolymeric based electrospun scaffold towards in vitro Human Trabecular Meshwork model"*.
- Samuel Ruipérez Campillo, *et ál. "Parametrización del Vectorcardiograma Auricular Para la Caracterización de Distintos Tipos De Flutter"*.
- Javier González Alonso, *et ál. "Validación de prototipo con sensores vestibles para captura y análisis de movimientos del cuerpo humano con aplicación en medicina"*.
- Leandro A. Hidalgo Torres, *et ál. "Predicting 3D Photon Interaction in a Hexagonal Positron Emission Tomography Detector: A Deep Learning Approach"*.
- Víctor Gutiérrez de Pablo, *et ál. "Redes de asociación: una nueva forma de explorar la huella fisiopatológica de la enfermedad de Alzheimer"*.
- Álvaro Huerta Herraiz, *et ál. "Comparación de Redes Neuronales Preentrenadas para la Evaluación de la Calidad en Registros Electrocardiográficos"*.

Competición de Estudiantes de Grado en Ingeniería Biomédica

Miércoles 25 de noviembre, 12:30 h – Sala Virtual 1

CASEIB 2020 acoge un año más la Competición de Estudiantes de Grado en Ingeniería Biomédica coordinada por la SEIB. En esta edición del CASEIB, hay nueve 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. Dr. Enrique J. Gómez Aguilera. *Universidad Politécnica de Madrid.*
- Dra. Elisabete Aramendi Ecenarro. *Universidad del País Vasco.*
- Dr. Luis Javier Reina Tosina. *Universidad de Sevilla.*

Trabajos seleccionados

- María Inmaculada García-Briega, *et ál. "Caracterización de una plataforma electroactiva de hidrogelmicroesferas para la diferenciación osteogénica de células madre mesenquimales"*, Universitat Politècnica de València.
- Alejo Pérez Gómez, *et ál. "Generative Adversarial Networks for Peripheral Cell Blood Images Standardization"*, Universitat Politècnica de Catalunya.
- Marina Moro López, *et ál. "Effect of cardiac stem cell derived extracellular vesicles on cardiomyocytes"*, Universidad Carlos III.
- Claudia Aymerich Moreno, *et ál. "Design, construction and validation of a low-cost device to measure maximal inspiratory and expiratory pressures"*, Universidad de Barcelona.
- Rodrigo Enríquez Medina, *et ál. "Development of a Virtual Reality Application for Wrist, Elbow and Shoulder Rehabilitation Using Oculus Quest"*, Universidad CEU San Pablo.
- Aitor Díaz Zelaia, *et ál. "On the use of texturized electrospun wound dressings for absorption maximization"*, Universidad de Mondragón.
- Roberto Holgado Cuadrado, *et ál. "Análisis de la importancia de las características de la esteatohepatitis no alcohólica aplicando modelos Random Forest"*, Universidad Rey Juan Carlos.
- Mario Corral Bolaños, *et ál. "Design and implementation of predictive models based on radiomics to assess response to immunotherapy in non-small-cell lung cancer"*, Universidad Politécnica de Madrid.
- Xabier Bujanda Cundin, *et ál. "Diseño y fabricación de un sistema para el cultivo celular rotatorio en paredes de dispositivos microfluídicos con canales circulares embebidos"*, TECNUN Universidad de Navarra.

ACTOS

Acto inaugural

Miércoles 25 de noviembre, 8:30 h – Sala Virtual 1

La XXXVIII edición del Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2020) contará en su inauguración con la presencia de las siguientes autoridades, en representación de, la Universidad de Valladolid, la Junta de Castilla y León, la Sociedad Española de Ingeniería Biomédica y el Comité Organizador de CASEIB 2020:

- D. Antonio Largo Cabrerizo, *Excmo. y Rector Magnífico de la Universidad de Valladolid*.
- Dña. Rocío Lucas Navas, *Excma. Consejera de Educación de la Junta de Castilla y León*.
- D. Raimon Jane Campos, *Presidente de la Sociedad Española de Ingeniería Biomédica*.
- D. Roberto Hornero Sánchez, *Presidente del Comité Organizador del XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2020)*.

Asamblea de la SEIB

Jueves 26 de noviembre, 17:45 h – Sala Virtual 1

Durante el CASEIB 2020 se celebrará la asamblea anual de la Sociedad Española de Ingeniería Biomédica.

Acto de clausura

Viernes 27 de noviembre, 13:45 h – Sala Virtual 1

El XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica será clausurado por los representantes de la SEIB y del Comité Organizador de CASEIB 2020. Durante la ceremonia de clausura se hará entrega de las distinciones de la Competición de Estudiantes de Grado en Ingeniería Biomédica y del Premio José María Ferrero Corral:

- Prof. Dr. Raimon Jané Campos. *Presidente de la Sociedad Española de Ingeniería Biomédica*.
- Prof. Dr. Enrique J. Gómez Aguilera. *Secretario General de la Sociedad Española de Ingeniería Biomédica*.
- Dr. Jesús Poza Crespo. *Secretario del Comité Organizador de CASEIB 2020*.

ÍNDICE DE SESIONES CIENTÍFICAS

Miércoles, 25 de noviembre de 2020, a las 9:30

Sesión oral: Señales Biomédicas I

Moderadores: José J. Rieta Ibáñez y Jesús Poza Crespo

Discriminación Mejorada Entre Personas Sanas e Hipertensas Combinando Fotopletismografía y Electrocardiografía 2

Retardo entre QT y RR en registros de prueba de esfuerzo como indicador de la heterogeneidad de la repolarización ventricular 6

Predicción de muerte súbita en pacientes con insuficiencia cardíaca crónica mediante el estudio de la dinámica periódica de la repolarización 10

Análisis multidimensional de las 12 derivaciones del ECG estándar para predecir el éxito de la cardioversión eléctrica en fibrilación auricular persistente 14

Sesión oral: Bioinstrumentación

Moderadores: David Naranjo Hernández y Javier García Casado

Artificial synovial fluid and artificial cerebrospinal fluid development and tests using QCR based biosensor for medical diagnosis 19

In vitro human skeletal muscle platform for the evaluation of functional outcomes..... 23

Transistores de grafeno en tatuaje como sensores para EEG 27

Dispositivo inteligente para el estudio de la viabilidad de riñones para trasplante..... 31

Sesión oral: Biomateriales

Moderadores: Ramón Farré Ventura y Elisabete Aramendi

Primeros pasos para el desarrollo de retina biohíbrida 3D basada en fibroína de seda para la degeneración macular relacionada con la edad (DMAE) 36

Modelización in-silico de la Rigidez y Resistencia de Estructuras Porosas Biodegradables 40

Estudio de la influencia del diámetro y temperatura del nozzle en el grosor del filamento para bioimpresión 44

Freeform Reversible Embedding of Suspended Hydrogels for the Manufacturing of Biomedical Applications..... 48

Miércoles, 25 de noviembre de 2020, a las 10:45

Sesión de pósteres

Moderadores: Javier Gómez Pilar, Gonzalo C. Gutiérrez Tobal y Daniel Álvarez González

Optical coherence tomography imaging of colorectal neoplastic polyps developed in genetically modified rats 53

Método de estadiaje del sueño basado en información electrocardiográfica y electromiográfica 57

Non-invasive Monitoring of Physiological Signals in an Ovo Model 61

Protocolo biomecánico de miembro superior para estudio del lanzamiento de pitcher en jugadores de béisbol 65

Diseño de concepto de componente protésico de miembro superior para ciclismo 69

Compensación de radiación dispersa en radiografía digital a través del aprendizaje automático: resultados preliminares 73

Diseño e implementación de clasificadores 2D para la detección de la enfermedad de Alzheimer mediante el uso de redes neuronales convolucionales en imágenes de resonancia magnética 77

Sistema de adquisición de datos de detector de radiación de bajo costo para la educación 81

Herramienta de pegado de múltiples camas para tomografía computarizada en 3D mediante dispositivos GPU 85

Photo vector extraction using Deep Learning techniques for Cognitive Map Index calculation 89

Estudio de sensibilidad de algoritmos de análisis de la variabilidad de la frecuencia cardiaca . 93

Diseño, desarrollo y evaluación de un sistema de detección de estrés en conductores basado en la señal de actividad electrodermal 97

Cambios en los loops de las ondas P del vectocardiograma tras el aislamiento de las venas pulmonares 101

Efecto de la Ablación por Catéter Sobre Ciertas Estructuras Auriculares de Pacientes con Fibrilación Auricular Paroxística: El Caso del Seno Coronario 105

Distinción Entre Electrogramas de Fibrilación Auricular Paroxística Frente a Persistente para Evaluación del Sustrato Auricular en Procedimientos de Ablación por Catéter 109

Estimación del “jitter” fisiológico en trenes de potenciales de unidad motora utilizando parámetros de “jiggle” 113

Predicción del fracaso temprano de la cardioversión eléctrica en fibrilación auricular sobre las 12 derivaciones del ECG estándar 117

Estudio de la asociación entre los patrones de activación local y de sincronización global en el EEG en la enfermedad de Alzheimer 121

Aplicación para usuarios intolerantes a la fructosa 125

Dispositivo laparoscópico mecánico articulado: Análisis de la curva de aprendizaje 129

Modelos Predictivos en el Tratamiento de Pacientes con Alcoholismo 133

Clasificación de bloqueadores de Ikr basada en protocolos de voltage clamp y técnicas de machine learning 137

Arrhythmogenicity quantification of two genetic defects affecting IKr channel in AF patients 141

Feature Selection and Machine Learning for Predicting Multi-drug Resistance just After ICU Admission 145

Construcción de pupilómetro binocular para estudio de lentes intraoculares 149

Miércoles, 25 de noviembre de 2020, a las 12:30

Sesión oral: Competición de Estudiantes de Grado en Ingeniería Biomédica

Tribunal: Enrique J. Gómez Aguilera, Elisabete Aramendi Ecenarro y Luis Javier Reina Tosina

Caracterización de una plataforma electroactiva de hidrogel-microesferas para la diferenciación osteogénica de células madre mesenquimales 153

Generative Adversarial Networks for Peripheral Cell Blood Images Standardization 157

Effect of cardiac stem cell derived extracellular vesicles on cardiomyocytes 161

Design, construction and validation of a low-cost device to measure maximal inspiratory and expiratory pressures 165

Development of a Virtual Reality Application for Wrist, Elbow and Shoulder Rehabilitation Using Oculus Quest 169

On the use of texturized electrospun wound dressings for absorption maximization 173

Análisis de la importancia de las características de la esteatohepatitis no alcohólica aplicando modelos Random Forest 177

Design and implementation of predictive models based on radiomics to assess response to immunotherapy in non-small-cell lung cancer 181

Diseño y fabricación de un sistema para el cultivo celular rotatorio en paredes de dispositivos microfluídicos con canales circulares embebidos 185

Miércoles, 25 de noviembre de 2020, a las 15:45

Sesión oral: Ingeniería de Rehabilitación

Moderadores: Ana de los Reyes Guzmán y Roberto Hornero Sánchez

Virtualización del test clínico Jebsen-Taylor Hand Function basado en Leap Motion Controller 190

Development of a motion measurement system of a long cane for Visually Impaired People in rehabilitation 194

Optimización de la focalidad espacial en tDCS mediante modelos computacionales bioelectromagnéticos: un enfoque alternativo 198

Comparison of electrode positions for a multi-field FES device for compensation of foot drop 201

Sesión oral: Sistemas de Ayuda a la Decisión I

Moderadores: Raúl Alcaraz Martínez y Gloria Bueno García

Método de reconstrucción tomográfica con información a priori obtenida con aprendizaje profundo 206

Implementation of a neural-network-based Surgical Workflow Analysis algorithm for surgical educational environments 210

Desarrollo de un Índice de Privación y un Algoritmo de Aprendizaje Automático para la Identificación de Pacientes Vulnerables 214

Detección Automática de Tos para la Monitorización de Pacientes con Enfermedad Pulmonar Obstructiva Crónica 218

Sesión oral: Telemedicina y E-Salud

Moderadores: Javier Reina Tosina y Daniel Sánchez Morillo

Plataforma mHealth ALIBIRD para el apoyo en el tratamiento y seguimiento de pacientes con cáncer 223

Implementación de mecanismos de seguridad en una plataforma IoT para la monitorización remota de la actividad física 227

Análisis emocional de tweets en castellano para personalizar tratamientos clínicos en telemedicina 231

Validación de la electrónica y ubicación de un sistema para la obtención ambulatoria de constantes vitales en el oído externo 235

Jueves, 26 de noviembre de 2020, a las 8:30

Sesión oral: Señales Biomédicas II

Moderadores: Pablo Laguna Lasaosa y Carlos Gómez Peña

Influence of voluntary contractions on the basal sEMG activity of the pelvic floor muscles 240
Decodificación no lineal de los movimientos de la mano en tiempo real mediante un sistema Brain-Computer Interface 244
Choosing strategies to deal with artifactual EEG data in children with cognitive impairment .. 248
Análisis del flujo aéreo nocturno mediante wavelets para la ayuda en el diagnóstico de la apnea del sueño infantil 252
Análisis de la respuesta electrodermal para la identificación de pacientes con depresión 256
Evaluación de la robustez de la red neuronal funcional en la demencia por enfermedad de Alzheimer 260

Sesión oral: Imágenes Biomédicas I

Moderadores: María García Gadañón y Carmen Serrano Gotarredona

Citometría de flujo con hoja plana para el conteo de bacterias de tuberculosis 265
Reconstrucción y validación de imágenes electrocardiográfica: Problema inverso 269
Efecto de los parámetros geométricos de adquisición en tomosíntesis digital lineal 273
Detection of PD-L1 expression in non-small cell lung cancer with ^{18}F-FDG PET imaging 277
Geometry performance study on scintillator crystal for hexagonal SiPM application 281
Corrección del Artefacto de Truncamiento en TAC mediante Aprendizaje profundo 285

Sesión oral: Modelado de Sistemas Médicos

Moderadores: F. Javier Saiz Rodríguez y Beatriz F. Giraldo Giraldo

Estimulación medular: optimización de la distancia entre polos de estimulación a partir de modelos de simulación 290
Modelado de la propagación óptica a través de tejidos biológicos anisotrópicos utilizando las funciones de Green 294
Computational Simulations of Atrial Electrophysiology Remodelling under Metabolic Syndrome Driven by Up-regulation of Potassium Repolarizing Ionic Currents 298
Prototipo de modelo anatómico impreso en 3D para exploración del oído humano 302

Predicción de la activación neuronal en estimulación tónica en la terapia de EME: uso de modelos 3D de médula espinal personalizados al paciente frente a modelos generalizados 306

Resultados del reto “Biomarker and Endpoint Assessment to Track Parkinson’s Disease (BEAT-PD) DREAM Challenge” 310

Jueves, 26 de noviembre de 2020, a las 10:15

Sesión oral: Señales Biomédicas III

Moderadores: Raimon Jané Campos y Unai Irusta Zarandona

Análisis de flujo aéreo y saturación de oxígeno en sangre mediante transformada wavelet para la detección de la apnea obstructiva del sueño infantil 315

Análisis de la variabilidad temporal de la onda P para predecir el resultado de la ablación por catéter en fibrilación auricular paroxística 319

Diseño de un prototipo de manga sensorizada para el registro de electromiografía de alta densidad en el antebrazo 323

Predicción a Largo Plazo del Resultado de la Ablación Quirúrgica en Fibrilación Auricular Mediante el Uso de una Única Derivación del Electrocardiograma Preoperatorio 327

Estudio de la relación entre el gen MAPT y la conectividad cerebral en la demencia por enfermedad de Alzheimer 331

Evaluación de la Calidad en Registros Electrocardiográficos Procedentes de Pacientes con Fibrilación Auricular Mediante una Red Neuronal Convolucional 335

Sesión oral: Informática Biomédica

Moderadores: Enrique J. Gómez Aguilera y Félix Fanjul Vélez

Modelo de deep learning basado en la arquitectura Inception para el diagnóstico de la apnea del sueño infantil mediante la señal de oximetría 340

Caracterización in silico de la duración de la repolarización y su variabilidad en el síndrome de QT Largo Tipo1 bajo estimulación beta adrenérgica 344

Segmentación automática de instancias aplicada a retinografías para la segmentación del disco óptico y la excavación 348

Algoritmo multietapa para la detección de ventilaciones en la impedancia torácica durante la resucitación cardiopulmonar 352

Caracterización de la morfología foveal: parametrización, diferencias de sexo y efectos de la edad 356

Determinación del efecto de la ablación en tejido porcino mediante un láser Nd:YAG en aplicaciones quirúrgicas 360

Sesión oral: Simulación y Planificación Quirúrgica

Moderadores: José Luis Arce Diego y María Elena Hernando Pérez

ARHealth: App que combina Realidad Aumentada e Impresión 3D para Mejorar el Flujo Quirúrgico 365

Diseño e implementación de un sistema electrónico para el control de un simulador de trauma 369

Fabricación de moldes mediante impresión 3D: Viabilidad de machos solubles para órganos artificiales huecos 373

Estudio del efecto combinado de la Oxacina 4 Perclorato y la irradiación en la citotoxicidad de células U87 376

Simulador de Examen Clínico Objetivo Estructurado para la evaluación de la cadena de supervivencia basado en realidad virtual 380

Real-Time Workflow Detection using Video Streams in Craniosynostosis Surgery 384

Jueves, 26 de noviembre de 2020, a las 15:30

Sesión oral: Premio José María Ferrero Corral

Tribunal: Raimon Jané Campos, Fco. Javier Saiz Rodríguez y Manuel Desco Menéndez

Characterization of polycaprolactone based electrospun scaffold towards in vitro human trabecular meshwork model 389

Parametrización del Vectorcardiograma Auricular Para la Caracterización de Distintos Tipos De Flutter 393

Validación de prototipo con sensores vestibles para captura y análisis de movimientos del cuerpo humano con aplicación en medicina 397

Predicting 3D Photon Interaction in a Hexagonal Positron Emission Tomography Detector: A Deep Learning Approach 401

Redes de asociación: una nueva forma de explorar la huella fisiopatológica de la enfermedad de Alzheimer 405

Comparación de Redes Neuronales Preentrenadas para la Evaluación de la Calidad en Registros Electrocardiográficos 409

Viernes, 27 de noviembre de 2020, a las 9:00

Sesión oral: Señales Biomédicas IV

Moderadores: Elisabete Aramendi Ecenarro y Gema Prats Boluda

RHRVEasy: Heart Rate Variability Analysis Made Easy 414

Aprendizaje máquina para la predicción del éxito de la desfibrilación mediante el análisis de la forma de onda de la fibrilación ventricular 418

Verificación Sobre la Estabilidad de Electrogramas Fragmentados para la Caracterización del Sustrato Auricular en Pacientes con Fibrilación Auricular 422

Modelo predictivo del retorno de circulación espontánea en la parada cardiorrespiratoria utilizando el ECG y la impedancia torácica 426

Confiabilidad de las Ondas de Activación Local para Evaluar el Sustrato Auricular de Pacientes con Fibrilación Auricular Paroxística 430

Monitorización en diálisis de la Concentración de Potasio en Sangre mediante los cambios en la morfología multi-lead de la onda T: Comparación Entre usar la transformación en Componentes Periódicas y Principales 434

Sesión oral: Imágenes Biomédicas II

Moderadores: Patricia Sánchez González y Juan José Vaquero López

Nuevo método para la obtención de imágenes TAC libres de endurecimiento de haz vía aprendizaje automático 439

Deep convolutional encoder-decoder network for semantic segmentation of atrial cavity 443

Estimación del desplazamiento horizontal del detector en un sistema de rayos X utilizando Aprendizaje por Transferencia 447

Diferenciación entre variantes de afasia progresiva primaria mediante el análisis cuantitativo de neuroimágenes de ^{18}F-FDG PET 451

Sistema para el reconocimiento automático de los cambios en el tamaño de la pupila en pequeños roedores 455

Método de calibración geométrica para tomosíntesis de tórax 459

Sesión oral: Sistemas de Ayuda a la Decisión II

Moderadores: Erik Alonso González y Begoña Acha Piñero

Feature Selection and Tree-based Models to Predict Multidrug-Resistance 464

Detección de criterios dermatoscópicos del carcinoma basocelular mediante aprendizaje profundo 468

Diagnóstico de esclerosis múltiple mediante el análisis de imágenes de tomografía de coherencia óptica con redes neuronales convolucionales 472

Predicción del Éxito de la Ablación por Catéter en Fibrilación Auricular Persistente Mediante Características Tiempo-Frecuencia de las Ondas Fibrilatorias 476

Transformada wavelet estacionaria y redes neuronales para la detección de pulso en episodios de parada cardiorrespiratoria 480

Detección automática de patología en imágenes de fondo de ojo utilizando técnicas de deep learning 484

Señales Biomédicas I

Miércoles, 25 de noviembre de 2020

Discriminación Mejorada Entre Personas Sanas e Hipertensas Combinando Fotopletismografía y Electrocardiografía

J. Cano Serrano[1], R. Alcaraz Martínez[2], J.J. Rieta Ibáñez[1*]

[1] BioMIT.org, Departamento de Ing. Electrónica, Universitat Politecnica de Valencia, España. {jecaser, jjrieta}@upv.es
[2] Grupo de Inv. en Electrónica, Telecom. y Bioingeniería, Univ. de Castilla-La Mancha, España. raul.alcaraz@uclm.es

Resumen

Las enfermedades cardiovasculares son una de las principales causas de muerte, siendo la hipertensión arterial (HTA) su principal factor de riesgo. Sus complicaciones pueden evitarse con un tratamiento temprano, pero dado que estos pacientes no presentan ninguna sintomatología, a menudo la HTA se detecta en estadios muy avanzados. Este trabajo presenta un modelo de estimación de la presión arterial (PA) a partir de señales de electrocardiografía (ECG) y fotopletismografía (PPG), que pueden obtenerse del paciente con facilidad mediante dispositivos llevables de monitorización continua. Se han analizado señales de 86 pacientes con registros de ECG, PPG y PA, siendo 35 normotensos, 26 prehipertensos y 25 hipertensos. Los registros fueron limpiados de ruido para mejorar la obtención de sus puntos característicos. A partir de ellos se obtuvieron parámetros discriminatorios como los tiempos de llegada de pulso (PAT) y características morfológicas de la señal PPG como el ancho del pulso e intervalo entre picos sistólicos, con los que se entrenarán modelos de clasificación buscando mejorar los resultados de discriminación entre pacientes. Los resultados obtenidos fueron del 85,44 % de F_1 score comparando a pacientes normotensos con prehipertensos e hipertensos, logrando una mejora del 7 % respecto a la utilización de parámetros predictivos y clasificadores propuestos en estudios previos. La utilización de las señales PPG y ECG ha demostrado ser capaz de discriminar entre personas sanas e hipertensas y, por tanto, podría ser usada para la prevención de la hipertensión implementando esta técnica en dispositivos llevables.

1. Introducción

La hipertensión arterial (HTA) es un factor de riesgo importante de muchas enfermedades cardiovasculares, donde se incluyen enfermedades del corazón, enfermedades vasculares del cerebro y enfermedades de los vasos sanguíneos [1]. El control regular de la presión arterial (PA) es crucial para las personas que ya sufren de HTA, ya que son particularmente vulnerables a una presión arterial elevada. Los estilos de vida saludables, la detección temprana, la evaluación de los niveles de presión arterial y obtener un diagnóstico adecuado son beneficiosos para la prevención y el control de la HTA [2].

Para estimar la PA de forma no invasiva, los dispositivos convencionales de medición basados en manguito ofrecen buena precisión. Sin embargo, no están concebidos para ser llevables, no son compatibles con una medida continuada a lo largo del día, son incómodos y su procedimiento de medida es algo tedioso y requiere la atención del paciente [3]. A esto hay que añadir que la mayoría de los pacientes con HTA no tienen síntomas en la etapa de presión arterial elevada e incluso en la hipertensión, por tanto, muchas personas pierden, por ignorancia y ausencia de control médico, la oportunidad de un tratamiento temprano y experimentan complicaciones cardiovasculares que podrían evitarse [1].

Debido a los factores mencionados, el trabajo en este campo se centra en el desarrollo de sistemas de estimación de la PA robustos y discretos, que puedan ofrecer al usuario actualizaciones periódicas del nivel de PA casi en tiempo real [4]. El principal avance que ha propiciado el desarrollo de estos sistemas es la presencia cada vez mayor de dispositivos llevables, como pulseras o relojes inteligentes [5], capaces de monitorizar señales fisiológicas que cambian en función del nivel de PA, como son el electrocardiograma (ECG) y el fotopletismograma (PPG). Los cambios morfológicos de las señales fisiológicas reflejan principalmente cambios en el estado de funcionamiento del corazón y del sistema vascular, por lo que la información morfológica del PPG podría utilizarse para evaluar la hipertensión [6].

El presente trabajo utiliza el análisis y procesado de señales ECG y PPG para desarrollar un sistema mejorado de discriminación entre pacientes normotensos, prehipertensos e hipertensos. El objetivo último es poder identificar, mediante dispositivos llevables, pacientes hipertensos sin síntomas aparentes o monitorizar pacientes de riesgo para así conseguir prevenir futuras enfermedades cardiovasculares, de las cuales la HTA es su principal factor de riesgo.

2. Materiales

Las señales que se han utilizado han sido obtenidas de la base de datos MIMIC, la cual contiene información de pacientes procedentes de la Unidad de Cuidados Intensivos (UCI) admitidos en el Beth Israel Deaconess Medical Center en Boston, EEUU [7]. Se ha elegido esta base de datos ya que contiene señales de ECG, PPG y de PA (ABP) registradas simultáneamente en UCI.

De la base datos MIMIC se ha seleccionado un total de 86 pacientes que contuvieran los tres tipos de registro mencionados anteriormente y en condiciones aceptables. De los pacientes seleccionados, 35 eran normotensos, 26 prehi-

ISBN: 978-84-09-25491-0, pags. 2- 5

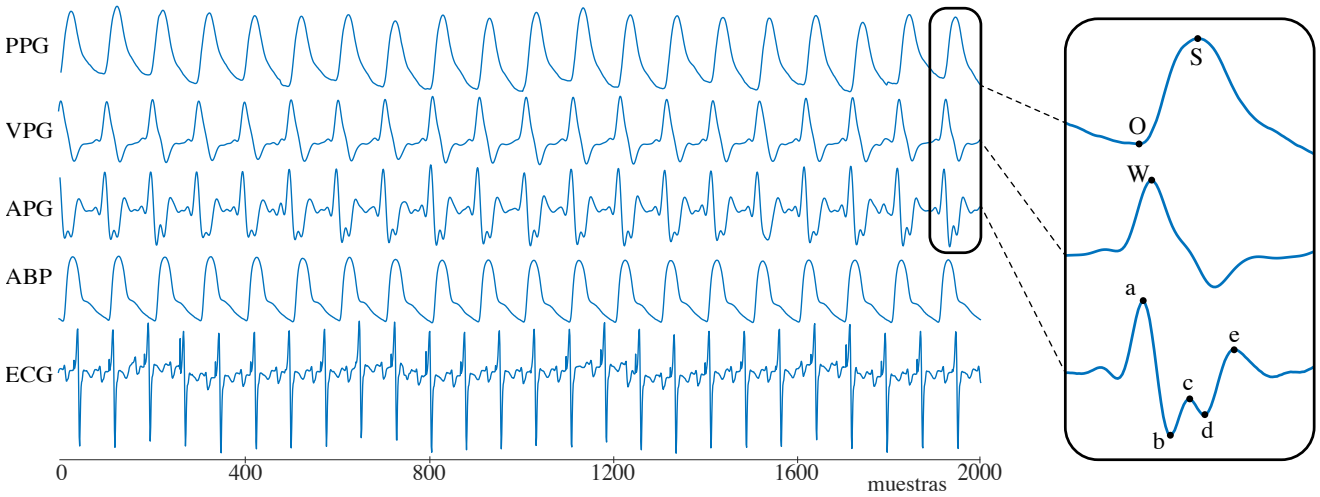


Figura 1. *Fragmento de 2000 muestras que ilustra la morfología de las señales utilizadas en el trabajo, junto con la representación de los puntos característicos de las señales PPG, VPG y APG definidos.*

pertensos y 25 hipertensos. Las señales fueron registradas todas ellas de forma simultánea con una duración de 60 segundos y con una frecuencia de muestreo común de 125 Hz y una resolución de 8-10 bits [8].

3. Métodos

3.1. Preprocesado y puntos característicos

Con el objetivo de eliminar el ruido y mejorar la calidad de la señal, a la señal PPG en bruto se le aplicó un filtro paso banda Chebyshev II de cuarto orden entre 0.5 y 10 Hz [9]. A continuación, se eliminaron las fluctuaciones en la línea base ajustando los mínimos del pulso a cero a partir de la diferencia entre la señal y su envolvente inferior. De esta forma, se facilita una protección para la obtención correcta de las amplitudes de los puntos característicos del PPG.

A partir de la señal PPG, que indica el movimiento del flujo de sangre en el lugar de medida, se obtuvo su primera y segunda derivada, denominadas señal de velocidad fotopletismográfica (VPG) y de aceleración fotopletismográfica (APG), respectivamente [10]. Por último, se han obtenindo los puntos característicos del PPG, VPG y APG, buscando máximos y mínimos locales que han sido calculados en base a umbrales en cada uno de los pulsos de las señales, que se ilustran en la Figura 1. Se han extraído los picos sistólicos de las tres señales (S,W,a), las muescas diastólicas de la señal PPG (O) y dos máximos y mínimos locales de la señal APG (b,c,d,e) [10].

Las señales de PA eran claras y no necesitaron preprocesado. Solo se debió detectar el máximo absoluto de presión arterial sistólica (PAS), que se utilizó como referencia de PA para el resto de cálculos. Finalmente, a cada ECG se le aplicó un preprocesado estándar [11]. Así, se filtró paso alto con frecuencia de corte de 0.5 Hz para eliminar la línea base y, a continuación, se filtró paso bajo con frecuencia de corte de 50 Hz para reducir el ruido muscular de alta frecuencia y eliminar la interferencia de red eléctrica, en este caso, de 60 Hz [11]. Después se aplicó un detector de picos R para obtener la posición de cada latido [12].

Normotenso vs Prehipertenso	Normotenso vs Hipertenso	Norm. + Prehip. vs Hipertenso	Normotenso vs Prehip. + Hip.

Tabla 1. *Tipos de pacientes comparados en cada matriz*

3.2. Obtención de parámetros discriminatorios

Una vez obtenidos los puntos característicos, se obtuvieron dos matrices de datos con las que se han entrenado los modelos de clasificación. La primera matriz se formó con los 13 parámetros de los que se ha reportado una mayor correlación con los niveles de PA en trabajos previos [13], como son los tres tiempos de llegada de pulso (PAT) que se calcularon como el retardo entre el pico R del ECG y un punto característico de la señal PPG, PAT_{peak}, $\text{PAT}_{derivate}$, PAT_{foot} que ha mostrado buenos resultados en estudios previos [14], intervalos de tiempo entre el pico S y los puntos c y d (ver Fig. 1), pendientes entre puntos característicos y ratios entre amplitudes y áreas bajo el pulso. La otra con 24 parámetros utilizados en la literatura basados en los tres PAT y características de la señal PPG, como son las áreas antes y después del pico sistólico, el ratio entre ellas además del área total, intervalos entre picos sistólicos, intervalo total del pulso, tiempo de ascenso y PTT, intervalo entre el pico PAS y el pico sistólico, amplitud de los picos sistólicos en las señales PPG y VPG y ratios entre los puntos característicos de la señal APG con el objetivo de mejorar los resultados de trabajos previos [10].

Debido a que se disponía de tres grupos diferentes de pacientes, se han realizado las cuatro comparativas pareadas con más lógica que se muestran en la Tabla 1. Por otro lado, para poder efectuar un correcto análisis de los datos, se ha realizado un tratamiento de los valores extremos obtenidos u outliers. Para ello se ha desarrollado un método que genera un array lógico cuyos elementos son verdaderos cuando se detecta un outlier. El criterio utilizado ha sido la desviación absoluta de la mediana (DAM), el cual detecta y etiqueta como outliers aquellos valores que están a una distancia superior a tres desviaciones absolutas medias escaladas de la mediana [15].

	AdaBoost			Logistic Regresion			KNN			Bagged		
	Se	Sp	F_1	Se	Sp	F_1	Se	Sp	F_1	Se	Sp	F_1
Normo vs Pre	61,54 %	74,29 %	62,75 %	57,69 %	68,57 %	57,69 %	42,31 %	80,00 %	50,00 %	46,15 %	74,29 %	51,06 %
Normo vs Hiper	80,00 %	74,29 %	74,07 %	52,00 %	77,14 %	56,52 %	40,00 %	94,29 %	54,05 %	76,00 %	82,86 %	76,00 %
Normo+Pre vs Hiper	88,00 %	81,97 %	75,86 %	36,00 %	90,16 %	45,00 %	36,00 %	100,00 %	52,94 %	52,00 %	90,16 %	59,09 %
Normo vs Pre+Hiper	76,47 %	71,43 %	78,00 %	60,78 %	34,29 %	59,05 %	62,75 %	68,57 %	68,09 %	72,55 %	57,14 %	71,84 %

Tabla 2. *Desempeño de los cuatro modelos de clasificación analizados con los parámetros característicos que han reportado una mayor correlación con los niveles de PA en trabajos previos [13].*

	Naive Bayes			SVM cubic			SVM quadratic			Coarse Tree		
	Se	Sp	F_1	Se	Sp	F_1	Se	Sp	F_1	Se	Sp	F_1
Normo vs Pre	57,69 %	91,43 %	68,18 %	57,69 %	74,29 %	60,00 %	61,54 %	91,43 %	71,11 %	50,00 %	80,00 %	56,52 %
Normo vs Hiper	64,00 %	85,71 %	69,57 %	68,00 %	82,86 %	70,83 %	60,00 %	88,57 %	68,18 %	40,00 %	80,00 %	47,62 %
Normo+Pre vs Hiper	64,00 %	91,80 %	69,57 %	64,00 %	91,80 %	69,57 %	68,00 %	93,44 %	73,91 %	48,00 %	77,05 %	47,06 %
Normo vs Pre+Hiper	82,35 %	82,86 %	84,85 %	54,90 %	65,71 %	61,54 %	76,47 %	80,00 %	80,41 %	86,27 %	77,14 %	85,44 %

Tabla 3. *Desempeño de los modelos de clasificación con los nuevos parámetros característicos y modelos de clasificación propuestos.*

Después de localizar los outliers, éstos se han sustituido por la mediana de los valores de la característica donde se ha obtenido el outlier de los pacientes que comparten su misma etiqueta de clasificación. Esta nueva matriz será la que se utilice para entrenar los modelos de clasificación.

3.3. Modelos de clasificación

Los modelos de clasificación utilizados fueron elegidos buscando una diversidad amplia en su metodología de operación. Inicialmente se aplicaron modelos de Regresión Logística, AdaBoost Tree, K Vecinos más próximos (KNN) y Bagged Tree, que representan teorías de clasificación como la regresión, árboles de decisión, cluster y árbol de decisión embolsado [16]. La razón principal de su elección se fundamenta en que han reportado buenos resultados en la discriminación de pacientes con HTA a partir del análisis de señales PPG y ECG en trabajos previos [13].

Por otro lado, también se han aplicado otros cuatro modelos de clasificación adicionales buscando mejorar el porcentaje de acierto en la clasificación. Para ello, se han utilizado hasta 37 estrategias de clasificación diferentes, como son varios tipos de árboles de decisión, análisis discriminante, regresión logística, Naive Bayes, máquinas de vectores soporte (SVM), KNN y clasificadores de conjuntos [16]. Tras utilizar las cuatro matrices de datos con las diferentes agrupaciones de pacientes, se seleccionaron los modelos SVM cúbica y cuadrática, Naive Bayes utilizando kernels y Coarse Tree por haber obtenido los mayores porcentajes de acierto clasificatorio.

3.4. Evaluación de los modelos de clasificación

Para estimar la precisión de los modelos se ha utilizado la estrategia de validación cruzada dejando uno fuera [17]. Finalmente, se han utilizado tests estadísticos para realizar una evaluación de los resultados obtenidos en los modelos de clasificación. En este trabajo se emplean la sensibilidad (Se) o capacidad de detectar la enfermedad en sujetos enfermos, especificidad (Sp) o capacidad de dar como negativos pacientes sanos y F_1 score, que es la media armónica de precisión de detectar falsos positivos y sensibilidad [18].

4. Resultados

El estudio estadístico comenzó utilizando los parámetros predictivos que han reportado una mayor correlación con los niveles de PA en trabajos previos [13] y siguiendo las cuatro agrupaciones de pacientes de la Tabla 1. Los resultados obtenidos se presentan en la Tabla 2. Se puede apreciar que las mejores clasificaciones se obtienen comparando los pacientes normotensos con los prehipertensos e hipertensos en el modelo AdaBoost, donde se obtiene un score F_1 en torno al 78 %. Los demás modelos no realizan correctamente la discriminación y tienden a clasificar como sanos tanto a pacientes que realmente están sanos como a los enfermos. Esto se puede apreciar en el desbalance que existe entre Se y Sp y que se generaliza en estos modelos.

De forma análoga, la Tabla 3 presenta los resultados mejorados de este estudio, en donde se puede observar que el empleo de los nuevos parámetros predictivos, junto con lo modelos de clasificación alternativos, permiten mejorar significativamente los resultados de clasificación. Así, el valor de F_1 para los modelos de clasificación de Naive Bayes y Coarse Tree supera el 84 %.

5. Discusión

El estudio de la variación en la morfología de la señal PPG tiene gran cantidad de información fisiológica que puede utilizarse para estudiar la actividad cardiovascular [19]. Existen trabajos en los que únicamente se utiliza el PAT para estimar la PA [6, 14], pero cuya eficacia ha sido discutida en trabajos posteriores al combinarlo con características adicionales del PPG, que permitieron reportar una mayor correlación con los niveles de PA [13]. En este sentido, el mayor porcentaje de acierto se obtuvo al combinar el PAT con características del PPG, ya que ambas reflejan una información fisiológica diferente. Así, el PAT indica la transmisión de la onda arterial en el vaso sanguíneo, mientras que las características del PPG indican el cambio de estado del tejido vascular y del volumen sanguíneo [13].

Hasta el momento, los últimos trabajos que han utilizado la señal PPG tanto para clasificar a los pacientes en sanos

o hipertensos [20] como para obtener el nivel de PA [21] no se han puesto de acuerdo en cuanto a los parámetros predictivos ni a los modelos de clasificación que se deben utilizar, ya que esta elección depende en gran medida de las señales utilizadas en los estudios y existe disparidad en cuanto a la selección de pacientes, modo de adquisición y calidad de las señales [20, 21].

Cabe destacar que, en todos los modelos analizados en el presente trabajo, los mejores resultados se obtienen cuando se compara a pacientes normotensos con prehipertensos e hipertensos. Este hecho es muy relevante, ya que quiere decir que los valores de los parámetros discriminatorios en los pacientes prehipertensos se asemejan más a los hipertensos que a los normotensos. Además, hay que tener en cuenta que los pacientes prehipertensos no muestran síntomas evidentes hasta que se encuentran en estados muy avanzados de la enfermedad, causando graves problemas cardiovasculares, por lo que alertar como pacientes enfermos a este grupo es de gran interés.

Finalmente, este trabajo presenta ciertas limitaciones. El número de pacientes no es demasiado elevado, no se tiene información de factores que pueden suponer un mayor riesgo de padecer hipertensión como la condición física, el sexo o la edad y sería deseable utilizar señales con una frecuencia de muestreo superior a 125Hz, lo cual podría proporcionar una mejora sustancial en la identificación de los puntos fiduciales.

6. Conclusiones

Este trabajo ha mostrado que el análisis combinado de señales PPG y ECG, junto con la definición de nuevas características morfológicas, como el ancho del pulso e intervalo entre picos sistólicos, así como el empleo de modelos de clasificación alternativos, permite discriminar mejor entre personas sanas y pacientes prehipertensos o hipertensos. La implantación de estas metodologías en dispositivos llevables puede posibilitar la prevención de la hipertensión y sus patologías cardiovasculares asociadas.

Agradecimientos

El presente trabajo ha sido cofinanciado por los proyectos de investigación competitiva DPI2017-83952-C3 de MINECO-AEI-FEDER-UE, SBPLY/17/180501/000411 de la JCCLM y AICO/2019/036 de la GVA.

Referencias

[1] Unger T, Borghi C, Charchar F, Khan NA, et al. 2020 international society of hypertension global hypertension practice guidelines. Hypertension Jun 2020;75(6):1334–1357.

[2] Levy D, Larson MG, Vasan RS, Kannel WB, Ho KK. The progression from hypertension to congestive heart failure. Jama 1996;275(20):1557–1562.

[3] Frese EM, Fick A, Sadowsky HS. Blood pressure measurement guidelines for physical therapists. Cardiopulmonary physical therapy journal 2011;22(2):5.

[4] Kario K. Management of hypertension in the digital era: Small wearable monitoring devices for remote blood pressure monitoring. Hypertension Sep 2020;76(3):640–650.

[5] MHealth W. New horizons for health through mobile technologies. Geneva WHO 2011;.

[6] Mukkamala R, Hahn JO, Inan OT, Mestha LK, Kim CS, Töreyin H, Kyal S. Toward ubiquitous blood pressure monitoring via pulse transit time: theory and practice. IEEE Transactions on Biomedical Engineering 2015;62(8):1879–1901.

[7] Johnson AE, Pollard TJ, Shen L, Li-wei HL, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, Mark RG. Mimic-iii, a freely accessible critical care database. Scientific data 2016;3:160035.

[8] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. circulation 2000;101(23):e215–e220.

[9] Liang Y, Elgendi M, Chen Z, Ward R. An optimal filter for short photoplethysmogram signals. Scientific data 2018; 5:180076.

[10] Elgendi M. On the analysis of fingertip photoplethysmogram signals. Curr Cardiol Rev Feb 2012;8(1):14–25.

[11] Sörnmo L, Laguna P. Biomedical Signal Processing in Cardiac and Neurological Applications. Elsevier Academic Press, 2005.

[12] Martínez A, Alcaraz R, Rieta JJ. Application of the phasor transform for automatic delineation of single-lead ECG fiducial points. Physiological Measurement sep 2010; 31(11):1467–1485.

[13] Liang Y, Chen Z, Ward R, Elgendi M. Hypertension assessment via ecg and ppg signals: An evaluation using mimic database. Diagnostics 2018;8(3):65.

[14] Liang Y, Abbott D, Howard N, Lim K, Ward R, Elgendi M. How effective is pulse arrival time for evaluating blood pressure? challenges and recommendations from a study using the mimic database. Journal of clinical medicine 2019;8(3):337.

[15] Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of Experimental Social Psychology 2013;49(4):764–766.

[16] Shalev-Shwartz S, Ben-David S. Understanding machine learning: From theory to algorithms. Cambridge university press, 2014.

[17] Wong TT. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. Pattern Recognition 2015;48(9):2839 – 2846.

[18] Altman DG, Bland JM. Statistics notes: Diagnostic tests 1: sensitivity and specificity. BMJ jun 1994;308(6943):1552–1552.

[19] Allen J. Photoplethysmography and its application in clinical physiological measurement. Physiological measurement 2007;28(3):R1.

[20] Riaz F, Azad MA, Arshad J, Imran M, Hassan A, Rehman S. Pervasive blood pressure monitoring using photoplethysmogram (ppg) sensor. Future Generation Computer Systems 2019;98:120–130.

[21] Chen S, Ji Z, Wu H, Xu Y. A non-invasive continuous blood pressure estimation approach based on machine learning. Sensors 2019;19(11):2585.

Retardo entre QT y RR en registros de prueba de esfuerzo como indicador de la heterogeneidad de la repolarización ventricular

C. Pérez Martínez[1], E. Pueyo Paules[1,2], JP. Martínez Cortés[1,2], J. Viik[3], P. Laguna Lasaosa[1,2]

[1]Grupo BSICoS, I3A, IIS Aragón, Universidad de Zaragoza, Zaragoza, España {cperez, epueyo, jpmart, laguna}@unizar.es

[2]CIBER-Bioingeniería, Biomateriales y Nanomedicina (CIBER BBN), España

[3]Faculty of Medicine and Health Techonology, Tampere University of Technology, Tampere, Finland jari.viik@tuni.fi

Resumen

El aumento de la heterogeneidad espacio-temporal en la repolarización ventricular está relacionado con arritmias cardiacas. Una adaptación lenta del intervalo QT en respuesta a cambios abruptos del ritmo cardiaco (HR), modelada como la constante de tiempo de un sistema de primer orden con entrada escalón, se ha identificado como marcador de riesgo arrítmico. Un cambio abrupto de HR es difícil de producir en pacientes, por lo que este trabajo se centra en el estudio de los cambios de HR en registros de prueba de esfuerzo modelados como un sistema de primer orden con entrada rampa. La constante de tiempo se calcula como el retardo entre la serie de QT y la serie de QT estimada usando el valor instantáneo de HR ajustando un modelo hiperbólico, evaluado separadamente en las fases de ejercicio y recuperación. Este procedimiento se ha aplicado en 251 sujetos con diferente riesgo de enfermedad en las arterias coronarias (CAD). Los resultados de la constante de tiempo obtenidos se encuentran entre 17s y 49s, y son comparables con los retardos reportados en estudios previos donde los cambios de HR se modelaron como un escalón. Los retardos en las fases de ejercicio y recuperación, así como la diferencia de ambos, se pueden considerar marcadores con capacidad de estratificar pacientes con distinto riesgo de sufrir CAD.

1. Introducción

En los últimos años se han propuesto diferentes biomarcadores para estratificar pacientes según su riesgo de sufrir arritmias ventriculares y Muerte Súbita Cardiaca (SCD). Concretamente, se destacan los que se basan en la dispersión espacio-temporal de la repolarización para predecir SCD [1]. La presencia de heterogeneidades electrofisiológicas en el miocardio ventricular puede ocasionar un aumento de la dispersión en la repolarización ventricular, ya que las células ventriculares presentan diferentes patrones de repolarización en respuesta a cambios del ritmo cardiaco (HR). Durante esta situación de cambio, la probabilidad de manifestarse arritmias ventriculares aumenta.

El tiempo de adaptación del intervalo QT a cambios abruptos en HR se ha identificado como un biomarcador de riesgo arrítmico [2], cuyos mecanismos básicos se han esclarecido por medio de estudios electrofisiológicos tanto experimentales como simulados [3]. Este fenómeno ocurre en la variabilidad de latido a latido, que se cuantifica normalmente bajo condiciones estacionarias, y puede ofrecer información complementaria [4]. Estudios previos también han destacado la importancia de definir rangos normales y anormales de las dinámicas de adaptación del intervalo QT en respuesta a cambios bruscos en HR como una posible vía para caracterizar el riesgo de sufrir arritmias y SCD [5].

Sin embargo, producir cambios bruscos en HR para medir el retardo de tiempo no es sencillo en registros Holter. De modo que, de forma alternativa, se propone estudiar este tiempo en señales tipo rampa procedentes de señales de pruebas de esfuerzo realizadas en la rutina clínica, ya que los cambios en el sistema cardiaco son más suaves y aproximadamente lineales tanto en la fase de ejercicio como de recuperación. Teóricamente, la constante de tiempo de un sistema de primer orden ante una entrada tipo rampa es la misma que en este mismo sistema, pero con una entrada tipo escalón [6]. De modo que, hipotéticamente, estudiar el retardo en señales tipo rampa debe aportar la misma información clínica.

2. Métodos

2.1. Datos

Para este estudio se han evaluado 251 señales de ECG muestreadas a 500Hz, registradas durante una prueba de esfuerzo en una bicicleta ergométrica realizada en *Tampere University Hospital* (Finlandia), con el propósito de caracterizar sujetos con alto riesgo de sufrir morbilidad y mortalidad cardiovascular. Los pacientes se clasifican acorde a su probabilidad de padecer una enfermedad de las arterias coronarias (CAD). Basándose en su historia clínica, los resultados de la prueba de esfuerzo y de una angiografía coronaria se diferencian cuatro grupos: *noCAD* (110 sujetos), *bajo* (39 sujetos), *medio* (12 sujetos), y *alto* riesgo (90 sujetos).

2.2. Preprocesado de señal

Primero, se obtiene una nueva señal ECG resultante de transformación espacial, aplicando Análisis por Componentes Principales (PCA) [7] sobre las 8 derivaciones estándar independientes. Sobre ella se calculan las series de *RR(i)* y *QT(i)* aplicando un detector de picos, inicios y finales de ondas [8] para localizar dichos intervalos en cada i-ésimo latido. Se combina con una interpolación tipo spline para tener en cuenta valores de intervalos perdidos. Para obtener las series de *RR(i)* y *QT(i)* lo más limpias posibles, se califican como *outliers* aquellos valores que se desvían del umbral ±5% de la mediana de una ventana móvil de 50 latidos, y su valor se reemplaza por el correspondiente a dicha mediana. Por último, se

ISBN: 978-84-09-25491-0, pags. 6- 9

interpolan a 4Hz estas series, trabajando así con las series temporales uniformemente muestreadas de *RR(n)* y *QT(n)*.

2.3. Serie QT estimada dependiente de HR

Para obtener una serie temporal de QT sin memoria correspondiente a la que se daría si el HR de ese momento fuera estacionario, es decir, aquella serie, retardada respecto a la real, que representa al valor de QT que ya habría convergido al estacionario para su HR ($\widehat{QT}(n)$), se estiman los parámetros α y β de cada uno de los modelos de regresión estudiados. Se ajustan los pares de datos [*QT(n), RR(n)*] de tres ventanas distintas definidas a lo largo de las series: una al inicio (40 segundos), otra en el pico de la señal (20 segundos) y la última al final (40 segundos). En la Fig.1a los rectángulos representan las ventanas seleccionadas, las cuales se asume que pertenecen a zonas estacionarias y que son representativas de la dependencia *QT-RR* de cada sujeto. Para que las tres ventanas tengan el mismo peso en el ajuste, los pares de valores de la ventana del pico de la señal se duplican.

Los cuatro modelos de regresión considerados fueron:

Parabólico log/log (Par)	$\widehat{QT}(n) = \beta(RR(n))^{\alpha}$
Lineal (Lin)	$\widehat{QT}(n) = \beta + \alpha\, RR(n)$
Hiperbólico (Hip)	$\widehat{QT}(n) = \beta + \frac{\alpha}{RR(n)}$
Logarítmico (Log)	$\widehat{QT}(n) = \beta + \alpha \ln(RR(n))$

Asumir que la ventana establecida en el pico de la señal pertenece a una zona estable es cuestionable, pero se incluye para cubrir todo el rango de valores de *RR* y así poder evaluar más precisamente la dependencia *QT-RR*. Por otro lado, esta ventana es simétrica alrededor del pico, de modo que se incluyen tanto valores de la fase ascendente como descendiente de *RR*, así que la relación media *QT-RR* no se alejan de un comportamiento estacionario.

2.4. Estimación del tiempo de retardo

En la Fig.1b se muestran las series temporales $\widehat{QT}(n)$ y *QT(n)* de un sujeto. El retardo se obtiene calculando el error cuadrático medio (MSE) entre las zonas de rampa de la señal *QT(n)* y $\widehat{QT}(n-\tau^*)$, siendo τ^* el retardo con el que se consigue el menor MSE. El retardo se calcula independientemente en la rampa de la fase de ejercicio, τ_e, y en la rampa de la fase de recuperación, τ_r. Un ejemplo de la delimitación de estas rampas se presenta en la Fig.1b: $n = n_{e,o}$ y $n = n_{r,o}$ marcan el inicio de la rampa de ejercicio y de recuperación, respectivamente, y $n = n_{e,e}$ y $n = n_{r,e}$ el final de la rampa de ejercicio y recuperación.

Delimitación de las rampas de ejercicio y recuperación

El punto que determina el inicio de la rampa en ejercicio, $n_{e,o}$ (y el punto final de la rampa de recuperación, $n_{r,e}$), se define como el punto donde se minimiza la suma de los residuos al cuadrado del ajuste lineal de los puntos predecesores -la parte plana (inclinada)- y del ajuste lineal de los puntos siguientes -la parte inclinada (plana)- de la serie $\widehat{QT}(n)$ [9]. Es decir, en el punto $n_{e,o}$:

$$n_{\{e,o\}} = \arg min_k(J(k)) + 60$$

Donde

$$J(k) = \sum_{n=m}^{k-1} (\widehat{QT}(n) - f_p(n))^2 + (\widehat{QT}(n) - f_s(n))^2$$

Tanto $f_p(n) = a_p + b_p n$ y $f_s(n) = a_s + b_s n$ son las estimaciones lineales de $\widehat{QT}(n)$ predecesora y siguiente al punto *k*, respectivamente, siendo *m=1* y *N* igual a la muestra correspondiente al pico de intensidad máxima de la prueba de esfuerzo. Análogamente, para $n_{r,e}$ se realiza la misma minimización, pero ahora *m* es igual a la muestra correspondiente al pico de intensidad máxima de la prueba de esfuerzo y *N* es el punto final de $\widehat{QT}(n)$. Se suma un valor adicional de 60 muestras para evitar situar el punto $n_{e,o}$ buscado justo en la zona de cambio (para el punto $n_{r,e}$ se resta).

Para posicionar el punto final de la rampa de ejercicio, $n_{e,e}$, primero se realiza el ajuste lineal desde $n_{e,o}$ y desde la muestra para la que $QT(n) = \widehat{QT}(n_{e,o})$ de las series $\widehat{QT}(n)$ y $QT(n)$, respectivamente, hasta la muestra correspondiente al pico de la prueba de esfuerzo, debido a que en este intervalo ambas series siguen una tendencia tipo rampa. Luego, se define $n_{e,e}$ como $n_{e,o} + 0.65(n_{e,c} - n_{e,o})$, donde $n_{e,c}$ es el punto donde ambas rectas del ajuste lineal se encuentran más próximas. De forma similar, el punto inicial de la rampa de recuperación, $n_{r,o}$, se calcula como $n_{r,c} + 0.45(n_{r,e} - n_{r,c})$, siendo $n_{r,c}$ el punto donde las rectas del ajuste de ambas series en el área de recuperación están más próximas. Los valores para ambos umbrales (0.65 y 0.45) se escogen de manera que maximizan la significancia entre los cuatro grupos de pacientes definidos de acuerdo a su riesgo de sufrir CAD.

Se utiliza el test U de Mann-Whitney para comprobar la independencia de los resultados entre dos grupos tanto en la fase de ejercicio como de recuperación de forma independiente.

3. Resultados

El mejor ajuste para la relación QR-a-RR (raíz del error cuadrático medio, ε_{rms}) se consigue con el modelo hiperbólico, por lo que solamente se evalúan los resultados obtenidos con este. En la Tabla 1 se muestran los valores medios y la deviación estándar de los parámetros α y β, empleados para obtener la serie $\widehat{QT}(n)$. Se observa una moderada variabilidad intersujeto en el parámetro α para cada uno de los grupos. Los valores medios máximo y mínimo de HR de cada grupo registrados en la prueba de esfuerzo también se adjuntan en la Tabla 1. El valor medio máximo, HR_{w2}, en el grupo noCAD y de bajo riesgo es considerablemente más alto que en los grupos de más riesgo.

Un ejemplo del retardo calculado entre las series temporales $\widehat{QT}(n)$ y *QT(n)* se muestra en la Fig.1b. El

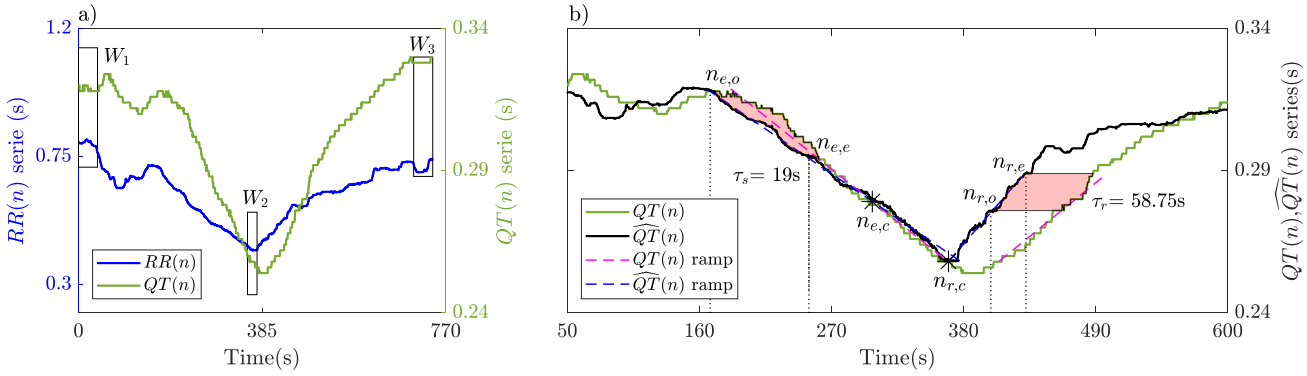


Figura 1. **a)** Selección de los pares de datos QT-RR de las tres ventanas definidas (W_1, W_2 y W_3) para estimar los parámetros del modelo hiperbólico y calcular la serie temporal QT dependiente de HR. **b)** Ejemplo de la delimitación de las rampas de ejercicio (siendo $n_{e,o}$ el inicio y $n_{e,e}$ el final) y de recuperación (donde $n_{r,o}$ es el inicio y $n_{r,e}$ el final), y sus correspondientes retardos obtenidos con el criterio de MSE entre $QT(n)$ y $\widehat{QT}(n-\tau^*)$.

	HR_{w1}	HR_{w2}	HR_{w3}	α	β	ε_{rms}
noCAD	77.5 ± 13.3	159.9 ± 14.5	94.0 ± 13.6	-0.08 ± 0.01	0.47 ± 0.03	4.6 ± 3.1
Bajo	71.2 ± 14.9	142.6 ± 23.2	83.9 ± 15.8	-0.09 ± 0.02	0.48 ± 0.04	5.8 ± 4.4
Medio	74.1 ± 11.4	128.1 ± 19.7	81.7 ± 14.7	-0.10 ± 0.02	0.50 ± 0.04	5.6 ± 5.2
Alto	65.5 ± 11.1	115.8 ± 18.6	72.7 ± 12.0	-0.10 ± 0.02	0.50 ± 0.04	5.4 ± 4.1

Tabla 1. *Media y desviación estándar de HR (ppm) de cada ventana, y α, β y ε_{rms}(ms) del modelo hiperbólico.*

retardo $\bar{\tau}_e$ (reflejado en la Fig.2) es claramente menor en los grupos noCAD y de bajo riesgo comparado con el de medio ($p=0.01$, $p<0.01$) y alto ($p<0.01$, $p<0.01$) riesgo, siendo acorde a la hipótesis de que una adaptación más lenta está relacionada con el incremento de riesgo cardiaco, aunque el valor medio del retardo en ejercicio del grupo de bajo riesgo es ligeramente menor que el del grupo de noCAD. Por otro lado, en el grupo de medio riesgo se aprecia que el retardo en la fase de recuperación es menor que en la de ejercicio, siendo este comportamiento contrario al resto de grupos. Se puede deber al bajo número de pacientes que componen este grupo. Además, la diferencia entre el tiempo de adaptación entre la fase de ejercicio y de recuperación es mayor en los pacientes noCAD y de bajo riesgo que en los de medio ($p<0.01$, $p=0.01$) y alto ($p<0.01$, $p<0.01$) riesgo, sugiriendo $\Delta_\tau = \tau_r - \tau_e$ también como un marcador de riesgo.

4. Discusión

Cuando se analiza el comportamiento de las series $\widehat{QT}(n)$ y $QT(n)$ se percibe que se superponen (sin retardo entre ellas) cuando se aproximan al pico de estrés en la fase de ejercicio (Fig.1b). Este fenómeno está en concordancia con un estudio reciente [10], el cual muestra que el tiempo de adaptación de la repolarización ventricular a estímulos de la rama simpática disminuye progresivamente conforme más aumentan los niveles de estimulación de los beta-adrenoreceptores. Las simulaciones electrofisiológicas realizadas en dicho estudio incluyen cambios en el Sistema Nervioso Autónomo (ANS) y en el HR, como puede corroborar el comportamiento entre $\widehat{QT}(n)$ y $QT(n)$ observado en este trabajo. Por ello, para el cálculo del retardo, tanto la rampa de ejercicio como la de recuperación se delimitan alejadas del pico de esfuerzo donde la acción del ANS es más regular, de modo que se evite incluir esta zona donde el solapamiento es bastante marcado. Tanto el retardo calculado en ejercicio como en recuperación es bastante similar para los grupos de noCAD y bajo riesgo, y medio y alto riesgo. De modo que estos biomarcadores podrían ser útiles solo para discernir entre riesgo leve o elevado de sufrir CAD.

Los valores medios de los retardos, presentados en la Fig.2, son del mismo orden de magnitud que los reportados en [2], donde la constante de tiempo se estimó en zonas donde HR experimentaba un cambio abrupto (tipo escalón) alcanzando valores entre 36 y 58 segundos, dependiendo del grupo de pacientes estudiado. Estos resultados refuerzan la hipótesis de que el tiempo de retardo medido en este estudio, es decir, en zonas donde los cambios de HR son menos bruscos (tipo rampa) pueden ser equivalentes, representando la misma constante de tiempo. Además, todos los pacientes de [2] eran supervivientes de infarto agudo de miocardio mientras que la característica común de los pacientes de este estudio es que sufren cierto riesgo de padecer CAD, de modo que estamos ante sujetos con patologías de origen diferente. Focalizando en los pacientes de noCAD, $\bar{\tau}_e = 20.96 \pm 22.86$s y $\bar{\tau}_r = 49.05 \pm 23.64$s, donde $\bar{\tau}$ es el valor medio del retardo correspondiente. Estos resultados siguen el mismo comportamiento que los de [11], donde se mide el tiempo medio de adaptación del intervalo QT ante cambios tipo escalón, tanto en aceleración como deceleración de HR,

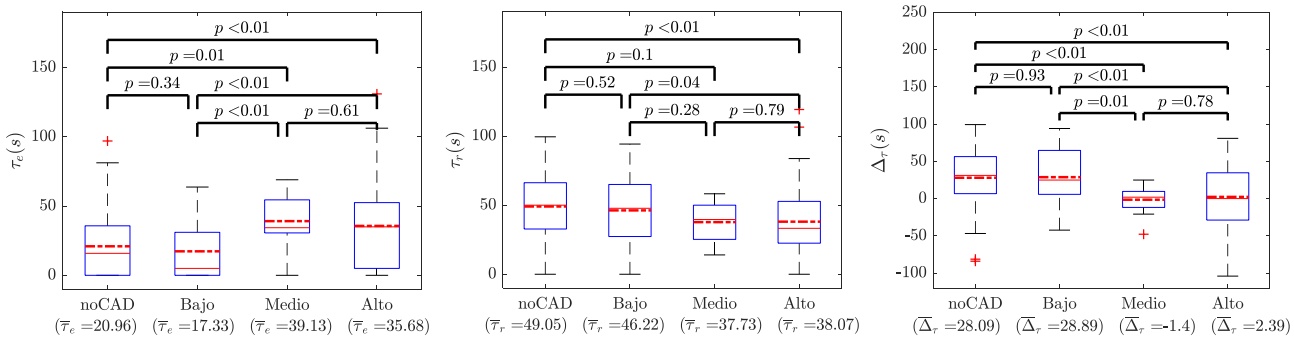


Figura 2. *Box plots, mediana e intercuartiles para cada uno de los retardos entre las series temporales $\widehat{QT}(n)$ y QT(n): en ejercicio (τ_e) y en recuperación (τ_r). Δ_τ corresponde a la diferencia entre los retardos de ejercicio y recuperación definidos. La línea discontinua es el valor medio de τ_e, τ_r o Δ_τ calculado para cada uno de los cuatro grupos, cuyo valor se incluye en el eje horizontal.*

obtenidos mediante cambios postulares en sujetos control: $\bar{\tau}_e = 34.79 \pm 13.63$ y $\bar{\tau}_r = 48.40 \pm 25.70$ s.

5. Conclusión

Se conoce que la histéresis del intervalo QT se puede modelar como la contante de tiempo de un sistema de primer orden en respuesta a una entrada tipo escalón de la señal HR. En este estudio se ha demostrado que este retardo también se puede medir como la constante de tiempo de un sistema de primer orden donde la señal de entrada HR sigue una tendencia tipo rampa, la cual se observa en señales registradas en una prueba de esfuerzo. Este retardo se ha estudiado independientemente para las fases de ejercicio y de recuperación entre las series temporales *QT(n)* y la de QT dependiente simplemente del valor instantáneo de HR ($\widehat{QT}(n)$), representando el valor de QT que ya habría convergido al estacionario para su HR. Esta última señal se ha obtenido ajustando un modelo hiperbólico, empleando valores estables de *RR(n)* y de *QT(n)* dependientes. El retardo en la fase de ejercicio (τ_e) aumenta conforme el riesgo del paciente es mayor, mientras que un comportamiento inverso se observa en los valores del retado en la fase de recuperación (τ_r), pues decrece cuando el riesgo aumenta. De modo que la diferencia entre ambos retardos (Δ_τ) también se propone como marcador, ya que es claramente mayor en el grupo de bajo riesgo y más pequeño conforme el riesgo aumenta.

Agradecimientos

Este trabajo está respaldado por los proyectos ERC-StG 638284 (Consejo Europeo de Investigación), PID2019-105674RB-I00 y PID2019-104881RB-I00 (Ministerio de Ciencia e Innovacion) y el proyecto LMP124-18 y grupo de referencia T39-20R (Gobierno de Aragón co-financiado por FEDER 2014- 2020 “Construyendo Europa desde Aragón”). Los cálculos se realizaron mediante ICTS NANBIOSIS (unidad HPC de la University of Zaragoza).

Referencias

[1] Laguna P, Martínez JP, Pueyo E. Techniques for ventricular repolarization instability assessment from the ECG. Proceedings of the IEEE 2016;104(2):392–415.

[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):1511–1520.

[3] Pueyo E, Husti Z, Hornyik T, Baczkó I, Laguna P, Varró A, Rodríguez B. Mechanisms of ventricular rate adaptation as a predictor of arrhythmic risk. American Journal of Physiology Heart and Circulatory Physiology 2010;298(5):H1577–H1587.

[4] Porta A, Tobaldini E, Gnecchi-Ruscone T, Montano N. RT variability unrelated to heart period and respiration progressively increases during graded head-up tilt. American Journal of Physiology Heart and Circulatory Physiology 2010; 298(5):H1406–H1414.

[5] Kligfield P, Lax KG, Okin PM. QT interval-heart rate relation during exercise in normal men and women: definition by linear regression analysis. Journal of the American College of Cardiology 1996;28(6):1547–1555.

[6] Ogata K. Modern Control Engineering. Instrumentation and controls series. Prentice Hall, 2010. ISBN 9780136156734.

[7] Castells F, Laguna P, Sörnmo L, Bollmann A, Roig JM, et al. Principal component analysis in ecg signal processing. EURASIP Journal on Advances in Signal Processing; 2007.

[8] Martínez JP, Almeida R, Olmos S, Rocha AP, Laguna P. A wavelet-based ecg delineator: evaluation on standard databases. IEEE Transactions on biomedical engineering 2004;51(4):570–581.

[9] Killick R, Fearnhead P, Eckley IA. Optimal detection of changepoints with a linear computational cost. Journal of the American Statistical Association 2012;107(500):1590–1598.

[10] Sampedro-Puente DA, Fernandez-Bes J, Szentandrássy N, Nánasi PP, Taggart P, Pueyo E. Time course of low-frequency oscillatory behavior in human ventricular repolarization following enhanced sympathetic activity and relation to arrhythmogenesis. Frontiers in Physiology 2019;10:1547.

[11] Pueyo E, Malik M, Laguna P. A dynamic model to characterize beat-to-beat adaptation of repolarization to heart rate changes. Biomedical Signal Processing and Control 2008;3(1):29–4.

Contacto:

Cristina Pérez (cperez@unizar.es)
C/María de Luna 1, 50018 Zaragoza, España.

Predicción de muerte súbita en pacientes con insuficiencia cardíaca crónica mediante el estudio de la dinámica periódica de la repolarización

S. Palacios[1], I. Cygankiewicz[2], A. Bayés-de-Luna[3], J.P. Martínez[1,4], E. Pueyo[1,4]

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

[2] Department of Electrocardiology, Medical University of Lodz, Lodz, Poland

[3] Institut Català de Ciències Cardiovasculars, Santa Creu i Sant Pau Hospital, Barcelona, Spain

[4] CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Zaragoza, Spain

Resumen

La insuficiencia cardíaca crónica (ICC) es un síndrome clínico con una alta mortalidad debida a la muerte por fallo de bomba (MFB), y al desarrollo de arritmias ventriculares que pueden provocar muerte súbita cardíaca (MSC). Los pacientes con ICC experimentan un desequilibrio en el sistema nervioso autónomo que podría reflejarse en el electrocardiograma (ECG). La Dinámica Periódica de la Repolarización ("Periodic Repolarization Dynamics", PRD) cuantifica las oscilaciones de baja frecuencia en la onda T del ECG y se ha relacionado con la modulación simpática de la repolarización ventricular. Este estudio evalúa la capacidad predictiva de PRD para MFB y MSC en una población con ICC. Se analizaron ECGs de 20 minutos y 3 derivaciones de 569 pacientes con ICC sintomática y ritmo sinusal normal. PRD se midió en segmentos de 5 minutos, con solapamientos de 4 minutos, asignándose a cada registro el mínimo valor entre ellos. Se encontró que el PRD presentaba valores más elevados en víctimas de MSC en comparación al resto de pacientes, aunque las diferencias no fueron estadísticamente significativas. Tomando como umbral la mediana de los valores de PRD en la población global, se definieron grupos de bajo y alto riesgo. Se obtuvo un cociente de riesgo de MSC [Intervalo de confianza al 95%] en el análisis univariado de Cox de 1.808 [1.031-3.169] grados (p=0.039). En conclusión, PRD predice el riesgo de MSC en pacientes con ICC, presentando las víctimas de MSC una mayor magnitud de las oscilaciones de la repolarización ventricular inducidas por la activación simpática.

1. Introducción

La insuficiencia cardíaca crónica (ICC) es un síndrome cardíaco complejo que conlleva una elevada mortalidad y un deterioro importante en la calidad de vida de los pacientes que la padecen. En todo el mundo, alrededor de 23 millones de personas sufren ICC y para el 2030 se estima que este número habrá aumentado en más de 8 millones [1]. La ICC está caracterizada por anomalías estructurales y funcionales del corazón, las cuales reducen su capacidad para bombear y satisfacer la demanda de sangre del cuerpo humano. Se ha demostrado que existe un desequilibrio autonómico en los pacientes con ICC, consistente en un aumento de la actividad simpática y una reducción de la actividad parasimpática, que supone un peor pronóstico de la ICC [2].

La muerte súbita cardíaca (MSC), definida como la muerte provocada por la pérdida de la función cardíaca, representa la principal causa de mortalidad entre pacientes con ICC. Otra causa común de muerte asociada a la ICC es la muerte por fallo de bomba (MFB), que se produce cuando hay una disminución progresiva del gasto cardíaco y de la capacidad de perfusión de los órganos por el bombeo del corazón. La predicción de estas dos formas de muerte es de gran importancia para el diseño de terapias personalizadas que permitan mejorar la calidad de vida de los pacientes.

Para estratificar el riesgo de MSC y MFB en pacientes con ICC, se han propuesto distintos marcadores no invasivos derivados del electrocardiograma (ECG). Entre estos marcadores, cabe señalar el intervalo QT [3], el intervalo Tpeak-to-Tend [4] o las alternancias de la onda T [5]. Estudios recientes han propuesto la Dinámica Periódica de la Repolarización (en inglés "Periodic Repolarization Dynamics", PRD) para evaluar la modulación simpática de la repolarización ventricular midiendo las oscilaciones de baja frecuencia (menores a 0.1 Hz) en el vector de la onda T [6]. Se ha demostrado la capacidad del índice PRD para predecir la mortalidad total y por causas arrítmicas, así como para identificar el beneficio de la implantación de desfibriladores automáticos en pacientes que han sufrido infarto de miocardio [6,7]. El objetivo de este trabajo es evaluar la capacidad del índice PRD para estratificar pacientes con ICC según el riesgo de fallecer por MFB o MSC.

2. Población de estudio

La población de estudio se compone de 569 pacientes (63±12 años, 28,1% mujeres) con ICC sintomática, clases II y III de la New York Heart Association (NYHA), participantes del estudio MUSIC (Muerte Súbita en Insuficiencia Cardíaca). MUSIC es estudio prospectivo, multicéntrico y longitudinal diseñado para evaluar distintos predictores de riesgo de muerte cardíaca y MSC en pacientes ambulatorios con ICC [8].

Se adquirieron registros ECG de alta resolución (1000 Hz) y de 20 minutos de duración en estado basal por medio de los dispositivos SpiderView (ELA medical, SorinGroup, París, Francia). Para cada sujeto se obtuvieron y analizaron tres derivaciones ortogonales.

Se realizó un seguimiento clínico a los pacientes cada 6 meses de forma ambulatoria durante un promedio de 48

ISBN: 978-84-09-25491-0, pags. 10- 13

meses. El grupo de estudio incluía 53 víctimas de MSC, 53 víctimas de MFB, 37 víctimas de muertes no cardíacas y 426 supervivientes. La muerte se definió como MSC si correspondía a: (i) una muerte con testigos dentro de los 60 minutos desde el inicio de nuevos síntomas, a menos que sea obvia una causa distinta a la cardíaca; (ii) una muerte sin testigos (< 24 horas) en ausencia de insuficiencia circulatoria progresiva preexistente u otras causas de muerte; o (iii) una muerte durante un intento de resucitación. Las causas de defunción fueron revisadas y clasificadas por un comité de expertos del estudio MUSIC. El protocolo de estudio fue aprobado por los comités institucionales y todos los pacientes dieron su consentimiento informado por escrito.

3. Métodos

3.1. Preprocesado

El preprocesamiento de las señales de ECG consistió en la eliminación de ruido mediante un filtro elimina-banda a 50 Hz, la detección y el etiquetado de latidos cardíacos mediante el software Aristotle ECG [9]. Dado que el análisis posterior se centró en la onda T, se aplicó un filtro paso-bajo de 40 Hz para eliminar el ruido sin que la forma de la onda T se viese alterada. Finalmente, se aplicó interpolación mediante splines cúbicos para estimar y eliminar la línea de base.

Se definió el inicio y final (T_{on_i} y T_{off_i} respectivamente) de ventanas que incluyeran las ondas T en cada latido. Estos valores se han definido, según las siguientes reglas, tomando como referencia las marcas de QRS (QRS_i) y considerando los intervalos RR correspondientes (RR_i), donde *i* denota el índice del latido:

- Si $RR_i < 720$ ms,

$$T_{on_i} = QRS_i + 90 \text{ ms}$$

$$T_{off_i} = QRS_i + \min\left(360, \frac{2}{3} RR_i\right) \text{ms}$$

- Si $RR_i \geq 720$ ms,

$$T_{on_i} = QRS_i + 90 \text{ ms}$$

$$T_{off_i} = QRS_i + 360 \text{ ms}$$

Se estimó el nivel de ruido en cada onda T a partir de las componentes de alta frecuencia (por encima de 15Hz) [10]. Cuando un latido presentaba una baja calidad de señal (nivel de ruido > 140μV), tanto el inicio como el final de la ventana de la onda T se han descartado, pero la posición del complejo QRS correspondiente se ha mantenido.

En la detección de las marcas de QRS se produjeron dos tipos de incidencias: complejos QRS que se identificaron erróneamente donde no había latido y latidos no detectados. Para solucionar estas incidencias se utilizó el algoritmo automático descrito en [11]. Posteriormente, si dos marcas consecutivas se encontraban separadas entre sí por menos de 50 ms, la segunda marca se descartó.

3.2. Dinámica Periódica de la Repolarización

Se calcularon los ángulos entre ondas T consecutivas a partir de las señales ECG preprocesadas utilizando un método basado en el descrito originalmente en [6] pero actualizado según se propone en [12]. Los pasos para el cálculo de PRD fueron los siguientes:

1. Se definieron los límites de las ventanas de la onda T según lo indicado en la sección anterior.
2. Se restó un valor constante en cada derivación de la ventana de la onda T, de tal manera que la amplitud del final de la ventana fuera 0 mV.
3. Se calculó un vector promedio para cada una de las ventanas de la onda T. Se calculó el ángulo dTº entre dos ondas T consecutivas mediante el producto escalar entre los dos vectores promedios de las ondas T. Este ángulo representa el grado instantáneo de inestabilidad en la repolarización.
4. Se aplicó un filtro de mediana de orden 10 sobre la serie temporal dTº para atenuar los valores atípicos y evitar los cambios excesivamente abruptos.

Para obtener el valor de PRD a partir de las series dTº se usó un método basado en *Phase-Rectified Signal Averaging* (PRSA) [13], que consiste en:

5. Se definieron puntos de anclaje en la serie dT a partir de la comparación de las medias de M=9 valores anteriores y posteriores a los candidatos a puntos de anclaje x_i. Un latido *i* se considera como punto de anclaje si:

$$\frac{1}{M}\sum_{j=0}^{M-1} x_{i+j} > \frac{1}{M}\sum_{j=1}^{M} x_{i-j}$$

   El uso de M=9 permite detectar frecuencias en el rango de interés para este trabajo (entre 0.025 y 0.1 Hz).
6. Se consideraron ventanas de W=40 muestras alrededor de cada punto de anclaje. Si un punto de anclaje se encontraba cerca del principio o del final de la serie dTº, no hay suficientes muestras a su alrededor para tener una ventana de tamaño W, procedía a descartarse ese punto de anclaje.
7. Se obtuvo la serie de PRSA mediante el promediado entre todas las ventanas de longitud W definidas anteriormente.

Se definió PRD como la diferencia entre los valores máximo y mínimo de la serie de PRSA. Cada registro ECG de 20 minutos de duración se dividió en segmentos de 5 minutos, con 4 minutos de solapamiento. Para cada segmento, se calculó el valor de PRD. Finalmente, a cada paciente se le asoció el mínimo valor de PRD entre todos los segmentos calculados.

3.3. Análisis estadístico

Los valores de PRD obtenidos para los diferentes subgrupos de pacientes se presentan en términos de mediana [rango intercuartílico]. Se utilizó la prueba U de Mann-Whitney (o prueba de suma de rangos de Wilcoxon) para las comparaciones univariadas de los valores de PRD entre grupos de pacientes. Utilizando el método de Kaplan-Meier se ha estimado la probabilidad de supervivencia a lo largo del tiempo, representándola gráficamente. La comparación entre los tiempos de supervivencia se ha realizado mediante la prueba logrank. Se ha utilizado el

modelo de regresión de riesgos proporcionales de Cox para la predicción de MSC, presentando los cocientes de riesgo y los intervalos de confianza al 95%. Se han considerado estadísticamente significativos los resultados asociados con p-valores menores de 0,05. Las variables que mostraron diferencias significativas en la predicción de los tiempos de supervivencia se utilizaron posteriormente en el modelo de regresión de Cox. Para el análisis de los datos se utilizaron MATLAB R2017a (9.2) y SPSS (versión 24.0).

4. Resultados

4.1. Asociación con la mortalidad cardíaca

En la figura 1 se muestran los diagramas de cajas correspondientes a las distribuciones estadísticas de PRD para tres subgrupos de pacientes: víctimas de MSC, víctimas de MFB y el resto de pacientes (supervivientes y víctimas de muerte no cardíaca). Aunque no se han encontrado diferencias estadísticamente significativas entre ninguno de los grupos, puede observarse que PRD tiende a tomar valores más altos en las víctimas de MSC (con valor de mediana [y rango intercuartílico] de 1,57 [1,21] grados) que en las víctimas de MFB (1,41 [0,91]) e incluso más altos que en el subgrupo de los supervivientes y víctimas de muerte no cardíaca (1,24 [1,28] grados).

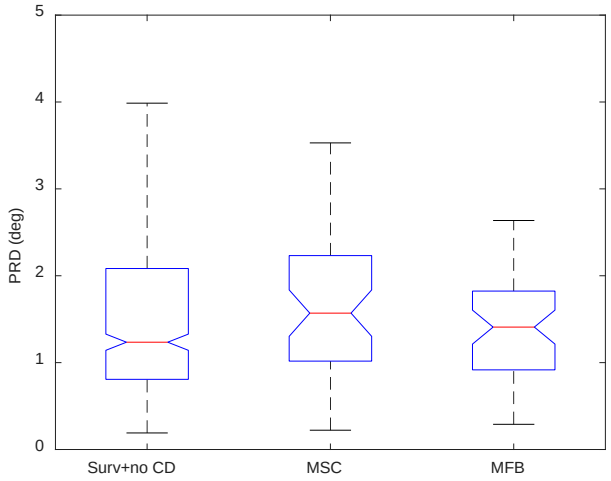


Figura 1.*Diagramas de caja para tres subgrupos de pacientes: MSC, MFB y el resto de pacientes (supervivientes y víctimas de muerte no cardíaca).*

4.2. Análisis de supervivencia

Se tomó el valor mediano de PRD en la población de estudio como umbral para definir grupos de bajo riesgo (PRD <1,76 grados) y de alto riesgo (PRD ≥ 1,76 grados). Cuando se consideró la MFB como causa de defunción, no se encontraron diferencias significativas en los tiempos de supervivencia entre los dos grupos. En el caso de MSC, las diferencias en los tiempos de supervivencia entre los grupos de bajo y alto riesgo fueron estadísticamente significativas (p=0,036). En la figura 2 se presentan las curvas de probabilidad de supervivencia a MSC a lo largo del seguimiento.

En el modelo univariado de regresión de Cox que utiliza el marcador PRD como variable predictiva, el cociente de riesgo [intervalo de confianza del 95%] para MSC fue 1,808 (1,031-3,169), p=0,039.

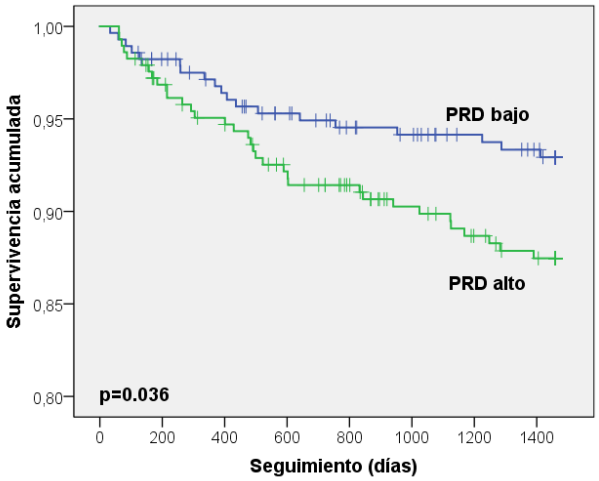


Figura 2. *Probabilidad estimada de supervivencia a MSC para dos subgrupos de pacientes dicotomizados según el valor de PRD.*

5. Discusión y conclusiones

En este trabajo se han medido de forma no invasiva las oscilaciones de baja frecuencia en la onda T del ECG, cuantificando el marcador PRD en registros ECG de 20 minutos pertenecientes a pacientes con ICC. El PRD ha sido asociado en la literatura con la modulación simpática de la actividad eléctrica ventricular [6]. Estudios clínicos, experimentales y teóricos previos han demostrado que la activación simpática conduce a un incremento en la magnitud de las oscilaciones de baja frecuencia de la repolarización ventricular, evaluadas a partir de la onda T del ECG, de los intervalos de recuperación-activación de los electrogramas y de los potenciales de acción celulares [7, 14-16]. Se ha postulado que los posibles mecanismos subyacentes a esos aumentos en las oscilaciones inducidas por la activación simpática involucran efectos β-adrenérgicos y mecánicos que actúan de forma sinérgica [14,17].

Cuando se han comparado los valores de PRD por separado en las víctimas de MFB, víctimas de MSC y el resto de pacientes, no se han encontrado diferencias estadísticamente significativas, pero se ha observado que los valores más altos de PRD se hallan en víctimas de MSC, lo que posiblemente podría ser indicativo de una mayor modulación simpática de la repolarización en estos pacientes. Después de dicotomizar la población de estudio en grupos de bajo y alto riesgo según la mediana de PRD, se determinó que la probabilidad de fallecer por MSC durante el periodo de seguimiento es casi el doble en el grupo de alto riesgo que en el grupo de bajo riesgo (p=0.036). Esto concuerda con estudios clínicos y teóricos que han mostrado una relación entre el aumento de PRD y el aumento del riesgo arrítmico [6,7,14,17].

Una característica interesante de este trabajo es que PRD se ha calculado en registros ECG de 20 minutos de duración, grabados en reposo, lo que facilita la generalización de la evaluación de PRD para la estratificación del riesgo arrítmico. Otros índices derivados

de ECG, como la turbulencia de la frecuencia cardíaca, la capacidad de desaceleración [18] o las alternancias de la onda T [5] suelen medirse a partir de registros ECG ambulatorios o bien siguiendo protocolos específicos como pruebas de esfuerzo, que no se realizan de forma tan generalizada como los ECGs de reposo.

En conclusión, este estudio indica que PRD, un marcador no invasivo de la inestabilidad de la repolarización medido a través de las oscilaciones de baja frecuencia de la onda T, tiene potencial para predecir MSC en pacientes con ICC a partir de registros cortos de ECG en reposo. Valores elevados de PRD son indicativos de un mayor riesgo de fallecer por MSC.

Agradecimientos

Este trabajo ha sido realizado con el apoyo de los proyectos ERC-StG 638284 (European Research Council), PID2019-105674RB-I00 y PID2019-104881RB-I00 (Ministerio de Ciencia e Innovación) y el proyecto LMP124-18 y el grupo de referencia BSICoS T39-20R (Gobierno de Aragón cofinanciado por el FEDER 2014-2020 “Construyendo Europa desde Aragón”). Los cálculos computacionales fueron realizados en la ICTS NANBIOSIS (HPC Unit at University of Zaragoza).

Referencias

[1] Lippi G, Sanchis-Gomar F. Global epidemiology and future trends of heart failure. *AME Medical Journal*, vol 5, 2020, pp 1-6.

[2] Kishi T. Heart failure as an autonomic nervous system dysfunction. *Journal of Cardiology*, vol 59, no 2, 2012, pp 117-22.

[3] Galinier M et al. QT interval dispersion as a predictor of arrhythmic events in congestive heart failure: Importance of aetiology. *European Heart Journal*, vol 19, no 7, 1998, pp 1054-62.

[4] Chua KCM et al. Tpeak-to-Tend interval corrected for heart rate: A more precise measure of increased sudden death risk? *Heart Rhythm*, vol 13, no 11, 2016, pp 2181-85.

[5] Daniłowicz-Szymanowicz L et al. Microvolt T-wave alternans and autonomic nervous system parameters can be helpful in the identification of low-arrhythmic risk patients with ischemic left ventricular systolic dysfunction. *PLOS ONE*, vol 13, no 5, 2018, pp 1-15.

[6] Rizas KD et al. Sympathetic activity–associated periodic repolarization dynamics predict mortality following myocardial infarction. *The Journal of Clinical Investigation*, vol 124, no 4, 2014, pp 1770-80.

[7] Rizas KD et al. Prediction of sudden and non-sudden cardiac death in post-infarction patients with reduced left ventricular ejection fraction by periodic repolarization dynamics: MADIT-II substudy. *European Heart Journal*, vol 38, no 27, 2017, pp 2110-8.

[8] Vazquez R et al. The MUSIC Risk score: a simple method for predicting mortality in ambulatory patients with chronic heart failure. *European Heart Journal*, vol 30, no 9, 2009, pp 1088-96.

[9] Moody GB, Mark RG. Development and evaluation of a 2-lead ECG analysis program. *Computers in Cardiology*, vol 9, 1992, pp 39-44.

[10] Mincholé A et al. Quantification of restitution dispersion from the dynamic changes of the T-wave peak to end, measured at the surface ECG. *IEEE Transactions on Biomedical Engineering*, vol 58, no 5, 2011, pp 1172-82.

[11] Mateo J, Laguna P. Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal. *IEEE Transactions on Biomedical Engineering*, vol 50, no 3, 2003, pp 334-43.

[12] Palacios S et al. Long-term microgravity exposure increases ECG repolarization instability manifested by low-frequency oscillations of T-wave vector. *Frontiers in Physiology*, vol 10, 2019, pp 1510.

[13] Bauer A et al. Phase-rectified signal averaging detects quasi-periodicities in non-stationary data. *Physica A: Statistical Mechanics and its Applications*, vol 364, 2006, pp 423-34.

[14] Pueyo E et al. Interactive effect of beta-adrenergic stimulation and mechanical stretchon low-frequency oscillations of ventricular action potential duration in humans. *Journal of Molecular and Cellular Cardiology*, vol 97, 2016, pp 93-105.

[15] Porter B et al. Beat-to-beat variability of ventricular action potential duration oscillates at low frequency during sympathetic provocation in humans. *Frontiers in Physiology*, vol 9, 2018, pp 147.

[16] Duijvenboden SV et al. Complex interaction between low-frequency apd oscillations and beat-to-beat apd variability in humans is governed by the sympathetic nervous system. *Frontiers in Physiology*, vol 10, 2020, pp 1582.

[17] Sampedro-Puente DA et al. Mechanisms underlying interactions between low-frequency oscillations and beat-to-beat variability of celullar ventricular repolarization in response to sympathetic stimulation: Implications for arrhythmogenesis. *Frontiers in Physiology*, vol 10, 2019, pp 916.

[18] Al-Zaiti SS et al. The role of heart rate variability, heart rate turbulence, and deceleration capacity in predicting cause-specific mortality in chronic heart failure. *Journal of Electrocardiology*, vol 52, 2019, pp 70-4.

Análisis multidimensional de las 12 derivaciones del ECG estándar para predecir el éxito de la cardioversión eléctrica en fibrilación auricular persistente

E.M. Cirugeda Roldán[1], S. Calero Núñez[2], E. Plancha Burguera[3], J. Enero Navajo[2], J.J. Rieta Ibáñez[4], R. Alcaraz Martínez[1]

[1] Grupo de Investigación en Ingeniería Electrónica, Biomédica y de Telecomunicación, Universidad de Castilla-La Mancha, Cuenca, España, {eva.cirugeda, raul.alcaraz}@uclm.es
[2] Unidad de Arritmias, Complejo Hospitalario Universitario de Albacete, Albacete, España
[3] FISABIO – Servicio de Medicina Interna, Hospital Francesc de Borja, Gandía, España
[4] BioMIT.org, Dep. de Ingeniería Electrónica, Universitat Politecnica de Valencia, Valencia, España

Resumen

La Sociedad Europea de Cardiología recomienda la cardioversión eléctrica (CVE) para el control del ritmo cardíaco en situaciones de fibrilación auricular (FA) persistente. Aunque es capaz de restaurar inicialmente el ritmo sinusal (RS) en la mayoría de los pacientes, su tasa de éxito a medio plazo es limitada, presentando la arritmia una alta probabilidad de recurrencia. En este contexto, es de gran interés poder prever el resultado de la CVE, ya que de esta forma se podría tratar individualizadamente a los pacientes. Con este objetivo, durante los últimos años se han propuesto distintos índices que caracterizan la actividad auricular (AA) sobre el electrocardiograma (ECG), tales como la amplitud de las ondas fibrilatorias (f), su frecuencia dominante, o su regularidad. Aunque estos índices han presentado resultados prometedores, solo se han empleado para caracterizar las ondas f en una única derivación, generalmente V1, descartando así la información espacial que, registrada en otras derivaciones, puede resultar de gran interés. Así pues, este trabajo evalúa si una extensión multidimensional de dichos índices es capaz de incrementar la capacidad de predicción del resultado de la CVE. En los resultados se observa que las extensiones multidimensionales de los índices propuestos incrementan la predicción de la CVE hasta casi un 6 %, lo que sugiere que las correlaciones espaciales en la AA de las distintas derivaciones del ECG obtienen información predictiva relevante sobre la recurrencia de la FA.

1. Introducción

A pesar de los recientes avances en el tratamiento de la fibrilación auricular (FA), esta arritmia sigue siendo una de las mayores causas de morbilidad y mortalidad en el mundo [1]. Además, es responsable de un deterioro en la calidad de vida de los pacientes y presenta altas tasas de hospitalización [2]. La Sociedad Europea de Cardiología recomienda el uso de técnicas orientadas al control del ritmo cardíaco para restaurar el ritmo sinusal (RS) en pacientes con FA y, consecuentemente, reducir la sintomatología asociada, mejorar la calidad del vida del paciente, y preservar la función cardíaca [2]. Entre las diferentes técnicas existentes, la cardioversión farmacológica es capaz de revertir la FA en un 50 % de los pacientes [2]. Sin embargo, la cardioversión eléctrica (CVE) ha demostrado ser más efectiva, restaurando inicialmente el RS en casi el 90 % de los pacientes [2, 3], a la vez que reduce el número de hospitalizaciones. Así pues, no resulta extraño encontrar una tendencia creciente en la aplicación de este procedimiento en medicina ambulatoria [4].

Aunque la capacidad de la CVE para inicialmente restaurar el RS es muy alta, las tasas de recurrencia a FA también lo son. Hasta un 30 % de los pacientes lo hacen durante las 4 primeras semanas, y casi el 80 % antes del primer año [5]. Así pues, identificar si existe una alta probabilidad de recurrencia es crucial para tratar de forma óptima a todos los pacientes [6]. De hecho, la CVE se podría evitar en aquellos sujetos cuya condición proarrítmica tras el procedimiento todavía fuera alta y, así prevenir el desarrollo de patologías asociadas, así como los riesgos propios inherentes al procedimiento. Aunque la CVE es un proceso no invasivo que no presenta grandes riesgos, está directamente relacionada con ciertos efectos secundarios, tales como hipotensión, bradicardias y tromboembolismos [5].

Con el objetivo de evaluar la condición proarrítmica de un paciente y así poder anticiparse al resultado de la CVE, hasta la fecha se han propuesto distintos parámetros que caracterizan las ondas fibrilatorias (f) presentes en la actividad auricular (AA) extraída a partir del electrocardiograma (ECG). Entre ellos, la amplitud de las ondas f (A_f), su frecuencia dominante ($Fdom$), o su regularidad, estimada a través de la entropía muestral (Em), se han utilizado para caracterizar las ondas f en la derivación V1, con resultados bastante prometedores [7, 8]. No obstante, aunque V1 se considera que registra las ondas f de mayor amplitud comparadas con otras componentes ventriculares, dada su proximidad a la aurícula derecha [9], estos índices descar-

ISBN: 978-84-09-25491-0, pags. 14- 17

tan la información espacial relativa a la actividad cardíaca presente en el resto de derivaciones. Así pues, este trabajo explora si una extensión multidimensional de estos índices es capaz de estimar de forma más precisa la condición proarrítmica del paciente y, por tanto, predecir de forma más eficaz el resultado de la CVE.

2. Materiales y métodos

2.1. Base de datos

Para alcanzar el objetivo previsto, se han analizado 58 pacientes diagnosticados con FA persistente. Durante la CVE los pacientes se monitorizaron mediante un registro de ECG de 12 derivaciones, adquirido con una frecuencia de muestreo de 1024 Hz y 16 bits de resolución.

El protocolo de CVE consistió en la aplicación de un máximo de cuatro descargas eléctricas sobre el torso del paciente con una intensidad de 200, 300, 360 y 360 J, respectivamente. Para ello, una pala se colocó parasternalmente en el segundo espacio intercostal derecho, y la otra lateralmente a lo largo de la línea axilar media. Todos los pacientes revirtieron a RS tras la CVE, pero 31 de ellos recurrieron a FA tras un seguimiento de 4 semanas. La información clínica de estos sujetos se muestra en la Tabla 1.

2.2. Extracción de la actividad auricular

La AA se extrajo de forma independiente sobre cada una de las derivaciones del ECG mediante un algoritmo de cancelación adaptativa de la actividad ventricular [10]. Cada derivación del ECG, de al menos 90 segundos de duración, fue previamente preprocesada para eliminar la línea de base, la interferencia de red, y el ruido de alta frecuencia. La detección de los complejos QRS [11], y su posterior corrección manual, se realizó también de forma independiente sobre ellas. Los latidos se clasificaron en base a su morfología [12], tal que la cancelación de los complejos QRST se realizó de forma secuencial, comenzando por aquellos perteneciendo a la morfología con menos componentes. La duración del complejo QRST se estableció como la mínima entre 470 ms (valor típico) y el 90 % del intervalo RR mediano. La señal resultante, se filtró paso alto a 3 Hz para eliminar cualquier actividad ventricular residual.

2.3. Caracterización de la actividad auricular

Considerando que $f(n)$ es la señal de AA que contiene las ondas f de una derivación, con una longitud de N muestras (es decir, $n = 1, 2, \ldots, N$), su amplitud (A_f) se calculó como [7]:

$$A_f = \sqrt{\frac{1}{N}\sum_{n=1}^{N} f(n)^2}. \qquad (1)$$

Parámetro	FA no recidiva	FA recidiva
Número de Pacientes	27	31
Hombres / Mujeres	15 / 12	18 / 13
Patología cardíaca subyacente	9	10
Diámetro aurícula izquierda (mm)	47.32± 4.76	44.72 ± 7.32

Tabla 1. *Características clínicas de los pacientes analizados.*

Su extensión multiderivacional (MA_f) se realizó de acuerdo a [8]. Así pues, $f(n)$ se aproximó por su descomposición en componentes principales de primer rango. Entonces, la MA_f se estimó como el valor mediano resultante de aplicar la ecuación (1) a cada una de las 12 derivaciones.

La estimación unidimensional de la frecuencia dominante, $Fdom$, se obtuvo para cada una de las 12 derivaciones, como la frecuencia en la cual la densidad espectral de potencia media ($\bar{P}$) fue máxima en la banda 3–12 Hz [7]:

$$Fdom = \arg\{\max_{f_k=3-12\ \mathrm{Hz}}\{\bar{P}(f_k)\}\}. \qquad (2)$$

Se debe mencionar que $\bar{P}$ denota la media de aquellos periodogramas de Welch cuyo coeficiente de correlación cruzada era superior a 0.7, de forma que se eliminó el efecto del ruido en su estimación. Los periodogramas, de 4096 muestras se calcularon sobre segmentos de la AA de 6 segundos de duración con un solapamiento de 4 segundos, considerando una ventana de Hamming y un 50 % de solape. La extension multiderivacional de este parámetro frecuencial, M$Fdom$, se obtuvo aplicando la ecuación (2) sobre la envolvente espectral, obtenida a partir de las 12 derivaciones mediante el procedimiento descrito en [13].

La organización de las ondas f se estimó a partir de la entropía muestral, Em, la cual es un estadístico diseñado para contabilizar la regularidad interna que existe en una serie aparentemente no lineal, asociando valores de entropía más altos a señales dinámicamente más complejas. Este índice se define como la probabilidad condicional de que dos secuencias que son similares para m muestras, con tolerancia r, lo sigan siendo para una longitud de una muestra más [14]. Como en trabajos previos [7], la regularidad de las ondas f se estimó aplicando la Em sobre la onda auricular principal, $ff(n)$, la cual se obtuvo filtrando $f(n)$ paso banda en torno a la $Fdom$ con un ancho de banda de 5 Hz. La Em se estimó entonces según [14], es decir,

$$Em(ff, m, r, N) = -\ln \frac{B^{m+1}(r)}{B^m(r)}, \qquad (3)$$

donde $B^m(r)$ y $B^{m+1}(r)$ representan la probabilidad de encontrar subsecuencias de longitud m y $m+1$, respectivamente, que son similares con una tolerancia r. Dicha probabilidad se calculó de acuerdo a:

$$B^m(r) = \frac{1}{N-m}\sum_{i=1}^{N-m} B_i^m(r). \qquad (4)$$

$B^{m+1}(r)$ se calcula de igual forma, pero para secuencias de longitud $m+1$, teniendo en cuenta que el número de secuencias similares responde a:

$$B_i^m(r) = \frac{1}{N-m-1}\sum_{\substack{j=1\\ j\neq i}}^{N-m} (d_{ij}(m) < r), \qquad (5)$$

donde $d_{ij}(m)$ representa la distancia de Chebyshev existente entre dos subsecuencias de longitud m, y responde

a la diferencia máxima entre los valores escalares de cada subsecuencia, es decir:

$$d_{ij}(m) = \max_{k} \{|ff(i+k) - ff(j+k)|, k=0:m-1\}. \quad (6)$$

La extensión multiderivacional (MEm) de esta entropía presenta el mismo proceso de cálculo que la Em, pero al tratarse de una señal multiderivacional hay que tener en cuenta varios aspectos: (i) la amplitud de las ondas f varía con respecto a las derivaciones, (ii) los subsegmentos de m muestras pueden evolucionar en 12 direcciones distintas, y (iii) al tratarse de una medida relativa es necesario realizar una generalización multivariable del umbral de la tolerancia r. Por todas estas cuestiones, se normalizó la AA de cada derivación a un rango unitario ([0,1]), y el umbral de tolerancia r se estableció como la varianza total de la matriz que contenía la AA de las 12 derivaciones. Una descripción más detallada del proceso para calcular la MEn se puede encontrar en [15]. La selección de los parámetros se hizo siguiendo las recomendaciones en [14, 15], de forma que N = 30 segundos, m = 2 y r = 0.2.

2.4. Análisis estadístico

La capacidad de predicción del resultado de la CVE de cada uno de los índices propuestos se analizó a través de una curva ROC. Esta curva ofrece información sobre el número de pacientes que revirtiendo a FA fueron clasificados como tal, es decir de la sensibilidad (Se), y del ratio de pacientes que siendo capaces de mantener un RS fueron clasificados de forma correcta, es decir de la especificidad (Sp), para cada posible umbral de clasificación. El umbral óptimo se determinó a partir del criterio de Youden, y se calculó también el área bajo la curva (AROC), ya que se considera una estimación global de a capacidad predictiva de un índice.

Desde un punto de vista estadístico, el test de Kolmogorov-Smirnov se utilizó para determinar si los datos se ajustaban a una distribución normal. En tal caso, la prueba t-Student se utilizó para evaluar la capacidad del índice para proveer diferencias estadísticamente significativas entre los pacientes que revirtieron a FA y aquellos que mantuvieron un RS. En caso contrario, se utilizó la prueba de Wilcoxon.

3. Resultados

La Figura 1 presenta las distribuciones de los índices propuestos. La mayor diferencia observada entre los índices unidimensionales y multidimensionales es que estos últimos presentaron distribuciones más uniformes. No obstante, excepto para la A_f, el comportamiento de los índices fue similar en ambos casos, aunque se apreciaron valores superiores de $Fdom$ y Em que de sus extensiones multidimensionales. De hecho, en ambos casos se observaron valores más altos para los pacientes con FA recidiva que para aquellos que consiguieron mantener un RS. En el caso de la A_f, se observaron distintas tendencias en los datos. Así pues, mientras que la A_f obtuvo valores medianos inferiores para los pacientes que mantuvieron el RS, en el caso de MA_f se obtuvieron valores medianos superiores. También cabe destacar que el índice MEm presentó valores inferiores con respecto a la Em. Todas las métricas

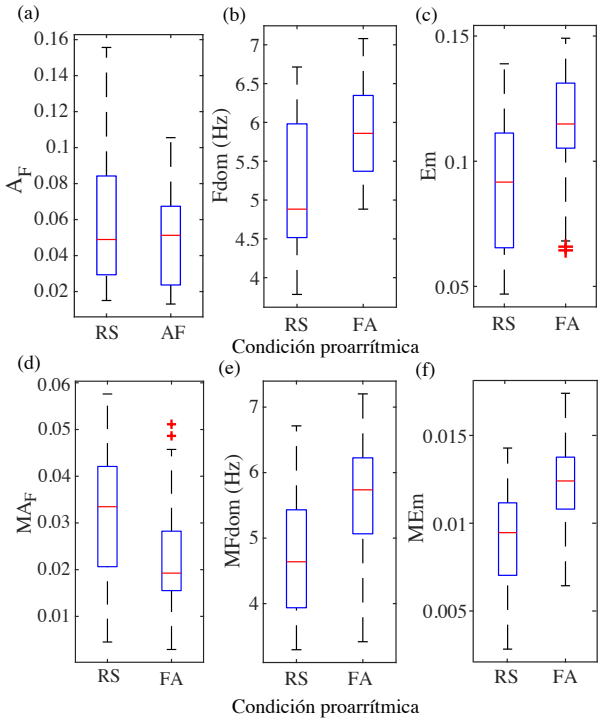


Figura 1. *Distribución de los índices (a) A_f, (b) $Fdom$ y (c) Em, y su extensión multiderivacional (d) MA_f, (e) M$Fdom$ y (f) MEm, en función de la condición proarrítmica del paciente.*

analizadas, excepto la A_f sobre V1, reportaron diferencias estadísticamente significativas (p <0.05).

En la Tabla 2 se muestra la caracterización estadística, los valores medianos y los rangos intercuartílicos, así como la habilidad para predecir el resultado de la CVE en términos de las curvas ROC, para cada uno de los índices analizados. Cabe destacar que se observó un incremento en los valores de la AROC para los índices multiderivacionales con respecto a los uniderivacionales, notándose también un mejor balance entre Se y Sp.

4. Discusión y conclusiones

De todas las métricas propuestas, únicamente la A_f no ha reportado diferencias estadísticamente significativas, además de presentar un comportamiento distinto respecto a su extension multiderivacional. Mientras que los valores medianos de A_f fueron mayores para los pacientes con FA recidiva, la MA_f mostró valores medianos inferiores para el mismo grupo. En trabajos anteriores [7, 8, 16], este índice se ha utilizado con resultados bastante prometedores y en todos ellos se ha asociado una mayor A_f a una mayor probabilidad de mantener un RS, algo que aquí solo se observa para la MA_f. Este resultado puede deberse a las diferencias en los algoritmos considerados, ya que en [8] se utilizó un algoritmo basado en envolventes y en [7, 16] se realizó una inspección visual, mientras que aquí A_f se estimó de forma no supervisada. En el caso de la MA_f, sí se obtuvieron diferencias estadísticamente significativas y un comportamiento similar al citado en [7, 8, 16], a la vez que se incrementó la capacidad de clasificación un 2 % y se obtuvieron valores de Se y Sp más balanceados.

Index	Uniderivacional					Multiderivacional				
	RS	FA	Se(%)	Sp(%)	AROC(%)	RS	FA	Se(%)	Sp(%)	AROC(%)
A_F	0.049 (0.055)	0.051 (0.044)	70.4	58.1	70.2	0.033 (0.021)	0.019 (0.013)	62.9	77.4	72.5
$Fdom$	4.761 (1.465)	6.104 (0.977)	80.6	63.0	76.6	4.639 (1.221)	5.859 (0.824)	77.4	74.1	77.4
Em	0.089 (0.043)	0.115 (0.026)	80.6	63.0	74.6	0.009 (0.004)	0.013 (0.003)	74.2	77.8	80.5

Tabla 2. *Valores medianos y rangos intercuartílicos en función de la condición proarrítmica. Capacidad clasificatoria, sensibilidad (Se), especificidad (Sp) y exactitud (AROC) de los índices uniderivacionales y su extensión multiderivacional.*

La M$Fdom$ no ha reportado una mejora significativa con respecto a la $Fdom$, pero sí presenta, al igual que la MA_f, una curva ROC con un mejor balance entre Se y Sp. El comportamiento en ambos casos es similar, asociando frecuencias más altas a una mayor probabilidad de recidividad, al igual que en estudios previos [7]. La M$Fdom$ también ha presentado rangos intercuartílicos ligeramente inferiores a los de la $Fdom$, a la vez que más balanceados. Esto quizás pueda deberse a que la M$Fdom$ se estimó sobre la envolvente espectral, la cual de acuerdo a [13] puede interpretarse como la envolvente que contiene a las $\bar{P}$ de las 12 derivaciones del ECG. Esto también podría justificar que los valores de Se y Sp aparezcan más balanceados.

En cuanto a las medidas de entropía, la MEm incrementó la correcta clasificación hasta un 80 %, lo que supone un incremento de casi el 6 % respecto a la Em, y al igual que en los casos anteriores, la Se y Sp en el umbral óptimo presentaron valores más parejos. De forma similar a algunos trabajos previos [7], se han observado valores mayores de Em para los pacientes que recaen a FA tras la CVE, sugiriendo que una mayor organización de las ondas f puede estar asociada con un menor remodelado del sustrato arritmogénico, y en consecuencia, con una mayor probabilidad de mantener un RS. Cabe destacar que la MEm presentó valores en torno a un orden de magnitud inferior a los obtenidos por la Em. En este caso, la MEm [15] establece que los subvectores pueden evolucionar en 12 direcciones distintas, así pues, al haber un mayor número de ellos disponibles, sería lógico pensar que el número de similitudes también se podría ver en cierta medida incrementado, y según [14], a mayor número de secuencias de longitud m similares, menor será la entropía resultante.

Finalmente, de acuerdo a los resultados obtenidos y a su discusión, puede concluirse que la extensión multiderivacional de los índices lineales anchamente usados para caracterizar las ondas *f*, es decir, de la A_f y la $Fdom$, no han conseguido mejorar la predicción del resultado de la CVE de forma notable. Al contrario, la MEm ha alcanzado una mejora predictiva del 6 % con respecto a su versión unidimensional, lo que sugiere que las correlaciones espaciales presentes en las distintas derivaciones del ECG ofrecen nueva información que puede ayudar a entender los mecanismos responsables de mantener la FA tras la CVE. No obstante, se debe tener en cuenta que el número de pacientes analizado en este estudio es limitado y se requieren más análisis en el futuro para confirmar este resultado.

Agradecimientos

Este trabajo se ha financiado con los proyectos de investigación competitiva DPI2017-83952-C3 de MINECO-AEI-FEDER-UE, SBPLY/17/180501/000411 de la JCCLM y AICO/2019/036 de la GVA.

Referencias

[1] Lane DA, et al. Temporal trends in incidence, prevalence, and mortality of atrial fibrillation in primary care. J Am Heart Assoc 2017;6(5).

[2] Kirchhof P, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Europace 2016;18(11):1609–1678.

[3] January CT, et al. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. J Am Coll Cardiol 2014;64(21):e1–76.

[4] Fried AM, et al. Electrical cardioversion for atrial fibrillation in the emergency department: A large single-center experience. Am J Emerg Med 2020;.

[5] Brandes A, et al. Cardioversion of atrial fibrillation and atrial flutter revisited: current evidence and practical guidance for a common procedure. Europace 2020;.

[6] Fujimoto Y, et al. Advanced interatrial block is an electrocardiographic marker for recurrence of atrial fibrillation after electrical cardioversion. Int J Cardiol 2018;113–117.

[7] Alcaraz R, et al. Noninvasive time and frequency predictor of long-standing atrial fibrillation early recurrence after electrical cardioversion: predictor of cardioversion outcome. Pacing Clin Electrophysiol 2011;34(10):1241–1250.

[8] Meo M, et al. Spatial variability of the 12-lead surface ECG as a tool for noninvasive prediction of catheter ablation outcome in persistent atrial fibrillation. IEEE Trans Biomed Eng 2013;60(1):20–7.

[9] Bollmann A, et al. Analysis of surface electrocardiograms in atrial fibrillation: techniques, research, and clinical applications. Europace 2006;8(11):911–26.

[10] Alcaraz R, et al. Adaptive singular value cancelation of ventricular activity in single-lead atrial fibrillation electrocardiograms. Physiol Meas 2008;29(12):1351–69.

[11] Vest AN, et al. An open source benchmarked toolbox for cardiovascular waveform and interval analysis. Physiol Meas 2018;39(10):105004.

[12] Corino V, et al. Signal processing methods for information enhancement in atrial fibrillation: Spectral analysis and non-linear parameters. Biomed Signal Process Control 2006;1(4):271–281.

[13] McDougall A, et al. Optimal transformations and the spectral envelope for real-valued time series. J Stat Plan Inference 1997;57(2):195–214.

[14] Richman JS, et al. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol 2000;278(6):H2039–49.

[15] Ahmed M, et al. Multivariate multiscale entropy: A tool for complexity analysis of multichannel data. Phys Rev E 2011;84(6):1069–1084.

[16] Nault I, et al. Clinical value of fibrillatory wave amplitude on surface ECG in patients with persistent atrial fibrillation. J Interv Card Electrophysiol 2009;26(1):11–9.

Bioinstrumentación

Miércoles, 25 de noviembre de 2020

Artificial synovial fluid and artificial cerebrospinal fluid development and tests using QCR based biosensor for medical diagnosis

A. Miranda Martínez[1], M. X. Rivera González[1], M. Zeinoun[1], L. A. Carvajal Ahumada[2], J. J. Serrano Olmedo[1,3]

[1] Centre for Biomedical Technology (CTB), Universidad Politécnica de Madrid (UPM), Madrid, España, {andres.miranda, marco.rivera, michael.zeinoun}@ctb.upm.es

[2] Faculty of Engineering and Basic Sciences, Universidad Central, Bogotá, Colombia, luis.carvajal@ctb.upm.es

[3] Networking Research Centre of Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Universidad Politécnica de Madrid (UPM), Madrid, España, josejavier.serrano@ctb.upm.es

Abstract

Quartz Crystal Resonators are devices with high capabilities in sensing systems due to their accuracy, efficiency and low cost. In this work we present a prototype of biosensor based on the use of QCR with the objective of measure the viscosity of small samples of biological fluids for medical diagnosis. In order to observe the response of the system in a biomedical application, two artificial fluids were developed for simulate healthy and pathological fluids. These measurements were compared with a professional viscometer to observe the similarity between both measurement methods. These tests may be used to observe the viability of the device and to calibrate the system before testing with real fluids.

1. Introduction and objective

One physical property present in fluids is the viscosity, which can be defined as the movement resistance of a fluid [1]. This property has been studied in biological fluids with the aim of relating changes on it to certain pathologies, two examples of this are: Synovial Fluid (SF) [2, 3] and Cerebrospinal Fluid (CSF) [4]. In the SF case the decrease of viscosity is a sign for Rheumatic Arthritis and Osteoarthritis, while in CSF case, an increase of viscosity points to meningitis, both bacterial and viral.

The most common test to detect the physical condition of SF is obtaining the sample through puncture in the joint, extending the sample with a needle on a surface and measuring the length before the thread is cut [5, 6], this also can be done using the thumb and index fingers, if the length is greater than 3 cm, the sample is healthy (high viscosity).

For the analysis of CSF there are not physical tests like the case of SF, the most commonly tests to diagnose meningitis are Gram stain and culture [7,8].

On the other hand, quartz crystal resonators (QCR) are piezoelectric elements that vibrate at a specific frequency with high precision when excited with a sinusoidal signal. at that resonance frequency their impedance will reach a minimum value. These properties allows QCR to have applications in the field of biomedicine [9], one of its main uses is as a microbalance (Quartz Crystal Microbalance) allowing to detect very small masses such as cells or tiny objects of this nature. QCR can also be used to determine the viscosity of biological fluids, this offers many advantages as they require a smaller amount of sample which adds to its simplicity, accuracy and efficiency.

For this reason a biosensor prototype is proposed which is based on QCR and is capable of measuring viscosity using a small amount of sample (50 µl approximately) by measuring the displacement of the resonance frequency which occurs when the sample is placed in contact with the crystal.

With this prototype is possible to measure the frequency shift and relate it with the viscosity of an specific biological fluid like SF and CSF and then do the medical diagnosis.

To achieve this objective, is necessary to develop artificial synovial fluid (aSF) and an artificial cerebrospinal fluid (aCSF) and test them before make experiments with real fluids. In this paper the two artificial fluids are proposed and tested using the prototype developed in the Laboratorio de Bioinstrumentación y Nanomedicina of the Centro de Tecnología Biomédica of Universidad Politécnica de Madrid.

2. Materials

2.1. Sensor system

This prototype is designed to measure the frequency when the peak of conductance is, this is achieved by producing a frequency sweep near the resonance frequency of the crystal and measuring the voltage, current and susceptance on each frequency step. The device consists of a series of electronic components, one of which is a module which generates the sinusoidal signal that excites the QCR. Also operational amplifiers and digitally controlled potentiometers are used with the aim of provide an amplification stage at the entrance of the crystal and on the sensing ports.

The control and data acquisition stage is achieved using an Arduino DUE board and LabView software. Two frequencies are established close to the resonance frequency of the crystal to perform a sweep where, for each frequency, the values of voltage, current and susceptance will be obtained, those measures are used to obtain the

ISBN: 978-84-09-25491-0, pags. 19- 22

conductance curve, thus, the resonance frequency at a maximum of conductance (fs), necessary to calculate viscosity. In [10] more details of the prototype are described. A blocks diagram of the system is shown in figure 1.

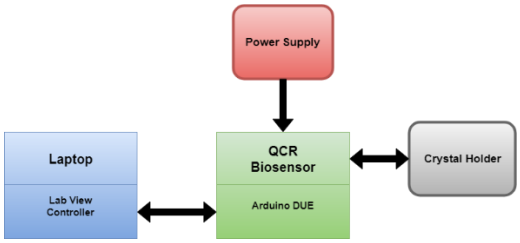


Figure 1. *System blocks diagram*

2.2. Crystals

The QCR were purchased form Krystaly (Hradec Králové, Czech Republic; http://www.krystaly.cz/). Crystals (figure 2) have a fundamental resonance frequency at 10 MHz, gold electrodes on both sides with 5 mm of diameter and the whole crystal is mounted on HC-51 holder.



Figure 2*: Quartz Crystal*

One Crystal Holder was printed on 3D with aim on have the QCR on a static and stable position, and to have the entrance where the sample will be pipetted.

The whole system can be observed on figure 3.

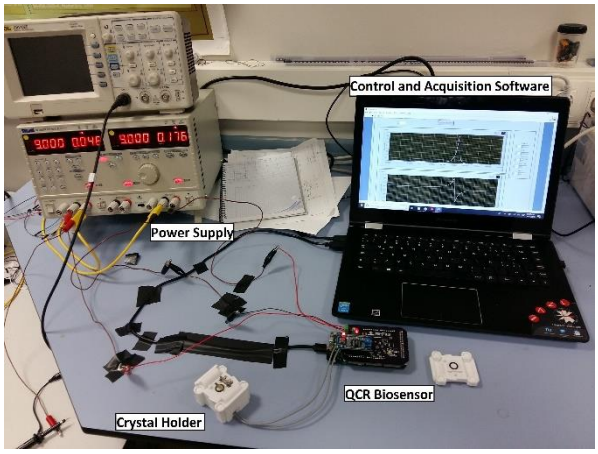


Figure 3: *QCR based biosensor system*

2.3. Artificial fluids

Synovial fluid is a mixture of plasma and hyaluronic acid (HA), this one brings the viscoelastic properties to the fluid, thus, viscosity of the fluid depends on hyaluronic acid concentration [11]. Three dilutions where made changing hyaluronic acid concentration, one for healthy fluid (3.5 mg/ml), one for osteoarthritis (1.3 mg/ml) and the last one for rheumatoid arthritis (0.84 mg/ml) [12]. In table 1 the concentrations for the aSF are shown. The mixtures were made with milli-q water and weighing the necessary amount of HA with an analytical balance. After add HA into water, gently stirred was necessary and finally store the sample in the refrigerator for a few hours in order to eliminate bubbles that form when mixing.

HA (Mw = 1.5 MDa) was supplied by Acros Organics.

Fluid	**H. A. concentration (mg/ml)**
Healthy aSF	3.5
aSF with osteoarthritis	1.3
aSF with rheumatoid arthritis	0.84

Table 1. *Artificial Sinovial Fluids compositions*

CSF is a complex fluid mainly composed of *Na*, *Cl* and HCO_3 with lesser amounts of *K*, *Mg*, *Ca* [13, 14]. Some artificial cerebrospinal fluid (aCSF) are reported in [15, 16], where *NaCl* and $NaHCO_3$ are the elements with more quantity. According this, healthy artificial CSF was developed with the concentrations listed in table 2. In order to increase viscosity in aCSF and simulate abnormal fluid, albumin was used since is the protein with higher concentration in pathological CSF [17, 18]. For the dilutions, an analytical balance was used to weigh the elements.

Fluid	**NaCl concentration (mg/ml)**	**NaHCO3 concentration (mg/ml)**	**Albumin concentration (mg/ml)**
Healthy aCSF	7.25	2.18	0
Viral aCSF	7.25	2.18	0.5
Bacterial aCSF	7.25	2.18	3

Table 2. *Artificial cerebrospinal fluids compositions*

2.4. Reference Viscometer

As a reference, rotational viscometer Alpha series by Fungilab was used in order to have viscosity measurements obtained by a professional equipment. LCP adapter was used, this allows low measures of viscosity up to 2 Pa*s, the quantity needed to work is about 16 ml.

3. Methods

To start the measurements with QCR biosensor first the crystal is mounted on the holder and connected to the device. Measures with the crystal without sample are done to obtain a first frequency (f_0) value, frequency at maximum conductance. Then, the sample is pipetted on the crystal and tested in order to obtain a second frequency value (f_s) which will be shifted in frequency. With this two values ΔF (frequency displacement due to a sample on the crystal) is calculated, since $\Delta F = f_s - f_0$. Both, f_s and ΔF are the parameters that are used when viscosity is calculated. This is observed in equation 1 exposed by Kanazawa et al. [19].

$$\Delta f = -\sqrt{n} f_0^{3/2} \sqrt{\frac{\rho_L \eta_L}{\pi \rho_q G_q}} \quad (1)$$

Where $\rho_q = 2.648\ g\ cm^{-3}$ and Gq = 2.947 x $10^{10}\ Nm^{-2}$ are the specific density and the shear modulus of quartz respectively, f0 is the fundamental resonance frequency of the quartz, ρ_L is the density of the fluid, η_L is the viscosity of the fluid and n is the overtone number (fundamental frequency is employed, thus n =1). The density of each sample was measured using an analytical balance.

At the end of the process, the crystal is rinsed with abundant deionized water and alcohol (isopropanol) in order to clean the surface of the crystal. Finally the electrode surface is dried with air (compressed air) or using an air blower (lens cleaning). It is important to avoid mechanical contact with the electrode.

The process is repeated at least five times with each sample.

Results and discussion

Following the methodology above, artificial synovial fluid tests were done, in figure 4, Δf obtained is showed. Using the value of Δf, viscosity is calculated using equation 1 in order to obtain the viscosity measure using the sensor, in table 3 these parameters are presented. Is observed that as the concentration value increases, the Δf also increases

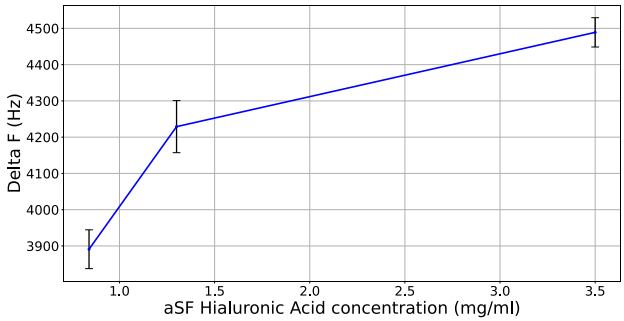


Figure 4. *Frequency shift in aSF*

Fluid	Δf (Hz) Measured	Density (mg/ml)	Viscosity (mPa*s)	Viscosity calculated with Δf (mPa*s)
Healthy aSF	4489 ±40	980	1184	5.04
aSF with oteoarthritis	4228 ±72	994	252	4.4
aSF with rheumatoid arthritis	3891 ±53	992	139	3.7

Table 3. *aSF parameters*

The results of measurements using aCSF shows the capacity of the sensor to discriminate the three types of artificial fluids which are having a slight increase in the albumin concentration. In figure 5, the frequency shift is observed for each case of aCSF, and in table 4 the parameters of the fluids with the frequency shift are exposed.

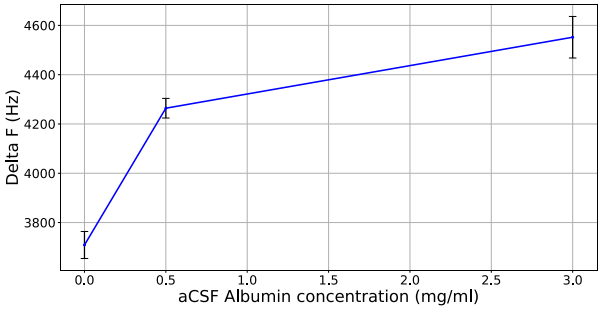


Figure 5. *Frequency shift in aCSF*

Fluid	Δf (Hz) Measured	Density (mg/ml)	Viscosity (mPa*s)	Viscosity calculated with Δf (mPa*s)
aCSF	3709 ±55	1008	1.02	3.3
aCSF viral	4264 ±40	1010	1.03	4.4
aCSF bacterial	4552 ±84	1015	1.04	5

Table 4. *aCSF parameters*

For both artificial fluids, comparing the viscosity obtained with the viscometer and the viscosity calculated from equation 1 using Δf, is observed that using the method proposed, a difference between viscosities is present. One explanation of this difference is mentioned in [20, 21] where is stated that the relationship explained in equation 1 have been derived theoretically for an infinite quartz plate without interfacial slip and for the case of a perfect smooth surface, consequently the use of this equation has not proved to be reliable when measuring density-viscosity product.

On the other hand, HA concentrations are no newtonian fluids with a pseudoplastic behavior, this means that at higher shear rate the viscosity will decrease. In this way, the rotational viscometer measures at 0.3 rpm, this allows to measure the higher viscosity since fluid is handled with low share rate, and the crystals are vibrating near 10 MHz which can be considered a high shear rate, thus, the viscosity will result to be minor. Notwithstanding, the sensor is capable of discriminate between the three fluids.

For these reasons, it is necessary to calibrate the system in order to avoid these viscosity difference.

Finally, the capability of the prototype to measure the variation of viscosity was demonstrated, since the results shows how the frequency shift grows while the concentration in both artificial fluids increase. Furthermore, is observed that the variation in each measure does not interfere between adjacent samples, this allows to discriminate between each sample used. With the results obtained it will be possible to adjust the system in order to

have quantitative results for future tests having as a reference a rotational viscometer.

Acknowledgements

This work was performed with the grant CONACYT (Consejo Nacional de Ciencia y Teconlogía) from México, CVU Nº 769250.

References

[1] LeBlanc GE, Secco RA, Kostic M. Viscosity measurement. *The Measurement, Instrumentation, and Sensors*. CRC Press, 1999.

[2] Brannan SR, Jerrar, DA. Synovial fluid analysis. J. Emerg. Med. 2006, 30, 331–339.

[3] Stafford C, Niedermeier W, Holley H, Pigman W. Studies on the concentration and intrinsic viscosity of hyaluronic acid in synovial fluids of patients with rheumatic diseases. Ann. Rheum. Dis. 1964, 23, 152–157.

[4] Yetkin F, Kayabas U, Ersoy Y. Cerebrospinal fluid viscosity: a novel diagnostic measure for acute meningitis. Southern Medical Journal, 01 Sep 2010, 103(9):892-895

[5] Martínez-Castillo A, Núñez C, Cabiedes J. Synovial fluid analysis. Reumatol. Clín. 2010, 6, 316–321.

[6] Rojas C. Estudio del Liquido Sinovial; Guías de Procedimientos en Reumatología: Bogotá, Colombia, 2011; pp. 41–47.

[7] Wright BLC, Lai JTF, Sinclair AJ. Cerebrospinal fluid and lumbar puncture: A practical review. Journal of Neurology. Volume 259, Issue 8, August 2012, Pages 1530-1545.

[8] Abro AH, Abdou AS, Ali H, Ustadi AM, Hasab AAH. Cerebrospinal fluid analysis acute bacterial versus viral meningitis. Pakistan Journal of Medical Sciences, Volume 24, Issue 5, October 2008, Pages 645-650.

[9] Nwankwo E, Durning C. Mechanical response of thickness-shear mode quartz-crystal resonators to linear viscoelastic fluids. Sensors Actuators A Phys. 1998, 64, 119-124.Doarn CR, Adilova F, Lam D. Review of telemedicine in Uzbekistan. *Journal of Telemedicine and Telecare*, vol 11, sup 3, 2005, pp 135-9 (ISSN: 1357-633X).

[10] Carvajal LA, Peña N, Herrera OL, del Pozo F, Serrano JJ. A new way to find dielectric properties of liquid sample using the quartz crystal resonator (QCR). Sens. Actuators A Phys. 2016, 239, 153–160.

[11] Barnett CH. Measurement and Interpretation of Synovial Fluid Viscosities. Ann. Rheum. Dis. 1958, 229–234

[12] Stafford C, Niedermeier W, Holley H, Pigman W. Studies on the concentration and intrinsic viscosity of hyaluronic acid in synovial fluids of patients with rheumatic diseases. Ann. Rheum. Dis. 1964, 23, 152–157.

[13] Spector R, Snodgrass SR, Johanson CE. "A balanced view of the cerebrospinal fluid composition and functions: Focus on adult humans," Exp. Neurol., vol. 273, pp. 57–68, 2015, doi: 10.1016/j.expneurol.2015.07.027.

[14] Sakka L, Coll G, Chazal J. "Anatomy and physiology of cerebrospinal fluid," Eur. Ann. Otorhinolaryngol. Head Neck Dis., vol. 128, no. 6, pp. 309–316, 2011, doi: 10.1016/j.anorl.2011.03.002.

[15] Saini M, Sadhu KK. "Two instantaneous fluorogenic steps for detection of nanomolar amyloid beta monomer and its interaction with stoichiometric copper(II) ion," Sensors Actuators, B Chem., vol. 303, no. October 2019, p. 8, 2020, doi: 10.1016/j.snb.2019.127086.

[16] Rudell JB, Rechs AJ, Kelman TJ, Ross-Inta CM, Hao S, Gietzen DW. "The anterior piriform cortex is sufficient for detecting depletion of an indispensable amino acid, showing independent cortical sensory function," J. Neurosci., vol. 31, no. 5, pp. 1583–1590, 2011, doi: 10.1523/JNEUROSCI.4934-10.2011.

[17] Reguera RM. Interpretation of cerebrospinal fluid. Anales de Pediatria Continuada. Volume 12, Issue 1, January-February 2014, Pages 30-33

[18] Codina MG, De Cueto M, Vicente D, Echevarría JE, Prats G. Microbiological diagnosis of central nervous system infections. Enfermedades Infecciosas y Microbiologia Clinica. Volume 29, Issue 2, February 2011, Pages 127-134

[19] Kanazawa K, Gordon JG. The oscillation frequency of a quartz resonator in contact with liquid. Anal. Chim. Acta 1985, 175, 99–105.

[20] Johannsmann, D. Viscoelastic, mechanical, and dielectric measurements on complex samples with the quartz crystal microbalance, Physical Chemistry Chemical Physics 10 (2008) 4516–4534.

[21] Cassiède M, Daridon JL, Paillol JH, Pauly J. Characterization of the behaviour of a quartz crystal resonator fully immersed in a Newtonian liquid by impedance análisis. Sens. Actuators A Phys. 2011, 167, 317-326.

In vitro human skeletal muscle platform for the evaluation of functional outcomes

Camila Vesga-Castro[1, 2], Ainara Vallejo-Illarramendi[2, 3] and Jacobo Paredes[1]

[1]Tissue Engineering group, Department of Biomedical Engineering and Science, TECNUN, Universidad de Navarra, Manuel Lardizabal 13, 20018 Donostia / San Sebastián, Spain.
[2]Group of Neurosciences, Department of Paediatrics, Hospital Donostia - IIS Biodonostia, UPV/EHU, 20014 San Sebastian, Spain
[3]CIBERNED, Instituto de Salud Carlos III, Ministry of Science, Innovation and Universities, 28031 Madrid, Spain
Corresponding author: e-mail jparedes@tecnun.es; ainaravallejo@yahoo.es

Abstract

Skeletal muscle is a striated tissue whose main function is to produce the force needed to support locomotion, posture, respiration, and other metabolic processes. Nowadays, most in vitro myotube cultures do not fully recapitulate muscle characteristics, and therefore, these studies are not physiologically relevant to human muscle. This work presents a new experimental approach to achieve a high-throughput platform of 2D mature functional myotubes. First, cell cultures were grown onto PDMS-grooves (µGrooves) on a glass substrate. Different sizes and shapes were tested to confine the myoblasts and promote their alignment within the micro-well. Electrical stimulus (4V, 6ms, 0.1Hz) was applied for 24 hours to increase myotube differentiation and maturation. Myotubes were tested for orientation, maturation, fusion index, cell density, calcium handling, and, contractile function. Our results demonstrate the capability of this platform to obtain highly homogeneous cultures composed of individual myotubes. These cultures replicate native characteristics of the skeletal muscle regarding excitation-contraction capacity and appropriate responsiveness to different stimuli. Furthermore, our protocol minimizes the variability of maturation and cytosolic calcium levels within cultures, which are significant limitations of traditional myotube cultures. In this regard, we found that µGrooves reduced the coefficient of variation of intracellular calcium levels. Therefore, this platform enables robust high content assays of calcium levels and contractile readouts with increased sensitivity compared to standard cultures. Finally, in addition to mechanistic studies, this platform could be used in combination with myotubes from patients with different muscle pathologies for high throughput drug screening assays.

1. Introduction

Skeletal muscle represents approximately 40% of human body mass and participates in several metabolic processes. The main function of this tissue is the production of force, which is the result of the sum of several intracellular pathways that occur within the muscle fiber. This process may be affected due to aging, major injuries, volumetric muscle loss, or muscular dystrophies [1]. One of the main characteristics of healthy skeletal muscle is its capacity of self-repair from small tears (i.e. daily activities and exercise). This is due to the presence of muscle stem cells, in an adjacent niche to the muscle fibers [2],[3]. However, muscle self-repair capacity is impaired in some conditions [4], for example, aging or muscular dystrophies such as Duchenne muscular dystrophy [5] and calpainopathy [6].

Calcium (Ca^{2+}) dysregulation is a feature observed in many forms of muscular dystrophies, since muscle contraction depends directly on the release of Ca^{2+} to the sarcoplasmic reticulum. Thus, when this process is affected, force production is inadequate to sustain the human body's function and eventually impacts normal behavior.

For years, cell cultures and animal models have been used to understand muscle environment, development, and function. Despite the numerous discoveries based on these models, there is still a wide gap between cellular models and human muscles [2],[7]. The lack of proper models to carry out screening of potential drugs for muscle diseases has unveiled the need to find alternatives to the well-known animal models and standard cell cultures. There is a wide range of applications where these new muscle platforms may fit, from drug discovery [9] to gene therapy [10] and muscle implantation or reconstructive surgery [11].

The latest studies in muscular tissue development are focusing their efforts on mimicking the natural microenvironment to direct myogenesis. More important, finding proper models that simulate the natural microenvironment of skeletal muscle (e.g. scaffolds, substrates, stimulation, vascularization) and give proper support throughout all the myogenesis, is becoming a subject of great interest in the last decades [8].

An *in vitro* muscular platform requires a setup with compatibility to perform standard cell analysis, for instance, protein quantification, cell density, western blot, immunocytochemistry, etc. Additionally, it must be able to allow the performance of functional assays in homogeneous cultures, such as contractile function (Twitch, Tetanus, Frequency respond) and calcium handling (Ca^{2+}). More important, it is necessary to be able to perform all these assays in a high-throughput manner to evaluate the efficacy and variability of the experiment.

In this work, we present a 2D platform that has demonstrate their suitability to perform human skeletal muscle myotubes cultures and the evaluation of their functional outcomes. The aim of this new 2D-platform focused on the creation of proper cues throughout all the myogenesis, so it would promote better development of the skeletal muscle myotubes, allowing the improvement of the functional outcomes.

2. Materials and Methods

2.1. µWell Substrate Fabrication

The culture platform consists of a glass slip substrate covered with a PDMS membrane containing the µGrooves features. This membrane was fabricated following an

ISBN: 978-84-09-25491-0, pags. 23- 26

adapted protocol [12]. Briefly, 80µm-height SU-8 (2025 MicroChem) masters were fabricated through photolithography and then silanized overnight in a vacuum desiccator with Trichloro (1H, 1H, 2H, 2H-perfluorooctyl) silane (Sigma). At the same time, a 3mm-thick PDMS block was prepared (Poly (dimethylsiloxane) sylgard 184 silicone, Dow corning) with a ratio of 15:1 and was also silanized as previously described. Then uncured PDMS (10:1 ratio) was poured onto the silanized PDMS and immediately put in contact with the SU-8 master. Weights of 600g were placed atop the sandwich to ensure homogenous contact between the layers. Polymerization was achieved overnight at 60ºC. Besides, the weight on top of the PDMS and the master need to be homogeneous over the entire surface. Delamination was then carefully carried out to remove the silicon master from the PDMS.

Glass coverslips were cleaned 70% ethanol. Oxygen plasma was used to bind the PDMS membrane containing the µGrooves features to the coverslip. The resulting pieces were put on a hot plate at 60ºC overnight to enhance the bonding. The next day, the layers of PDMS were carefully separated leaving the coverslip with the intact features.

µGrooves substrates were sterilized by a series of washing steps with 70% ethanol, dry with compressed air, and placed under UV inside the culturing hood for 20 minutes. Before seeding, µGrooves were 48h in incubation at 37ºC with growth media.

2.2. Cell Cultures.

Immortalized human myoblast (8220) were used to generate human skeletal myotubes. Cells were grown in proliferation media (Skeletal growth media SGM, Promo cell, with supplemented mix, inactivated 10% fetal bovine serum (FBS), 50µg/ml gentamicin and 1% glutamax (Gibco-Invitrogen)). Cells were kept in cell culture T75 Flask until reaching 70-80% confluence.

In order to prepare µGrooves substrates for cell culture, µGrooves coverslips were placed into 24-well plates and coated with 25µg/ml fibronectin (FN, sigma) in Sodium bicarbonate solution (pH 8.5) for one hour, at room temperature was performed. An initial density of 500.000 cells were seeded per µGrooves substrate.

To induce differentiation, SGM was changed to basic differentiation media (bDM) when cultures reach confluence (bDM: DMEM, 100µg/ml apotransferine, 10µg/ml insulin, and 50ng/ml gentamicin). When fusion of myoblast was observed, and multinucleated myotubes emerged, bDM was changed to complete differentiation media. (cDM): (cDM: Neurobasal medium, B27 1X, glutamax 1X, 20ng/ml brain-derived neurotrophic factor (BDNF), 50ng/ml sonic hedgehog (shh), 10ng/ml insulin-like growth factor 1 (IGF-1), 5ng/ml ciliary neurotrophic factor (CNTF), 20ng/ml neurotrophin 3 (NT-3), 4mg/ml laminin, 100ng/ml agrin and 50mg/ml gentamicin). Medium was refreshed every 3 days.

2.3. Alignment verification

Since cell alignment is one of the main characteristics of the skeletal muscle, this feature was analyzed with the ImageJ plugin OrientationJ. We selected the entire geometry as a region of interest and the angle was determined for the different cultures. Alignment was compared between cells grown in µGrooves and controls. Cells were considered aligned when the angles were ≥ 10º [13].

2.4. Calcium Transients Analysis.

To determinate Calcium handling under several conditions, the myotubes that were grown and differentiated onto µGrooves substrates were loaded with 2mM Fura 2-AM and 20% Pluronic acid (Thermo Fisher) in cDM for 30 minutes at 37ºC. Then, cells were left in IB buffer 1X for 30 more minutes at room temperature (RT) to remove nonhydrolyzed fluorophore and complete de-esterification of the dye. Experiments were performed using an ECLIPSE Ti-S/L100 microscope (Nikon) equipped with a 10x S-Fluor objective and attached to a lambda-DG4 illumination system. Images were acquire using an Orca-Flash2.8 camera (Hammamatsu) with the NisElements-AR software. Ca^{2+} transients were elicited first by stimulating the myotubes with 20V, 6ms pulse, and 0.1Hz. Then, the response to an increase of voltage was performed by changing from 4V to 40V, with 20-s intervals. Finally, tetanic stimulation was induced with 40V, varying the frequency between 10 and 66Hz, with 20-s intervals. Intracellular calcium concentration was estimated by the ratio of Fura 2-AM fluorescence intensities at excitation wavelength 340 and 380nm (ΔF/F).

2.5. Contractility

To assess the main function of skeletal muscle, after 3 days post differentiation, myotubes were stimulated for 24h to improve their contractile capacity at 4V, 6ms, 0.2Hz using the C-PACE EP unit (IonOptix, Westwood, MA 02090 USA). First, twitch contraction was induced at 20V, 6ms pulse, and 0.1Hz. Then, as previously mentioned for Calcium transients, increasing voltages and tetanic stimulation were performed. Between each experiment, the myotubes were left 3 minutes to rest. Myotubes contraction was recorded in bright field for further analysis by using the ImageJ Macro MUSCLEMOTION [14]. (Figure 1).

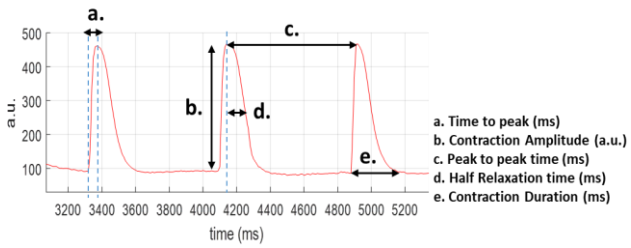


Figure 1*. Contractility Assay. Five variables were determinate as a response of electrical stimulation*

2.6. Immunostaining

Cultures were washed with DPBS and fixed for 10 minutes at room temperature with paraformaldehyde 4% (PFA 4%). Samples were pretreated with blocking buffer solution, (2% BSA, 2% normal donkey serum, 0.5% Triton-X100, 0.02% NaN_3) in PBS 1X for one hour, and incubated with α-actinin antibody (A7811, 1:500 dilution)

and with α-laminin antibody (1:60) overnight at 4ºC. Then samples were washed with PBS, followed by one hour-incubation with secondary antibodies (1:400), followed by incubation with Myosin Heavy Chain-FITC (1:50), and DAPI (1:1000) in the dark at room temperature for one hour. Cells images were taken with ECLIPSE Ti-S/L100 microscope (Nikon) and analyzed using ImageJ software.

2.7. Statistics

Data is presented as mean ± SEM. P values were calculated using Student's *t*-test. P≤0.05 was considered statistically significant. Alignment and masters characterization data was analyzed using Ordinary one-way ANOVA.

3. Results and Discussion

3.1. µGrooves fabrication and characterization

Figure 2 presents a schematic representation of the fabrication of the substrates previously described. In Table 1 the characterization of the µGrooves are described. Before any use, SU8 masters were tested to verify their height of approximately 80µm.

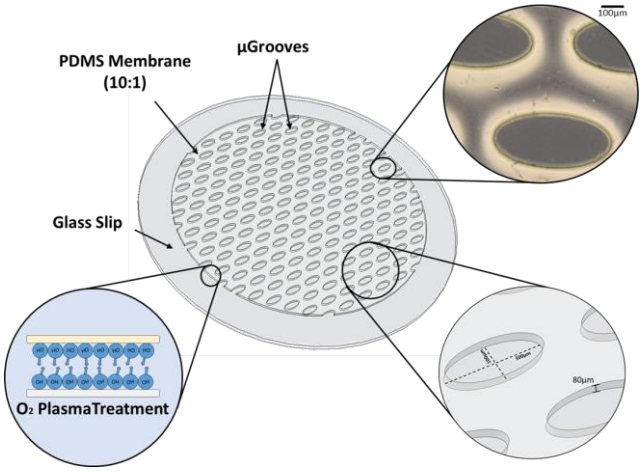


Figure 2. a. *µGrooves Substrates Characteristics. Cells are confined within these microwells. µGrooves masters fabrication was designed to obtain features of 80µm-height. This condition was confirmed by stylus profilometer.*

µGroove	µGroove height	Nº µGroove/glass slip	µGroove Area (mm²)	Cells/µGroove
500/100	75.642 ± 7.047µm	263	0,039	148
500/200	75.684 ± 4.554µm	211	0,078	295
700/100	78.078 ± 5.523µm	257	0,054	205
700/200	77.12 ± 6.147µm	161	0,109	413

Table 1. µGrooves Substrates types characterization and number of myoblast per type

3.2. Myotubes growth within µGrooves

Cellular concentration was also under optimization, from 50.000 up to 600.000 cells per µGrooves. Increasing the number of cells per well, reduces the time of differentiation from 48-72 hours to 24 hours, decreasing the time of culture. In addition, this eludes the problem of cell lifting from the substrate. Furthermore, the number of individual myotubes formed was improved by high concentration cultures rather than in low (50.000 cells/µGrooves= 3 myotubes vs 600.000 cells/µGrooves = 10 myotubes).

Fiber alignment was confirmed by the determination of cellular orientation within the microwells (Figure 3). Each day, images of the culture were taken to observe their growth under the platform cues. As observed in the images, cells appeared aligned with the axis of the features and reached confluence. Regardless of the geometries size, all four structures allow cell aligment within 24 hours of seeding. This result is a very important variable, as the degree of alignment of the myoblast up to the formation of the myotube will influence its contraction outcome. Moreover, unlike controls with no µGrooves, having independent myotubes will allow a better understanding of tissue behavior.

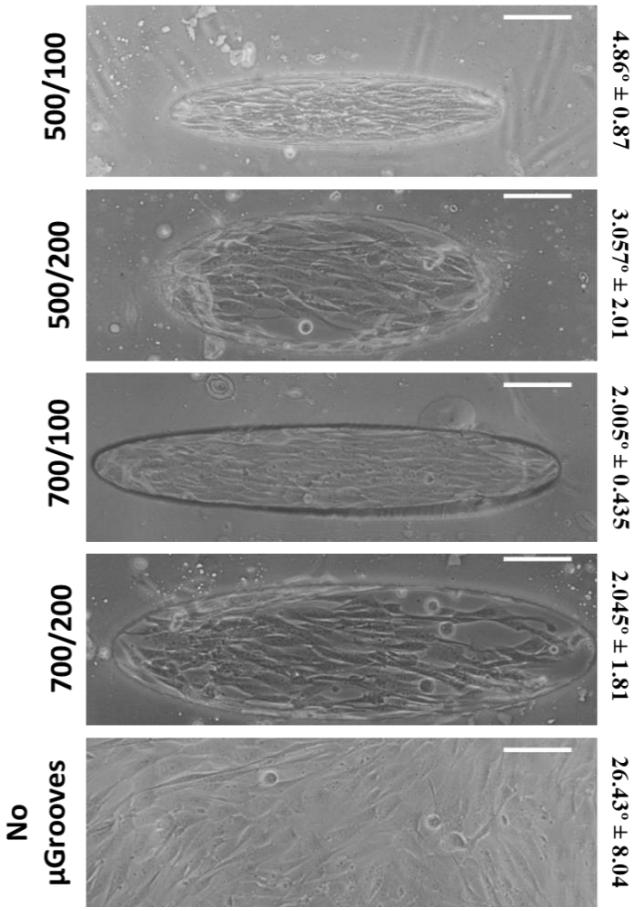


Figure 3. *Determination of cell Alignment. Brightfield images were taken 24 hours after seeding. Myoblast were aligned in the geometry axis within its area. Scale bar=100µm*

3.3. Contractile Function

One of the main advantages of this µGrooves platform is that by having individual myotubes, contraction was easily detected, as well as the calculation of the myotubes width was easier. Compare to standard cultures, in our system we can assure that the contraction value only depends on a specific myotube, and not on the effect of several myotubes, that could be related or be independent of the myotube of interest (Figure 4). These cultures respond to electrical stimulation, but also present spontaneous

contraction, which is a positive signal of the degree of myotube maturation, which was confirmed also by the presence of striation. From the curves obtained we determine, Peak amplitude, and time to peak: 366.327 ± 5.589; 430.667 ± 0.943, respectively.

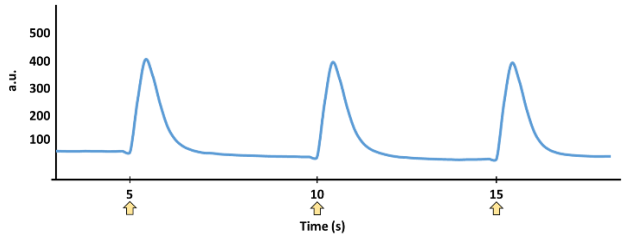


Figure 4. Contractile respond for the feature 500/200. Electrical stimulation of 4V, 6ms pulse, 0.20Hz. Yellow arrows indicate the moment of the stimulation.

3.4. Calcium handling

Calcium handling was also evaluated, with Fura 2-AM. This is a critical variable for normal skeletal muscle function. First, basal cytosolic calcium was determined comparing controls vs. µGrooves substrates (Figure 5). The coefficient of variation of intracellular calcium levels varied from 13.83% to 8.904%. This first approximation to the Ca^{2+}, demonstrate the suitability of the platform to perform high content assays of calcium levels with increased sensitivity rather than in standard cultures. Change of voltage and frequencies as well as induced twitch and tetanus stimulation was also performed. Increasing fluorescence (ΔF) was observed as the amplitude of voltage went from 4V to 40V, however, a clear and synchronous (0.1Hz) respond was detected for voltages of 20V and 40V.

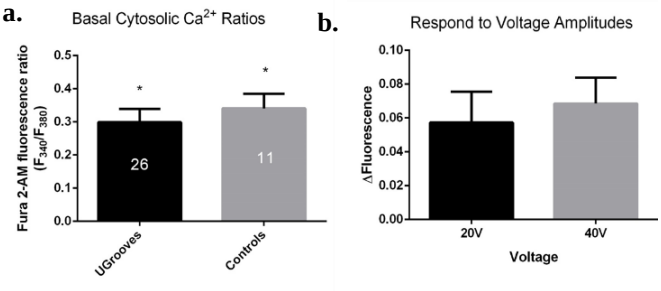


Figure 5. *Calcium handling. a. Basal Cytosolic Calcium. Ratios were determined without any electrical or chemical stimuli for µGrooves substrates and controls. (*P<0.05, n= total numbers of myotubes recorded are shown in the graph). b. Calcium transients in response to electrical twitches of different voltage amplitudes. (n= 5 myotubes)*

4. Conclusions

Here, we present an approach for a high-throughput platform of 2D human skeletal muscle. The high degree of fiber alignment within 24 hours of seeding, as well as the positive response to electrical stimuli by the determination of the contractile properties and calcium transients (Ca^{2+}), demonstrate the suitability of this platform to induce the formation of individual and homogeneous human myotubes. Calcium handling as well as alignment was clearly improved by the growth of cells in the µGrooves, which are results that get closer to the native characteristics of the muscle. Furthermore, with this work, dystrophic cells could also be used to study their mechanism and respond to different stimuli.

References

[1] F. Yusuf and B. Brand-Saberi, "Myogenesis and muscle regeneration," *Histochem. Cell Biol.*, vol. 138, no. 2, pp. 187–199, 2012.

[2] J. Wang, A. Khodabukus, L. Rao, K. Vandusen, N. Abutaleb, and N. Bursac, "Engineered skeletal muscles for disease modeling and drug discovery," *Biomaterials*, 2019.

[3] K. S. E. Widmaier, H. Raff, *Human Physiology - The Mechanisms of Body Function*. 2008.

[4] I. Y. Shadrin, A. Khodabukus, and N. Bursac, "Striated muscle function, regeneration, and repair," *Cell. Mol. Life Sci.*, vol. 73, no. 22, pp. 4175–4202, 2016.

[5] A. Vallejo-Illarramendi, I. Toral-Ojeda, G. Aldanondo, and A. López de Munain, "Dysregulation of calcium homeostasis in muscular dystrophies," *Expert Rev. Mol. Med.*, vol. 16, no. October 2014, p. e16, 2014.

[6] E. Taghizadeh, M. Rezaee, G. E. Barreto, and A. Sahebkar, "Prevalence, pathological mechanisms, and genetic basis of limb-girdle muscular dystrophies: A review," *J. Cell. Physiol.*, vol. 234, no. 6, pp. 7874–7884, 2019.

[7] S. Swasdison and R. Mayne, "Formation of highly organized skeletal muscle fibers in vitro. Comparison with muscle development in vivo," *J. Cell Sci.*, vol. 102, no. 3, pp. 643–652, 1992.

[8] A. M. Kasper, D. C. Turner, N. R. W. Martin, and A. P. Sharples, "Mimicking exercise in three-dimensional bioengineered skeletal muscle to investigate cellular and molecular mechanisms of physiological adaptation," *J. Cell. Physiol.*, vol. 233, no. 3, pp. 1985–1998, 2018.

[9] H. Vandenburgh *et al.*, "Drug-screening platform based on the contractility of tissue-engineered muscle," *Muscle Nerve*, vol. 37, no. 4, pp. 438–447, Apr. 2008.

[10] H. Vandenburgh *et al.*, "Tissue-Engineered Skeletal Muscle Organoids for Reversible Gene Therapy," *Hum. Gene Ther.*, vol. 7, no. 17, pp. 2195–2200, 1996.

[11] J. H. Kim *et al.*, "3D Bioprinted Human Skeletal Muscle Constructs for Muscle Function Restoration," *Sci. Rep.*, vol. 8, no. 1, pp. 1–15, 2018.

[12] D. Huh *et al.*, "Microfabrication of human organs-on-chips," *Nat. Protoc.*, vol. 8, no. 11, pp. 2135–2157, 2013.

[13] L. Altomare, M. Riehle, N. Gadegaard, M. Tanzi, and S. Farè, "Microcontact printing of fibronectin on a biodegradable polymeric surface for skeletal muscle cell orientation," *Int. J. Artif. Organs*, vol. 33, no. 8, pp. 535–543, 2010.

[14] L. Sala *et al.*, "MUSCLEMOTION," *Circ. Res.*, vol. 122, no. 3, pp. 5–16, Feb. 2018.

Transistores de grafeno en tatuaje como sensores para EEG

V. Cepero Escribano

Máster en Ingeniería Biomédica, Universidad Internacional de Valencia (VIU) – en colaboración con el Instituto de Microelectrónica de Barcelona- Centro Nacional de Microelectrónica (IMB-CNM) – victor.cep.95@gmail.com

Resumen

Estudios recientes en el Grupo de Aplicaciones Biomédicas (GAB) han demostrado el potencial de los microtransistores de grafeno como alternativa a los electrodos para medidas de la actividad cerebral. El propósito de este trabajo ha sido explorar la escalabilidad de esta tecnología para registrar ondas de baja amplitud con métodos no invasivos. Para ello se ha realizado el prototipado de nuevos dispositivos de mayor tamaño buscando obtener una resolución capaz de detectar señales del orden de decenas de microvoltios. Los transistores han demostrado valores de Vrms inferiores a 10^{-5}V siendo este el límite del sistema de detección y demostrando la capacidad de estos transistores de detectar ondas de baja amplitud.

1. Introducción

El cerebro es una de las estructuras más complejas del cuerpo. La electrofisiología nos permite estudiar la actividad cerebral mediante transductores que convierten las caídas de potencial de las corrientes iónicas en una señal eléctrica que puede ser registrada. Los transductores más habituales son los electrodos, pero el GAB ha propuesto transistores de grafeno (graphene solution gated field effect transistors o gSGFET) como una alternativa con resultados prometedores en métodos invasivos.

Los transductores en métodos no invasivos requieren una preparación de la piel y el uso de un gel conductor en la interfaz electrodo-piel con tal de reducir la impedancia. El desarrollo de tatuajes electrónicos permite un contacto total sin adhesivos artificiales que permite una adaptación completa a la morfología de la piel.

En este trabajo se ha hecho una prueba de concepto fabricando un prototipo de transistor para medidas no invasivas y se han evaluado sus propiedades eléctricas para dar un primer paso necesario para un proyecto mayor. El objetivo de este proyecto es crear transistores de grafeno como un dispositivo flexible, biocompatible y adaptable a la morfología de la piel capaz de detectar las mismas figuras que la tecnología actual sin la necesidad de gel conductor.

2. Conceptos introductorios

2.1. Electrofisiología

La naturaleza eléctrica de la actividad neuronal hace posible la detección señales desde cierta distancia de la fuente. Sin embargo, las medidas se ven afectadas por la distancia en sí y las capas existentes entre la muestra y el transductor. Por ese motivo es necesario determinar las capacidades de detección de los dispositivos y su funcionamiento para entender cómo se registran digitalmente estas señales.[1]

La interfaz cuerpo electrodo suele describirse con un modelo eléctrico en el que las capas del cuerpo se implementan como impedancias en serie. Uno de los objetivos en la fabricación de dispositivos es una baja impedancia, ya que esta supone una mayor relación señal-ruido (SNR).

En el caso de los microelectrodos, las mediciones a bajas frecuencias se ven comprometidas por la relación inversa entre la impedancia y el área del electrodo. [2] En métodos no invasivos se usan electrodos de mayor tamaño pero la señal se ve atenuada en su lugar por las capas entre el córtex cerebral y el electrodo, haciendo que la resolución espacial de los métodos no invasivos sea pobre en comparación con los invasivos.[3]

Para mejorar esta resolución se usan geles conductores en la interfaz piel-electrodo que disminuyen la impedancia que generarían los huecos de aire en electrodos secos. Sin embargo, el secado de estos geles puede generar artefactos en la señal además de precisar de la limpieza tanto del electrodo como de la piel tras su uso.[4]

La implementación de estos electrodos en forma de tatuajes electrónicos puede ser una solución para medidas a largo plazo (días). Su adaptación conformacional con la piel reduciría la impedancia de contacto y directamente aumentaría la SNR. Además, permitiría registrar señales electrofisiológicas con menos susceptibilidad al movimiento y al sudor en un dispositivo transpirable y casi transparente. [5]

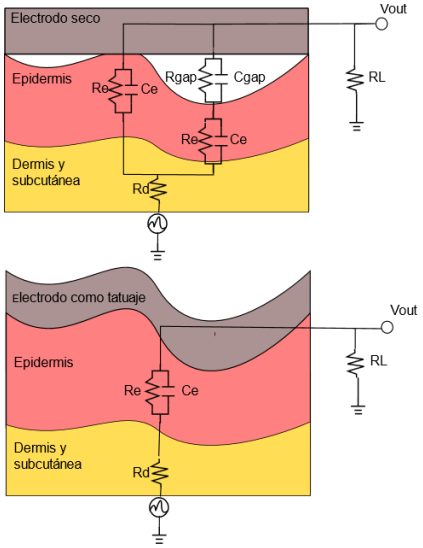


Figura 1. *Comparación entre electrodo seco sobre la piel y electrodo en tatuaje electrónico.[6]*

ISBN: 978-84-09-25491-0, pags. 27- 30

2.2. Transistores de grafeno o gSGFET

En un SGFET el voltaje del Gate (Vg) se aplica mediante un electrodo de referencia sumergido en un electrolito. La configuración habitual es con el Source en común. (figura 2). Se mide la corriente Ids mediante las variaciones de Vgs con un Vds fijo. La figura obtenida con forma de "V" es la que se puede observar en la figura 2. Este curioso efecto ambipolar es característico del grafeno, que dependiendo de su nivel energético conduce mediante electrones en la banda de conducción o huecos en la banda de valencia. Para valores de negativos de Vgs respecto al punto de Dirac (punto de neutralidad de carga o CNP) los huecos en la banda de valencia mejoran la conductividad. En los valores positivos los electrones se inyectan en la banda de conducción. [7]

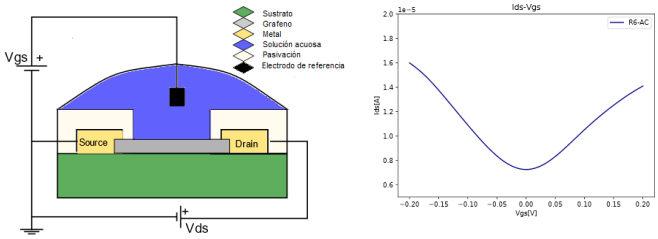


Figura 2. *Ilustración esquemática de gSGFET y curva Ids-Vgs, mostrando el comportamiento ambipolar del grafeno. La curva ha sido normalizada en el punto de neutralidad de carga (CNP)*

El sistema de medida de señales electrofisiológicas consiste en la recolección de distintos rangos de frecuencias que se amplifican con distintas ganancias. La señal calibrada se obtiene de la conversión corriente voltaje.

2.3. Caracterización de µ-gSGFETs

Para la caracterización de las muestras se sumerge el sensor en una solución de PBS junto con el electrodo de referencia. Las propiedades eléctricas de los transistores se obtienen mediante un barrido de Vgs. En este caso cada muestra contaba con 16 canales-transistores (figura 3). Para poder detectar ondas del orden de decenas de microvoltios el valor Vrms debe ser menor a 10^{-5}V.

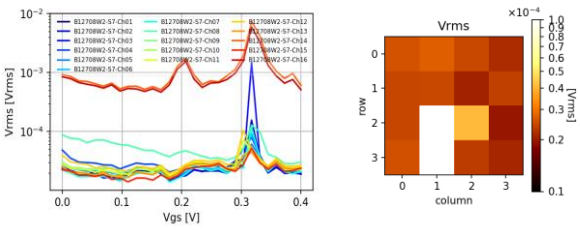


Figura 3. *Valores Vrms de muestra con 16 µSGFETs y respectivo diagrama de posición de cada uno de los canales y su funcionamiento.*

3. Metodologías

3.1. Fabricación de transistores de grafeno

Para la fabricación del prototipo se utilizó una base de PET donde se depositaría la lámina de grafeno que en su aplicación estaría en contacto directo con la piel. El tamaño de la lámina de grafeno se define mediante un grabado de ion reactivo (RIE) y para los contactos metálicos se utiliza evaporación metálica de Ti/Au. Finalmente, para la capa de pasivación que aísla los contactos metálicos del PBS se utiliza una capa adhesiva de polipropileno adhesivo por una cara (PSA).

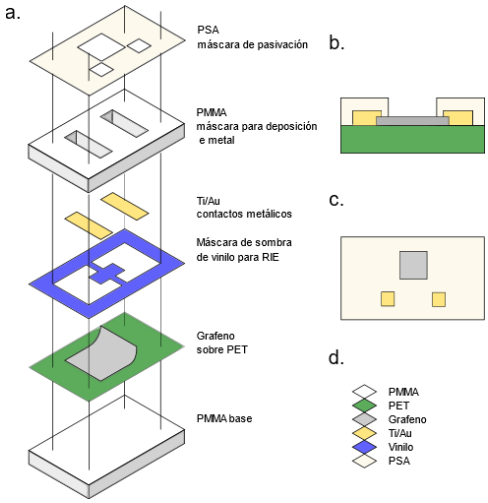


Figura 4. *A. Proceso de fabricación de los prototipos de gSGFET. B. Vista lateral del transistor. C. Vista aérea del transistor. D. Leyenda*

3.2. Diseño del sustrato

Para el prototipo su utilizo una lámina de PET de 100µm de grosor como sustrato. Se escogió este material debido a su alta resistencia a la corrosión, su resistencia química y térmica, su transparencia y su estabilidad en condiciones externas. El grosor fue elegido para su fácil manipulación durante la fabricación y caracterización.

3.3. Crecimiento de grafeno

La capa de grafeno fue crecida en una lámina de cobre de 25µm de grosor mediante deposición química de vapor (CVD) en el instituto de microelectrónica IMB-CNM, CSIC, Barcelona. Para caracterizar este crecimiento se realizó una espectroscopia Raman usando un láser de iones de Ar a 514 nm para la excitación de los electrones.

Los picos de interés en la espectroscopia son el D (~1285cm^{-1}), G (~1580cm^{-1}) y el 2D (~2680cm^{-1}). El D indica el número de defectos en la lámina de grafeno, aumentando su altura con la concentración de defectos. La ratio 2D/G ilustra el número de capas de grafeno. El grafeno monocapa de buena calidad se considera cuando G es pequeño o inexistente, al menos una ratio 2D/G superior a 3. [8]

3.4. Transferencia de grafeno

Para dar estabilidad mecánica al grafeno, antes de la delaminación, se recubrió con una capa de PMMA de ~500nm por centrifugado a 3000rpm. Las láminas de grafeno se extrajeron por delaminación electroquímica siguiendo el proceso descrito por [9]. Añadiendo un voltaje a la lámina de cobre mientras se sumerge en un electrolito, se producen burbujas de hidrogeno entre el cobre y el grafeno separando ambos materiales. Tras la neutralización del grafeno y la eliminación de impurezas se deposita en el substrato de PET.

3.5. Evaporación de metal

Para crear los contactos metálicos que conectaran el grafeno con el Source y el Drain se utilizó evaporación de rayo de electrones en una aleación de Ti/Au, depositando una capa de entre 20-200nm en los bordes laterales del grafeno.

3.6. Adaptación de la geometría del grafeno

Para homogenizar las medidas de la lámina se usó la técnica de grabado por ion reactivo (RIE) usando máscaras de vinilo sobre el prototipo. La efectividad del grabado se comprobó con espectroscopía Raman mostrando presencia de grafeno en la zona cubierta por la máscara y ausencia en la parte expuesta al RIE.

3.7. Pasivación

El proceso de pasivación consiste en la deposición de un material inerte sobre un metal para evitar su interacción con agentes externos. Las máscaras de pasivación se recortaron en un adhesivo de una cara de PSA por su excelente adhesión a substratos de baja energía superficial, su adhesivo inerte de silicona, su alto rango de temperatura de operabilidad, su sistema de adhesivo flexible y especialmente su resistencia adhesiva a soluciones diluidas de DMSO (dimetil sulfóxido), ácidos, alcoholes y solvente orgánicos. Las máscaras dejaban expuestos los contactos y la lámina de grafeno.

3.8. Registro de señales

Para validar el funcionamiento se intercambió el electrodo de referencia por un generador de ondas y así poder comprobar el registro de estas. Se aisló el sistema de detección en una caja de Faraday para evitar las fuentes externas de ruido y se hizo un filtro pasa banda de 3Hz a 45Hz.

4. Resultados y discusión

4.1. Proceso de fabricación

El proyecto empezó con 10 láminas de grafeno con el fin de fabricar 10 gSGFETs, aunque 2 de ellas tuvieron que descartase por errores humanos en la transferencia de grafeno. De las 8 muestras restantes, 4 se metalizaron y caracterizaron directamente y 4 se sometieron al RIE previamente. Del primer grupo una de las muestras no funcionó y de las sometidas al RIE solo una de ellas funcionó después. En definitiva, se fabricaron con éxito 4/10 dispositivos.

Es difícil determinar el punto crítico en la fabricación, ya que el funcionamiento del transistor no se puede comprobar hasta una vez finalizado el proceso. La transferencia de grafeno es un proceso que se realiza habitualmente en el CNM. La metalización no debería afectar a este ya que la lámina queda protegida por la máscara.

Esto da a pensar que el conflicto reside en el RIE, pero la caracterización por Raman indicó su correcto funcionamiento. Probablemente el motivo por el cual fallaron varias de las muestras fue por los daños sufridos en el grafeno durante su manipulación. Los arañazos pueden dañar el canal de grafeno y evitar el paso de corriente a través. Esto, en conjunto con los restos de adhesivo de las máscaras de vinilo pudo haber dañado el dispositivo. Pese a todo, el transistor que sobrevivió a todo el proceso completo fue el que mostró mejores figuras de mérito.

4.2. Caracterización

Para la caracterización de los prototipos se utilizó el mismo proceso descrito en el apartado (2.3) mostrando figuras de valores similares a los µ-gSGFETs, demostrando así su correcto funcionamiento y validando la escalabilidad del dispositivo.

Las principales figuras de mérito de los gSGFETs son la transconductancia, que indica la sensibilidad del transistor a los cambios en Vgs y el Vrms, que es el ruido efectivo en el Gate. Una mayor transconductancia puede obtenerse incrementando Vds y modificando la geometría del canal para tener una mayor relación W/L. Esto nos lleva a una mayor sensibilidad a las variaciones de Vds y en consecuencia a dispositivos más sensibles.

$$V_{RMS} = \sqrt{\int_{f0}^{f1} \frac{S_I}{g_m^2} df} = \frac{I_{RMS}}{g_m^2} \qquad (1)$$

Se busca también que el valor de Vrms sea lo más bajo posible para conseguir una mejor SNR, por lo que aquellos transistores con menor Irms y mayor transconductancia se consideran los mejores para la toma de medidas (fórmula 1). Todos los transistores fabricados mostraron valores Vrms inferiores a $2 \cdot 10^{-5}$ V y el mejor de los fabricados mostró valores de hasta $4 \cdot 10^{-6}$V. Estos valores son suficientes para detectar señales electrofisiológicas del orden de decenas de microvoltios, sin embargo, estos valores no distan tanto de los obtenidos en los micro SGFETs. Considerando que la influencia del ruido es inversamente proporcional al área del dispositivo, dispositivos mayores deberían suponer una menor influencia del ruido. [10]

Para comprobar esto, se realizó una densidad espectral de potencia (PSD) para analizar la influencia del ruido. La figura 5 compara la PSD de un µ-SGFET y la de uno de los fabricados. En el primer caso podemos ver como el sistema se ve limitado por la tolerancia al ruido del transistor. La relación 1/f nos muestra el conocido como

"flicker noise", que se muestra como el principal factor limitante en la medida de señales de baja amplitud y frecuencia. En el caso de los transistores grandes, no se observa esta relación 1/f. Esto nos da a entender que el factor limitante en este caso el ruido intrínseco de la electrónica, que se muestra como una línea plana en frecuencias más altas.

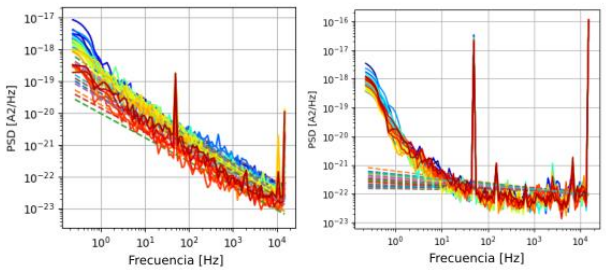


Figura 5. *PSD de µ-SGFETs (izq) y SGFET fabricados (der).*

Esto se puede ver también en la figura 6, donde los valores de Irms modulan y van a valores más bajos a medida que se acercan al CNP, pero en algunos transistores como en el R4 y R6, se muestra claramente una línea plana sin modulación. Esto significa que Irms no puede bajar a valores inferiores ya que está limitada por el sistema de registro.

Para conseguir bajar este nivel de ruido intrínseco se precisaría una modificación de la electrónica. Una primera opción sería modificar el factor de amplificación de Vsig, que además de amplificar la señal también lo hace con el ruido. Se debería buscar el factor óptimo para reducir el ruido con la menor perdida de señal posible.

Esta reducción del factor de ganancia debería compensarse con un cambio en la geometría del transistor. Teniendo en cuenta la relación directa de Irms con W/L, combinado con un incremento de Vds resultaría en una mayor transconductancia y una menor influencia del ruido intrínseco de la electrónica.

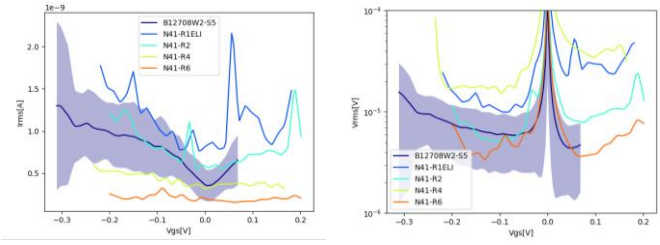


Figura 6. *(Izq.) Valores Irms – Vgs. Los transistores R4 y R6 se ven limitados por la electrónica. (Der.) Valores Vrms-Vgs. Podemos observar valores del orden de 10^{-5}V e incluso menores en muestras como N41-R6 y N41-R2.*

5. Conclusiones

El objetivo del proyecto era implementar los transistores de grafeno como tatuajes electrónicos para registros de EEG. Para ello, será necesario optimizar Vds y la geometría de la lámina de grafeno para conseguir menor ruido, adaptando también la electrónica para que el ruido intrínseco no cubra los valores de interés. Seguidamente, sería conveniente realizar medidas electrofisiológicas para compararlo con los dispositivos actuales.

También quedaría por elegir un material de substrato ideal para su implementación como tatuaje electrónico, ya que al no ser el PET un polímero transpirable no sería la opción óptima. Estudios recientes sugieren laminas finas de PMMA pueden ofrecer mejores prestaciones para este tipo de aplicación. Finalmente quedaría la adaptación del proceso de fabricación a métodos de sala blanca para mejorar la reproducibilidad y promover la escalabilidad de producción.

Las propiedades eléctricas y mecánicas del grafeno pueden proporcionar diferencias interesantes en su implementación en comparación con los dispositivos actuales. Obteniendo medidas con suficiente calidad de señal, los tatuajes de grafeno pueden ofrecer una alternativa cómoda y barata que pueda permitir medidas en tiempos prolongados (días) [6] y que puedan ser realizadas hasta en actividades deportivas, sin la preocupación del secado de geles ni del movimiento del electrodo en sí.

6. Referencias

1. Obien, M.E.J., et al., *Revealing neuronal function through microelectrode array recordings.* Frontiers in Neuroscience, 2015. **8**(423).
2. Masvidal-Codina, E., et al., *High-resolution mapping of infraslow cortical brain activity enabled by graphene microtransistors.* Nat Mater, 2019. **18**(3): p. 280-288.
3. Burle, B., et al., *Spatial and temporal resolutions of EEG: Is it really black and white? A scalp current density view.* International Journal of Psychophysiology, 2015. **97**(3): p. 210-220.
4. Xing, X., et al., *A High-Speed SSVEP-Based BCI Using Dry EEG Electrodes.* Scientific Reports, 2018. **8**(1): p. 14708.
5. Yeo, W.H., et al., *Multifunctional Epidermal Electronics Printed Directly Onto the Skin.* Advanced materials (Deerfield Beach, Fla.), 2013. **25**.
6. Kabiri Ameri, S., et al., *Graphene Electronic Tattoo Sensors.* ACS Nano, 2017. **11**(8): p. 7634-7641.
7. Hébert, C., et al., *Flexible Graphene Solution-Gated Field-Effect Transistors: Efficient Transducers for Micro-Electrocorticography.* Advanced Functional Materials, 2018. **28**(12): p. 1703976.
8. Malard, L.M., et al., *Raman spectroscopy in graphene.* Physics Reports, 2009. **473**(5): p. 51-87.
9. Lockhart de la Rosa, C.J., et al., *Frame Assisted H2O Electrolysis Induced H2 Bubbling Transfer of Large Area Graphene Grown by Chemical Vapor Deposition on Cu.* Applied Physics Letters, 2012. **102**.
10. Simoen, E. and C. Claeys, *The low-frequency noise behaviour of silicon-on-insulator technologies.* Solid-State Electronics, 1996. **39**(7): p. 949-960.

Dispositivo inteligente para el estudio de la viabilidad de riñones para trasplante

D. Naranjo Hernández[1], J. Reina Tosina[1], L.M. Roa Romero[1], M.Á. Pérez Valdivia[2,3], M. Salgueria Lazo[4,3,1], R.A. Medina López[2,3]

[1] Grupo de Ingeniería Biomédica, Dep. de Teoría de la Señal y Comunicaciones, Universidad de Sevilla, Sevilla, España, {dnaranjo, jreina, lroa}@us.es

[2] Unidad de Gestión Clínica de Urología y Nefrología, Hospital Universitario Virgen del Rocío, Sevilla, España, {miguel.perez.sspa, rantonio.medina.sspa}@juntadeandalucia.es

[3] Instituto de Biomedicina de Sevilla (IBIS), Sevilla, España

[4] Servicio de Nefrología, Hospital Universitario Virgen Macarena, Sevilla, España, mercedes.salgueira.sspa@juntadeandalucia.es

Resumen

El presente trabajo propone una primera aproximación a un dispositivo sensor inteligente basado en la técnica de espectroscopia de bioimpedancia para el estudio de la viabilidad de riñones para trasplante. El dispositivo es capaz de realizar medidas de bioimpedancia sobre un riñón en múltiples frecuencias configurables, procesar los datos para obtener el módulo y la fase de bioimpedancia en cada una de las frecuencias y transmitir de forma inalámbrica la información procesada. Un algoritmo de procesado determina los parámetros del modelo de Cole-Cole que mejor se ajusta a los valores de bioimpedancia, siendo estos parámetros el resultado final y la información útil en la medida de bioimpedancia. El trabajo se centra en describir las adaptaciones realizadas a un prototipo preliminar para su aplicación a medidas sobre riñones en un entorno quirúrgico: identificación de alternativas y diseño de la sonda de medida teniendo en cuenta la necesidad de esterilización, mejoras en la etapa de inyección de corriente para una mayor precisión, calibración, ajuste de ganancia, entre otras. Se presentan además unos primeros resultados de medidas sobre riñones de porcino y humanos que confirman la viabilidad de la tecnología propuesta.

1. Introducción

Actualmente hay cerca de 4000 personas en España esperando un trasplante de riñón [1]. El número de pacientes en tratamiento sustitutivo renal está aumentando de forma progresiva. Sin embargo, la cantidad de riñones disponibles no si incrementa con la misma velocidad [2]. Para abordar este problema, en las últimas décadas se han considerado a donantes con edades superiores, a pesar de que la edad es un factor que puede influir en la función del injerto renal, su supervivencia o la supervivencia del paciente receptor [3].

La biopsia renal pre-implantación, y el informe anatomo-patológico resultante, es un protocolo clínico que se ha introducido para garantizar la viabilidad del injerto renal. Sin embargo, la implantación de este protocolo ha derivado en un aumento en el porcentaje de órganos descartados, 27% en Andalucía y 25% en España en 2019, a pesar de tener una funcionalidad adecuada para el trasplante renal [1]. Esta situación puede explicarse si se tiene en cuenta el grado de subjetividad del procedimiento [4]: existe una gran variabilidad inter-observador en la interpretación de los resultados de la biopsia, las técnicas de procesamiento del tejido renal para el análisis histológico no son uniformes y tienden a sobrestimar la glomeruloesclerosis, el sistema de puntuación es subjetivo y depende del especialista que lo realice.

Esta situación justifica la investigación de métodos complementarios a la biopsia pre-implantación que permitan evaluar de forma objetiva la viabilidad de los injertos renales para reducir los factores subjetivos del procedimiento actual.

La tecnología de bioimpedancia podría ser aplicada con tal propósito si se tiene en cuenta que proporciona información sobre las condiciones fisiológicas de los tejidos vivos [5], y su dinámica temporal es sensible a variaciones estructurales y morfológicas [6]. Algunos estudios han aplicado la tecnología de bioimpedancia en la monitorización del estado de isquemia en órganos [7], la detección de tejidos tumorales, o la evaluación objetiva de la viabilidad de cultivos celulares y tejidos [8]. Existen estudios que han evaluado el uso de la bioimpedancia en el estudio de la viabilidad de riñones para trasplante de hígado [9], pero hasta el conocimiento de los autores todavía no ha sido empleada en el trasplante renal. Las propiedades conductivas y de bioimpedancia de un riñón vivo no se conocen en detalle, a pesar de que la variabilidad de la bioimpedancia en el tiempo puede tener una relación directa con la evolución de la isquemia [10].

El objetivo del presente trabajo es describir las adaptaciones realizadas sobre un dispositivo sensor inteligente de bioimpedancia patentado por el Grupo de Ingeniería Biomédica de la Universidad de Sevilla [11, 12], para que pueda ser aplicado en el estudio de la viabilidad de riñones para trasplante. También se presentan los primeros resultados obtenidos con un prototipo preliminar.

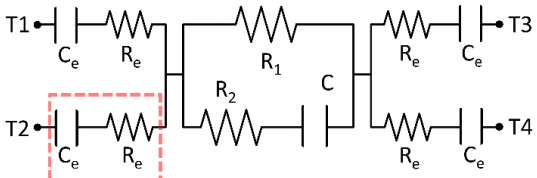
2. Materiales y métodos

La adaptación del dispositivo de bioimpedancia ha seguido una metodología de diseño centrado en usuario, siguiendo el concepto de modularidad funcional para facilitar la integración de nuevas tecnologías y funcionalidades, y

ISBN: 978-84-09-25491-0, pags. 31- 34

sobre la base del empleo de estándares. Para la simulación de los componentes electrónicos se ha utilizado el programa TINA de Texas Instruments (versión 9.3). Para el diseño electrónico se empleó el paquete software Cadence OrCAD (versión 16.0). Algunos de los requisitos de diseño han sido bajo costo, precisión, inocuidad, inmunidad frente a interferencias electromagnéticas, portabilidad, comunicación inalámbrica y un diseño compacto y esterilizable para facilitar su uso en quirófano.

Para una primera validación de la etapa de instrumentación del dispositivo se ha diseñado un patrón circuital que permite emular diferentes valores de bioimpedancia en un rango afín a los que se obtendrían en medidas de bioimpedancia sobre un riñón humano (los experimentos realizados en riñones definieron el rango de estudio). La Figura 1 muestra el esquemático del patrón circuital y las diferentes configuraciones que permite. El patrón incluye también un modelo de electrodo formado por un condensador C_e y una resistencia R_e en serie para una mejor aproximación a la realidad, siendo T1 y T3 los terminales de inyección de corriente y T2 y T4 los de medida de tensión. Los valores de los componentes fueron definidos según la experiencia de los investigadores. Para evaluar la precisión del dispositivo se empleó la distancia euclídea en el plano complejo de bioimpedancia como medida del error, analizando 22 frecuencias distribuidas de forma logarítmica entre 5 kHz y 1 MHz.



Configuración	R_1 (Ω)	R_2 (Ω)	C (nF)
1	400	1000	3
2	500	1500	2
3	600	2000	1
4	700	2500	0.5

Figura 1. *Patrón circuital empleado en la validación.*

Para una primera evaluación del dispositivo de medida sobre medios biológicos se utilizaron riñones procedentes de cerdos sacrificados para consumo humano, dada la semejanza anatómica entre el riñón humano y el porcino. Las primeras medidas fueron realizadas a los pocos minutos de la extracción del órgano, y fueron repetidas en diferentes instantes de tiempo para permitir un análisis de la evolución temporal de los parámetros de bioimpedancia.

Para confirmar la viabilidad del dispositivo en medidas sobre riñones humanos se realizó una segunda evaluación con riñones recién extirpados procedentes de pacientes sometidos a cirugía de nefrectomía completa. Dicho estudio fue aprobado por el Comité Ético del Hospital Universitario Virgen del Rocío de Sevilla.

3. Resultados

3.1. Diseño y desarrollo

El sistema de medida de bioimpedancia presenta la siguiente arquitectura de diseño modular (ver Figura 2):

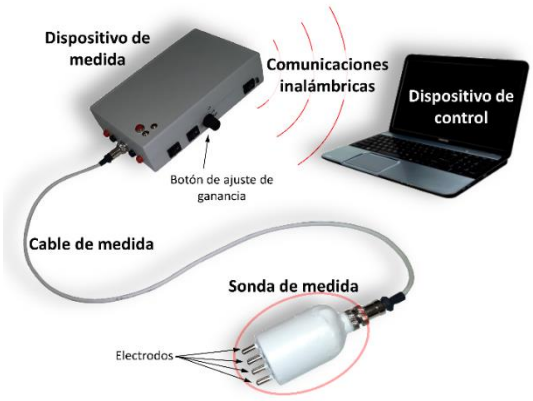


Figura 2. *Esquema del sistema de bioimpedancia propuesto.*

- **Sonda de medida**: Para mantener la inocuidad del escenario de medida se ha considerado que los electrodos sean incorporados en una o dos sondas que puedan ser fácilmente esterilizadas.

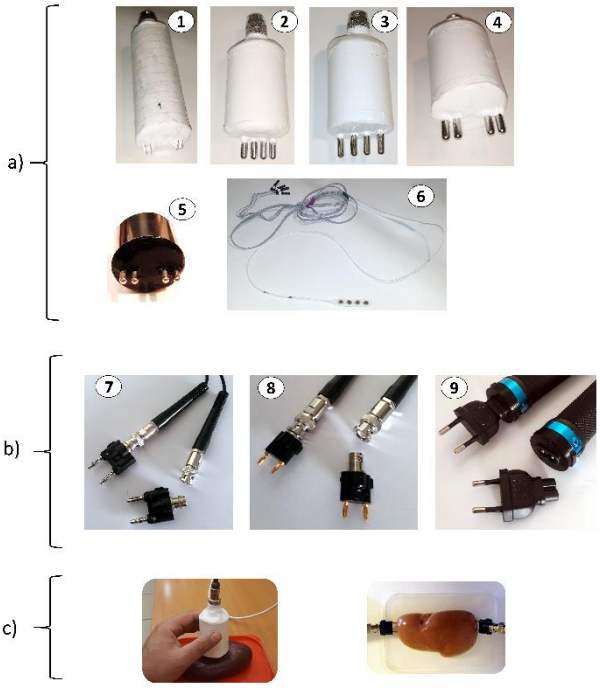


Figura 3. *Sondas investigadas y desarrolladas: a) configuraciones basadas en una sonda; b) configuraciones con dos sondas; c) ejemplos de medidas sobre riñones.*

Se decidió emplear un montaje de cuatro electrodos frente a la solución estándar y más sencilla de dos electrodos [12], ya que permite eliminar el efecto de la impedancia de contacto entre el electrodo y el tejido renal, la cual puede ser muy elevada y variable, mejorando así la precisión del dispositivo. Los equipos clínico y técnico analizaron la posibles ubicaciones y montajes de la/s sonda/s, derivando una solución que permite alcanzar un compromiso entre sensibilidad, especificidad y usabilidad. Factores como la distancia, forma y dimensiones de los electrodos también fueron investigados, así como aspectos mecánicos y físicos de la/s sonda/s como el peso, la presión de medida o el área de sujeción. La Figura 3 muestra algunas de las sondas que se analizaron. La sonda 6 en la Figura 3 muestra un ejemplo basado en electrodos comerciales desechables y/o

esterilizables (electrodos subdurales intraoperatorios de neurofisiología, MS04R-IP10X-0JH de Ad-Tech). La sonda 5 fue seleccionada como la más adecuada entre las configuraciones de una única sonda ya que puede ser esterilizada mediante calor y porque proporciona un mayor valor de bioimpedancia debido a la mayor distancia entre los electrodos de medida de tensión. Actualmente se están analizando las configuraciones de dos sondas.

- **Cable de medida**: Un cable separa la sonda del dispositivo de medida para mejorar la ergonomía del proceso y facilitar la esterilización de los materiales. Conectores RS Pro de la serie ZC de 4 pines (tipo BNC en el caso de sondas duales) fueron empleados en sus extremos por sus características: apantallados, permiten que el montaje sea sencillo y tienen una conexión segura de bloqueo por roscado/bayoneta. El cable de medida de 170 cm de longitud está apantallado para evitar problemas de ruido electromagnético.

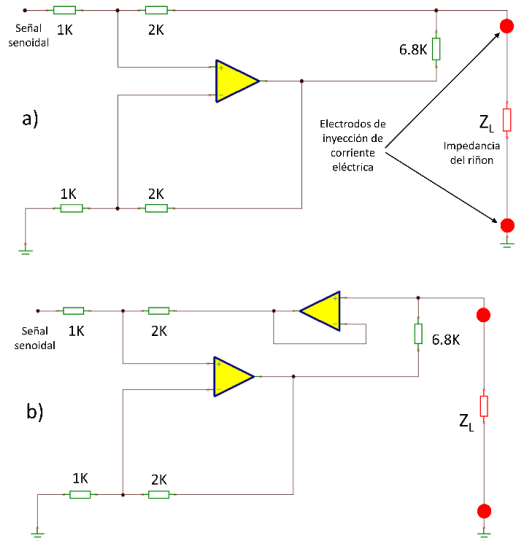


Figura 4. *Esquemáticos de la fuente de corriente: a) dispositivo previo [11, 12]; b) fuente de corriente propuesta.*

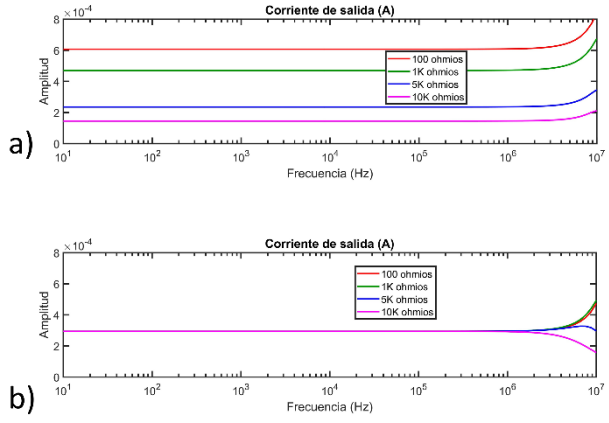


Figura 5. *Amplitud de la corriente eléctrica inyectada en función de la impedancia de carga: a) dispositivo previo [11, 12]; b) fuente de corriente propuesta.*

- **Dispositivo de medida**, encargado de realizar la medida de bioimpedancia propiamente dicha. Respecto del dispositivo previo [11, 12] se propone una mejora en la etapa de generación de corriente eléctrica para hacerla independiente del valor de la impedancia de carga (bioimpedancia del riñón). La Figura 4 ilustra el esquema previo y el propuesto y la Figura 5 pone de relieve las mejoras obtenidas en los resultados. Aunque en el prototipo preliminar la calibración ajustaba la estimación de la bioimpedancia en función de la carga, la etapa de generación de corriente propuesta proporciona una mayor precisión en la medida y permite asegurar el valor de corriente eléctrica que circulará por el riñón.

  Por otro lado, la ganancia del amplificador de instrumentación es variable y configurable manualmente mediante un botón de ajuste para prevenir los problemas derivados de la variabilidad de los valores de bioimpedancia en riñones.

- **Comunicaciones**: Se ha utilizado el módulo RN42 de Microchip para el desarrollo de las comunicaciones inalámbricas del dispositivo de medida con el dispositivo de control conforme al perfil SPP del estándar de Bluetooth v2.1.

- **Dispositivo de control**: El cual se encarga de realizar un segundo procesado empleado el algoritmo descrito en [12] para determinar los parámetros del modelo de Cole-Cole que se corresponde con los valores de bioimpedancia. Los parámetros del modelo son el resultado final y la información útil en la medida de bioimpedancia. Estos parámetros, y las relaciones entre ellos pueden ser sensibles al estado morfológico y fisiológico del órgano y su evolución [10, 12]. El dispositivo de control también se encarga del almacenamiento de la información y sirve de interfaz con el usuario. El software del dispositivo de control ha sido implementado en un ordenador portátil Satellite P50 de Toshiba sobre el entorno de programación MATLAB (versión R2018a).

Ante el elevado número de alternativas estudiadas (diferentes sondas, diferentes cables y longitudes, diferentes condiciones de ganancia en la medida de tensión, etc.), y como cada alternativa precisa de una calibración diferente del sistema, se empleó el procedimiento de auto-calibración semi-automático descrito en [12]. Una vez calibrado el dispositivo, una medida de bioimpedancia multifrecuencia puede ser realizada en ocho segundos.

4. Validación

Para cada alternativa de diseño (sonda, longitud de cable, ganancia, etc.), y una vez calibrado el dispositivo, se evaluaron todas las configuraciones del patrón circuital. El error cometido en todas las estimaciones fue inferior al 1%.

La Figura 6 muestra, a modo de ejemplo, algunos de los resultados obtenidos en los dos estudios realizados para evaluar la viabilidad del dispositivo para la realización de medidas sobre riñones. La Figura 6.a muestra los valores de bioimpedancia en uno de los experimentos realizados sobre riñones de cerdo. La Figura 6.b representa los valores de bioimpedancia correspondientes a una medida sobre un

riñón humano quince minutos después de ser extirpado en una cirugía de nefrectomía completa.

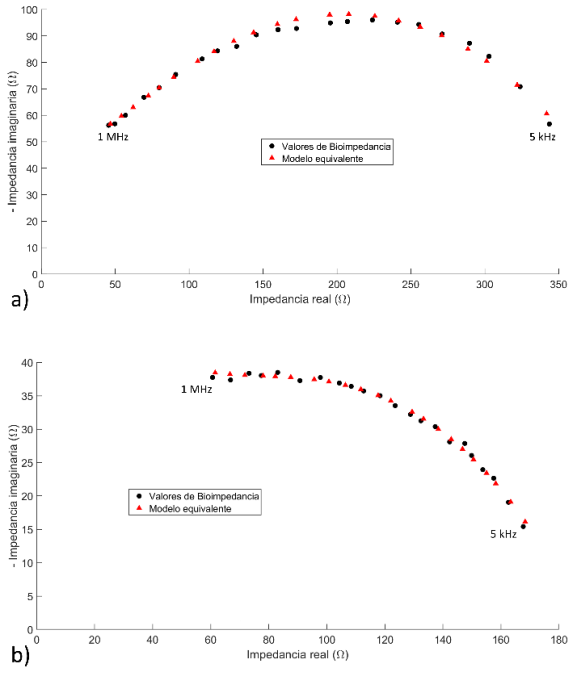


Figura 6. *Ejemplos de valores de bioimpedancia obtenidos con la sonda 5: a) riñón de cerdo; b) riñón humano.*

5. Discusión

Los resultados de precisión, con un error inferior al 1% sobre un patrón circuital, son comparables a los obtenidos en sistemas similares [13]. El uso de dicho patrón, con resistencias y condensadores de parámetros conocidos, es necesario para evaluar las estimaciones, ya que los valores de bioimpedancia del riñón no pueden ser conocidos a priori. Los primeros resultados sobre muestras biológicas confirman la viabilidad de la tecnología propuesta para la realización de medidas de bioimpedancia en riñones y animan a continuar las investigaciones.

6. Conclusiones

Se ha mostrado un primer prototipo de un dispositivo inteligente de bioimpedancia para su aplicación en el estudio de la viabilidad de riñones para trasplante. Los resultados obtenidos con este prototipo preliminar han mostrado la precisión y viabilidad de la solución propuesta. El dispositivo propuesto puede aportar una medida que complemente a la biopsia para objetivar la toma de decisiones en el proceso de trasplante, optimizar su eficiencia aumentando las posibilidades de los pacientes de recibir un injerto renal viable, reducir la tasa de órganos desechados y minimizar las listas de espera, lo que supondría una mejora tanto social como económica para el Sistema Sanitario Público.

Agradecimientos

Queremos agradecer al Dr. José Luis Rocha Castilla, al Dr. Juan José Egea-Guerrero y a la Dra. Ana Rodríguez-Rodríguez su colaboración en el trabajo. También queremos mostrar nuestro agradecimiento a la empresa Almevan S.L. por facilitarnos una muestra de los electrodos subdurales y a la empresa GEPIELMA S.L. por proporcionarnos los riñones de cerdo.

Este trabajo ha sido financiado en parte por la Fundación Mutua Madrileña, bajo el proyecto VÍA-RENAL, en parte por la Fundación Progreso y Salud, Gobierno de Andalucía, bajo el proyecto PIN-0394-2017, y el Programa Eurostars de la Comisión Europea (proyecto FRAIL).

Referencias

[1] O. N. de Trasplantes. Memoria actividad donación y trasplante renal. *Publicaciones del Ministerio de Sanidad. Gobierno de España*, 2019.

[2] López-Sánchez P, Portolés J, Rodríguez L, et al. Impact of first year renal replacement therapy on the hospital admissions of a regional public health system. *Nefrologia*, vol 39, no 6, 2019, pp 653–663.

[3] Martins J, Barreto S, Bravo P, Santos J, Ferreira M, Oliveira C, Ramos A. Kidney transplant from elderly donors: A center experience. *Transplantation Proceedings*, vol 52, no 5, 2020, pp 1265-1268.

[4] Reeve J, Böhmig G, Eskandary F, Einecke G, Lefaucheur C, Loupy A, Halloran P. Assessing rejection-related disease in kidney transplant biopsies based on archetypal analysis of molecular phenotypes. *JCI insight*, vol 2, no 12, 2017.

[5] Yao J, Wang L, Liu K, Wu H, Wang H, Huang J, Li J. Evaluation of electrical characteristics of biological tissue with electrical impedance spectroscopy. *Electrophoresis*, vol 41, no 16-17, 2020, pp 1425-1432.

[6] Emran S, Lappalainen R, Kullaa A, Myllymaa S. Concentric ring probe for bioimpedance spectroscopic measurements: Design and ex vivo feasibility testing on pork oral tissues. *Sensors*, vol 18, no 10, 2018.

[7] Strand-Amundsen R, Tronstad C, Kalvoy H, Ruud T, Hogetveit J, Martinsen O, Tonnessen T. Small intestinal ischemia and reperfusión - Bioimpedance measurements. *Physiological Measurement*, vol 39, no 2, 2018.

[8] Amini M, Hisdal J, Kalvøy H, Applications of bioimpedance measurement techniques in tissue engineering. *Journal of Electrical Bioimpedance*, vol 9, no 1, 2018, pp 142–158.

[9] Wang H, He Y, Yan Q, You F, Fu F, Dong X, Shi X, Yang M. Correlation between the dielectric properties and biological activities of human ex vivo hepatic tissue. *Physics in Medicine and Biology*, vol 60, no 6, 2015, pp 2603–2617.

[10] Zhu J, Shi X, You F, Wang H, Wang H, Cai Z, Dong X. Preliminary research on relationship between dielectric properties and microstructure of rabbit liver. IFMBE Proceedings, Beijing, China, vol 39, pp 643–646, 2013.

[11] Naranjo D, Reina J, Roa LM, Barbarov G, Cejudo P, Ortega F. Dispositivo inteligente de espectroscopía de bioimpedancia para la estimación de la composición corporal. *XXXVII Congreso Anual de la Sociedad Española de Ingeniería Biomédica*. Santander, España, 2019.

[12] Naranjo-Hernández D, Reina-Tosina J, Roa LM, Barbarov-Rostán G, Aresté-Fosalba N, Lara-Ruiz A, Cejudo-Ramos P, Ortega-Ruiz F. Smart bioimpedance spectroscopy device for body composition estimation. *Sensors*, vol 20, no 1, 2020.

[13] Allegri D, Donida A, Malcovati P, Barrettino D. CMOS-based multifrequency impedance analyzer for biomedical applications. *IEEE Transactions on Biomedical Circuits and Systems*, vol 12, no 6, 2018, pp 1301–1312.

Biomateriales

Miércoles, 25 de noviembre de 2020

Primeros pasos para el desarrollo de retina biohíbrida 3D basada en fibroína de seda para la degeneración macular relacionada con la edad (DMAE)

N. Jemni-Damer[1,2,§], A. Guedan-Duran[1,2,3,§], J. Cichy[1,4], P. Lozano-Picazo[5], D. Gonzalez-Nieto[5,6,7], J. Perez-Rigueiro[5,6,7,8], F. Rojo[5,6,7,8], G. Guinea[5,6,7,8], A. Virtuoso[9], G. Cirillo[9], M. Papa[9], F. Armada-Maresca[10], C. Largo-Aramburu[11], S. D. Aznar-Cervantes[12], J. L. Cenis[12], F. Panetsos[1,2,7,*]

[1] Neuro-computing & Neuro-robotics Research Group, Complutense University of Madrid, Spain

[2] Innovation Research Group, Institute for Health Research San Carlos Clinical Hospital (IdISSC), Madrid, Spain

[3] Department of Biomedical Engineering, Tufts University, Medford, Massachusetts 02155, United States

[4] Faculty of Biology. Christian Albrechts University, Kiel, Germany

[5] Center for Biomedical Technology. Polytechnic University of Madrid, Pozuelo de Alarcón, Madrid, Spain

[6] Biomedical Research Networking Center in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain

[7] Silk Biomed SL, Madrid, Spain

[8] Dept. Materials Science, ETSI de Caminos, Canales y Puertos. Polytechnic University of Madrid, Spain

[9] Division of Human Anatomy – Neuronal Networks Morphology Lab, Dept. Mental, Physical Health and Preventive Medicine, University of Campania "Luigi Vanvitelli", Naples, Italy

[10] Ophthalmology Service, La Paz University Hospital, Madrid, Spain

[11] Experimental Surgery. University Hospital, La Paz (IdiPAZ), Madrid Spain

[12] Instituto Murciano de Investigación y Desarrollo Agrario y Alimentario (IMIDA), Department of Biotechnology La Alberca (Murcia), E-30150, Spain

[§] Equal contribution

[*] Corresponding author

E-mail: fivos@ucm.es

Resumen

La degeneración macular relacionada con la edad es una enfermedad neurodegenerativa crónica incurable, que provoca una pérdida progresiva de la visión central e incluso ceguera. Todos los tratamientos actuales tienen como objetivo ralentizar la progresión de la enfermedad. En el presente artículo proporcionamos evidencia in vitro de la viabilidad de retinas biohíbridas en capas construidas con fibroína del gusano de seda y cultivos de diferentes tipos de células: epitelio pigmentario retiniano, neuronas retinianas, células madre mesenquimales y de Müller. Las capas de la retina biohíbrida están pegadas con hidrogel de fibroína de seda. Las pruebas in vitro muestran que las neuronas de cultivo primario crecen en biofilms de fibroína de seda solas, o en presencia de otras células cultivadas en el mismo biomaterial, la organización de las células y los fenotipos celulares se mantienen durante los 7 días del experimento.

1. Introducción

La degeneración macular asociada con la edad (DMAE) es una patología neurodegenerativa crónica que provoca la pérdida progresiva de visión central conduciendo, en su estadío más avanzado, a la ceguera. Además, es considerada una de las principales causas de pérdida irreversible de la visión entre sujetos mayores de 65 años siendo sus mecanismos de patogénesis heterogéneos y aún no bien conocidos [1]. La membrana de Bruch (MBr) es una capa de 2-5 µm que se encuentra entre la coriocapilar y el epitelio pigmentario de la retina (EPR), constituyendo la lamina basal de ambos. En condiciones normales, desempeña un papel fundamental para el correcto funcionamiento de las células del EPR.

Sin embargo, se sabe que con la edad [2], el estrés oxidativo [3] y la genética [4], se debilita, engrosa y se hace más rígida (disminuye su elasticidad y flexibilidad debido a un aumento de la reticulación de colágeno) perdiendo así su permeabilidad y dinámica natural. Esto, va a ralentizar el intercambio molecular, lo que provoca la acumulación de lípidos, restos celulares y metabólicos que se originan en el EPR y el tejido circundante, dando lugar a la aparición de drusas entre la membrana basal del EPR y la MBr. En etapas más avanzadas de la enfermedad, estos depósitos también se pueden encontrar entre el EPR y la neurorretina [5]. Todos estos cambios, van a provocar perdida de la integridad de la barrera hematorretiniana externa y la muerte de las células del EPR, lo que conlleva la pérdida

ISBN: 978-84-09-25491-0, pags. 36- 39

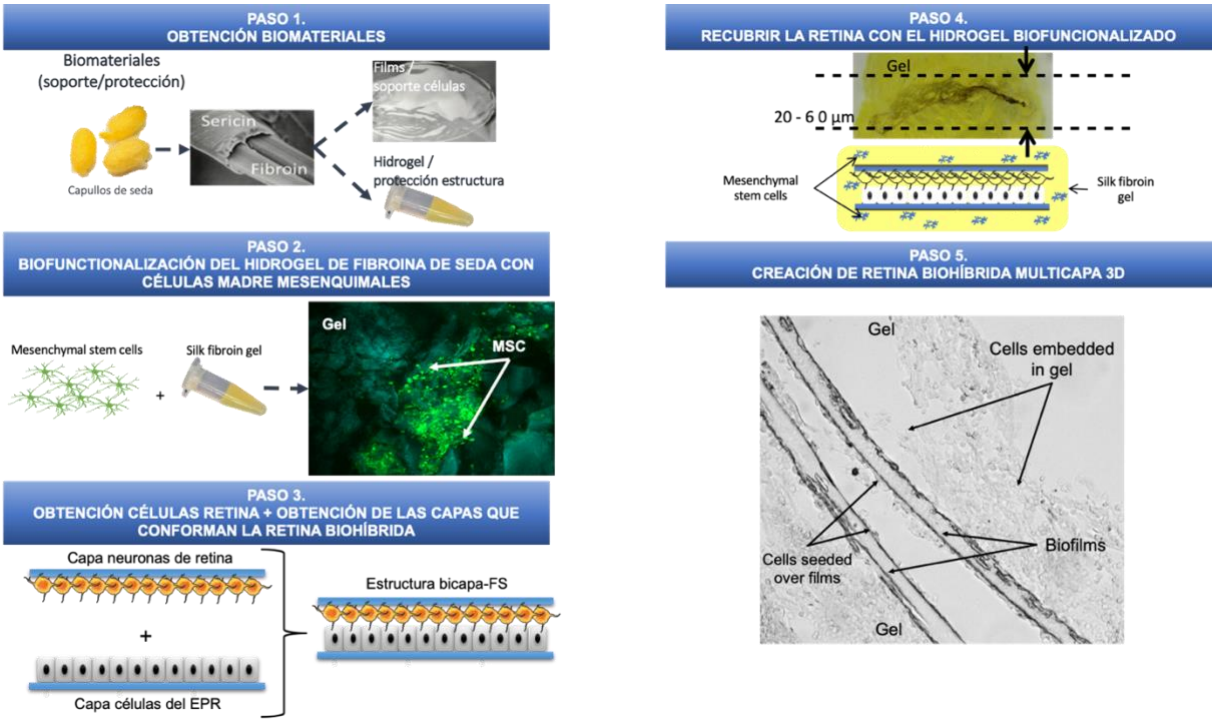


Figura 1. *Creación in vitro paso a paso de la retina biohíbrida multicapa 3D.*

de la función visual debido a la muerte de los fotorreceptores, neurona muy especializada de la retina encargada de la percepción visual, al transformar la luz en señales biológicas (debido a que los segmentos externos de los fotorreceptores forman una unión íntima con el EPR). También se producirán alteraciones a nivel de la coriocapilar, neovasos subcoroideos penetrarán al espacio subretinal, a través de los defectos existentes en la MBr y el EPR, dando lugar a la forma húmeda de la DMAE [6].

Se han investigado diversos scaffolds artificiales, con células del EPR sembradas, cuyo objetivo es imitar una MBr sana para ser trasplantada en sujetos con DMAE avanzada y de este modo reemplazar el tejido retiniano dañado. Estos biomateriales deben ser biocompatibles, a la vez que promover y mantener el fenotipo adecuado del EPR. Además, deben ser porosos, finos (menos de 10 µm, imitando la membrana de Bruch) y mecánicamente competentes para resistir la manipulación durante la cirugía.

En este trabajo mostramos el futuro prometedor y las aplicaciones del uso de la fibroína de seda como base de la fabricación de biomateriales. Gracias a la composición de nuestros biomateriales, las estructuras creadas son completamente biocompatibles, siendo totalmente inertes y no producen ningún tipo de rechazo inmunológico. Mostrando la versatilidad de este compuesto y las múltiples aplicaciones de los diferentes formatos que somos capaces de fabricar.

En el presente estudio describimos los primeros pasos para diseñar, caracterizar y desarrollar “in vitro” una retina biohíbrida 3D multicapa recubierta de hidrogel de fibroína de seda biofuncionalizado (Figura 1). Nuestros biohíbridos se basan en células del epitelio pigmentario y neuronas de la retina, cultivadas por separado sobre biofilms de fibroína de seda y luego unificadas mediante hidrogeles de seda enriquecidos con células de soporte (células de Müller y células madre mesenquimales).

2. Métodos

2.1. Fabricación de los biomateriales

Para la fabricación de nuestros biomateriales hemos empleado la fibroína de seda, siguiendo el protocolo descrito en [7]. La fibroína de seda se trata de un biomaterial biocompatible, biodegradable, inerte y esterilizable. Es el componente principal de las fibras de seda del gusano de seda y uno de los biomateriales más prometedores de la naturaleza debido a su alta biocompatibilidad con los tejidos del cuerpo y, en particular, con los del sistema nervioso central. La fibroína de seda se extrae a partir de los capullos del gusano *Bombyx mori*. Una vez limpios, estos se trocean en pequeños trozos y se les realiza una serie de procesos entre los cuales esta el desgomado, en el cual la presencia de sericina se elimina de los capullos, siendo la fibroína la única proteína restante. Una vez obtenida la fibroína de seda en polvo, se puede preparar de varias maneras para conseguir distintos formatos del material con estructuras y propiedades mecánicas diferentes. En nuestro estudio, hemos usado la fibroína en forma de Biofilm e Hidrogel.

2.2. Desarrollo in vitro

Obtención de células de cultivo primario

Una vez que los globos oculares de la rata se enuclean, se someten a diversas incubaciones con diferentes enzimas para lograr la separación de la retina neural, de la que se obtienen las neuronas y el complejo esclera-coroides-EPR del que se separan las células del EPR. Estas células del EPR y las neuronas de la retina se siembran sobre diferentes biofilms de fibroína de seda para crear las diferentes capas de la retina biohíbrida (Figura 1, paso 5). Las células del EPR se extrajeron siguiendo el protocolo descrito en [8] y las neuronas de [9]. Las células de soporte empleadas fueron la línea celular MU-PH1, una línea celular inmortalizada obtenida a partir de un cultivo de células de Müller de retinas de ratones adultos, y células madre mesenquimales (MSC) obtenidas según el protocolo descrito en [10].

Creación de retina biohíbrida multicapa 3D

Para conseguir el ensamblaje de las dos capas celulares y de este modo obtener una retina artificial multicapa 3D, estas se recubrieron con un hidrogel de fibroína de seda biofuncionalizado con células de müller y células madre mesenquimales. Para ello hemos seguido los siguientes pasos (Figura 1):

1) Obtención de los biomateriales que van a dar soporte y protección a la retina biohíbrida (Paso 1).
2) Obtención de las células mesenquimales a partir de la médula ósea de ratones, según el protocolo descrito en [10] y a continuación, biofuncionalización de los hidrogeles de fibroína de seda con células madre mesenquimales. En la imagen se observa el gel con las células mesenquimales verdes en su interior (Paso 2).
3) Obtención de las capas que conforman la retina biohíbrida. Sembrado de las células del EPR y las neuronas sobre los biofilms de fibroína de seda y su posterior ensamblaje (Paso 3). Lo que nos permitió que hubiera contacto entre las neuronas y las células del EPR sembradas.
4) Después del ensamblaje de las células del EPR y las neuronas a los biofilms de fibroína de seda, para dar forma, consistencia, protección y aportar factores neurotróficos de forma continuada, y de este modo mejorar la supervivencia y viabilidad de las células que conforman la retina biohíbrida, esta fue recubierta por hidrogel de fibroína de seda biofuncionalizado con células madre mesenquimales (Pasos 4-5).

2.3. Evaluación del tiempo de supervivencia celular in vitro

La viabilidad celular a los 7 días de la siembra y la conservación del fenotipo original se estudió mediante técnicas de inmunofluorescencia contra marcadores específicos (RPE-65 para células del epitelio pigmentario de la retina y NeuN para neuronas).

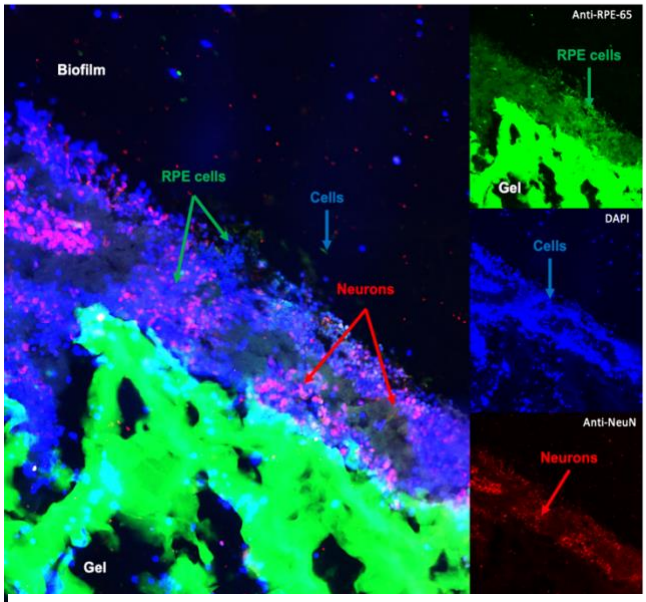


Figura 2 *Reconstrucción de secciones verticales de retina biohíbrida multicapa 3D que muestra la supervivencia celular después de siete días in vitro. Una vez que la retina biohíbrida se ensambla con el gel de fibroína de seda, las células se mantienen vivas en el interior.*

3. Resultados

Con los resultados obtenidos podemos decir que somos capaces de crear una retina biohíbrida utilizando cultivos celulares y sustratos artificiales. Las células utilizadas son de cultivo primario, lo que nos proporciona una fuente celular que se asemeja mucho a una retina natural y nos permite estudiar la interacción entre diferentes tipos celulares y entre dichas células y su entorno, sin necesidad de difíciles protocolos de diferenciación y obtención de células funcionales. Estas células necesitan un sustrato adecuado para su supervivencia y que facilite su adhesión. Además, el sustrato (biofilm) permite manipular la estructura creada en un futuro implante in vivo. Los sustratos creados se obtienen mediante la funcionalización de la fibroína de seda, que proporciona sitios de unión a las células retinianas y nos permite adaptar el sustrato para cada tipo celular. Pudimos crear la estructura de la retina 3D con dos biofilms adheridas, permitiendo así el contacto entre las neuronas y las células del EPR sembradas. Además, los biomateriales de fibroína de seda empleados no producen citotoxicidad favoreciendo la viabilidad celular y la adhesión. La idoneidad de los sustratos fabricados se demuestra por la supervivencia de las células obtenidas por cultivo primario durante largos períodos de tiempo, hasta 7 días in vitro (Figura 2).

La versatilidad de la fibroína de seda nos permite obtener biomateriales con diferentes formatos, aptos para diversas funciones. La posibilidad de crear un hidrogel nos permitió incluir células de soporte y crear un entorno favorable y protector para la retina biohíbrida. Además, este hidrogel proporciona un ambiente húmedo y con propiedades físicas similares a las del tejido en el que serán implantados, lo

que provocaría un menor rechazo del tejido huésped en caso de futuros estudios in vivo. Además, nos permite aislar las células implantadas del entorno hostil, producido por la cirugía de un cuerpo extraño, y además la producción de factores favorables para la supervivencia celular gracias a las células de soporte incluidas en el hidrogel.

Como se puede ver en la Figura 2, tras montar la estructura 3D, esta se mantuvo en el incubador de cultivos a 37º de temperatura y 5% de CO2 durante 1 día para estudiar la viabilidad de las células encapsuladas en su interior. En esta imagen, observamos las tinciones inmuno-histológicas de la reconstrucción de secciones verticales de la retina 3D in vitro obtenidas mediante microscopia fluorescente, donde se observa uno de los biofilms, las células encapsuladas y el hidrogel de fibroína de seda. En las imágenes pequeñas de la derecha se puede observar: el hidrogel teñido con Fluoresceine isocyanate (verde), DAPI marca el núcleo de las células vivas (azul) y NeuN marca núcleo de las neuronas (rojo).

4. Conclusiones

Los resultados obtenidos demuestran que la estructura biohíbrida 3D artificial sería un método terapéutico prometedor para restaurar la visión en estadios avanzados de patologías degenerativas de la retina. Nos da la posibilidad de ser implantado para aplicaciones in vivo. Gracias a las características biocompatibles de la fibroína de seda, el rechazo producido por el implante de la retina va a poder ser reducido al mínimo. Por ultimo, la creación de un ambiente protector alrededor de la retina artificial nos va a asegurar una ventana de supervivencia de la retina tras el implante, esencial para la posibilidad de que la retina sea integrada en el tejido huésped.

Referencias

[1] K. Michalska-Małecka, A. Kabiesz, M. Nowak, and D. Śpiewak, Age related macular degeneration – challenge for future: Pathogenesis and new perspectives for the treatment, *European Geriatric Medicine*, vol. 6, 2015, pp. 69–75 (ISSN: 18787649).

[2] V. Daneault, M. Dumont, É. Massé, G. Vandewalle, and J. Carrier, Light-sensitive brain pathways and aging, *Journal of Physiological Anthropology*, vol. 35, 2016, p. 9 (ISSN: 18806805).

[3] J. Hanus, C. Anderson, and S. Wang, RPE necroptosis in response to oxidative stress and in AMD, *Ageing Research Reviews*, vol. 24, 2015, pp. 286–298 (ISSN: 18729649).

[4] C. M. L. Lau, Y. Yu, G. Jahanmir, and Y. Chau, Controlled release technology for anti-angiogenesis treatment of posterior eye diseases: Current status and challenges, *Advanced Drug Delivery Reviews*, vol. 126, 2018, pp. 145–161 (ISSN: 18728294).

[5] J. Damián, R. Pastor, F. Armadá, and L. Arias, "Epidemiología de la degeneración macular asociada con la edad. Situación en España," *Atencion Primaria*, vol. 38, 2006, pp. 51–57 (ISSN: 02126567).

[6] M. F. Marmor and T. J. Wolfensberger, The retinal pigment epithelium: function and disease. Oxford University Press, 1998 (ISBN: 9780195109566).

[7] L. Fernández-García, N.Marí-Buyé, JA. Barios, R. Madurga, M. Elices, J. Pérez-Rigueiro, M. Ramos, GV. Guinea and D. González-Nieto, Safety and tolerability of silk fibroin hydrogels implanted into the mouse brain, *Acta Biomaterialia*, vol. 45, 2016, pp. 262–275 (ISSN: 18787568).

[8] J. Shen, J. He, and F. Wang, Isolation and Culture of Primary Mouse Retinal Pigment Epithelial (RPE) Cells with Rho-Kinase and TGFβR-1/ALK5 Inhibitor, *Medical science monitor*, vol. 23, 2017 pp. 6132–6136 (ISSN: 116433750).

[9] M. L. Rocco, B. O. Balzamino, P. P. Passeri, A. Micera, and L. Aloe, Effect of Purified Murine NGF on Isolated Photoreceptors of a Rodent Developing Retinitis Pigmentosa, *PLoS ONE*, vol. 10, 2015 pp. e0124810 (PMID: 25897972).

[10] Y. Martín-Martín, Martín-Martín, L. Fernández-García, MH. Sanchez-Rebato, N.Marí-Buyé, FJ. Rojo, J. Pérez-Rigueiro, M. Ramos, GV. Guinea and D. González-Nieto and F. Panetsos, Evaluation of Neurosecretome from Mesenchymal Stem Cells Encapsulated in Silk Fibroin Hydrogels, *Scientific Reports*, vol. 9, 2019, p. 8801 (ISSN: 20452322).

Modelización in-silico de la Rigidez y Resistencia de Estructuras Porosas Biodegradables

J. Aldazabal[1], G. Artola[2]

[1] TECNUN - Universidad de Navarra, Donostia - San Sebastián, España (jaldazabal@tecnun.es)

[2] AZTERLAN, Basque Research and Technology Alliance (BRTA). Durango, España (gartola@azterlan.es)

Resumen

Una estrategia muy empleada en la medicina regenerativa es el empleo de estructuras porosas, o scaffolds, biodegradables. Estas estructuras son las encargadas inicialmente de servir como soporte para la proliferación celular. A medida que el nuevo tejido va creciendo la estructura se degrada y es reabsorbida por el cuerpo. El proceso finaliza cuando las células han regenerado el tejido dañado y el scaffold se ha disuelto por completo. Las propiedades mecánicas del scaffold son clave para determinar el correcto crecimiento del tejido celular a lo largo de este proceso. En este trabajo se propone un modelo combinado para estudiar la degradación y la pérdida de propiedades mecánicas producidas en estos materiales porosos. Una primera parte del modelo se encarga de simular la degradación superficial que sufren estas estructuras en medios fisiológicos. La segunda parte de modelo permite estimar tanto la rigidez como la carga de rotura en cada momento de la degradación.

Los resultados del modelo permiten relacionar la evolución temporal de la microestructura, la pérdida de masa, el módulo elástico y la carga de rotura.

1. Introducción

Una estrategia ampliamente utilizada en la medicina regenerativa es el empleo de estructuras porosas, o scaffolds, biodegradables [1]. Estas estructuras son las encargadas inicialmente de servir como soporte, o scaffold, para la proliferación celular [2]. A medida que el nuevo tejido va creciendo, la estructura se degrada y es reabsorbida por el cuerpo. El proceso finaliza cuando las células han regenerado el tejido dañado y el scaffold se ha disuelto por completo. Las propiedades mecánicas del scaffold son clave para determinar el correcto crecimiento del tejido celular a lo largo de este proceso [3].

La simulación mediante computador de la degradación de dichas estrucuturas, así como de su rigidez, supone una gran ayuda para los ingenieros biomédicos responsables del diseño de dichos scaffolds. Estas simulaciones permiten un ahorro de tiempo en el desarrollo de un implante de este tipo [4].

En este trabajo se ha desarrollado una metodología para generar estructuras porosas bidimensionales de forma controlada, así como un algoritmo para poder estimar la rigidez y resistencia mecánica de dichas estructuras.

2. Descripción del modelo

En el siguiente apartado se expone el modelo desarrollado en este trabajo para la generación de estructuras porosas bidimensionales, así como su degradación y la estimación de su respuesta mecánica. En todo momento se ha trabajado con estructuras periódicas para minimizar los efectos de los bordes.

2.1. Microestructura inicial

Para generar la estructura bidimensional lo primero que se ha de considerar es su tamaño. Para ello, se fija la altura y la anchura del dominio rectangular.

Una vez determinado el tamaño del dominio se determina el número de huecos, o poros, a introducir en la estructura. Los poros se han de colocar de la forma más uniforme posible sobre el dominio definido. En un principio, lo ideal es escoger el tamaño del dominio que pueda acomodar todos los poros formando una red hexagonal regular. Para obtener geometrías más realistas es posible desplazar aleatoriamente los poros una pequeña distancia predeterminada.

Tras determinar la posición de los poros se ha realizado una teselación de Voronoi empleando dichos poros como semillas. Para realizar esta teselación se ha discretizado el dominio de trabajo empleando una red regular en M×N elementos, o píxeles. Tras dicha teselación, a los pixeles se les han asignado uno de tres siguientes valores:

- 1,0 para los elementos de las juntas entre dos celdas de Voronoi.
- 0,0 a los pixeles que contienen el centro de un poro.
- 0,5 al resto de los píxeles que no cumplen ninguna de las dos condiciones anteriores.

Una vez asignados dichos valores, se ha procedido a suavizar el campo resultante. Este suavizado no se ha aplicado sobre los puntos con valores 1 y 0. El resultado de este algoritmo es una estructura de Voronoi "suavizada", con una gama de valores que varía desde 0 en los centros de los poros hasta 1 en los bordes de las celdas de Voronoi.

Una vez generado este campo de valores, la estructura porosa se obtiene sin más que escoger los píxeles que tienen un valor mayor o igual que un umbral predeterminado. Cuanto más próximo sea dicho umbral a 0, se cogerán más elementos, y cuanto más próximo a 1 sea ese umbral, menos elementos serán considerados.

Una vez generada la microestructura y para determinar si se produce una rotura en algún punto del dominio, se asigna a cada elemento un valor de resistencia, r_{ij}, en términos del valor de tensión máxima que soporta antes de romperse. Con esto se construye una matriz de resistencia $\mathbf{R}_{M\times N}$. Los valores de resistencia se han asignado a los diferentes píxeles de forma aleatoria con un valor arbitrario promedio de 0,0011 y una dispersión de ±0,0001.

ISBN: 978-84-09-25491-0, pags. 40- 43

2.2. Degradación de la microestructura

En este trabajo se ha supuesto una degradación de la estructura superficial homogénea y a velocidad constante. Para la simulación de la degradación se ha empleado un algoritmo de erosión clásico [5]. El kernel empleado para la degradación ha sido cuadrado y de tamaño 3×3. En cada una de las etapas de degradación el kernel anteriormente mencionado se ha colocado centrado sobre cada uno de los píxeles no-sólidos y se ha eliminado todo el sólido cubierto por el kernel. De esta manera en cada erosión se ha eliminado una capa homogénea de material de espesor 1 píxel.

2.3. Estimación de las tensiones mecánicas

La componente de solicitaciones mecánicas se ha abordado desde la perspectiva de los desplazamientos. Para ello se ha discretizado el dominio de interés en una red homogénea y regular de M×N puntos, o píxeles, como se ha indicado anteriormente. A continuación, se ha asignado a cada nodo un escalar denominado "norma", *u*, que representa el estado de desplazamiento de cada elemento respecto a su posición inicial [6].

Esto se materializa en una matriz $\mathbf{U}_{M\times N}$, cuyos elementos valen inicialmente u_{ij}=0. Para simular la aplicación de una deformación sobre la estructura, se han impuesto condiciones de contorno sobre la cara superior e inferior del dominio. En la cara inferior, se ha fijado el valor de las normas durante todas las simulaciones como 0, para reproducir la fijación de la microestructura. Para un cierto nivel de deformación, se ha impuesto en la cara superior un valor positivo fijo de la norma, equivalente al desplazamiento macroscópico que se quiere simular.

Una vez aplicadas las condiciones de contorno, se deduce el equilibrio del sistema mediante un algoritmo que equilibra el reparto de la norma en todo el dominio. Este equilibrio se produce en las condiciones en las que el gradiente de la norma es mínimo en cada punto. La resolución de este equilibrio se ha implementado utilizando el algoritmo iterativo de difusión de diferencias finitas presentado en la expresión (1).

$$u_{i,j}^{s+1} = u_{i,j}^{s} + \beta\left(u_{i+1,j}^{s} + u_{i-1,j}^{s} + u_{i,j+1}^{s} + u_{i,j-1}^{s} - n u_{i,j}^{s}\right) \quad (1)$$

Mediante esta ecuación se calcula la evolución de la norma en cada elemento *i,j* de la matriz en la iteración *s*+1, a partir de los valores de las normas propia y de los *n* vecinos más cercanos (generalmente 4) en la etapa anterior, *s*. En el caso de que alguno de los vecinos no exista, como ocurre en los nodos adyacentes a superficies libres o grietas, su norma no es considerada y la cantidad de norma transferida a dicho vecino es nula.

El parámetro β, coeficiente de equilibrado de normas, determina la magnitud de la intensidad de transferencia de norma en cada etapa de cálculo. El algoritmo será estable siempre y cuando se asegure la desigualdad $\beta \leq 0{,}2$.

Como criterio de convergencia, se considera que se llega al equilibrio cuando no se produce ninguna transferencia de norma entre elementos de valor superior a un valor prefijado. La figura 1 representa el resultado de este equilibrado de normas en una muestra cuadrada a la que se ha practicado un agujero. La cara superior del dominio, coloreada de rojo, corresponde a la posición donde el desplazamiento, o norma, se ha impuesto como condición de contorno, mientras que el borde inferior, coloreado de azul, reproduce la condición de contorno de desplazamiento cero.

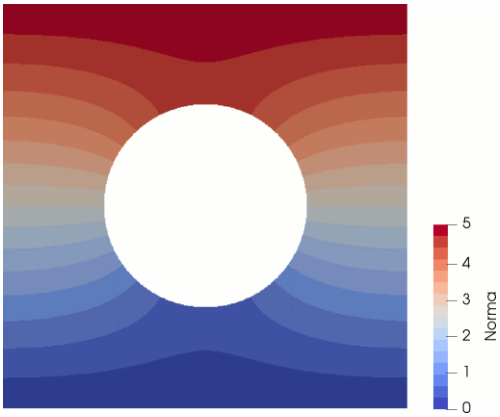


Figura 1. *Distribución de normas alrededor de un agujero circular.*

Al observar la figura 1 de distribución de normas, se detecta que el desplazamiento no se ha repartido homogéneamente en cada línea horizontal de la matriz **U**. Durante el equilibrado, se generan gradientes de desplazamiento en las direcciones horizontal y vertical por efecto del orificio central. Las matrices de gradiente horizontal y vertical, $\mathbf{H}_{M\times N}$ y $\mathbf{V}_{M\times N}$ respectivamente, se calculan según las expresiones (2) y (3).

$$h_{i,j} = \frac{u_{i+1,j} - u_{i-1,j}}{2} \quad (2)$$

$$v_{i,j} = \frac{u_{i,j+1} - u_{i,j-1}}{2} \quad (3)$$

Ambas expresiones se deben corregir ligeramente en las posiciones contiguas a las superficies libres, en cuyo caso se toma sólo la diferencia entre el valor $u_{i,j}$ de interés y el elemento contiguo del lado opuesto a la superficie libre.

A partir de las matrices de gradiente horizontal y vertical, **H** y **V**, operando término a término como se indica en (4) se introduce el escalar "deformación".

$$def_{i,j} = \sqrt{v_{i,j}^2 + h_{i,j}^2} \quad (4)$$

La matriz $\mathbf{DEF}_{M\times N}$ cuyos elementos son $def_{i,j}$, se interpreta como la deformación máxima soportada en cada punto, sin atender a su dirección.

A partir de la matriz de deformaciones, y suponiendo un material elástico lineal, se puede deducir una tensión elástica aproximada existente en cada punto del sistema simulado. En lo que resta de este trabajo, se denomina a esta tensión elástica aproximada simplemente "tensión". Para conocer la fuerza total que sufre una microestructura simulada se han considerado las tensiones de todos los elementos de una misma fila horizontal. La fila

considerada para calcular la carga aplicada es la adyacente a la de aplicación de las condiciones de contorno.

2.4. Condiciones de fractura

El modelo propuesto determina la posible rotura de un elemento de la microestructura en función de la tensión que sufre y la relación de esta tensión con la resistencia de dicho elemento y su entorno.

Para reproducir un comportamiento verosímil del modelo se distinguen dos situaciones diferentes: la nucleación de las grietas en superficies y su posterior crecimiento. La integración de estos fenómenos en el modelo se requiere la incorporar de algunos criterios operativos adicionales en el algoritmo, tanto para el proceso de nucleación de grietas como para su crecimiento. Estos criterios se basan en la consideración de un entorno de tipo Moore, es decir, considerando los elementos vecinos de arista y vértice.

Para iniciar una grieta se han de cumplir las siguientes condiciones:

- Dos elementos contiguos (horizontal-vertical o en diagonal) han de superar simultáneamente la tensión de rotura.
- Sólo uno de dichos elementos ha de estar en contacto con una superficie libre, es decir, la grieta ha de nuclear hacia el interior del material.
- El número de vecinos sólidos del que está en contacto con la superficie ha de ser 4 o 5.

El crecimiento de la grieta, una vez nucleada, se desarrolla de elemento en elemento, utilizando como criterio de rotura la superación de la tensión de rotura y las siguientes condiciones:

- El número de vecinos sólidos del elemento que ha superado la condición de tensión crítica debe ser 7.
- Las grietas no pueden crecer realizando giros de 90 o más grados.

3. Resultados

El modelo ha partido de una estructura con 36 poros dispuestos en un dominio periódico discretizado en 500×500 elementos. Para lograr una fracción volumétrica de sólido de 0,35 se ha utilizado un umbral de 0,944.

Una vez generada la estructura se ha procedido a su degradación. Tras cada una de las degradaciones se ha procedido al ensayo mecánico virtual de las estructuras resultantes hasta rotura.

3.1. Evolución microestructural

En la figura 2 se han presentado los puntos calculados hasta la aplicación de 4 erosiones consecutivas ya que, a partir de esa aplicación, y pese a que sigue existiendo una cierta fracción de sólido, la microestructura pierde su contigüidad.

3.2. Respuesta mecánica

La figura 3 muestra las curvas tensión-deformación obtenidas para cada una de las estructuras degradadas, correspondientes a cada etapa de erosión.

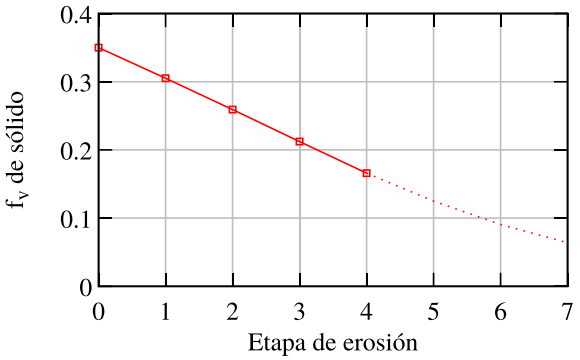


Figura 2. *Evolución de la fv de sólido durante la degradación.*

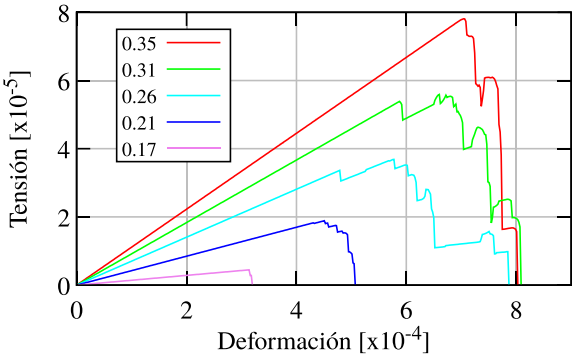


Figura 3. *Respuesta mecánica de las estructuras porosas en distintos momentos de la degradación. (Nótese que, con la definición utilizada en este trabajo, la tensión es adimensional)*

4. Discusión

En el siguiente apartado se va a proceder a discutir los resultados mecánicos obtenidos por el modelo propuesto.

En vista a los resultados mostrados en la figura 3, y como era de esperar, tanto la rigidez como la resistencia de las estructuras decae con la pérdida de masa.

La figura 4 muestra la evolución de la rigidez, medida como la pendiente en la parte lineal de la respuesta mecánica, y refleja su caída con la pérdida de masa.

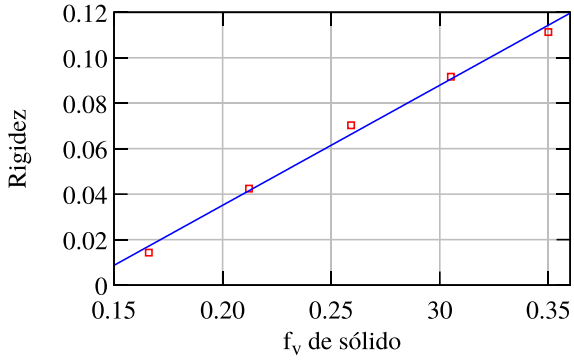


Figura 4. *Evolución de la rigidez de la microestructura con la fracción de sólido.*

La evolución de la rigidez se ha ajustado empleando mínimos cuadrados a la expresión lineal:

$$Rigidez = 0{,}5283\, f_v - 0{,}07056 \qquad (5)$$

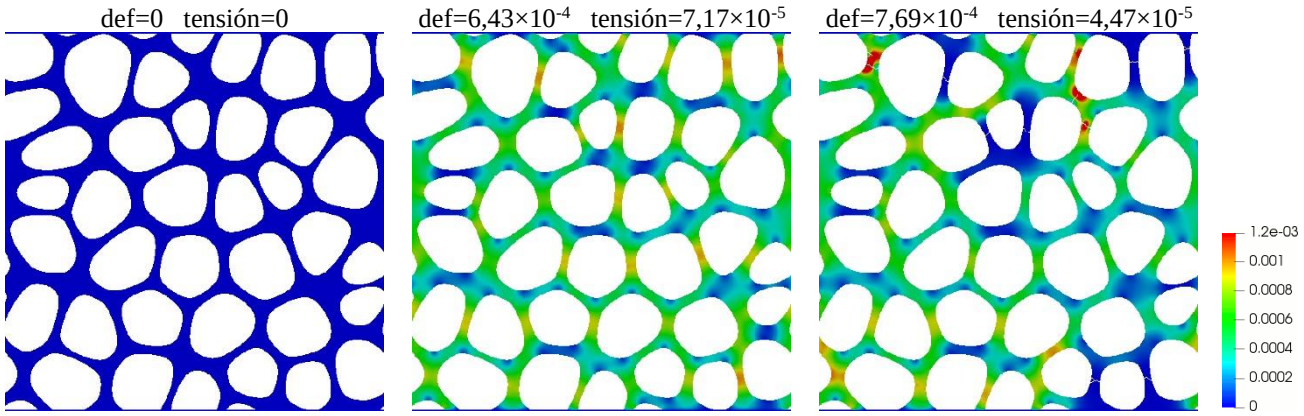


Figura 5. *Distribución de las tensiones en la microestructura según aumenta la deformación aplicada. En cada caso representado se indica la deformación instantanea impuesta y la tensión resultante obtenida.*

Al extrapolar este ajuste se podría predecir que cuando la fracción volumétrica de sólido es 0,13 la rigidez de la estructura es cero, es decir, pierde la continuidad. Asimismo, empleando el ajuste se esperaría una rigidez de 0,45 cuando la estructura fuera totalmente sólida. Para ese caso concreto se sabe que la rigidez de la estructura es 1. Dada la geometría de los poros este valor de 0,45 predicho es muy razonable. Esto es debido a que la estructura simulada se podría simplificar como una red ortogonal de ligamentos sólidos, en la cual solo la mitad de ellos, los alineados con la carga, aportarían rigidez al sistema.

Procediendo de forma análoga, se ha obtenido una relación entre fracción volumétrica de sólido y tensión máxima de:

$$T_{máx} = 3{,}999 \times 10^{-4}\, f_v - 6{,}454 \times 10^{-5} \qquad (6)$$

En este caso la resistencia 0, es decir, pérdida total de propiedades de la estructura, se daría con una fracción volumétrica de 0,16; valor ligeramente superior al de 0,13 predicho con la ecuación (5). Para una fracción volumétrica de 1 cabría esperar una tensión de rotura de la microestructura de 0,0011, ya que este es el valor inicialmente impuesto para los píxeles sólidos. Sin embargo, el valor extrapolado por la ecuación (6) sería de 0,00033, que es un valor inferior al esperado, pero también razonable dada la disposición de los ligamentos sólidos en las estructuras utilizadas.

La figura 5 muestra la evolución de las tensiones en el caso de la estructura no degradada en tres instantes del ensayo virtual a tracción. En la primera imagen de esta figura se ve la estructura en la etapa inicial. La estructura está descargada, es homogénea y se puede apreciar la periodicidad de la misma. La segunda imagen corresponde a una etapa intermedia donde se aprecia que la tensión no se distribuye de forma homogénea y que se concentra en los ligamentos alineados con la dirección de la carga aplicada a la microestructura. En la tercera y última etapa se pueden apreciar zonas relajadas (azules oscuras) en las proximidades de los ligamentos fracturados, así como la concentración de tensiones generadas en los frentes de las grietas (zonas rojas).

El modelo para el cálculo de tensiones aquí propuesto es capaz de predecir la distribución cualitativa de las mismas de forma satisfactoria, si bien no es capaz de ofrecer resultados cuantitativos precisos. La principal ventaja del modelo es su sencilla implementación, permitiendo su uso para estudios de fenómenos combinados, como por ejemplo la degradación de un scaffold en función de la distribución de tensiones mecánicas presentes.

5. Conclusiones

El algoritmo propuesto para la generación de estructuras porosas ha sido capaz de reproducir geometrías realistas de una forma controlada.

La aplicación de erosiones geométricas ha sido útil para simular una degradación homogénea de la micro-estructura.

El modelo de distribución de normas y tensiones se corresponde cualitativamente a lo esperado.

El modelo de fractura propuesto ha permitido reproducir de una forma realista tanto la nucleación como el crecimiento de grietas esperado en estructuras porosas.

Referencias

[1] A. Ataya. Regenerative medicine strategies. J. Pediatric Surgery (2012) 47:17-28.

[2] E.C. Novosel, C. Kleinhans, P.J. Kluger. Vascularization is the key challenge in tissue engineering. Adv. Drug Delivery Rev. (2011) 63:300-311.

[3] G.C. Reilly, A.J. Engler. Intrinsic extracellular matrix properties regulate stem cell differentiation. J. Biomech. (2010) 43:55-62.

[4] J.J. Mao, G. Vunjak-Navokovic, A.G. Mikos, A. Atala. Translational approaches in tissue engineering and regenerative medicine. Ed. Artech House Publishers. (2007) ISBN 1596931116.

[5] R. Fisher, S. Perkins, A. Walker, E.Wolfart. Hypermedia Processing Reference. Ed. John Wiley & Sons Ltd. (1996) ISBN 0471962430.

[6] G. Artola. Susceptibilidad a la fragilización por hidrógeno de aceros de alta resistencia: Comportamiento en ambientes marinos y modelización de patrones de agrietamiento. Tesis Doctoral, Universidad de Navarra: Donostia-San Sebastián 2019.

Estudio de la influencia del diámetro y temperatura del nozzle en el grosor del filamento para bioimpresión

E. Mancha Sánchez[1], J.C. Gómez-Blanco[1], A.C. Marcos Romero[2], M. Matamoros Pacheco[2], F.M. Sánchez-Margallo[1], J.B. Pagador Carrasco[1]

[1] Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España

{emancha, jcgomez, msanchez, jbpagador}@ccmijesususon.com

[2]Escuela de Ingenierías Industriales, Universidad de Extremadura, Badajoz, España,

{acmarcos, manuelmp }@unex.es

Resumen

La bioimpresión es uno de los campos con más recorrido dentro de la biomedicina. Permite generar estructuras como tejidos o incluso órganos con mucha precisión. Para asegurar esta precisión hay que tener en cuenta ciertos factores tanto de los materiales como de la bioimpresora. La temperatura y el diámetro del nozzle son dos de los elementos mas importantes a la hora de asegurar la precisión. En nuestro caso, para estudiar estos parámetros se han realizado simulaciones computacionales modificando la temperatura (15, 25 y 37 °C) y el diámetro del nozzle (20G, 22G y 25G) para estudiar el grosor del filamento extruido. Se han obtenido diferentes comportamientos en la evolución del grosor del filamento, dependiendo de la temperatura y el diámetro del nozzle. El grosor del filamento siempre es mayor que el diámetro del nozzle. Por lo que se hace necesario tener en cuenta estas dos condiciones a la hora de generar estructuras bioimpresas.

1. Introducción

La bioimpresión es la técnica que se sirve de la impresión 3D tradicional utilizando como material hidrogeles y células. El objetivo de esta técnica es replicar o crear estructuras, tejidos u órganos que puedan sustituir o reemplazar estos elementos de un paciente. Para posibilitar estas creaciones hay muchas disciplinas que han de coexistir, tales como la ingeniería de materiales, biología, electrónica o ingeniería tisular, entre otras.

Dos de los elementos más importantes en el campo de la bioimpresión son el material utilizado y las células. El material utilizado va a depender del tipo de construcción que se quiera realizar, si requiere más o menos rigidez, porosidad, degradación, etc. En este aspecto los materiales más utilizados son alginato, gelatina, y GelMA [1]. La modificación de las características de los materiales (viscosidad, densidad, etc) proporciona unas propiedades diferentes en la construcción realizada, pero además facilita o dificulta su impresión. La modificación de estas características puede hacer que las células no sobrevivan al proceso de impresión [2], haciendo inútil la utilización de este material.

Los parámetros de bioimpresión también se ven afectados por el material utilizado. Dentro de los parámetros más importantes destacan la temperatura de impresión, la velocidad, la presión de extrusión, y la altura de capa. Los tres primeros parámetros son de los más estudiados en bioimpresión, ya que están relacionados con la parte inicial de la impresión, algunas veces relacionada con la printabilidad [3]. El parámetro de la altura de capa es, a menudo, infravalorado, suponiendo en muchos casos un factor diferencial a la hora de conseguir una estructura fiable y estable [4]. Esto es debido a que una altura de capa demasiado alta puede provocar que las distintas capas no se adhieran entre sí, haciendo que la construcción sea inestable y colapse. Por otro lado, si la altura de capa es demasiado baja se genera exceso de material, lo que afecta a su precisión. Este parámetro va a estar determinado por el grosor del filamento impreso. La formación de este filamento ha sido estudiada según diferentes velocidades, de impresión o de presión de extrusión [5], pero siempre de manera empírica o mediante cálculo experimental [6].

El estudio de estos parámetros, así como su influencia en la viabilidad celular, es un proceso costoso, tanto a nivel de tiempo como de dinero. Estos inconvenientes pueden solventarse mediante simulación computacional. La dinámica de fluidos computacional (CFD, por sus siglas en inglés) nos permite modificar las características físicas de los materiales, para así poder estudiar su comportamiento. El CFD es utilizado ampliamente en muchos campos, desde la medicina [7], hasta la ingeniería [8]. La simulación computacional se ha utilizado en más ocasiones para estudiar parámetros de la bioimpresión, tanto de la extrusión en sí, como del comportamiento de los materiales, o su combinación.

En este trabajo se han realizado una serie de simulaciones computacionales del sistema de bioimpresión modificando la temperatura y el calibre del nozzle para estudiar el grosor del filamento obtenido.

2. Material y Métodos

Se crearon tres geometrías diferentes en COMSOL Multiphysics mediante un modelo 2D axisimétrico y la interfaz “Two-Phase flow level set”. Los tres modelos representados corresponden a tres puntas cónicas de diferente calibre más utilizadas en bioimpresión (20G, 22G, y 25G). Se consideraron dos dominios diferentes, aire y biotinta. El material para este último dominio fue Cellink Bioink, una biotinta comercial formada por alginato y nanocelulosa. De las que sus características reológicas (viscosidad y densidad) fueron obtenidas de Cellink.

ISBN: 978-84-09-25491-0, pags. 44- 47

El método de *level set* es una ecuación de transporte que se añade a las ecuaciones de Navier-Stokes para seguir la interfaz entre dos líquidos inmiscibles, en este caso la biotinta y el aire, estas ecuaciones son:

$$\rho \frac{\partial u}{\partial t} + \rho\,(u \cdot \nabla)u = \nabla \cdot [-pI + \mu(\nabla u + \nabla u^T)] + F + \rho g$$

$$\nabla \cdot u = 0$$

donde, ρ es la densidad, u la velocidad del fluido, p la presión, I la matriz identidad, μ es la viscosidad del fluido, F son todas las otras fuerzas externas, y g la fuerza de la gravedad.

En la ecuación del *level set*, la biotinta se expresa como $\phi = 0$, el aire como $\phi = 1$, y la línea de contorno como $\phi = 0.5$. La ecuación de level set se puede ver como el porcentaje de volumen de líquido en el flujo de dos fases formado por el aire-líquido. De este modo, la ecuación de migración de la interfaz gas-líquido es:

$$\frac{\partial \phi}{\partial t} + u \cdot \nabla\phi = \gamma \nabla \epsilon_{ls} \nabla\phi + \phi 1 - \phi \frac{\nabla\phi}{\nabla\phi}$$

Donde, ϕ es la línea de contorno de la interfaz de del flujo de gas-liquido, γ es el parámetro de reinicialización para resolver la ecuación, ϵ es el grosor de la interfaz que controla. Aquí, el material de bioimpresión corresponde al dominio donde $\phi < 0.5$, y el aire corresponde al dominio donde $\phi > 0.5$.

Dominios. Para cada simulación se consideraron dos dominios diferentes. El primero es el relacionado con el nozzle y la jeringa, donde se encuentra el material de bioimpresión (biotinta), el segundo es la salida del nozzle, el aire donde el nozzle eyecta el material (Figura 1). El material que se utilizó fue una tinta comercial (Cellink Bioink), compuesta por alginato con fibrillas de nanocelulosa. Este material presenta una densidad de 1000 kg/m3, y una tensión superficial calculada de 55.8 mN/m y una viscosidad no linear descrita por la ley potencial:

$$\mu = m\,\dot{\gamma}^{n-1}$$

Donde, μ es la viscosidad dinámica (Pa·s), m es el índice de consistencia del fluido, $\dot{\gamma}$ es la tensión cortante (s-1) y n el índice de comportamiento del fluido. Se utilizaron tres temperaturas distintas para cada nozzle (15, 25, y 37 °C). Como fluido no-newtoniano, la viscosidad se obtuvo por parte de Cellink y los parámetros de la ley potencial son los que se muestran en la tabla 1

	15 °C	25 °C	37 °C
$m\ (Pa \cdot s^n)$	92.735	102.530	87.906
n (adimensional)	0.146	0.170	0.208

Tabla 1. *Valores de* m *y* n *de la biotinta a 15, 25, y 37 °C*

Se colocó una entrada de flujo en la parte superior del nozzle, con un valor de 15 kPa, presión recomendada por Cellink para realizar la bioimpresión. Además, se estableció un contorno abierto en la parte superior del dominio de aire para permitir a la biotinta llenar ese dominio mientras permite al aire salir. Se estableció una condición de interfaz inicial al final del nozzle para definir el contorno donde el fluido (biotinta) termina y empieza el aire (Figura 1).

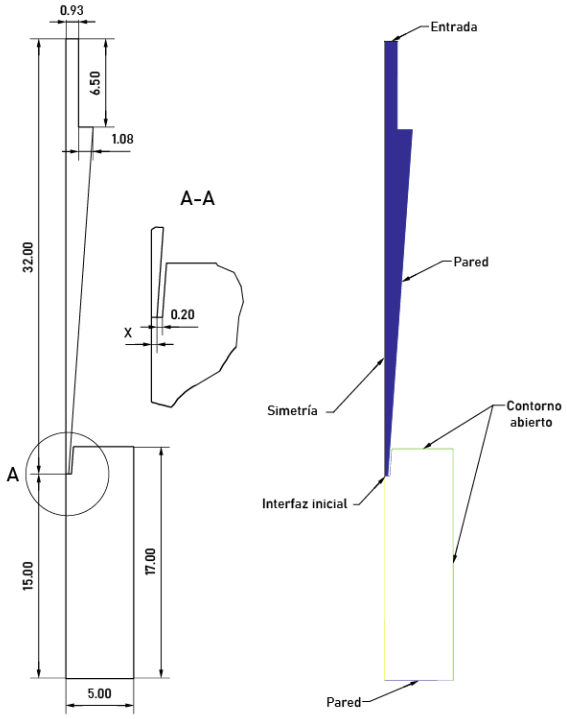


Figura 1.*Geometría y condiciones de contorno del modelo*

Simulación. Se realizaron 9 simulaciones, una para cada temperatura y nozzle, de 10 s con un paso de 1 ms. Cada simulación estaba compuesta de dos pasos de estudio: Inicialización de fase y temporal. La inicialización de fase resuelve la distancia inicial de la interfaz. Luego la temporal utiliza esta condición inicial para la función de level set de acuerdo con las siguientes expresiones:

$$\phi_0 = \frac{1}{1 + e^{D_{wi}/\epsilon}}$$

En los dominios inicialmente rellenos con material de bioimpresión, y:

$$\phi_1 = \frac{1}{1 + e^{-D_{wi}/\epsilon}}$$

En dominios rellenos inicialmente con aire.

Donde ϕ es el dominio de referencia (fracción volumétrica) y ϵ el grosor de la interfaz (m). La fracción volumétrica se definió como la cantidad de cada material durante la simulación, donde $\phi_0 = 1$ significa el 100% del hidrogel y $\phi_0 = 0$ significa 100% de aire (interfaz del material igual a 0.5)

Para calcular el grosor del filamento se usó la fracción volumétrica de una parte del dominio inicialmente relleno de aire, que posteriormente se llenaría de biotinta. Para ello, se dividió este dominio en una parte superior de 2 mm. Sobre ella se calculó mediante una sonda de integración volumétrica la fracción volumétrica en ese dominio y se

dividió por los 2 mm que tiene de alto el dominio para obtener el grosor del filamento en esa zona.

La metodología completa puede encontrarse en trabajos previos [9].

3. Resultados y discusión

3.1. Evolución del grosor del filamento

En la Figura 2 se observa el grosor del filamento obtenido y los distintos tipos de comportamiento en la evolución del grosor del filamento según la temperatura y diámetro del nozzle.

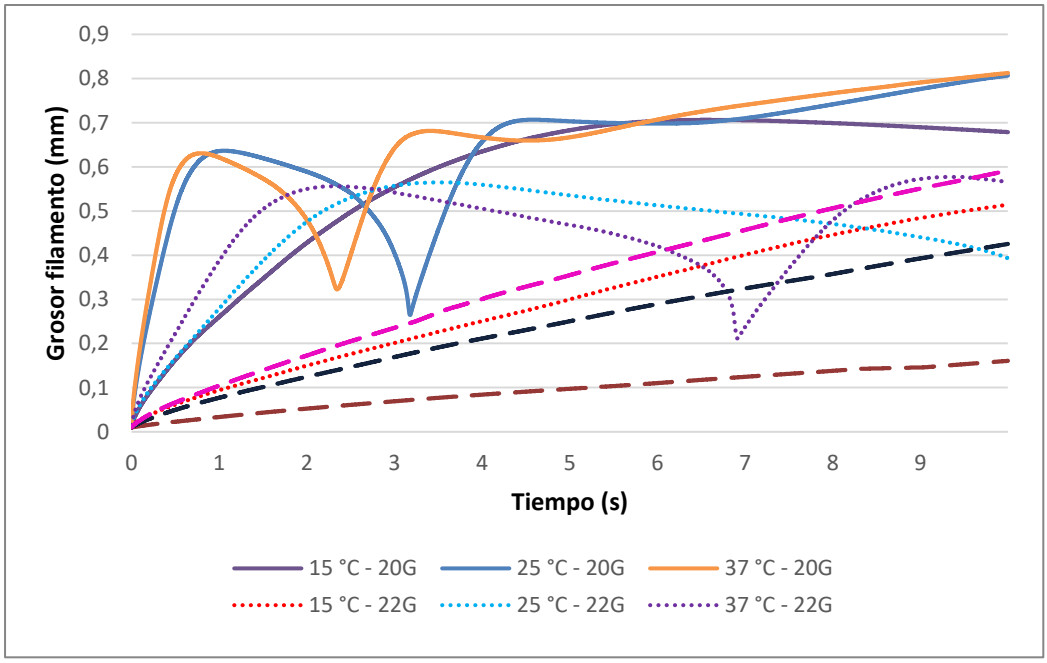


Figura 2. *Evolución del grosor del filamento durante una simulación de 10 s para tres diámetros de nozzle (20G, 22G y 25G) y tres temperaturas (15, 25 y 37 °C).*

condiciones descritas previamente, una disminución de la viabilidad celular. Esta disminución de la viabilidad es debida a que, al aumentar la presión de extrusión, la tasa de cizalladura que presenta el material en la pared del nozzle aumenta, provocando así un mayor daño en la pared celular.

En el caso en el que el material fluya con mayor facilidad, ya sea por mayor temperatura o por un mayor diámetro del nozzle (37 °C con 20 y 22G, y a 25 °C con 20G), se observa que el grosor del filamento sufre una caída, para posteriormente aumentar de nuevo. En un principio el filamento aumenta de grosor y va cayendo, hasta que, debido a su masa, se parte. Después de esa ruptura, parte del material que no ha caído vuelve hacia arriba debido a la tensión superficial y la necesidad termodinámica de mantener la forma con menos energía. Así, aumenta de nuevo el grosor, tanto por la subida de nuevo del material ya extruido, como por el material que continúa extruyéndose.

El último tipo de comportamiento que se da en el grosor del nozzle es aquel en el que alcanza un máximo y a partir de ese punto comienza a decaer. Esta forma de actuar es intermedia a lo que se producía, en los otros dos supuestos. Se da en aquellas condiciones en las que la gota alcanza toda la distancia de medición, pero no llega a romperse. Por lo que es de esperar que, pasado un tiempo mayor del simulado, el comportamiento se acerque al del segundo supuesto, forme una gota, cayendo el grosor del filamento, para posteriormente aumentar, como se comentó anteriormente.

Cuando las temperaturas son bajas y el diámetro del nozzle es reducido (25G a cualquier temperatura y 22G a 15 °C) el grosor del filamento no deja de aumentar durante los 10 s de simulación. Esto se debe a que el diámetro del calibre de 25G, 0.26 mm, dificulta el flujo adecuado de la biotinta. Con un aumento de la temperatura, en la que la viscosidad se viese reducida drásticamente, se producirían flujos controlados con calibres más reducidos. Sin embargo, un aumento por encima de los 38-40 °C vería seriamente comprometida la viabilidad celular. Por lo tanto, para provocar un flujo más rápido a lo largo de un nozzle de 27G sería necesario aumentar la presión de entrada, suponiendo, al igual que con las

La generación de un filamento estable podría conseguirse aumentando el flujo de material. Este aumento puede provocarse de distintas maneras: 1) aumentando la temperatura/reduciendo la viscosidad, 2) aumentando la presión de extrusión, 3) reduciendo el calibre del nozzle. Todos estos cambios tienen implicaciones en la bioimpresión. La mayoría generando cambios en las presiones o viscosidad, afectando directamente a la viabilidad celular.

3.2. Grosor del filamento a la salida del nozzle

La tabla 2 muestra el grosor máximo que se obtiene en cada una de las simulaciones realizadas. En las simulaciones

donde el filamento no extruye más de 2 mm no es posible realizar una medida correcta del grosor máximo. A pesar de esto, es bastante evidente la tendencia que presentan todos los grosores de los filamentos, y es que todos los filamentos obtenidos son de mayor diámetro que el calibre del nozzle.

Las simulaciones del nozzle de calibre 25G no llegan a producir un filamento lo suficientemente estable como para poder observar de manera fiable su grosor. Del mismo modo a 15 °C con el nozzle de 22G no llega tampoco a formarse un filamento estable. En aquellas simulaciones donde sí se forma un filamento estable se pude observar cómo el grosor del filamento es mayor que el del nozzle por el que sale. Concretamente en las que el nozzle es de 20G, el grosor es de un 17% más que el calibre del nozzle para 15 °C, y un 34 y 35% para 25 y 37 °C, respectivamente. Con el nozzle de 22G las diferencias se mantienen en porcentaje, siendo de un 37 y un 40% para las simulaciones de 25 y 37 °C, respectivamente. Las diferencias mostradas entre el grosor del nozzle y el grosor del filamento, siempre mayor el segundo, pueden ser debidas a la tensión superficial que presenta la biotinta. Haciendo que esta fuerza pegue el material a la pared externa del nozzle, provocando así un aumento en el grosor filamento obtenido.

Este aumento ha de tenerse siempre en cuenta a la hora de elegir tanto nozzle como los parámetros de impresión. Contar únicamente con que el filamento va a ser extruido con un grosor parejo al grosor del nozzle puede provocar que a la hora de elegir la altura de capa de la construcción subestimemos la cantidad de material que es extruido. Esto puede provocar en construcciones con exceso de material, con el coste que ello conlleva, y menos precisas, lo que reduce la personalización objetivo de las construcciones.

	Grosor (mm)		
Temp	**20 G**	**22 G**	**25 G**
15 °C	0.706	-	-
25 °C	0.806	0.564	-
37 °C	0.812	0.577	-

Tabla 2. *Máximo grosor de filamento (mm) para cada una de las temperaturas y diámetros del nozzle.*

4. Conclusión

Por lo tanto, se puede concluir que con la configuración actual y en los 10s de simulación el grosor máximo del filamento siempre es mayor al del nozzle que lo extruye. Además, a menor temperatura y menor diámetro del nozzle, se hace más difícil generar un filamento continuo y estable. Además, las condiciones de este estudio se han simplificado para permitir una aproximación inicial al comportamiento de bioimpresión por extrusión. Para futuros trabajos sería necesario realizar simulaciones de mayor periodo temporal, así como, añadir el movimiento del extrusor sobre la superficie de deposición y simular la interacción del filamento sobre esta superficie. De este modo las condiciones serían más cercanas a una bioimpresión real con el objetivo de generar un filamento continuo y estable.

Agradecimientos

Esta investigación ha sido cofinanciada por la Consejería de Economía, Ciencia y Agenda Digital de la Junta de Extremadura, proyecto número IB16200 y beca predoctoral número PD16067. Cofinanciado por Unión Europea/FEDER y FSE.

Referencias

[1] E. Mancha Sánchez *et al.*, “Hydrogels for Bioprinting: A Systematic Review of Hydrogels Synthesis, Bioprinting Parameters, and Bioprinted Structures Behavior,” *Front. Bioeng. Biotechnol.*, vol. 8, no. August, 2020.

[2] A. Blaeser, D. F. Duarte Campos, U. Puster, W. Richtering, M. M. Stevens, and H. Fischer, “Controlling Shear Stress in 3D Bioprinting is a Key Factor to Balance Printing Resolution and Stem Cell Integrity,” *Adv. Healthc. Mater.*, vol. 5, no. 3, pp. 326–333, 2016.

[3] S. Kyle, Z. M. Jessop, A. Al-Sabah, and I. S. Whitaker, “‘Printability’’ of Candidate Biomaterials for Extrusion Based 3D Printing: State-of-the-Art,’” *Adv. Healthc. Mater.*, vol. 6, no. 16, p. 1700264, Aug. 2017.

[4] M. P. Serdeczny, R. Comminal, D. B. Pedersen, and J. Spangenberg, “Experimental validation of a numerical model for the strand shape in material extrusion additive manufacturing,” *Addit. Manuf.*, vol. 24, pp. 145–153, Dec. 2018.

[5] S. Krishnamoorthy, Z. Zhang, and C. Xu, “Biofabrication of three-dimensional cellular structures based on gelatin methacrylate–alginate interpenetrating network hydrogel,” *J. Biomater. Appl.*, vol. 33, no. 8, pp. 1105–1117, 2019.

[6] J. M. Lee and W. Y. Yeong, “A preliminary model of time-pressure dispensing system for bioprinting based on printing and material parameters: This paper reports a method to predict and control the width of hydrogel filament for bioprinting applications,” *Virtual Phys. Prototyp.*, vol. 10, no. 1, pp. 3–8, Jan. 2015.

[7] S. N. Doost, D. Ghista, B. Su, L. Zhong, and Y. S. Morsi, “Heart blood flow simulation: A perspective review,” *Biomed. Eng. Online*, vol. 15, no. 1, pp. 1–28, 2016.

[8] C. Wang *et al.*, “Numerical simulation on hydraulic characteristics of nozzle in waterjet propulsion system,” *Processes*, vol. 7, no. 12, pp. 1–21, 2019.

[9] J. C. Gómez-Blanco, E. Mancha-Sánchez, A. C. Marcos, M. Matamoros, A. Díaz-Parralejo, and J. B. Pagador, “Bioink temperature influence on shear stress, pressure and velocity using computational simulation,” *Processes*, vol. 8, no. 7, Jul. 2020.

Freeform Reversible Embedding of Suspended Hydrogels for the Manufacturing of Biomedical Applications

A. Tejo-Otero[1], A. Colly[2], E.J. Courtial[2], F. Fenollosa-Artés[1,3], I. Buj-Corral[3], C.A. Marquette[2]

[1] Centre CIM, Universitat Politècnica de Catalunya (CIM UPC), Barcelona, Spain, {atejo,ffenollosa}@cimupc.org

[2] 3d.FAB, Univ Lyon, Université Lyon1, CNRS, INSA, CPE-Lyon, ICBMS, UMR 5246, 43, Bd du 11 novembre 1918, 69622, Villeurbanne cedex, France {arthur.colly, edwin.courtial, christophe.marquette}@univ-lyon1.fr

[3] TECNOFAB, Department of Mechanical Engineering, UPC, Barcelona, Spain, irene.buj@upc.edu

Abstract

Additive Manufacturing (AM) can be applied in different medical applications. Amongst all the AM technologies, DIW (Direct Ink Writing), FDM (Fused Deposition Modelling) and SLS (Selective Laser Sintering) are some of the most popular. Nevertheless, in case of biomedical applications, DIW was identified as the best option since it easily handles hydrogels and silicones, two great materials for biomedical applications. However, these materials present some difficulties in being 3D printed due to their low yield stress values which lead to the 3D printed sample to fall apart. To overcome these problems, the FRESH (Freeform Reversible Embedding of Suspended Hydrogels) technique has been developed and implemented. In this technique, a supporting gel bath is used as the support material, and removed after the material curing process. Additionally, and in particular conditions, this support bath can be reused multiple times, making the technique a cost-effective approach. In the present study, several simple and complex silicone samples were manufactured using FRESH, confirming not only the success of this AM technique for biomedical applications, but also its effectiveness.

1. Introduction

During the last years, AM has gone through an enormous transformation. It has been used not only in industry for aeronautics or automotive, but also in healthcare. The latter mentioned field has bloomed up thanks to 3D printing. For example, it has been possible to 3D print scaffolds for tissue engineering [1], manufacture surgical planning prototypes [2], prepare 4D bioprinted objects [3] and print living tissues [4].

The AM technologies which are mainly used in medical applications are DIW, FDM SLS, etc. Regarding the manufacturing of implantable devices and phantoms, DIW technique is the most popular, taking in charge soft materials like silicones and hydrogels. Nevertheless, these materials present five major challenges that need to be addressed [5]: (1) it is difficult to print soft materials that have an elastic modulus below 100 kPa using AM techniques; (2) soft materials usually require support to avoid collapsing during printing and curing; (3) the materials mechanical properties need to match the real tissue; (4) the supports need to be removed; and (5) cells need to be kept alive in case of biological purposes.

Therefore, it is necessary to find a solution to overcome these drawbacks and one viable solution was found in the FRESH technique, which is an embedded 3D printing technique (see Figure 1). It uses a thermoreversible support bath, also known as reservoir, which enables the deposition of materials for manufacturing complex structures. It is important to highlight that the support bath has a Bingham plastic behaviour during the printing [5]. This means that the support material tends to behave as a hard body under low shear stresses, while it acts like a viscous fluid at higher shear stresses [5]. In this way, the nozzle can move throughout the bath, and although there is a small mechanical resistance, the extruded material can be deposited in the appropriate place [5]. Additionally, as the nozzle moves along the supporting bath, an empty space can be created, which must be filled with a filler fluid layer [6], or by adjusting the rheology of the matrix [7].

Embedded 3D printing is an avant-garde technology, and was mentioned for the first time in 2011 [6]. Since then, different supporting gels have been developed and studied such as carbopol [8], Pluronic® F-127 [9,10] and gelatin [5,11].

Amongst the different applications, two can be highlighted which were seldom studied using FRESH technique: medical devices such as implantable epidermal devices [12] or patient-specific pulse oximeters [13], and phantoms for surgical planning [5,14].

Therefore, the aim of the present study is to show the benefits of the embedded 3D printing for different biomedical applications such as implantable medical devices and phantoms production. To do so, simple structures were first manufactured and used for validation. Then, more complex structures, i.e. a bifurcation and a kidney, were produced to evaluate and demonstrate the full capacity of the technique. For process fidelity validation purpose, the 3D printed parts were measured and 3D scanned, and the obtained surfaces compared to the original STL ("Standard Triangle Language") file.

2. Materials and Methods

2.1. Thermoreversible Support Bath

Different materials were used for creating the thermoreversible Pluronic® F-127 based support bath supplied by Sigma Aldrich, USA. The exact amount and name of the rest of materials is not mentioned due to intellectual properties.

ISBN: 978-84-09-25491-0, pags. 48- 51

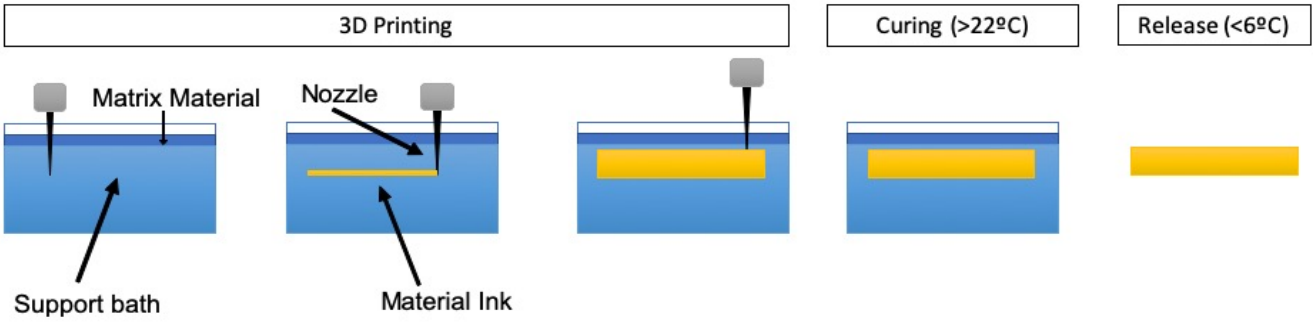


Figure 1. *FRESH 3D printing process scheme.*

2.2. Printing Ink

The printed material is a monocomponent silicone supplied by Elkem Silicones, France (AMSil™ 20101). Additionally, in some printings a blue colorant was used to enhance the visibility of the object during printing.

2.3. 3D Printing

Different 3D printed objects and samples were manufactured. A cube and a pyramid were designed using a CAD (Computer Aided Design) software. Then, a cylinder with a small inclination (10º) was obtained from [15], and modified using a specific software in order to hollow it. Regarding the bifurcation and the kidney, both were obtained from [16] and the bifurcation was modified using a specific software to hollow it out. Once STL files are prepared, a g-code generator software is needed for the 3D printing. Finally, a 3D printing sofware was used for the control of the printer and printing process. The 3D printer used was a TOBECA 333 (France).

2.4. Scanning

Once the samples were 3D printed, cured and removed from the support bath, the dimensions were measured with a calliper and analysed with an image processing software by using a reference value for the scale bar. Then, the 3D printed samples were scanned using a 3D scanner. Once the scanning was completed, a detailed quantitative analysis was carried out in which both STL (original and scanned) files were manually aligned by selecting different referential points in each mesh. Then, the distance between the meshes was computed using the cloud-to-cloud distance (Hausdorff distance algorithm). By this way, the software calculated three different parameters:

- RMS (Root Mean Square): it is defined as the square root of mean square where mean square is the arithmetic mean of the squares of numbers. If the value is close to 0 that means both surfaces are identical.
- MAD (Mean Absolute Deviation): is the average measured between each value and the mean. As before, if the value is close to 0 that means both surfaces are identical.
- MoSD (Mean of Signed Distances): is the mean of the distances between the coordinates of a point in space (scanned object mesh points) and some surface (STL file). The sign of the return value indicates if the point is inside or outside the surface. On the one hand, a negative sign indicates the point is inside the surface. On the other hand, a positive value corresponds to a point outside of the surface. If the value is a 0, the point is in the surface.

3. Results

3.1. 3D Printed Samples STL Files

The STL files used for the 3D printing are depicted in Figure 2-STL. The first three are geometrical figures using basic shapes. Then, the other two are typical biomedical cases. If a biocompatible material is used for manufacturing the bifurcation, it could be implanted in the human body. For example, in the bifurcation of the abdominal aorta into the left and right primitive iliac arteries.

3.2. FRESH 3D Printing

All the 3D printed samples using FRESH technique were successfully manufactured as it can be seen in Figure 2-FRESH 3D Printing. These good results were obtained thanks to the supporting effect of the FRESH technique, in which gravity does not affect the printed material.

However, when DIW is used, it is necessary either to look for the best position for printing or using support material because collapse might take place. The latter solution implies more material consumption and waste, unlike FRESH.

3.3. Dimensions

Regarding the dimensions of the 3D printed samples obtained using the e processing software analysis (Figure 2-Dimensions), it seems that they are quite similar to the STL files. An in-depth analysis of the error of the samples can be seen in the section 3.5. In terms of all the samples, it seems that the best one was the cylinder. This can be due to the simplicity of the shape.

3.4. Scanning

The scanning of the 3D printed parts was not achieved in the locations where the scanner was not able to reach. For example, both cylinder and bifurcation have flat surfaces in the holes due to this. The location of that surface is the maximum place the scanner could reach. Some of the scans seemed to be smoother than other due to the simplicity of the objects. For instance, the triangle and cube shapes are simple.

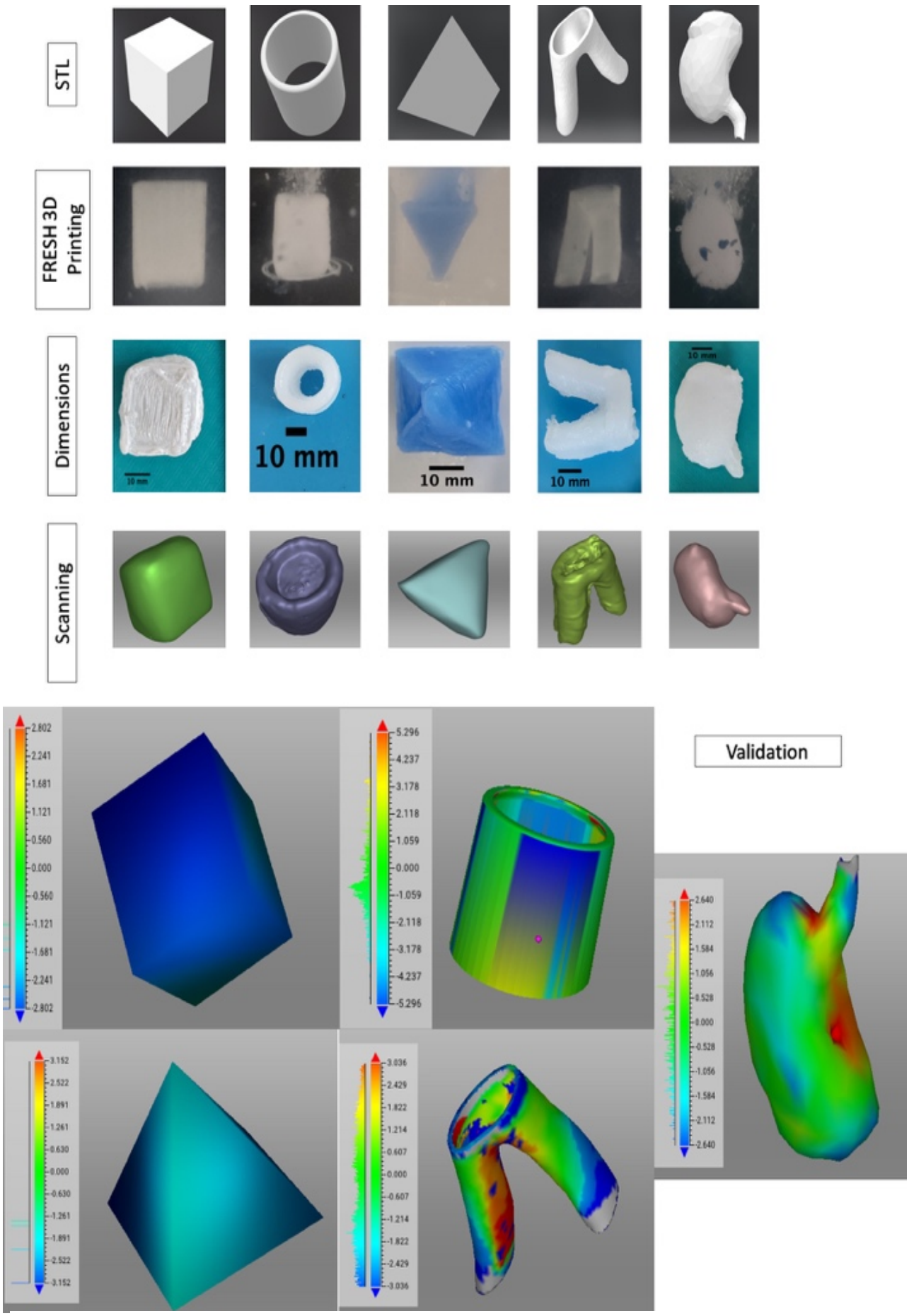


Figure 2. *FRESH 3D printing process results. The scale bar in the dimensions corresponds to 10 mm.*

3.5. Validation

As can be seen in Figure 2-Validation, all the samples seem to exhibit fidelity error with low errors, which were quantified (Table 1). Amongst the five samples, the best one was the cylinder due to the mean of signed distances close to 0. This means that most of the points of the scanned part were in the surface of the STL and, consequently, the 3D printing of the cylinder was effective.

Samples	RMS [mm]	MAD [mm]	MoSD [mm]
Cube	0.71	2.24	-2.24
Cylinder	4.06	2.25	0.04
Pyramid	2.37	3.20	-3.20
Bifurcation	3.41	2.65	-0.36
Kidney	5.47	3.69	-0.57

Table 1. *Validation values of the samples.*

Plain objects are more likely to be low fidelity since material flow is the most difficult parameter to be controlled (the worse is the cube and pyramid). Angular geometries are the most likely to be low fidelity and smooth surface as bioinspired one (tube, bifurcation, kidney) are much easier to print. Despite these minor issues, it can be concluded that the FRESH technique can successfully reproduce the morphology of the samples.

4. Discussion

This study demonstrates that FRESH technique can solve problems of the DIW like the fact that gravity makes impossible to 3D print some complex shapes, because the support bath acts as support material Additionally, the gel used can be reused for several times.

Regarding the 3D printed samples, similar shapes were 3D printed in Hinton et al. [11], but no validation was carried out, unlike in our study. In addition, our samples were cured at room temperature (22ºC) for few days or at 60ºC for 24 hours, although the UV curing is more common [11,14]. This depends on the material used. On the one hand, it is worth highlighting the possibility of using biocompatible silicones. In this way, implantable medical devices like a stent can be manufacture for the aorta or a biological part for the regeneration of a tissue.

Finally, the cost of the process FRESH technique is lower in comparison with the tissue fabrication platforms that have 3D printers that may cost up to $100.000, because it is built on an open-source hardware and software.

5. Conclusion

In conclusion, in the present study promising results were achieved and this technique is just starting to bloom up in the bioengineering field.

References

[1] Buj-Corral, I., Bagheri, A., & Petit-Rojo, O. (2018). 3D printing of porous scaffolds with controlled porosity and pore size values. Materials, 11(9), 1532.

[2] Tejo-Otero, A., Buj-Corral, I., & Fenollosa-Artés, F. (2020). 3D printing in medicine for preoperative surgical planning: a review. Annals of biomedical engineering, 48(2), 536-555.

[3] Devillard, C. D., Mandon, C. A., Lambert, S. A., Blum, L. J., & Marquette, C. A. (2018). Bioinspired Multi-Activities 4D Printing Objects: A New Approach Toward Complex Tissue Engineering. Biotechnology Journal, 13(12), 1800098.

[4] Pourchet, L. J., Thepot, A., Albouy, M., Courtial, E. J., Boher, A., Blum, L. J., & Marquette, C. A. (2017). Human skin 3D bioprinting using scaffold-free approach. Advanced Healthcare Materials, 6(4), 1601101.

[5] Hinton, T. J., Jallerat, Q., Palchesko, R. N., Park, J. H., Grodzicki, M. S., Shue, H. J., ... & Feinberg, A. W. (2015). Three-dimensional printing of complex biological structures by freeform reversible embedding of suspended hydrogels. Science advances, 1(9), e1500758.

[6] Wu, W., DeConinck, A., & Lewis, J. A. (2011). Omnidirectional printing of 3D microvascular networks. *Advanced materials*, *23*(24), H178-H183.

[7] O'Bryan, C. S., Bhattacharjee, T., Hart, S., Kabb, C. P., Schulze, K. D., Chilakala, I., ... & Angelini, T. E. (2017). Self-assembled micro-organogels for 3D printing silicone structures. Science advances, 3(5), e1602800.

[8] Bhattacharjee, T., Zehnder, S. M., Rowe, K. G., Jain, S., Nixon, R. M., Sawyer, W. G., & Angelini, T. E. (2015). Writing in the granular gel medium. Science advances, 1(8), e1500655.

[9] Muth, J. T., Vogt, D. M., Truby, R. L., Mengüç, Y., Kolesky, D. B., Wood, R. J., & Lewis, J. A. (2014). Embedded 3D printing of strain sensors within highly stretchable elastomers. Advanced materials, 26(36), 6307-6312.

[10] Liu, W., Heinrich, M. A., Zhou, Y., Akpek, A., Hu, N., Liu, X., ... & Zhang, Y. S. (2017). Extrusion bioprinting of shear-thinning gelatin methacryloyl bioinks. Advanced healthcare materials, 6(12), 1601451.

[11] Hinton, T. J., Hudson, A., Pusch, K., Lee, A., & Feinberg, A. W. (2016). 3D printing PDMS elastomer in a hydrophilic support bath via freeform reversible embedding. ACS biomaterials science & engineering, 2(10), 1781-1786.

[12] Tan, W. S., Juhari, M. A. B., Shi, Q., Chen, S., Campolo, D., & Song, J. (2020). Development of a new additive manufacturing platform for direct freeform 3D printing of intrinsically curved flexible membranes. Additive Manufacturing, 101563.

[13] Abdollahi, S., Markvicka, E. J., Majidi, C., & Feinberg, A. W. (2020). 3D Printing Silicone Elastomer for Patient-Specific Wearable Pulse Oximeter. Advanced healthcare materials, 9(15), 1901735.

[14] Zhao, J., Hussain, M., Wang, M., Li, Z., & He, N. (2020). Embedded 3D printing of multi-internal surfaces of hydrogels. Additive Manufacturing, 32, 101097.

[15] Cilindro de 20 x 20 x 20 mm by Inven Technologies 3D https://www.thingiverse.com/thing:2513108 (Access: 03/09/2020)

[16] Heart, Aorta and Kidney CAD Model by chris@printhuman.org https://3dprint.nih.gov/discover/3dpx-000906 (Access: 03/09/2020)

Pósteres

Miércoles, 25 de noviembre de 2020

Optical coherence tomography imaging of colorectal neoplastic polyps developed in genetically modified rats

C. Reyes[1], B. Corbett[1], J.F. Ortega Morán[2], J. Bote[2], B. Roycroft[1]

[1] Tyndall National Institute, Cork, Ireland, {carlos.reyes,brian.corbett,brendan.roycroft}@tyndall.ie

[2] Jesús Usón Minimally Invasive Surgery Centre, Cáceres, Spain, {jfortega,jbote}@ccmijesususon.com

Summary

The high incidence and mortality of colorectal cancer requires the need to develop new technologies for the early detection and diagnosis of disease in early stages. In this way, cure rates of patients could be increased. Despite the imaging techniques used today for the detection and diagnosis of colorectal cancer, there is still an unmet need for a technique that could provide images of tissue below the visible surface with a higher resolution, and with the possibility of distinguish adenomatous polyps and allow rapid diagnosis in situ previous to the histopathological examination. One of these emerging techniques with great potential is Optical Coherence Tomography (OCT). In this work, a developed neoplastic polyp in the colon of a genetically modified PIRC F344/NTac-Apcam rat has been imaged using OCT in order to observe its subsurface structure. The polyp was imaged with a Thorlabs commercial OCT system at 1300 nm obtaining an axial and lateral resolution of 3 μm and 4 μm respectively. Results show that the neoplastic polyp demonstrated irregular glands with enlarged nuclei in the epithelial layer showing different morphology that affect the underlying homogenous layer structure of healthy colon tissue. 3D rendering of the sample provided by the OCT system allows obtaining an "optical biopsy" for better analysis of the surface and subsurface structure of the colonic polyp. The inclusion of a scanning OCT probe through an endoscopic working channel could be used to perform in-vivo imaging of polyps during colonoscopy as a support for an improved diagnosis.

1. Introduction

Colorectal cancer (CRC) is the third most common cause of cancer death, and represents about 10% of all cancer cases [1-2]. The majority of CRC develops from adenomatous polyps, which are premalignant lesions that develop over a period of years. These polyps are identified using colonoscopy, the gold standard method for detection of polyps. It is an endoscopic examination that provides direct visualization of the colonic mucosa using white-light and narrow-band imaging (NBI) based on light absorption of hemoglobin [3]. Detected polyps are removed by endoscopic polypectomy for further histological analysis to determine the classification of polyps and depth of invasion. However, such a process is invasive and the detection of small polyps (< 5 mm) has a miss rate between 15% and 26% [4-5].

In order to complement the information provided by colonoscopy, optical coherence tomography (OCT) has been progressing towards an 'optical biopsy' of lesions as it provides subsurface polyp image information regarding thickness of mucosal layers and its optical properties [6-7]. With the purpose of developing this research field, ex-vivo colon OCT imaging studies have been performed to determine parameters that differentiate between benign and adenomatous polyps [8-9].

In this sense, the PICCOLO project [10] aims to design and develop a new compact, hybrid and multimodal photonics endoscope probe based on OCT and Multi-Photon Tomography (MPT) for improving in vivo diagnosis and clinical decision support. By combining the outstanding structural information from OCT with the precise functional information from MPT, this innovative endoscope would provide gastroenterologists immediate and detailed in situ identification of colorectal neoplastic lesions and facilitate accurate and reliable in vivo diagnostics, with additional, grading capabilities for colon cancer as well as in-situ lesion infiltration and margin assessment. For that, human representative animal models are used to generate imaging biomarkers that allow automated detection, assessment and grading of disease.

Although the above-mentioned PICCOLO project aims to use jointly MPT and OCT photonics technologies, only OCT technology is addressed in this work. Concretely, a first study has been carried out where neoplastic polyps developed in genetically modified rats have been imaged with the purpose of observing its development as well as proposing the use of OCT as a complementary imaging modality for the analysis of ex-vivo colon samples.

2. Methods

2.1. Optical coherence tomography

OCT is a non-invasive, high-resolution imaging technology capable of providing cross-sectional and three-dimensional images to visualize the structure of a biological sample [7]. OCT systems are based on partially coherent light interferometry (PCLI) by measuring the magnitude and echo time delay of backscattered light from a sample using a broadband light source with a short coherence length. The time delay represents the longitudinal location of reflection points. The axial and lateral resolution define image quality and for OCT imaging both resolutions are independent from each other.

ISBN: 978-84-09-25491-0, pags. 53- 56

The basic setup for an OCT is shown in Figure 1. It basically consists of: 1) a light source at a center wavelength λc, 2) a wavelength range δλ, 3) two optical arms created by the beam splitter including the reference and sample. The light travels through both arms creating an interference pattern if the optical distance difference between both arms is within the coherence length of the light source. The interference pattern is then detected with a photodetector and finally digitized with an acquisition card.

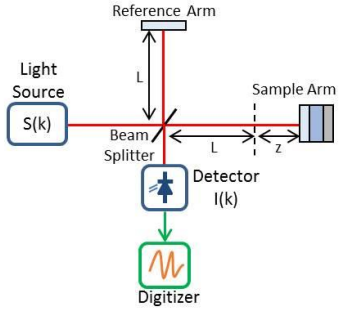


Figure 1. *OCT basic setup*

The samples imaged in this article were imaged using a commercial Thorlabs OCT system Ganymede GAN620 with an OCT-LK2-BB lens at a λc = 900 nm, with an axial resolution of 3 μm and lateral resolution of 4 μm.

2.2. Rodent models

Rodent models are used to simulate disease progression for the validation of optical imaging tools due to their accurate modeling of human pathology. In this case in particular, the large intestine of rodents and humans are similar in mural structure [11]. For this paper, ex-vivo tissue samples were sections of the colon of "polyposis in rat colon" (PIRC F344/NTac-Apcam [12-13]) rats that developed neoplastic polyps. Ethical approval for the samples was provided by the CCMIJU Animal Welfare Committee and by the Regional Government of Extremadura.

For the ex-vivo sample preparation, the colon from the rat was excised and a protocol for sample preservation was followed. First, the rat colon was divided in three regions following the ascending, transverse and descending anatomical structures of the colon. The regions were then immersed in 4% formaldehyde overnight at a temperature of 4°C for 14 hours. Finally, the samples were submerged in phosphate buffered saline (PBS) and 0.1% of sodium azide after three washes with PBS each 30 minutes after at least 14 hours and kept in refrigeration at 4°C.

OCT is used to obtain cross-sectional images to measure mucosal thickness at the site of adenoma (2x normal mucosal thickness) and a mild protrusion above the basal mucosal height (<50% total thickness). Signal attenuation with depth has also been observed, and the underlying tissue boundaries are fainter as compared to healthy colon tissue [14].

3. Results

A neoplastic polyp (Figure 2A) was imaged across the indicated arrow in order to differentiate the subsurface structure of healthy against the polyp regions. Figure 2B,C shows four regions of the sample with different optical and morphological properties. These include regions where the polyp has affected the homogenous structure and regions that were not affected.

Regions 1 and 4 are normal tissue next to the developed polyp. Mucosal and submucosal layers can be clearly identified where the muscularis mucosa appears as a scattering band separating each layer for both regions. Both layers show a homogeneous structure, however in the orthogonal direction in region 1 polyp development has affected the substructure. The layering is not as homogeneous as well as presenting areas with either higher or lower backscattered light, hence highlighting the relevance of analyzing different transversal cuts when classifying a polyp. Region 4 presents the most homogenous layering in both directions as it has not been fully invaded by the polyp growth.

Regions 2 and 3 show the developed neoplastic polyp whose total length across the horizontal axis is ≈ 0.4 mm. Compared to regions 1 and 4, the subsurface structure homogeneous layering has been lost, and instead there is a heterogeneous structure distribution of the crypt pattern. The polyp has developed invasive irregular glands with enlarged nuclei in the epithelial layer that have different morphology as well as different optical scattering properties compared to the homogeneous mucosal and submucosal layering.

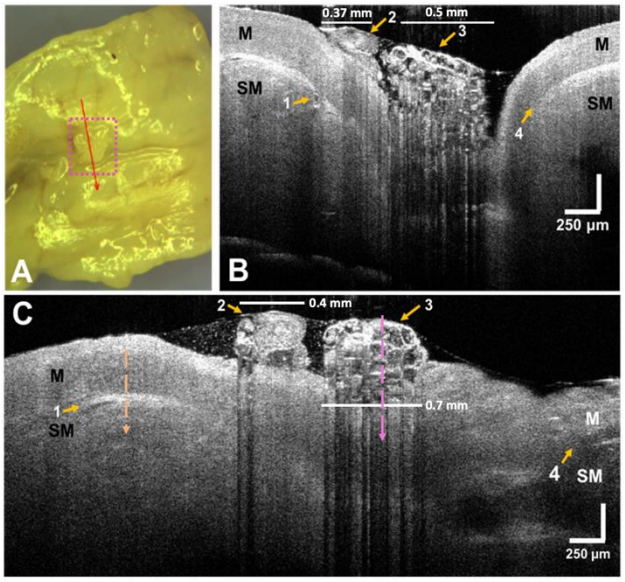


Figure 2*. Neoplastic polyp imaging. A) Neoplastic colon sample. B) Orthogonal B-scan of the polyp. C) B-scan across the red arrow. Four regions are identified: (1) Homogenous subsurface structure layering has been affected. (2-3) Large polyp showing a heterogeneous substructure presenting glands of different shapes and different scattering properties. (4) Healthy tissue with the most homogenous subsurface structure showing a clear layering. M: Mucosa; SM: Sub-mucosa*

A B-scan is obtained by stacking adjacent A-scans next to each other. Two examples of an A-scan are taken in the orange and pink lines from Figure 2C. With the A-scan it is possible to measure the thickness of the different layers as well as the microstructures. Figure 3A (orange) shows the A-scan of healthy tissue with two layers of 75 and 60 µm separated by a distance of 125 µm. Compared to the healthy tissue, the polyp structure (Figure 3B) presents irregular glands with enlarged nuclei in the epithelial layer showing different morphology with thicknesses that vary between 13 and 45 µm. This feature makes OCT different from colonoscopy and it is comparable to histology in terms of axial resolution without the need of slicing the tissue.

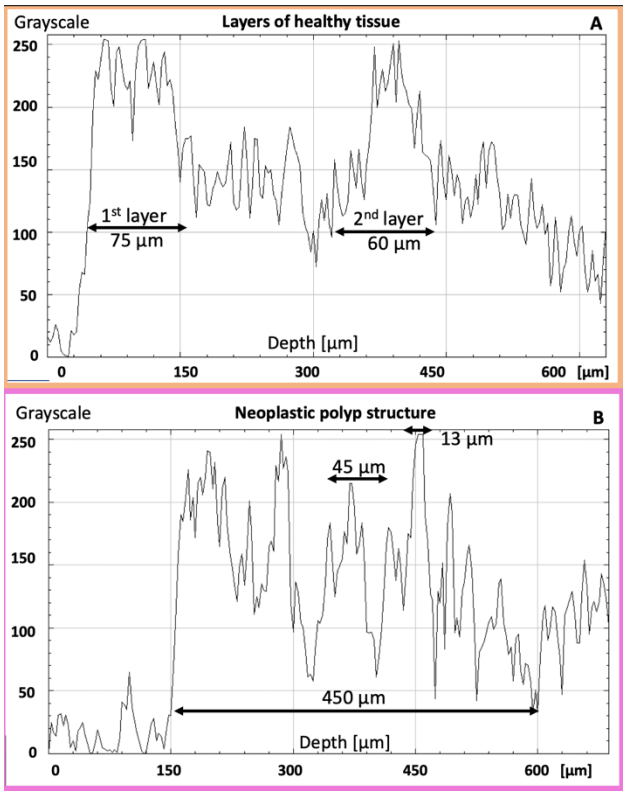


Figure 3. *A-scans A) Two layers identified in healthy tissue. B) Several micro features are observed in the neoplastic polyp*

With the obtained cross-sectional images it is possible to render a volume of the sample by stacking adjacent B-scans. With this, it is possible to retrieve 'virtual cuts' or 'optical biopsies' where the sample can be sliced at a specific depth in order to observe its internal layering and homogeneity. This term is a C-scan. Two examples are shown in Figure 4A,B including the surface of the polyp and an internal cut which is equivalent to an 'optical biopsy' as the cut is made in depth of the sample.

An interesting aspect of OCT imaging compared to other imaging modalities can be seen when comparing region 1/4. When observed at the top region 1 shows a clear structure with 'stripes' as well as a homogeneous structure, however for an internal slice it can be seen that their homogeneity is different since region 1 has been affected by the polyp growth.

Regions 2 and 3 show an irregular crypt structure in terms of its shape and heterogenous distribution on the surface, compared to the 'stripes' of healthy tissue. The subsurface rendering of the polyp shows regions with clear high backscattered properties, as well as inhomogeneous distributions, with glands of different shapes and sizes.

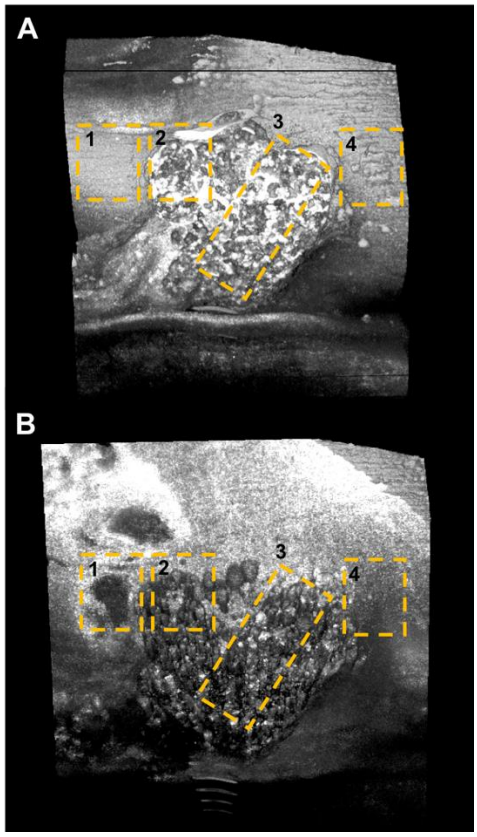


Figure 4. *OCT C-scan 'optical biopsy' A) Reconstruction of the surface of the sample. B) Sub-surface C-scan showing internal structure distribution. The same regions as in Figure 2 are identified.*

4. Discussion and conclusion

The proposed method for sample preparation by extracting the rat colon and then dividing it into three regions with healthy and polyp tissue across the transversal plane provided clear imaging regions in order to differentiate the sub-surface morphology and observe how a developed polyp affects normal tissue.

The developed neoplastic polyp has caused an abnormal growth of healthy tissue in terms of a heterogeneous structure with irregular glands of different sizes, altered backscattering properties and an epithelium that no longer has a symmetrical curvature. 3D rendering of the sample provides a better picture of the surface and subsurface structure of the sample, as well as comparing internal cuts that have shown heterogeneity of neoplastic crypt structures.

OCT is an imaging modality with a great potential that provides subsurface image data allowing ex-vivo polyp samples studies in order to observe the abnormal growth and compare it against healthy tissue. Compared to other diagnostic methods, OCT could be considered as a generator of a "virtual histology", since the use of

exogenous fluorophores or other markers is not necessary to support the obtaining of a diagnosis, providing highly useful images of tissue penetration and spatial resolution [15]. In contrast to other optical image modalities, OCT is able to image non-invasively and without the need for tissue preparation [16]. Surface imaging modalities, such as chromoendoscopy, virtual chromoendoscopy, magnification endoscopy and endocytoscopy, or autofluorescence imaging, are limited to superficial lesions and improving magnification would mean restricting the field of view, making thus more difficult the screening and surveillance for multi-focal lesions. And within the subsurface imaging modalities, such as endoscopic ultrasonography, confocal endomicroscopy or OCT, the latter combines nearly microscopic resolution with volumetric and subsurface real-time imaging capabilities [17]. Further analysis to a broader sample of polyps should be performed as future works to strengthen the conclusions drawn in this paper.

5. Acknowledgment

This study is funded by EU H2020 project PICCOLO, grant agreement 732111: 'Multimodal highly-sensitive PhotonICs endoscope for improved in-vivo COLOn Cancer diagnosis and clinical decision support' (piccolo-project.eu) and Irish Photonic Integration Centre (12/RC/2276_P2).

This information reflects only the institution's view, so the Agencies and the Commission are not responsible for any use that may be made of the information it contains.

References

[1] R. Siegel, K. Miller, and A. Jemal, "Cancer statistics, 2019," Cancer Journal, Am. Cancer Soc. **69**, 7–34 (2019).

[2] R. Siegel, K. Miller, A. Goding, S. Fedewa, L. Butterfly, J. Anderson, A. Cercek, R. Smith, and A. Jemal, "Colorectal cancer statistics, 2020," Cancer Journal, Am. Cancer Soc. 0, 1-20 (2020).

[3] R. Ahmed, K. Santhirakumar, H. Butt, and A. Yetisen, "Colonoscopy technologies for diagnostics and drug delivery," Med. devices sensors, Wiley 2, No. 3, 1-16 (2019)

[4] S. Bong, D. Soo, J. Ho, T. Jun, J. Pyo, and C. Soo, "The miss rate for colorectal adenoma determined by quality-adjusted, back to back colonoscopies" Gut Liver 6, No 1, 64-70 (2012)

[5] J. van Rijn, J. Reitsma, J. Stoker, P. Bossuyt, S. van Deventer, and E. Dekker, "Polyp miss rate determined by tandem colonoscopy: a systematic review," Am. Journal gastroenterology 101, No. 2, 343-350 (2006).

[6] J. Fujimoto, and W. Drexler, *Introduction to OCT* in Optical Coherence Tomography technology and applications (Springer reference, 2015 2nd edition).

[7] W. Drexler, M. Liu, A. Kumar, T. Kamali, A. Unterhuber, and R. Leitgeb, "Optical coherence tomography today: speed, contrast, and multimodality," Biomed. Opt. 19, No. 7, 071412:1-34 (2014).

[8] C. Wang, Q. Zhang, X. Wu, T. Tang, H. Liu, S. Zhu, B. Gao, and X. Yuan, "Quantitative diagnosis of colorectal polyps by spectral domain optical coherence tomography," Biomed research international pp. 1-7 (2014)

[9] X. Yu, Y. Luo, X. Liu, S. Chen, X. Wang, S. Chen, and L. Liu, "Toward high-speed imaging of cellular structures in rat colon using micro-optical coherence tomography," IEEE photonics journal 8, No. 4, 3900308 (2016).

[10] PICCOLO project. www.piccolo-project.eu/. (Accessed: September 2020).

[11] P. Treuting, S. Dintzis, and K. Montine, Comparative anatomy and histology a mouse, rat and human atlas (Academic press, Washington, 2018 (2nd edition))

[12] A. Irving, K. Yoshimi, M. Hart, T. Parker, L. Clipson, M. Ford, T. Kuramoto, W. Dove, and J. Amos-Landgraf, "The utility of apc-mutant rats in modeling human colon cancer," Dis. Models mechanisms 7, 1215-25 (2014)

[13] S. Hansen, M. Hart, S. Busi, T. Parker, A. Goerndt, K. Jones, J. Amos-Landgraf, and E. Bryda, "Fischer-344 tp53-knockout rats exhibit a high rate of bone and brain neoplasia with frequent metastasis," Dis. models mechanisms **9**, 1139–46 (2016).

[14] L. Hariri, Z. Qiu, A. Tumlinson, D. Besselsen, E. Gerner, N. Ignatenko, B. Povazay, B. Hermann, H. Sattman, J. McNally, A. Unterhuber, W. Drexler, and J. Barton, "Serial endoscopy in azoxymethane treated mice using ultra-high resolution optical coherence tomography," Proc. SPIE 6, Issue 11, 1-10 (2007).

[15] B. Glover, J. Teare, and N. Patel, "The Status of Advanced Imaging Techniques for Optical Biopsy of Colonic Polyps," Clinical and Translational Gastroenterology, 11(3) (2020).

[16] L. van Manen, J. Dijkstra, C. Boccara, E. Benoit, A. L. Vahrmeijer, M. J. Gora, and J. S. D. Mieog, "The clinical usefulness of optical coherence tomography during cancer interventions," *Journal of cancer research and clinical oncology*, *144*(10), 1967-1990 (2018).

[17] N. S. Samel, and H. Mashimo, "Application of OCT in the gastrointestinal tract." *Applied Sciences*, *9*(15), 2991 (2019).

Método de estadiaje del sueño basado en información electrocardiográfica y electromiográfica

G.M. Besné [1,2], J. Artieda[1,2], M. Alegre[1,2] M. Valencia[1,2]

[1]IdiSNA, Instituto de Investigación Sanitaria de Navarra, Pamplona, España;

[2]Universidad de Navarra, CIMA, Pamplona, España

gbesne@alumni.unav.es, mvustarroz@unav.es

Resumen

Esta comunicación presenta un método automático para realizar el estadiaje de sueño utilizando única y exclusivamente datos de ritmo cardíaco y actimetría. El método ha sido desarrollado utilizando la base de datos de registros polisomnográficos (PSG) DREAMS e incluye herramientas para la detección de la onda R del electrocardiograma (ECG) en registros provenientes de la rutina clínica. Una vez detectados los instantes de tiempo en los que aparecen las ondas R, se utilizan para desarrollar un sistema personalizado de estadiaje de sueño automático basado en métodos de Machine Learning (ML). A continuación, se estudia si la inclusión de parámetros extraídos de la actimetría (concebida como actividad motora a partir de las señales de electromiografía (EMG)) mejora el rendimiento del clasificador. Los resultados demuestran que es posible predecir las fases del sueño a partir de la señal de ECG con una exactitud media del 96.04%. La inclusión de la actimetría obtenida a partir del EMG mejora de forma estadísticamente significativa ($0.0083 < p < 0.00002$) la capacidad del clasificador sin necesidad de utilizar señal electroencefalográfica (EEG) llegando al 97.91% en exactitud media para todos los registros de la base de datos DREAMS.

1. Introducción

La posibilidad de llevar a cabo un estudio y estiadiaje del sueño en el ámbito ambulatorio sin necesidad de utilizar registros de EEG abre la posibilidad de ampliar el seguimiento y tratamiento de pacientes con trastornos del sueño y/u otras patologías. Si bien algunas soluciones de electrónica de consumo ofrecen estimaciones de la calidad del sueño en base a medidas de actimetría y fotopletismografía (PPG), la realidad es que muchos especialistas del sueño no confían en estos sistemas y ponen en entredicho la capacidad real de estos dispositivos para proveer de información útil para la práctica clínica. En este trabajo presentamos una metodología inspirada en esta aproximación aplicada sobre la base de datos de PSG DREAMS [1] y se demuestra cómo es posible predecir las fases de sueño con un alto grado de precisión.

El trabajo se divide en dos partes; detección del latido cardiaco en ECG rutinario (proveniente de un registro de PSG) mediante la localización del ápice de la onda R en el complejo QRS, y una segunda parte dedica a la obtención del modelo personalizado para el estadiaje automático de las fases de sueño en cada paciente.

La primera parte responde a la necesidad de detectar las ondas R en registros provenientes de la rutina clínica y que no han sido específicamente diseñados para la captura de la señal de ECG, con lo que los artefactos o incluso las inversiones del complejo QRS resultan frecuentes.

2. Material y métodos

La base de datos publica DREAMS consta de 20 PSGs realizados en sujetos sanos. Incluye registros de ECG, EEG, EMG, pulsioximetría, posición del paciente y parámetros respiratorios entre otros. Dichas señales están muestreadas a 200Hz con una duración mínima de 7h. A sí mismo, incluye la anotación de las fases del sueño según los protocolos R&K y AASM [2]. Sin embargo, no incluye marcadores de la presencia de latidos cardiacos

2.1. Localización del latido cardiaco

Para la localización del latido cardiacos, se adaptó e integró un detector de la presencia de la onda R en registros no óptimos de ECG (como los que posiblemente se pueden hallar en PSGs). El algoritmo realiza una primera detección de posibles candidatos a onda R y un segundo análisis para la confirmación y localización precisa del ápice de dicha onda.

Tras la importación de los archivos a Matlab, se eliminan las componentes lentas con un filtrado paso-alta de 0.5 Hz. Basándose en [3], se aplica el *Empirical Mode Descomposition* (EMD) en ventanas que incluyen un total de tres complejos QRS solapándose 1.5 segundos dando lugar a la detección de posibles puntos en los que se encuentra una onda R.

Sobre estos candidatos se realiza un análisis de correlación cruzada en ventanas de la misma amplitud que el análisis anterior. Dicha correlación se realiza entre el tramo de interés y el promedio de las tres ondas R detectadas anteriormente ausentes de ruido. Este último análisis proporciona una estimación más certera de la posición de la onda R de interés, así como un segundo candidato para la estimación de la próxima onda R a integrar con la obtenida por el EMD.

ISBN: 978-84-09-25491-0, pags. 57- 60

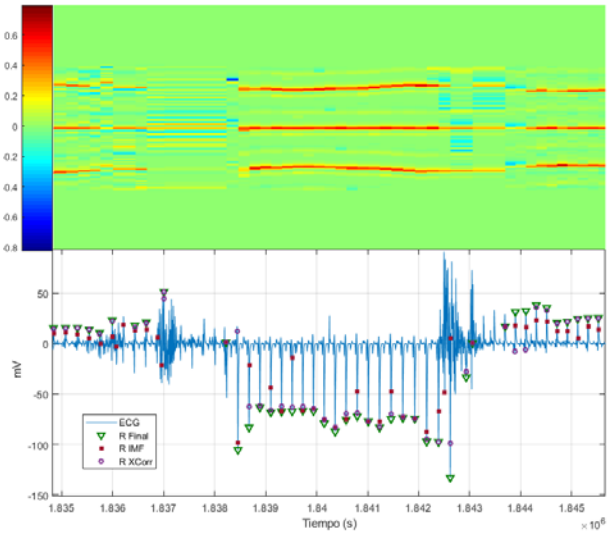


Figura 1 *Superior) imagen formada con la superposición de las correlaciones cruzadas según el instante de tiempo en que fueron tomadas, siendo la escala de color la intensidad de cada punto. Inferior) señal original indicando las distintas estimaciones obtenidas a lo largo del algoritmo*

2.2. Características del HRV

Con las posiciones de las ondas R determinadas, se procede al cálculo de las características relacionadas con la señal de ECG. Para ello se calcula el intervalo RR (iRR) como el tiempo que sucede entre consecutivos latidos. A partir de estos valores, se calculan un total de 29 características repartidas en 4 grupos [4-8]:

1. Dominio temporal
   a. Número de latidos
   b. Media, mediana, moda y desviación estándar del iRR
   c. Moda de iRR consecutivos
   d. Porcentaje de iRR consecutivos de más de 25, 50 y 100 ms (pNN25/50/100)
   e. Cuadrado de la raíz media de la unión de los iRR adyacentes (RMSSD)
   f. Variabilidad de la frecuencia cardiaca (SDNN)
   g. Desviación estándar de consecutivos iRR (SDSD)
   h. Covariancia de la media del iRR
2. Dominio frecuencial, donde se toman 2 bandas de frecuencias: altas frecuencias (Hf) de 0.5 a 0.15 Hz y baja frecuencia (Lf) de 0.04 a 0.15 Hz.
   a. Potencia media total (de 0.04 a 0.5Hz)
   b. Potencia media normalizada en Hf y Lf
   c. Ratio de potencias Hf/Lf
   d. Potencia normalizada máxima en Hf y Lf
   e. Frecuencia a potencia máxima en Hf y Lf
3. Métodos no lineales basados en la representación Pointcaré de las muestras de iRR
   a. Eje transversal (SD1), eje longitudinal (SD2) y ratio entre SD1 y SD2
   b. Área de la elipse
   c. Alfa corto y largo
4. Entropía
   a. Entropía de las muestras de iRR

2.3. Características de la actimetría por EMG

Para las características derivadas del EMG se han extraído un total de 8 características. La primera característica representa el tono del segmento como la media de la envolvente de la señal de EMG [9]. Las 7 características restantes se relacionan con picos de actividad muscular en el segmento de señal: número de picos, media de la altura y anchura de los picos, máximo de la altura y anchura de los picos, y energía media y máxima de los picos.

2.4. Implementación del modelo

El hipnograma R&K se ha simplificado evaluándose tres niveles de detalle diferentes, resultando en hipnogramas con los siguientes estados/fases a clasificar [10]:

a. Hipnograma de 4 estados (4S): Despierto (W). sueño profundo (DS), sueño ligero (LS) y sueño REM (REM).
b. Hipnograma de 3 (3S) estados: Despierto (W). sueño no REM (nREM) y sueño REM (REM).
c. Hipnograma de 2 estados (2S): Despierto (W) y sueño (S).

Con el fin de evaluar si la inclusión de la actimetría partir del EMG incrementa las prestaciones del clasificador, se generan dos conjuntos de características: uno únicamente con características derivadas del HRV y otro con características a partir del HRV y la actimetría por EMG. En ambos casos se hará uso de todas las características explicadas anteriormente: 29 para el HRV y 8 para la actimetría.

Además, se han evaluado ventanas temporales de 3, 4 y 5 minutos con solapamientos de 1, 2 y 3 minutos respectivamente, para el cálculo de las características haciendo que la clasificación efectiva se de cada 2 minutos para estos 3 subgrupos de características.

Teniendo en cuenta lo anteriormente descrito se obtiene un total de 18 sets de características distintas con las que entrenar distintos modelos para cada paciente. Finalmente se evaluará el rendimiento de los 9 modelos con características del HRV frente al rendimiento de los 9 modelos con características del HRV y en EMG combinadas.

Para la selección del algoritmo para la generación del modelo a implementar, se ha utilizado la herramienta *Clasification Learner* de Matlab, que permite realizar el testeo simultáneo de diferentes algoritmos para el mismo set de datos y categorías. Para esta prueba se han extraído las características de 3 sujetos considerando ventanas de 3 minutos con solapamiento de 1 minuto y el hipnograma de 4 estados del sueño, resultando en 3 sets de datos diferentes.

Al introducir dichas características en la herramienta, se obtienen valores de precisión para cada uno de los modelos implementados. Finalmente seleccionamos el modelo que mejor predice las fases de sueño para los 3 sujetos. Una vez seleccionado se implementa en un sistema automático para la generación de los modelos en los 20 sujetos que comprenden la base de datos.

Como se plantea este sistema como una solución personalizada para cada paciente se ha optado por generar modelos únicos utilizando un 20% de cada estado del sueño para el entrenamiento y un 80% para el testeo por medio de una validación cruzada de 5 veces. Este procedimiento genera un total de 5 modelos entrenados cada uno con el 20% de los datos seleccionados aleatoriamente. Por último, este entrenamiento será repetido 5 veces, resultando en un total de 25 modelos. De ellos aquel que mejor realice la clasificación en el 80% restante de los datos según la pi de Scott será seleccionado como el modelo representativo para este sujeto y ese set de características concreto.

Como resultado final, por cada paciente se obtiene un total de 18 modelos entrenados con distintos sets de datos en función del hipnograma, las ventanas de tiempo y el tipo de características.

3. Resultados

La técnica combinada de EMD y correlación cruzada para la detección de latidos cardiaco permite detectar con exactitud la presencia de los latidos a pesar de posibles interferencias como artefactos por movimiento o inversiones de polaridad. Tras la revisión de las anotaciones se ha obtenido una precisión media del 91.85%.

En la Figura 2 se muestran los resultados de la exactitud y la pi de Scott de los distintos modelos que se han generado para la clasificación a partir de datos de ECG. Para estos sets de datos se ha utilizado un SVM polimnomial de segundo orden. Como puede apreciarse, el rendimiento de los modelos decrece a medida que aumenta el número de estados de sueño considerados y se reduce el tamaño de la ventana para calcular las características.

En los 3 grupos de modelos para distinto número de estados del sueño, la aproximación de ventanas de 5 minutos solapadas 3 minutos muestra la mejor clasificación obteniendo a) para 2 estados de sueño una pi de Scott media de 79.38% y una exactitud del 96.04%, b) para 3 estados del sueño una pi de Scott media de 74.60% y una exactitud del 90.01% y c) para 4 estados del sueño una pi de Scott media de 68.39% y una exactitud del 83.20%.

La Figura 3 muestra el rendimiento de los modelos al entrenarse con las características del HRV y las de la actimetría por EMG. Para estos sets de datos se ha utilizado un SVM polimnomial de tercer orden. En esta aproximación se sigue apreciando la caída del rendimiento de la clasificación conforme aumenta el número de estados y disminuye la ventana. Sin embargo, los resultados son mejores detectándose diferencias estadísticamente significativas en el rango $0.0083 < p < 0.00002$ al comprar ambas modalidades: a) para 2 estados de sueño una pi de Scott media de 92.32% y una exactitud del 97.91%, b) para 3 estados del sueño una pi de Scott media de 86.38% y una exactitud del 94.01% y c) para 4 estados del sueño una pi de Scott media de 81.49% y una exactitud del 89.29%.

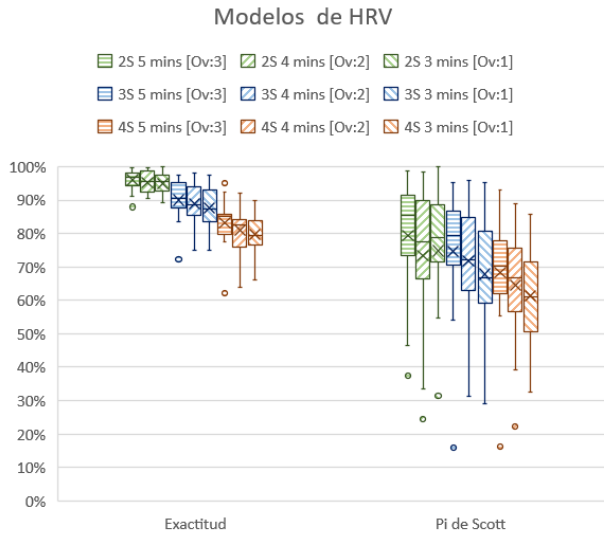


Figura 2 *Representación de boxplot de la exactitud y la pi de Scott para cada una de las nueve condiciones para los modelos considerando características del HRV. En la leyenda, 2S 5mins [0v:3], corresponde a: clasificación de 2 estados de sueño (2s) en ventanas de 5 minutos (5mins) con un solapamiento de 3 minutos ([Ov:3]).*

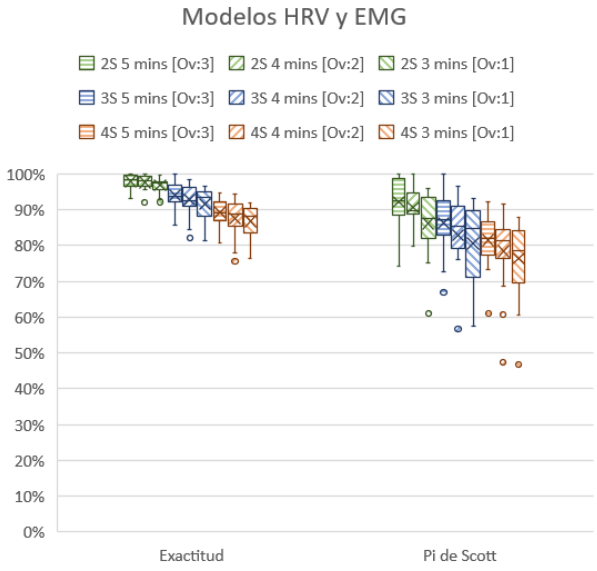


Figura 3 *Representación de boxplot de la exactitud y la pi de Scott para cada una de las nueve condiciones para los modelos considerando características del HRV y de la actimetría por EMG.*

4. Discusión

La metodología de entrenamiento que se ha propuesto ha sido capaz de discernir entre sueño ligero, sueño profundo y sueño REM con una precisión adecuada si se le presentan únicamente características derivadas del HRV. Sin embargo, este rendimiento se ve superado con creces al añadirle las características relacionadas con la actimetría basada en EMG. Este incremento se ve sobre todo al considerar la pi de Scott, alcanzando incrementos de hasta un 13% para 4 estados del sueño, mientras que la exactitud incrementa un 6%. Con la configuración adecuada, Se alcanzan valores de rendimiento con una pi de Scott de hasta 92.31% y exactitudes de 94.71% según el sujeto.

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 y la Asociación de Amigos de la Universidad de Navarra por financiar mis estudios doctorales.

Referencias

[1] Kerhofs M, Devuyts S, Dutoit T, Dreams subjects database, University of MONS – TCTS Laboratory and Universit Libre de Bruxelles – CHU de Charleroi Sleep Laboratory.

[2] Moser D, Anderer P, Gruber G, Paraptics S, Loretz E, Boek M, Kloesch G, Heller E, Schmidt H Danker-Hopfe H, Saletu B, Zeitlhofer J, Dorffner G. Sleep Classification According to AASM and Rechtschaffen & Kales: Effects on Sleep Scoring Parameters. Sleep 2009, 32 2 139.149.

[3] Hadj ZE, Naït-Ali A. QRS complex detection using Empirical Mode Decomposition. Digital Signal Processing, vol 20, 2010, 1221-1228

[4] Xiao M, Yan H, Song J, Yang Y, Yang X Sleep stages classification based on heart rate variability and random forest. Biomedial Signal Proecssing and Control 8 (2013) 624-633

[5] Yoon H, Hwang SH, Choi JW, Lee YJ, Jeong DU, Park KS. REM sleep estimation based on autonomic dynamics using R-R intervals. Physiol Meas. 2017;38(4):631-651.

[6] Singh J, Sharma RK, Gupta AK A method of REM-NREM sleep distinction using ECG signal for unobtrusive personal monitoring. Computers in Biologyand Medicine 78 (2016) 138-143.

[7] Surantha N, Isa SM, Lesmana TF and Setiawan IMA, "Sleep stage classification using the combination of SVM and PSO," 2017 1st International Conference on Informatics and Computational Sciences (ICICoS), Semarang, 2017, pp. 177-182

[8] Lesmana, TF, Isa SM, Nico S. Sleep Stage Identification Using the Combination of ELM and PSO Based on ECG Signal and HRV. 3rd International Conference on Computer and Communication Systems (ICCCS)

[9] Cooray N, Andreotti F, Lo C, Symmonds M, Hu M, De Vos M. Detection of REM sleep behavior disorder by automated polysomnigraphy analysis. Clinical Neurophysiology 130 (2019) 505-514.

[10] Bresler M, Seffy K, Pillar G, Preiszler M, Herscovici S Differentiating between light and deep sleep stages using an ambulatory device based on peripheral arterial tonometry. Physiological measurement 29 (2018). 571-84.

Non-invasive Monitoring of Physiological Signals in an Ovo Model

Lucía Cubero Gutiérrez[1], Leandro A. Hidalgo-Torres[1], Evan Balaban[1], Javier Pascau[1,2], Juan José Vaquero[1,2]

[1] Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid

[2] Instituto de Investigación Sanitaria Gregorio Marañón, Madrid

Abstract

A widely used animal model in foetal and neonatal studies is the avian embryo. Some of their embryonic variables can be currently monitored, including the heart rate, which is a natural indicator of cardiac functioning and general development. Different non-invasive acquisition systems to detect the heart rate have been designed, most of them based on ballistocardiography. This method is founded on the fact that the motions generated by the cardiovascular system induce some micro-vibrations in the eggshell that can be recorded by external sensors. The main objective of this study is to present an innovative acquisition system to measure the heart rate of chicken embryos non-invasively. The methodology is based on the detection of ballistocardiographic signals employing a digital accelerometer. The wavelet transform is used to process the obtained ballistocardiogram. With the described tools, it was possible to detect cardiac and non-cardiac embryonic signals. The heart rate was successfully measured from the ballistocardiogram, and the results agreed with the ones found in the literature.

1. Introduction

Prenatal behavioural development is a subject of constant interest. Neonatal complications are often related to irregularities in the growth of the respiratory, circulatory, and central nervous systems. Avian embryos have noticeably similar developmental patterns compared to human foetuses [1] and are currently widely used due to their easy availability and cost. Numerous research projects have been conducted to analyse their biological processes and characterize *in ovo* the embryological development of their central nervous, cardiovascular and respiratory systems. Because of this, they have also greatly contributed to the fields of immunology, cancer, drug response, neurodevelopment, developmental biology and others.

In this context, the embryonic heart rate (HR) is a relevant foetal parameter as one of the first signs of activity is cardiac work. It is strictly related to diverse developmental conditions during incubation as well as to foetal movement. To study the embryo's HR, invasive, semi-invasive and non-invasive methodologies have been developed [1-4]. Non-invasive techniques, including ballistocardiography, have been gradually developed to compensate for the limitations that the invasive methods present. The ballistocardiogram (BCG) is recorded from the minute movements of the embryo's body induced by the recoil and spread of the heart contraction, which is propagated from the embryo to the amniotic and allantoid fluids, to the rest of the egg contents, and finally to the eggshell. The BCG records these micro-vibrations of the eggshell, which intrinsically contain the cardiac rhythms. The BCG has been demonstrated to be a trustworthy technique to detect the HR and successfully obtained with different sensors, including a phonograph cartridge [2], laser speckles [3] and an electromagnetic induction coil [4]. It is now known that the waves obtained with different instruments differ from each other and that the BCG also varies between embryos of the same age and even between recordings of the same embryo, as a consequence of changes in the embryo's posture.

The BCG provides information about various cardiovascular parameters such as the HR, cardiac output or changes in the heart's continuity, as well as information about non-cardiac activities. Some authors have noted that the BCG cannot be detected before the 7th embryonic day (E7) [2] (this notation refers to the day of incubation and will be employed throughout the paper), with the detection rate being highly improved as incubation proceeds. Moreover, somatic motions of the embryo also contaminate the signal, hiding the cardiac rhythms. Therefore, adequate signal processing is fundamental to obtain a clear pattern [2,3]. In particular, the chicken embryo's HR has been observed to vary in the range of 160 to 300 beats per minute (bpm) [1-4]. Diverse developmental phenomena are related to variations in their HR, and external stimuli like mechanical injury or exposure to cold significantly alter the HR too.

Non-cardiac motions overlap in many cases with the cardiac signal but also provide valuable information about the embryo's development. These movements can be periodic or non-periodic and usually have a noticeably wider amplitude than the cardiac signal. They are divided into general activity and coordinated rotatory movements. The former refers to movements of various body parts that appear in an unpredictable order and pattern, and can have any shape, duration and rate. They include beak clapping, limb movement and swallowing. On the contrary, a coordinated rotatory movement corresponds to the motion by which the embryo turns around itself. They can be observed from E17 and appear more frequently as incubation continues, usually characterizing certain stages of development related to hatching.

Considering all of the above, this study aims to present a new and innovative technique to measure the embryonic HR of chicken embryos non-invasively. The system is based on the acquisition of the BCG, introducing the digital accelerometer as the sensor to record these rhythms. Accelerometers provide an accurate measurement of linear

ISBN: 978-84-09-25491-0, pags. 61- 64

acceleration, vibration, tilt, inclination and shock. In the present study, they are used to detect acceleration, measured in gravitational forces (g). To filter the ballistocardiographic signal and obtain the heart rate, the wavelet transform (WT) is implemented alongside other mathematical tools.

2. Methods

2.1. Acquisition system

To acquire the BCG, a non-invasive system was designed. There is no record in the literature of a system with such characteristics. The structure consists of an electronic board containing an accelerometer and a set of springs that support both the board and the egg (Figure 1).

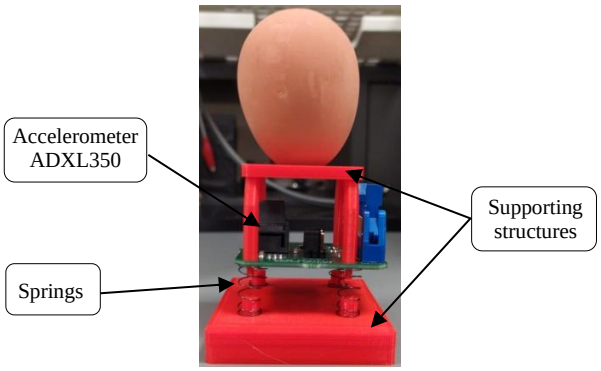


Figure 1. *Set up of the acquisition system*

The two supporting structures (depicted in red) were designed in SolidWorks and 3D-printed with polylactic acid, and are fundamental to ensure the stability of the system. The bottom part supports the four springs, and the top piece pierces the electronic board and is sustained by the springs. The top part also has a small circular hole where the egg is placed. Therefore, any embryonic motion is transmitted to the springs and recorded by the accelerometer in the board.

The electronic system chosen was the EVAL-ADXL350ZM from Analog Devices, which provides a real-time measurement and characterization of the ADXL350Z-M digital accelerometer. This system was expected to provide good outcomes in the set operational range. The ADXL350 is a thin, small and low-power 3-axis accelerometer that records the dynamic acceleration caused by motion or shock, as well as the static acceleration of gravity. It is a polysilicon surface-micromachined system hanging over a silicon wafer by polysilicon springs, which enables the sensor resistance against acceleration forces. The evaluation of the accelerometer was done using the Graphical User Interface provided. The parameters chosen for every experiment were a full resolution, a sampling rate of 100 Hz and a measurement range of ±8g. The board was always calibrated prior to its use to establish the threshold for measuring acceleration. The sensor z-axis was oriented aligned with the gravity 1g field, and the x-axis and y-axis remained in the 0g field.

The springs also have a fundamental role as they enable the acquisition system to move freely in response to any motion generated by the embryo. They allow the egg to be

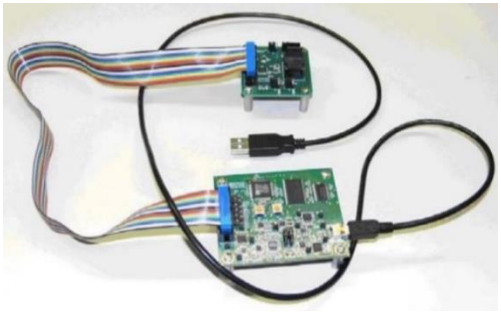


Figure 2. *Typical set up of the EVAL-ADXL350. The satellite board contains the accelerometer ADXL350 and is connected through a ribbon cable to the main board, which transmits the data to the PC. Extracted from the Analog Devices guide*

in a state of zero gravity, as its weight is balanced out. To select the most appropriate spring, three different helical compression models were purchased, all of them made of music steel wire and with dimensions in agreement with the 3D printed parts. Their mechanical characteristics were studied and finally the spring that provided the most accurate measurements was selected. First, an open-source simulator was used to calculate the maximum load that each spring could support under normal loading conditions and under the conditions set by the experiment. This was combined with visual and signal processing experiments. A rubber egg was placed on top of the system, and a servo motor together with an Arduino UNO were used to deliver rhythmic stimulus to the egg. In this way, the coil-recoil motion of the springs was recorded and processed by a custom-like program written in MATLAB (The Math Works, Inc.). The frequency spectrum was calculated to obtain the spring's dominant frequency, and the spectrogram was computed to analyse the time taken by each spring to recoil to its original conformation. The simulator results and the frequency analysis allowed to select the most adequate spring, which was used during the whole experiment.

2.2. Data acquisition

The experiments were carried out initially in 24 chicken eggs, but due to the Covid pandemic the trials had to be stopped before all the embryos could be recorded. Because of this, only a reduced number of 39 signals from 9 different embryos could be obtained.

The eggs were incubated with their round end up, and nonfertilized eggs were excluded from E5 after being candled. Two different incubators were used and were set to maintain a constant temperature of 37.5°C and air humidity of 55%. The main incubator was programmed to rotate the eggs every 30 minutes. The embryos were kept in this incubator under the described conditions and were only taken out during their recording span. It was decided that the embryos would be studied from E12 to E20. During the recording period, one egg was taken out at a time and placed in the experimental incubator, where a vibration isolation platform had been placed to reject any external vibrations. The acquisition system was located above the centre of this platform, and the egg was placed on top of it. The accelerometer ADXL350 was then calibrated and the recording was started. After the first 20

minutes, the egg was manually moved and then the recording continued for another 20 minutes.

2.3. Signal processing

The recorded signals were processed with MATLAB. First, the data generated by each of the accelerometer's axes was studied in the time domain. After some filtering stages, the Discrete Fourier Transform (DFT) of the z-axis signal was computed and represented. Surprisingly, no dominant frequency appeared in the frequency spectrum, and therefore it was not possible to extract the cardiac frequency with this mathematical tool. Instead, a new method was approached: the WT. The wavelets are broadly used for multiresolution decomposition of continuous-time signals, and as they provide a superior time-frequency localization than the DFT, they yield particularly good results in pattern recognition. To implement the WT, a custom-like program was used together with the Wavelet Analyzer App of MATLAB. After processing the signals with different families of wavelets and cross-validating the results, the mother wavelet that better denoised the recordings was the symlet7 with a level of decomposition 5, and was therefore employed along the rest of the study. Then, some conditions of height and distance were set to isolate each heartbeat and discriminate its main peak. The HR was finally computed by dividing the number of heartbeats by the length in minutes of the signal.

3. Results and Discussion

A total of 9 embryos could be studied, giving rise to 39 recordings. 4 embryos were recorded for 3 days, 1 for 4 days, 1 for 5 days, and 3 for 6 days. The HR range of each incubation day is depicted in table 1. Due to the small sample size, no statistical analysis could be performed, and no general conclusions could be made. Nevertheless, it was observed that each embryo had a characteristic cardiac rhythm that was maintained in a narrow range. Because of this, conclusions can be drawn from individual analysis and some examples will be now explained.

Day of Incubation (E)	Number of embryos	Heart Rate range (bpm)
E12	4	211 – 265
E14	5	247 – 273
E15	9	220 – 273
E16	9	224 – 273
E17	4	223 – 237
E18	4	226 – 238
E19	4	224 – 276

Table 1. *Heart Rate analysis*

External stimuli were observed to alter the HR. First, various embryos showed a significantly higher HR during the first recorded minutes, probably a consequence of the embryo's excitement after being translated into a different environment. Secondly, the temperature of the experimental incubator slightly decreased throughout each recording, what resulted on an average drop of 6 bpm of the embryos' HR. This result was expected as many research studies have addressed the negative impact that a decrease in temperature has in the chicken embryos' HR. A reduction of 1°C can derive in a drop of 20 bpm. An example of a denoised cardiac signal is exposed in Figure 3. The small discrepancies among the size of the heartbeats are a consequence of the denoising process by the WT, but these did not entail any impediment for cardiac detection.

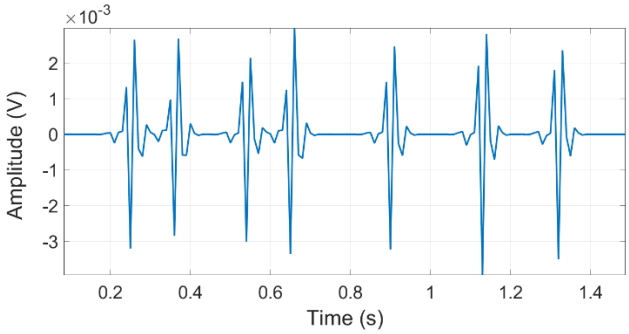


Figure 3. *Filtered cardiac signal of a 19-day embryo*

Moreover, although non-cardiac movements were observed in every recording, the embryos that generated more non-cardiac movements (mainly in E18 and E19) exhibited a dominant frequency in their spectrum. This peak corresponded to the characteristic frequency of the springs used and appeared as a result of the load that the non-cardiac motions enforced on the springs. This frequency was filtered for HR detection but was useful for the study of non-cardiac activity. A comparison between heartbeats and general movements can be found in Figure 4, where two subsequent motions are depicted. As it can be observed, non-cardiac movements have a significantly wider amplitude and rate than cardiac rhythms, and are also recognizable because of their random shapes.

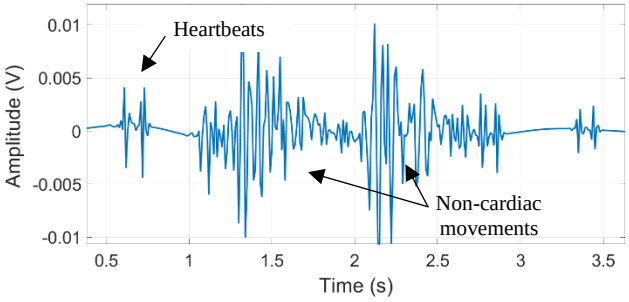


Figure 4. *Comparison between cardiac and non-cardiac movements*

More vigorous motions were not noticeable until E17. Figure 5 represents the non-cardiac movements observable in a 19-day embryo. From top to bottom, it depicts the general activity of a recording, 40 seconds of the signal in which 5 different movements appear, and 3 subsequent movements observable with a higher degree of detail. In the latter image, the diverse motions have completely different patterns. The first one probably corresponds to a coordinated rotatory movement of the embryo, due to its symmetrical shape [5]; whereas the last two have a smaller amplitude and a more irregular pattern, and possibly are just random general shifts of different body parts. Another coordinated rotatory movement is clearly observable in Figure 6.

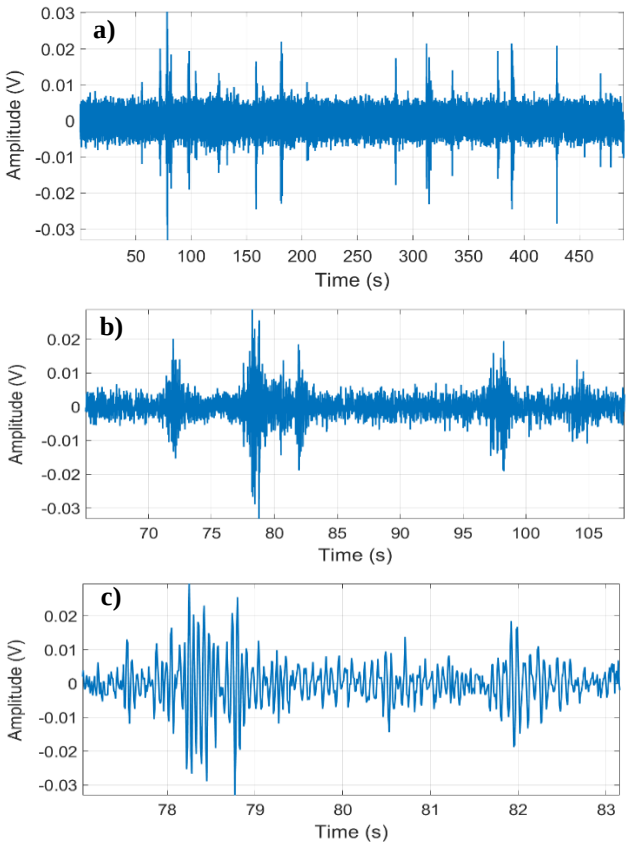


Figure 5. *Non-cardiac activity of a 19-day embryo.* ***a)*** *General activity of a 500 s recording.* ***b)*** *Enlargement of the previous signal in which 40 s of the recording are depicted.* ***c)*** *3 subsequent non-cardiac motions appearing along 6 s*

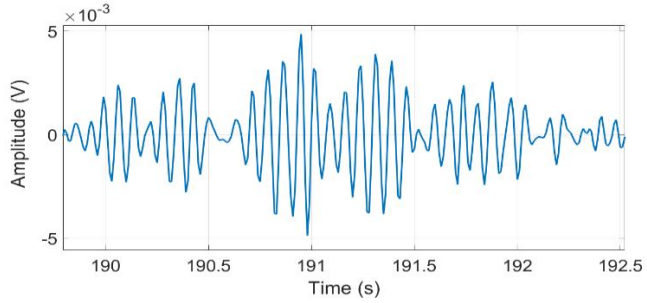


Figure 6. *Coordinated rotatory movement of a 19-day embryo*

Some limitations of the study must be highlighted at this point. In the first place, the characteristic frequency of the spring is removed from the signals to avoid the interference of the coil-recoil motion, what might hide some interesting patterns of the embryonic signals. Another main restriction is the frequency sampling utilized, which only had a value of 100 Hz. For the detection of the HR as a wave it is sufficient, but a higher value would provide much better outcomes and could even reveal a clearer shape of each heartbeat. This constraint is easily solved by increasing its value. Finally, an inherent limitation of the ballistocardiography technique is that it is always perturbed by diverse sources of noise and various interferences. A lot of signal processing is necessary to compensate for these circumstances, and still these noisy components usually appear in the filtered BCGs.

4. Conclusions

Despite of the limitations explained, the objective of the study, which was to characterize the accelerometer as an appropriate sensor to measure the chicken embryo's HR, was successfully accomplished. These tools, together with the WT, could detect HR variations and development throughout incubation. The HR ranges detected agreed with the literature [1-4], although no statistical analysis could be performed. On the contrary, it did not provide such good results in capturing the shape of each heartbeat, as they mostly suffered from various interferences.

Furthermore, the system proposed provided a good knowledge about non-cardiac movements. The accelerometer could detect many movements of different amplitudes, strengths and durations, from unlocalized minuscule motions to coordinated rotatory movements. Nonetheless, these motions usually interfered with the cardiac rhythms, and had to be adequately filtered to acquire the HR.

The acquisition system presented is only a first prototype and will undergo various modifications and improvements in future studies. Moreover, the frequency sampling must be increased in further research to achieve better heartbeat characterization. Other processing methods can also be tried to obtain better outcomes than the ones provided by the WT.

Acknowledgments

Ministerio de Ciencia, Innovación, PID2019-109820RB-I00; Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación, TEC2016-78052-R / MINECO/FEDER, UE; This project has received funding from 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.

Bibliography

[1] E. Balaban, M. Desco, and J.-J. Vaquero, "Waking-like brain function in embryos," Current Biology, vol. 22, no. 10, pp. 852–861, 2012.

[2] Y. Suzuki, H. Musashi, and H. Tazawa, "Noninvasive heart rate monitoring system for avian embryos based on the ballistocardiogram," Medical & Biological Engineering & Computing, vol. 27, pp. 399–404, 1989.

[3] Y. Hashimoto, T. Narita, and H. Tazawa, "Cardiogenic ballistograms of chicken eggs: Comparison of measurements," Medical & amp; biological engineering & amp; computing, vol. 29, no. 4, pp. 393–397, Jul.1991.

[4] H. Ono, R. Akiyama, Y. Sakamoto, J. T. Pearson, and H. Tazawa, "Ballistocardiogram of avian eggs determined by an electromagnetic induction coil," Medical and Biological Engineering and Computing, vol. 35, no. 4, pp. 431–435, Jul. 1, 1997.

[5] J. K. Kovach, "Development and mechanisms of behaviour in the chick embryo during the last five days of incubation," Journal of Comparative and Physiological Psychology, vol. 73, no. 3, pp. 392–406, 1970

Protocolo biomecánico de miembro superior para estudio del lanzamiento de pitcher en jugadores de béisbol

D. Díaz Gómez, J. Niño Merchán, J. Vega Beltrán, L. Amado Forero

Ingeniería Biomédica, Universidad Autónoma de Bucaramanga, Colombia, {ddiaz223, jnino518, jvega72, lamado175}@unab.edu.co

Resumen

Las lesiones articulares son un riesgo cada vez más común entre los jugadores de beisbol, esto debido a los movimientos incorrectos y el sobreesfuerzo realizado durante los gestos deportivos tales como el lanzamiento o "pitching", es por ello, que surge la necesidad de crear un protocolo de análisis biomecánico de miembro superior para beisbolistas durante el lanzamiento. La reconstrucción de cada una de las etapas que componen el gesto deportivo se realizó por medio de marcadores optoelectrónicos, sistemas de tracking, canales de electromiografía y software de procesamiento, los cuales en conjunto permitieron la recolección de las variables más relevantes durante el desarrollo del gesto deportivo y la reconstrucción tridimensional del mismo, y de esta forma se lograron identificar los errores y factores de riesgo para lesiones potenciales. Luego, el presente documento contiene la composición detallada del protocolo de análisis biomecánico, desde las herramientas necesarias para la adquisición de variables más relevantes, los pasos a seguir para su correcta implementación y el tratamiento de la información obtenida en el desarrollo del gesto deportivo, para así obtener un análisis biomecánico completo.

1. Introducción

Recientemente se ha observado un aumento en la aparición de ciertas lesiones en los jugadores de béisbol a raíz de los movimientos utilizados en la posición de pitcher. En atletas jóvenes se ha observado que aproximadamente el 30% ha sufrido de dolor en las articulaciones del miembro superior durante las temporadas [1]; el 25% de los jugadores de la **major league of baseball** han requerido de una reconstrucción del ligamento ulnar colateral (UCL) [2], y cuando se habla de jugadores de secundaria ha habido un aumento que supera el 13% en las operaciones de UCL debido a lesiones por sobre exposición [3].

El lanzamiento o "pitching" en inglés es el resultado de una cadena cinética, en la cual la estabilidad y movilidad de la zona pélvica y escapular es importante para la transferencia de energía desde la parte proximal de los miembros inferiores hasta la parte distal del miembro superior [4]; el pitching se realiza a lo largo de 6 etapas: cuerda, preparación temprana, preparación tardía, aceleración, desaceleración y seguimiento [5], donde la mala realización de estas fases aumenta el riesgo de lesión [6-7] , pese a esto el 40% de las mecánicas incorrectas durante el pitching pueden ser corregidas una vez sean detectadas [8].

Se plantea un protocolo para el estudio biomecánico del pitching enfocado en el miembro superior, para esto, se hace uso de un laboratorio de biomecánica, por medio de diversas tecnologías allí dispuestas, como lo son las plataformas de presión, marcadores optoelectrónicos y electromiogramas.

2. Metodología

2.1. Creación de protocolo

Con base a la información del gesto deportivo *Anexo 1* se establecen dos protocolos para la ejecución del estudio.

1. **Protocolo para el posicionamiento de marcadores**

Para este protocolo biomecánico se ubican marcadores optoelectrónicos sobre los ejes de acción de mayor relevancia en el gesto deportivo, tal que se definen 13 puntos de interés como se muestra en la *Figura 1*, siendo estos las articulaciones de C7 y S2, para que dichos marcadores detecten la torsión que hace el tronco en cada una de las fases del movimiento del pitcher; en la articulación acromioclavicular para tener línea de seguimiento y sea detectado el movimiento como la elevación y rotación del hombro; también en las articulaciones del codo, muñeca, metacarpofalángica y en la segunda falange del segundo dedo de la mano dominante para detectar la ubicación en los ejes x, y, z; en las crestas iliacas para complementar la observación del movimiento que realiza el tronco y por último, dos marcadores en el tronco inferior, uno en la rodilla y el otro en el tobillo de la pierna contraria al brazo dominante.



Figura 1. *Ubicación de los marcadores optoelectrónicos en los puntos articulares [9]*

Con el fin de obtener el registro de los movimientos realizados por el jugador de béisbol durante el lanzamiento, se utilizan 6 cámaras optoelectrónicas

ISBN: 978-84-09-25491-0, pags. 65- 68

SMART-DX6000 [10] dispuestas en la parte superior (techo del laboratorio) en dos arreglos de 3 camaras con enfoque al punto medio del mismo.

2. Protocolo para el posicionamiento de los canales de EMG

Adicionalmente, se ubican los canales de electromiografía BTS FREEEMG [10] sobre los músculos superficiales relacionados en la *Tabla 1* con el fin de obtener información de la actividad eléctrica de los músculos más relevantes en la ejecución de la práctica deportiva. Para el posicionamiento de los canales de EMG se asignan cada uno de los siete disponibles a un músculo de relevancia en la acción deportiva.

Canal EMG	Movimiento	Músculo
1 y 2	Elevación	Trapecio y deltoides
3	Flexión	Bíceps
4	Abducción	Pectoral mayor
5 y 6	Rotación	Oblicuo externo izquierdo y derecho
7	Extensión	Tríceps

Tabla 1. *Ubicación de los canales de EMG en los músculos con mayor activación*

Mediante el conjunto de dispositivos utilizados junto con los protocolos se identifica si el movimiento de los músculos accionados por el pitcher es el correcto o si por el contrario hay que realizar algún tipo de corrección en el gesto deportivo, esto con el fin de evitar posibles lesiones en el jugador a nivel muscular.

2.2. Toma de medidas antropométricas

Se realiza la toma correspondiente a las diferentes medidas antropométricas del paciente.

Medida	Medida
Peso	Anchura lateral de brazos
Estatura	Largo de brazo desde la vertical
Alcance vertical	Largo de puño
Altura de hombros	Ancho de hombros
Altura de codos	Ancho de pecho
Altura espina iliaca derecha	Circunferencia máxima del pecho
Altura espina iliaca izquierda	Circunferencia de la cintura
Altura rodilla	Circunferencia de la cadera
Altura a la muñeca	Largo de la mano
Altura al tercer dedo	Largo palma de la mano
Anchura de codos	Ancho palma de la mano

Tabla 2. *Medidas antropométricas.*

2.3. Preparación del paciente

En esta etapa se posicionan los elementos de sensado sobre el paciente, siguiendo los protocolos de posicionamiento de los marcadores optoelectrónicos y los canales de EMG.



Figura 2*. Vistas del paciente con los marcadores optoelectrónicos y los canales de EMG*

2.4. Indicaciones al paciente

Una vez realizada la toma de medidas antropométricas y el posicionamiento de los canales de EMG, junto con los marcadores optoelectrónicos, se revisan con el paciente las diversas acciones que componen el gesto deportivo a analizar, así mismo, se le da información sobre la ubicación espacial del laboratorio durante la ejecución del gesto deportivo y el rango de movimiento del cual dispone.

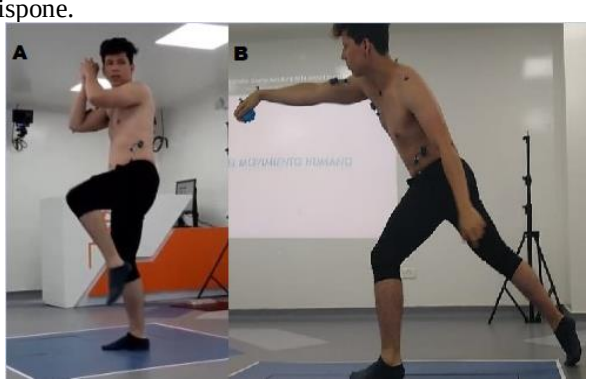


Figura 3*. A. Vista frontal, inicio del gesto deportivo. B. Vista lateral de la ejecución del gesto deportivo.*

2.5. Procesamiento de datos

Después de extraer los datos del gesto deportivo se realiza el procesamiento de las señales obtenidas por los canales de EMG, usando el software Matlab, donde se diseña un filtro pasa bandas de 0.5 Hz a 150 Hz ya que los canales de EMG realizan un filtro pasa banda Butterworth.

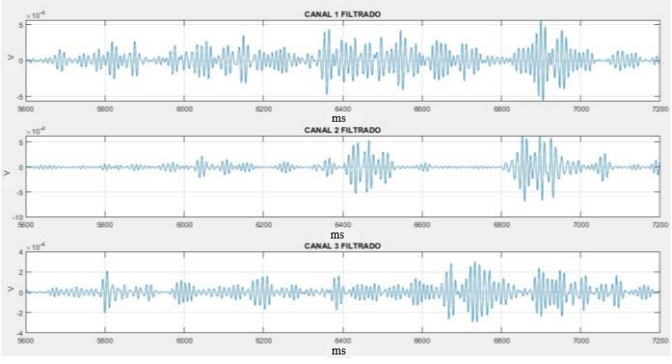


***Figura 4**. Señales filtradas de los canales 1, 2 y 3*

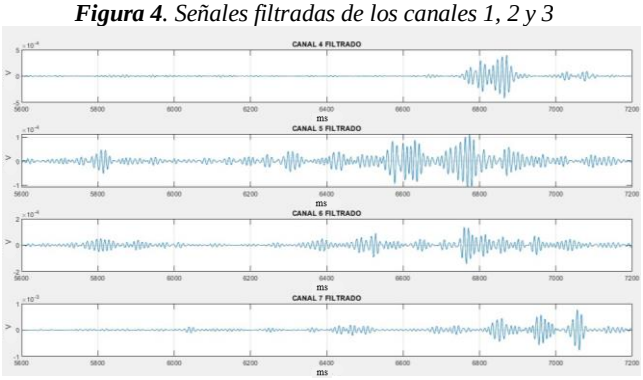


***Figura 5**. Señales filtradas de los canales 4, 5, 6 y 7*

2.6. Reconstrucción del gesto deportivo

Por medio del software Smart Tracker se adiciona un protocolo con la toma de datos correspondiente a los marcadores optoelectrónicos, y así se reconstruye el modelo 3D del gesto deportivo.

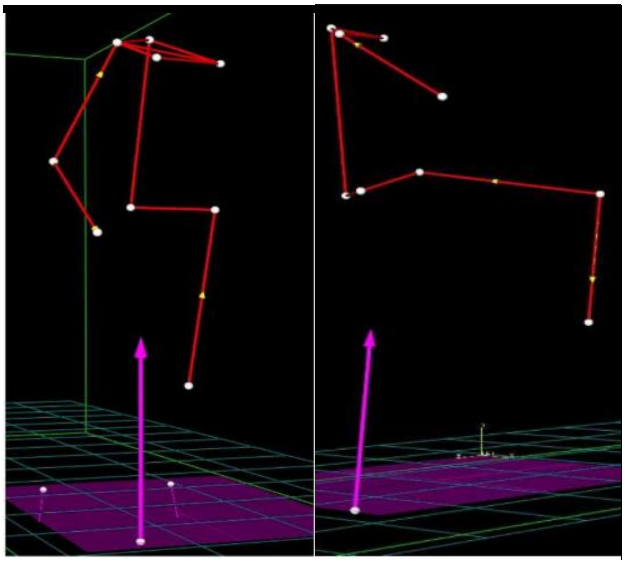


***Figura 6**. Reconstrucción 3D.*

El modelo 3D permite observar los puntos de interés del gesto deportivo a través del movimiento completo, y así mismo, se relacionan los vectores entre los puntos de los marcadores para obtener información sobre el recorrido angular de las articulaciones de mayor relevancia; tal que se opta por graficar la variación de ángulos para 3 articulaciones como se puede ver a continuación.

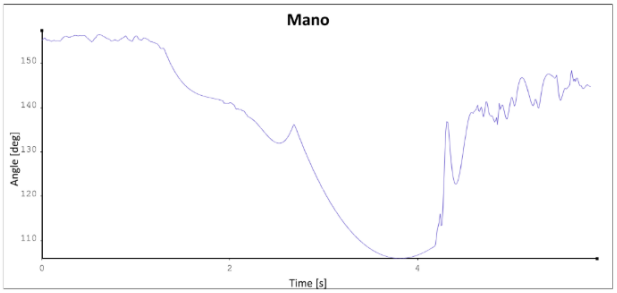


***Figura 7**. Recorrido angular de la articulación de la mano*

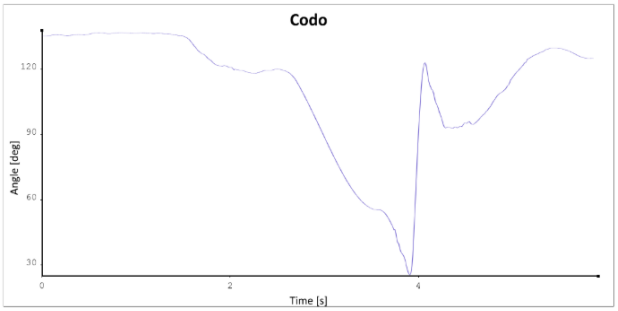


***Figura 8**. Recorrido angular de la articulación del codo.*

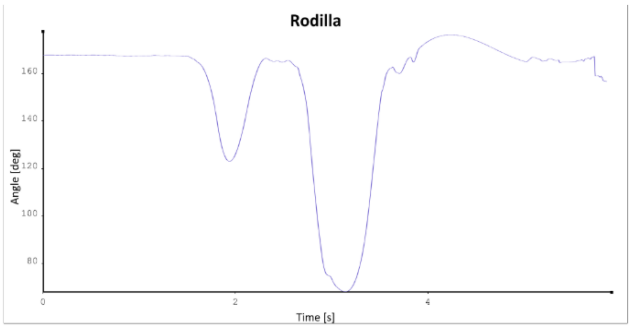


***Figura 9**. Recorrido angular de la articulación de la rodilla.*

3. Resultados y conclusiones

Con el fin de generar un buen análisis biomecánico se deben realizar varias tomas para reducir el margen de error, tal que se pueda determinar cuál de ellas presenta un menor ruido (marcadores clonados y no existentes), conservando los trazos (recorridos) necesarios para la reconstrucción 3D del gesto deportivo, por lo cual, se realizaron un total de 52 tomas y se seleccionó la toma que a pesar de que no contenía la totalidad de los trazos requeridos presentaba la mayor posibilidad de reconstrucción.

Para realizar una correcta reconstrucción del gesto deportivo es necesario tener un software robusto para el procesado de los datos, puesto que en muchas ocasiones se pierde parte de la información y depende de la capacidad del software la interpolación de los datos perdidos, en este caso puntual el software Smart Tracker permitió interpolar los datos perdidos en los gráficos sobre los ángulos logrando así una mejor reconstrucción 3D.

Como resultado del desarrollo del estudio, se obtuvo la creación de un protocolo para el estudio biomecánico del gesto deportivo de pitching, mediante el uso de marcadores optoelectrónicos y canales de EMG.

Este protocolo permite la reconstrucción del movimiento y la obtención de variables importantes como, la respuesta electrica en los musculos implicados y las trayectorias angulares de las articulaciones de mayor incidencia durante el movimiento.

4. Anexos.

Anexo 1. Etapas del pitching

1. Cuerda o wind-up: Esta fase comienza desde la posición estática del pitcher frente al bateador, y finaliza cuando la rodilla alcanza su altura máxima "balance point", hace referencia al momento en el que el pitcher comienza a retirar la pelota del guante para preparar el lanzamiento[11-12], esta fase es la que presenta menor riesgo de lesión de las 6 fases, pese a esto la ubicación del centro de gravedad se encuentra desviado hacia la parte anterior o posterior la cadena cinética se transmite hacia la extremidad superior, predisponiendo las articulaciones de hombro y codo a lesiones [12]

2. Preparación temprana: Esta etapa comienza una vez la rodilla alcanza su altura máxima y finaliza cuando la pierna elevada (dominante) toca de nuevo el suelo, durante esta fase el centro de gravedad cae frente a su posición anterior y es acelerado en dirección a la parte posterior, permitiendo el aumento de la distancia linear y angular del tronco, generando una mayor transmisión de energía al miembro superior [12-13]

3. Preparación tardía: Esta fase se inicia cuando el pie dominante hace contacto con el suelo y finaliza con la posición de máxima rotación externa del hombro de lanzamiento [14], en esta fase el tronco rota de atrás hacia adelante (oblicuos externos) al igual que la cadera, la articulación de la rodilla dominante se flexiona para dar soporte al movimiento, el miembro superior y el hombro son abducidos, mientras que el codo se flexiona [12]

4. Aceleración: Esta fase inicia desde la máxima rotación externa del hombro y finaliza con la liberación de la bola, durante esta fase la región axial del cuerpo sigue en rotación y el codo pasa de una flexión de 90 grados a 120 grados para devolverse hasta los 25 grados como acción de la fuerza centrífuga. [15]

5. Desaceleración: Esta fase inicia con la liberación de la pelota y termina con la finalización de la rotación humeral (0 grados), la abducción máxima del hombro y la abducción horizontal del brazo a los 35 grados [11]. Esta es la fase más violenta del pitching ya que pone la mayor cantidad de cargas sobre las articulaciones del miembro superior.

6. Seguimiento: Esta fase abarca los movimientos generados por las fuerzas una vez finalizada la acción (abducción horizontal de 60 grados) [11], y culmina con el pitcher en su posición normal de campo.

Referencias

[1] Lyman S, Fleisig G, Waterbor J, et al. Longitudional study of elbow and shoulder pain in youth baseball pitchers. Med Sci Sports Exerc. 2001;33:1803–1810.

[2] Conte SA, Fleisig GS, Dines JS, Wilk KE, Aune KT, PattersonFlynn N, et al. Prevalence of ulnar collateral ligament surgery in professional baseball players. Am J Sports Med. 2015;43(7):1764– 9.

[3] Petty DH, Andrews JR, Fleisig GS, et al. Ulnar collateral ligament reconstruction in high school baseball players: clinical results and injury risk factors. Am J Sports Med. 2004;32:1158–1164.

[4] Oliver, G. D., Plummer, H., Henning, L., Saper, M., Glimer, G., Brambeck, A., & Andrews, J. R. Effects of a Simulated Game on Upper Extremity Pitching Mechanics and Muscle Activations Among Various Pitch Types in Youth Baseball Pitchers. Journal of Pediatric Orthopaedics. 2017.

[5] Moyer JE, Brey JM.Shoulder Injuries in Pediatric Athletes. Orthop Clin North Am. 2016 Oct;47(4):749-62.

[6] Chalmers. PN, Wimmer MA, Verma NN, Cole BJ, Romeo AA, Cvetanovich GL, et al. The relationship between pitching mechanics and injury: a review of current concepts. Sports Health. 2017;9(3):216–21.

[7] Shane T. Seroyer, MD,* Shane J. Nho, MD.The Kinetic Chain in Overhand Pitching Its Potential Role for Performance Enhancement and Injury Prevention.Sports Health. 2010 Mar; 2(2): 135–146.

[8] Fleisig GS, Diffendaffer AZ, Ivey B, Aune KT. Do baseball pitchers improve mechanics after biomechanical evaluations? Sports Biomech. 2018;17(3):314-321.

[9] "Cuerpo humano fuerte y flexible | Mundo Entrenamiento", Mundo Entrenamiento. [Online]. Available: https://mundoentrenamiento.com/cuerpo-humano-fuerte-y-flexible/

[10] "BTS GAITLAB | Integrated Motion Analysis Systems | BTS Bioengineering", BTS Bioengineering. [Online]. Available:https://www.btsbioengineering.com/es/products/freeemg/

[11] Meister K. Injuries to the shoulder in the throwing athlete. Part one: biomechanics/pathophysiology/classification of injury. Am J Sports Med. 2000;28(2):265-275

[12] Gary J. Calabrese.PITCHING MECHANICS, REVISITED.Int J Sports Phys Ther. 2013 Oct; 8(5): 652–660.

[13] Dillman CJ, Fleisig GS, Andrews JR. Biomechanics of pitching with emphasis upon shoulder kinematics. J Orthop Sports Phys Ther. 1993;18(2):402-408.

[14] Atwater AE. Biomechanics of over arm throwing movements and throwing injuries. Exerc Sports Sci Rev. 1979;7:43–85.

[15] Fleisig GS, Escamilla RF, Barrentine SW. Biomechanics of pitching: mechanism and motion analysis. In: Andrews JR, Wilk KE, editors. , eds. Injuries in Baseball. Philadelphia, PA: Lippincott-Raven; 1998:3-22.

Diseño de concepto de componente protésico de miembro superior para ciclismo

I. Vaz Pato[1], O. Pelliccioni Monrroy[2]

[1] Ottobock Ibérica, Madrid, España

[2] Departamento de Ciencia y Tecnología, Universidad Internacional de Valencia, Valencia 46002, España, orlandojose.pelliccioni@campusviu.es

Resumen

Se describen las etapas del desarrollo de una solución protésica para un ciclista amputado transhumeral, solicitada por la carencia de componentes en el mercado para necesidades en el sector deportivo. El producto consiste en un kit de componentes que conjugan movilidad y seguridad en las actividades en la que se desempeña el usuario con la bicicleta, cumpliendo además con las normativas dictadas por la Unión de Ciclistas Internacional. Guiado el diseño bajo el enfoque de ingeniería concurrente orientado a la calidad, se describen los requerimientos del usuario, los técnicos y de las normativas, así como el prototipo propuesto y otros componentes protésicos. Para su validación, primero se construyó un prototipo funcional, mejorado posteriormente con el uso de herramientas de diseño y análisis asistido por ordenador. Hasta la fecha, el kit continúa en uso por el usuario en diversas pruebas competitivas en España.

1. Introducción

El deporte en el mundo de la amputación se considera de gran importancia para la rehabilitación del individuo y su consecuente adaptación a la prótesis. El ciclismo se introdujo como un deporte paralímpico en 1984, en Inglaterra, en el cual sólo se involucraron atletas con parálisis cerebral. Sin embrago, no fue hasta 1992 que atletas con amputaciones pudieron participar en las Paralimpiadas en la categoría de ciclismo[1]. La aplicación de una prótesis adaptada se verifica frecuentemente cuando el usuario presenta necesidades específicas para realizar una determinada actividad o tarea, en contexto laboral o de deporte y ocio. En este grupo de prótesis pueden encontrarse adaptaciones para agarrar determinadas herramientas en el entorno laboral, montar a caballo, en bicicleta, jugar al golf o tocar un instrumento musical. A pesar de los avances en el diseño y la composición de las prótesis de miembro superior/inferior, no hay evidencia científica de cuál de los diseños existentes es el más adecuado o el mejor para un deporte en particular o un grupo de amputados. Actualmente, la elección de una prótesis de deporte también se basa en el individuo, soluciones individuales basadas en opiniones de expertos en el deporte y en la rama de la ortopedia [2].

Cuando se habla del diseño de un dispositivo de uso médico, y cuando la adaptación deportiva en cuestión es para una amputación, deben tomarse en consideración las distintas necesidades del paciente y la regulación asociada a este tipo de adaptaciones en el deporte en concreto.

De esta forma, el propósito de la ejecución de una metodología de diseño concurrente consiste en la mejora de tiempos de producción y de costes del diseño en el desarrollo de nuevos productos en una empresa, calidades que pueden suponer una ventaja competitiva en el mercado. Este tipo de metodología permite abarcar de forma simultánea las necesidades de los distintos grupos involucrados en el desarrollo del dispositivo, ya sea el diseño del producto, los materiales o el usuario final de los mismos. Para ello también se solicita la colaboración de profesionales que asesoren en el desarrollo del proyecto, y se analizan distintas herramientas de modelaje, simulación y cálculo, o incluso prototipado rápido.

Una de las perspectivas del diseño concurrente, denominada diseño para la calidad, busca más allá de la conformidad de un producto con las especificaciones previstas, incidiendo en el mismo diseño para que éste sea más apto para la calidad. Esta perspectiva tiene en cuenta el grado de satisfacción que se proporciona a las expectativas del usuario, así como la rentabilización de los recursos, eliminando las pérdidas asociadas. Para concretar este proceso de calidad a través del diseño, ésta se evalúa por medio de la función de calidad (QFD).

2. Tecnología del ciclismo adaptado

La variabilidad presente entre los atletas paralímpicos, ya sea por su nivel de amputación o por su grado de discapacidad, representa múltiples desafíos para todos los involucrados en el entrenamiento de un atleta y su adaptación al deporte. Esto no solo influye en la forma de abordar a cada atleta desde el punto de vista de las mejoras de su rendimiento sino también desde la capacidad de incrementar esos resultados con el mínimo riesgo de lesión.

En general, las prótesis de miembro superior presentan capacidades de agarre del manillar y utilización de frenos y cambios. Sin embargo, su utilización en la modalidad exigida, y con base a las características del usuario, no requiere todas estas funcionalidades, cuyo control se traslada al miembro contralateral. Las prótesis de ciclismo deben ser ligeras, capaces de posicionarse en distintas posiciones durante la prueba y específicas para cada modalidad del ciclismo a la que se destinan [3-6].

Con las necesidades planteadas, se realizó una revisión de productos disponibles actualmente con la intención de seleccionar las tecnologías más relevantes para el estudio.

ISBN: 978-84-09-25491-0, pags. 69- 72

La metodología utilizada para la recopilación de soluciones disponibles consistió en la revisión de los aspectos técnicos disponibles que definen los dispositivos actuales, su utilización y características pertinentes para aplicación a nivel competitivo. Cabe destacar que ninguno de estos dispositivos está diseñado con la finalidad aquí planteada: la aplicación a nivel transhumeral para la práctica de ciclismo de pista; sin embargo, se consideraron como competencia al dispositivo planteado para el proyecto, puesto que esta necesidad no ha sido cubierta hasta el momento en España.

3. Requerimientos de calidad

Para poder avanzar durante las fases de diseño, es necesario comenzar con el concepto, estableciendo las especificaciones que se reflejarán en el producto final. Bajo un enfoque de diseño concurrente se rescataron y estudiaron los requerimientos establecidos por el usuario, detalles técnicos involucrados y normativa existente. Su análisis y correlación permitió establecer criterios ponderables que ayudaron en la valoración y selección de las soluciones de mayor impacto, identificar las limitaciones para su manufactura y reducir los costes, todo dirigido a maximizar la calidad del producto en comparación con productos de la competencia en el mercado actual.

Para lograr ese grado de empatía necesario entre el diseñador y el usuario, se mantuvieron conversaciones con un grupo de muestra de usuarios potenciales contactados a través de relación de los autores con Ottobock Ibérica. Algunos requerimientos que pueden mencionarse se relacionan al interés en que el producto tenga gran movilidad, y esto deriva en aumento de comodidad y mejora en la postura en la práctica deportiva. También se solicita aerodinamia que no afecte en el rendimiento deportivo, resistencia de las piezas al desgaste y caídas, debe inspirar sensación de seguridad y debe ser fácil de maniobrar con una sola mano.

Los requerimientos técnicos rescatados con mayor puntaje se centran en reusabilidad de las piezas, poco peso, durabilidad, fácil manutención y limpieza. Por su parte, normativas internacionales relacionadas a deportes de competición adaptados destacan requerimientos de seguridad, fácil manipulación (en este caso con el miembro disponible), correcta posición durante la prueba del individuo y estos dispositivos diseñados no deben afectar las medidas de la bicicleta. La seguridad del usuario es el punto central [2,7].

4. Requerimientos funcionales

Continuando con la metodología QFD se tuvieron en consideración factores de fabricación y montaje del sistema, y posturales del usuario: ▪ Los requerimientos posicionales asociados a la optimización de la aerodinámica implican que el dispositivo presente movilidad, garantizando que se comporte de forma idéntica a la del codo, auxiliando al ciclista durante su uso. ▪ Las simulaciones por ordenador permiten verificar el dimensionamiento de las piezas, estudiar su resistencia y asistir en la selección de materiales. ▪ Adaptabilidad del producto a distintas bicicletas, en diferentes posiciones, también permitiendo la correcta posición postural del usuario. ▪ Adaptabilidad del producto a distintas prótesis, miembro amputado y distintos niveles de amputación independientemente de su origen. ▪ Los materiales deben disponerse fácilmente en el mercado. Su selección debe aportar beneficios al diseño, minimizando peso sin afectar su desempeño mecánico. ▪ La disminución en el número de partes facilitaría su fabricación y montaje. ▪ Se valora un diseño sencillo que no comprometa la calidad y durabilidad. ▪ El dispositivo debe permitir seguridad y movilidad al usuario, y desacople rápido cuando se encuentre fuera de los límites de seguridad establecidos. ▪ El dimensionamiento del dispositivo debe cumplir con las normativas establecidas por UCI relacionadas con las medidas de una bicicleta y con su aprobación previa al uso en competencias.

Las especificaciones de mayor peso que permiten asistir de forma eficaz a un usuario con una amputación de desarticulación de codo o transhumeral pasan por garantizar que el sistema protésico le ofrezca maniobrabilidad para adecuarse a las posturas necesarias en competición. Seguidamente, la prótesis debe ser resistente ante solicitaciones generadas por el atleta, y finalmente, deben cumplir con la normativa decretada por la UCI, garantizando su seguridad.

Para una prueba de pista se consideró que el usuario adopta dos posturas distintas: la postura de arranque, y la postura de prueba. En la postura de arranque se tuvieron en cuenta dos posibilidades: el arranque sin que la prótesis contacte con la pieza de acople y en particular, el arranque desde una postura intermedia, ya ajustada a la pieza de acople. Se solicitó al usuario que adoptase la postura en la cual se encontrara mejor posicionado para la realización de su prueba, para elegir la posición y orientación de la pieza distal de la prótesis y de la pieza de acople.

Analizando los movimientos del usuario, que van variando de acuerdo con el balanceo habitual del miembro superior originado durante la prueba, se admite que la fuerza máxima presente ronda los 737 N y la fuerza máxima aplicada al dispositivo los 316 N. Estos valores son reportados por Soden y Ayedefa [8] para la postura de arranque de un individuo con 64,8 kg. Para su análisis en movimiento continuo, en superficies horizontales, las fuerzas correspondientes representan el 15,3% y el 10,76% del peso del individuo. En este punto en particular la literatura relacionada con la práctica de ciclismo de pista resulta escasa.

5. La casa de la calidad o el Análisis QFD

Para entender mejor el funcionamiento de este análisis se identifican las zonas de esta representación gráfica del diseño concurrente (ver Figura 1). En el lado izquierdo se encuentran reunidos todos los requerimientos de calidad, y poseen un puntaje asociado que refleja su peso o importancia, fijada previamente por el levantamiento y estudio de los requerimientos: La parte superior “de la casa” simboliza los requerimientos funcionales, que dan respuesta a la calidad necesaria. Estos también poseen un peso o importancia que se obtiene de las relaciones que se

verifican entre los requerimientos de calidad y los funcionales, en la parte central de la casa. Se identifican los factores más críticos a través de la atribución de una puntuación más elevada en la parte inferior del esquema. También se atribuye una ponderación de dificultad a los requerimientos funcionales, buscando exponer su complejidad directamente relacionada con los conocimientos del equipo de trabajo y el mismo proyecto. De esta manera, se pueden verificar qué costes se podrán tener asociados al diseño y los tiempos que conllevarán.

La parte derecha de la casa de la calidad representa un análisis de los mismos requerimientos con otros productos comerciales que podrían ser posibles competidores con la solución propuesta. Esta sección permite evaluar en qué medida el nuevo producto puede introducirse al mercado y cuáles son sus principales mejoras o debilidades.

La Figura 1 resume los resultados del nuevo producto propuesto y toda la investigación realizada para el diseño de concepto. El cumplimiento de las normativas existentes es lo que más se destaca, a la par del diseño sencillo y aerodinámico del mismo. Luego se valora la movilidad, la facilidad de manejo y la durabilidad ante un uso repetitivo.

El estudio de materiales, a pesar de su menor valoración, influye en cierta medida para cumplir con los requerimientos de calidad para garantizar alta resistencia y seguridad. Otros criterios como el peso reducido y que sus componentes sean reutilizables, aunque describen características del producto, presentan menor puntaje.

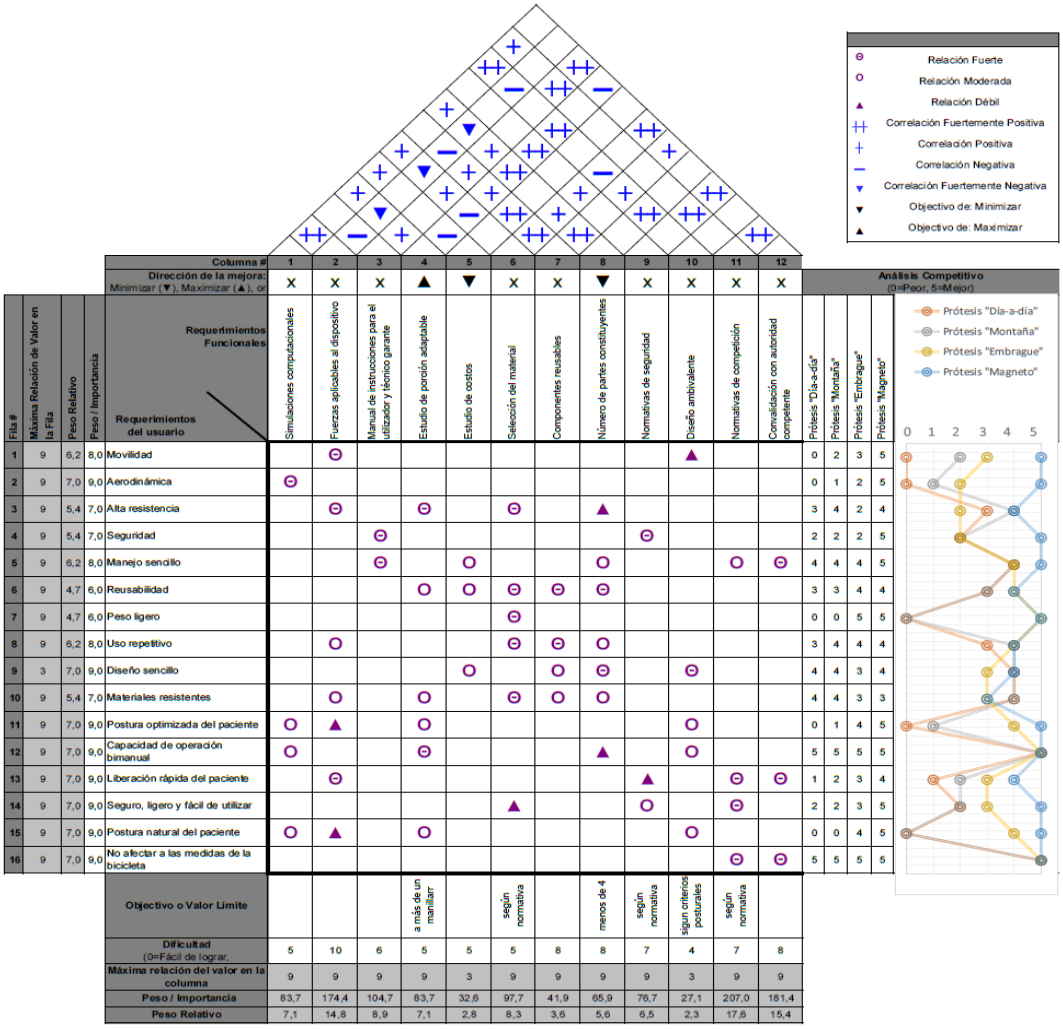
En la parte superior de la casa se pueden observar correlaciones negativas entre los requerimientos funcionales. Las solicitaciones a las que se verán sometidas las piezas son un criterio que influye en gran parte para el resto de los requerimientos y esto se debe a la alta dependencia que puede esto tener para el estudio detallado por componentes.

La propuesta de un diseño ambivalente y el estudio de porción adaptable también causa conflictos con otros requerimientos.

6. Prototipo propuesto

La solución presenta un sistema de acople encaje/manubrio ajustado a las necesidades del desempeño deportivo del usuario con amputación transhumeral y modelado a través de herramientas de diseño asistido por ordenador. Un esquema se presenta en la Figura 2.

Se considera un dispositivo de grado médico clase I, por no ser un producto invasivo [2] que permite el manejo de una bicicleta de competición en pruebas de pista. Su



Fila #	Máxima Relación de Valor en la Fila	Peso Relativo	Peso / Importancia	Requerimientos del usuario	Simulaciones computacionales	Fuerzas aplicables al dispositivo	Manual de instrucciones para el utilizador y técnico garante	Estudio de porción adaptable	Estudio de costos	Selección del material	Componentes reusables	Número de partes constituyentes	Normativas de seguridad	Diseño ambivalente	Normativas de competición	Convalidación con autoridad competente	Prótesis "Día-a-día"	Prótesis "Montaña"	Prótesis "Embrague"	Prótesis "Magneto"
				Columna #	1	2	3	4	5	6	7	8	9	10	11	12				
				Dirección de la mejora: Minimizar (▼), Maximizar (▲), or	x	x	x	▲	▼	x	x	▼	x	x	x	x				
1	9	6,2	8,0	Movilidad		⊖								▲			0	2	3	5
2	9	7,0	9,0	Aerodinámica	⊖												0	1	2	5
3	9	5,4	7,0	Alta resistencia		⊖		⊖		⊖		▲					3	4	2	4
4	9	5,4	7,0	Seguridad			⊖						⊖				2	2	2	5
5	9	6,2	8,0	Manejo sencillo			⊖		○			○			○	⊖	4	4	4	5
6	9	4,7	6,0	Reusabilidad				○	○	⊖	⊖	⊖					3	3	4	4
7	9	4,7	6,0	Peso ligero						⊖							0	0	5	5
8	9	6,2	8,0	Uso repetitivo		○				⊖	⊖	○					3	4	4	4
9	3	7,0	9,0	Diseño sencillo					○		○	⊖		⊖			4	4	3	4
10	9	5,4	7,0	Materiales resistentes		○		○		⊖	○	○					4	4	3	3
11	9	7,0	9,0	Postura optimizada del paciente	○	▲		○						○			0	1	4	5
12	9	7,0	9,0	Capacidad de operación bimanual	○			⊖				▲		○			5	5	5	5
13	9	7,0	9,0	Liberación rápida del paciente		⊖							▲		⊖	⊖	1	2	3	4
14	9	7,0	9,0	Seguro, ligero y fácil de utilizar						▲			○		⊖		2	2	3	5
15	9	7,0	9,0	Postura natural del paciente	○	▲		○						○			0	0	4	5
16	9	7,0	9,0	No afectar a las medidas de la bicicleta											⊖	⊖	5	5	5	5
				Objectivo o Valor Límite				a más de un manillar		según normativa		menos de 4	según normativa	según criterios posturales	según normativa					
				Dificultad (0=Fácil de lograr,	5	10	6	5	5	5	8	8	7	4	7	8				
				Máxima relación del valor en la columna	9	9	9	9	3	9	9	9	9	3	9	9				
				Peso / Importancia	83,7	174,4	104,7	83,7	32,6	97,7	41,9	65,9	76,7	27,1	207,0	181,4				
				Peso Relativo	7,1	14,8	8,9	7,1	2,8	8,3	3,6	5,6	6,5	2,3	17,6	15,4				

Figura 2. *Gráfico QFD del producto propuesto*

funcionamiento ha sido pensado para permitir la unión segura del usuario al manillar sin perder la capacidad de liberación cuando alcance un límite de posible riesgo. También permite mantener una postura adecuada del usuario y su total libertad de movimientos durante la prueba [9].

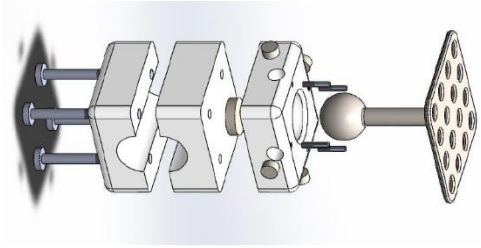


Figura 2. *Vista explosionada del conjunto*

La Figura 2 presenta las componentes del sistema: La pieza A debe ser fijada al encaje/miembro residual del usuario, y permitirá el acople al manillar a través de la pieza B. Esta segunda pieza incluye insertos de imanes que facilitarán el continuo contacto con la pieza A. Seguidamente las piezas C1 y C2, son utilizadas como pieza de fijación modular al manillar de la bicicleta y soporte fijo para la pieza B.

El ejercicio de un primer dimensionamiento de las piezas y selección de los materiales cumplió en primer lugar las restricciones establecidas en la fase de levantamiento de requerimientos, incluyendo talla y posición de los usuarios, aspectos normativos de la UCI, reglas vigentes a nivel de la UE y EE. UU., relacionadas con el deporte adaptado y, más concretamente, de ciclismo de competición, que determinan las exigencias para la práctica segura y reglamentada de este deporte. En segundo lugar, se adaptaron los primeros resultados al filtro establecido por el análisis QFD. Finalmente, para la elaboración de un primer prototipo de prueba funcional, se verificó su resistencia mecánica bajo un modelo de análisis estático equivalente a la energía de deformación elástica que absorbería ante una caída libre desde la altura en que se posiciona el manillar en su posición de equilibrio.

Para la fase de prototipado y evaluación de la respuesta del usuario ante su uso y manipulación se seleccionó un voluntario adulto de prueba. El individuo disponía de una prótesis no adaptada la cual utilizaba en la práctica deportiva, y se tuvo en consideración como modelo a mejorar al principio del estudio. La pieza A o componente distal del encaje se fabricó a partir de una placa perforada, barra y esfera de aceros soldadas. Para el acople con la bicicleta, una pieza recortada con perfil de encaje de esfera disponible en el taller se incrustó en las piezas C1 y C2 mecanizadas a partir de una barra de nylon. Finalmente, para magnetizar la conexión de las piezas A y B se colocaron 4 insertos imantados alrededor del orificio de acople y uno más justo en la zona de contacto del acoplamiento de A con B.

El tope de seguridad establecido para el sistema ensamblado se basa en la capacidad de liberación del dispositivo distal de la porción semiesférica de la pieza de acople en posturas en las cuales la pieza A se encuentre a aproximadamente a 30º o menos del plano medial de la semiesfera, independientemente del lado del cuerpo al que se adapte; estos valores se analizaron con el usuario durante su prueba.

La capacidad adaptativa de las piezas C1 y C2 es aún muy reducida en esta primera versión y serán atendidas próximamente. Sin embargo, de modo apreciativo en general, el usuario de prueba manifestó una buena receptividad del sistema de acople, expresando estabilidad durante su uso y seguridad al momento de desconectarse del manillar. Además, las pruebas también reportaron comodidad y una mejor postura en la bicicleta en movimiento. Estos primeros resultados manifiestan una viabilidad del diseño y el avance para nuevas fases en la ingeniería de detalle.

El enfoque de diseño bajo ingeniería concurrente permitió en una primera prueba de verificación establecer detalles referidos a las dimensiones, funcionamiento y funcionalidad del sistema e incluso adelantar propuestas para la selección de materiales y procedimientos de manufactura considerando desde un principio las necesidades y deseos del usuario con muy pocos cambios.

Referencias

[1] Lusardi MM, Jorge M, Nielsen C. Orthotics & Prosthetics in Rehabilitation. Elsevier, 2013 (ISBN: 9781437719369).

[2] Walter Reed Army Medical Center Borden Institute, Lenhart M, Pasquina P, Cooper R (Eds.) *Care Combat Amputee*. 2010 (ISBN: 0160840775).

[3] Deijs m, Bongers RM, Ringeling - van Leusen NDM, van der Sluis CK. Flexible and static wrist units in upper limb prosthesis users: functionality scores, user satisfaction and compensatory movements. *Journal of NeuroEngineering and Rehabilitation*, vol 13, 2016, pp 13-26 (ISSN 1743-0003).

[4] Bragaru M, Dekker R, Geertzen J HB. Sport prostheses and prosthetic adaptations for the upper and lower limb amputees: an overview of peer reviewed literature. *Prosthetics and Orthotics International*, vol 36(3), 2012, pp 290-296 (ISSN: 0309-3646).

[5] Tiele A, Soni-Sadar S, Rowbottom J, Patel S, Mathewson E, Pearson S, Hutchins D, Head J, Hutchins S. Design and Development of a Novel Upper-Limb Cycling Prosthesis. *Bioengineering*, vol 4(4), 2017, pp 89-100 (ISSN: 2306-5354).

[6] Dyer B. The importance of aerodynamics for prosthetic limb design used by competitive cyclists with an amputation: An introduction. *Prosthetics and Orthotics International*, vol 39(3), 2014, pp 232–237 (ISSN: 0309-3646).

[7] Union Cycliste Internationale - UCI Cycling Regulations. http://www.uci.ch/mm/Document/News/Rulesandregulation/16/80/73/1-GEN-20160101-E_English.pdf (Consultada: Octubre 2020).

[8] Soden PD, Adeyefa BA. Forces applied to a bicycle during normal cycling. *Journal of Biomechanics*, vol 12(7), 1979, pp 527–41 (ISSN: 0021-9290).

[9] Atlas of Limb Prosthetics: Surgical, Prosthetic, and Rehabilitation Principles. http://www.oandplibrary.org/alp/chap01-01.asp (Consultada: Octubre 2020).

Compensación de radiación dispersa en radiografía digital a través del aprendizaje automático: resultados preliminares

N. Sakaltras[1,2], F. A. Tovar[1], C. Martinez[1,2], C. F. Del Cerro[1,2], M. Desco [1,2,3,4], M. Abella [1,2,4]

[1] Dept. Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, España

[2] Instituto de Investigación Sanitaria Gregorio Marañón, Madrid

[3] Centro de investigación en red en salud mental (CIBERSAM), Madrid

[4] Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, España

Resumen

La dispersión de los rayos X reduce significativamente la resolución de contraste de la imagen en radiografía digital de tórax. La estrategia convencional para la reducción de la radiación dispersa es el uso de rejillas antidifusoras que, aunque mejoran la calidad de la imagen, aumentan la dosis de radiación absorbida por el paciente y plantean problemas en técnicas no estándar. En este trabajo, proponemos un método de corrección de la radiación dispersa basado en técnicas de aprendizaje profundo, que adopta una red neuronal convolucional de arquitectura U-net con 4 bloques tanto en el codificador como en el decodificador. Debido a la falta de pares de adquisiciones reales con y sin rejilla antidifusoras, se realizaron simulaciones de Monte Carlo para generar los datos de entrenamiento. El presente estudio demuestra el potencial del método propuesto, con un error inferior al 5%.

1. Introducción

En radiología digital de tórax, la radiación dispersa reduce significativamente la calidad de la imagen en términos de disminución de la resolución de contraste, limitando seriamente el uso clínico de la imagen. El uso de rejillas antidifusoras [1], actual estándar en el entorno clínico, permite la eliminación de una parte significativa de los fotones dispersos que llegan al detector a costa de un aumento de dosis de entre 3 y 6 veces. Además, cada rejilla está optimizada para una distancia y posición específicas entre la fuente y el detector, lo que dificulta su uso para otras geometrías de adquisición, como en tomosíntesis [2] o en radiografía portátil.

La mayoría de los trabajos para la corrección de radiación dispersa se centran en la tomografía computarizada (TAC), donde se dispone de información 3D de la muestra. Enfoques basados en hardware como en los trabajos propuestos en [2-4] hacen uso de un *Beam-Stopper* (BS), una placa con agujeros de un material con un alto coeficiente de atenuación a los rayos X, que se coloca entre la fuente de rayos X y el paciente, estimándose el campo de radiación dispersa a partir de la región sombreada, que se sustraerá de los datos completos. Alternativamente, en los métodos basados en software como [5, 6], los autores proponen el uso de convoluciones con *kernels* de distinto tamaño y valor, basándose en la observación de que la radiación de dispersión se puede aproximar con una versión paso bajo de la radiación total (primaria + dispersa). La optimización del tamaño, forma y valor de los *kernels* se realiza adquiriendo maniquíes sencillos. Sin embargo, al no simular las estructuras complejas del cuerpo humano, la búsqueda de los *kernels* no es óptima, lo que impide la obtención de un mapa de radiación disperso perfecto. En radiografía plana, también se ha utilizado la idea del BS, basada en una corrección hardware, para la estimación de la radiación dispersa. En [4], se realizan dos adquisiciones, con y sin el BS, estimando el componente de radiación dispersa del objeto y restándolo de la imagen total. Otros trabajos utilizan una adquisición con un BS detrás del detector para adquirir la señal de dispersión [7], pero a costa de disminuir la señal a ruido de la imagen.

Los métodos software más comunes para la corrección de radiación dispersa en radiografía plana, al igual que en TAC, se basan en convolucionar la imagen original para obtener el mapa de radiación dispersa y substraerlo [8, 9]. Sin embargo, estos métodos no son capaces de eliminar completamente la radiación dispersa sin reducir la resolución de la imagen original. En los últimos años se ha sustituido la utilización de convoluciones por el uso de algoritmos de optimización basados en la norma L1 [10, 11] para preservar los bordes y mejorar el mapa de radiación dispersa obtenido. El más interesante es el propuesto por Kim et al. [12], en el que se consigue una mejora sustancial del contraste. Sin embargo, el esquema de corrección necesita de un parámetro que da mayor o menor peso según los valores de atenuación. En imagen de tórax, en los que el rango dinámico de la imagen es tan amplio, desde valores muy bajos en los pulmones a valores muy altos en la zona columna, no es posible obtener un parámetro óptimo para toda la imagen.

Los recientes avances en el campo de la inteligencia artificial y el desarrollo de las unidades de procesamiento gráfico (GPU) han proporcionado nuevas herramientas para abordar la corrección de la radiación dispersa. En TAC, varios autores [13, 14] utilizan redes neuronales convolucionales para estimar el componente de radiación dispersa y eliminarlo de las proyecciones originales. En [13], Maier et al. proponen el método *Deep Scatter Estimation* (DSE), que usa la arquitectura U-net para obtener el mapa de radiación dispersa y muestra resultados prometedores. En este trabajo, se presenta una aproximación más directa para la corrección de la radiación dispersa basada en el uso de la arquitectura U-net. En lugar

ISBN: 978-84-09-25491-0, pags. 73- 76

de estimar primero la radiación dispersa y restarla de la adquisición original, como en DSE [13], se obtiene directamente la imagen de radiación primaria. Los resultados preliminares muestran mejoras significativas con respecto DSE y el método que usa una optimización basada en la norma L1 de Kim et al. [12].

2. Métodos

2.1. Descripción de la red

La red utilizada es una modificación del modelo U-Net propuesto por Ronneberger et al. [15], que se compone de tres bloques principales: el codificador, el cuello de botella y el decodificador. El codificador está formado por cuatro bloques compuestos por dos capas convolucionales con un tamaño de *kernel* de 3×3, una capa de ReLU y una capa de 2×2 *max-pooling* (no se utiliza la normalización de cada conjunto ya que los datos de entrada tienen el mismo rango de valores [16]). Después de cada bloque, el número de filtros se duplica para compensar la pérdida de resolución espacial causada por las capas de *max-pooling* y aprender características más complejas. Después de la última capa de *max-pooling* del codificador, se forma el cuello de botella de la U-net con una doble capa convolucional. El decodificador está formado por cuatro bloques compuestos por una capa 2×2 de *up-sampling* concatenada con el correspondiente mapa de características de la parte codificadora y una doble capa de convolución seguida de una activación ReLU. La idea principal del decodificador es recuperar la resolución espacial original utilizando la información contextual extraída en la parte del codificador. La arquitectura U-net utilizada se muestra en la Figura 1.

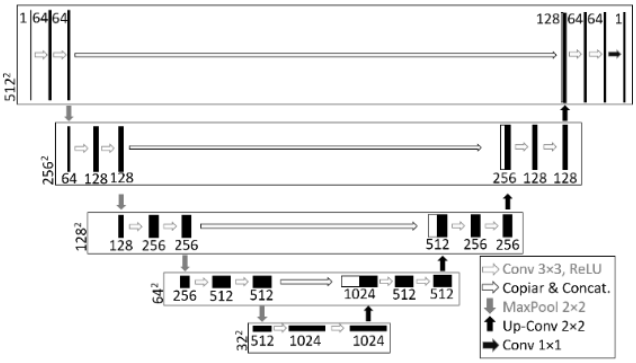


Figura 1. *La arquitectura de U-net.*

El número de filtros iniciales utilizados es 64, duplicándose en cada bloque subsiguiente. Para el entrenamiento de la red se usa un tamaño de lote (*batch-size*) de 8 imágenes, el optimizador Adam con una tasa de aprendizaje de 10^{-4}, y un número de épocas de 125. La función de coste utilizada durante el entrenamiento fue el error cuadrático medio (ECM).

2.2. Conjunto de datos y simulaciones

Las simulaciones se han hecho con MC-GPU v1.5b [17], basadas en la geometría estándar de adquisición de tórax, con un espectro de 120 kVp y un filtro de aluminio de 3 mm, obtenidos con la herramienta Spektr [18]. La distancia fuente-detector (DFD) se fijó en 1800 mm, con una separación entre paciente y el detector de 50 mm. El tamaño del detector es de 512×512 píxeles con un tamaño de píxel de 0,1 mm×0,1 mm.

Las simulaciones se basan en cinco volúmenes de TAC de tórax descargados del *National Biomedical Imaging Archive* (NBIA). Estos volúmenes se convirtieron de unidades Hounsfield (HU) a mapas de densidad y se segmentaron en tres tejidos con valores de densidad de 0,001205 g/cm3 (aire), 1,060 g/cm3 (tejidos blandos) y 1,920 g/cm3 (hueso) según la base de datos del *National Institute of Standard and Technology* (NIST) [19]. Un ejemplo de esta segmentación se muestra en la Figura 2.

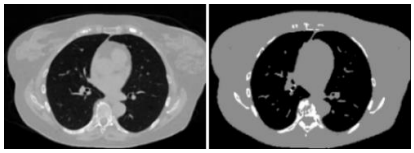


Figura 2. *Corte axial de un paciente (izquierda) y segmentación del mismo corte con las correspondientes densidades asociadas (derecha).*

Para el entrenamiento y evaluación de la red se generaron 82 proyecciones de cada volumen cubriendo un rango de ±20 grados, en pasos de un grado, alrededor de las posiciones radiológicas antero-posterior (AP) y postero-anterior (PA). En total se generaron 410 proyecciones, 328 para el conjunto de entrenamiento y 82 para el conjunto de evaluación. Para cada proyección, se obtuvieron los componentes de radiación primaria, radiación dispersa y Rayleigh. A partir de estas componentes se crearon dos imágenes: total (radiación primaria + Rayleigh + radiación dispersa) y referencia (radiación primaria + Rayleigh). La radiación producida por el efecto Rayleigh se añadió a la radiación primaria para formar la imagen referencia, debido a que los ángulos de desviación son mínimos y consiguen pasar a través de las rejillas antidifusoras. En la Figura 3 se muestra un ejemplo de las imágenes de salida.



Figura 3. *Imagen total (izquierda), imagen referencia (centro) y radiación dispersa (derecha).*

La Figura 4 muestra la evolución del error en los conjuntos de entrenamiento y validación, que indica que la red está aprendiendo y que el número de muestras utilizadas son suficientes para el estudio preliminar llevado a cabo.

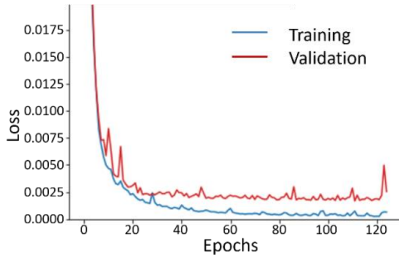


Figura 4. Errores de entrenamiento y validación de la red propuesta.

3. Evaluación

El método propuesto se comparó, usando la proyección AP central del volumen de evaluación, con dos métodos recientes: DSE [13], basado también en redes neuronales, y el método propuesto por Kim et al. [12], basado en la optimización de la norma L1. Para el método de post-procesamiento de Kim et al., usamos dos valores diferentes para el parámetro de peso, $A_{Pulmones}$= 1 and $A_{Columna}$= 2.5, para las regiones de los pulmones y la columna vertebral, respectivamente, resultando en dos imágenes corregidas diferentes. Como medidas cuantitativas se obtuvieron perfiles y la raíz del error cuadrático medio (RECM) entre las imágenes corregidas y la referencia en regiones de interés situadas en los pulmones y la columna vertebral (Figura 5).

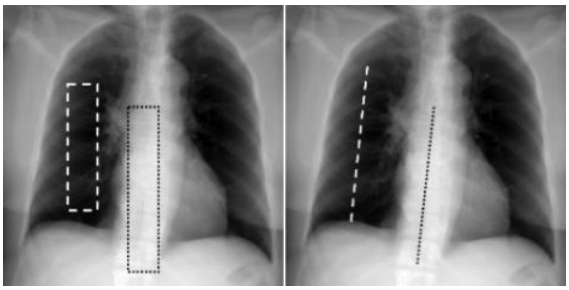


Figura 5. *ROI para el cálculo del RECM (izquierda) y los perfiles elegidos para la evaluación (derecha).*

4. Resultados

En la Figura 6, podemos ver un aumento de la resolución de contraste al utilizar el método propuesto, especialmente en la región de la columna vertebral, similar a la imagen referencia. También los bronquios se vuelven más visibles. De manera similar, el método DSE recupera el contraste en ambas regiones. El método de Kim et al. recupera el contraste de los pulmones y la columna vertebral, sin embargo, no es posible mejorar el contraste de ambas regiones en una misma imagen.

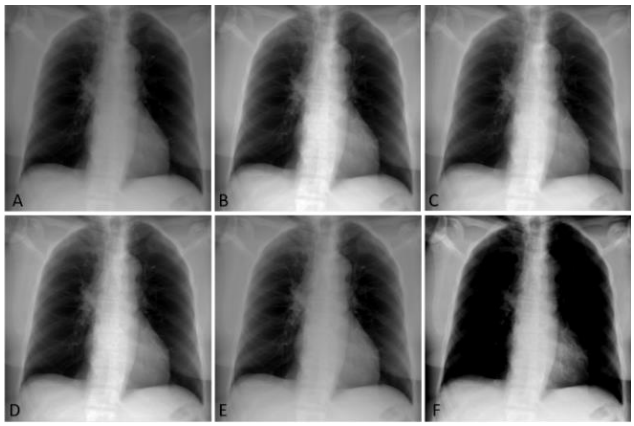


Figura 6. *Imagen total (A), imagen referencia (B), imagen corregida con el método propuesto (C), imagen corregida con DSE (D), imagen corregida con el método de Kim et al. optimizado para los pulmones (E) e imagen corregida con el método de Kim et al. optimizado para la columna (F).*

Estas diferencias pueden verse de forma más cuantitativa en los perfiles de que se muestran en la Figura 7.

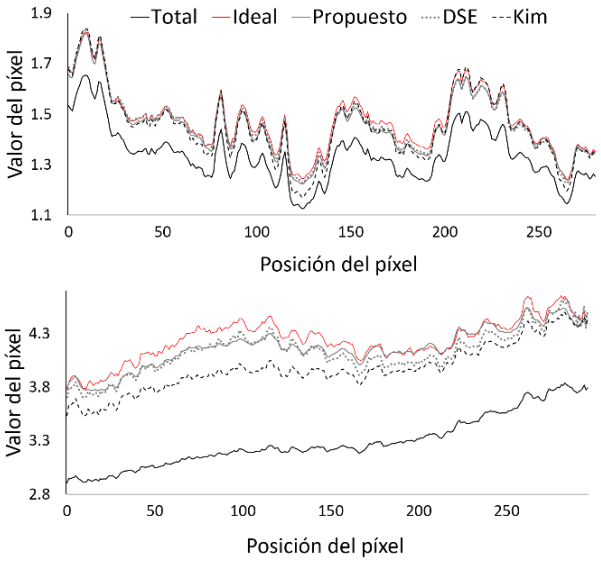


Figura 7. *Perfiles trazados en los pulmones (arriba), y la columna vertebral (abajo).*

La Tabla 1 muestra la RECM de todos los métodos en comparación con la imagen referencia, donde se puede ver que el método propuesto resulta en el menor error, especialmente en la región de los pulmones.

RECM	Método propuesto	DSE	Kim et al.
Pulmones	0.009	0.036	0.093
Columna	0.066	0.082	0.214

Tabla 1. *Errores de los distintos métodos en comparación con la referencia.*

5. Discusión y conclusiones

En este trabajo, se presenta un método de corrección de radiación dispersa en radiografía digital de tórax basado en la estimación de la radiación primaria a través de redes convolucionales con arquitectura U-net. Los resultados preliminares, en datos simulados, muestran una recuperación de la resolución de contraste similar al resultado de usar una rejilla antidifusora. La arquitectura U-net utilizada no incorpora las normalizaciones por conjuntos de la original propuesta por Ronneberger et al. [15], eliminando un paso adicional que consume tiempo, debido a que los valores de nuestro conjunto de datos se encuentran en el mismo rango.

Comparando con el método de software propuesto por Kim et al., nuestro método logró una mejor corrección de la radiación dispersa sin necesidad de elegir una región específica, como los pulmones o la columna vertebral. El método propuesto mostró resultados similares que el método DSE, pero con la ventaja de la estimación directa de la radiación primaria, que permite un menor error de los valores obtenidos en la zona de los pulmones y la columna.

Las imágenes de radiografía plana de tórax tienen un tamaño superior a 3000×3000 píxeles con un tamaño de píxel en el detector de alrededor de 0.1 mm. Sin embargo, en este trabajo hemos utilizado imágenes de 512×512 pixeles, debido a las limitaciones actuales en la memoria

de la GPU. En futuros trabajos se hará uso de imágenes con el tamaño original, lo que podría lograrse mediante el uso de parches de la imagen para entrenar la red. Por otra parte, la red neuronal se ha entrenado con imágenes simuladas con un voltaje de fuente y una geometría de adquisición específicos. El trabajo futuro también incluye el entrenamiento con un conjunto de datos más grande y diverso, lo que podría conducir a la optimización del número de capas y filtros, reduciendo los tiempos de entrenamiento. Por último, actualmente la red se entrena con un ritmo de aprendizaje constante, siendo posible la utilización de una tasa de aprendizaje variable a lo largo de las épocas [20] en un futuro. Para reducir el error también se considerará el uso de funciones de pérdida más sofisticadas como *perceptual loss* [21] o la inclusión de arquitecturas basadas en *Generative Adversarial Networks* (GAN) [22].

6. Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia, Innovación y Universidades (Instituto de Salud Carlos III, proyecto DTS17/00122; Agencia Estatal de Investigación, proyecto DPI2016-79075-R-AEI/FEDER, UE), cofinanciado por Fondos de la Unión Europea (FEDER), "A way of making Europe". Además, ha sido financiado por el Programa de apoyo a la realización de proyectos interdisciplinares de I+D para jóvenes investigadores de la Universidad Carlos III de Madrid 2019-2020 en el marco del Convenio Plurianual Comunidad de Madrid- Universidad Carlos III de Madrid (proyecto DEEPCT-CM-UC3M). El CNIC está financiado por el Ministerio de Ciencia, Innovación y Universidades y la fundación PRO-CNIC y es un centro de excelencia Severo Ochoa (SEV-2015-0505).

7. Referencias

1. Neitzel, U., *Grids or air gaps for scatter reduction in digital radiography: a model calculation.* Medical physics, 1992. **19**(2): p. 475-481.
2. Dong, X., et al., *Low-dose and scatter-free cone-beam CT imaging using a stationary beam blocker in a single scan: phantom studies.* Computational and mathematical methods in medicine, 2013. **2013**.
3. Maltz, J.S., et al., *Focused beam-stop array for the measurement of scatter in megavoltage portal and cone beam CT imaging.* Med Phys, 2008. **35**(6): p. 2452-62.
4. Sakaltras, N., et al., *Scatter Correction for planar radiography with a Beam-hole array: preliminary results*, in *IEEE Nuclear Science Symposium & Medical Imaging Conference (NSS/MIC)*. 2019: Manchester.
5. Star-Lack, J., et al., *Efficient scatter correction using asymmetric kernels.* 2009. **7258**: p. 72581Z.
6. Sun, M. and J.M. Star-Lack, *Improved scatter correction using adaptive scatter kernel superposition.* Phys Med Biol, 2010. **55**(22): p. 6695-720.
7. Lo, J.Y., et al., *Scatter compensation in digital chest radiography using the posterior beam stop technique.* Medical physics, 1994. **21**(3): p. 435-443.
8. Kruger, D.G., et al., *A regional convolution kernel algorithm for scatter correction in dual-energy images: Comparison to single-kernel algorithms.* Medical physics, 1994. **21**(2): p. 175-184.
9. Love, L.A. and R.A. Kruger, *Scatter estimation for a digital radiographic system using convolution filtering.* Med Phys, 1987. **14**(2): p. 178-85.
10. Kim, K., et al., *A new software scheme for scatter correction based on a simple radiographic scattering model.* Medical & biological engineering & computing, 2019. **57**(2): p. 489-503.
11. Meng, G., et al. *Efficient image dehazing with boundary constraint and contextual regularization*. in *Proceedings of the IEEE international conference on computer vision*. 2013.
12. Kim, K.S., et al., *Improvement of radiographic visibility using an image restoration method based on a simple radiographic scattering model for x-ray nondestructive testing.* NDT & E International, 2018. **98**: p. 117-122.
13. Maier, J., et al., *Deep scatter estimation (DSE): Accurate real-time scatter estimation for X-ray CT using a deep convolutional neural network.* Journal of Nondestructive Evaluation, 2018. **37**(3): p. 57.
14. Nomura, Y., et al., *Projection-domain scatter correction for cone beam computed tomography using a residual convolutional neural network.* Medical physics, 2019. **46**(7): p. 3142-3155.
15. Ronneberger, O., P. Fischer, and T. Brox. *U-net: Convolutional networks for biomedical image segmentation.* in *International Conference on Medical image computing and computer-assisted intervention*. 2015. Springer.
16. Ioffe, S. and C. Szegedy, *Batch normalization: Accelerating deep network training by reducing internal covariate shift.* arXiv preprint arXiv:1502.03167, 2015.
17. Badal, A. and A. Badano, *Accelerating Monte Carlo simulations of photon transport in a voxelized geometry using a massively parallel graphics processing unit.* Medical physics, 2009. **36**(11): p. 4878-4880.
18. Siewerdsen, J., et al., *Spektr: A computational tool for x-ray spectral analysis and imaging system optimization.* Medical physics, 2004. **31**(11): p. 3057-3067.
19. Hubbell, J. and S. Seltzer, *NIST standard reference database 126.* Gaithersburg, MD: National Institute of Standards and Technology, 1996.
20. Smith, L.N. *Cyclical learning rates for training neural networks.* in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 2017. IEEE.
21. Johnson, J., A. Alahi, and L. Fei-Fei. *Perceptual losses for real-time style transfer and super-resolution.* in *European conference on computer vision*. 2016. Springer.
22. Wolterink, J.M., et al., *Generative adversarial networks for noise reduction in low-dose CT.* IEEE transactions on medical imaging, 2017. **36**(12): p. 2536-2545.

Diseño e implementación de clasificadores 2D para la detección de la enfermedad de Alzheimer mediante el uso de redes neuronales convolucionales en imágenes de resonancia magnética.

I. Aleixandre Soriano[1], I. Bosch Roig[2], E. Camacho Ramos[3], J. Sánchez García[4]

[1] Estudiante, Universitat Politècnica de València, Valencia, España, iralso@etsii.upv.es

[2] Instituto de Telecomunicaciones y Aplicaciones Multimedia, Universitat Politècnica de València, Valencia, España, igbosroi@dcom.upv.es

[3] Científico de análisis de imágenes, QUIBIM, Valencia, España, educamacho@quibim.com

[4] Científico de datos, QUIBIM, Valencia, España, josesanchez@quibim.com

Resumen

La enfermedad de Alzheimer (AD) es el principal tipo de demencia a nivel mundial. Actualmente, no existe cura para dicha patología, pero sí tratamientos capaces de retrasar o paliar sus síntomas si es detectada a tiempo. Este estudio ha abordado el problema mediante el uso de redes neuronales convolucionales bidimensionales. Para ello, se han aplicado técnicas de pre-procesado de imagen médica con el objetivo de desarrollar herramientas capaces de diferenciar, a partir de imágenes de resonancia magnética cerebral, entre muestras patológicas y no patológicas. Adicionalmente, también se ha conseguido diferenciar entre muestras sanas (CN), con deterioro cognitivo leve (MCI) y AD. Los resultados demuestran que es posible abordar el problema utilizando imágenes 2D reduciendo así el coste temporal y computacional. El principal elemento de confusión se da para muestras con MCI, pero, pese a ello, los valores de exactitud obtenidos han sido superiores al 90%.

1. Introducción

La enfermedad de Alzheimer (AD, por sus siglas en inglés) es una patología crónica de naturaleza degenerativa y evolución progresiva. Se caracteriza neuropatológicamente por la aparición de depósitos de la proteína β-amiloide extracelulares y ovillos neurofibrilares compuestos por la proteína τ hiperfosforilada en la zona intracelular. [1]

En etapas tempranas, ataca al sistema límbico, siendo la estructura del hipocampo la mayormente afectada. Posteriormente ataca a la corteza cerebral siguiendo una trayectoria temporal-parietal-frontal. En etapas avanzadas, puede llegar a afectar el tronco encefálico. [2-4]

Todos estos cambios conllevan una disminución del volumen o atrofia de las estructuras mencionadas, dejando espacio libre que debe ser compensado para preservar la presión intracraneal. Por este motivo, también se observa un aumento del volumen de líquido cefalorraquídeo (LCR) sobre todo en los ventrículos. [5], [6]

Los cambios anteriormente mencionados son apreciables en neuroimagen estructural. Esta técnica ofrece una visión estática del cerebro aportando información meramente anatómica. Para la realización de este estudio se han empleado imágenes de resonancia magnética (RM) estructural potenciadas en T1, donde T1 largos indican mayor dificultad en la liberación energética implicando una disminución de intensidad. [7]

Actualmente, la mayor parte de las investigaciones se centran en el uso de innovadoras técnicas basadas en inteligencia artificial (IA). Más concretamente, en técnicas de aprendizaje profundo centradas en redes neuronales convolucionales (CNN, por sus siglas en inglés) tridimensionales, capaces de procesar directamente los volúmenes resultado de la RM. Las CNN son un tipo determinado de red neuronal artificial dentro del campo del aprendizaje supervisado. Se trata de algoritmos en los cuales las transformaciones se realizan utilizando la operación de convolución. [8]

Las técnicas de pre-procesado más comunes son, por su parte, el *skull stripping* [9-11], la normalización de intensidades [9-11] y la aplicación de filtros de reducción de ruido. [9], [11], [12]

En cuanto a bases de datos se refiere, la mayormente utilizada es ADNI (*Alzheimer's Disease Neuroimaging Initiative*). En esta base de datos se pueden encontrar datos e imágenes de RM estructural de 2.562 sujetos, clasificados como cognitivos normales (CN), con deterioro cognitivo leve (MCI, por sus siglas en inglés) o diagnosticados con Alzheimer (AD).

La motivación principal de este trabajo es, por tanto, la intención de ayudar a los pacientes de AD y también al sistema sanitario, ya que se desconoce si tal y como está diseñado hoy en día podrá afrontar las situaciones futuras.

El objetivo fue desarrollar diferentes clasificadores a partir de CNN 2D, haciendo uso de la base de datos ADNI y considerando distintas técnicas de pre-procesado.

2. Materiales y métodos

En este estudio se ha utilizado un subconjunto de 150 imágenes procedentes de la base de datos ADNI, en formato NIfTI de RM cerebral potenciadas en T1 adquiridas con un equipo de 3-T.

Los experimentos iniciales han considerado los 3 grupos de diagnóstico (CN, MCI, AD) teniendo en cuenta los 150 sujetos anteriormente mencionados (ver tabla 1).

ISBN: 978-84-09-25491-0, pags. 77- 80

Etiquetas	Sujetos
CN	46
MCI	71
AD	33
Total	150

Tabla 1. *Clasificación multiclase. Número de sujetos asociados a cada etiqueta.*

Posteriormente, se han realizado tres experimentos finales: un clasificador binario, un clasificador multiclase y un clasificador multiclase en cascada.

Para la clasificación binaria las etiquetas asignadas han sido Normal y Anormal. La primera de ellas recoge aquellas muestras sanas (CN) mientas que la segunda abarca los grupos MCI y AD (ver tabla 2).

Etiquetas	Sujetos
Normal	46
Anormal	104
Total	150

Tabla 2. *Clasificación binaria. Número de sujetos asociados a cada etiqueta.*

Para la aplicación de las técnicas de pre-procesado de imagen médica se ha hecho uso de MatLab como lenguaje de programación junto con el *software* SPM (*Statistical Parametric Mapping*).

En cuanto al desarrollo del clasificador se ha trabajado con la librería de código abierto *Fastai* para aprendizaje profundo construida sobre *PyTorch* en *Python 3*. Esta librería cuenta con modelos pre entrenados de *torchvision*, un paquete que incluye datos populares, arquitecturas de modelos y transformaciones comunes de imágenes.

Con el objetivo de evaluar la calidad del clasificador se emplearon 7 métricas. [13]

- Matriz de confusión. Contiene tantas filas (valores reales) y columnas (valores predichos) como clases. Permite conocer los verdaderos positivos (VP), verdaderos negativos (VN), falsos positivos (FP) y falsos negativos (FN).
- Exactitud. Relación entre el número de predicciones correctas y el número total de muestras de entrada.
- Precisión. Razón entre los VP y el total de valores reales positivos.
- Sensibilidad. VP entre la suma de VP y FN.
- Especificidad. Cociente entre VP y la suma de VP y FN.
- F1 *score*. Cociente entre el producto y la suma de la precisión y sensibilidad, multiplicado por ambos.
- Área bajo la curva ROC (AUC). Probabilidad de que el clasificador sea capaz de clasificar una muestra positiva elegida al azar con una probabilidad mayor que una muestra negativa escogida al azar.

2.1. Pre-procesamiento

En este estudio se ha trabajado con CNN bidimensionales. Por tanto, el primer paso ha sido obtener imágenes 2D y, a continuación, se han aplicado distintas técnicas de pre-procesado con el objetivo de mejorar los resultados.

Para la realización de los experimentos las imágenes 2D fueron divididas en tres subconjuntos: entrenamiento (70%), validación (20%) y *test* (10%). En todo caso se han balanceado previamente de los datos de entrada, quedando el número de observaciones determinado según el número de imágenes del grupo de menor tamaño.

a) Creación de imágenes 2D

La primera etapa del pre-procesado consistió en la transformación de cada uno de los volúmenes 3D NIfTI en imágenes 2D correspondientes a cortes axiales de dichos volúmenes, de forma que todas las imágenes de un mismo sujeto contuviesen la etiqueta asociada al volumen 3D de éste. Con ello, se obtuvo un total de 37.168 imágenes 2D.

b) Zero padding

Seguidamente, se realizó un redimensionado a 256x256 píxeles para conseguir un tamaño de imagen cuadrada. Con el objetivo de no alterar las estructuras anatómicas, se optó por realizar un relleno con ceros a izquierda y derecha (técnica *zero padding*) a las imágenes 2D.

c) Skull stripping

Haciendo uso de SPM12 se segmentaron los tres tejidos principales: sustancia gris, sustancia blanca y LRC. Seguidamente, se emplearon las tres máscaras obtenidas para extraer el parénquima cerebral del resto de elementos no deseados presentes en la imagen (principalmente cráneo, vasos sanguíneos, grasa y aire).

d) Normalización de intensidad

Las imágenes 2D presentaron rangos de intensidad distintos. Para solventar el problema, se normalizó la intensidad de todas las imágenes en el rango [0, 255] haciendo uso de la librería *OpenCV* disponible en *Python*.

e) Eliminación de cortes axiales con ausencia de tejido cerebral

De un estudio pormenorizado de las imágenes se observó gran cantidad de ellas con ausencia de tejido cerebral, lo cual aporta información contradictoria a la CNN. Tras su eliminación el resultado fue un total de 35.462 imágenes.

f) Eliminación de cortes axiales con una media de intensidad menor a 0.05.

Tras eliminar los cortes con ausencia de tejido cerebral se siguieron encontrando cortes irrelevantes que contenían muy poca información y no eran clasificables. Se estableció como umbral un valor medio de intensidad de imagen de 0.05, resultando en un total de 32.692 muestras.

g) Eliminación de cortes axiales sin información anatómica relevante para la detección de AD.

Se decidió considerar solo aquellos cortes axiales que mostraban las estructuras anatómicas cerebrales de interés

para la AD, eliminando los cortes axiales donde solo se observaban estructuras del lóbulo occipital o cerebelo.

Para ello se hizo uso del atlas *Harvard-Oxford*. En primer lugar, se normalizó espacialmente el atlas del espacio MNI al espacio nativo de cada sujeto. En segundo lugar, se generó una máscara binaria para identificar las estructuras anatómicas mencionadas. Para finalizar, se identificaron los cortes axiales máximo y mínimo de interés para cada sujeto. Se obtuvo un total de 20.703 muestras (ver tabla 3).

Etiqueta	Número de imágenes
CN	6.378
MCI	9.907
AD	4.418
Normal	6.387
Anormal	14.325
Total	20.703

Tabla 3. *Clasificador multiclase. Número de imágenes asociadas a cada etiqueta.*

3. Resultados y discusión.

En primer lugar, se llevaron a cabo una serie de experimentos iniciales, de 5 *epochs* de duración manteniendo las primeras capas congeladas. Cabe recordar el número de *epochs* se define como el número de veces que el algoritmo de aprendizaje funciona con todos los datos del conjunto de entrenamiento. Se utilizaron la arquitectura CNN *Resnet34* con un *batch size* de 64 con todas las imágenes. Las conclusiones extraídas de cada uno de ellos fueron aplicadas en el experimento siguiente siempre buscando mejorar los resultados y extraer conclusiones aplicables a los clasificadores finales. Los resultados se pueden observar en la tabla 4.

Experimentos previos		Exactitud final
Zero padding	**Aplicado**	**0.8273**
	Sin aplicar	0.7390
Skull stripping	Aplicado	0.7022
	Sin aplicar	**0.7390**
Tamaño de imagen	**128x128**	**0.7850**
	128x128 → 256x256	0.7232
Normalización de intensidad	**Aplicado**	**0.8084**
	Sin aplicar	0.7850
Eliminación de cortes axiales con ausencia de tejido cerebral	**Aplicado**	**0.8255**
	Sin aplicar	0.8084
Eliminación de cortes axiales con una media de intensidad inferior a 0.05	**Aplicado**	**0.8369**
	Sin aplicar	0.8255
Eliminación de cortes axiales sin información anatómica relevante para la detección de AD	**Aplicado**	**0.8623**
	Sin aplicar	0.8369
Elección de la arquitectura CNN	ResNet34	0.8623
	ResNet50	0.9041
	SqueezeNet1_1	0.7574
	DenseNet121	**0.9451**
Elección del *batch size*	**64**	**0.9451**
	32	0.9331
	128	0.8466

Tabla 4*. Resumen de los valores de exactitud obtenidos para cada uno de los experimentos previos realizados.*

Se puede comprobar cómo las técnicas de pre-procesado de *zero padding*, normalización de intensidad y eliminación de los cortes axiales sin información anatómica relevante, permitieron obtener mejores resultados de clasificación. Sin embargo, el *skull stripping* no consiguió mejoras significativas. La arquitectura CNN que obtuvo mejores resultados fue *DenseNet121*, con un tamaño de imagen 128x128 y un *batch size* de 64.

Los subconjuntos de datos asociados al clasificador binario, tras el balanceo de los datos, se pueden observar en la tabla 5. El grupo minoritario sería el formado por la etiqueta *Normal* (ver tabla 3).

Subconjunto	Número de imágenes
Entrenamiento	9.184
Validación	2.296
Test	1.276
Total	12.756

Tabla 5. *Clasificador binario. Número de imágenes que forman cada subconjunto de datos.*

Se obtuvieron buenos resultados, con una exactitud del 90% y una AUC de 0.90. Además, se observa que el valor de especificidad conseguido es mayor, aproximadamente en un 10%, al valor de sensibilidad. Este hecho es importante, ya que prueba que el clasificador binario tiene una mayor capacidad para identificar correctamente muestras patológicas que no patológicas.

Sin embargo, el principal problema de este clasificador binario se asocia a la detección precoz de la AD. Sería preferible que fuese capaz de identificar el MCI. Por consiguiente, se llevaron a cabo los dos experimentos finales asociados a clasificadores multiclase. Los subconjuntos de datos asociados se pueden consultar en la tabla 6. En este caso, el grupo minoritario sería el correspondiente a la etiqueta AD (ver tabla 3).

Subconjunto	Número de imágenes
Entrenamiento	9.543
Validación	2.385
Test	1.326
Total	13.254

Tabla 6. *Clasificador multiclase. Número de imágenes que forman cada subconjunto de datos.*

Para evaluar cada una de las clases por separado se prepararon tres subconjuntos de *test* a partir del mostrado en la tabla 6, cada uno de ellos formado por 884 imágenes.

Se preparó también un clasificador en cascada donde cada muestra pasa por un primer clasificador (patológica vs. no patológica), de forma que si la muestra es clasificada como no patológica es etiquetada como CN. De no ser así, pasa por un segundo clasificador (MCI vs. AD) para diferenciar de qué tipo de patología se trata.

A continuación, se analizaron los resultados finales para los clasificadores multiclase, cuyos resultados se muestran de forma resumida en la tabla 7.

Modelo de *test*	CN vs MCI+AD		MCI vs CN+AD		AD vs CN+MCI	
Cascada	No	Sí	No	Sí	No	Sí
Exactitud	**0.9355**	0.9162	**0.9061**	0.8993	**0.9513**	0.9117
Precisión	**0.9529**	0.8898	**0.9761**	0.7259	**0.9271**	0.7455
Sensibilidad	0.9162	**0.9502**	0.8325	**0.9444**	**0.9796**	0.9723
Especificidad	**0.9547**	0.8823	**0.9796**	0.8847	**0.9230**	0.8920
F1 score	**0.9342**	0.9190	**0.8986**	0.8209	**0.9526**	0.8440
AUC	**0.9355**	0.9162	0.9061	**0.9145**	**0.9513**	0.9322

Tabla 7*. Resumen de las métricas obtenidas con los subconjuntos de datos de test para los clasificadores multiclase.*

De la tabla 7 se puede deducir cómo los valores de exactitud son mayores para el caso del clasificador sin cascada para todos los modelos. Esto mismo ocurre para las métricas de precisión, especificidad y f1 *score*, aunque no es así para el caso de la sensibilidad.

Inicialmente, puede parecer trascendental que los valores de sensibilidad sean menores que los de especificidad para dos de los subconjuntos de *test*, ya que indicaría que el clasificador es capaz de identificar correctamente en mayor medida muestras no patológicas que patológicas. Sin embargo, la preferencia por mejorar el valor de una u otra métrica depende de la finalidad de la herramienta desarrollada. Sí se puede valorar el balance entre ambas con la métrica f1 *score*, siendo ésta en todo caso superior para el clasificador multiclase sin cascada.

Otra muestra del buen funcionamiento del clasificador multiclase sin cascada son los valores del AUC, donde se puede observar que en todo caso son valores superiores a 0.9, mostrando que existe una probabilidad mayor de clasificar correctamente una muestra positiva elegida al azar que una negativa. Este valor es fundamental sobre todo a la hora de la detección de la clase CN, ya que es de especial importancia evitar que un paciente patológico sea identificado como no patológico. En la tabla 7 puede observarse que dicho valor es superior para el clasificador multiclase con respecto al clasificador multiclase en cascada y también con respecto al clasificador binario.

Finalmente, es necesario recordar que la mayor parte de los trabajos en nuestro ámbito de aplicación se focalizan en imágenes 3D. Sin embargo, existe un pequeño número de investigaciones que tienen en cuenta imágenes 2D. Cabe destacar uno de ellos [11], el cual hace uso de CNN 2D obteniendo los mejores resultados al considerar únicamente el corte axial 40, con una exactitud del 73%.

4. Conclusiones y líneas futuras

Durante la realización de este estudio se ha podido comprobar, por una parte, la gran cantidad de técnicas de pre-procesado de imagen médica que son aplicables a las imágenes de RM cerebral y también la gran labor que el campo de la IA puede aportar a la detección de la AD.

Los resultados muestran una nueva línea de investigación que implica el uso de imágenes 2D en los clasificadores. Esto reduce en gran medida el coste computacional obteniendo resultados superiores al 90% de exactitud. El principal problema de confusión en la clasificación de imágenes se da en sujetos que presentan MCI.

De forma habitual, un mayor número de observaciones, o en este caso imágenes, podrían mejorar los resultados ya que permiten conseguir un mejor ajuste de los pesos de las CNN. Del mismo modo, también podrían aplicarse técnicas de pre-procesado más complejas.

Sin embargo, existe la necesidad de abordar el problema desde un punto de vista tridimensional, ya que el objetivo de estos clasificadores no debería ser clasificar imágenes 2D si no a los propios sujetos. Una línea de estudio futura, por tanto, sería estudiar el número de cortes consecutivos y por sujeto que son clasificados bidimensionalmente como MCI o AD y, a partir de ello, establecer umbrales que permitiesen obtener una sola clasificación por sujeto.

Referencias

[1] Donoso A. La enfermedad de Alzheimer. Rev. Chil. Neuro-Psiquiatr., vol 41, 2003, pp 13-22 (ISSN: 0717-9227).

[2] Cambios del cerebro en la enfermedad del Alzheimer. https://www.dementiacarecentral.com/caregiverinfo/cambios-del-cerebro-en-la-enfermedad-del-alzheimer/ (Consultada: Mayo 2020).

[3] Rueda-O. A. del P, Enríquez-S. LF. A review of basic neuroimaging techniques in diagnosis of neurodegenerative diseases. Biosalud, vol 17, no 2, 2018 ,pp 59-90 (ISSN: 1675-9550).

[4] Arana Fernández de Moya E. Demencias e imagen: lo básico. Radiología, vol 52, no 1, 2010, pp 4-17 (ISSN: 0033-8338).

[5] Masdeu Puche JC. La neuroimagen en la enfermedad de Alzheimer: perspectiva actual. Rev. Neurol., vol 38, no 12, 2004, pp 1156-1165 (ISSN: 0210-0010).

[6] Orellana D. I. et al. Trabajos en cartel. XXXVI curso anual de radiología SMRI. Anales de radiología México, vol 1, nº 3, 2002, pp 529-533 (ISSN: 2604-2053).

[7] Gili J. Introducción biofísica a la resonancia magnética en neuroimagen. Vol. 1. Centre Diagnóstic Pedralbes, 1993.

[8] Prakash D, Madusanka N, Bhattacharjee S, . Park HG, Kim CH, y Choi HK. A Comparative Study of Alzheimer's Disease Classification using Multiple Transfer Learning Models. Journal of Multimedia Information System, vol 6, nº 4, 2019, pp 209-216 (ISSN: 2383-7632).

[9] Hosseini-Asl E, Keynton R, y El-Baz A. Alzheimer's disease diagnostics by adaptation of 3D convolutional network. En *2016 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2016, pp 126-130.

[10] Korolev S, Safiullin A, Belyaev M, y Dodonova Y. Residual and Plain Convolutional Neural Networks for 3D Brain MRI Classification. En *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE, 2017. pp 835-838.

[11] Convolutional Neural Networks for alzheimer disease classification from 2D brain images. ResearchGate. https://www.researchgate.net/publication/324747235_Convolutional_Neural_Networks_for_alzheimer_disease_classification_from_2D_brain_images (Consultada: Junio 2020).

[12] Chaplot S, Patnaik LM, y Jagannathan NR. Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. Biomed. Signal Process. Control, vol 1, no 1, 2016, pp 86-92 (ISSN: 1746-8094).

[13] Metrics to Evaluate your Machine Learning Algorithm. https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234 (Consultada: Mayo 2020).

Sistema de adquisición de datos de detector de radiación de bajo costo para la educación

S. Jiménez[1], R. Chil[1], D. Pérez Benito[1], J. Pascau[1,2], J.J. Vaquero[1,2]

[1] Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid

[2] Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), Madrid

Resumen

Los detectores de radiación son el elemento más crítico de los tomógrafos PET y las señales que producen requieren sistemas de adquisición capaces de muestrearlas con resolución de nanosegundos, lo que hace que estos sistemas sean costosos. Este trabajo presenta un sistema de adquisición de bajo coste con el objetivo de poder utilizarlo en enseñanza.

El sistema propuesto se basa en un proceso de "integración de carga analógica" para pre-procesar los pulsos del detector seguido de una placa de desarrollo Launchpad F28379D de Texas Instruments para muestrear la señal de carga integrada.

El sistema propuesto permite realizar la caracterización completa de un detector de radiación en menos de 5 minutos y con calidad aceptable. Este sistema también permite realizar prácticas de programación concurrente para sistemas de tiempo real crítico sobre un sistema físico real dado que dispone de múltiples CPUs. Se trata de una solución económica, sencilla y versátil con la que los alumnos podrían aprender el funcionamiento de los detectores de radiación empleados en medicina nuclear.

1. Introducción

La medicina nuclear es una de las modalidades de más utilizadas en los estudios de imagen funcional. Las imágenes nucleares se obtienen gracias a la detección de radiaciones emitidas por núcleos radiactivos inestables. Su uso principal es el diagnóstico de enfermedades cardiológicas, oncológicas, neurológicas, infecciosas, entre otras. Su principal característica frente al resto de modalidades de imagen médica es su alta especificidad.

En función del tipo de isótopo utilizado para el estudio podemos distinguir distintas modalidades de imagen nuclear, siendo las dos principales las realizadas por un fotón único o tomografía computarizada por emisión de fotón único (SPECT), y la tomografía por emisión de positrones (PET). Aunque ambas modalidades se basan en la detección de partículas gamma, la diferencia radica en los distintos procesos de decaimiento radiactivo [1].

El componente crítico de los sistemas de imagen nuclear son los detectores de radiación. Dichos detectores generan señales que pueden llegar a durar menos de 200 nanosegundos, con tiempos de subida de 10 nanosegundos. Las características de estas señales hacen que sea necesario el uso de sistemas de adquisición de datos especiales para poder adquirir y procesar los pulsos procedentes de los detectores. Por lo general estos detectores hacen uso de cuatro señales para detectar la posición de los eventos de radiación (X+, X-, Y+, Y-). Esta codificación es comúnmente conocida como lógica de Anger y fue desarrollada en los 60 [2].

En este trabajo se presentará el desarrollo de un sistema de adquisición de bajo coste que podría ser utilizado para que los estudiantes de la universidad realicen prácticas en las que poder familiarizarse con el funcionamiento de estos sistemas, sin que ello conlleve una inversión de capital muy elevado permitiendo así adquirir varios de estos sistemas de adquisición para una misma clase.

1.1. Sistemas de Adquisición en Imagen Nuclear

Las señales de los detectores de radiación son pulsos muy rápidos y que además, dependiendo de la dosis inyectada al paciente, pueden alcanzar tasas de conteo (*count rate*) de 10^6 eventos por segundo y por detector.

Con el objetivo de poder adquirir la mayor cantidad de eventos sin que ello suponga una pérdida en la calidad de la información adquirida se han desarrollado diversos métodos de adquisición [3]:

- Tiempo sobre el umbral (*time over threshold*, ToT): se mide el tiempo que está la señal por encima de un determinado umbral como un subrogado de la amplitud del pulso, que es a su vez proporcional a la energía del mismo. La principal ventaja de este método es que es más fácil de implementar al no requerir ADCs, siendo posible trabajar todo el tiempo con señales digitales. Este esquema se suele implementar con múltiples ASICs [4].
- Muestreo de Pulsos: debido a las características temporales de estas señales es necesario utilizar conversores de más de 100 MSPs, lo que hace que estos sistemas sean bastante costosos. Por otra parte, el disponer del pulso entero muestreado permite realizar operaciones de procesamiento digital sobre la señal adquirida mejorando notablemente el resultado final [5].
- Integración de Carga: esta modalidad surge como alternativa a la anterior y se basa en la realización de una integración de carga en la etapa analógica, medida que es proporcional a la energía

ISBN: 978-84-09-25491-0, pags. 81- 84

del evento detectado. Con este procedimiento solo es necesario muestrear la señal conteniendo que representa la carga ya integrada de cada pulso radioactivo recibido. La principal ventaja es una reducción de coste y consumo, a cambio se pierde la posibilidad de procesamiento digital del pulso [6].

- Detección de Pico: Al igual que el caso anterior es un intento de simplificar el proceso de adquisición. En este caso se almacena tan solo la información del valor máximo del pico del pulso del evento registrado. Supone también un menor coste y consumo de potencia. Sin embargo, este método ofrece peores resultados que la integración de carga dado que es muy sensible al ruido [6].

Teniendo en cuenta que la reducción de coste es uno de los objetivos de este trabajo, se ha decidido utilizar la integración de carga, al ser un método ya empleado previamente en nuestro grupo de investigación con buenos resultados.

2. Metodología

2.1. Sistema de Adquisición y procesamiento

Con el objetivo de poder realizar la conversión de carga se estudió qué opciones había en el mercado que pudiesen cumplir con nuestros requerimientos de disponer cuatro canales de conversión simultáneos de al menos 1 MSPs. También era necesario poder controlar el comienzo de la conversión, para lo que es necesario un ADCs que permita iniciar la conversión de forma asíncrona.

El entorno de desarrollo seleccionado fue la placa de desarrollo Launchpad F28379D de Texas Instruments (Fig. 1). Además de dos microporcesadores, la placa dispone de cuatro ADCs independientes de 3,5 MSPs con 12 bits de resolución. Al tratarse además de una placa de desarrollo permite minimizar el número de elementos externos, lo que reduce considerablemente el coste, la complejidad y la susceptibilidad a ruido. Las dos CPUs que resultan muy útiles para poder tener una ejecución eficiente asignado a cada CPU una operación diferente. Dispone además de dos aceleradores de reglas de control (CLA), capaces de procesar eventos en tiempo real descargando de esta tarea a las CPUs.

Un inconveniente de esta placa de desarrollo es que la conexión USB es a través de un chip de FTDI que solo soporta JTAG y emulación de puerto serie sobre USB, siendo por lo tanto imposible realizar trasnferencias masivas de datos (*bulk transfers*) mediante el USB. Esta limitación nos impone una restricción en la tasa máxima de conteo al permitir una velocidad de comunicación máxima de 115200 bps.



Figura 1. *Texas Instruments F28379D Launchpad [7].*

2.2. Electrónica de Acondicionamiento para integración de carga

Una parte fundamental de un sistema basado en integración de carga es la electrónica necesaria para el acondicionamiento del pulso. El sistema utilizado en el presente proyecto se basa en otro desarrollado previamente por nuestro grupo para un dosímetro personal [8].

La primera etapa de amplificación adapta la señal para soportar pulsos procedentes de distintos detectores. A continuación, la señal se retrasa con un retardo analógico de 50 ns para dar tiempo a la apertura de la puerta de integración sin perder ninguna parte del pulso. La apertura de la puerta de integración se realiza utilizando un flip-flop activado por un comparador que detecta la llegada de un nuevo evento (Fig. 2). Dicho flip-flop se resetea posteriormente desde el microprocesador una vez haya terminado la adquisición del evento actual, preparando así el sistema para la adquisición de un nuevo evento.

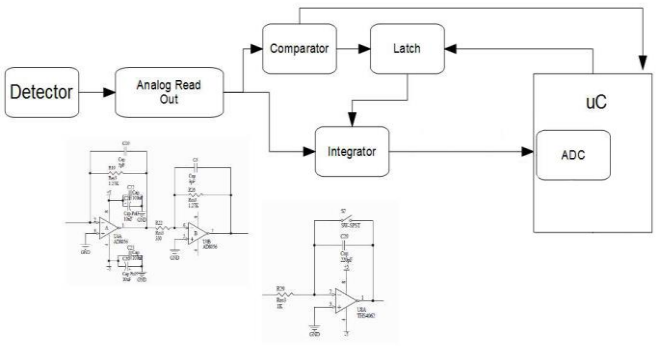


Figura 2.*Sistema de Integración de Carga*

2.3. Programa de Adquisición

El programa de adquisición consta de dos hilos de ejecución separados, cada uno ejecutándose en una de las CPUs del microcontrolador (Fig. 3). Durante el proceso de inicialización ambas CPUs configuran adecuadamente el acceso a los distintos bloques de memoria RAM del microcontrolador, para garantizar el acceso de ambos

hilos de ejecución a una zona de memoria compartida, donde serán almacenados los eventos adquiridos. A cada CPU le serán asignados los periféricos, pines e interrupciones correspondientes.

Durante el bucle principal de ejecución, el hilo ejecutado en la CPU 2 estará a la espera de la detección de un pulso. Al detectar un nuevo evento mediante una interrupción externa conectada una GPIO se activarán de forma simultánea los cuatro ADCs del microcontrolador para comenzar la conversión de las señales del detector. Los valores adquiridos se almacenan directamente en la RAM compartida entre CPUs. Se utilizará un buffer en modo ping-pong para alternar entre parte alta y parta baja de dicho buffer. De forma tal que mientras una CPU va llenando una mitad con nuevos datos, la otra va enviando los datos anteriores a través del USB/serie. Es por lo tanto necesaria la utilización de mensajes IPC entre ambas CPUs para coordinar que CPU trabaja sobre qué parte del buffer en cada momento.

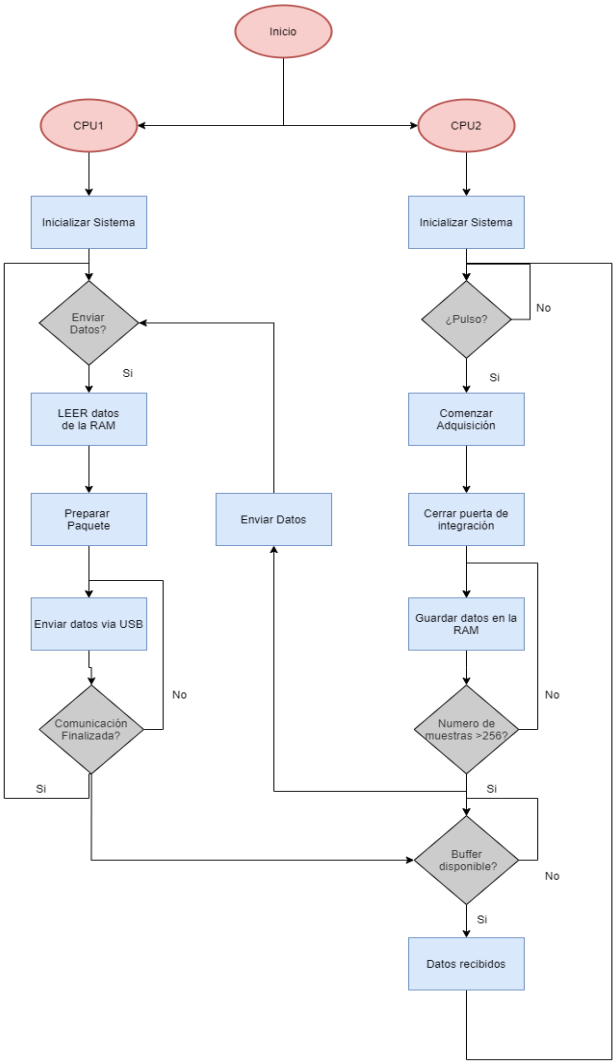


Figura 3. *Flujograma*

Al tratarse de un sistema para ser utilizado por estudiantes, solo se podrán utilizar fuentes radiactivas para docencia (con muy poca actividad), y por lo tanto producen tasa de conteo bajas. Esto permite garantizar que siempre será más rápido el envío de datos a través del puerto USB/serie que la adquisición de nuevos eventos. En el caso de que no fuese posible enviar todos los datos, se dejaría de adquirir nuevos eventos hasta tener espacio disponible en la memoria interna, o lo que es lo mismo, hasta que terminase el presente envío de datos.

2.4. Pruebas de laboratorio

Las pruebas de funcionamiento de este sistema se habían programado en realizarlas con detectores de radiación desarrollados previamente en nuestro grupo. En cambio, no se ha podido realizar de esta manera por las circunstancias marcadas por la pandemia. Por ello se ha realizado la simulación de los pulsos mediante un generador de funciones (Fig. 4).

3. Resultados preliminares

El objetivo es una adquisición de una imagen de llenado de campo y espectros para cada uno de los canales.

Provisionalmente se ha podido comprobar que el sistema alcanza una tasa de conteo máxima de 1200 eventos por segundo. Esta limitación viene impuesta por la conexión serie utilizada y que limita la velocidad de transferencia. Para ensayos posteriores se intentará solucionar este problema y así poder aumentar la tasa de conteo máxima. Esta la tasa de conteo permite obtener en cinco minutos imágenes de “llenado de campo” con la calidad suficiente como para poder caracterizar el detector de radiación.

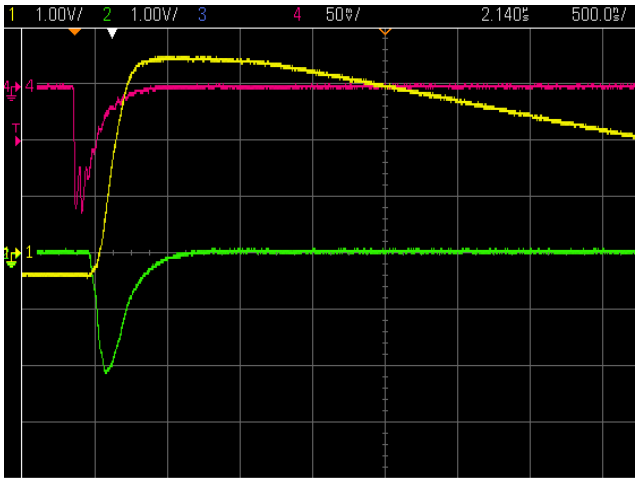


Figura 4. *Señales del sistema de integración de carga. Ch1: Integral del pulso, Ch2: Pulso amplificado y retrasado, Ch3: Pulso de entrada.*

3.1. Potenciales casos de uso

- Prácticas de Grado: utilizando este sistema los alumnos pueden adquirir pulsos de radiación con los que generar espectros de radiación o imágenes de llenado de campo. Una práctica con la que comprender como funciona la caracterización de un detector PET utilizando un detector real. Los alumnos podrían modificar algunos parámetros del

detector como la alta tensión utilizada y el umbral de comparación empleado para la detección de eventos.

- Prácticas de Máster: tratándose de alumnos con mayor nivel de conocimiento. Estos alumnos podrían modificar el código presente en el microcontrolador para en lugar de enviar los datos de los pulsos adquiridos realizar el procesado de estos datos en el microcontrolador y posteriormente enviar la información al PC. En esta práctica los alumnos podrían estudiar el funcionamiento de los ADCs, de las interrupciones externas, del envío de datos, de la comunicación entre procesadores y la programación concurrente, entre otros.

4. Conclusión

En este trabajo demostramos que es viable la utilización desarrollar sistemas de adquisición de señales de detectores de medicina nuclear de bajo coste, lo cual permite su utilización utilizarlas en ámbitos con un presupuesto limitado, como puede ser la docencia.

Dentro del ámbito académico encontramos más ventajas que la de su reducido coste respecto a los sistemas convencionales. El empleo de un microcontrolador programable con un entorno de desarrollo fácil de utilizar, da mucha versatilidad para la modificación y mejora del sistema, ya sea modificando la sección analógica del circuito, o programando algoritmos más eficaces para la caracterización del evento. Al disponer de múltiples hilos este sistema permite además realizar pruebas de programación concurrente sobre un sistema real de tiempo crítico.

A su vez este tipo de sistemas permiten ilustrar muy eficazmente cómo se comporta y funciona un detector de radiación de un tomógrafo PET o SPECT. Por ello, este tipo de desarrollo resulta muy interesante en el entorno académico tanto para alumnos de grado como de máster.

Agradecimientos

Ministerio de Ciencia, Innovación, Proyecto "PID2019-109820RB-I00"; Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación, proyecto "TEC2016-78052-R / MINECO/FEDER, UE"; This project has received funding from 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.

Referencias

[1] S. R. Cherry, J. A. Sorenson, and M. E. Phelps, *Physics in nuclear medicine*. Elsevier/Saunders, 2012.

[2] H.Anger, *Scintillation Camera with Multichannel*, pp. 515–531, 1964.

[3] G. F. Knoll, *Radiation detection and measurement*. John Wiley, 2010.

[4] J. Castilla *et al.*, "Evaluation of the FlexToT ASIC on the readout of SiPM matrices and scintillators for PET," in *IEEE Nuclear Science Symposium Conference Record*, 2013, pp. 1–4.

[5] Nutaq, "PicoDigitizer 125-Series | Nutaq." [Online]. Available: https://www.nutaq.com/products/picodigitizer-125-series. [Accessed: 07-Jun-2019].

[6] S. Ahmad *et al.*, "Triroc: A Multi-Channel SiPM Read-Out ASIC for PET/PET-ToF Application," *IEEE Trans. Nucl. Sci.*, vol. 62, no. 3, pp. 664–668, Jun. 2015.

[7] "LAUNCHXL-F28379D Overview User's Guide LAUNCHXL-F28379D Overview," 2016.

[8] R. Chil et al., Advanced networked modular personal dosimetry system, vol. 41. Springer, Cham, 2014.

Herramienta de pegado de múltiples camas para tomografía computarizada en 3D mediante dispositivos GPU

Javier Garcia Blas[1], Pablo Brox[1], Jesus Carretero[1], Manuel Desco[2,3,4,5] and Monica Abella[2,3,5]

[1] *Departamento de Informática, Universidad Carlos III de Madrid, Madrid, España*

[2] *Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), Madrid, España*

[3] *Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III, Madrid, España*

[4] *Centro de Investigación en Red de Salud Mental (CIBERSAM, CIBER CB07/09/0031), 28007 Madrid, España*

[5] *Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, España*

Resumen

En sistemas de tomografía axial computarizada (TAC) es común hacer varias adquisiciones tomográficas consecutivas para distintas posiciones de la cama, que posteriormente se combinan para ampliar el campo de visión en la dirección longitudinal. Para esta combinación es necesaria la calibración geométrica del movimiento de la cama para evitar dobles bordes en la zona de pegado. Esta calibración se realiza de forma periódica usando un maniquí específico. Este trabajo presenta una herramienta novedosa de pegado de cama automático para TAC basada en correlación. Nuestra propuesta explota el paralelismo masivo que ofrecen las GPUs y cuenta con un modelo de memoria optimizado que permite pegar grandes volúmenes en tiempo casi real. La evaluación en estudios de roedor demuestra, no solo que la implementación ofrecida es capaz de pegar los estudios tomográficos en un tiempo reducido, sino también que hace un uso eficiente de los recursos de memoria disponibles.

1. Motivación

Con los nuevos avances en el campo de la imagen médica, la calidad de los detectores en tomografía computarizada ha aumentado significativamente al obtener un mayor número de proyecciones por estudio. Esta mejora se traduce en una mayor resolución y cantidad de datos para procesar [3], exigiendo cada vez más potencia de computación. Este desafío, combinado con la necesidad de obtener los resultados en poco tiempo [4] motiva la necesidad de aceleradores para mejorar los cálculos. El enfoque más empleado es el uso de dispositivos de GPU, que ofrecen un buen equilibrio entre rendimiento y uso. Además, se pueden acoplar al hardware existente con poco esfuerzo.

La tomografía axial computarizada (TAC) es uno de los procedimientos de diagnóstico y evaluación de imágenes médicas más importantes. Se basa en la combinación de múltiples imágenes de rayos X tomadas desde diferentes ángulos alrededor del objeto para obtener un volumen tridimensional [1]. Cuando el campo de visión es insuficiente para abarcar la estructura anatómica de interés, es necesario generar múltiples volúmenes (camas) solapados, que se concatenan una sola salida. Este paso de concatenación se denomina pegado de cama e implica la fusión de la parte común de las dos camas, que denominamos "zona de solape". En este sentido, en [2] se introdujo un software de reconstrucción basado en FDK que incluye pegado de camas basado en una calibración previa.

Hay dos contribuciones principales presentadas en este trabajo. La primera es portar una aplicación de pegado automático de múltiples camas, proporcionando una nueva implementación paralela de los algoritmos incluidos utilizando dispositivos GPU. Se presenta un enfoque basado en módulos, que incluye un nuevo modelo de memoria que reduce significativamente los recursos necesarios, lo que permite el pegado de mayores volúmenes en menos tiempo. Además, el diseño elegido abre la puerta a una mayor implementación de nuevos algoritmos y modos de operación. El segundo es integrar esta aplicación en un software de reconstrucción basado en FDK [6].

Las principales contribuciones de este trabajo están relacionadas con la mejora e integración del prototipo dado y pueden dividirse en tres. Primero, identificar los diferentes procesos que conforman el prototipo y estructurarlos en diferentes módulos, facilitando la adición de nuevos algoritmos o modos de ejecución. Segundo, redefinir el modelo de memoria de la aplicación reduciendo la memoria consumida por los algoritmos implementados. Por último, presentar la implementación y evaluación de una versión de la aplicación de pegado de camas que aproveche la potencia intrínseca de los dispositivos GPU para optimizar los cuellos de botella, reduciendo el tiempo global de ejecución y reduciendo la posibilidad de introducir errores de forma manual.

2. Método propuesto

Para obtener un volumen final una vez solapadas cada una de las camas, tenemos que abordar principalmente dos obstáculos: los cortes superpuestos y la desviación de ángulo entre camas consecutivas. En la Figura 1 muestra un ejemplo de dos camas no alineadas del mismo volumen, el área sombreada representa los cortes que aparecen en ambos sub-volúmenes.

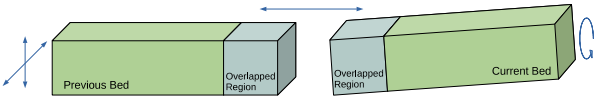


Figura 1.*Diferentes camas tomadas del mismo sujeto.*

ISBN: 978-84-09-25491-0, pags. 85- 88

Las etapas principales del método son dos. Primero, a través de tres pasos de correlación entre camas adyacentes se caracterizan los parámetros de registro: (i) desplazamiento axial (eje z), que corresponde con el corte de comienzo de la porción superpuesta en la cama actual; (ii) rotación alrededor del eje axial; (iii) desplazamiento coronal y (iv) desplazamientos sagital. Segundo, la fusión de los sub-volúmenes o camas.

El flujo de ejecución se presenta en la Figura 2. El proceso se divide en cuatro submódulos.

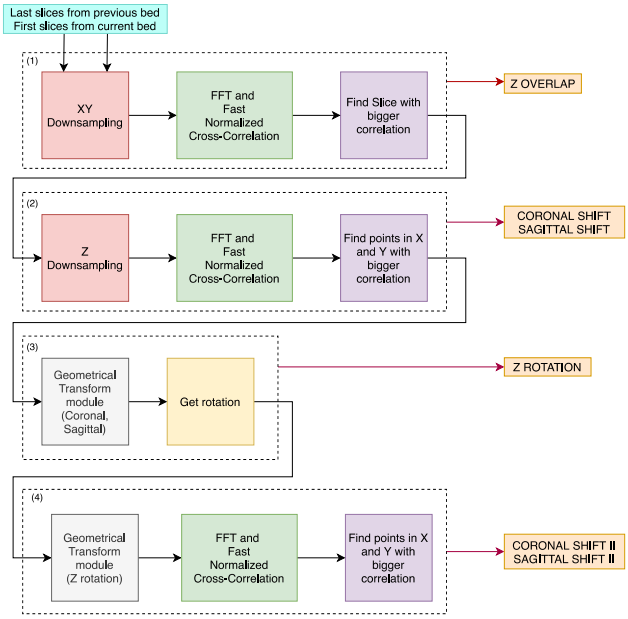


Figura 2. *Flujo de ejecución de pegado automático de camas.*

En el Módulo 1, a través de una primera correlación se caracteriza el desplazamiento en el eje Z, que se corresponde con el comienzo de la zona de solape con la cama anterior. Para poder acelerar el proceso se permite realizar un submuestro en el plano XY. Una vez aplicado dicho desplazamiento, el Módulo 2 busca mediante una segunda correlación el desplazamiento en los ejes X e Y. Una vez más, se puede aplicar un submuestro en el eje Z para acelerar este proceso. Estos desplazamientos se aplican mediante una transformación geométrica, para en un siguiente paso, calcular el ángulo de rotación entre camas. Finalmente, en el Módulo 4 se refinan los valores de desplazamiento en X e Y una vez aplicada la rotación anterior.

En la última etapa, se realiza una transformación geométrica a la cama actual con los parámetros calculados anteriormente. Para ello, se hace la fusión de la parte común de las dos camas tomando el valor del vóxel que sea máximo entre ambas camas.

La escritura al archivo de salida se hace de forma gradual. Al principio, el archivo de salida sólo contiene el primer volumen completo, a continuación, al final de cada iteración, el sistema coloca el puntero del archivo en el sub-volumen donde comienza la superposición y escribe la correspondiente cama con la zona de solape corregida.

3. Optimizaciones aplicadas

El objetivo de este trabajo no sólo es ofrecer una nueva implementación del prototipo, más fácil de usar e integrable con otros sistemas de reconstrucción, sino también disminuir el tiempo de ejecución y el consumo de memoria. Como estamos trabajando principalmente con volúmenes gráficos, la computación en GPU nos permite paralelizar los algoritmos usados aliviando los cuellos de botella con una baja sobrecarga.

La solución seleccionada para interactuar con la GPU es NVIDIA CUDA. No sólo por su flexibilidad y facilidad de uso, sino también porque proporciona algunas librerías útiles (cuFFT [7] o NPP [6]) que implementan algoritmos útiles para el sistema, como se verá en las siguientes secciones.

3.1. Fast Normalized Cross-Correlation

El algoritmo de Correlación Cruzada Rápida Normalizada introducido por Lewis [9] requiere que sus entradas estén en el dominio de la frecuencia, y para lograrlo es necesario aplicar varias FFT (en ambas direcciones) a los volúmenes. El proceso de implementación de FFT en la GPU es complejo, sin embargo, CUDA ofrece la librería cuFFT que proporciona funciones para realizar estas transformaciones dentro de la GPU, reduciendo la complejidad del código.

A pesar de la mejora en el tiempo de computación obtenido con esta librería, el algoritmo que aplica la FFT consume mucha memoria. Como la memoria de los dispositivos GPU es inferior a la memoria principal del computador, se ha implementado un mecanismo de control que comprueba la memoria disponible antes de llamar a la librería y transfiere a la memoria principal los resultados anteriores en caso de que no haya suficiente espacio.

3.2. Aceleración de transformaciones geométricas

A pesar del esquema de paralelización aplicado al procedimiento de interpolación, los accesos se realizan a vóxeles vecinos, realizando un uso ineficiente de la caché. Los accesos a la memoria de texturas son a través de una caché de sólo lectura que proporciona localidad espacial, por lo que los vóxeles se almacenan cerca de sus vecinos en lugar de en posiciones consecutivas en memoria.

Para solventar los anteriores problemas, hemos empleado la librería de primitivas de rendimiento de NVIDIA, que proporciona funciones de procesamiento de imágenes y señales aceleradas por la GPU e interpola utilizando la memoria de texturas. Esta solución permite obtener los valores interpolados directamente del hardware del dispositivo, evitando su cálculo.

3.3. Gestión de memoria

Con el objetivo de reducir el consumo de memoria del sistema, se aplicaron varios pasos de submuestreo de los volúmenes en el módulo de parámetros de pegado. Para ello, utilizamos la librería de primitivas de rendimiento de NVIDIA, que aceleran el proceso, aprovechando la potencia de GPU. Estas transformaciones mejoran no sólo el uso de la memoria sino también la velocidad del sistema.

4. Evaluación experimental

El hardware utilizado para llevar a cabo los experimentos consiste en un ordenador personal equipado con un procesador Intel Core i7-7700, 32 GB de RAM y una GPU NVIDIA GeForce GTX TITAN Black.

En este trabajo comparamos el rendimiento del prototipo inicial basado en MatLab con la versión del sistema implementada en lenguaje C. Un prototipo del sistema planteado en este trabajo se ha integrado en RAPTOR, un algoritmo basado en FDK [5].

Para ello, ambas versiones se ejecutan con dos estudios de roedor descritos en la Tabla 1.

Estudio	Camas	Dimensiones (vóxeles)	Resultado (vóxeles)
1	3	512×512×572	512×512×1170
2	2	516×516×574	516x516×970

Tabla 1. *Descripción de los conjuntos de datos empleados.*

4.1. Consumo de memoria

Para cada ejecución, se midió el consumo de memoria y el tiempo de ejecución. Las Figuras 3 y 4 muestran los resultados del experimento para cada conjunto de datos para la versión en Matlab y la propuesta en este trabajo.

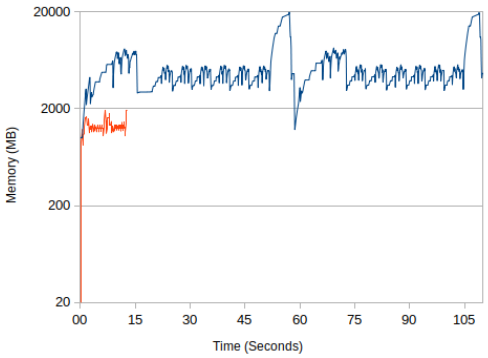


Figura 3. *Tiempo de ejecución en comparación del consumo de memoria del Estudio 1 (3 camas). El color azul corresponde a la versión basada en Matlab y el naranja a la implementada para este trabajo.*

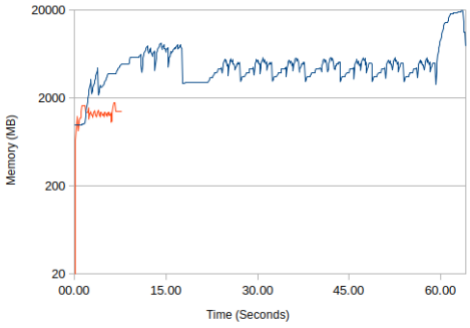


Figura 4. *Tiempo de ejecución en comparación del consumo de memoria del Estudio 2 (2 camas). El color azul corresponde a la versión basada en Matlab y el naranja a la implementada para este trabajo.*

Podemos observar que la versión de MatLab utiliza una media de 4 veces más memoria RAM que la versión propuesta para ambos estudios, comportamiento que es reproducible durante toda la ejecución. Además, en los experimentos hemos observado picos de consumo de memoria a lo largo de la ejecución. Esos picos son más significativos para la versión MatLab y aparecen al final del proceso de cada cama (el primer estudio tiene tres camas, y por lo tanto, dos regiones superpuestas e iteraciones) indican que un proceso está mal optimizado.

En términos generales, se emplea en promedio un 23,5% del total de la memoria RAM disponible en la versión de MatLab y el porcentaje cae a un 9,3% cuando se observa el pico de consumo de memoria. En cuanto al tiempo de ejecución, la solución propuesta obtiene una aceleración de 6,85× para el primer estudio y 8,18× para el segundo, resultando en un promedio de aceleración de 7,51×.

4.2. Tiempo de ejecución

El propósito de este experimento es determinar esa sobrecarga.

Estudio	Pegado	Tiempo de ejecución (seg)
1	Manual	20,020
	Auto	26.101
2	Manual	17,540
	Auto	19,645

Tabla 2. *Comparativa del tiempo de ejecución en modos manual y automático.*

Observamos en la Tabla 2 que el pegado de camas automático presenta una sobrecarga de 6,08 segundos para el primer estudio (30% más lento) y 2,105 segundos para el segundo (12% más lento). La principal diferencia entre el primer y el segundo conjunto de datos es el número de camas, lo que afecta al número de iteraciones realizadas en la fase de pegado automático y explica la diferencia porcentual. En promedio, la integración del mecanismo automático añade un tiempo de sobrecarga del 21% frente al tiempo total de ejecución.

5. Conclusiones y discusión

En este trabajo hemos presentado un nuevo sistema de pegado automático de camas que aprovecha la capacidad de cómputo de los dispositivos GPU. La versión inicial de la solución implementada en MatLab hacía inviable la puesta en producción de dicho método. A pesar de contar con un alto número de optimizaciones, MatLab carece de mecanismos de alto rendimiento para el procesamiento de transformadas de Fourier y el submuestreo de volúmenes. El prototipo presentando en este trabajo acelera de forma significativa este proceso mediante es uso de bibliotecas especializas adaptadas a la plataforma hardware, exprimiendo todo su potencial de cómputo.

La evaluación preliminar demuestra que la implementación propuesta no sólo es hasta 8 veces más rápida que la versión implementada en Matlab, sino que su consumo de memoria es un 76,5 % menor en promedio y

no presenta picos notables. Además, la integración con software de reconstrucción implica una sobrecarga aceptable (21% en promedio) en comparación con el ahorro de tiempo, esfuerzo y la eliminación de artefactos producidos por un pegado basado en una etapa previa de calibración.

El pico de mayor consumo de memoria de la aplicación se alcanza dentro de la llamada de la biblioteca de cuFFT, lo que hace que la aplicación pueda tener dificultades a la hora de procesar volúmenes de gran tamaño. Una posible solución es gestionar el consumo de memoria dentro de la biblioteca mediante el particionamiento de cada una de las camas en varios subvolúmenes.

En promedio, la integración del mecanismo automático añade un tiempo de sobrecarga del 21% frente al tiempo total de ejecución. No obstante, elimina la calibración geométrica necesaria en un esquema convencional, que debe realizarse con mucho cuidado para obtener resultados válidos.

En este trabajo no se ha considerado el empleo de más de un dispositivo de GPU a la vez. Implementar un enfoque de multi-GPU que permita aprovechar todos los dispositivos instalados redundará en una disminución del tiempo de ejecución.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia, Innovación y Universidades (Instituto de Salud Carlos III, proyecto DTS17/00122; Agencia Estatal de Investigación, proyecto DPI2016-79075-R-AEI/FEDER, UE), cofinanciado por Fondos de la Unión Europea (FEDER), "A way of making Europe". Además, ha sido financiado por el Programa de apoyo a la realización de proyectos interdisciplinares de I+D para jóvenes investigadores de la Universidad Carlos III de Madrid 2019-2020 en el marco del Convenio Plurianual Comunidad de Madrid- Universidad Carlos III de Madrid (proyecto DEEPCT-CM-UC3M) y por CRUE Universidades, CSIC y el Banco Santander (Fondo Supera, proyecto RADCOV19). El CNIC está financiado por el Ministerio de Ciencia, Innovación y Universidades y la fundación PRO-CNIC y es un centro de excelencia Severo Ochoa (SEV-2015-0505).

Referencias

[1] T. B. R. Company. (Jul. 2018). Veterinary industry outlook: Top global trends and statistics, [Online]. Available: https://blog.marketresearch.com/veterinary-industry-outlook-top-global-trends-and-statistics.

[2] M. Abella et al., "Software architecture for multi-bed fdk-based reconstruction in x-ray ct scanners", Computer methods and programs in biomedicine, vol. 107, no. 2, pp. 218–232, 2012.

[3] N. De Briyne and D. Iatridou, "Challenges seen with treatment of exotic pets in veterinary practice exotic pets in veterinary practice", Dec. 2016.

[4] H. Turbell, "Cone-beam reconstruction using filtered backprojection", PhD thesis, Linköping University Electronic Press, 2001.

[5] J. G. Blas, M. Abella, F. Isaila, J. Carretero, and M. Desco, "Surfing the optimization space of a multiple-gpu parallel implementation of a x-ray tomography reconstruction algorithm", Journal of Systems and Software, vol. 95, pp. 166–175, 2014.

[6] J. Ghorpade, J. Parande, M. Kulkarni, and A. Bawaskar, "GPGPU processing in CUDA architecture", arXiv preprint arXiv:1202.4347, 2012.

[7] CUDA Toolkit 4.2 CUFFT Library programming guide, NVIDIA Corporation, Mar. 2012. [On- line]. Available: https://developer.download.nvidia.com/compute/DevZone/docs/ html/CUDALibraries/doc/CUFFT_Library.pdf.

[8] NVIDIA 2D Image and Signal Performance Primitives (NPP), [Online]. https://docs.nvidia.com/cuda/npp/index.html.

[9] J. Lewis, "Fast normalized cross-correlation, 1995", in *Vision Interface*, vol. 2010, 2010, pp. 120–123.

Photo vector extraction using Deep Learning techniques for Cognitive Map Index calculation

M. Jaén-Vargas[1], M. Fernández Cerezo[1], K. Reyes Leiva[1,2], F. Ortiz-Gomez[3], A. Cobo[1,4],

J. Serrano Olmedo[1,4]

[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

[3]Information Processing and Telecommunications Center, Universidad Politécnica de Madrid, Madrid, España

[4] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, (CIBER-BBN), Madrid, España

Abstract

The main purpose of this project is to automate the manual mechanism of the acquisitioning of 3 vectors from the images taken of an experiment executed by impaired people. These vector values are implied to calculate de CMQ index which intervenes to obtain how well a person mentally understands their exterior space by creating cognitive maps. The index mentioned was obtained in a more efficient and less time-consuming process.
The system includes a smartphone application that aims users to identify Point of Interests (POI) in a scenario previously virtualized. Importantly, Deep Learning techniques were applied to automate the process, principally, two supervised learning approaches: regression and classification. While classification helps to improve the model, the *regression approach is closer to the expected result. However, neither of the two types of algorithms provides the desired results since the model presents a high error in the prediction, although in turn, they show a clear capacity for improvement. Regarding the technologies used in the study, Python is the main language for coding besides Google Collab where compiling models in the cloud. The results indicate that the tested model represents a first approximation that requires a greater number of samples and further development to obtain a more accurate prediction and be implemented, but the path is appropriate.*

1. Introduction

Sighted and blind people perceive outer space in different ways [1]. This fact has been a very important source of information when developing tools for spatial learning. Cognitive maps are called to the mental representation which individuals use to acquire, code, store, recall, and decode information about the relative locations and attributes of phenomena in their everyday or metaphorical spatial environment [2].
Ultimately, our research line is focused on study how people with no sight understand its spatial environment calculating the CMQ (Cognitive Map Quality) index or SUQI (Spatial Understanding Quality Index). Hence, this project is based on previous work, where the mentioned index was calculated considering the experience of the blind after using an application to recognize objects in a specific room [3]. This process was manually done by volunteers that interpreted a photo taken from a mock-up where the users had to locate the objects they remembered were in the game. In the present research, it was proposed to use Deep Learning algorithms to automatize the current manual vector extractions process to obtain the CMQ index mentioned above. Consequently, we are able to know how well the user creates a map representation in its brain.

2. Materials and Methods

Mainly, the automation process is carried out through two software: Microsoft Excel and Google Collaboratory.
First, Microsoft Excel is used to create a database with the calculations for the vector extractions and its final values.
Second, an iPhone 6S plus mobile phone is the tool used to obtain the images and a cardboard box is adapted to create a mock-up of the original environment.
Third, the elaboration of the model is done with Python through Google Collaboratory software.
To accelerate the current manual process of vectors extraction and find the SUQI, this study was divided into three main parts: the scene recreation; the database creation; and the algorithm coding.

3.1 Scene recreation: As many images were required to recreate the bar scene shown in Figure 1. Then, a mock-up was created using a box within its corresponding structures such as 4 walls, 2 windows, and 3 doors.

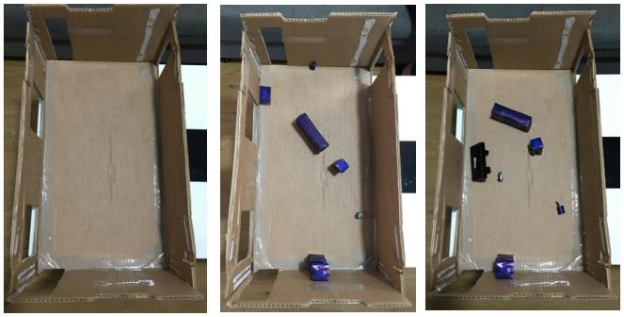


Figure 1*. Bar scene mock-up.*

Also, 6 blue-painted objects were added to differentiate the ones who belong to the correct scene: a stair, a porta umbrella, a chair, a bar, a dartboard, and an extinguisher. At the same time, objects who did not belong to the bar

ISBN: 978-84-09-25491-0, pags. 89- 92

scenario were black painted. Besides, it was established the exact position of the objects in the bar scene following the steps defined in the previous work [3-4]. The initial experiment has 3 scenes, and the user could be confused including objects that do not belong to the bar scene chosen for this project. Thus, we decide differentiating the bar objects because CMQ calculation included the two class of objects, and for image processing the algorithm extracts the features easy using different colors.

3.2 Database creation: After recreating the bar scene a 500 photos database was built. The object's position was varied to simulate the results obtained after the blind users experimented with the app. This program oriented the user to find where the objects were located. The first 250 photos were taken including or not, the 6 objects that belong to the bar scene or making combinations between the 6-blue and the black objects. Besides, the other half were taken to replicate the experiment results and get more samples.
Second, the images were processed manually. It was calculated 3 of the 4 vectors necessary to get SUQI through *Mahalanobis measure* [5]. The smaller the resulting Mahalanobis distance between two scenes, the better the quality of the cognitive map.
Third, the vectors which working with are *Identification, Absolute position, and Relative position* [5]. Additionally, the vector who is involved with the structure it is not taking into consideration because we do not have a mock-up to compare each other.

3.3 Algorithm coding:

The algorithms of Convolutional Neural Networks (CNN) with regression were implemented for extracting the vectors of interest. This is because regression predicts a previously known output value [6].
CNN were chosen because are recommended for working with complex images due to the good prediction. Thus, this class of networks can reduce the number of intern parameters and reutilize the weights [7]. Also, when using CNN, a good image understand can be obtained but reducing the processing time.
Regarding the experiment type, the photos taken have 3 channels (RGB), so it will require a huge amount of connections and intern parameters in the network. Thus, the prediction will be better and, in less time, using DL than using ML. In contrast, if it were used ML, the images should have been converted to binary (2 channels) where the process to predict did not extract all the characteristics.

Once the database is made, the outputs of the network were selected. In this case, the algorithm is focused on predicting the result of the 3 vectors of interest. In the beginning, a single vector at a time was needed to find the correct algorithm and then train the network with the 3 vectors' information.
After choosing the type of AI method, the code in Google Collab was made considering setup the following steps:
First, the images were reduced to 256x256 because it is the size recommended for working with CNN. Thus, a table with the picture's name was created to add all the values of 3 vectors information that was calculated manually, before. The algorithm matches the picture name, which is the input with the same name written in the table with its corresponding values.
Second, the train and test sets were randomly generated using a weight of 75% and 25%, respectively. Also, the "labels" which are the output values, were printed. In this case, the labels are the ones that the algorithm will predict.
Third, a convolutional network was created as an architecture taking into consideration the next parameters: depth (2 to 3 channels); filters, which were progressively larger so our network can learn more features; and the number of convolutional layers using the activation function.
Finally, Batch Normalization and Pooling layer were settled. When using "batch", it tends to stabilize training, and facilitates adjustment hyperparameters although it can increase processing times.
Usually, CNN is used for classification problems where the final output is a class result. However, in this case, fits a regression problem where the result expected is a numerical value between 0 to 1, we changed the SoftMax layer at the end for a linear function.
For the validation stage, 20 new images were taken to do the prediction and were not included in the 500 photos of the whole data frame.
To evaluate the algorithm's performance, 7 important parameters were tuned to obtain better predictions:

1. The loss function: Average Absolute Error was chosen since offers a good robustness to outliers [8].
2. The optimization algorithm: "Adam" algorithm which its high computational efficiency, the minimum memory requirement and ease of implementation [9].
3. The number of epochs: There is not an optimal value for all cases, but the appropriate value must be found so that there is neither underfitting nor overfitting [10].
4. The size of batches.
5. The learning rate: In general, a ratio of high learning allows the model to learn faster, at the cost of obtaining a set of worse quality final parameters. And conversely, a smaller learning rate may allow the model to acquire optimal weights but with a longer processing. However, it is possible that having a small learning ratio the algorithm never converges or even paralyze [11].
6. The activation functions: ReLU function has great computational efficiency, and is not linear, although it has behavior as such, thus allowing to go through the database backward ("backpropagation") [12].
7. The number of convolutional layers: the more layers the network has, it will be able to recognize more complex shapes. But also, an excessive number of layers have a high computational cost associated with it and can lead to overfitting. In addition, the layers are associated with a certain number of filters, so it is important to find the right combination of both elements.

3. Results and Discussion

Once the parameters were adjusted, the prediction was evaluated as shown in Table 1.

Table 1. Validation values comparison using regression.

Real values	Predicted values
0,067	0,225
0,133	0,288
0,200	0,084
0,133	0,240
0,067	0,131
0,333	0,298
0,267	0,359
0,200	0,256

The model error indicates how well the algorithm is learning from the input. After compiling all the epochs, the model error was 0.062, hence it was too large. For improving the model performance, it was decided to approach the case from another point of view: classification, a supervised learning technique. For this, a new codification was needed. It was observed that the clue thing to calculate the 3 vectors was the blue-objects localization in the box. So, the output data frame was changed for the decimal values (results of vector calculations) to a codification of 0 or 1 depending on how many blue objects were in the box. Thus, we managed 7 classes (from L0-L6). The network will try to identify the number of blue objects in the experiments, as shown in Table 2.

Table 2. Codification for Classification experiment.

Photo ID	Blue Objects	L0	L1	L2	L3	L4	L5	L6
A	**0**	1	0	0	0	0	0	0
B	**1**	0	1	0	0	0	0	0
C	**2**	0	0	1	0	0	0	0
D	**3**	0	0	0	1	0	0	0
E	**4**	0	0	0	0	1	0	0
F	**5**	0	0	0	0	0	1	0
G	**6**	0	0	0	0	0	0	1

After the codification table experiment, 2 CNN architectures were selected for doing the prediction: ConvNet and AlexNet. When analyzing the results, it was noticed that there was an incongruent number of the total samples resulting in 4 objects the bigger quantity. Due to this, the network was unable to do the prediction, accurately. However, the code was changed from 7 classes to 2, only selected the cases which were 4 objects in the box; and compile the model again. Then, we figured out that the problem was overfitting as appreciated in Figure 2, probably due to a low number of samples [13]. Thus, the classification approach allows identifying the problem observed in the regression type.

Due to overfitting, it was determined to make 300 samples more resulting in 800 photos and compiling the regression model tried at first, to get the numerical prediction for the identification vector. As a result, the model error was reduced from 0.062 to 0.042.

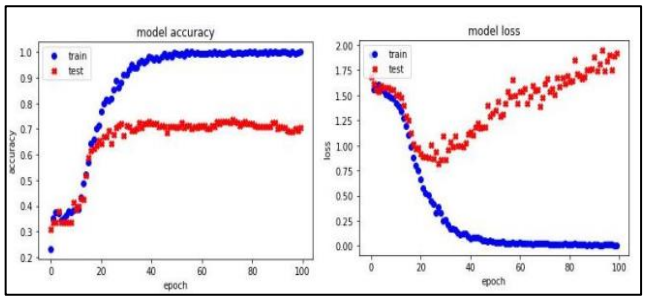


Figure 2. *Model Accuracy vs. Epochs and Loss vs epochs for 2 classes*

After obtained a diminution of the error, it was considered to calculate the 3 vectors predictions one by one. Thus, for the *Identification vector* some parameters varied as shown in Figure 3.

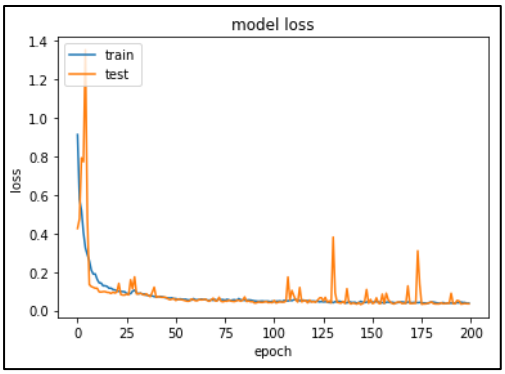


Figure 3. *Model Loss vs. Epochs for Identification Vector*

Also, it was set up the loss function as the Mean Absolute Error, the optimization algorithm as "Adam", 200 epochs were applied, the batch size was 8 samples, and finally, the function of activation was ReLU. It was obtained an improvement of 0.125 to 0.031 in the absolute error.
For the *Absolute position vector*, the same parameters were fixed but these results were not accurate as the Identification vector. After making some adjustments changing 8 layers for 16 these results improved obtaining an error from 0.081 to 0.073 as shown in Figure 4.

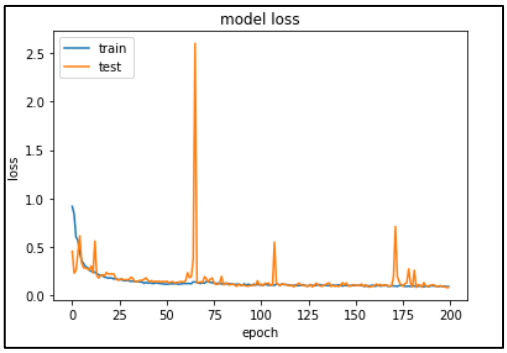


Figure 4. *Model Loss vs. Epochs for Absolute Position Vector*.

Regarding *Relative position vector* the smallest error found was 0.121 and occurred when varying the number of layers

convolutional to 7. Then the following sequence of filters was applied: 8, 16, 18, 32, 64, 128, 128. In this case, this vector has a laborious level of calculation because depends on the placement of the object, its relationship with the position of the objects remaining, as well as their relationship in the reference scene. The result is shown in Figure 5.

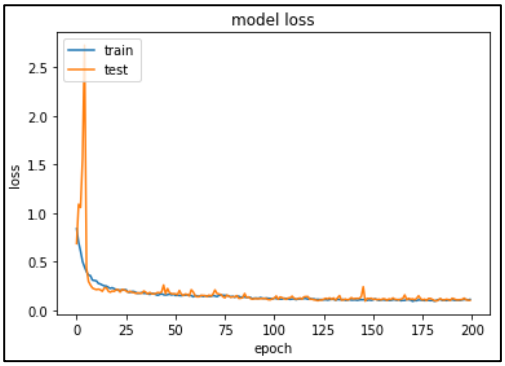


Figure 5. *Model Loss vs. Epochs for Relative Position Vector*

Once the results of the 3 vectors had been obtained separately, then, the algorithm was applied to all the vectors together. In this way, the error of the model was determined, capable to simultaneously predict the values necessary for the calculation of the CMQ index from a single photograph, which turns out to be 0.106. The contrast shown in Table 3 demonstrates the similitudes between the real vector value and the predicted one.

Table 3. *Real Values vs Predicted Values.*

	dIdentification		**pAbsolute**		**pRelative**	
Photo ID	Real	Predicted	Real	Predicted	Real	Predicted
IMG_1209	0,067	0,191	0,100	0,161	0,300	0,240
IMG_1211	0,200	0,177	0,500	0,417	1,000	0,549
IMG_1215	0,267	0,223	0,333	0,294	0,667	0,427
IMG_1229	0,067	0,137	0,600	0,480	0,700	0,603
IMG_1231	0,400	0,250	1,000	0,837	1,000	0,940

According to the current manual process used to calculate the CMQ index, an improvement is observed to get the vector's values. Using this technique of DL, we will be able to include it in the application and give us the possibility to know if the person created an accurate representation of the scene through the cognitive maps only introducing the mock-up pictures. Thus, the volunteers that helped us, will no longer be needed, and this will reduce time delays.

4. Conclusions

To summarize, Convolutional Neural Networks have turned out to be the indicated architecture for feature extraction from the bar scene picture. Even though the problem was initially posed only since the regression approach, a new point of view using classification was raised. Thus, it was observed that they are not exclusive supervised learning methods but also the two paths are complementary.

On the other hand, using regression demands to have many samples from which the network can learn and thus generate a better prediction, while in classification there must be class balancing in the input database.

Regarding regression method, in the case of the Identification vector the proposed model had an error of 0.062 for a database of 500 photographs while for 800 it is reduced to 0.031. Moreover, the remaining ones showed low performance for the proposed algorithm since they suppose a greater calculation and complex extraction. The model errors were 0.073 and 0.121 for the vector of Absolute position and vector 'Relative position', respectively. In addition, when applying the algorithm to obtain the 3 vectors simultaneously from the images, it is concluded with an error of 0.106.

In short, the tested model represents a first approximation that requires a greater number of samples and further development to be implemented, however, the results indicate that this is the appropriate path. Hence, more samples will be needed to obtain a more accurate prediction.

References

[1] Andrews, S. K. (1983). Spatial cognition through tactual maps. In Proceedings of the 1st international symposium on maps and graphics for the visually handicapped (pp. 30e40). Washington, DC: Association of American Geographers.

[2] Tolman, Edward C. (1948). "Cognitive maps in rats and men". Psychological Review. 55 (4): 189–208. doi:10.1037/h0061626. PMID 18870876.

[3] Cobo, Antonio & Guerron, Nancy & Martín, Carlos & Del Pozo Guerrero, Francisco & Serrano Olmedo, José. (2017). Differences between blind people's cognitive maps after proximity and distant exploration of virtual environments. Computers in Human Behavior. 77. 10.1016/j.chb.2017.09.007.

[4] Cobo, Antonio & Serrano Olmedo, José & Guerron, Nancy. (2018). Methodology for building virtual reality mobile applications for blind people on advanced visits to unknown interior spaces.

[5] Cobo Sánchez de Rojas, A. Proposal of a cognitive interface for the distant exploration of virtual environments as an aid to blind people during the creation of cognitive maps, PhD Thesis, E.T.S.I. Telecommunication (UPM), Madrid, 2018.

[6] A. Rosebrock. "Pyimagesearch." [Online]. Available: https://www.pyimagesearch.com/2019/01/28/keras-regression-and-cnns/ [Accessed May 2020].

[7] S. Saha, "Towards Data Science." [Online]. Available: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neuralnetworks-the-eli5-way-3bd2b1164a53. [Accessed May 2020].

[8] P. Grover, "Heart Beat." [Online]. Available: https://heartbeat.fritz.ai/5-regression-loss-functions-all-machine-learners-should-know-fb140e9d4b0. [Accessed June 2020].

[9] Kingma, D. and Ba, J. (2014). Adam: A Method for Stochastic Optimization (cite arxiv:1412.6980 Comment: Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015).

[10] J. Brownlee, "Machine Learning Mastery," [Online]. Available: https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch/. [Accessed May 2020].

[11] Goodfellow, I., Bengio, Y. and Courville, A. (2016). "Chapter 8: Optimization for Training Deep Models - Deep Learning."

[12] W. H. Lopez-Pinaya, S. Vieira, R. Garcia-Dias, and A. Mechelli (2020). "Chapter 10 - Convolutional neural networks," in Machine Learning. Cambridge, MA, USA: Academic.

[13] Simon S. Du, Yining Wang, Xiyu Zhai, Sivaraman Balakrishnan, Ruslan Salakhutdinov, and Aarti Singh. (2018). How many samples are needed to estimate a convolutional neural network? In Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18). Curran Associates Inc., Red Hook, NY, USA, 371–381.

Estudio de sensibilidad de algoritmos de análisis de la variabilidad de la frecuencia cardiaca

A. Pérez-Porro Segarra, C.A. García Martínez, G. Caffarena Fernández, A. Otero Quintana

Departamento de Tecnologías de la Información, Escuela Politécnica Superior, Universidad San Pablo-CEU,

Campus Montepríncipe, Boadilla del Monte, 28668 Madrid, España. a.perezporro@usp.ceu.es, {constantino.garciama, gabriel.caffarena, aotero}@ceu.es

Resumen

El análisis de la Variabilidad de la Frecuencia Cardiaca (VFC) es un método no invasivo que permite hallar marcadores de un gran número de patologías a partir de un registro de electrocardiograma. El objetivo de este trabajo es analizar la sensibilidad de los algoritmos utilizados para el análisis de la VFC evaluando tanto los métodos en el dominio del tiempo como en el de la frecuencia. Los algoritmos analizados son los implementados en el paquete RHRV de R. Se observa que en el análisis en tiempo los índices sensibles a las variaciones de sus parámetros son SDANN (la desviación estándar de los promedios de intervalos RR en periodos cortos de tiempo), SDNN index (el promedio de las desviaciones estándar de los intervalos RR en periodos cortos de tiempo) y TINN (la interpolación triangular del histograma de RRs). En frecuencia, el análisis utilizando la Transformada Corta de Fourier indica que las variables estudiadas no son sensibles a las variaciones de los parámetros. En el análisis a través de la Transformada Wavelet, las componentes de frecuencia LF (Low Frequencies) y HF (High Frequencies) son sensibles a los parámetros de esta transformada, entre los que destaca la madre Wavelet con la que se analiza la señal.

1. Introducción

La variabilidad de la frecuencia cardiaca (VFC) se define como la variación que ocurre en el intervalo de tiempo entre dos latidos, concretamente, entre 2 ondas R consecutivas [1]. Una de las características del análisis de la VFC es que permite estudiar el funcionamiento del sistema nervioso autónomo de una forma simple y no invasiva, concretamente, a partir de un registro de electrocardiograma (ECG). Además, diversos índices derivados de la VFC han sido empleados como marcadores de distintas patologías [2-4].

La VFC incluye métodos tanto en el dominio del tiempo como de la frecuencia. Los primeros suelen emplear técnicas estadísticas, mientras que los segundos suelen basarse en la Transformada Corta de Fourier (TCF) o en la Transformada Wavelet (TW) [5]. Los algoritmos que llevan a cabo estos análisis tienen un gran número de parámetros de configuración, lo que a menudo dificulta la reproducibilidad de los resultados de otros autores [6]. Además, no siempre está claro cómo variaciones en dichos parámetros afectan al resultado final del análisis (de ahora en adelante, resultados o índices). Por una parte, pueden existir índices robustos, en el sentido de que variar sus parámetros apenas afecte al resultado final; y por otra parte pueden existir índices sensibles, en los que pequeñas variaciones en sus parámetros de configuración resulten en resultados significativamente distintos. Cuantificar cómo variaciones en los parámetros afectan a los índices permitiría identificar aquellos parámetros a los cuales el análisis de la VFC es más sensible, y por tanto aquellos parámetros a los que el investigador debe prestar mayor atención a la hora de seleccionar sus valores.

El presente artículo estudia cómo las variaciones de algunos parámetros afectan a los índices más comunes en un análisis de la VFC. Para ello se aplican variaciones de modo sistemático en los parámetros y se realiza un estudio estadístico sobre cómo estos afectan a los índices resultantes. El estudio se centra en los algoritmos implementados en RHRV, un paquete de R [7].

El resto del artículo se estructura de la siguiente forma. En la sección 2 se detallan los pasos que se han seguido para llevar a cabo el análisis de la sensibilidad. La sección 3 muestra los resultados obtenidos. Por último, en la última sección se discuten los resultados obtenidos, se extraen conclusiones y se apuntan posibles líneas de trabajo futuro.

2. Materiales y Métodos

Para llevar a cabo el análisis de sensibilidad se empleará una base de registros electrocardiográficos disponible públicamente en Physionet que incluye 18 registros de ECG de diferentes pacientes del Hospital Boston's Beth de Israel [8]. Los pacientes no presentan ninguna patología significativa. Los registros se obtuvieron de 5 hombres de entre 26 y 45 años, y 13 mujeres de entre 20 a 50 años, con una frecuencia de muestreo de 128 Hz y tienen una duración aproximada de 10 horas.

Para realizar el análisis en el dominio del tiempo se han empleado todos los registros, sin embargo, en el dominio de la frecuencia se han descartado 3 registros (16272, 19088 y 19830) debido a errores a la hora de procesarlos. Esto se debe a la existencia de intervalos sin detección de latidos por el elevado ruido en el electrocardiograma y eso provoca que la interpolación de la serie temporal no se realice de forma correcta. Esto impide aplicar las técnicas del dominio de la frecuencia en las que se centra este artículo. Aunque existen diversas formas de afrontar este problema, se ha optado por no incluir estos registros para evitar introducir una fuente de variación más al estudio (la estrategia empleada para corregir la presencia de huecos).

2.1. Procedimiento

Se realizan tres análisis de sensibilidad diferentes. Uno estudia los índices derivados del análisis en el dominio del tiempo, y en los otros dos los índices derivados del análisis

ISBN: 978-84-09-25491-0, pags. 93- 96

en el dominio de la frecuencia. Entre estos últimos, el primero se centra en la TCF y el siguiente en la TW.

Como parte del análisis en tiempo se hace un estudio de la frecuencia cardiaca por ventanas y se construye un histograma con los intervalos RR, por lo que se varía la anchura de la ventana y el número de intervalos que tiene el histograma.

Por otro lado, en el dominio de la frecuencia se utiliza la TCF para extraer las frecuencias de la señal en secciones localizadas en tiempo. Para ello, la TCF emplea una ventana de Hamming que selecciona partes de la señal. Esta ventana se desplaza hasta realizar la TCF completa. En el análisis de la sensibilidad se varían la frecuencia de interpolación, el tamaño de la ventana y el desplazamiento de esta.

En el dominio de la frecuencia también es posible emplear la TW. La TW sigue un procedimiento similar a la TCF, pero en vez de ventanas se utilizan diferentes Wavelets madre. Estas Wavelets madre se desplazan dependiendo del intervalo temporal a analizar y se contraen o expanden en función de la frecuencia de interés. En el análisis con la TW se varía el tipo de Wavelet madre utilizada y el margen de error a la hora de realizar el análisis; este margen de error está relacionado con la aproximación de los intervalos de frecuencia correspondiente con las bandas de frecuencia ULF (*Ultra Low Frequencies:* 0-0.03 Hz), VLF (*Very Low Frequencies*: 0.03-0.05 Hz), LF (*Low Frequencies*: 0.05-0.15 Hz) y HF (*High Frequencies*: 0.15-0.4 Hz) de la VFC a partir de la TW [5].

Tras la obtención de los índices para los análisis de tiempo y frecuencia se lleva a cabo un estudio estadístico para determinar qué parámetros producen variaciones en los índices. Siguiendo a [9], para concluir que un índice es sensible a un parámetro se exigen dos condiciones: el parámetro debe ser estadísticamente significativo (medido por el p-valor) y el tamaño del efecto debe ser no despreciable. Por ejemplo, si un parámetro *P* es estadísticamente significativo, pero produce variaciones muy pequeñas en el índice *I*, se concluirá que *I* no es sensible a *P*.

A continuación, se indica más detalladamente qué índices se han estudiado y qué parámetros se han variado para cada uno de los tres análisis de sensibilidad.

2.2. Análisis en el dominio del tiempo

Para realizar el análisis en el dominio del tiempo se ha creado un script en R y se han importado los datos de todos los registros de ECG. En concreto, se cargan las posiciones de los latidos. A continuación, se calculan las distancias entre cada latido para construir los RRs. Antes de realizar los análisis, y con el fin de eliminar valores atípicos, se filtra la señal utilizando la función *FilterNIHR* [2]. Esta función emplea umbrales adaptativos para localizar y eliminar RRs con valores fisiológicamente inaceptables.

A continuación, se varían sistemáticamente los parámetros que afectan directamente al análisis. Tras cada ejecución se almacenan los índices resultantes. Estos índices incluyen *SDANN* (la desviación estándar de los promedios de intervalos RR en periodos cortos de tiempo), *SDNN index* (el promedio de las desviaciones estándar de los intervalos RR en periodos cortos de tiempo) y *TINN* (la interpolación triangular del histograma de RRs) [2].

Los parámetros que afectan al análisis en el dominio temporal son *size* y *numofbins*. *Size* indica la anchura de la ventana que se usa para analizar segmentos de la serie temporal RR. El valor por defecto es de 300 segundos, tal y como se sugiere en [1]. A este parámetro se le aplican incrementos del 1%, 2%, 5%, 20% y 50%. Este parámetro afecta directamente a los valores de los índices *SDNN index* y *SDANN*.

Numofbins indica el número de *bins*, es decir, el número de rangos que forman el histograma a partir del cual se calcula *TINN*. RHRV usa un número diferente de *bins* dependiendo del registro. El valor por defecto de cada registro se calcula a partir de los RR máximo y mínimo y un tamaño de *bin* por defecto, que es de 7.8125 ms. Al valor resultante se le aplican incrementos del 1%, 2%, 5%, 20% y 50%.

Seguidamente, se verifica si los resultados para cada índice siguen una distribución normal. Dado que los datos obtenidos no son normales, se emplea el test de Friedman para comparar el valor de cada índice con sus valores de referencia (los obtenidos al realizar los cálculos con los valores por defecto). El test de Friedman es una prueba no paramétrica comúnmente empleada cuando los datos son no-normales y dependientes entre sí. Si el test de Friedman es significativo implica que al menos dos grupos de entre los comparados son significativamente diferentes, pero no indica cuáles. Para saberlo es necesario comparar los grupos entre sí. Esto se realiza con un test Post-Hoc, en concreto el test de Conover [10] con corrección de Bonferroni.

2.3. Análisis en el dominio de la frecuencia

2.3.1 Transformada Corta de Fourier

Tras construir la serie RR y filtrar la señal, comienza el proceso de análisis específico para el dominio de la frecuencia. Primero se lleva a cabo la interpolación de la serie RR, ya que para las técnicas empleadas en este estudio se necesita una serie de frecuencia cardiaca muestreada uniformemente. Para ello se realiza una interpolación lineal con una determinada frecuencia de muestreo. Se realizan 4 interpolaciones con frecuencias de 2 Hz, 4 Hz, 8 Hz y 16 Hz. La elección de estos valores se basa en [1], que indica que el rango de frecuencias de una señal de VFC cubre aproximadamente 0-4 Hz.

Para realizar el análisis en el dominio de la frecuencia se utiliza la TCF variando los parámetros *size* y *shift*, que son respectivamente el tamaño de la ventana temporal empleada en la TCF, y el desplazamiento que se lleva a cabo entre ventanas temporales. Para el parámetro *size* se aplican incrementos del 1%, 2%, 5%, 20% y 50% respecto al valor por defecto (300 s). Para el parámetro *shift*, que indica el desplazamiento en segundos de la ventana, se prueban cuatro valores diferentes. Estos 4 valores están

equiespaciados entre 1s hasta *size*/2. Tras cada ejecución se almacenan los resultados del análisis.

Cabe destacar que los resultados que devuelve el análisis son las potencias en las bandas ULF, VLF, LF y HF en ventanas temporales de tamaño *size*. Por tanto, dichos índices son series temporales. Para simplificar el análisis, se extrae un único valor de cada serie usando una media. A continuación, se compara cada índice obtenido al variar los parámetros con los índices de referencia (los resultados obtenidos con los parámetros por defecto). A pesar de que los datos no son normales (test de *Shapiro-Wilks*), se decidió no obstante emplear un análisis de la varianza (ANOVA), el cual no asume normalidad. Esta decisión se tomó en base a la falta de implementaciones *software* sencillas de métodos no-paramétricos multifactoriales. Además, existen diversos estudios que demuestran que el test ANOVA es robusto en términos de error de tipo I (nivel de significancia) frente a la no-normalidad de los datos [11, 12]. Para simplificar el análisis Post-Hoc, y a pesar de que cuenta con las mismas limitaciones que ANOVA, se ha empleado el test de Tukey con corrección de Bonferroni. En este sentido, el análisis Post-Hoc debe considerarse una aproximación. Para medir el tamaño del efecto se usa el valor η^2, que describe el ratio de la varianza explicada en el índice por un parámetro en concreto.

2.3.2 Transformada Wavelet

Los pasos iniciales de construcción de la serie temporal y filtrado de latidos no normales se realizan del mismo modo que en el caso anterior. Para la interpolación de la serie temporal, se prueban tres valores de la frecuencia de interpolación (parámetro *freqhr*); 2, 4 y 8 Hz. También se varía el parámetro *wavelet* que indica el tipo de Wavelet madre utilizado para el análisis. Se usan los siguientes tipos: Haar ("haar"), Wavelet de Daubechies de fase extrema ("d4", "d6", "d8", "d16") y Wavelet de Daubechies de mínima asimetría ("la8", "la16", "la20"). El valor por defecto para este parámetro en RHRV es "d4".

El tercer parámetro que se varía es *bandtolerance.* Al realizarse la TW se busca extraer la energía en las bandas ULF, VLF, LF y HF. Sin embargo, dado que las Wavelets descomponen la señal en bandas de frecuencia diádicas (del tipo $[\frac{f_s}{2^{j+1}}, \frac{f_s}{2^j}]$, siendo j un entero y f_s la frecuencia de muestreo), no siempre es posible extraer las bandas de VFC de forma exacta. El parámetro *bandtolerance* indica el máximo error permitido (en Hz) al aproximar una banda de VFC con Wavelets [4]. El valor por defecto es de 0.01 Hz. En el análisis de sensibilidad se emplearán los valores 0.01, 0.005 y 0.002, 0.001 y 0.0001 Hz.

Tras cada ejecución se calcula la potencia media en cada banda de frecuencia y se almacenan los resultados. A continuación, se compara cada índice obtenido con los índices de referencia. Al realizar el test estadístico se observa que los datos no siguen una distribución normal (*Shapiro-Wilks*) pero por las razones ya mencionadas, se aplica el test ANOVA [11, 12].

3. Resultados

3.1. Análisis en el dominio temporal

Al aplicar la prueba de Friedman a los índices estudiados (*SDANN, SDNN index* y *TINN*) se obtienen p-valores muy bajos, lo que indica que la variación de los parámetros afecta de manera significativa a los índices resultantes. Además, el tamaño de los efectos es grande. Sin embargo, se desconoce a partir de qué variación obtenemos índices significativamente diferentes a los índices de referencia. Para ello se aplica un test Post-Hoc de Conover que indica a partir de qué porcentaje de variación la diferencia es significativa, como muestra la Tabla 1.

Índice	Primera variación significativa	P-valores ajustados
SDANN	10%	4.0×10^{-4}
SDNN index	2%	7.0×10^{-3}
TINN	1%	5.5×10^{-8}

Tabla 1. *Resultados de la prueba Post-Hoc para los estadísticos en el dominio del tiempo.*

3.2. Análisis en dominio de la frecuencia

Se emplea un nivel de significancia del 0.01. Al aplicar el análisis realizado con la TCF, los factores que pueden afectar de forma significativa a los resultados son los parámetros *size*, *shift* y *freqhr*. Para todas las bandas de frecuencia se obtienen dos resultados comunes: el parámetro *shift* no es significativo y el parámetro *size* tiene un efecto muy pequeño (la suma de sus cuadrados es aproximadamente 0) aunque ocasionalmente significativo.

El parámetro *freqhr* afecta de forma significativa a las componentes VLF, LF y HF, si bien los resultados del análisis Post-Hoc son diferentes en cada una de estas bandas. En la prueba Post-Hoc para las bandas VLF y LF se observa que las diferencias obtenidas respecto al valor por defecto no son significativas (es decir, hace falta comparar valores extremos como 2 y 8 Hz para obtener diferencias significativas). En cambio, la componente HF muestra diferencias significativas para el factor *freqhr* a partir del doble de su valor por defecto. A pesar de que los resultados para *freqhr* son estadísticamente significativos, cabe destacar que su efecto es realmente muy pequeño. De hecho, la suma de sus cuadrados es prácticamente 0.

Por otro lado, en el análisis con la TW los parámetros que se tienen en cuenta en la prueba ANOVA son *wavelet* y *bandtolerance.* Al variar el parámetro *wavelet* las conclusiones obtenidas para las bandas espectrales ULF, VLF y LF es similar; las variaciones aplicadas sí producen cambios significativos, pero la suma de los cuadrados es aproximadamente 0, lo que significa que el efecto es muy pequeño. No obstante, en el caso de la componente HF todos los cambios producen un efecto no despreciable.

Respecto al parámetro *bandtolerance*, este afecta de forma diferente a las distintas bandas de frecuencia. En el caso de la banda ULF, la suma de los cuadrados es aproximadamente 0; es decir, tiene un efecto pequeño en

los resultados. Para VLF el efecto no es despreciable, excepto cuando se reduce el valor por defecto del parámetro a la mitad (0.005 Hz). Para LF y HF los cambios en el parámetro *bandtolerance* siempre tienen un efecto no despreciable. En cuanto al parámetro *freqhr* los resultados muestran que no tiene un impacto significativo en los resultados del análisis.

4. Discusión y conclusiones

En el análisis en tiempo, *TINN* es el índice más sensible ya que con tan solo una variación del 1% se obtiene una variación significativa en su valor. Lo mismo ocurre también con el índice *SDNN index*, cuyo resultado es significativo a partir de un 2% de variación. Por último, *SDANN* no varía significativamente hasta que la variación en el parámetro es de un 10%. De esto se concluye que al realizar un análisis en el dominio del tiempo debe prestarse atención a la elección de los parámetros de configuración, ya que variaciones pequeñas pueden afectar a las conclusiones del estudio.

En cuanto a la sensibilidad de los parámetros de la TCF, cabe destacar que las sumas de los cuadrados para los parámetros estudiados son todas aproximadamente 0. El hecho de que existan diferencias significativas entre algunos de sus valores debe atribuirse a que la variación producida por estos, aunque casi despreciable, tiene una varianza muy pequeña. Por tanto, podemos considerar que el análisis de la frecuencia con la TCF no es sensible a los parámetros empleados.

En la TW se observa que la banda de HF es sensible a los cambios en el parámetro *wavelet*. Este resultado es esperado ya que distintas madres Wavelet tienen distintas resoluciones espectrales y esta característica de las wavelets se manifiesta más claramente en las frecuencias más altas (cuanto más alta es la frecuencia analizada más baja es la resolución espectral de la TW). Por otra parte, los resultados de ANOVA muestran que las únicas diferencias no significativas se dan cuando se emplean Wavelets con resolución espectral similar (por ejemplo, "la16"- "d16", "la8"- "d8"). Respecto a *bandtolerance*, las bandas de frecuencia VLF, LF y HF son sensibles a cambios en este parámetro. Aunque significativo, el efecto en ULF es pequeño, por lo que puede considerarse que la potencia en ULF no es sensible a *bandtolerance*. Por último, el parámetro *freqhr* no parece afectar a los índices calculados.

Por tanto, a la hora de realizar un análisis mediante la TW, debemos prestar atención a la selección del parámetro *bantolerance* y a la selección de la madre Wavelet. El uso de madres Wavelets con resoluciones espectrales distintas puede producir resultados distintos en la componente HF de la frecuencia. Sin embargo, cambiar una madre wavelet por otra de características espectrales similares no afectará a los índices resultantes.

En ambos análisis en frecuencia se han utilizado pruebas ANOVA para simplificar el análisis y por ser robusto en términos de error de tipo I frente a la no-normalidad de los datos. Como trabajo futuro deberían rehacerse estos tests usando modelos avanzados que permitan capturar la no-normalidad de los datos. También sería interesante realizar el estudio con más muestras y sobre distintas patologías para así poder obtener conclusiones generalizables a cualquier estudio de VFC.

Agradecimientos

Esta investigación ha sido financiada por el Ministerio de Ciencia, Innovación y Universidades, y por la Unión Europea a través del Fondo Europeo de Desarrollo Regional mediante el proyecto RTI2018-095324-B-I00.

Referencias

[1] Camm A, Malik M, Bigger J, Breithardt G, Cerutti S, Cohen R, Coumel P, Fallen E, Kennedy H, Kleiger R y otros. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Annals of Noninvasive Electrocardiology*, vol. 1(2), 1996, pp 151-181.

[2] García CA, Otero A, Vila XA, Touriño MJ, Rodríguez-Liñares L, Presedo JM, Penín AJ. Heart rate variability analysis with the R package RHRV. Springer International Publishing, 2017.

[3] Acharya U, Joseph K, Kannathal N, Lim C, Suri J. Heart rate variability: a review. *Medical and biological engineering*

[4] Thayer JF, Åhs F, Fredrikson M, Sollers III JJ, Wager TD. A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. *Neuroscience & Biobehavioral Reviews*, 36(2), 2012, pp 747-756.

[5] García CA, Otero A, Vila XA, Márquez, D. A new algorithm for wavelet-based heart rate variability analysis. *Biomedical Signal Processing and Control*, 8(6), 2013, pp 542-550.

[6] Lado MJ, Vila XA, Rodríguez-Liñares L, Penín AJ, Olivieri DN, Félix P. Detecting sleep apnea by heart rate variability analysis: assessing the validity of databases and algorithms. *Journal of medical systems*. 35(4), 2011, pp 473-481.

[7] Página web del paquete RHRV en CRAN. https://cran.r-project.org/web/packages/RHRV/index.html. [Consultada: Octubre 2020].

[8] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*. 101(23), 2003, pp 215-220

[9] Sullivan GM, Feinn R. Using effect size—or why the P value is not enough. *Journal of graduate medical education.* 4(3), 2012, pp 279-282.

[10] Zimmerman D, Zumbo B. Relative power of the Wilcoxon test, the Friedman test, and repeated-measures ANOVA on ranks. *The Journal of Experimental Education.* 62(1), 1993, pp 75-86.

[11] Blanca M, Alarcón R, Arnau J, Bono R, Bendayan R. Non-normal data: Is ANOVA still a valid option? *Psicothema*, 29(4), 2017, pp 552-557.

[12] Glass G, Peckham P, Sanders J. Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of educational research*, 42(3), 1972, pp 237-288.

Diseño, desarrollo y evaluación de un sistema de detección de estrés en conductores basado en la señal de actividad electrodermal

L. Valenzuela López, M. Barrachina Fernández, C. Sánchez Ávila

Departamento de Matemática Aplicada a las Tecnologías de la Información y las Comunicaciones,
ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, España,
laura.valenzuelal@alumnos.upm.es
mariadelasmercedes.barrachina.fernandez@alumnos.upm.es
carmen.sanchez.avila@upm.es

Resumen

Es frecuente que las personas experimenten situaciones de estrés en su día a día. Sin embargo, dichas situaciones pueden resultar perjudiciales para la salud. Por ello, resulta útil evaluar el nivel de estrés al que una persona se encuentra sometida. El objetivo del presente trabajo es analizar la viabilidad de utilizar una única señal fisiológica, la señal de actividad electrodermal (EDA), para la detección del estrés psicológico en conductores. Para ello, se han utilizado los datos de EDA pertenecientes a la base de datos pública "driverdb" del MIT. Tras un procesamiento previo de las señales de EDA, se han desarrollado e implementado dos algoritmos clasificadores, máquinas de vector soporte (SVM) y análisis discriminante lineal (LDA), con el fin de predecir el nivel de estrés de cada registro de las señales de EDA. Los resultados mostraron llegar a exactitud del 81,40% para SVM, lo que demuestra que la señal EDA puede ser utilizada como única señal fisiológica para detectar niveles de estrés bajos y altos con un buen grado de precisión, e invita a continuar con la línea de investigación estudiada para predecir el nivel de estrés.

1. Introducción

En la actualidad es cada vez más frecuente el interés en mejorar la salud de las personas, tanto física como mental. El estrés, tiene un efecto directo sobre la salud mental, sin embargo, también es responsable de numerosas afecciones cardiovasculares, digestivas y neurológicas, entre otras. Por otro lado, el estrés también puede ocasionar situaciones de riesgo de forma indirecta, por ejemplo, accidentes de tráfico [1]. Sólo en España, el estrés afecta a casi 18 millones de personas de manera continuada, llegando a sufrirlo el 96% de la población en algún momento del año, según [2], por lo que muchos han bautizado al estrés como "la enfermedad del siglo XXI".

Debido a ese mayor interés por la salud, hoy en día se realizan numerosos controles de detección y diagnóstico de enfermedades a la ciudadanía. Muchos de estos controles conllevan procesos complejos y costosos, por lo que resulta inevitable la necesidad de simplificarlos para así conseguir procesos más eficientes y asequibles a una mayor parte de la población.

Por otro lado, para llevar a cabo ese seguimiento y control de la salud de las personas, cada vez es más habitual recurrir a la información proporcionada por tecnologías biomédicas, especialmente por señales biomédicas como son el electrocardiograma, encefalograma, etc. El estudio de estos datos, obtenidos a través de la monitorización del paciente, también es cada vez más frecuente para poder entender mejor todo lo que sucede en el organismo. La señal de actividad electrodermal o EDA, es un ejemplo de señal biomédica que está siendo cada vez más utilizada en múltiples estudios, especialmente en estudios relacionados con el estrés, aunque también en muchos otros relacionados con el dolor, la depresión o el reconocimiento de emociones.

Por lo tanto, un mejor entendimiento de esta señal, así como un análisis más detallado de sus principales características, en diferentes situaciones (reposo, estrés, dolor, enfado, etc.) podría ayudar enormemente a optimizar procesos de detección y diagnóstico de diferentes alteraciones fisiológicas en las personas.

Al tratarse el estrés de un estado generalmente subjetivo, su manejo y medición no es sencillo. Sólo pueden obtenerse medidas objetivas y precisas a partir de mediciones fisiológicas de cada individuo. La señal de actividad electrodermal o actividad electrodérmica (EDA) es un término que engloba todas las propiedades eléctricas de la piel, las glándulas sudoríparas y sus apéndices. La secreción de estas glándulas, que ocurre especialmente cuando un individuo se encuentra sometido a situaciones de estrés, es la causante de variaciones de resistencia y conductividad en la piel [3]. Por otro lado, la EDA se considera como una medida de la actividad del sistema nervioso autónomo (SNA). El SNA es el encargado del control de las funciones viscerales involuntarias del organismo, lo cual realiza a través de dos vías antagónicas: el sistema nervioso simpático y el sistema nervioso parasimpático. La activación de la rama simpática del SNA se encarga de mediar en la respuesta de estrés hormonal, conduciendo a una excitación generalizada de los órganos diana. Por el contrario, la rama parasimpática se encarga de mantener un estado de relajación tras la situación de estrés

ISBN: 978-84-09-25491-0, pags. 97- 100

o la realización de un esfuerzo [4]. Todos los órganos del cuerpo están conectados a nervios tanto simpáticos como parasimpáticos que funcionan simultáneamente, pero de manera antagónica, para mantener el equilibrio del organismo. Sin embargo, las glándulas sudoríparas y los vasos sanguíneos de la piel se encuentran únicamente inervados por el sistema nervioso simpático, siendo por tanto la piel una excepción en el funcionamiento del SNA [5]. Como consecuencia de ello, el SNA producirá un mayor nivel de sudoración ante la presencia de un estresor, a través de la inervación sudomotora, pero esta acción no se detendrá con la misma rapidez cuando el estímulo cese. Esto hace que la conductancia de la piel, una medida de EDA, sea un medio ideal para cuantificar la activación simpática del organismo [4]. Por todo ello, estudios recientes destacan el uso de la señal de EDA como medida crucial en los sistemas de detección de estrés.

El principal objetivo de este trabajo es analizar la viabilidad de utilizar una única señal fisiológica, la señal de EDA, para la detección del nivel de estrés en conductores. En situaciones de conducción de vehículos, el estrés sufrido por el conductor supone un incremento en el riesgo de sufrir accidentes de tráfico [1]. Por ello, su detección y evaluación temprana es clave para evitar este tipo de accidentes. Para ello, se propone el empleo de dos algoritmos clasificadores (SVM y LDA) como métodos para la detección y clasificación objetiva del estrés, basándose en los datos obtenidos únicamente a partir de señales de EDA.

Como objetivos secundarios, necesarios para poder alcanzar la finalidad de este trabajo, se establecen dos puntos fundamentales: en primer lugar, el análisis de diferentes bases de datos para obtener y formar los datos de entrada de los algoritmos; y en segundo lugar, el diseño y desarrollo de un sistema de detección de estrés basado solo en la señal de EDA, lo que a su vez requiere de un adecuado procesado de las señales seleccionadas, poniendo en práctica técnicas de filtrado, segmentación y análisis de características, y posteriormente su correcta clasificación.

2. Materiales y métodos

Para realizar este estudio, se ha utilizado la base de datos pública "Driver Stress Data", generada por los investigadores Jennifer Healey y Rosalind W. Picard, pertenecientes al grupo de investigación "Affective Computing" del Instituto Tecnológico de Massuchussets (MIT). Esta también ha sido utilizada en otros artículos, como [6] y [7], debido a que es una de las principales bases de datos existentes relacionadas con el estrés estudiado en conductores.

Con estos datos, se llevaron a cabo tres pasos principales: la obtención de los registros de EDA, el tratamiento de la señal y su clasificación a través de la implementación de los diferentes algoritmos.

2.1. Obtención de los registros de EDA

Los datos del estudio original del MIT [8] se obtuvieron a través de la realización de un protocolo de conducción específico en el área metropolitana de Boston, que incluía tareas de descanso, conducción en ciudad y conducción en autopista, a realizar por 15 conductores. De todos los datos recopilados en dicho estudio, en el presente trabajo se han utilizado únicamente las señales de EDA emitidas por un sensor Empatica E4 colocado en un tobillo de cada conductor, junto con una señal de validación del nivel de estrés que posibilita diferenciar las distintas etapas de la ruta en las que se hallaba el conductor en cada momento. Dicho protocolo de conducción consistía en una ruta con 7 secciones diferenciadas. El primer y el último tramo de la ruta correspondían a las situaciones de descanso. En los 5 tramos intermedios se alternaban los dos tipos de conducción (comenzando y acabando en la ciudad). Así, se realizó una asociación entre la situación de la ruta de conducción y el nivel de estrés, para la posterior clasificación, en la cual los periodos de descanso corresponden a niveles de estrés bajo, mientras que la conducción tanto en ciudad como en autopista corresponden a niveles de estrés alto. Esta asociación permitirá confirmar la presencia o ausencia de estrés en cada situación.

Una vez obtenida la información de cada fichero de datos, se realizó una evaluación de los datos obtenidos de cada conductor. De los 15 conductores del estudio inicial, sólo se han utilizado las señales de 11 de ellos, ya que no se había obtenido la señal de EDA y/o información clara de los marcadores de 4 de los conductores. En la *Figura 1* se muestra la representación de la señal EDA obtenida del conductor 4 durante el protocolo de conducción.

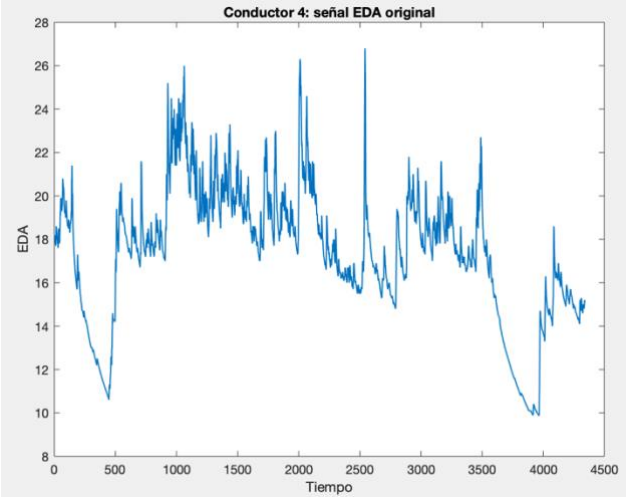


Figura 1. *Ejemplo de la representación de la señal EDA del conductor 4.*

Las señales de los 11 conductores fueron unidas de forma consecutiva obteniendo una única señal de EDA para facilitar su posterior tratamiento.

2.2. Tratamiento de la señal EDA

La señal EDA puede sufrir diversos tipos de artefactos, ruido e interferencias. Especialmente, pueden encontrarse artefactos no estereotipados, que son aquellos que se producen sobretodo debido a un mal contacto entre el electrodo y la piel y movimientos de los electrodos [9]. Por ello, es necesario realizar un procesamiento previo para filtrar la señal con el fin de atenuar dichos artefactos y obtener una señal más útil para el estudio. En primer lugar, se aplicaron distintos filtros paso bajo de tipo Butterworth

a diferentes frecuencias de corte, obteniendo resultados óptimos a una frecuencia de corte de 0.05Hz. La señal de salida obtenida, a continuación, fue pasada por un filtro suavizado.

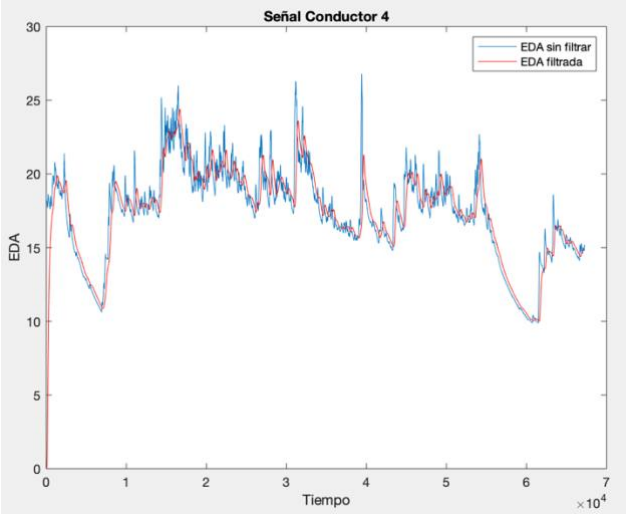


Figura 2. *Señal EDA del conductor 4 antes y después de la eliminación de artefactos.*

Tras el filtrado, la señal de EDA, junto con los marcadores de ruta, se segmentó en fragmentos de 12 segundos, cada uno asociado a un único tipo de ruta, con el fin de utilizar segmentos más pequeños de la señal EDA que permitan conseguir una óptima clasificación por los algoritmos. Posteriormente, de cada segmento se extrajeron las características más destacables de la señal en el dominio del tiempo para poder realizar la clasificación mediante los diferentes algoritmos. Dichas características son: media, mediana, desviación estándar, valor mínimo, valor máximo, media de la primera derivada, mediana de la primera derivada, desviación estándar de la primera derivada, valor mínimo de la primera derivada, valor máximo de la primera derivada, media de la segunda derivada, mediana de la segunda derivada, valor mínimo de la segunda derivada, valor máximo de la segunda derivada.

2.3. Implementación de los algoritmos

Las Máquinas de Vectores Soporte (Support Vector Machine, SVM) son una técnica de aprendizaje automático o "machine learning" que permite encontrar la mejor separación posible de datos entre clases diferentes. Utiliza un espacio de hipótesis de funciones lineales en un espacio de características de alta dimensión para realizar una clasificación supervisada. En él, busca encontrar el hiperplano de separación óptimo entre distintas clases basándose en casos de entrenamiento previos [10]. Como entradas al modelo de SVM se han utilizado una matriz de datos predictores, es decir, la matriz formada por las características obtenidas de la señal, y una matriz de etiquetas que corresponde con el vector caracterizado por el nivel de estrés para cada tramo de la ruta. Se han dividido los datos en dos subconjuntos (entrenamientos y test) y se han probado tres tipos de núcleo, lineal, de base radial y polinómico, para evaluar con cuál de ellos se obtiene una mejor clasificación.

El Análisis Discriminante Lineal de Fisher, más comúnmente conocido como Análisis Discriminante Lineal (Linear Discrimiant Analysis, LDA) es una técnica estadística utilizada en reconocimiento de patrones y aprendizaje automático cuya finalidad consiste encontrar una combinación lineal de características que permiten separar dos o más clases de datos en función de si existen diferencias significativas entre ellas [11]. En este caso, las entradas al algoritmo coinciden con las del SVM. Además, también se han utilizado tres tipos de discriminación: lineal, diaglineal y pseudolineal [12].

En ambos modelos, la medida que se ha utilizado para evaluar el rendimiento de los algoritmos es la exactitud o tasa correcta ("Correct Rate"), que muestra el número de muestras predichas correctamente dividido por la cantidad total de muestras. Así mismo, se han obtenidos los datos de sensibilidad y especificidad para cada caso.

3. Resultados

Las señales de EDA de 11 sujetos experimentales se recopilaron y dividieron en fragmentos de 12 segundos, cada uno de ellos asociado a un nivel de estrés, obteniendo un total de 4.301 registros. A continuación, se determinaron las 15 características de cada uno de los fragmentos que se detallan en la sección 2.2, que constituyeron la entrada de datos para los algoritmos clasificadores. Tanto el algoritmo SVM como el algoritmo LDA, se entrenaron para construir modelos de clasificación binarios que predicen si un individuo está sometido a estrés o no.

Para evaluar la capacidad predictiva de ambos algoritmos se utilizó la exactitud, además de los valores de sensibilidad y especificidad alcanzados en la predicción.

Se han probado los tres tipos de núcleo para el caso del SVM, y los tres tipos distintos de discriminación para el caso del LDA, obteniéndose los resultados que aparecen en la Tabla 1 y *Tabla 2* respectivamente.

Núcleo	Exactitud	Sensibilidad	Especificidad
Lineal	75.35%	61.33%	93.95%
Base radial	81.40%	63.75%	82.61%
Polinómico	65.70%	83.69%	54.44%

Tabla 1. *Resultados del algoritmo SVM en función del tipo de núcleo.*

Tipo de discriminación	Exactitud	Sensibilidad	Especificidad
Lineal	76.51%	58.18%	87.92%
Diag_Lineal	67.09%	70.94%	64.44%
Pseudo_Lineal	73.14%	53.85%	83.42%

Tabla 2. *Resultados del algoritmo LDA en función del tipo de discriminación.*

4. Discusión

Los resultados obtenidos a través del algoritmo SVM están en línea con los obtenidos en [13], donde obtuvieron una

precisión de entre el 75.0% y el 87.7% (dependiendo de la longitud del segmento de señal utilizado) a través de un núcleo lineal, mientras que en nuestro estudio la precisión ha sido del 75.35%. Por otro lado, en el caso del LDA, los resultados son muy similares a los obtenidos en [6], donde el mismo método logró una exactitud de en torno al 80%, siendo en nuestro caso del 76.51%. Esto, por tanto, sugiere continuar trabajando en esta dirección.

Así mismo, respecto a los niveles de estrés detectados, debido a que el tipo de clasificación empleado es biclase, se ha considerado que los dos tipos de conducción (en ciudad y en autopista) proporcionan un nivel de estrés alto, mientras que los periodos de descanso proporcionan niveles de estrés bajo, consiguiendo detectar así la presencia o ausencia de estrés. Sin embargo, para futuras investigaciones sería conveniente distinguir entre el nivel de estrés experimentado en la conducción en autopista y en la conducción en ciudad.

5. Conclusiones

El objetivo del presente trabajo ha sido analizar la viabilidad de utilizar la señal de EDA como única señal fisiológica para la detección del estrés psicológico en conductores, sin la necesidad de utilizar señales fisiológicas adicionales. Para ello, a través de las señales de EDA de 11 conductores, se ha realizado una clasificación biclase para detectar la presencia o ausencia de estrés a través de dos algoritmos clasificadores: SVM y LDA.

El algoritmo SVM logró una exactitud en la predicción del estrés del 81,40%. Por otra parte, el algoritmo LDA logró una exactitud del 76,51%. Estos resultados demuestran que la señal EDA, como única señal fisiológica, puede ser utilizada para detectar niveles de estrés bajos y altos con un buen grado de precisión, siendo la detección de estrés, de esta manera, más simple y práctica, pudiéndose realizar en tiempo real e integrarse, por ejemplo, en sistemas de ayuda a la conducción, alertando al conductor de un posible nivel alto de estrés. Por tanto, los resultados obtenidos sugieren continuar con la línea de investigación estudiada y utilizar el algoritmo SVM como método de detección de estrés utilizando únicamente los datos obtenidos de la señal EDA.

En futuras investigaciones sería conveniente analizar la viabilidad de estos métodos en la detección de más niveles de estrés. Para ello, crear conjuntos de datos de EDA más grandes, con más sujetos, consolidaría los resultados hallados en este trabajo y permitiría detectar una diversidad mayor de niveles de estrés. Además, podrían utilizarse características adicionales de la señal (a los ya utilizados en este trabajo) para mejorar las tasas de predicción. También se podrían utilizar características de la señal en el dominio de la frecuencia o utilizar otros algoritmos de aprendizaje automático como, por ejemplo, las redes neuronales o los árboles de decisión. Adicionalmente, también se podrían añadir a las características de entrada de los algoritmos seleccionados, determinadas características de otras señales fisiológicas (tasa cardiaca, electrocardiograma, etc.), para mejorar la predicción relativa al estado de estrés. Finalmente, una clara línea de trabajo es la confección de una nueva base de datos con un número suficientemente amplio de usuarios que permita extraer conclusiones estadísticas más significativas.

Referencias

[1] Holt, P. L. (1981). Stressful life events preceding road traffic accidents. *Injury*, *13*(2), 111-115.

[2] Página web del VII Estudio CinfaSalud. https://cinfasalud.cinfa.com/p/estudio-cinfasalud-estres/

[3] Boucsein, W. (2012). Electrodermal activity: Second edition. In *Electrodermal Activity: Second Edition*, 9781461411260. Springer US.

[4] Anusha, A. S., Joy, J., Preejith, S. P., Joseph, J., & Sivaprakasam, M. (2017). Differential effects of physical and psychological stressors on electrodermal activity. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 4549–4552.

[5] Setz, C., Arnrich, B., Schumm, J., La Marca, R., Tröster, G., & Ehlert, U. (2010). Discriminating stress from cognitive load using a wearable eda device. *IEEE Transactions on Information Technology in Biomedicine*, 14(2), 410–417.

[6] Liu, Y., & Du, S. (2018). Psychological stress level detection based on electrodermal activity. *Behavioural brain research*, *341*, 50-53.

[7] A. Ghaderi, J. Frounchi and A. Farnam, "Machine learning-based signal processing using physiological signals for stress detection," 2015 22nd Iranian Conference on Biomedical Engineering (ICBME), Tehran, 2015, pp. 93-98, doi: 10.1109/ICBME.2015.7404123.

[8] Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, *6*(2), 156–166.

[9] Granero, A. C., Ornedo, V. N., Provinciale, J. G., & Alcañiz, M. (2014). Adquisición , Acondicionamiento y Análisis de Señales Fisiológicas para la Evaluación de Contenidos Audiovisuales. Noviembre.

[10] Tzotsos, A., & Argialas, D. (2008). Support vector machine classification for object-based image analysis. *Lecture Notes in Geoinformation and Cartography*, *9783540770572*, 663–677.

[11] Zhai, J., Barreto, A. B., Chin, C., & Li, C. (2005). Realization of stress detection using psychophysiological signals for improvement of human-computer interactions. *Conference Proceedings - IEEE SOUTHEASTCON*, 415–420.

[12] Xu, P., Brock, G. N., & Parrish, R. S. (2009). Modified linear discriminant analysis approaches for classification of high-dimensional microarray data. *Computational Statistics & Data Analysis*, *53*(5), 1674-1687.

[13] Sánchez-Reolid, R., Martínez-Rodrigo, A., & Fernández-Caballero, A. (2019, June). Stress identification from electrodermal activity by support vector machines. In *International Work-Conference on the Interplay Between Natural and Artificial Computation* (pp. 202-211). Springer, Cham.

Cambios en los loops de las ondas P del vectocardiograma tras el aislamiento de las venas pulmonares

N. Ortigosa[1], O. Cano[2], F. Sandberg[3]

[1] I.U. Matemática Pura y Aplicada, Universitat Politècnica de València, nuorar@upvnet.upv.es

[2] Servicio de Cardiología, Hospital Universitari i Politècnic La Fe, Valencia cano_osc@gva.es

[3] Department of Biomedical Engineering, Lund University, Suecia

Resumen

La fibrilación auricular es la arritmia más frecuente en la práctica clínica. Además de los tratamientos antiarrítmicos, el aislamiento de las venas pulmonares mediante ablación por catéter es uno de los tratamientos utilizados acorde a las guías clínicas para ayudar a mantener el ritmo sinusal de forma indefinida en el paciente.

El propósito de este estudio es analizar cómo dicho procedimiento afecta a los patrones de conducción eléctrica en las aurículas, y cómo estos se ven modificados en los registros electrocardiográficos registrados al paciente. Para ello, se ha comparado la morfología de las ondas P de los loops extraídos de los vectocardiogramas calculados a partir del electrocardiograma de superficie, antes y después del procedimiento en pacientes con fibrilación auricular paroxística. Estos pacientes tienen episodios autolimitados, permaneciendo gran parte del tiempo en ritmo sinusal y sin un alto grado de evolución de la arritmia. El primer y segundo autovalor de los loops, así como una medida de la redondez de los mismos señalaron cambios significativos tras el procedimiento, mostrando que el aislamiento de las venas pulmonares induce cambios morfológicos en los loops de las ondas P.

1. Introducción

La fibrilación auricular (FA) es la arritmia más frecuente en la práctica clínica diaria y, además, su prevalencia va en aumento. Los estudios apuntan a que en la Unión Europea habrá de 14 a 17 millones de pacientes con FA en el año 2030 [1]. Esta arritmia se caracteriza por un ritmo cardiaco rápido e irregular, donde las aurículas laten de forma desincronizada respecto de los ventrículos [2].

Los tratamientos de la FA se basan en un control del ritmo y de la frecuencia cardiaca. Incluyen medicación antiarrítmica, pero también técnicas de ablación por catéter para aislar las venas pulmonares, sobre las que se ha demostrado que se origina la recirculación eléctrica que es origen de la arritmia [3].

Las tasas de éxito del aislamiento de las venas pulmonares alcanzan el 80% para los pacientes con FA paroxística (es decir, aquellos en los que los episodios de arritmia son recurrentes y autolimitados, con duración inferior a 7 días). Sin embargo, se ha visto que el éxito de la intervención disminuye a medida que pasa el tiempo durante las sucesivas revisiones [4].

De este modo, el objetivo del aislamiento de las venas pulmonares en pacientes con FA es el restablecimiento del ritmo sinusal en el paciente, así como su mantenimiento a largo plazo [5], siendo de interés estudiar qué cambios produce dicho procedimiento para que resulte duradero o no.

Estudios previos han analizado cómo el aislamiento de las venas pulmonares se ve reflejado en cambios en la señal electrocardiográfica, fundamentalmente afectando a la duración de las ondas P [6-7], así como a su amplitud y dispersión de la duración [8-10].

En el presente trabajo proponemos un estudio del vectocardiograma (VCG) para analizar cómo el aislamiento de las venas pulmonares afecta al patrón de conducción eléctrico en las aurículas. Para ello, a partir del electrocardiograma de superficie de pacientes con fibrilación auricular paroxística (con tramos en ritmo sinusal), extraemos el vectocardiograma y, tras delinear las ondas P, comparamos los loops asociados antes y después del procedimiento.

2. Materiales

La cohorte de este estudio está conformada por 10 pacientes que padecían fibrilación auricular paroxística, y que se sometieron al aislamiento de las venas pulmonares para tratársela en la Unidad de Arritmias del Hospital Universitari i Politècnic La Fe de Valencia.

El electrocardiograma de superficie de 12 derivaciones fue registrado antes y media hora después del procedimiento, almacenando registros de 60 segundos de duración. Dichos registros tenían las ondas P visibles, ya que los pacientes estaban en ritmo sinusal al ser intervenidos. De esta manera, las ondas P del electrocardiograma eran visibles en los registros adquiridos.

3. Métodos

3.1. Preprocesado de la señal

Los registros electrocardiográficos fueron preprocesados para eliminar el ruido y artefactos presentes en ellos. Para eliminar el ruido de red, un filtro Notch a 50Hz fue aplicado en primer lugar. Después, el ruido de base fue suprimido aplicando interpolación cúbica por splines, acorde al método presentado en el trabajo [11].

ISBN: 978-84-09-25491-0, pags. 101- 104

A continuación, el vectocardiograma fue calculado mediante la utilización de la matriz de Kors [12], obteniéndose así las derivaciones ortogonales de Frank.

Tanto el inicio como el final de las ondas P fueron automáticamente delineadas, siendo posteriormente revisadas las marcas de forma manual. Para el estudio, únicamente se han considerado las ondas P que tienen una calidad de señal suficiente. Así, para ser incluidas en el estudio, las ondas P deben tener una correlación cruzada superior a 0.9 con el promedio de las ondas P. Las ondas P que no cumplan este requisito serán excluidas del estudio.

3.2. Análisis del vectocardiograma

En primer lugar se construyeron los loops asociados a las ondas P del vectocardiograma. Para cada loop, se calcularon los tres autovectores de los mismos mediante descomposición en valores singulares. Cada autovector está asociado a un autovalor, que representa la energía correspondiente en dicha dimensión.

Dichos autovalores se denotan por λ_i, donde $i = 1, 2, 3$, y los autovalores se denotan en orden descendente ($\lambda_1 > \lambda_2 > \lambda_3$). De este modo, el primer autovalor (λ_1) representa la magnitud del loop en la dirección de la máxima variación. En la Figura 1 se ha incluido un ejemplo de dos loops de ondas P (antes y después del aislamiento de las venas pulmonares), así como de los respectivos primeros autovectores.

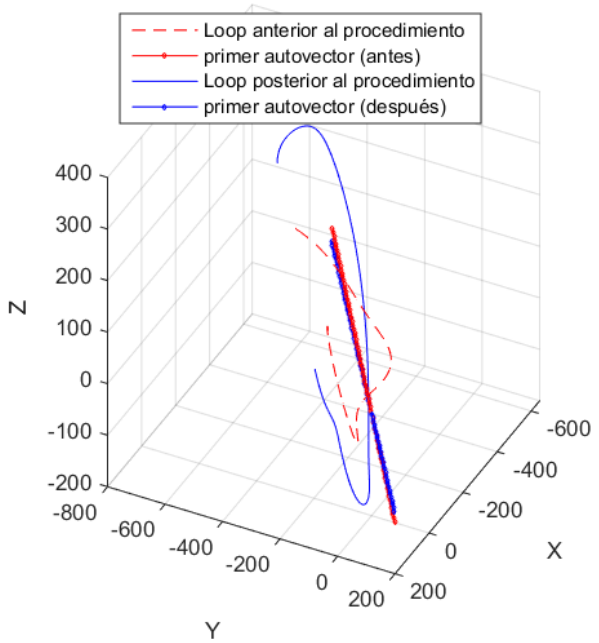


Figura 1. *Ejemplos de dos loops correspondientes a las ondas P de un paciente antes (en rojo) y después (en azul) del aislamiento de las venas pulmonares. Se incluyen también los respectivos primeros autovectores.*

En este trabajo se han analizado el primer y segundo autovalor de los loops de las ondas P (λ_1, λ_2), así como la relación del segundo autovalor frente al primero ($\rho = \lambda_2/\lambda_1$). Esta ratio describe la morfología del loop y cómo se distribuye el mismo a lo largo de sus dos diagonales principales, aportando una medida de cuán redondo es.

4. Resultados

Al eliminar las ondas P cuya calidad de señal no era suficientemente buena (aquellas cuya correlación cruzada era inferior a 0.9), quedaron para incluir en el estudio el 69% de las ondas P inicialmente presentes en los registros electrocardiográficos anteriores al aislamiento de las venas pulmonares, y el 75% de las ondas P procedentes de los registros posteriores al procedimiento.

En la Figura 2, a modo de ejemplo, se han representado gráficamente los loops de las ondas P de un paciente antes (primera gráfica) y después (segunda gráfica) del aislamiento de las venas pulmonares. Además, en la tercera gráfica, para facilitar la comparativa visual, se ha incluido un detalle de 10 loops de las subfiguras anteriores.

En estas figuras, se puede observar que no hay cambios en la fase de los vectores de los loops al comparar antes y después, pero sí hay cambios en la morfología. Estos cambios se reflejan en las características analizadas: tanto el primer como el segundo autovalor se incrementan tras el procedimiento, y la medida de la morfología de los loops se reduce, siendo estos menos redondos después del mismo.

A continuación, se muestra la Tabla 1, en la que se indican los resultados numéricos de las características analizadas, el promedio para los pacientes junto con la desviación estándar, así como el p-valor calculado mediante el test de t-Student pareado, que indica si dichos cambios son estadísticamente significativos, ya que las variables bajo estudio siguen una distribución normal.

Características	Antes	Después	p-valor
λ_1	5800±3000	7600±3200	0.0023*
λ_2	1700±500	1900±700	0.3241
ρ	0.35±0.19	0.30±0.20	0.0378*

Tabla 1. *Resultados promedio para los pacientes incluidos en la cohorte bajo estudio. Los valores de las características se comparan antes y después del aislamiento de las venas pulmonares (promedio±desviación estándar). El símbolo * indica diferencias estadísticamente significativas (p-valor < 0.05).*

En la Figura 3 se han incluido también tres gráficas con los diagramas de caja para observar el comportamiento de cada una de las características analizadas en los pacientes de la cohorte bajo análisis.

Las gráficas y la tabla muestran que el primer autovalor (λ_1) es significativamente mayor tras el aislamiento de las venas pulmonares, reflejándose así un incremento de interés en la diagonal mayor de los loops de las ondas P. Asimismo, también se puede observar que la medida de la

morfología de los loops (ρ) decrece significativamente tras el procedimiento de ablación, indicando que los loops de las ondas P son más alargados y menos redondos al finalizar dicho procedimiento.

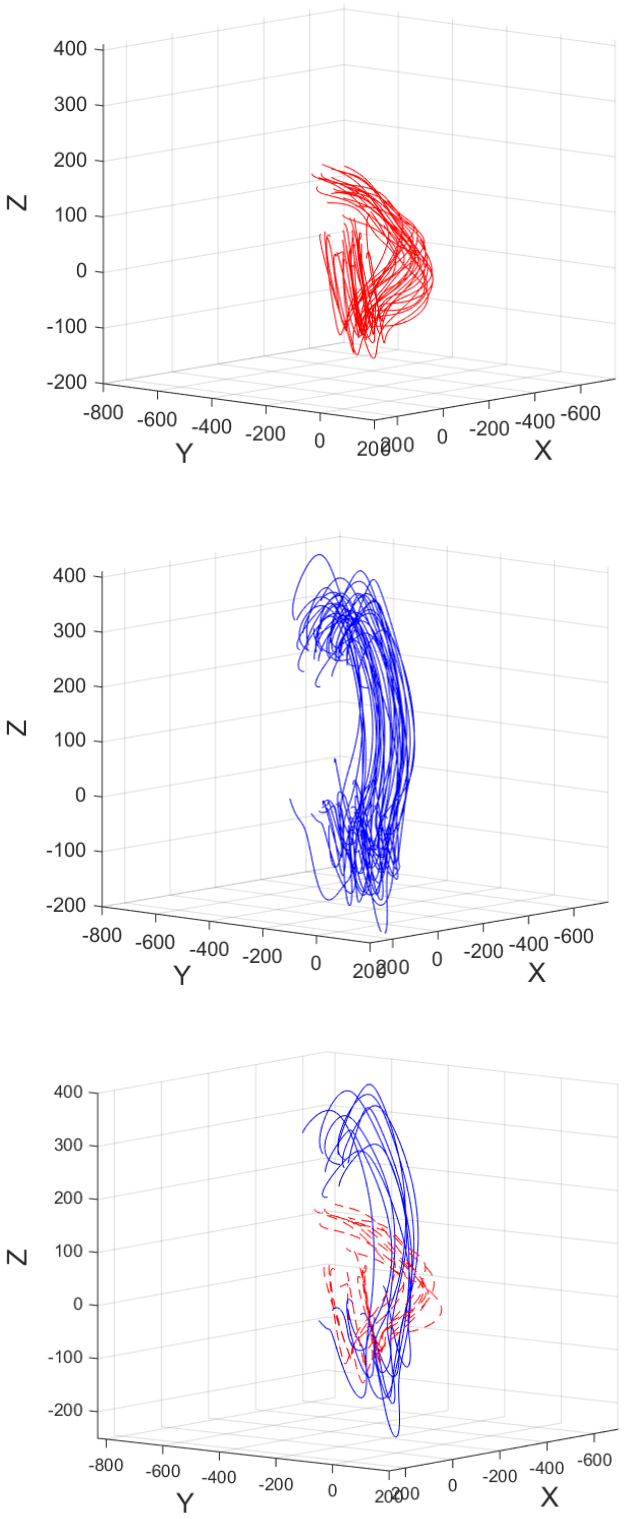


Figura 2. *Loops correspondientes a las ondas P de un paciente del estudio. Primera y segunda gráficas: antes/después del aislamiento de las venas pulmonares, respectivamente. Tercera gráfica: detalle de los primeros 10 loops antes (línea discontinua en color rojo) y después (línea continua en color azul).*

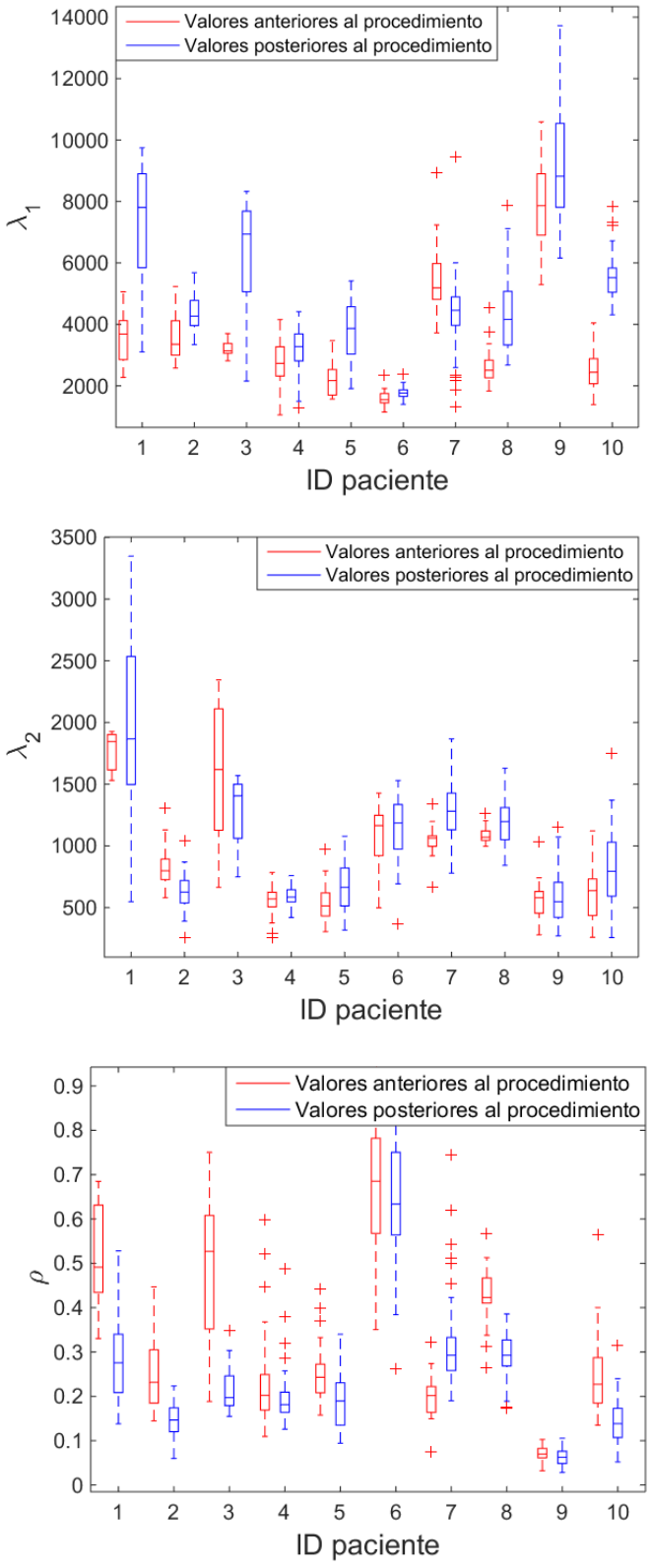


Figura 3.*Diagramas de cajas para las características analizadas: primer autovalor (primera gráfica), segundo autovalor (segunda gráfica), medida de redondez (tercera gráfica).*

5. Discusión y conclusiones

El objetivo del presente estudio ha sido analizar las diferencias morfológicas en los loops correspondientes a las ondas P tras el procedimiento de aislamiento de venas pulmonares mediante ablación por catéter en pacientes con fibrilación auricular paroxística.

La mayor parte de las referencias que analizan cómo se ve reflejado el aislamiento de las venas pulmonares en el electrocardiograma de superficie y en sus recurrencias se han centrado en la duración de las ondas P, así como en el decremento de la variabilidad de la morfología latido a latido.

En este estudio, sin embargo, se han analizado los cambios que se producen en los loops de las ondas P del vectocardiograma tras la ablación por catéter. Se han mostrado el análisis del primer y segundo autovalores de los loops, así como una medida de la redondez de los mismos, que han demostrado que el procedimiento introduce cambios morfológicos en la señal eléctrica. Los loops de las ondas P son significativamente menos redondos, con forma más elipsoidal al finalizar dicho procedimiento.

Estos resultados indican que el aislamiento de las venas pulmonares modifica el patrón de propagación eléctrico en las aurículas, y que dichos cambios pueden visualizarse en el vectocardiograma.

El trabajo futuro irá enfocado a aumentar la cohorte de pacientes bajo estudio, incluir otras características del vectocardiograma en el análisis, e investigar las posibles relaciones entre las características analizadas y la aparición de arritmias tras el procedimiento de aislar las venas pulmonares.

Agradecimientos

N. Ortigosa agradece el apoyo del Ministerio de Ciencia, Innovación y Universidades en el marco del Programa Estatal de Promoción del Talento y su Empleabilidad en I+D+i -Subprograma Estatal de Movilidad- con la beca CAS19/00168, de la Generalitat Valenciana con el proyecto Prometeo/2017/102 y del MINECO con el proyecto MTM2016-76647-P.

Referencias

[1] Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *European Heart Journal*, vol 37, sup 38, 2016, pp 2893-2962 (ISSN: 0195-668X).

[2] January C, Wann L, Alpert J, Calkins H, et al. 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation. A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *Journal of the American College of Cardiology,* vol 64, sup 21, 2014, pp. e1–e76 (ISSN: 0735-1097).

[3] Gill J. How to perform pulmonary vein isolation. *Europace*, vol 6, sup 2, 2004, pp 83–91 (ISSN: 1099-5129).

[4] Kis Z, Muka T, Franco O, Bramer W, de Vries et al. The Short and Long-Term Efficacy of Pulmonary Vein Isolation as a Sole Treatment Strategy for Paroxysmal Atrial Fibrillation: A Systematic Review and Meta-Analysis. *Current Cardiology Reviews*, vol 13, sup 3, 2017, pp 199-208 (ISSN: 1573-403X).

[5] Jared T, Cutler M. Is pulmonary vein isolation still the cornerstone in atrial fibrillation ablation? *Journal of Thoracic Disease,* vol 7, sup 2, 2015, pp 132–141 (ISSN: 2072-1439).

[6] Ogawa M, Kumagai K, Vakulenko M, Yasuda T, et al. Reduction of P-wave duration and successful pulmonary vein isolation in patients with atrial fibrillation. *Journal of Cardiovascular Electrophysiology*, vol 18, sup 9, 2007, pp 931–938 (ISSN: 1045-3873).

[7] Kaypakli O, Koca H, Sahin D, et al. Association of P wave duration index with atrial fibrillation recurrence after cryoballoon catheter ablation. *Journal of Electrocardiology*, vol 51, sup 2, 2018, pp 182-187 (ISSN: 0022-0736).

[8] Dilaveris P, Gialafos J. P-Wave Dispersion: A Novel Predictor of Paroxysmal Atrial Fibrillation. *Annals of Noninvasive Electrocardiology*, vol 6, sup 2, 2001, pp 159-165 (ISSN: 1542-474X).

[9] Censi F, Corazza I, Reggiani E, Calcagnini G, Mattei E, Triventi M, Boriani G. P-wave Variability and Atrial Fibrillation. *Scientific Reports*, vol 6, 2016, pp 1-7 (ISSN: 2045-2322).

[10] Ortigosa N, Cano O. P-wave analysis in paroxysmal atrial fibrillation patients before and after pulmonary vein isolation. *Computing in Cardiology*, vol 44, 2017, pp 1-4. (ISSN: 2325-887X).

[11] Meyer CR, Keiser HN. Electrocardiogram baseline noise estimation and removal using cubic splines and state space computation techniques. *Computers and Biomedical Research*, vol 10, 1977, pp 459-470 (ISSN: 0010-4809).

[12] Kors J, van Herpen G, Sittig A, van Bemmel J. Reconstruction of the Frank vectorcardiogram from standard electrocardiographic leads: diagnostic comparison of different methods. *European Heart Journal*, vol 11, sup 12, 1990, pp 1083-1092 (ISSN: 0195-668X).

Efecto de la Ablación por Catéter Sobre Ciertas Estructuras Auriculares de Pacientes con Fibrilación Auricular Paroxística: El Caso del Seno Coronario

A. Vraka[1], V. Bertomeu-González[2], F. Hornero Sos[3],
A. Quesada Dorador[4], R. Alcaraz Martínez[5], J. J. Rieta Ibáñez[1]

[1] BioMIT.org, Departamento de Ingeniería Electrónica, Universitat Politecnica de Valencia, España, {aivra, jjrieta}@upv.es
[2] Departamento de Medicina Clínica, Universidad Miguel Hernández, España, vbertomeu@umh.es
[3] Servicio de Cirugía Cardiovascular, Hospital Clínico Universitario de Valencia, España, hornero_fer@gva.es
[4] Servicio de Cardiología, Hospital General Universitario de Valencia, España; quesada_aur@gva.es
[5] Grupo de Inv. en Electrónica, Telecom. y Bioingeniería, Univ. de Castilla-La Mancha, España, raul.alcaraz@uclm.es

Resumen

Numerosos estudios previos han reportado alteraciones significativas de varias características de la onda P en el electrocardiograma (ECG) de superficie tras un procedimiento de ablación por catéter (AC) de fibrilación auricular paroxística (FAP). Sin embargo, no se ha reportado resultado alguno sobre una de las estructuras auriculares más empleada como referencia en el registro de señales invasivas durante procedimientos de AC, como es el seno coronario (SC). Para indagar en este asunto, este estudio utiliza registros bipolares del SC de 26 pacientes sometidos a AC de FAP. Los registros fueron grabados en ritmo sinusal (RS) y duraron 5 minutos. El trabajo incluye la detección y el análisis de las siguiente características de ondas de activación local (OALs): duración, amplitud, área positiva y negativa, variabilidad morfológica (VM), dispersión y el número de tramos crecientes y decrecientes (NTCD). Las características mencionadas se compararon antes y después de la AC usando el test U de Mann-Whitney, calculando también sus valores medianos. Duración, amplitud y áreas positivas y negativas marcaron un moderado decrecimiento (de $-1{,}42\,\%$ a $-14{,}28\,\%$). La VM y la dispersion se incrementaron (de $+1{,}20\,\%$ a $+64{,}81\,\%$). No se encontró ninguna alteración en el NTCD. En consecuencia, los resultados obtenidos coincidieron en su mayor parte con hallazgos similares reportados sobre la onda P del ECG en trabajos previos. No obstante, las características de las OALs en registros bipolares del SC no sufrieron una alteración significativa por causa de la AC ($p > 0{,}05$). En consecuencia, las mayores alteraciones observadas en estudios sobre la onda P del ECG deben provenir, posiblemente, de otras áreas auriculares alejadas del SC o son una superposición de alteraciones de menor nivel que tienen lugar en múltiples zonas auriculares.

1. Introducción

Siendo la fibrilación auricular (FA) la arritmia cardiaca más extendida a nivel mundial y presentando además una tasa elevada de riesgo de mortalidad [1], la evaluación de las modificaciones del sustrato de auricular ha sido objeto de investigación en numerosos estudios de los últimos años [2]. Aunque la ablación por catéter (AC) de las venas pulmonares (VPs) sigue siendo el tratamiento estrella de la FA [3, 4], sus efectos en el comportamiento eléctrico auricular están en constante revisión [5], intentando lograr un conocimiento mayor y más integrado sobre los mecanismos de la FA [6].

En este contexto, se han publicado numerosos estudios que, partiendo del análisis de electrocardiogramas (ECGs) de superficie, han intentado detectar variaciones en las características de la onda P tras AC de pacientes con FA paroxística (FAP). Los hallazgos publicados ofrecen aspectos comunes entre sí, como la reducción de la duración de la onda P [7–9], de su amplitud y del área [7, 8], así como un aumento de la variabilidad morfológica (VM) [10].

Las antedichas modificaciones han sido atribuidas a la alteración de sustrato auricular que produce la AC, así como a la participación de una menor cantidad efectiva de tejido en la despolarización auricular, teniendo como consecuencia un valor de amplitud de la onda P menor [7, 8]. Por su parte, una duración más corta de la onda P puede explicarse por la eliminación de bloqueos de conducción causados por el sustrato proarrítmico que origina activaciones prolongadas [7–9]. Por lo que respecta a la variabilidad morfológica (VM), esta característica es un indicador de la organización de la FA y de la normalidad de tejido, por lo que su incremento implica el éxito del aislamiento eléctrico de las áreas que muestran alta actividad fibrilatoria y tejido alterado por el remodelado auricular [8].

A pesar de que las anteriores observaciones pueden formar una idea muy concreta sobre cómo la AC impacta en la dinámica de la activación auricular durante el ritmo sinusal (RS) de pacientes con FAP, el análisis de la onda P desde el ECG de superficie superopone el estudio de la actividad eléctrica de ambas aurículas, considerándolas como una sola entidad. Sin embargo, las antedichas modificaciones de la onda P, ocasionadas por la AC, pueden originarse

ISBN: 978-84-09-25491-0, pags. 105- 108

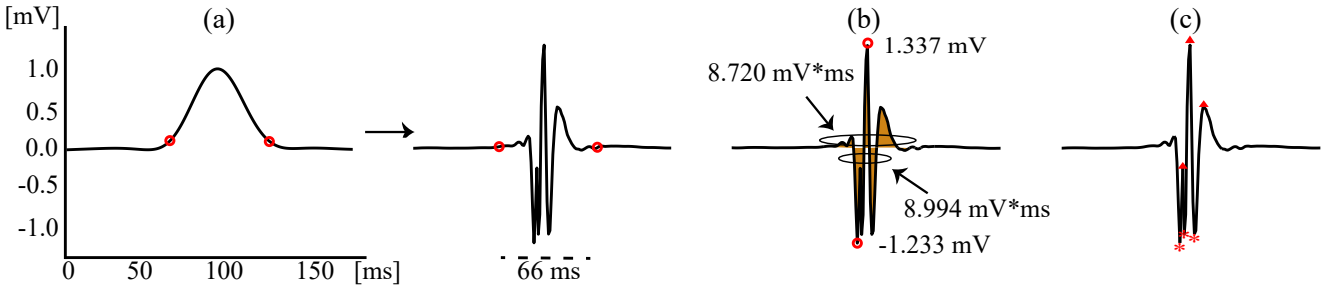


Figura 1. *(a) Curva obtenida del preprocesado de Botteron [11] indicando los puntos fiduciales que delimitan el comienzo y el final de la OAL (izquierda) y con los mismos puntos superpuestos sobre el EGM original (derecha). (b) Definición de amplitud máxima y mínima, área positiva y negativa. (c) Máximos locales (indicados como triángulos en la parte superior) y mínimos locales (indicados como asteriscos en la parte inferior) que determinan segmentos crecientes y decrecientes en la zona principal de la OAL.*

fundamentalmente en áreas auriculares concretas, que son, posiblemente, favorecidas o afectadas afectadas con mayor intensidad por el proceso ablativo y cuya localización contribuiría de manera extensa a una comprensión más amplia de los mecanismos de mantenimiento y terminación de la FA. En este caso, sería necesario el análisis de registros invasivos, es decir de electrogramas (EGMs), para que aquellas alteraciones que podrían producirse a nivel local en ciertas zonas auriculares, pudieran ser rastreadas e investigadas. Con el afán de aportar información adicional sobre este aspecto clave, que conlleve una mejor comprensión de los efectos que producen las técnicas ablativas sobre el sustrato auricular de pacientes con FAP, el presente estudio utiliza registros invasivos del seno coronario (SC), que por su ubicación topológica entre la aurícula izquierda y derecha es una amplia referencia utilizada en la evaluación del sustrato auricular durante AC [12]. Así, será posible poder observar durante el RS aquellas alteraciones de las características electrofisiológicas del SC que puedan ser ocasionadas por el proceso ablativo en pacientes con FAP.

2. Materiales

La base de datos ha se compuso con registros bipolares multicanal del SC correspondientes a 26 pacientes com FAP sometidos por primera vez a aislamiento circunferencial de VPs mediante AC con energía de radiofrecuencia (RF). Los registros se obtuvieron mediante un sistema LabSystem® Pro (Boston Scientific, Marlborough, MA, USA) con una frecuencia de muestreo de 1 kHz y con una duración de cinco minutos antes y después del procedimiento de ablación.

3. Métodos

Para cada paciente, se seleccionó por inspección visual el canal menos ruidoso de los cinco registrados en el SC, teniendo en cuenta que se debía analizar el mismo canal antes y después de la AC. En primer lugar se realizó la cancelación de posibles interferencias ventriculares mediante un método de plantillas adaptativas [13]. Después se redujo el ruido de red y de alta frecuencia con una estrategia de filtrado wavelet [14]. Finalmente, se detectó cada una de las ondas de activación local (OALs) en cada registro mediante un método que evalúa el fraccionamiento presente en los electrograms [11].

3.1. Características fundamentales

Para cada OAL, se generó un perfil conteniendo sus características principales, que fueron definidas buscando paralelismo con las características más relevantes publicadas sobre la onda P. Así, se definieron las siguientes:

Duración: La distancia temporal entre el inicio y el fin de cada OAL, calculada con la ayuda de una curva normalizada entre 0 y 1, obtenida por una técnica alternativa a Botteron [11]. La OAL fue suavizada primeramente con un filtro de media móvil de 5 puntos y luego convertida a una curva normalizada. El primer punto de muestreo de la curva con amplitud mayor de 0.1 mV marcó el tiempo de inicio, mientras el último punto de muestreo cumpliendo el mismo requisito definió el final. La figura 1.a ilustra el proceso de cálculo de la duración de la OAL.

Amplitud: La amplitud máxima fue definida por el pico positivo más alto, mientras la amplitud mínima por el pico negativo más bajo. La amplitud pico a pico (PP) quedó definida como la distancia vertical entre las dos amplitudes anteriores, como se puede observar en la Figura 1.b. Finalmente, el valor cuadrático medio (RMS) de cada OAL también fue calculado.

Área: Inicialmente, se marcaron los puntos de cruce por cero para definir las partes de señal positiva y negativa. El área positiva total se obtuvo como la integración de todas partes de señal positiva dentro de la OAL. De la misma manera, el área negativa total se obtuvo como la suma de áreas de señal negativa. La Figura 1.b muestra la definición de área positiva y negativa de la OAL.

Número de tramos crecientes y decredientes: La señal original fue normalizada entre ± 1, definiendo además dos lineas base auxiliares al $25\,\%$ de la amplitud positiva y negativa. El número de tramos crecientes y decrecientes (NTCD) se obtuvo mediante el cómputo del número de segments de cada OAL que cruzaban las líneas auxiliares, creciendo o decreciendo. Además se tuvo la precaución de evitar cómputos duplicados para aquellos segmentos que pudiesen cruzar ambas líneas auxiliares. De esta manera, solo se contempló el NTCD de la zona central o principal de cada OAL, despreciando aquellas posibles oscilaciones de baja amplitud que, en muchas ocasiones, suelen estar presentes en los extremos de la activación local, tal como se puede observar en la Figura 1.c.

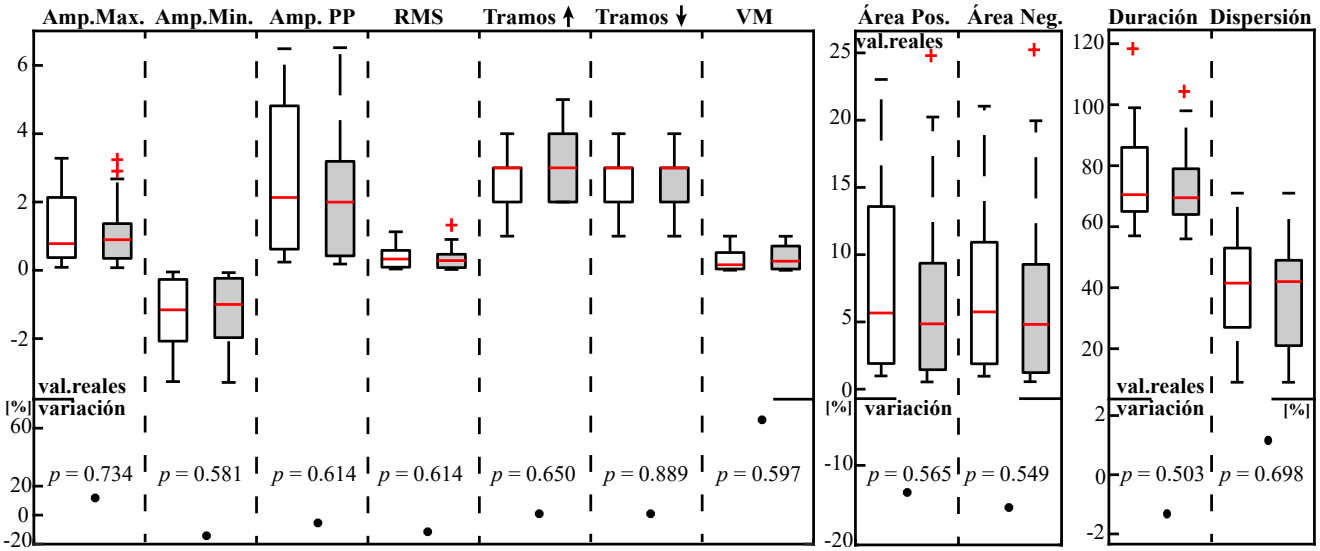


Figura 2. *En la parte superior se muestran los diagramas de cajas y bigotes para todas las características pre-AC (izquierda) y post-AC (sombreadas a la derecha). En la parte inferior se indica la variación porcentual correspondiente que experimenta cada característica debido a la AC (puntos negros) y se presenta también la significancia estadística p.*

3.2. Características extendidas

Además de las características fundamentales de cada OAL definidas anteriormente, se realizó la definición de otras adicionales que han sido utilizadas en trabajos previos para evaluar dinámicas fibrilatorias así como la condición del sustrato auricular. Fueron las siguientes:

Dispersión: Se definió como la diferencia entre duración máxima y mínima de las OALs, analizando todos los canales válidos de cada registro. Dispersiones elevadas en la onda P están relacionadas con heterogeneidad del tejido auricular y peor pronóstico de efectividad en AC [15].

Variabilidad morfológica: Para cada registro se calculó una onda de referencia obtenida como promedio las 20 OAL más similares. Después, cada OAL se comparó con la de referencia utilizando un índice adaptativo de correlación con signo (IACS) con un $12\,\%$ de tolerancia [16]. La variabilidad morfológica (VM) de cada registro se obtuvo como el porcentaje de OALs con parecido superior al $95\,\%$ respecto de la OAL de referencia.

3.3. Evaluación del método

La normalidad de los datos se exploró con el test de Kolmogorov-Smirnov y, de acuerdo a los resultados, la comparación de las características antes y después de la AC se realizó con el test de Kruskal-Wallis. Para poder seguir la dirección de variación de cada característica, también se calcularon los correspondientes valores medianos antes y después de la AC.

4. Resultados

En la Figura 2 se muestran los diagramas de cajas y bigotes de todas las características analizadas, tanto para antes como para después de la AC, junto con el porcentaje de variación que cada característica marcó después de la AC en función de su valor original y los valores de significancia estadística (p) que corresponden a cada variación.

No se observó ninguna variación en NTCD en valores medianos después de la AC ($+0{,}0\,\%$). En cambio, la duración de las OALs, su amplitud mínima, valores PP, valores RMS y las áreas disminuyeron ($-1{,}42\,\%$ a $-16{,}17\,\%$). Por otro lado, también se observó un ligero ($+1{,}2\,\%$) a moderado ($+64{,}81\,\%$) aumento en la amplitud máxima, la dispersión y la VM. Ninguna de estas variaciones fue estadísticamente significativa ($p > 0{,}05$). Aún así, se observó que la alteración de la mayoría de características estaba en concordancia con trabajos previos llevados a cabo analizando onda P desde el ECG de superficie [7–9, 15].

5. Discusión

El SC constituye un puente de conexión interauricular ampliamente aceptado como referencia del comportamiento eléctrico de las aurículas [12] y su análisis electrofisiológico se ha revelado como clave en la evaluación del sustrato auricular en FA [17]. Este trabajo presenta por primera vez un estudio encaminado a descifrar la alteración de ciertas propiedades electrofisilógicas del SC causadas por el aislamiento circunferencial de las VPs en pacientes con FAP. En el estudio se ha intentado mimetizar características similares analizadas previamente sobre la onda P del ECG de superficie. Los resultados demuestran que la AC de la FAP no altera de manera significativa la función electrofisiológica del SC, lo que implica que las alteraciones reportadas por trabajos previos sobre la onda P deben tener su origen en zonas diferentes las aurículas.

Aunque los resultados no presentan variaciones estadísticamente significativas causadas por la AC de los parámetros analizados en el SC, debe destacarse que la mayoría de las alteraciones observadas concuerdan con resultados de estudios previos sobre onda P. La duración de las OALs, su amplitud negativa y PP y sus áreas disminuyeron tras la AC. Estudios previos han presentado reducciones similares en características paralelas de la onda P [7–9, 18, 19]. Además, se ha reportado que valores bajos de VM están

asociados a defectos de conducción auricular [10], por lo que el aumento de esta característica tras la AC también coincide con estudios previos que analizaron el ECG [10]. La amplitud máxima marcó la misma tendencia, que está asociada a un descenso en la recurrencia de FA en pacientes con FAP sometidos a Crioablación [15].

Por otro lado el incremento de la dispersión, asociado a la presencia de tejido auricular con comportamiento heterogéneo, coincidió con estudios recientes [19], pero discrepó con anteriores [15]. En este sentido debe decirse, aparte de que la alteración fue muy sutil ($0{,}5$ ms; $+1{,}20\,\%$), que la evaluación de OALs en registros invasivos es un proceso bastante complicado. La línea base en estos registros contiene muy a menudo grandes oscilaciones, incluso durante RS. Además, aunque en el preprocesado se aplicó cancelación de ruido y se descartaron canales demasiado ruidosos, algunos de los canales restantes continuaban presentando muy bajos valores de amplitud, lo que complica el proceso de delineación precisa de OALs y, por tanto, de cálculo de la dispersión.

Aunque las observaciones antedichas no fueron estadísticamente significativas, sí que se observó un comportamiento paralelo entre las variaciones en onda P reportadas en estudios previos y las de OALs del SC obtenidas en este trabajo. Por consiguiente, como las ondas P ofrecen una perspectiva más global de la función auricular, es probable que las alteraciones mencionadas en trabajos anteriores se deban a la superposición de microalteraciones a nivel local en múltiples zonas auriculares, siendo el SC una de ellas.

6. Conclusiones

El efecto de la AC en registros del SC durante de pacientes con FAP es limitado, aunque paralelo al efecto reportado en onda P por estudios anteriores. Es posible que el impacto de la AC sea mayor en otras zonas auriculares o que las variaciones de onda P se deban a la contribución simultánea de microalteraciones procedentes de múltiples áreas auriculares.

Agradecimientos

El presente trabajo ha sido cofinanciado por los proyectos de investigación competitiva DPI2017-83952-C3 de MINECO-AEI-FEDER-UE, SBPLY/17/180501/000411 de la JCCLM y AICO/2019/036 de la GVA.

Referencias

[1] January CT, Wann LS, et al. 2019 AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation. Circulation 07 2019;140(2):e125–e151.

[2] Kottkamp H, Bender R, Berg J. Catheter ablation of atrial fibrillation: how to modify the substrate? J Am Coll Cardiol Jan 2015;65(2):196–206.

[3] Ramirez FD, Reddy VY, Viswanathan R, Hocini M, Jaïs P. Emerging technologies for pulmonary vein isolation. Circ Res 06 2020;127(1):170–183.

[4] Haissaguerre M, Jaïs P, Shah DC, Takahashi A, Hocini M, Quiniou G, Garrigue S, Le Mouroux A, Le Métayer P, Clémenty J. Spontaneous initiation of atrial fibrillation by ectopic beats originating in the pulmonary veins. New England Journal of Medicine 1998;339(10):659–666.

[5] Zink MD, Chua W, Zeemering S, et al. Predictors of recurrence of atrial fibrillation within the first 3 months after ablation. Europace Jul 2020;.

[6] Lau DH, Linz D, Schotten U, Mahajan R, Sanders P, Kalman JM. Pathophysiology of paroxysmal and persistent atrial fibrillation: Rotors, foci and fibrosis. Heart Lung Circ Sep 2017;26(9):887–893.

[7] Van Beeumen K, Houben R, Tavernier R, Ketels S, Duytschaever M. Changes in p-wave area and p-wave duration after circumferential pulmonary vein isolation. Europace 2010;12(6):798–804.

[8] Maan A, Mansour M, Ruskin JN, Heist EK. Impact of catheter ablation on p-wave parameters on 12-lead electrocardiogram in patients with atrial fibrillation. Journal of electrocardiology 2014;47(5):725–733.

[9] Yanagisawa S, Inden Y, Okamoto H, Fujii A, Sakamoto Y, Mamiya K, Tomomatsu T, Shibata R, Murohara T. Electrocardiogram characteristics of p wave associated with successful pulmonary vein isolation in patients with paroxysmal atrial fibrillation: Significance of changes in p-wave duration and notched p wave. Annals of Noninvasive Electrocardiology 2019;e12712.

[10] Huo Y, Holmqvist F, Carlson J, Gaspar T, Hindricks G, Piorkowski C, Bollmann A, Platonov PG. Variability of p-wave morphology predicts the outcome of circumferential pulmonary vein isolation in patients with recurrent atrial fibrillation. Journal of electrocardiology 2015;48(2):218–225.

[11] Osorio D, Alcaraz R, Rieta JJ. A fractionation-based local activation wave detector for atrial electrograms of atrial fibrillation. In 2017 Computing in Cardiology (CinC). IEEE, 2017; 1–4.

[12] Antz M, Otomo K, Arruda M, Scherlag BJ, Pitha J, Tondo C, Lazzara R, Jackman WM. Electrical conduction between the right atrium and the left atrium via the musculature of the coronary sinus. Circulation 1998;98(17):1790–1795.

[13] Alcaraz R, Rieta JJ. Adaptive singular value cancelation of ventricular activity in single-lead atrial fibrillation electrocardiograms. Physiol Meas Dec 2008;29(12):1351–69.

[14] Martínez-Iniesta M, Ródenas J, Rieta JJ, Alcaraz R. The stationary wavelet transform as an efficient reductor of powerline interference for atrial bipolar electrograms in cardiac electrophysiology. Physiol Meas 07 2019; 40(7):075003.

[15] Kizilirmak F, Demir GG, Gokdeniz T, Gunes HM, Cakal B, Guler E, Karaca İO, Omaygenç MO, Yılmaz F, Olgun FE, et al. Changes in electrocardiographic p wave parameters after cryoballoon ablation and their association with atrial fibrillation recurrence. Annals of Noninvasive Electrocardiology 2016;21(6):580–587.

[16] Alcaraz R, Hornero F, Martínez A, Rieta JJ. Short-time regularity assessment of fibrillatory waves from the surface ecg in atrial fibrillation. Physiological measurement 2012; 33(6):969.

[17] Ahmed N, Perveen S, Mehmood A, Rani GF, Molon G. Coronary sinus ablation is a key player substrate in recurrence of persistent atrial fibrillation. Cardiology 2019;143(3-4):107–113.

[18] Liu Z, Hayano M, Hirata T, Quin Y, et al. Abnormalities of electrocardiographic p wave morphology and the relationship to electrophysiological parameters of the atrium in patients with idiopathic paroxysmal atrial fibrillation. Journal of cardiology 1998;32(3):189–196.

[19] Nakatani Y, Sakamoto T, Yamaguchi Y, Tsujino Y, Kataoka N, Kinugawa K. Coefficient of variation of p-wave duration measured using an automated measurement system predicts recurrence of atrial fibrillation. Journal of electrocardiology 2019;53:79–84.

Distinción Entre Electrogramas de Fibrilación Auricular Paroxística Frente a Persistente para Evaluación del Sustrato Auricular en Procedimientos de Ablación por Catéter

E. Finotti[1], V. Bertomeu González[2], F. Hornero Sos[3],
A. Quesada Dorador[4], R. Alcaraz Martínez[5], J. J. Rieta Ibáñez[1]

[1] BioMIT.org, Departamento de Ingeniería Electrónica, Universitat Politecnica de Valencia, España, {aivra, jjrieta}@upv.es
[2] Departamento de Medicina Clínica, Universidad Miguel Hernández, Alicante, España, vbertomeu@umh.es
[3] Servicio de Cirugía Cardiovascular, Hospital Clínico Universitario de Valencia, España, hornero_fer@gva.es
[4] Servicio de Cardiología, Hospital General Universitario de Valencia, España; quesada_aur@gva.es
[5] Grupo de Inv. en Electrónica, Telecom. y Bioingeniería, Univ. de Castilla-La Mancha, España, raul.alcaraz@uclm.es

Resumen

Para contribuir a un abordaje optimizado de la ablación por catéter (AC) de fibrilación auricular (FA), durante los últimos años se han introducido complicados índices destinados a discriminar FA paroxística de persistente (FApar vs. FAper) aplicados sobre electrogramas auriculares fraccionados complejos (CFAE). Sin embargo, los electrofisiólogos exigen el uso de métodos de clasificación simples y de comprensión directa. Por ello, el presente trabajo explota la longitud de ciclo de FA (AFCL), la frecuencia dominante (DF), la entropía muestral (SE) y el determinismo (DET) del análisis de cuantificación recurrente, aplicado a registros de FA con CFAEs, para crear modelos sencillos de discriminación entre FApar y FAper. El AFCL y la DF se calcularon sobre los registros enteros, mientras que SE y DET se calcularon utilizando segmentos de 1, 2 y 4 s. La información redundante se eliminó umbralizando sucesivamente matrices de correlación y el algoritmo de Random Forests ordenó las variables por relevancia. A continuación, un árbol de clasificación combinó de manera óptima los índices con alto nivel clasificatorio y se probaron con validación cruzada dejando uno fuera. Después de analizar todas las combinaciones posibles, el mejor resultado obtuvo una Precisión (Acc) del 88,2 % para discriminar FApar de FAper, mientras que DET proporcionó la mejor Acc individual de 82,4 %. Como conclusión, una selección cuidada y reducida de características puede facilitar un modelo de clasificación sencillo capaz de discriminar con precisión entre CFAEs de FApar y FAper para mejorar el abordaje terapéutico de la FA.

1. Introducción

La fibrilación auricular (FA) es la arritmia cardíaca más común, caracterizada por contracciones rápidas y caóticas de las aurículas originadas por una activación eléctrica descoordinada [1]. El tratamiento de la FA se realiza mediante cardioversión eléctrica [2] y farmacológica [3], pero desde que Haissaguerre *et al.* reportaron la crucial importancia de las venas pulmonares (PV) en el inicio y mantenimiento de la FA [4], el procedimiento de ablación por catéter (AC) dirigido al aislamiento de focos arritmogénicos en las PVs, se ha convertido en el tratamiento estrella de la FA [5].

Dado que se ha demostrado que la AC es superior a cualquier tratamiento farmacológico de FA en cuanto a su reducida tasa de mortalidad por cualquier causa [6], se está realizando un gran esfuerzo investigador para personalizar los tratamientos de AC gracias al mapeo del sustrato electrofisiológico auricular. El objetivo consiste en identificar aquellos focos de arritmogenia responsables de las actividades eléctricas anormales [7].

En la práctica clínica, una de las finalidades de la caracterización del sustrato auricular tiene por objetivo discriminar pacientes en FA paroxística (FApar) frente a persistente (FAper), ya que se ha demostrado que la tasa de éxito de la AC es estrictamente dependiente del área ablacionada y del tipo de FA [8]. De hecho, en la FApar se alcanza una elevada tasa de éxito con la simple ablación de las PVs, mientras en la FAper se requieren ablaciones adicionales para lograr resultados similares [7]. En consecuencia, esto define un desafío para una caracterización precisa del sustrato auricular con el fin de guiar de forma óptima la AC. Por ello, las metodologías para distinguir entre electrogramas auriculares fraccionados complejos (CFAE) de FApar y FAper serían muy interesantes y útiles en el mapeo rápido y eficiente del sustrato auricular [9].

Hasta el momento, se han propuesto diversas estrategias de clasificación de CFAEs para discriminar FApar de FAper. Así, Ciaccio *et al.* midieron la repetitividad de los CFAEs [9] y cuantificaron el grado de heterogeneidad morfológica en las deflexiones de los mismos [10]; Acharya *et al.* utilizaron gráficos de recurrencia, análisis de cuantificación de recurrencia (RQA) y medidas de entropía [11]; Ndrepepa *et al.* y Ravi *et al.* usaron la longitud del ciclo de fibrilación auricular [12, 13] y Sanders *et al.* emplearon el mapeo de frecuencias dominantes [14].

ISBN: 978-84-09-25491-0, pags. 109- 112

Sin embargo, las estrategias mencionadas suelen definir parámetros de elevada complejidad que, suelen compatibilizar difícilmente con la práctica clínica, donde se requieren modelos de clasificación sencillos y fácilmente interpretables. El presente trabajo propone explotar métodos no lineales con modelos clasificatorios sencillos para discriminar entre CFAEs de FApar y FAper de pacientes sometidos a AC. Se ha analizado la frecuencia dominante (DF), la longitud del ciclo de FA (AFCL), la entropía muestral (SE) y el determinismo (DET) de RQA. Los índices extraídos se han procesado y seleccionado para alimentar posteriormente sencillos árboles de clasificación. Esto ha posibilitado poder discriminar con precisión entre CFAEs de FApar y FAper para evaluar el sustrato auricular, contribuyendo a mejorar las decisiones terapéuticas sobre manejo de FA.

2. Materiales

2.1. Base de datos y preprocesado

Se analizó un total de 204 electrogramas de 16 segundos muestreados a 977 Hz de pacientes con FApar y FAper sometidos a AC en el laboratorio de electrofisiología cardíaca. Para los pacientes con FApar en ritmo sinusal, la FA se indujo mediante estimulación rápida en el seno coronario o en la pared lateral de la aurícula derecha. Se analizaron 90 y 114 CFAEs de FApar y FAper, respectivamente, tomando registros de puntos estratégicos que fueron las cuatro venas pulmonares: superior izquierda (LSPV), inferior izquierda (LIPV), derecha superior (RSPV) e inferior derecho (RIPV); y la pared libre anterior (ANT) y posterior (POS) de la aurícula izquierda.

Los CFAEs se dividieron en segmentos de 1, 2 y 4 s, creando tres conjuntos de datos de 3264, 1632 y 816 elementos , respectivamente. Antes de la digitalización, todas las señales se filtraron paso-banda con el sistema de adquisición (0,5-450 Hz, CardioLab, GE Healthcare, Waukesha, WI), que eliminó la deriva de la línea de base y cierto ruido muscular de alta frecuencia. Después se los registros se remuestrearon a 1 KHz y se filtraron con un método basado en wavelets para la reducción adicional de ruido [15].

2.2. Entropía muestral

Se ha elegido la SE por su capacidad acreditada para delinear la dinámica de la FA, así como para evaluar el nivel del remodelado auricular [16]. La SE asignó a cada secuencia un valor no negativo que refleja su complejidad, con valores más grandes correspondientes a una mayor irregularidad en los datos [17]. La longitud máxima de la plantilla m se fijó en 2 muestras y la tolerancia r en 0,35 veces la desviación estándar del segmento [16].

2.3. Medidas de cuantificación de recurrencias

Se utilizaron gráficos de recurrencia (GR) para encontrar patrones dentro de las secuencias analizadas. Para reconstruir los GRs a partir de los CFAEs, sus trayectorias se proyectaron en un espacio fásico d-dimensional obtenido mediante el teorema de Takens, con una dimensión de reconstrucción d y un retardo en el tiempo τ, definidos con los métodos de vecinos más próximos e información mutua, respectivamente. A continuación, se produjeron los GR representando cada par de estados de trayectoria como un punto negro, si su distancia estaba por debajo de un umbral especificado, y blanco en caso contrario. El umbral se fijó en el 10 % de la media del diámetro del espacio fásico [18]. La medida RQA seleccionada para el estudio ha sido el determinismo (DET), que cuantifica la predictibilidad de la secuencia, midiendo el porcentaje de puntos de recurrencia que pertenecen a líneas diagonales del GR con longitud mínima de 50.

2.4. Longitud del ciclo de fibrilación auricular

El AFCL se midió detectando ondas de activación local en los CFAEs [19], invirtiendo las distancias entre las activaciones consecutivas individuadas y, en fin, promediando. El AFCL se calculó en los CFAEs de 16 s de longitud de todos los sitios de registro.

2.5. Frecuencia dominante

La DF se ha medido identificando el pico espectral más alto en el rango electrofisiológico de interés de 3 a 12 Hz. Se ha demostrado que la DF es generalmente mayor en FAper y menor en FApar [10]. La DF se calculó para la longitud completa de cada CFAEs y para cada punto de registro usando el periodogrma de Welch con una ventana Hamming de 4096 muestras, un solapamiento del 50 % entre ventanas consecutivas y una FFT de 8192 puntos.

2.6. Evaluación de características estadísticas

La prueba de Mann-Whitney verificó la hipótesis nula de que los índices eran similares para FApar y FAper, con un valor de significancia establecido en $p < 0{,}05$. La prueba se realizó para los índices en cada zona de registro. Para los tres conjuntos de datos, se promediaron los resultados de p en cada canal para proporcionar un valor final.

2.7. Selección de características y clasificación

Este proceso se utilizó para eliminar variables redundantes, minimizando la pérdida de información [20]. La estrategia adoptada combinó dos pasos. Primeramente se umbralizó a la matriz de correlación, eliminando las variables más redundantes con un valor de corte de 0,60. Después se realizó un ranking de relevancia de variables mediante un algoritmo de Random Forest, asignado puntuaciones (impureza de Gini) de 0 a 100, con valores crecientes según la relevancia de la variable. Empíricamente se seleccionaron las variables con puntuaciones superiores a 40.

La clasificación entre FApar y FAper se realizó utilizando árboles de decisión junto con validación cruzada dejando uno fuera. Cada modelo de árbol se construyó utilizando una de las posibles combinaciones de variables obtenida tras el proceso de selección antes mencionado. Para todas las longitudes de segmento se analizaron los posibles subconjuntos de características para encontrar el árbol de mayor precisión. Además, para simplificar aún más la clasificación, se creó un modelo de árbol utilizando una sola característica de entrada.

	SE		DET	
	FApar	FAper	FApar	FAper
1s	0.123±0.042	0.148±0.057	0.608±0.135	0.524±0.160
2s	0.120±0.043	0.145±0.056	0.656±0.140	0.556±0.177
4s	0.117±0.043	0.143±0.057	0.688±0.135	0.585±0.179

Tabla 1. *Valores medios y desviación estándar de entropía muestral (SE) y determinismo (DET) para los tres conjuntos de CFAEs analizados de 1, 2 y 4 s de longitud según el tipo de FA.*

canal	SE	DET	AFCL	DF
LSPV	0.0006	0.0001	>0.05	>0.05
LIPV	0.0253	0.0080	>0.05	0.0287
RSPV	<0.0001	<0.0001	0.0192	0.0007
RIPV	>0.05	>0.05	>0.05	0.0035
ANT	0.0035	0.0016	>0.05	0.0387
POS	0.0002	0.0006	>0.05	>0.05

Tabla 2. *Valores de significancia estadística p de la prueba de Mann-Whitney para cada punto de registro. Se han promediando los valores p para las tres longitudes de segmento analizadas.*

3. Resultados y discusión

Las medias y los valores de desviación estándar que resultaron de la aplicación de SE y DET a los tres conjuntos de datos se presentan en la Tabla 1, divididos por tipo de FA. Los valores medios de SE disminuyeron con la longitud del segmento, obteniendo mayores valores en FAper en comparación con FApar, reflejando así un mayor grado de desorganización en FAper. Al revés, los valores de DET aumentaron al crecer la longitud del segmento, mostrando mayores valores en FApar que en FAper, por lo que la primera es más predecible que la segunda.

El análisis sobre la longitud completa de los CFAEs, proporcionó valores de AFCL con un promedio y una desviación estándar de 7.01±1.46 para FApar y 7.75±1.56 para FAper, respectivamente. Por su parte, la DF obtuvo un promedio y desviación estándar de 5.59±1.36 para FApar y de 6.20±1.08 para FAper, respectivamente.

Como se muestra en la Tabla 2 para SE y DET, la prueba de Mann-Whitney rechazó la hipótesis nula en la mayoría de los canales, con valores de p particularmente bajos en la vena pulmonar superior derecha (RSPV), mientras en la vena pulmonar inferior derecha (RIPV) se aceptó la hipótesis nula, revelando un remodelado auricular similar en ambos tipos de FA. Para el AFCL, los valores del p resultaron siempre más altos que la significación estadística fijada en 0.05 a excepción de RSPV. Por ultimo, las pruebas para DF fueron estadísticamente diferenciadoras entre FApar y FAper, salvo en la vena pulmonar superior izquierda (LSPV) y en la pared posterior libre de la aurícula izquierda (POS).

La matriz de correlación obtenida promediando los valores correspondientes para todos los segmentos analizados de 1, 2 y 4 s se muestra en la Figura 1. En particular, se observó una fuerte correlación negativa (puntos grandes y granates) entre los pares de índices SE-DET y AFCL-DET, tomados en los mismos puntos de registro. También, se observó una fuerte correlación positiva (puntos grandes y azules) entre el par SE-AFCL medido en el mismo canal.

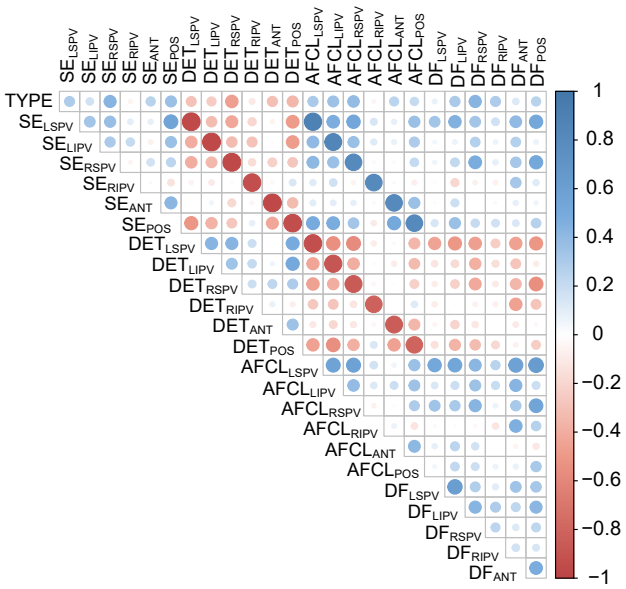


Figura 1. *Matriz de correlación indicativa del grado de relación entre índices aplicados a cada lugar de registro estudiado. La variable TYPE puede tomar valor de 0 si el paciente es paroxístico, y de 1 si el paciente es FA persistente.*

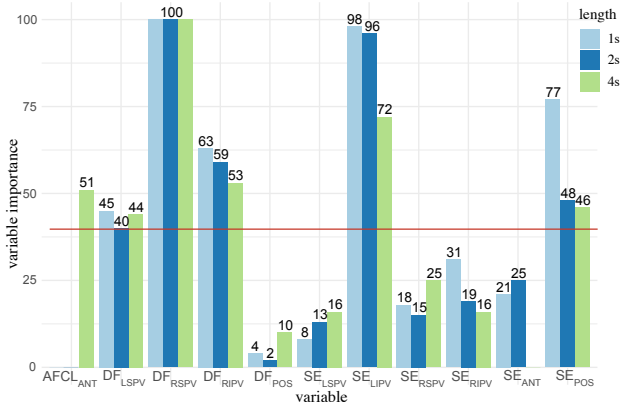


Figura 2. *Importancia de las variables clasificadas con el algoritmo Random Forest para los segmentos de CFAEs de 1, 2 y 4 s. El valor de corte de 40 se muestra en rojo.*

La umbralización de la matriz de correlación en los CFAEs de 1 y 2 s eliminó el índice DF_{LIPV} y por completo los índices de DET y AFCL, debido a su fuerte correlación con SE. Para los CFAEs de 4 s, el subconjunto de índices mantenidos fueron SE_{LSPV}, SE_{LIPV}, SE_{RSPV}, SE_{RIPV}, SE_{POS}, $AFLC_{ANT}$, DF_{LSPV}, DF_{RSPV}, DF_{RIPV}, DF_{ANT} y DF_{POS}.

Para los tres conjuntos de datos, la puntuación de las variables proporcionada por Random Forest se presenta en la Figura 2. Para los conjuntos de datos de 1 y 2 s de longitud, las variables clasificadas con puntuación superior a 40 fueron DF_{LSPV}, DF_{RSPV}, DF_{RIPV}, SE_{LIPV} y SE_{POS}. En el conjunto de datos de 4 s, además de las anteriores, también se incluyó $AFLC_{ANT}$ ya que superó el valor umbral.

Después de probar todas las combinaciones posibles de las características con alto ranking, el grupo SE_{POS}, DF_{LSPV}, DF_{RSPV} y DF_{RIPV} proporcionó el mejor resultado de clasificación obteniendo una precisión del 88,2 % para todas las longitudes de los segmentos. Los valores de precisión de los demás modelos construidos con las restantes combi-

naciones posibles de las características mejor clasificadas (puntuación > 40) ofrecieron una media y una desviación estándar de 70,7$\pm$8,7 % para segmentos de CFAE de 1 s, 70,8$\pm$8,3 % para 2 s y 69,2$\pm$8,8 % para 4 s. Lo cual es una reducción significativa en la precisión clasificatoria.

La precisión más alta lograda por un solo índice fue proporcionada por DET_{LIPV} con un 82.4 % para cualquier longitud de segmento. En cambio, el promedio de las precisiones alcanzadas, con el modelo a una variable, por los otros índices fue 60.2$\pm$11 % para 1 s de longitud, 54.7$\pm$13. % para 2 s y 57,3$\pm$11,4 % para 4 s de longitud.

Para terminar, cabe decir que el estudio del poder discriminatorio de DF_{RIPV} mediante prueba estadística, llevó a la conclusión errónea de que el índice no tiene capacidad discriminatoria entre FApar y FAper, ya que se aceptó la hipótesis nula. Por el contrario, la selección de características clasificó al DF_{RIPV}, junto con SE_{POS}, DF_{LSPV}, DF_{RSPV}, como el conjunto de variables más importantes para clasificar los tipos de FA con los pacientes analizados, alcanzando la mayor precisión de 88,2 %.

4. Conclusiones

Una cuidada selección de combinaciones limitadas de características utilizada para la construcción de modelos de clasificación sencillos puede ser capaz de discriminar con precisión aceptable entre CFAEs de FApar y FAper, pudiendo así proporcionar ayuda para mejores decisiones terapéuticas sobre el manejo de la FA e interpretaciones más sencillas sobre la evaluación del sustrato auricular en FA.

Agradecimientos

El presente trabajo ha sido cofinanciado por los proyectos de investigación competitiva DPI2017-83952-C3 de MINECO-AEI-FEDER-UE, SBPLY/17/180501/000411 de la JCCLM y AICO/2019/036 de la GVA.

Referencias

[1] January CT, Wann LS, et al. 2019 AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation. Circulation 2019; 140(2):e125–e151.

[2] Kwon CH. External electrical cardioversion is an easy and safe intervention for rhythm control in persistent atrial fibrillation. Korean Circ J Jun 2020;50(6):524–526.

[3] Capucci A, Cipolletta L, Guerra F, Giannini I. Emerging pharmacotherapies for the treatment of atrial fibrillation. Expert Opin Emerg Drugs 03 2018;23(1):25–36.

[4] Haïssaguerre M, Jaïs P, Shah DC, Takahashi A, Hocini M, Quiniou G, Garrigue S, Le Mouroux A, Le Métayer P, Clémenty J. Spontaneous Initiation of Atrial Fibrillation by Ectopic Beats Originating in the Pulmonary Veins. The New England Journal of Medicine 1998;339(10):659–666.

[5] Hesselson AB. Catheter ablation in the treatment of atrial fibrillation. Int J Angiol Jun 2020;29(2):108–112.

[6] Turagam MK, Garg J, Whang W, Sartori S, Koruth JS, Miller MA, Langan N, Sofi A, Gomes A, Choudry S, Dukkipati SR, Reddy VY. Catheter ablation of atrial fibrillation in patients with heart failure: A meta-analysis of randomized controlled trials. Ann Intern Med 01 2019;170(1):41–50.

[7] Kottkamp H, Bender R, Berg J. Catheter ablation of atrial fibrillation: how to modify the substrate? J Am Coll Cardiol Jan 2015;65(2):196–206.

[8] Latchamsetty R, Morady F. Catheter ablation of atrial fibrillation. Heart Fail Clin Apr 2016;12(2):223–33.

[9] Ciaccio EJ, Biviano AB, Whang W, Vest JA, Gambhir A, Einstein AJ, Garan H. Differences in repeating patterns of complex fractionated left atrial electrograms in longstanding persistent atrial fibrillation as compared with paroxysmal atrial fibrillation. Circ Arrhythm Electrophysiol Aug 2011;4(4):470–7.

[10] Ciaccio EJ, Biviano AB, Whang W, Gambhir A, Garan H. Different characteristics of complex fractionated atrial electrograms in acute paroxysmal versus long-standing persistent atrial fibrillation. Heart Rhythm Sep 2010;7(9):1207–15.

[11] Acharya UR, Faust O, Ciaccio EJ, Koh JEW, Oh SL, Tan RS, Garan H. Application of nonlinear methods to discriminate fractionated electrograms in paroxysmal versus persistent atrial fibrillation. Comput Methods Programs Biomed Jul 2019;175:163–178.

[12] Ndrepepa G, Karch MR, Schneider MAE, Weyerbrock S, Schreieck J, Deisenhofer I, Zrenner B, Schömig A, Schmitt C. Characterization of paroxysmal and persistent atrial fibrillation in the human left atrium during initiation and sustained episodes. J Cardiovasc Electrophysiol Jun 2002; 13(6):525–32.

[13] Ravi KC, Krummen DE, Tran AJ, Bullinga JR, Narayan SM. Electrocardiographic measurements of regional atrial fibrillation cycle length. Pacing Clin Electrophysiol Mar 2009;32 Suppl 1:S66–71.

[14] Sanders P, Berenfeld O, Hocini M, Jaïs P, Vaidyanathan R, Hsu LF, Garrigue S, Takahashi Y, Rotter M, Sacher F, Scavée C, Ploutz-Snyder R, Jalife J, Haïssaguerre M. Spectral analysis identifies sites of high-frequency activity maintaining atrial fibrillation in humans. Circulation Aug 2005; 112(6):789–97.

[15] Martínez-Iniesta M, Ródenas J, Rieta JJ, Alcaraz R. The stationary wavelet transform as an efficient reductor of powerline interference for atrial bipolar electrograms in cardiac electrophysiology. Physiol Meas 07 2019; 40(7):075003.

[16] Alcaraz R, Abásolo D, Hornero R, Rieta JJ. Optimal parameters study for sample entropy-based atrial fibrillation organization analysis. Comput Methods Programs Biomed Jul 2010;99(1):124–32.

[17] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol Jun 2000;278(6):H2039–49.

[18] Schinkel S, Dimigen O, Marwan N. Selection of recurrence threshold for signal detection. The European Physical Journal Special Topics 10 2008;164:15–53.

[19] Osorio D, Alcaraz R, Rieta JJ. A fractionation-based local activation wave detector for atrial electrograms of atrial fibrillation. In 2017 Computing in Cardiology (CinC). IEEE, 2017; 1–4.

[20] Guyon I, Elisseeff A. An introduction of variable and feature selection. J Machine Learning Research Special Issue on Variable and Feature Selection 2003;3.

Estimación del "jitter" fisiológico en trenes de potenciales de unidad motora utilizando parámetros de "jiggle"

Tunca Nüzket[1], A. Malanda Trigueros[1], J. Navallas Irujo[1]

[1] Departamento de Ingeniería Eléctrica, Electrónica y de Comunicación, Universidad Pública de Navarra, Pamplona, España, {malanda,tunca.nuzket,javier.navallas}@unavarra.es

Resumen

El "jitter" neurofisiológico es la variabilidad temporal de las descargas eléctricas sucesivas producidas en las fibras de una unidad motora (UM). El "jitter" aparece aumentado en patologías neurógenas como la ELA o en enfermedades de la unión neuromuscular, como la miastenia gravis. El "jitter" hace variar la forma y tamaño de los potenciales de unidad motora (PUMs), efecto conocido como "jiggle". Por ello tiene sentido tratar de estimar el "jitter", difícil de medir, usando parámetros de "jiggle". En el presente trabajo se comparan dos de estos parámetros, ampliamente utilizados: las diferencias de amplitud consecutivas (DAC) y los coeficientes de correlación cruzada de descargas consecutivas (CCC).

En el trabajo se simularon PUMs utilizando un reconocido modelo de potenciales de fibra simple. Se realizaron cuatro estudios diferentes en los que se variaron, respectivamente: el nivel de ruido, el número de fibras de las UMs, el número de descargas sucesivas de PUMs, y la distancia entre los dos grupos de fibras que componen la UM modelada. En todos ellos se contemplaron distintos niveles de "jitter".

Se definió el error de superposición como una ratio para medir la precisión del DAC y el CCC como estimadores del "jitter". El CCC obtuvo errores de superposición de alrededor de 0.5 en todos los casos estudiados. El DAC tuvo un mejor rendimiento, con errores de superposición siempre menores que los del CCC, y más bajos para valores de "jitter" bajos. En conclusión, el DAC ha resultado un estimador del "jitter" más preciso que el CCC.

1. Introducción

La unidad motora (UM) es la unidad básica de contracción muscular y consta de una motoneurona, el axón de esta motoneurona, los extremos terminales que salen de este axón y las fibras musculares inervadas por estos extremos terminales. El axón mielinizado permite que el potencial eléctrico se propague rápidamente a los extremos terminales [1]. Con la activación de la UM, el potencial eléctrico viaja a lo largo de la motoneurona y se distribuye a lo largo de sus extremos terminales. La acetilcolina (Ach) liberada en la unión neuromuscular se une a sus receptores y desencadena un potencial de acción (AP) en las fibras musculares de la UM. La suma de estos potenciales registrada en un electrodo cercano se denomina potencial de unidad motora (PUM) [1]. Mientras la UM está activada, este proceso se repite varias veces generándose un tren de PUMs.

En las enfermedades de la unión neuromuscular, la liberación cuántica de Ach en el lado presináptico o la sensibilidad de la placa terminal a Ach en el lado postsináptico puede estar disminuida [2]. Con una afectación crítica de la placa terminal, la debilidad muscular se hace patente. Para la identificación de tales condiciones patológicas se emplean diferentes métodos de análisis diseñados para examinar las señales electrofisiológicas involucradas.

El disparo de las diferentes fibras de una MU presenta pequeñas diferencias temporales ("jitter") [3]. En individuos sanos, el valor de "jitter" varía entre 21 y 45 ↘s [4]. Algunas condiciones patológicas causan una función defectuosa de los extremos presinápticos o postsinápticos de la placa motora terminal, interrumpiendo la sincronización en el disparo de los PUMs, lo que aumenta el valor del "jitter". El "jitter" aumentado aparece en el caso de denervación-reinervación en enfermedades neurogénicas como la ELA o en enfermedades de transmisión neuromuscular, como la miastenia gravis.

La medida del "jitter" se utiliza como técnica exploratoria en enfermedades en las que se ve afectada la transmisión neuromuscular. Las mediciones del "jitter" se han obtenido tradicionalmente utilizando señales electromiográficas tomadas de electrodos de fibra simple (Figura 1.A). Este tipo de electrodo presenta una pequeña ventana abierta en la cánula, de la cual emerge su parte activa. El área de captación del electrodo es pequeña (del orden de 0.07 mm^2) por lo que es raro que más de dos fibras de la misma UM puedan estar contenidas en esta área. En la adquisición, el electromiografista mueve el electrodo insertado en el músculo, tratando de acercarse a una UM activa en la que se puedan observar dos potenciales de fibra simple (PFS). Si los dos potenciales están suficientemente separados en el tiempo (Figura 1.B) será posible medir su retardo (intervalo entre potenciales o IPI). El "jitter" es, en esencia, la variabilidad del IPI y se puede calcular como la diferencia media entre IPI consecutivos. También se puede especificar como la media de las diferencias en los tiempos de disparo de potenciales consecutivos (descargas) en un tren PUMs.

Los electrodos de fibra simple son caros y normalmente reutilizables. Para evitar el riesgo de infecciones, en muchos países se ha suprimido su uso, reemplazándolos

ISBN: 978-84-09-25491-0, pags. 113- 116

por electrodos de aguja concéntrica, que son más baratos y desechables.
Sin embargo, el uso de electrodos de aguja concéntricos también ha traído algunas desventajas. Con un área de captación grande, el electrodo puede captar potenciales de más de una fibra muscular en el registro del PUM. Esto hace que sea difícil distinguir los potenciales de cada fibra muscular, que aparecen en forma de picos agregados [6].

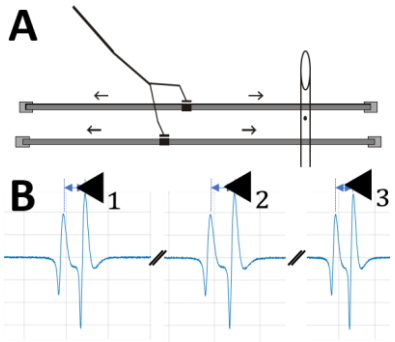


***Figura 1**. A) Registro de PUMs procedente de dos fibras tomado con electrodo de fibra simple. B) Variacion del IPI en PFSs.*

Los sucesivos potenciales de un tren de PUMs pueden superponerse de diferente manera, lo que dificulta la medición del "jitter". Hasta la fecha se han desarrollado algunas técnicas para implementar el cálculo de "jitter" a partir de registros EMG de aguja concéntrica con varios grados de automatización [7]-[8].
El "jitter" muestra directamente el estado de la conducción neuromuscular. Si el potencial de la placa terminal no logra alcanzar un valor umbral, la fibra muscular no puede crear un potencial de acción, fenómeno conocido como bloqueo del impulso. El "jitter" y el bloqueo causan fluctuaciones en la forma y tamaño de los sucesivos potenciales del tren de MUPs, a lo cual se le da el nombre de "jiggle" [9].
Aparte de estas condiciones fisiológicas, hay otros factores que afectan al "jiggle", como el ruido o el movimiento del electrodo de aguja durante el registro [10]. En cualquier caso, dada la fuerte relación entre "jiggle" y "jitter", puede ser posible estimar el valor de "jitter", que es relativamente difícil de calcular, usando parámetros de "jiggle". Esta es la hipótesis fundamental de este trabajo, en el que se calculan y comparan los valores de las diferencias de amplitud consecutivas (DAC), los coeficientes de correlación cruzada de descargas consecutivas (CCC) y "jitter" en trenes de PUMs simulados. Dado que no es posible conocer a ciencia cierta el "jitter" existente en un tren de PUMs, la simulación resulta ser la vía más adecuada para realizar estudios comparativos sobre métodos existentes de estimación del "jitter" o para probar nuevos métodos.

2. Materiales y Métodos

2.1. Estudios de simulación

Se utilizó un reconocido modelo de PFS [11] para simular trenes de PUMs. La UM simulada se componía de dos grupos de fibras (fascículos), ligeramente separadas en el eje longitudinal. Se establecieron varios parámetros fisiológicos y geométricos como valores predeterminados para las simulaciones: diámetro de la UM: 1 mm; número total de fibras de la UM: 20; longitud de las fibras musculares 120 mm; velocidad de conducción media: 3,5 ± 0,1 m/s; número de descargas: 50, nivel de "jitter": de 0 a 0,1 ms a intervalos de 0,01 ms; desviación estándar (SD) del nivel de ruido: 0,1 mV; distancia longitudinal de las placas motoras de las fibras: 40 ± 2,5 mm; distancia longitudinal de los dos grupos de fibras (DGF): 5 mm.
Se llevaron a cabo cuatro estudios de simulación diferentes con un nivel de "jitter" que variaba entre 0 y 0,1 ms, con un incremento de 0,01 ms. En el primer estudio, se contemplaron tres valores de DGF: 0, 5 y 10 mm, respectivamente. En el segundo estudio, se varió el nivel de ruido (SD) dando tres valores distintos: 0,025, 0,05 y 0,1 mV, respectivamente. En el tercer estudio se incluyó la variación del número total de fibras en los siguientes valores: 10, 20, 40 y 60. En el cuarto estudio, el número de potenciales del tren de PUMs varió de 10 a 80 en incrementos de 10. En los cuatro estudios el resto de los parámetros tomaron sus valores por defecto.
En cada estudio se simularon 100 trenes de PUMs para cada uno de los 11 valores de "jitter" y cada valor del parámetro variable bajo prueba. En total, se utilizaron 19.800 trenes de PUMs.

2.2. Alineamiento de trenes de PUMs

Una vez obtenidas las señales de la simulación, los trenes de PUMs deben ser alineados antes de calcular los parámetros de "jiggle". El retardo temporal de cada potencial se calculó como una estimación de los retardos con cada una de las descargas restantes. Los retardos temporales entre cada par de potenciales en un tren PUMs se calcularon buscando el máximo de la correlación cruzada aplicada a secciones de 5 ms de los PUMs centrados en sus picos más altos.

2.3. Parámetros de "jiggle"

Después de alinear los trenes de PUMs, se calcularon los parámetros DAC y CCC siguiendo la formulación de Stalberg y Soono [9]. El DAC representa el promedio de las diferencias de amplitud en potenciales consecutivos del tren de PUMs, normalizado por el área media del PUM en estos potenciales y restando un factor (C) que da cuenta de la variabilidad de la señal en intervalos donde no hay potencial. En este trabajo, a diferencia del método original, el factor C se calculó de acuerdo con la modificación de Campos [10]. La CCC se calcula utilizando la correlación cruzada entre PUMs consecutivos. Concretamente se calcula como el valor mediano de las correlaciones cruzadas entre todas las parejas de PUMs consecutivos del tren. El cálculo de los valores DAC y CCC se realiza dentro de una ventana de 5 ms de longitud centrada en máximo del potencial.

2.4 Figuras de mérito

El "jitter" es una variable de dimensión temporal, mientras que DAC y CCC, debido a su normalización, son adimensionales. Por tanto, se necesita una figura de mérito específica para evaluar ambos parámetros de "jiggle" como estimadores del "jitter". Para un valor de "jitter" dado (*Jit_i*) y un conjunto definido de parámetros de simulación, habrá 100 medidas de DAC distribuidas alrededor de un valor medio de DAC (distribución *DAC_i*) (Figura 2). Para el siguiente (y mayor) valor de "jitter" del estudio (*Jit_j*), también habrá una colección de 100 valores de *DAC* (*DAC_j*) distribuidos alrededor de otro valor de *DAC* central, supuestamente más grande que el anterior. Dependiendo de la proximidad de los dos valores de fluctuación, *DAC_i* y *DAC_j* pueden superponerse.

Esta superposición significa que algunas muestras de *DAC_i* son más grandes que algunas muestras de *DAC_j*, lo que debe interpretarse como errores en la estimación de "jitter", ya que en estos casos valores de "jitter" crecientes dan lugar a valores de "jiggle" decrecientes. Para medir la frecuencia con la que ocurren estos errores, se considera un umbral (*Th*) que separa las distribuciones *DAC_i* y *DAC_j*. Las muestras de *DAC_i* mayores que Th se consideran errores del tipo 1, mientras que las muestras de *DAC_j* inferiores a *Th* se consideran errores del tipo 2.

La frecuencia de cada uno de estos errores dependerá del valor de *Th*. Se busca un valor de *Th* que haga que el número de errores tipo 1 y tipo 2 sean iguales. Entonces, la ratio del número de errores frente al número total de muestras será la figura de mérito final (error de superposición) para cada pareja de valores consecutivos de "jitter" inspeccionados. Los errores de superposición en el CCC se calcularán de manera idéntica.

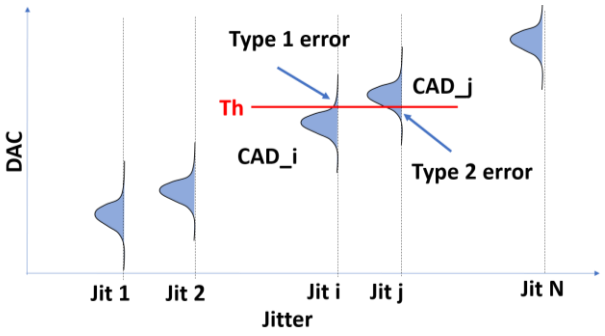


Figura 2. *Errores de solapamiento en las distribuciones de valores de CAD para valores consecutivos de "jitter".*

La segunda figura de mérito refleja si hay diferencias estadísticas significativas de los conjuntos de valores de DAC o CCC correspondientes a parejas contiguas de los valores de "jitter" evaluados. Para este análisis se utilizó la prueba U de Mann-Whitney, aceptándose como significativos los valores de p menores de 0,001.

3. Resultados

3.1 Errores de solapamiento

Los errores de superposición para DAC y CCC de los cuatro estudios se muestran en la Figura 3. Para cada uno de estos estudios sólo se presentan dos figuras, correspondientes a los valores extremos de los parámetros involucrados.

En todos los casos los errores de superposición del CCC se mantienen cerca de 0,5, con muy poca variación para todo el rango de valores de "jitter" inspeccionados. Por el contrario, el error de superposición en el DAC aumenta con el "jitter", tendiendo asintóticamente a un nivel de 0,4. Este mejor rendimiento del CAD sobre el CCC es coincidente con los resultados del trabajo de Campos et al. [10], aunque no con los de Stålberg y Soono [9].

El ruido hace aumentar el error de superposición en el DAC (Figuras 3.a-b). Esta sensibilidad es tanto mayor cuanto menor es el nivel de "jitter".

El número de fibras parece tener una ligera influencia en la medición del DAC para valores pequeños de jitter (<0.04 ms), haciendo que el error de solapamiento sea menor para UMs con mayor número de fibras (Figuras 3.c-d).

Por otra parte, los errores de superposición no cambian significativamente al aumento el DGF (Figuras 3.e-f).

Finalmente, aunque era esperable una disminución del error de solapamiento al aumentar el número de descargas del tren de PUMs, no se observó prácticamente variación alguna del rendimiento del DAC con esta variable (Figuras 3.g-h).

3.2 Diferencias estadísticas

Se encontraron diferencias estadísticamente significativas de DAC correspondientes a valores contiguos de "jitter" para todos los casos evaluados y valores de "jitter", excepto en el caso de nivel de ruido de 0.1 y valores de "jitter" de 0-0.01 ms y 0.02-0.03 ms (Figura 3).

Por el contrario, con el parámetro CCC no se hallaron diferencias significativas entre las distribuciones de este parámetro para niveles consecutivos de "jitter" para ninguno de los casos estudiados (Figura 3).

Por otro lado, se estudió la correlación lineal entre el "jitter" y los parámetros de "jiggle". Agrupando los datos de los cuatro estudios anteriores, el DAC obtuvo un coeficiente de correlación con el "jitter" superior a 0.999. Por su parte, el del CCC fue de -0.975. Al analizar la correlación de manera separada para los cuatro estudios, se pudo apreciar que la pendiente de la recta de correlación apenas variaba con el número de potenciales del tren de PAUMs, aunque sí lo hacía con el número de fibras, el DGF o el ruido.

La SD del CAD aumenta con los valores de "jitter" para los cuatro estudios realizados (el coeficiente de variación oscila entre 0.0112 e 5.5912). Sin embargo, para el CCC la SD no parece aumentar consistentemente con el "jitter" (fluctuando en el rango entre los 0.0111 y los 0.0401 ms).

4. Conclusiones

Los resultados de los distintos estudios muestran una correlación lineal muy alta del CAD y el CCC con el "jitter", aunque las pendientes de las rectas de correlación varían con el número de fibras, el DGF y el ruido. En todo caso, el rendimiento del parámetro CCC es muy pobre para estimar el "jitter". El error de superposición para CCC en todos los niveles de "jitter" es de alrededor de 0,5.

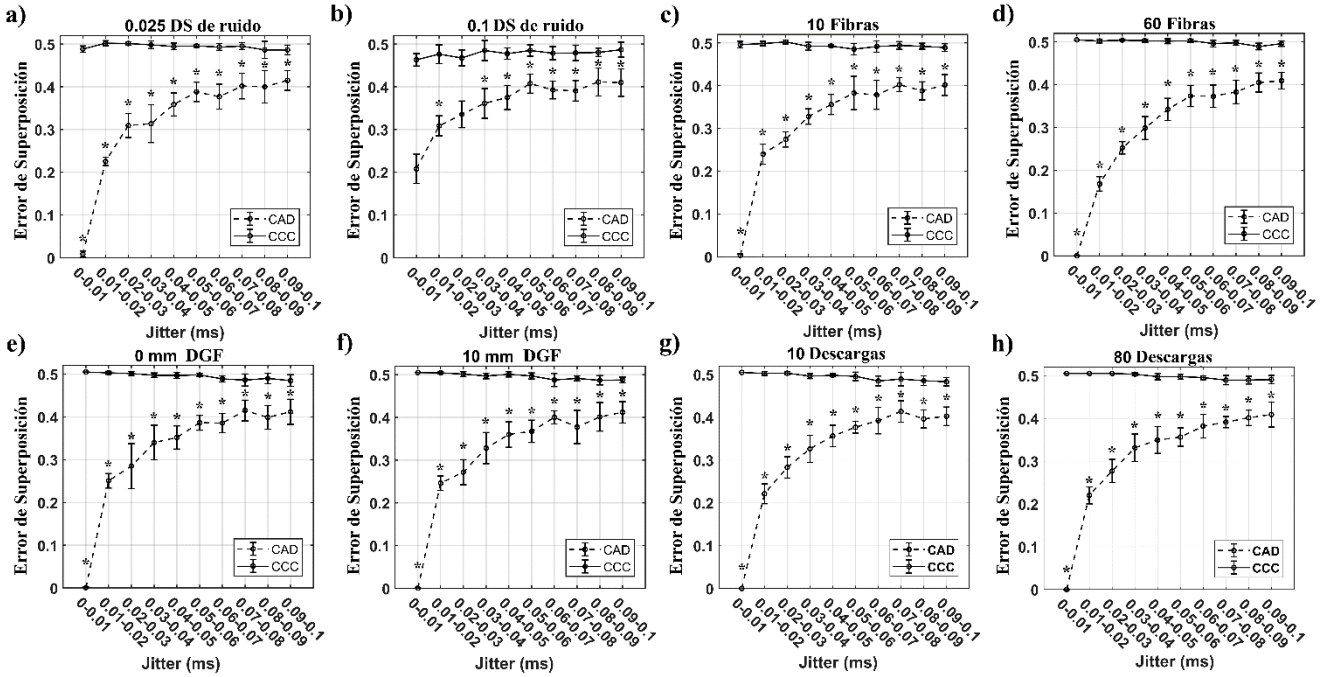


Figura 3. *Errores de solapamiento (media y STD) en DAC (línea continua) y CCC (línea discontinua) para distintos valores de "jitter" en los distintos estudios. a) Nivel de ruido (STD) 0.025 mV, b) Nivel de ruido (STD) 0.1 mV, c)10 fibras, d) 60 fibras, e) DGF de 0 mm, f) DGF de 10 mm, g) 10 descargas en el tren de PUMs, h) 80 descargas en el tren de PUMs. (Los asteriscos indican diferencias significativas para los valores de DAC y CCC correspondientes a niveles de "jitter" contiguos (p <0,001)).*

El CAD, por el contrario, tiene un rendimiento mejor, especialmente para valores de "jitter" más bajos. A medida que aumenta el nivel de "jitter", el error de superposición también aumenta, aunque arroja valores más bajos que el CCC en todos los casos estudiados. Por tanto, el DAC debe considerarse un parámetro más adecuado para estimar el "jitter", aunque su precisión no es muy alta, sobre todo en niveles altos de "jitter". Por todo ello puede resultar de utilidad buscar nuevos parámetros que mejoren la estimación del "jitter", así como realizar estudios que validen las conclusiones de este trabajo sobre señales patológicas reales.

Agradecimientos

Este proyecto ha recibido financiación del programa de investigación e innovación H2020 de la Unión Europea en virtud del acuerdo de subvención Marie Sklodowska-Curie n. 801586.

Referencias

[1] A. C. Guyton and J. E. Hall, *Medical Physiology*. 2005. (ISBN-10 : 1455770051)

[2] V. C. Juel, "Single fiber electromyography," in *Handbook of clinical neurology*, vol. 160, Elsevier, 2019, pp. 303–310. (ISBN: 0072-9752)

[3] E. Stalberg and J. Ekstedt, "Single Fibre EMG and Microphysiology of the Motor Unit in Normal and Diseased Human Muscle," in *New Concepts of the Motor Unit, Neuromuscular Disorders, Electromyographic Kinesiology*, 2015. (ISBN: 978-3-8055-1451-4)

[4] E. Stålberg *et al.*, "Reference values for "jitter" recorded by concentric needle electrodes in healthy controls: A multicenter study," *Muscle and Nerve*, 2016, doi: 10.1002/mus.24750. (ISSN: 10974598)

[5] J. V Trontelj, M. Mihelin, J. M. Fernandez, and E. Stålberg, "Axonal stimulation for end-plate jitter studies.," *J. Neurol. Neurosurg. Psychiatry*, vol. 49, no. 6, pp. 677–685, 1986. (ISSN: 0022-3050)

[6] E. Stålberg, ""jitter" analysis with concentric needle electrodes," *Ann. N. Y. Acad. Sci.*, vol. 1274, no. 1, pp. 77–85, 2012. (ISSN: 0077-8923)

[7] D. W. Stashuk. "Detecting single fiber contributions to motor unit action potentials". *Muscle Nerve*, vol. 22, pp. 218–29. 1999. (ISSN: 0148639X)

[8] M. Abdelmaseeh and D. W. Stashuk. "Motor unit potential "jitter": a new measure of neuromuscular transmission instability". *IEEE Trans Neural Syst Rehabil Eng*, vol. 25, no. 7, pp.1018-1025, Jul. 2017. (ISSN: 1534-4320)

[9] E. V Stålberg and M. Sonoo, "Assessment of variability in the shape of the motor unit action potential, the "'jiggle",' at consecutive discharges," *Muscle Nerve Off. J. Am. Assoc. Electrodiagn. Med.*, vol. 17, no. 10, pp. 1135–1144, 1994. (ISSN: 0148-639X)

[10] C. Campos, A. Malanda, L. Gila, V. Segura, I. Lasanta, and J. Artieda, "Quantification of "jiggle" in real electromyographic signals," *Muscle Nerve Off. J. Am. Assoc. Electrodiagn. Med.*, vol. 23, no. 7, pp. 1022–1034, 2000. (ISSN: 0148-639X)

[11] N. A. Dimitrova and G. V. Dimitrov, "Electromyography (EMG) Modeling," in *Wiley Encyclopedia of Biomedical Engineering*, 2006. (ISBN: 9780471249672)

Predicción del fracaso temprano de la cardioversión eléctrica en fibrilación auricular sobre las 12 derivaciones del ECG estándar

E.M. Cirugeda Roldán[1], S. Calero Nuñez[2], A. Quesada Dorador [3], V.M. Hidalgo Olivares[3], J.J. Rieta Ibáñez[4], R. Alcaraz Martínez[1]

[1] Grupo de Investigación en Ingeniería Electrónica, Biomédica y de Telecomunicación, Universidad de Castilla-La Mancha, Cuenca, España, {eva.cirugeda, raul.alcaraz}@uclm.es
[2] Unidad de Arritmias, Complejo Hospitalario Universitario de Albacete, Albacete, España
[3] Departamento de Cardiología, Hospital General Universitario, Valencia, España
[4] BioMIT.org, Dep. de Ingeniería Electrónica, Universitat Politecnica de Valencia, Valencia, España

Resumen

Dentro de las técnicas existentes para el control del ritmo cardíaco en fibrilación auricular (FA), una de las más extendidas es la cardioversión eléctrica (CVE). Esta técnica, de aplicación sencilla y bajo coste, resulta muy efectiva en la restauración del ritmo sinusal (RS) en pacientes con FA persistente. Aunque es capaz de restaurar el RS en la gran mayoría, un alto número de pacientes, alrededor del 40 %, todavía revierten a FA durante las 4 primeras semanas. Así pues, resulta de gran interés clínico la predicción del éxito de la CVE para poder tratar de la mejor forma a cada paciente. Con dicho objetivo, se han propuesto distintos índices para caracterizar las ondas fibrilatorias (f), generalmente considerando solo la derivación V1 del ECG, dada su proximidad a la aurícula derecha, y así descartando otra información relevante que pudiera obtenerse desde otras derivaciones. Este trabajo extiende dicha caracterización de las ondas f a las 12 derivaciones del ECG estándar para evaluar la capacidad de predicción del fracaso temprano de la CVE. Se han evaluado tres de los índices más conocidos, es decir, la amplitud de las ondas f, la frecuencia dominante y la regularidad, y su rendimiento se ha comparado con el obtenido sobre V1. Los resultados obtenidos indican que, mediante la caracterización de las ondas f sobre la derivación II, la predicción del resultado de la CVE aumenta hasta un 12 % con respecto a V1, sugiriendo que esta derivación es la mejor elección para evaluar la condición proarrítmica de un paciente tras la CVE.

1. Introducción

La fibrilación auricular (FA) es una de las arritmias cardíacas con mayor afectación mundial y, generalmente, está asociada a otras patologías, entre ellas, apnea, hipertensión, obesidad o cardiomiopatía [1]. Durante la FA, la excitación eléctrica de las aurículas y, en consecuencia, su actividad mecánica son extremadamente rápidas y descoordinadas, lo que a su vez genera que la contractilidad auricular sea ineficiente. Aun cuando la FA, por sí misma, no es un factor crítico, puede generar alteraciones hemodinámicas que aumentan la predisposición de un paciente a sufrir tromboembolismos o demencia cerebro-vascular [2]. Cuando la FA se vuelve persistente, su duración es superior a siete días, y su finalización requiere de una intervención clínica. Los protocolos actuales de actuación recomiendan el uso de medidas de control sobre la frecuencia o sobre el ritmo cardíaco [1]. Estas últimas actuaciones permiten reducir la sintomatología, a la vez que aumentan la calidad de vida del paciente, lo que unido a otros factores, tales como la edad, el cuadro clínico o la duración de la FA, pueden hacer que el cardiólogo se decante por ellas [1]. Entre las disponibles, la cardioversión eléctrica (CVE) es una de las técnicas más extendidas para restaurar el ritmo sinusal (RS), dado su bajo coste y efectividad, ya que es capaz de revertir inicialmente la FA en más del 90 % de los pacientes [3]. No obstante, su efectividad se ve drásticamente reducida a medio y largo plazo, ya que apenas un 70 % de ellos permanecen en RS tras cuatro semanas y poco más del 20 % lo hace tras el primer año [3]. Por otro lado, y aunque no suele tener grandes efectos adversos, puede ser causa de trombosis, bradicardia, hipotensión, e incluso arritmias ventriculares malignas [1]. Teniendo todo esto en cuenta, la predicción del fracaso temprano del procedimiento resulta de gran interés clínico para decidir si se trata del tratamiento más adecuado para el paciente [4].

Con esta idea en mente, se han propuesto distintos parámetros para caracterizar las ondas fibrilatorias (f) presentes en la actividad auricular (AA) extraída a partir del electrocardiograma (ECG) de superficie. La amplitud de las ondas f (A_f), su frecuencia dominante (*Fdom*), y su regularidad son tres de los índices más utilizados, pero casi siempre se han calculado desde la derivación V1, descartando así la información contenida por las restantes señales capturadas [5,6]. La selección de la derivación V1 viene motivada principalmente por su proximidad a la aurícula derecha [7], presentando por tanto ondas f de mayor amplitud que las otras derivaciones [7, 8]. No obstante, teniendo en cuenta la dependencia existente entre las derivaciones y las correlaciones espaciales del ECG, este trabajo presenta una evaluación de la condición proarrítmica del paciente sobre las distintas derivaciones del ECG, para intentar maximizar la predicción del resultado de la CVE.

ISBN: 978-84-09-25491-0, pags. 117- 120

Parámetro	FA no recidiva	FA recidiva
Número de Pacientes	27	31
Hombres	15	18
Patología cardíaca subyacente	9	10
Diámetro aurícula izquierda (mm)	47.32± 4.76	44.72 ± 7.32

Tabla 1. *Características clínicas de los pacientes analizados.*

2. Materiales y métodos

2.1. Base de datos

La base de datos analizada en este trabajo consta de 58 pacientes diagnosticados con FA persistente y bajo tratamiento antiarrítmico, a los que se les ha aplicado una CVE para restaurar el RS. El protocolo de actuación consistió en la aplicación de cuatro choques eléctricos sobre el tórax del paciente. Una de las palas se situó paraesternalmente en el segundo espacio intercostal derecho y la otra lateralmente a lo largo de la línea axilar izquierda. Se consideró una secuencia creciente de cuatro descargas eléctricas de 200, 300, 360 y 360 J.

Durante todo el protocolo de CVE se registró un ECG de 12 derivaciones a una frecuencia de muestreo de 1024 Hz y 16 bit de resolución. Todos los pacientes revirtieron a RS tras el primer choque eléctrico, pero solo 27 fueron capaces de mantenerlo tras un seguimiento de 4 semanas. Para el análisis se seleccionó un segmento de 90 segundos antes del primer choque eléctrico. Los datos clínicos de los sujetos se muestran en la Tabla 1.

2.2. Caracterización de la actividad auricular

El preprocesado del ECG para eliminar la línea de base, la interferencia de red, y el ruido de alta frecuencia se realizó de forma independiente en cada derivación. La detección de los complejos QRS se realizó utilizando un algoritmo previamente publicado [9], y se supervisó y corrigió (si fue necesario) manualmente en cada una de las derivaciones.

La AA se extrajo también de forma independiente en cada derivación mediante la sustracción de la actividad ventricular [10]. Para ello, los latidos se agruparon en base a la morfología del complejo QRS. Un latido se consideró perteneciente a una morfología si el coeficiente de correlación cruzada entre latido y morfología era superior a 0.8 [5]. Para cada una de las morfologías encontradas, se construyó un latido promedio. La cancelación se llevó a cabo de forma iterativa, comenzando por la morfología con menor número de complejos QRS hasta la cancelación de la clase dominante. La duración del complejo QRST se determinó como la duración mínima entre 470 ms (valor habitual) y el 90 % del intervalo RR medio. La señal resultante se filtró paso alto a 3 Hz para eliminar cualquier residuo ventricular todavía presente en la señal [11].

Para la caracterización de las ondas f se consideraron tres de los parámetros más habituales, es decir, la A_f, la ***Fdom***, y su organización mediante la entropía muestral (Em) [5]. La caracterización se realizó de forma independiente sobre cada derivación. Así pues, considerando $f(n)$ como la señal de AA, de longitud N muestras (es decir, n = 1,2 . . . , N), la A_f se determinó según [6]:

$$A_f = \sqrt{\frac{1}{N}\sum_{n=1}^{N}|f(n)|^2}. \tag{1}$$

La ***Fdom*** se obtuvo como la frecuencia en la cual la densidad espectral de potencia (P) en la banda 3–12 Hz [6] fue máxima, es decir:

$$Fdom = \arg\{\underset{f_k=3-12\ Hz}{\text{máx}}\{\bar{P}(f_k)\}\}, \tag{2}$$

donde $\bar{P}$ representa el periodograma de Welch promedio. Cabe indicar que dicho periodograma se calculó para tramos de 6 segundos con un protocolo de ventana deslizante de 2 segundos, considerando únicamente aquellos segmentos que presentaban un coeficiente de correlación cruzada superior a 0.7 para evitar ruido y artefactos todavía presentes en la señal preprocesada.

Finalmente, para caracterizar la organización de las ondas *f*, se calculó la Em [12] sobre la denominada onda auricular principal de la AA, $ff(n)$, la cual se estimó tras aplicar un filtrado paso banda de 5 Hz centrado sobre la ***Fdom*** a $f(n)$. Este índice es un estadístico diseñado para cuantificar la regularidad o predictibilidad de una serie numérica, de tal forma que a mayor regularidad se obtienen menores valores de Em [12]. Más concretamente, se define como el ratio de verosimilitud de que dos secuencias de longitud m muestras, que son similares para una distancia r, lo sigan siendo para una longitud de una muestra superior. Desde un punto de vista matemático, la Em se calcula a través de los siguientes pasos [12].

1. Se definen las secuencias de longitud m:

$$v_m(j) = \{ff(j{+}i) : 0{\leq}i{\leq}m-1\}. \tag{3}$$

2. Se estima la probabilidad de que dos secuencias de longitud m sean similares respecto a una distancia de Chebyshev r, $B^m(r)$, es decir,

$$B^m(r) = \frac{1}{N-m}\sum_{k=1}^{N-m} B_k^m(r), \text{ siendo} \tag{4}$$

$$B_k^m(r) = \frac{1}{N-m-1}\sum_{\substack{j=1\\ j\neq k}}^{N-m}(d_{jk}(m) < r), \text{y} \tag{5}$$

$$d_{jk}(m) = \text{máx}\{|v_m(j) - v_m(k)|\} \tag{6}$$

3. Se calculan otra vez las ecuaciones (3)–(6) para una longitud de $m{+}1$ muestras, y se obtiene la Em como:

$$Em(ff, m, r, N) = -\ln\frac{B^{m+1}(r)}{B^m(r)}. \tag{7}$$

Los valores de N, m, y r se han determinado de acuerdo a la directrices encontradas en [12], siendo por tanto N = 30 segundos, m = 2 y r = 0.2 la desviación estándar de $ff(n)$.

2.3. Análisis estadístico

La normalidad de los datos se estimó mediante el test de Kolmogorov-Smirnov, y la separabilidad estadística entre los pacientes que mantuvieron el RS y recayeron a FA durante el seguimiento se calculó mediante una prueba t-Student o mediante una prueba de rangos de Wilcoxon, en función de si la distribución de los datos fue o no normal.

La evaluación de la capacidad predictiva de cada parámetro se realizó mediante curvas ROC. Estas curvas indican, para distintos umbrales de clasificación, la sensibilidad (Se) y la especificidad (Sp) del clasificador, es decir, el porcentaje de pacientes que recayeron a FA y que mantuvieron el RS, respectivamente, que fueron clasificados de forma correcta. El umbral de clasificación óptimo se estimó de acuerdo al criterio de Youden. Finalmente, se calculó el área bajo la curva ROC (AROC), ya que se trata de un indicador global de la exactitud de la clasificación que tiene en cuenta todos los posibles umbrales.

3. Resultados

La Figura 1 muestra la distribución estadística de los tres índices propuestos mediante un diagrama de cajas y bigotes. Cabe indicar que se muestra únicamente la derivación que tuvo el mayor valor de AROC (ver Tabla 2), es decir, la derivación II. Se puede observar como la amplitud mediana de las ondas *f* es mayor en aquellos pacientes que tras 4 semanas consiguieron mantener un RS. Por contra, tanto la *Fdom* como la *Em* son superiores en aquellos pacientes que revirtieron a FA antes de las 4 semanas. Adicionalmente, también se observa una mayor dispersión en las tres variables para el grupo de pacientes que mantuvieron el RS, siendo mucho más evidente para la *Fdom* y la *Em*.

La capacidad de predicción de los parámetros se muestra en la Tabla 2 en términos de AROC, donde en negrita se ha identificado la derivación para la cual dicha área fue máxima. Como se puede observar, para los tres parámetros el mejor resultado de predicción del éxito de la CVE se obtuvo sobre la derivación II, donde la A_f alcanzó una exactitud alrededor del 70 % (Se=58.1 % y Sp=70.4 %), presentando además valores de Se y Sp más balanceadas que V1 (AROC=57.0 %, Se=87.1 % y Sp=40.7 %). La *Fdom* también mostró una capacidad de predicción superior a V1 (AROC=76.6 %, Se=80.6 % y Sp=63.0 %), alcanzando casi un 82 % de exactitud y una notablemente mejor Sp (Se=80.6 % y Sp=74.1 %). Finalmente, la *Em* también fue capaz de mejorar la AROC casi un 7 % con respecto a la derivación V1 (AROC=74.6 %, Se=80.6 % y Sp=63.0 %), alcanzando una correcta predicción del resultado de la CVE en un 81.4 % (Se=81.5 % y Sp=67.8 %), y obteniendo mayores valores de Se y Sp.

Los resultados relativos a las diferencias estadísticas entre los dos grupos de pacientes también se muestran en la Tabla 2. En ella se puede observar que la *Fdom* obtuvo diferencias estadísticamente significativas (p <0.05) sobre la totalidad de las derivaciones, mientras que la Em no lo hizo en tres de ellas (I, V4, V5), y la amplitud de las ondas *f* en nueve (I, aVF, aVL, V1, V2, V3, V4, V5 y V6).

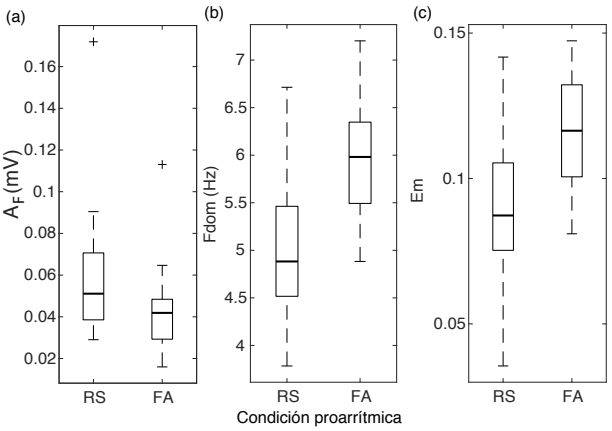


Figura 1. *Diagrama de cajas y bigotes para la distribución de (a) la amplitud, (b) la frecuencia dominante y, (c) la entropía muestral de las ondas f estimadas sobre la derivación II.*

4. Discusión y conclusiones

Este trabajo ha presentado una evaluación de tres de los índices más utilizados para caracterizar las ondas *f* presentes en la AA de las 12 derivaciones del ECG, haciendo hincapié en su capacidad de predecir la condición proarrítmica de un paciente y, en consecuencia, el resultado de la CVE.

El hallazgo más notable es que la mejor predicción del resultado de la CVE se ha obtenido sobre la derivación frontal II, aun cuando la mayoría de autores prefieren el uso de V1 [6, 8]. Esta preferencia se podría achacar a criterios anatómicos, ya que se considera que dicha señal es la que más cerca está de la aurícula derecha y, por tanto, presenta las ondas *f* de mayor amplitud con respecto a otras componentes ventriculares, tal como el complejo QRS [7, 8]. No obstante, la derivación frontal II está alineada con el tabique interauricular [13], y por tanto puede no solo registrar las ondas *f* de la aurícula derecha, sino también aquellas originadas en la aurícula izquierda. Este hecho sugiere que las ondas *f* de la derivación II pueden mostrar una AA más global procedente de ambas aurículas, redundando así en una mejor predicción de la condición proarrítmica del paciente. También, a diferencia de la derivación unipolar V1, la II se caracteriza por ser bipolar, lo que a su vez contribuye a la obtención de una señal de mejor calidad, con una mejor relación señal a ruido [13]. Además, algunos autores [5] consideran que la onda P monofásica de mayor amplitud se encuentra en la derivación II, por lo que teniendo en cuenta que durante la FA las ondas P se reemplazan por ondas *f*, es lógico pensar que esta señal también presenta ondas *f* de gran amplitud.

Respecto al comportamiento de los índices analizados, los resultados sugieren que la presencia de ondas *f* con mayor amplitud incrementan la probabilidad de mantener el RS tras la CVE. Este hallazgo es recurrente en la literatura [6, 8], aunque los valores de AROC que se han obtenido aquí son ligeramente inferiores a los obtenidos en otros trabajos previos. Esto puede estar asociado al algoritmo de computación, ya que tanto en [6] como en [8], la amplitud de las ondas *f* fue supervisada y corregida manualmente, mientras que aquí se ha estimado de forma automática.

Estadístico	Parámetro	I	II	III	aVL	aVR	aVF	V1	V2	V3	V4	V5	V6
AROC	A_f	0.593	**0.702**	0.651	0.621	0.682	0.657	0.570	0.569	0.517	0.540	0.612	0.625
	$Fdom$	0.670	**0.819**	0.725	0.744	0.760	0.790	0.766	0.754	0.698	0.688	0.750	0.674
	Em	0.628	**0.814**	0.657	0.663	0.731	0.753	0.745	0.706	0.644	0.650	0.736	0.683
p-valor	A_f	0.230	0.008	0.049	0.115	0.018	0.041	0.366	0.374	0.827	0.607	0.147	0.105
	$Fdom$	0.027	0.001	0.003	0.001	0.001	0.001	0.001	0.001	0.010	0.015	0.001	0.024
	Em	0.095	0.001	0.041	0.034	0.003	0.001	0.001	0.007	0.061	0.051	0.002	0.017

Tabla 2. *Capacidad discriminante en términos de AROC y significancia estadística (p) para la amplitud de las ondas f (A_f), la frecuencia dominante (Fdom), y la organización temporal (Em) sobre las 12 derivaciones estándar del registro de ECG.*

Por otro lado, la *Fdom* en pacientes con FA recidiva también ha mostrado ser superior que en aquellos que mantienen el RS, tal y como se podía esperar de acuerdo con trabajos previos [6, 8]. Al igual que en el caso de la A_f, la *Fdom* predice el resultado de la ECV con mayor exactitud sobre la derivación II, incrementándola en aproximadamente un 6 %, a la vez que aumenta la Sp. Esto puede deberse, al igual que con la A_f, a que si la amplitud de las ondas f es mayor, y el algoritmo de cancelación del complejo QRST es capaz de estimar morfologías más homogéneas, se puede realizar una mejor reducción de las interferencias ventriculares, lo que se traduce en una mejor estimación frecuencial de la *Fdom*.

En cuanto a la regularidad de las ondas *f*, la Em presenta un comportamiento similar al de la *Fdom*, donde se obtuvo un mayor valor mediano para pacientes con FA recidiva. En trabajos previos también se ha encontrado una tendencia similar [6, 8], sugiriendo que los pacientes con señales a priori más desorganizadas presentan una menor capacidad de mantener un RS a medio o largo plazo. Aunque no se conocen en profundidad los mecanismos de generación y mantenimiento de la FA, una mayor irregularidad de las ondas f puede ser indicativa de un mayor grado de remodelado del sustrato arritmogénico, lo que puede resultar a su vez en una mayor predisposición a desarrollar FA [6, 8]. De igual forma que con los índices anteriores, la mejor predicción se obtuvo sobre la derivación frontal II.

Finalmente, destacar que los resultados obtenidos en el estudio son similares a los presentados en otros trabajos previos [11] o [14], aun cuando existen diferencias significativas entre ellos en lo que se refiere a las bases de datos analizadas y a las técnicas de control de ritmo cardíaco usadas. De hecho, en [11] se utilizaron los índices propuestos en el presente trabajo para predecir el resultado de la cardioversión farmacológica, y en [14] el de la ablación por catéter. No obstante, se puede concluir que para obtener una estimación más precisa de la condición proarrítmica del paciente tras la CVE parece ser más conveniente el uso de la derivación frontal II que la precordial V1, ya que permite una mejor predicción del fracaso temprano del procedimiento en pacientes con FA persistente.

Agradecimientos

Este trabajo ha sido cofinanciado por los proyectos de investigación competitiva DPI2017-83952-C3 de MINECO-AEI-FEDER-UE, SBPLY/17/180501/000411 de la JCCLM y AICO/2019/036 de la GVA.

Referencias

[1] January C, et al. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. Circ 2014;130(23).

[2] Duarte R, et al. Thrombin generation and other hemostatic parameters in patients with atrial fibrillation in use of warfarin or rivaroxaban. J Thromb Thrombolysis 2020;.

[3] Duytschaever M, et al. Factors influencing long term persistence of sinus rhythm after a first electrical cardioversion for atrial fibrillation. Pacing Clin Electro 1998;21(1):284–287.

[4] Fujimoto Y, et al. Advanced interatrial block is an electrocardiographic marker for recurrence of atrial fibrillation after electrical cardioversion. Int J Cardiol 2018;272:113–117.

[5] Corino V, et al. Signal processing methods for information enhancement in atrial fibrillation: Spectral analysis and non-linear parameters. Biomed Signal Process Control 2006;1(4):271–281.

[6] Alcaraz R, et al. Noninvasive time and frequency predictor of long-standing atrial fibrillation early recurrence after electrical cardioversion: predictor of cardioversion outcome. Pacing Clin Electrophysiol 2011;34(10):1241–1250.

[7] Bollmann A, et al. Analysis of surface electrocardiograms in atrial fibrillation: techniques, research, and clinical applications. EP Europace 2006;8(11):911–926.

[8] Lankveld T, et al. The ECG as a tool to determine atrial fibrillation complexity. Heart 2014;100(14):1077–1084.

[9] Vest A, et al. An open source benchmarked toolbox for cardiovascular waveform and interval analysis. Physiol Meas 2018;39(10):105004.

[10] Alcaraz R, et al. Adaptive singular value cancelation of ventricular activity in single-lead atrial fibrillation electrocardiograms. Physiol Meas 2008;29(12):1351–69.

[11] Zeemering S, et al. The electrocardiogram as a predictor of successful pharmacological cardioversion and progression of atrial fibrillation. Europace 2018;20(7):e96–e104.

[12] Richman J, Moorman J. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol 2000;278(6):H2039–H2049.

[13] Park J, et al. Early differentiation of long-standing persistent atrial fibrillation using the characteristics of fibrillatory waves in surface ECG multi-leads. Sci Rep 2019;9(1):2746.

[14] Lankveld T, et al. Atrial fibrillation complexity parameters derived from surface ECGs predict procedural outcome and long-term follow-up of stepwise catheter ablation for atrial fibrillation. Circ Arrhythm Electrophysiol 2016;9(2).

Estudio de la asociación entre los patrones de activación local y de sincronización global en el EEG en la enfermedad de Alzheimer

V. Rodríguez-González[1], C. Gómez[1,2], M. Revilla-Vallejo[1], V. Gutiérrez-de Pablo[1], M.A. Tola-Arribas[2,3], M. Cano del Pozo[3], R. Hornero[1,2,5], J. Poza[1,2,5]

[1] Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, victor.rodriguez@gib.tel.uva.es
[2] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España
[3] Servicio de Neurología, Hospital Universitario Río Hortega, Valladolid, España
[4] Servicio de Neurofisiología Clínica, Hospital Universitario Río Hortega, Valladolid, España
[5] IMUVA, Instituto de Investigación en Matemáticas de la Universidad de Valladolid, Valladolid, España

Resumen

Dos de los métodos más utilizados para estudiar la actividad electroencefalográfica (EEG) consisten en estimar la densidad espectral de potencia (PSD) para cuantificar los patrones de activación local y en realizar un análisis de red que resuma los patrones de sincronización entre diferentes regiones cerebrales. Ambas aproximaciones se han empleado para caracterizar las alteraciones neurofisiológicas de la enfermedad de Alzheimer (EA): la primera muestra una lentificación de la actividad neuronal y la segunda una desconexión global. Estudios recientes apuntan a que ambas metodologías están relacionadas; sin embargo, esta asociación no ha sido explorada en la EA. En este estudio, se analiza por tanto la relación entre los patrones de activación local y de sincronización global. Para ello, se introduce una nueva metodología para generar redes funcionales con alta resolución en frecuencia, lo cual posibilita el cálculo de la distribución en frecuencia de los parámetros de red y una comparación directa con la PSD. Los resultados muestran que el coeficiente de agrupamiento sigue una evolución en frecuencia similar a la PSD, con valores de correlación elevados entre ambos. Además, en el grupo de enfermos con EA se produce una disminución de esta asociación. En definitiva, la relación observada apunta a que existen procesos neuronales subyacentes que modulan el balance entre los patrones de activación local y de sincronización global, y el cual se ve alterado por la EA.

1. Introducción

La enfermedad de Alzheimer (EA) es una enfermedad neurodegenerativa que provoca daños cognitivos, funcionales y conductuales [1]. Se trata de la causa de demencia más común, con una prevalencia especialmente alta en los países occidentales, donde tiene un gran impacto económico y social [2]. La EA produce alteraciones en la actividad cerebral que pueden ser medidas con diferentes técnicas: resonancia magnética funcional (fMRI), magnetoencefalografía (MEG) o electroencefalografía (EEG) [3]. En este estudio nos hemos centrado en la actividad EEG, ya que permite adquirir de forma no invasiva las señales eléctricas del cerebro, tiene una alta resolución temporal, su coste es bajo y es una técnica portable [4]. Estas ventajas han propiciado que su uso esté muy extendido en entornos clínicos [3], [4].

Una de las metodologías más extendidas para caracterizar los registros EEG en pacientes con EA consiste en analizar su contenido en frecuencia mediante la densidad espectral de potencia (PSD, *power spectral density*,). Este tipo de análisis miden la potencia de la sincronización local de los generadores neuronales. Estos análisis han permitido observar una lentificación de la actividad cerebral en esta patología [3], [4]. En los últimos años, se han desarrollado nuevos métodos que permiten obtener una caracterización más exhaustiva de las alteraciones que provoca la EA. En este sentido, diversas investigaciones han observado alteraciones en la complejidad de la actividad cerebral, así como cambios en su sincronización global, reflejados en los patrones de conectividad y en las propiedades de la red neuronal [4]–[6]. Entre los parámetros de red, el coeficiente de agrupamiento (CA) ha sido ampliamente usado para cuantificar la segregación de la red a escala global. Se ha visto previamente que la EA provoca descensos del CA incluso en su estado preclínico [7]. A pesar de que las alteraciones del EEG han sido obtenidas con metodologías muy diversas y que atienden a diferentes escalas de análisis (local y global), unos pocos estudios han planteado la hipótesis de que pueden no ser totalmente independientes [5], [8]. Así, se ha sugerido que existe una relación entre la lentificación del EEG y la pérdida de complejidad en la EA [5]. La relación entre el contenido espectral y diversos parámetros de red también se ha observado en modelos computacionales y registros EEG de sujetos sanos [8], [9]. Sin embargo, son necesarios nuevos análisis que permitan confirmar estas asociaciones y permitan entender su naturaleza.

En el presente trabajo se plantea la hipótesis de que la EA altera las funciones cognitivas de alto nivel que dependen del balance entre la activación local y la sincronización global de la actividad cerebral. Como medida de activación local se empleó la PSD, mientras que para cuantificar la segregación de la red a escala global se empleó el CA. Dado que las métricas extraídas de la PSD y el CA presentan ciertas analogías a diferente nivel (local y global), sería plausible plantear que existe cierta asociación entre ambos y que ésta estaría alterada como consecuencia de la EA. Para evaluarlo, este estudio tiene como objetivo estudiar la relación entre la PSD y la distribución en frecuencia del CA, calculados a partir del EEG de sujetos sanos y enfermos con demencia por EA.

2. Materiales

2.1. Sujetos

Con el objetivo de poder replicar los resultados, se han incluido dos bases de datos diferentes. Cada una estaba formada por 100 sujetos y fueron recogidas gracias a la colaboración del Hospital Universitario Río Hortega de Valladolid (HURH) y al proyecto POCTEP ‘Análisis y correlación entre el genoma completo y la actividad

ISBN: 978-84-09-25491-0, pags. 121- 124

Base de datos	Datos	Controles	EA leve
POCTEP	Sujetos (H:M)	25:25	21:29
	Edad (años)	80.1 ± 7.1	80.5 ± 6.9
	Formación (A:B)	33:18	36:14
	MMSE	28.8 ± 1.1	22.5 ± 2.3
HURH	Sujetos (H:M)	14:31	26:29
	Edad (años)	76.3 ± 3.9	79.7 ± 5.9
	Formación (A:B)	17:28	37:18
	MMSE	28.8 ± 1.1	23.0 ± 2.0

Tabla 1. Datos sociodemográficos y clínicos de las dos bases de datos. Los valores se expresan como: media ± desviación estándar. EA leve: pacientes con EA leve; H: hombre; M: mujer; A: educación primaria o inferior; B: secundaria o superior; MMSE: Mini-Mental State Examination (rango: [0 30]).

cerebral para la ayuda en el diagnóstico de la enfermedad de Alzheimer' (Programa 'Interreg V-A España-Portugal, 2014-2020'). Sus características se resumen en la Tabla 1.

En ambos casos los enfermos fueron diagnosticados mediante los criterios del *National Institute of Aging and Alzheimer's Association* (NIA-AA) [10], mientras que los controles eran personas de edad avanzada, cognitivamente sanas y sin una historia clínica con enfermedades neurológicas o psiquiátricas. Los participantes y cuidadores, en su caso, dieron su consentimiento escrito e informado. Los Comités Éticos de la Universidad de Porto (Portugal) y del HURH (España) aprobaron el estudio de acuerdo con el Código Ético de la Asociación Médica Mundial (Declaración de Helsinki).

2.2. Registros electroencefalográficos

Las señales EEG se registraron mediante sistemas de 19 canales (POCTEP: Neurofax JE921A, Nihon Kohden; HURH: XLTEK®, Natus Medical). Los electrodos se colocaron siguiendo el sistema internacional 10-20: Fp1, Fp2, Fz, F3, F4, F7, F8, Cz, C3, C4, T3, T4, T5, T6, Pz, P3, P4, O1 y O2. Las frecuencias de muestreo fueron de 500 para POCTEP y de 200 Hz para HURH. Se registraron 5 minutos de actividad en reposo mientras los sujetos estaban despiertos, relajados y con los ojos cerrados. Para prevenir la somnolencia, se monitorizaron los registros en tiempo real. A continuación, las señales se preprocesaron de manera similar [11]: (i) filtrado a 50 Hz para la eliminar la componente de red; (ii) filtrado FIR paso banda entre [1 70] Hz; (iii) análisis de componentes independientes (ICA) para eliminar componentes ruidosas; y (iv) selección visual de épocas de 5 s libres de artefactos.

3. Metodología

3.1. sLORETA

Con el fin de minimizar los efectos de conducción de volumen, los análisis se llevaron a cabo a nivel de fuente. Debido al limitado número de sensores disponibles, la estimación de la activación de las fuentes cerebrales se llevó a cabo mediante el algoritmo *standardized Low Resolution Brain Electromagnetic Tomography* (sLORETA). Este método restringe las soluciones asumiendo máxima correlación entre generadores neuronales vecinos; es decir tiene la suposición *a priori* de que la actividad de los generadores neuronales vecinos está correlada [12]. En este estudio se ha trabajado con 15000 fuentes neuronales, las cuales han sido agrupadas en 68 regiones de interés (ROIs) de acuerdo con el atlas de Desikan-Killiany [13].

3.2. Estimación de la densidad espectral de potencia

A partir de la activación temporal de cada una de las ROIs bajo estudio, se calculó la PSD. Para ello, se utilizó el método de Blackman-Tuckey, con una ventana rectangular, de ancho 5 s y sin solapamiento. Finalmente, cada PSD se normalizó por su potencia total, obteniendo así la PSD normalizada (PSDn).

3.3. Construcción de la red de alta resolución

Existen numerosas medidas que pueden ser utilizadas para calcular la conectividad entre cada par de ROIs. En este estudio hemos empleado el *Phase Lag Index* (PLI), que mide el acoplamiento entre series temporales basado en la fase [14]. Concretamente, el PLI evalúa la asimetría de la distribución de diferencias de fase entre dos series temporales [14]. Una de las ventajas del PLI es que es insensible a los efectos de conducción de volumen [14].

Típicamente, el PLI se calcula en determinadas bandas (delta: 1-4 Hz; zeta: 4-8 Hz; alfa: 8-13 Hz; beta: 13-30 Hz; y/o gamma: más de 30 Hz), lo cual permite evaluar su dependencia con la frecuencia [15]. Sin embargo, la resolución espectral obtenida es muy baja; esto evita que se pueda analizar con precisión las alteraciones en su distribución en frecuencia como consecuencia de una patología, como la EA. Para solventar este problema, en este trabajo se calculó el PLI en intervalos estrechos de frecuencia. Concretamente, se filtraron las señales previamente al cálculo del PLI empleando una ventana deslizante con un ancho de banda de 1 Hz (*finite impulse response*, ventana de Hamming, orden 300) y con un solapamiento del 50%. Esto permitió generar las matrices de conectividad de PLI con una alta resolución en frecuencia (red funcional de alta resolución). Estas matrices se usaron para calcular finalmente el coeficiente de agrupamiento (CA). El CA es un parámetro de red que mide cómo de conectado está un nodo con sus vecinos en la red [15]. Específicamente, el CA se define como la probabilidad de que los nodos vecinos de un nodo estén también conectados entre sí [15].

Por último, con el objetivo de capturar variaciones genuinas en las propiedades de la red, se generaron 500 modelos subrogados reordenando aleatoriamente la matriz de conectividad, calculándose el CA asociado a cada uno. Los valores de CA obtenidos inicialmente se normalizaron entonces dividiendo por la media de los valores de CA de los modelos subrogados (CAn).

4. Resultados

En la Figura 1 se muestran las PSDn y la distribución en frecuencia de valores del CAn para los dos grupos de estudio en las dos bases de datos analizadas. Se puede observar que las distribuciones en frecuencia de la PSDn y de CAn para cada grupo es muy similar para las dos bases datos. Además, la Figura 1 muestra la evolución diferencial

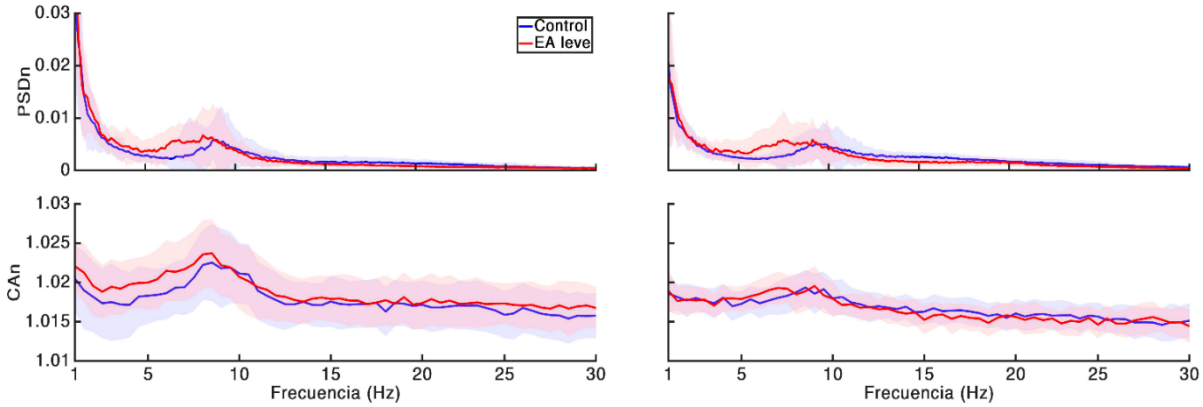


Figura 1. *PSDn y distribución en frecuencia de la CAn para los dos grupos (controles: azul; enfermos con EA leve: rojo) y las dos bases de datos (POCTEP: columna izquierda; HURH: columna derecha). La línea más gruesa se corresponde con el valor medio, mientras que el área sombreada representa la desviación estándar de la distribución.*

del grupo patológico, en el que se aprecia la típica lentificación de actividad EEG asociada a la EA [4]. Curiosamente, la distribución del CAn parece reflejar también un incremento de este parámetro de red para las bajas frecuencia conforme avanza la patología.

Con el objetivo de cuantificar la asociación entre las distribuciones de la Figura 1, se calculó el test de correlación de Spearman entre las distribuciones de valores de la PSDn y del CAn para cada sujeto. La Figura 2 muestra los histogramas correspondientes a la distribución de valores del coeficiente de correlación para cada grupo en las dos bases de datos. Se aprecia que una gran cantidad de sujetos presentan correlaciones significativas ($p < 0.05$, correlación de Spearman) para ambas bases de datos (POCTEP: controles 94%, pacientes con EA 74%; HURH: controles 89%, pacientes con EA 93%), con valores medianos de correlación elevados (POCTEP: controles 0.63, pacientes con EA 0.51; HURH: controles 0.69, pacientes con EA 0.65).

5. Discusión

Los resultados obtenidos sugieren que existe cierta relación entre los patrones de activación neuronal en frecuencia y la

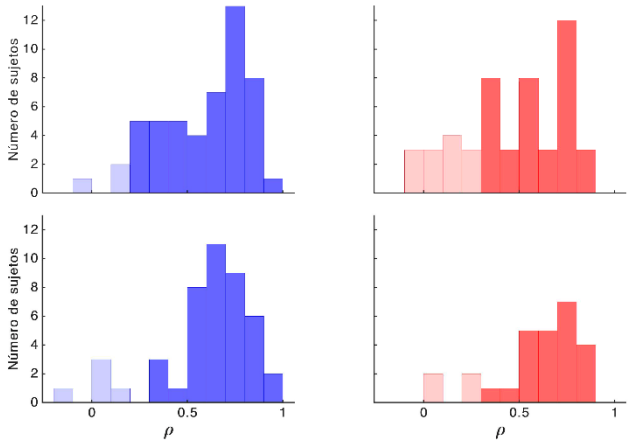


Figura 2. *Histogramas de los valores de correlación entre la PSDn y la distribución en frecuencia de CAn para los dos grupos (controles: azul; enfermos con EA leve: rojo) y las dos bases de datos (POCTEP: fila superior; HURH: fila inferior. Los sujetos cuyo valor de correlación es significativo (p<0.05, correlación de Spearman) se muestran con un color más oscuro.*

configuración de la red neuronal funcional. Además, esta asociación se ve reducida por la EA.

La relación entre la potencia en la banda alfa y el CA en registros EEG en reposo ya había sido sugerida en un estudio reciente realizado por Demuru *et ál*. [8]. Nuestros hallazgos confirman esta idea con dos bases de datos independientes y extienden estos resultados, ya que las redes de alta resolución presentadas en este estudio permiten analizar con precisión la distribución en frecuencia de los parámetros de red. Sin embargo, no está claro cuál es el mecanismo responsable de tal asociación. En nuestro estudio se ha utilizando el PLI, que cuantifica el acoplamiento en fase [14] y que, *a priori*, no se considera a la hora de estimar la PSD. Asimismo, la red funcional se construyó a partir de la activación de los generadores neuronales mediante sLORETA; la combinación de sLORETA y el PLI minimiza por tanto los posibles efectos de conducción de volumen y dispersión del campo eléctrico inherentes al EEG de superficie [14]. Otro aspecto que habría que descartar sería la frecuencia de muestreo (o tamaño de las series temporales), ya que los resultados con las dos bases de datos son similares, en línea con lo obtenido en [8]. En base a las ideas anteriores, podría considerarse que existe algún mecanismo cerebral subyacente que modula los patrones de activación local (medidos mediante la PSDn) y de red (medidos con el CAn). Dada la escala de análisis de la PSDn y del CAn, esto podría implicar que la sincronización neuronal se produce conjuntamente a escala local y global, por lo que ambos análisis estarían midiendo un balance diferente de la combinación entre activación local y sincronización global. Esta hipótesis va en línea con un estudio previo en el que se observó que la PSD reflejaba, no sólo patrones de activación local, sino también la sincronización entre generadores neuronales distantes [16].

Asimismo, se trabajó con dos grupos de sujetos (controles y pacientes con EA) para tratar de caracterizar las alteraciones que esta patología induce en la relación existente entre sincronización local y global, pues es muy común utilizar este tipo de métricas para tratar de caracterizar esta demencia [4], [14]. En este sentido, se ha podido observar que la aparición de la EA provoca una reducción de la relación entre sincronización local y global, pues para ambas bases de datos el valor mediano de correlación disminuye en pacientes con EA. Esto podría

deberse a la combinación entre la pérdida de generadores en la banda alfa [17] y la desconexión asociada a la EA [6].

Finalmente, existen ciertas limitaciones en este estudio que cabe mencionar. En primer lugar, sólo se ha usado un algoritmo de localización de fuentes, un parámetro de conectividad y uno de red; de cara a obtener resultados más robustos habría que replicar el estudio con mayor variedad de técnicas, e incluso con señales MEG. También sería interesante incluir otros grupos patológicos (enfermos con deterioro cognitivo leve y con estados más avanzados de demencia) para evaluar hasta qué punto la progresión de la EA afecta a la asociación entre los patrones de activación local y de sincronización global.

6. Conclusiones

Los análisis realizados en este estudio indican que los patrones de activación neuronal en frecuencia están relacionados con las propiedades de la red funcional. Esta relación parece estar modulada por procesos neuronales que influyen en los patrones de sincronización locales y globales. Asimismo, la EA atenúa la asociación entre sincronización local y global, pero sin llegar a eliminarla. Desentrañar la naturaleza de la asociación observada puede ayudar a entender de manera más profunda el funcionamiento del cerebro y a asociar de manera más precisa los resultados obtenidos con la fisiopatología de la EA.

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 PGC2018-098214-A-I00, por la Comisión Europea y el FEDER bajo los proyectos 'Análisis y correlación entre el genoma completo y la actividad cerebral para la ayuda al diagnóstico de la enfermedad de Alzheimer' y 'Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres' ('Programa de Cooperación Transfronteriza Interreg V-A España-Portugal POCTEP 2014-2020'), y por el CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto de Salud Carlos III cofinanciado con fondos FEDER. V. Rodríguez González es beneficiario de una ayuda PIF-UVa de la Universidad de Valladolid.

Referencias

[1] J. L. . Cummings *et ál.*, *La neuropsiquiatría de la enfermedad de Alzheimer y demencias relacionadas*. Barcelona: Atlas Medical Publishing, 2004.

[2] T. A. Salthouse, "What and When of Cognitive Aging," *Curr. Dir. Psychol. Sci.*, vol. 13, no. 4, pp. 140–144, Aug. 2004.

[3] C. Babiloni *et ál.*, "Fundamentals of Electroencefalography, Magnetoencefalography, and Functional Magnetic Resonance Imaging," *Int. Rev. Neurobiol.*, vol. 86, no. 5, pp. 67–80, 2009.

[4] J. Dauwels *et ál.*, "Slowing and Loss of Complexity in Alzheimer's EEG: Two Sides of the Same Coin?," *Int. J. Alzheimers. Dis.*, vol. 2011, no. 1, pp. 1–10, 2011.

[5] J. M. Buldú *et ál.*, "Frequency-based brain networks: From a multiplex framework to a full multilayer description," *Netw. Neurosci.*, vol. 2, no. 4, pp. 418–441, Oct. 2018.

[6] M. R. Brier *et ál.*, "Network Dysfunction in Alzheimer's Disease: Refining the Disconnection Hypothesis," *Brain Connect.*, vol. 4, no. 5, pp. 299–311, Jun. 2014.

[7] M. R. Brier *et ál.*, "Functional connectivity and graph theory in preclinical Alzheimer's disease," *Neurobiol. Aging*, vol. 35, no. 4, pp. 757–768, Apr. 2014.

[8] M. Demuru *et ál.*, "A comparison between power spectral density and network metrics: An EEG study," *Biomed. Signal Process. Control*, vol. 57, p. 101760, Mar. 2020.

[9] A. Daffertshofer *et ál.*, "On the Influence of Amplitude on the Connectivity between Phases," *Front. Neuroinform.*, vol. 5, no. 6, pp. 1–12, 2011.

[10] C. R. Jack *et ál.*, "NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease," *Alzheimer's Dement.*, vol. 14, no. 4, pp. 535–562, Apr. 2018.

[11] V. Rodriguez-Gonzalez, *et ál.*, "Assessing the influence of cognitive reserve in EEG signals through Alzheimer's Disease progression," in *Actas del XXXIV URSI*, 2019, pp. 1–4.

[12] R. D. Pascual-Marqui, "Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details.," *Methods Find. Exp. Clin. Pharmacol.*, vol. 24 Suppl D, pp. 5–12, 2002.

[13] R. S. Desikan *et ál.*, "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, Jul. 2006.

[14] P. Núñez *et ál.*, "Characterizing the fluctuations of dynamic resting-state electrophysiological functional connectivity: reduced neuronal coupling variability in mild cognitive impairment and dementia due to Alzheimer's disease," *J. Neural Eng.*, vol. 16, no. 5, p. 056030, Sep. 2019.

[15] M. Rubinov *et ál.*, "Complex network measures of brain connectivity: Uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, Sep. 2010.

[16] P. Tewarie *et ál.*, "Relationships Between Neuronal Oscillatory Amplitude and Dynamic Functional Connectivity," *Cereb. Cortex*, vol. 29, no. 6, pp. 2668–2681, Jun. 2019.

[17] J. Poza *et ál.*, "Characterization of the Spontaneous Electroencephalographic Activity in Alzheimer's Disease using Disequilibria and Graph Theory," in *35th Annual International Conference of the IEEE EMBS*, 2013, pp. 5990–5993.

Aplicación para usuarios intolerantes a la fructosa

M. Gimeno Tébar[1], N. Ortigosa[2,3]

[1] Escola Tècnica Superior d'Enginyeria, Universitat de València, cargite@alumni.uv.es

[2] I.U. Matemática Pura y Aplicada, Universitat Politècnica de València, nuorar@upvnet.upv.es

[3] Departament d'Informàtica, Universitat de València

Resumen

La intolerancia a la fructosa, a diferencia de la intolerancia a lactosa o al gluten, no dispone del etiquetado adecuado y específico en los productos comerciales. Por ello, el presente trabajo pretende solventar este problema mediante el desarrollo de una aplicación capaz de, analizando una fotografía de los ingredientes de un producto, indicar si el mismo es adecuado o no para los diferentes niveles de intolerancia a fructosa. Por otra parte, también incluirá un listado de alimentos ordenados por cantidad de fructosa por cada 100 gramos de alimento en el que el usuario pueda buscar directamente un producto (natural o procesado) y comprobar si su composición hace que sea adecuado para la intolerancia a la fructosa o no, indicándose mediante un código de colores tipo semáforo. La aplicación final es capaz de proporcionar una ayuda a usuarios que padezcan este tipo de intolerancia alimentaria, facilitándoles el reconocimiento de aquellos productos que contengan fructosa en alguno de sus ingredientes o edulcorantes, y funcionando con todo tipo de productos, incluso aquellos nuevos que se lancen al mercado.

1. Introducción

En los últimos diez años se ha duplicado el número de personas que padecen algún tipo de intolerancia o alergia alimentaria. Actualmente, más del 20% de la población padece alguna intolerancia alimentaria, y las intolerancias a la lactosa, al gluten o a la fructosa se hayan entre las más frecuentes [1].

Además, siguiendo este ritmo de crecimiento, se prevé que para el año 2025 casi el 50 por ciento de la población tenga algún tipo de intolerancia alimentaria [2].

La fructosa es un azúcar simple que se encuentra de forma natural en la miel, en muchas frutas (entre el 20% y el 40% de ellas) y otros alimentos, y que se ingiere principalmente en forma de sacarosa [3]. La sacarosa, también conocida como azúcar común, es una molécula compuesta por glucosa y fructosa y que, al ser ingerida, se separa en el intestino en sus dos moléculas. Otra posible fuente de fructosa común es el sorbitol, un edulcorante que al ser metabolizado por nuestro organismo genera este azúcar. La fructosa debe ser absorbida por las células de nuestro intestino y metabolizada por nuestro organismo para poder ser utilizada. En el momento en que existe algún problema en su absorción o metabolización es cuando surge la intolerancia a la fructosa.

La intolerancia hereditaria a la fructosa (IHF) es una enfermedad donde la enzima aldolasa B, que degrada la fructosa en hígado, intestino y riñones, no funciona de forma adecuada. Es una enfermedad hereditaria que se transmite de padres a hijos. Además, está considerada como una patología rara, ya que tiene una baja incidencia en la población (se estima que afecta a 1 de cada 20.000 personas) [4].

El único tratamiento que existe actualmente para los afectados de IHF es una dieta estricta sin fructosa en la que no se consuma más de 1-2 gramos al día ya sea en forma de fructosa, sorbitol o sacarosa.

Por otra parte, la malabsorción de la fructosa es una situación mucho más común que afecta a alrededor del 30% de la población en diferentes grados [5]. El tratamiento consiste en disminuir la ingesta de fructosa o sorbitol, así como los alimentos que contengan mezclas de estos [6], y restringir los alimentos que contengan más fructosa que glucosa. En este caso la sacarosa no suele ser un problema, pues al romperse en el intestino libera la misma cantidad de glucosa que de fructosa.

Existen diferentes aplicaciones móviles para analizar la composición de los productos del mercado a partir del código de barras de los mismos y que, como resultado, devuelven si el producto es apto o no según las diferentes intolerancias alimentarias. Por ejemplo, entre las más populares en España se encuentran *Facemovil* [7], utilizada por personas celiacas o que deban realizar una dieta sin gluten. Esta aplicación incluye un escáner del código de barras de un producto para consultar si está dentro de sus listas de alimentos permitidos. Además, contiene información sobre restaurantes, hoteles o puntos de venta que ofrecen opciones sin gluten.

Por otra parte, las aplicaciones *¿Qué puedo comer?* [8], y *¡Bueno para mí!* [9] contienen información alergénica y nutricional sobre gran diversidad de productos. En ambas, al escanear el código de barras de un producto mediante la cámara del teléfono móvil, se obtiene una ficha de este donde se detalla si su consumo es apto o no en función del perfil del usuario. En la segunda, además se aportan alternativas de otros artículos que sí se ajustan a la intolerancia alimentaria indicada por el usuario.

Actualmente, cuando a un paciente se le detecta la intolerancia a la fructosa, se le entrega un folleto informativo general y no específico sobre los productos que debe evitar. No obstante, existe un amplio listado de edulcorantes y aditivos empleados en la industria

ISBN: 978-84-09-25491-0, pags. 125- 128

alimentaria que no son bien tolerados por los pacientes con intolerancia a la fructosa.

Además, con frecuencia surgen dudas, ya sea porque en el etiquetado de los alimentos no conste con claridad cuáles son los edulcorantes utilizados, o porque alguno de los edulcorantes utilizados habitualmente puede contener fructosa o algún producto que se metabolice por la misma vía.

Dada la cantidad de población que demanda algún tipo de solución o ayuda con este problema y la poca oferta de soluciones tecnológicas que proporcionen información fiable que hay en el mercado en el caso de la intolerancia a la fructosa, existe una gran demanda por cubrir. Las aplicaciones generalistas que proveen información sobre el contenido de fructosa a partir del código de barras tienen múltiples carencias. Entre otras, por ejemplo, la no disponibilidad del código de barras de un producto en su base de datos, no pudiendo aportar información o indicaciones al usuario.

Por este motivo se ha decidido implementar una aplicación específica para las personas afectadas por la IHF, aunque también se ha tenido en cuenta a aquellos que padecen malabsorción a la fructosa, ya que el estudio de estas es muy similar, ambas intolerancias se basan en el control del mismo tipo de alimentos y edulcorantes, y la gente afectada por esta última es mucho mayor.

2. Materiales y métodos

La aplicación desarrollada plantea que, con solo tomar una fotografía al apartado de ingredientes de cualquier producto, se extraiga el texto de la imagen y se consulte en una base de datos previamente definida si alguno de los ingredientes contiene fructosa para mostrárselo al usuario.

Dependiendo del perfil del usuario (IHF o malabsorción, que será indicado en el primer paso de la aplicación), se utilizarán uno de los dos archivos JSON de las bases de datos creadas para tal fin. La información que contienen estos archivos ha sido obtenida y contrastada, especialmente con la base de datos de composición de alimentos del Departamento de Agricultura de Estados Unidos (USDA) [10], donde se detallan la cantidad de fructosa de cada alimento o ingrediente por cada 100gramos de este.

Para facilitar un amplio uso de la aplicación, esta debe poder utilizarse en cualquier dispositivo móvil, independientemente de su sistema operativo (Android y iOS). La interfaz debe ser sencilla e intuitiva, haciendo que los pasos para comprobar si un producto contiene fructosa sean lo menos tediosos posible. Además, la aplicación está pensada para que pueda funcionar en modo offline, aunque se recomienda tener una conexión a la red periódica para actualizar la base de datos en caso de que exista alguna modificación en la misma.

Por todo lo anterior, se ha utilizado *Visual Studio Code* con *Ionic Framework* para realizar la aplicación. Además de las librerías incluidas en la plataforma de *Ionic*, también se utilizarán las librerías *Tesseract* (para el reconocimiento óptico de caracteres de la imagen a analizar capturada por el usuario) y *RxJS* (para trabajar con el archivo JSON de la base de datos de forma óptima).

3. Resultados

La funcionalidad de la aplicación implementada es la siguiente:

1. Al entrar por primera vez en la aplicación, se muestra un breve tutorial.

2. Al finalizar el tutorial, el usuario debe especificar si padece IHF o malabsorción a la fructosa.

3. El usuario puede introducir los ingredientes de un producto mediante la adquisición o carga de una fotografía al apartado ingredientes de este.

4. El usuario también puede introducir los ingredientes (o buscar alguno en particular) mediante texto introducido manualmente.

5. En el caso de tomar una fotografía, se permite recortar o rotar la imagen para ayudar a su posterior lectura.

6. El sistema analiza la imagen para extraer el texto en ella y vuelca el resultado a un cuadro de texto.

7. El sistema comprueba automáticamente si alguno de los ingredientes existe en la lista de las bases de datos de ingredientes con contenido de fructosa, así como palabras muy similares en caso de posibles errores en el reconocimiento óptico.

8. La aplicación muestra, en caso de encontrar resultados, los ingredientes que contengan fructosa. En caso de no contenerlos, se indica que el producto es adecuado para el usuario.

9. El usuario, si lo desea, también puede consultar una lista de alimentos básicos y filtrarlos por diferentes grupos predefinidos, indicándose en cada caso su contenido en fructosa.

A continuación, se muestran algunas capturas de pantalla de la aplicación desarrollada (Figuras 1 a 4).



Figura 1.*Pantalla de bienvenida de la aplicación y ejemplo de pantalla del tutorial.*

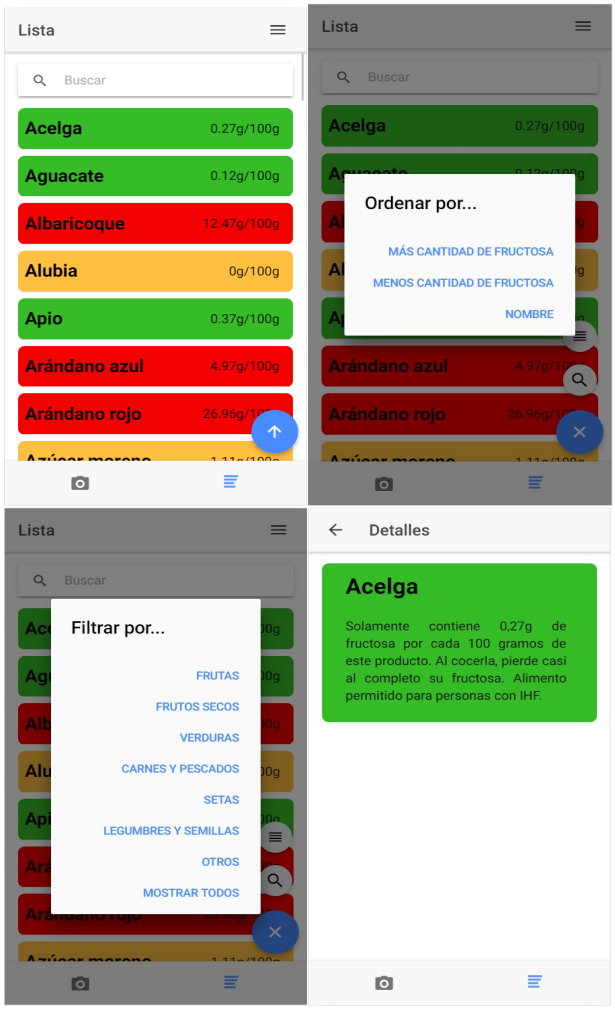


Figura 2. *Ejemplos de uso de la aplicación.*

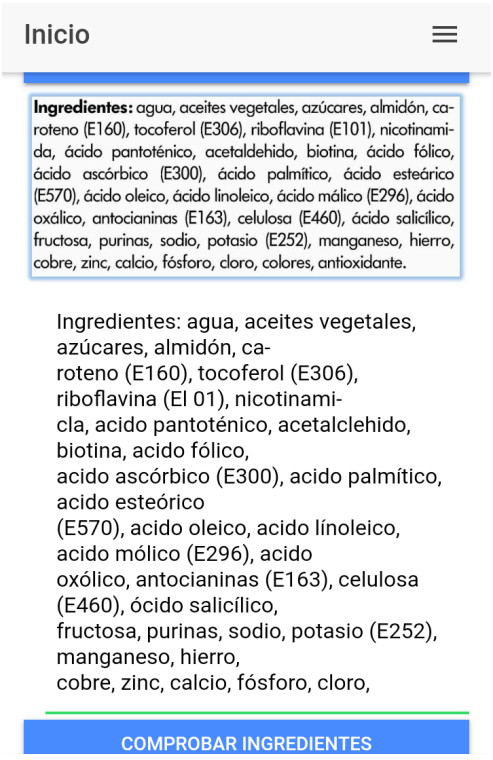


Figura 3. *Ejemplo de reconocimiento de texto en una imagen.*



Figura 4. *Resultados de la ventana "Ingredientes" tras realizar el análisis.*

Producto	% reconoc.	Muestra	Nº real
Mayonesa	100%	4	4
Bebida espelta	100%	1	1
Chocolate blanco	100%	3	3
Helado vainilla	100%	3	3
Barrita cereales	99.20%	4	4
Galleta de cacao	98.33%	3	3
Ketchup	94.74%	3	3
Yogur manzana	94.73%	3	3
Cereales	92.06%	3	3
Helado naranja	90.82%	8	9
Leche piña/coco	86.27%	5	5
Crema calabaza	74.19%	9	9
Mostaza	61.54%	1	2
Batido	47.62%	2	2
Mermelada	44.12%	2	2

Tabla 1. *Algunos ejemplos de productos utilizados en las pruebas de verificabilidad. Las columnas muestran el producto, el porcentaje de caracteres correctamente reconocidos, el número de ingredientes con fructosa que muestra la aplicación, y el número real de ingredientes con fructosa que contiene el producto.*

Para analizar los resultados de usabilidad de la aplicación, se han realizado diferentes tests tanto con un grupo diverso de usuarios (diferentes dispositivos móviles, diferentes sistemas operativos, diferentes grupos de edad), así como con imágenes tomadas en diferentes circunstancias (de iluminación, de enfoque, de encuadre, etc). Un total de 10 usuarios testearon la aplicación, tomando fotografías de un mínimo de 5 productos y un máximo de 10, comprobando los resultados obtenidos en cada caso.

En la Tabla 1 se muestran algunos de los resultados de reconocimiento de ingredientes problemáticos para

aquellas personas con distintos grados de intolerancia a la fructosa en diferentes conjuntos de productos alimentarios (de los que hemos omitido la marca), así como los porcentajes de reconocimiento del texto, el número de ingredientes mostrados por la aplicación como problemáticos, y el número que realmente contiene fructosa. No obstante, la aplicación ha sido testada en una amplia base de datos de imágenes (más de 50 imágenes de productos diferentes).

En global, sobre la base de datos de las imágenes analizadas, debía haberse mostrado un total de 138 ingredientes por contener fructosa. La aplicación ha mostrado 106 de esos ingredientes, pudiendo decir que su nivel de verificabilidad general es de un 76,81%. Este último porcentaje se obtiene al contar con todas las imágenes utilizadas por los usuarios, incluyendo aquellas en las que la calidad de la imagen capturada no es buena (mal enfoque, mala iluminación, brillos y reflejos). También aparecen errores de ingredientes no detectados cuando el usuario toma por ejemplo una fotografía de un envase cilíndrico, en el que los ingredientes no se ven en su totalidad o aparecen distorsionados. Este aspecto se puede solucionar tomando una nueva fotografía donde se puedan leer bien los mismos. Además, la aplicación también indica al usuario final el porcentaje de texto reconocido por lo que, si este es bajo, el usuario puede tomar una nueva fotografía de mejor calidad (mejor enfocada y con mejor iluminación) para obtener un mejor análisis.

Si únicamente se contabilizan aquellas imágenes en las que se ha reconocido más del 50% del número de ingredientes (las de una calidad de imagen suficiente), el número total de ingredientes que la aplicación debía mostrar son 96, de los cuales muestra 92, dejando así un nivel de verificabilidad del 95,83%. No obstante, para un usuario con intolerancia o malabsorción, lo importante es poder detectar que un producto contiene o no un ingrediente que le provocará reacción si lo ingiere. De esta forma, hay productos de la Tabla 1 (como pueden ser el helado de naranja o la mostaza, por ejemplo) de los que la aplicación no reconoce la totalidad de los ingredientes problemáticos para el usuario, pero sí detecta la presencia de al menos algunos de ellos. Por ello, a pesar de no haber reconocido el 100% de ingredientes, al haber detectado alguno, la aplicación lanzaría igualmente el aviso de que el alimento no es seguro para el usuario, pudiendo así este evitar trastornos posteriores debidos a su ingesta.

4. Conclusiones

En el presente trabajo se ha presentado una aplicación que, mediante la captura de una fotografía del listado de ingredientes de un producto, es capaz de proporcionar información acerca de la presencia de algún ingrediente que provoque una inadecuación del mismo para personas con distintos grados de intolerancia a la fructosa.

En base a las pruebas realizadas, la tasa de acierto al mostrar los ingredientes con fructosa en imágenes de calidad adecuada es superior al 95%. Además, el proyecto presentado elimina el problema de otras aplicaciones que necesitan que el código de barras del producto a analizar se encuentre registrado en la base de datos de la aplicación. Así, la aplicación presentada puede utilizarse sobre cualquier producto, incluyendo aquellos nuevos que se puedan ir lanzando al mercado.

Además, aunque la aplicación esté diseñada específicamente para gente con IHF o malabsorción a la fructosa, está implementada de forma que pueda usarse el mismo modelo para otro tipo de intolerancias alimentarias, como la intolerancia a la proteína de la leche, simplemente cambiando el contenido de la base de datos a la que hacemos las consultas.

Finalmente, hemos de indicar que el trabajo presentado en este artículo ha sido registrado bajo el número 2019-146150. Como trabajo futuro, se propone la difusión de la aplicación, tras haber realizado exhaustivos cuestionarios de un sistema de escalas de usabilidad.

Agradecimientos

N. Ortigosa agradece el apoyo de la Generalitat Valenciana con el proyecto Prometeo/2017/102 y del MINECO con el proyecto MTM2016-76647-P.

Referencias

[1] Lomer, MCE. Review article: the aetiology, diagnosis, mechanisms and clinical evidence for food intolerance. *Alimentary Pharmacology & Therapeutics,* vol 41, sup 3, 2015, pp 262-75 (ISSN:1365-2036).

[2] Ruiz Sánchez JG, Palma Milla S, Pelegrina Cortés B, et al. Una visión global de las reacciones adversas a alimentos: alergia e intolerancia alimentaria. *Nutrición Hospitalaria,* vol 35, sup 4, 2018, pp. 102-108 (ISSN: 0212-1611) .

[3] Ruiz Pons M, Sánchez-Valverde Visus F, Dalmau Serra J. Tratamiento nutricional de los errores innatos del metabolismo, 2015 (ISBN: 978-84-15950-60-8).

[4] Cornejo V, Raimann E. Alteraciones del metabolismo de la fructosa. *Revista chilena de nutrición,* vol 31, sup 2, 2004 (ISSN: 0717-7515).

[5] Pérez Alsa A, Riquelme A. Intolerancia a la fructosa-sorbitol. *Páginas informativas para pacientes de la comunidad hispanoparlante con síntomas gastrointestinales*. Asociación Española de Gastroenterología. Grupo CTO Editorial 2019.

[6] Izquierdo García E, Moreno Villares JM, León Sanz M. Edulcorantes en pacientes con intolerancia hereditaria a la fructosa. *Acta pediátrica española,* vol 72, 2014, pp 15-23 (ISSN : 0001-6640).

[7] Martín Cabrejas, I, Campos Caamaño A, Rivilla Lizcano L, et al. Manual de enfermedad celíaca: app Facemovil, 2017 (ISBN: 978-84-697-8507-2).

[8] ¿Qué puedo comer? https://www.quepuedocomer.es/ (Consultada: Octubre 2020).

[9] ¡Bueno para mí" http://www.buenoparami.com/ (Consultada: Octubre 2020).

[10] U.S. Department of Agriculture, Agricultural Research Service. FoodData Central: Food Composition Databases. https://ndb.nal.usda.gov/ndb/nutrients/index (Consultada: Octubre 2020).

Dispositivo laparoscópico mecánico articulado: Análisis de la curva de aprendizaje

J.A. Sánchez-Margallo[1], D. Durán Rey[2], F.M. Sánchez-Margallo[3]

[1] Unidad de Bioingeniería y Tecnologías Sanitarias, Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España, jasanchez@ccmijesususon.com

[2] Unidad de Laparoscopia, Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España, dduran@ccmijesususon.com

[3] Dirección Científica, Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España, msanchez@ccmijesususon.com

Resumen

En este trabajo presentamos el nuevo diseño de instrumento mecánico articulado para cirugía laparoscópica, en su versión de pinza de agarre y disector Maryland. Durante el estudio se analizó la evolución del rendimiento quirúrgico y la carga de trabajo de cinco cirujanos laparoscopistas noveles en el desarrollo de tareas básicas de aprendizaje laparoscópico. Los participantes llevaron a cabo varias tareas de coordinación y una tarea de disección y tres puntos intracorpóreos en tejido orgánico. Cada cirujano repitió cada tarea cinco veces. Para cada repetición, se evaluó el tiempo de ejecución, el número de errores y la carga de trabajo quirúrgico. Los resultados preliminares con el nuevo dispositivo muestran una curva de aprendizaje positiva con respecto al desempeño quirúrgico y el nivel de destreza con el uso del instrumental. La formación con el nuevo instrumento conlleva una reducción en la carga de trabajo experimentada por el cirujano y del número de errores cometidos durante la práctica laparoscópica.

1. Introducción

La cirugía laparoscópica plantea desafíos singulares a los cirujanos, como la reducción de la destreza, limitada libertad de movimientos y exigencias ergonómicas [1]. A medida que se utilizan técnicas mínimamente invasivas en procedimientos cada vez más complejos, las mejoras en la laparoscopia siguen estando a la vanguardia de la innovación quirúrgica.

Los múltiples avances en el campo de la robótica han dotado a los cirujanos de mejoras en la destreza y percepción visual del espacio de trabajo. A pesar de su novedad, la cirugía robótica ha sido objeto de críticas por el aumento de los costes y las necesidades de infraestructura. Del mismo modo, se han planteado diversos dispositivos articulados robotizados que ofrecen un incremento en los grados de libertad y un diseño ergonómicamente más adecuado [2,3]. Sin embargo, aunque mejorar la funcionalidad y la destreza es un objetivo común, lograrlo sin imponer un costo financiero significativo sigue siendo un desafío.

Recientemente, la empresa Innovative Medical Mannheim, (Mannheim, Alemania) ha presentado un nuevo diseño de instrumental laparoscópico flexible y con actuación puramente mecánica. Su objetivo es transferir los beneficios de la tecnología robótica a las manos del cirujano, proporcionando una alternativa asequible al alcance de todos los cirujanos. Actualmente, disponen de unas pinzas laparoscópicas de agarre y un disector Maryland. El portaagujas laparoscópico se encuentra en las últimas fases de desarrollo.

Por consiguiente, el objetivo de este trabajo es presentar el nuevo diseño de instrumento mecánico articulado para cirugía laparoscópica, en su versión de pinza de agarre y disector Maryland, y analizar la evolución del rendimiento quirúrgico y la carga de trabajo en el desarrollo de tareas básicas de aprendizaje laparoscópico.

2. Material y métodos

2.1. Instrumental laparoscópico

En este trabajo presentamos el diseño de instrumental laparoscópico de Innovative Medical Mannheim, en sus versiones de piza de agarre y disector Maryland (Figura 1). El diseño de instrumental tiene un diámetro de 5 mm y una punta que permite ser articulada hasta 90 grados y rotación de 360 grados. Los movimientos de la punta del instrumento se controlan con las dos anillas del mango.

2.2. Participantes y tareas

En este estudio participaron cinco cirujanos noveles (<10 procedimientos laparoscópicos realizados), cuatro diestros y un zurdo. Los participantes llevaron a cabo cuatro tareas básicas de formación en cirugía laparoscópica: (1) Una tarea de transferencia de objetos, (2) el paso de una aguja de sutura por un circuito de anillas, (3) una tarea de disección y (4) la realización de tres puntos intracorpóreos.

La tarea de transferencia de objetos consistió en coger tres objetos de color de un lado del tablero con su instrumental más próximo, transferir el objeto al instrumental opuesto y colocarlo al otro lado del tablero (Figura 2A). Una vez transferidos los tres objetos, el proceso es invertido. Para la siguiente tarea, se pidió a los participantes que condujeran una aguja recta a través de un circuito de anillas, distribuidas en diferentes ángulos (Figura 2B). La aguja fue manejada a través de las anillas usando la mano dominante y con el apoyo de la mano no dominante. El orden de las anillas y el sentido del paso de la aguja se

ISBN: 978-84-09-25491-0, pags. 129- 132

indicó con números y flechas en el tablero, respectivamente. Para las tareas anteriores, los cirujanos utilizaron la pinza de agarre articulada en su mano dominante y una pinza de agarre convencional en su mano no dominante.

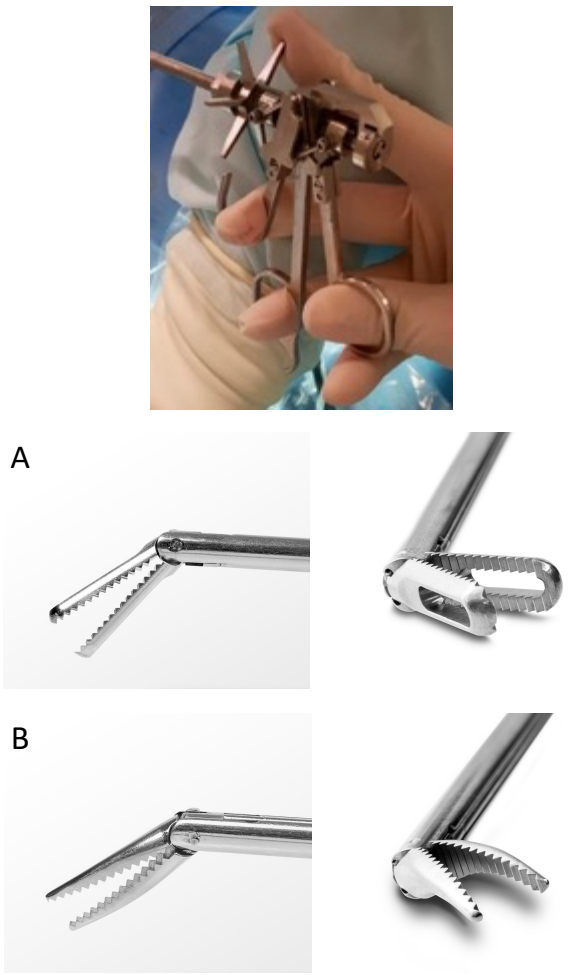


Figura 1. *Diseño del instrumental laparoscópico articulado: (A) Pinza de agarre y (B) disector Maryland.*

La tarea de disección consistió en separar la capa serosa de la capa muscular en un tejido orgánico (estómago porcino *ex vivo*) (Figura 2C). Para el desarrollo de la tarea, los cirujanos emplearon el disector articulado en su mano dominante y una pinza de agarre convencional en su mano no dominante. La tarea de sutura consistió en tres puntos simples en la incisión previamente disecada, formados por un nudo doble y dos simples, en sentidos opuestos (Figura 2D). Cada cirujano repitió cada tarea cinco veces y con un tiempo máximo de 10 min. para la realización de cada repetición. Todas las repeticiones de cada participante se realizaron el mismo día.

2.3. Evaluación

Para analizar la curva de aprendizaje, durante la primer (T1) y la última (T5) repetición, se evaluó el tiempo de ejecución y la carga de trabajo quirúrgico en las diferentes tareas laparoscópicas. Del mismo modo, para las tareas de coordinación (transferencia de objetos y circuito de anillas) se valuó el número de errores durante su desarrollo. Para la tarea de sutura intracorpórea también se valoró la calidad de la sutura laparoscópica.

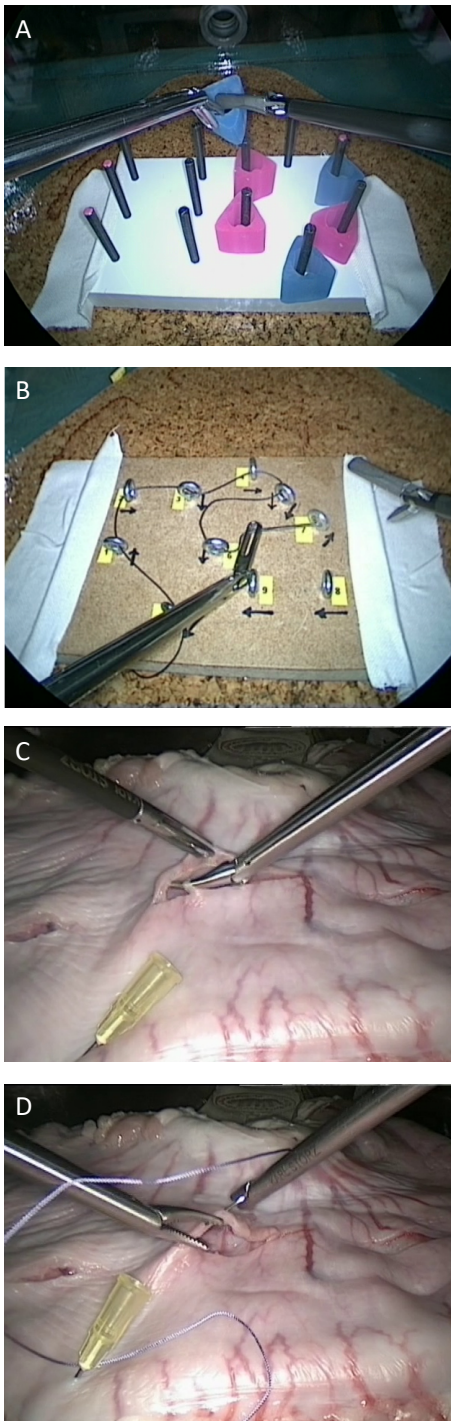


Figura 2. *Tareas de evaluación: (A) Transferencia de objetos; (B) Circuito de anillas; (C) Disección y (D) Sutura en tejido orgánico.*

Para la evaluación de la carga de trabajo quirúrgico, se empleó cuestionario subjetivo SURG-TLX [4], el cual es una adaptación para cirugía de la encuesta National Aeronautics and Space Administration Task Load Index (NASA-TLX). El cuestionario SURG-TLX evalúa seis dominios: Demanda mental, demanda física, demanda temporal, complejidad de la tarea, el estrés en el desarrollo de la tarea y el nivel de distracción. Por consiguiente, al final de cada repetición (T1 y T5), se solicitó a los participantes completar el cuestionario de evaluación de la carga de trabajo quirúrgico en la realización de cada tarea con el nuevo instrumental.

Para evaluar la calidad en la realización de cada sutura, se empleó una lista de comprobación para sutura laparoscópica previamente validada [5,6]. En esta lista de comprobación, se evalúan las características técnicas de la tarea de sutura como son la posición de la aguja y su manejo, la extracción de la aguja y la técnica y calidad de los nudos intracorpóreos. Este estudio no incluye nudos deslizantes, por lo que no se han considerado los puntos 25 (nudo que se deja suelto para que se deslice) y 26 (3 o <3 intentos de deslizamiento del nudo) de la lista de comprobación, lo que da como resultado una puntuación final máxima de 27. Un cirujano laparoscopista experimentado evaluó las tareas registradas de cada sujeto. El cirujano desconocía en todo momento la identidad y el nivel de experiencia de cada participante.

Para el análisis estadístico, se utilizó la Prueba de los rangos con signo de Wilcoxon para analizar la evolución del rendimiento de los cirujanos y su carga de trabajo. Todos los análisis estadísticos se realizaron utilizando el software R v4.0.0 (Fundación R para la Informática Estadística, Viena, Austria). Los resultados se muestran como valor medio y desviación estándar. Para todas las pruebas, $p<0,05$ se consideró estadísticamente significativo.

3. Resultados y discusión

En general, los cirujanos manifestaron una evolución positiva en el tiempo de ejecución de todas las tareas laparoscópicas analizadas, mostrando una reducción significativa para las tareas de paso de la aguja por el circuito de anillas, disección y sutura (Figura 3). Durante la tarea de sutura, los participantes consiguieron completar un máximo de dos puntos intracorpóreos dentro de los 10 min. de limite para la realización de la tarea.

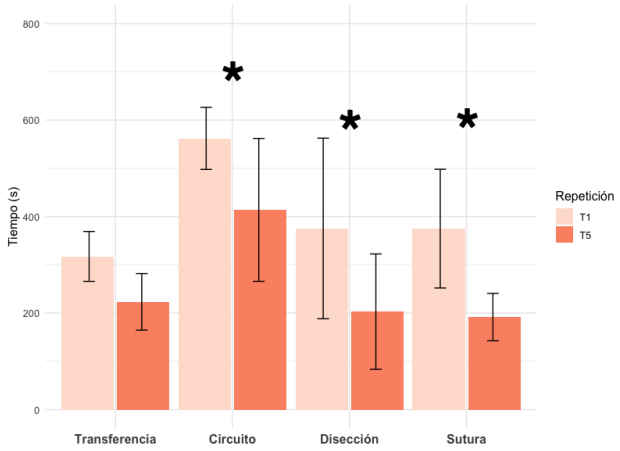


Figura 3. *Evolución del tiempo de ejecución para las diferentes tareas analizadas. *p<0,05.*

Respecto al análisis del número de errores cometidos durante las tareas de coordinación, se obtuvo una reducción significativa en la tarea de transferencia de objetos (Figura 4). Después del periodo de formación, los cirujanos mostraron una mejora en la destreza con el uso del nuevo instrumento articulado y en la coordinación óculo-manual durante el desarrollo de las tareas. Se observa que las tareas de coordinación más complejas (circuito de anillas) muestran una evolución más moderada en cuanto a la reducción en el número de errores cometidos con el manejo del nuevo instrumento.

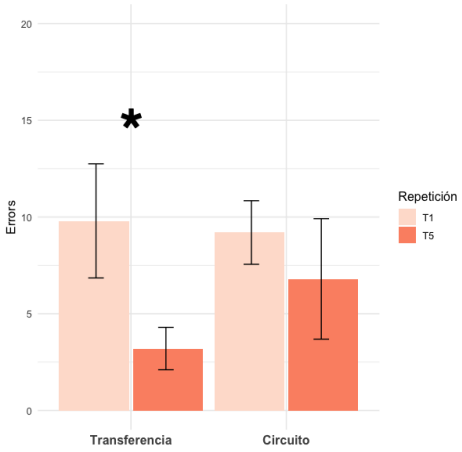


Figura 4. *Evolución del número de errores en las tareas de transferencia de objetos y circuito de anillas. *p<0,05.*

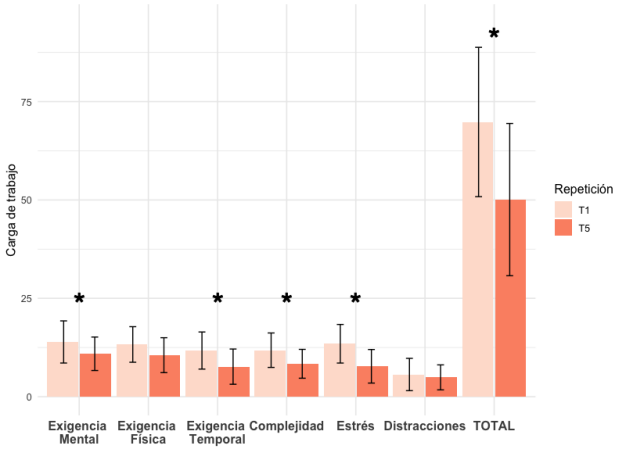


Figura 5. *Evolución de la carga de trabajo durante la ejecución de las tareas laparoscópicas. *p<0,05.*

El nivel de carga de trabajo quirúrgico experimentado por los cirujanos con el uso del dispositivo laparoscópico articulado disminuyó significativamente tras el periodo de formación (Figura 5). Esta reducción en la carga de trabajo quirúrgico se observó principalmente en la reducción de la exigencia mental y temporal durante el desarrollo de las tareas laparoscópicas, la complejidad de la tarea con el nuevo instrumental y el estrés experimentado durante la realización de las tareas.

Existen otros instrumentos laparoscópicos mecánicos flexibles en el mercado, centrados en incremental los grados de libertad y la maniobrabilidad durante la práctica laparoscópica. Algunos de estos ejemplos son el instrumental Radius r2 DRIVE (Tuebingen Scientific Medical GmbH, Tübingen, Alemania) y el dispositivo FlexDex® (FlexDex Inc., Brighton, MI, EE.UU.) [7], los cuales presentan características similares al instrumental actual, aunque con diferentes planteamientos en cuanto a su diseño. El instrumento r2 DRIVE presenta una punta flexible hasta 90 grados, proporcionando un total de siete grados de libertad. A

diferencia del instrumental actual, es necesario flexionar el mango del instrumento para controlar el movimiento de la punta. La rotación de la punta se controla con una rosca en el mango, la cual se acciona con el dedo pulgar e índice. Sin embargo, parece que el instrumento no resulta muy intuitivo, y requiere una amplia práctica para aprender a manejarlo de forma adecuada [8]. En el caso del instrumento FlexDex®, este dispositivo emplea un brazalete acoplado a la muñeca del cirujano para transmitir los movimientos de la mano, la muñeca y el brazo del cirujano a la punta del instrumental, a través de tiras de transmisión, poleas y cables [9]. A pesar del incremento en las posibilidades de movimiento de la punta del instrumental, la necesidad del adaptador para la muñeca hace que el dispositivo resulte aparatoso durante el desarrollo de una cirugía, en la cual el espacio de maniobrabilidad puede verse comprometido por el empleo de varios instrumentos quirúrgicos y la cámara laparoscópica. Por otro lado, los dos dispositivos anteriores son de un solo uso, lo cual puede llevar a un encarecimiento de los procedimientos quirúrgicos. El diseño de instrumental laparoscópico presentado en este trabajo está enfocado a poder ser esterilizado y facilitar su reutilización como material quirúrgico inventariable. Sin embargo, son necesarios más estudios con este nuevo dispositivo para poder determinar de forma precisa las posibilidades de esterilización, así como determinar un programa formativo hecho a medida para este nuevo dispositivo y que permita un aprendizaje eficaz para su aplicación a la práctica quirúrgica laparoscópica.

Este trabajo presenta una serie de limitaciones como el reducido número de participantes y la escasa complejidad de las tareas. En estudios futuros, se plantea ampliar el número de cirujanos participantes, incluyendo cirujanos con diferentes niveles de experiencia en cirugía laparoscópica. Los resultados obtenidos en el presente estudio nos permiten plantear nuevos trabajos con este nuevo diseño de instrumental laparoscópico en procedimientos laparoscópicos más complejos y poder incluir métodos de evaluación del desempeño quirúrgico como el Global Operative Assessment of Laparoscopic Skills (GOALS) [10]. Del mismo modo, se plantea ampliar este estudio incluyendo un mayor margen de tiempo entre repeticiones, de forma que podamos analizar si las habilidades y destrezas aprendidas se mantienen.

3.1. Conclusiones

Los resultados preliminares con el nuevo diseño de instrumento laparoscópico mecánico articulado muestran una curva de aprendizaje positiva de los cirujanos noveles con respecto al desempeño quirúrgico y el nivel de destreza con el uso del instrumental. La formación con el nuevo dispositivo conlleva una reducción en la carga de trabajo experimentada por el cirujano y con ello del número de errores cometidos durante la práctica laparoscópica.

Agradecimientos

Este trabajo ha sido financiado por la Junta de Extremadura (GR18199, TA18023, PD18077) (Cofinanciado por el Fondo Europeo de Desarrollo Regional).

Referencias

[1] Sánchez-Margallo JA, González González A, García Moruno L, Gómez-Blanco JC, Pagador JB, Sánchez-Margallo FM. Comparative Study of the Use of Different Sizes of an Ergonomic Instrument Handle for Laparoscopic Surgery. Appl Sci 2020;10(4):1526.

[2] Sánchez-Margallo FM, Sánchez-Margallo JA. Assessment of Postural Ergonomics and Surgical Performance in Laparoendoscopic Single-Site Surgery Using a Handheld Robotic Device. Surg Innov 2018;25(3):208–17.

[3] Sánchez-Margallo JA, Sánchez-Margallo FM. Initial experience using a robotic-driven laparoscopic needle holder with ergonomic handle: assessment of surgeons' task performance and ergonomics. Int J Comput Assist Radiol Surg 2017;12(12):2069–77.

[4] Wilson MR, Poolton JM, Malhotra N, Ngo K, Bright E, Masters RSW. Development and validation of a surgical workload measure: the surgery task load index (SURG-TLX). World J Surg 2011; 35: 1961–9.

[5] Kroeze SGC, Mayer EK, Chopra S, Aggarwal R, Darzi A, Patel A. Assessment of laparoscopic suturing skills of urology residents: a Pan-European study. Eur Urol 2009; 56(5):865–872

[6] Moorthy K, Munz Y, Dosis A, Bello F, Chang A, Darzi A. Bimodal assessment of laparoscopic suturing skills: construct and concurrent validity. Surg Endosc 2004; 18(11):1608–1612

[7] Sánchez-Margallo FM, Sánchez-Margallo JA, Szold A. Handheld Devices for Laparoscopic Surgery. In: New Horizons in Laparoscopic Surgery. InTech; 2018. p. 75–93.

[8] Ishikawa N, Watanabe G, Inaki N, Moriyama H, Shimada M, Kawaguchi M. The da Vinci surgical system versus the radius surgical system. Surgical Science. 2012;3:358-361

[9] Criss CN, Jarboe MD, Claflin J, Matusko N, Rooney DM. Evaluating a Solely Mechanical Articulating Laparoscopic Device: A Prospective Randomized Crossover Study. J Laparoendosc Adv Surg Tech 2019 Apr;29(4):542–50.

[10] Gumbs AA, Hogle NJ, Fowler DL. Evaluation of resident laparoscopic performance using global operative assessment of laparoscopic skills. J Am Coll Surg 2007;204:308–13.

MODELOS PREDICTIVOS EN EL TRATAMIENTO DE PACIENTES CON ALCOHOLISMO

L. Montero-Cordal[1], J. Tapia-Galisteo[1,2], A. Martínez-Maldonado[3,4,] I. Oropesa[1], Rosa Jurado-Barba[3,5], M.E. Hernando[1,2]

[1] Grupo de Bioingeniería y Telemedicina (GBT), Centro de Tecnología Biomédica, ETSI de Telecomunicación, Universidad Politécnica de Madrid, España, lucas.montero.cordal@alumnos.upm.es,{jtapia, ioropesa, elena}@gbt.tfo.upm.es

[2] Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, CIBER, Madrid, España

[3] Grupo de adicciones. Instituto de Investigación Biomédica H. 12 de Octubre, Madrid, España

[4] Universidad Rey Juan Carlos I, andresmartinezmaldonado@gmail.com

[5] Universidad Camilo José Cela, mrjurado@ucjc.edu

Resumen

Según la OMS, el consumo de alcohol es un problema de primer nivel, causando el 5.3% de las muertes a nivel mundial. Es también la droga más consumida en España, llegando a crear graves problemas de adicción. Entre los tratamientos disponibles contra el alcoholismo se encuentran las técnicas de modificación del sesgo cognitivo, mediante las cuales se entrena a los pacientes en el rechazo a los estímulos relacionados con el alcohol, tratando de devolver a rangos de normalidad ciertos parámetros cerebrales alterados en sujetos con esta dependencia, mejorando con ello el pronóstico de abstinencia.
El objetivo de este trabajo es crear modelos predictivos que permitan identificar la gravedad de la dependencia y el riesgo de recaída tras el tratamiento. Se han utilizado datos de 70 pacientes dependientes y 17 sujetos sanos, considerando datos demográficos, clínicos y cerebrales recogidos durante el tratamiento. Mediante técnicas de aprendizaje de máquina, se generan variantes de dos tipos de modelos predictivos: modelos regresivos para cuantificar la gravedad de la dependencia mediante la escala EIDA y clasificadores de recaídas a los tres meses de concluir el tratamiento. El mejor modelo regresivo de EIDA logra un coeficiente de correlación de 0,81. Los mejores modelos predictivos de recaída se logran mediante árboles J48 y redes neuronales, con precisiones de entre 0,92 y 0,93, exactitud del 86% y un área bajo la curva ROC (AUC) igual a 0,85. Estos modelos obtienen resultados de utilidad clínica sobre los que basar siguientes investigaciones.

1. Introducción

1.1. Uso de alcohol

Según la OMS, los problemas derivados del consumo de alcohol son uno de los principales factores de riesgo para la salud de la población mundial, siendo un problema de primer nivel presente en los distintos planes de acción, ya que causa el 5,3% de las muertes globales [1]

Esta adicción afecta tanto al consumidor como a su entorno. Entre las personas más cercanas al sujeto aumenta el riesgo de sufrir abusos físicos y sexuales, así como la probabilidad de abandono hacia la descendencia. Para el propio paciente aumenta también el riesgo de bajo rendimiento profesional, de depresión o de suicidio [2].

Según la Encuesta Domiciliaria sobre Abuso de Drogas (EDADES) [3], realizada por el Ministerio de Sanidad en 2017-2018 a 21.249 personas de entre 15 y 64 años, el 91,2% de los encuestados declara haber consumido bebidas alcohólicas a lo largo de su vida. De los encuestados el 62,7% en los últimos 30 días, siendo mayoritariamente hombres, con una media de edad de 40,8 años. Mientras que un 18,6% de los encuestados tuvieron alguna intoxicación etílica aguda en los últimos 12 meses. El alcohol es la droga legal más consumida con una prevalencia mayor del 70%, seguida del tabaco con una prevalencia en torno al 40%.

Existen diferentes tratamientos contra la dependencia al alcohol y entre ellos encontramos tres tipos principales [4]: tratamientos farmacológicos, grupos de apoyo, y tratamiento con personal médico.

La impulsividad es un aspecto clave a la hora de tratar a una persona alcohólica. Dicha impulsividad se puede medir a través de la inhibición, obtenida en la tarea de la señal de stop, Stop Signal Task (SST). En la SST se emplea una secuencia estándar de este tipo de tareas que consiste en presentar una serie de estímulos (visuales generalmente) en los que el individuo tiene que responder. Sin embargo, en un porcentaje reducido de ensayos se presenta de forma simultánea otro estimulo (visual o auditivo), el cual indica al participante que tiene que retener su respuesta en ese ensayo [5]. En este caso se emplea la SST modificada (MSST) de Jurado et al. [6] en donde además de los dos estímulos anteriores se emplea un estímulo relacionado con el alcohol, con un componente afectivo para las personas realizando la prueba. En nuestro caso se emplean pseudopalabras, palabras neutras y palabras relacionadas con el alcohol. Tras la presentación de una palabra se indica a la persona que pulse la tecla "-" y tras una pseudopalabra la tecla "z". El tiempo que transcurre entre la presentación del estímulo y la pulsación de la tecla es el tiempo de reacción a la señal de stop, Stop Signal Reaction Time (SSRT), que se cuantifica con el registro simultáneo del electroencefalograma (EEG).

Durante la prueba se obtienen dos potenciales evocados asociados al evento cerebral producido por un estímulo, Event Related Potential (ERP). Estos eventos se ven alterados en el caso de que el sujeto bajo estudio sea alcohólico o lo haya sido en el pasado, permitiendo medir la impulsividad del sujeto. Los sujetos alcohólicos tienen un peor desempeño en esta tarea que los sujetos sanos.

ISBN: 978-84-09-25491-0, pags. 133- 136

La actividad cerebral sincronizada se ve afectada en pacientes alcohólicos, lo que se puede observar mediante la alteración en las distintas bandas EEG de dicha actividad. Esto produce dificultades a la hora de procesar la información externa o al inhibir los mecanismos de alerta. En el estudio de Martínez-Maldonado et al. [7] se introduce una técnica de modificación del sesgo cognitivo, Cognitive-Bias Modification (CBM) [8]y de proceso de extinción de la recuperación de memoria como terapia para eliminar el comportamiento que induce al consumo. Con ello se consigue modificar la sincronización cerebral en reposo y las tendencias automáticas hacia el alcohol. En este complemento terapéutico se pide al paciente que realice una acción, ya sea, alejar un joystick cuando se le presenta un estímulo asociado al alcohol (estímulo positivo) o una acción de acercar el joystick cuando se presenta un estímulo neutro. Tras realizar este entrenamiento cuatro veces por semana se ha observado que la vuelta a valores no patológicos en personas alcohólicas se alcanza progresivamente.

1.2. Objetivos

El objetivo de este trabajo es el diseño y desarrollo de algoritmos de predicción en pacientes en tratamiento por alcoholismo a partir de datos psicológicos, EEG y de imagen LORETA [9]. En primer lugar, se plantea el desarrollo de un modelo predictor de la escala de intensidad de dependencia al alcohol (EIDA). En este modelo se tendrán en cuenta únicamente los valores previos al tratamiento con CBM. En segundo lugar, se plantea implementar un modelo predictor de recaídas a partir de los datos previos y posteriores a la terapia CBM. Con ello se pretende ampliar el conocimiento existente en la caracterización de pacientes con problemas de alcoholismo y apoyar en la toma de decisiones de los profesionales clínicos en la personalización de las rutas terapéuticas.

2. Materiales y métodos

2.1. Base de datos

Los datos empleados proceden de 87 sujetos voluntarios, de los que 70 son pacientes en abstinencia con tratamientos previos en el hospital y los 17 restantes son sujetos sanos. Uno de los sujetos abandonó el tratamiento tras tomar los primeros datos, contando con información completa de 69 pacientes. Los sujetos se dividen en cuatro grupos: (1) pacientes que reciben un tratamiento CBM aplicado mediante vídeo; (2) pacientes con tratamiento CBM aplicado con realidad virtual (RV); (3) pacientes sin tratamiento; (4) sujetos sanos. De cada sujeto se registra información referente a aspectos demográficos, clínicos, cuestionarios de personalidad, datos de EEG y de LORETA.

Realizando una descripción muestral de la base de datos se observa que de los 87 sujetos bajo estudio 71 (82%) son hombres y 16 (18%) mujeres. La edad está comprendida entre los 29 y los 71 años (media de edad de 51,05±8,91). El promedio de la edad de inicio de consumo es de 15,59±1,41 y la media de inicio de consumo patológico es de 35,84±13,62 años.

Del total de 69 pacientes, 60 (69%) de ellos declaran consumir otras sustancias entre las que se consideran la marihuana (32%), la cocaína (31%), y la heroína (6%). La puntuación EIDA tiene una media de 35,39±16,69. Para los sujetos control la media es de 0,74±1,64. En cuanto a los datos referentes a las recaídas en los meses posteriores al tratamiento de los pacientes con alcoholismo, encontramos 60 pacientes sin recaída (86%) y 9 con recaída (13%).

2.2. Cuestionarios psicológicos y datos cerebrales

Entre los cuestionarios psicológicos realizados a los pacientes se encuentran Hamilton Depression rating scale (HAM D), Alcohol Use Disorders Identification Test (AUDIT), Escala de Intensidad de Dependencia al Alcohol (EIDA), Baratt Impulsivity Scale (BIS-11), la puntuación agregada de las Reestructured Clinical Scales (RC), Cuestionario sobre el Deseo de Beber (CDB), Obsessive Compulsive Drinking Scale (OCDS), Sensibilidad al castigo y a la recompensa, Meta-cognitions Questionnaire (MCQ), Rumiative Responses Scale (RSS), Emotional Regulation Questionnaire (ERQ), Cognitive Emotion Regulation Questionnaire (CERQ) y Difficulties in Emotion Regulation Scale (DERS).

Como variables medidas en la SST se cuenta con el sesgo total, la Modified Stop Signal Task (MSST) y el SSRT para los distintos tipos de estímulos.

Los datos de EEG empleados en este trabajo se recogen simultáneamente a la realización de la MSST de Jurado et al. [6]. Se registran datos de tres electrodos (Fz, Cz y FCZ) en los que se describen los tiempos de latencia y la amplitud para las ondas N100 y P300 tanto antes como después del tratamiento. Las mediciones por electrodo se clasifican en función del tipo de palabra, positiva, neutra o pseudopalabra.

Los datos del análisis LORETA empleados se construyen en torno a la señal de stop identificando las distintas regiones de interés, el tipo de palabra que genera el estímulo y si es pre o post tratamiento.

Para el desarrollo de los modelos predictivos se ha empleado la herramienta gratuita Weka v.3.8 [10].

2.3. Métodos

El Test EIDA es un instrumento eficaz en la evaluación de pacientes con diagnóstico de dependencia al alcohol, y permite medir la intensidad de dicha dependencia mediante una escala numérica. Por este motivo se decidió desarrollar un modelo predictor siguiendo una metodología similar a la presentada en la publicación de O´Halloran et al. [11], para predecir el valor del test AUDIT, una prueba que califica las conductas de consumo.

En el presente trabajo se propone la generación de diferentes modelos regresivos que incluyan diferentes conjuntos de variables entre los disponibles. El algoritmo de regresión empleado será la regresión lineal penalizada. Esta regresión sigue el criterio de Akaike para realizar el modelo predictor y acepta instancias ponderadas. Para el desarrollo de los modelos se mantienen los parámetros por defecto que establece Weka.

Para cada modelo se mide el coeficiente de correlación medio y su desviación estándar, así como el error absoluto y su desviación estándar. La evaluación de los modelos se lleva a cabo mediante validación cruzada de 10 ramas. Para el mejor de los modelos resultantes se analizan los coeficientes β correspondientes al peso de los distintos parámetros en el modelo regresivo, siguiendo el ejemplo de O´Halloran [11].

En total se realizarán 3 conjuntos de modelos, el primero incluye datos de EEG, diferenciando dos submodelos, uno basado en ERP (1-ERP) y otro en LORETA (1-LORETA). El segundo utiliza las pruebas psicológicas, creando en este caso tres submodelos, uno incluyendo todos los campos medidos en dichas pruebas (2-Completo), otro con valores globales intermedios (2-Reducido) y otro en el que se incluyen únicamente los valores globales (2-Mínimo). Por último, en el tercer modelo se emplearán los datos combinados de los modelos anteriores, incluyendo los datos cerebrales, psicológicos, clínicos y demográficos.

En la segunda parte del trabajo se desarrolla un predictor de recaída tres meses después de recibir el tratamiento. Se emplea el modelo Zero Rule como punto de partida para la comparación con los modelos generados, siendo este el modelo más simple que se puede obtener puesto que clasifica a todos los sujetos con el valor de la clase mayoritaria. La evaluación de los modelos se lleva a cabo mediante validación cruzada de 5 ramas. Se ha elegido un valor de 5 ramas para forzar que la validación cruzada sea balanceada y que haya por tanto sujetos con y sin recaída en todos los conjuntos de entrenamiento de todas las ramas, ya que solo 9 pacientes de los disponibles presentaban alguna recaída en el periodo bajo estudio. Para estudiar la significación estadística se consideran p-valores de 0,05 y 0,01. Para la generación de los modelos se estudian cinco algoritmos de diferente tipología, de forma que pueda comprobarse cuál se ajusta mejor a los datos disponibles: Árboles J48 (J48); Máquina de vectores de soporte (SMO); Bosques aleatorios (Random Forest); redes neuronales (Multilayer perceptrón, MLP); y Bayes ingenuo (BI).

Para la generación de estos modelos se considerarán los datos pre y post disponibles en la base de datos. En diversas iteraciones se considerarán diferentes conjuntos de datos en base a los resultados obtenidos y a la información extraíble de los algoritmos con los que se trabajará. Siguiendo la línea de trabajo del predictor de EIDA, se trabajará también con los datos de EEG y LORETA para clasificar las recaídas. Como criterio de selección de algoritmos se tendrán en cuenta los parámetros de precisión, exactitud y AUC.

3. Resultados

3.1. Análisis EIDA

En la Tabla 1 puede verse el desempeño de cada uno de los modelos regresivos generados con los diferentes conjuntos de variables anteriormente mencionados.

De entre los modelos descritos, el de mejor desempeño es el modelo 2-Mínimo. Este modelo obtiene un valor de correlación de 0,82±0,14, y un error de 8,99±2,64. Particularizando en este modelo, si prestamos atención a los coeficientes β obtenidos, los valores con mayor peso y relevancia son, en orden de valor absoluto decreciente, el policonsumo, CERQ_perspectiva, el nivel de estudios, el total de OCDS, HAM A PRE, HAM D PRE y total RC con el mismo valor, el total RSS, la edad de inicio de consumo patológico, el total CDB, el valor de SSRT previa con un estímulo positivo y por último el valor de SSRT previa con un estímulo neutro.

Modelo	Coeficiente de correlación	Error medio absoluto
1-ERP	0,22 (0,35)	18,43 (4,50)
1-LORETA	-0,00 (0,10)	17,96 (4,40)
2-Completo	0,81 (0,16)	8,71 (2,71)
2-Reducido	0,79 (0,16)	10,48 (2,97)
2-Mínimo	0,82 (0,14)	8,99 (2,64)
3-Completo	0,72(0,28)	10,92(7,26)
3-Reducido	0,44(0,30)	25,74(5,49)
3-Mínimo	0,71(0,19)	12,35(3,18)

Tabla 1. *Coeficientes de correlación y error medio absoluto con desviaciones estándar para los modelos regresivos de EIDA generados con los diferentes conjuntos de variables.*

3.2. Clasificador de recaída

En la Tabla 2 pueden verse los mejores resultados obtenidos para el clasificador de recaída. Por motivos de brevedad solo se muestran los modelos con mejores prestaciones.

Tres de los modelos estudiados han obtenido mejoras estadísticamente significativas en lo referente a precisión respecto del modelo Zero Rule y cuatro muestran mejoras en AUC. Ninguno de los modelos ha obtenido una mejora estadísticamente significativa en cuanto a exactitud. Los mejores resultados son los obtenidos por el árbol J48 con todo el conjunto de datos pre y post tratamiento y por la MLP con el modelo 2 Mínimo pre y post tratamiento.

Mod.	Datos	Precisión	Exactitud	AUC
Zero Rule	Todos	0,87 (0,03)	0,87 (0,02)	0,5 (0,0)
J48	Todos-Post	0,92 ^{v} (0,05)	0,86 (0,08)	0,85 $^{v\,w}$ (0,09)
BI	ERP-Post	0,90 (0,07)	0,62 (0,11)	0,60 (0,11)
BI	LORETA-Post	0,93 ^{v} (0,06)	0,73 (0,07)	0,76 $^{v\,w}$ (0,08)
MLP	2 Mínimo-Post	0,93 $^{v\,w}$ (0,05)	0,86 (0,07)	0,85 $^{v\,w}$ (0,11)
J48	3 Mínimo-Post	0,90 (0,05)	0,84 (0,09)	0,81 $^{v\,w}$ (0,11)

Tabla 2. *Valores de precisión, exactitud y AUC para el modelo Zero Rule y los modelos generados. La mejora con significación estadística respecto del modelo Zero Rule se muestra con: v para p<0,05 o w para p<0,01.*

4. Discusión

4.1. Análisis EIDA

Los resultados obtenidos muestran que los submodelos del conjunto 1 creados a partir de los datos de EEG y de LORETA no tienen capacidad para predecir intensidad de dependencia al alcohol (EIDA). Ambos se caracterizan por bajos coeficientes de correlación y altos valores de error medio absoluto. Esto parece confirmar lo observado por O´Halloran [10], dado que en dicho estudio tampoco obtuvieron buenos resultados a partir de los datos de EEG.

Una limitación al utilizar variables de registro cerebral es la alta dimensionalidad de los parámetros obtenidos por cada paciente frente a la pequeña muestra poblacional empleada. Sería necesario investigar si un tratamiento previo de los datos que permitiera reducir la dimensionalidad de estos podría suponer un cambio en su utilidad predictiva.

Los resultados más relevantes se han obtenido al considerar los datos provenientes de las pruebas psicológicas. Los campos con los valores β más destacados ayudan a entender qué comportamientos favorecen más la gravedad de la dependencia. El policonsumo, la depresión o la ansiedad, por tanto, son factores claves a tratar para reducir la adicción.

4.2. Clasificador de recaída

La base de datos con la que se ha trabajado se encuentra desbalanceada contando con 9 casos de recaída frente a 60 de no recaída. Esta situación dificulta la obtención de modelos predictivos.

Los dos modelos con mejores prestaciones, el J48 Todos-Post y el MLP 2 Mínimo-Post, obtienen unos parámetros muy similares en lo referente a precisión e idénticos en cuanto a AUC. De igual forma, ambos emplean tanto los datos previos como posteriores al tratamiento, lo que significa que la terapia no solo ayudaría a los pacientes a mantener la abstinencia, si no que permitiría predecir con mayor precisión las posibles futuras recaídas. Esto sería de gran utilidad en la práctica clínica, puesto que permitiría seleccionar aquellos pacientes con mayores probabilidades de recaída y realizar un seguimiento más intensivo en dichos casos.

La principal diferencia entre ambos modelos es la capacidad de interpretación de los resultados, algo de gran importancia en la práctica clínica. Los árboles de decisión son modelos mucho más fácilmente interpretables mientras que las redes neuronales se caracterizan por su complejidad.

Si hubiera que elegir un modelo en base a los parámetros obtenidos, este sería el obtenido por el modelo MLP con el conjunto de datos 2 Mínimo pre y post tratamiento, el cual obtiene una mejora estadísticamente significativa con $p<0,01$ en ambos parámetros. Además, el conjunto de variables empleado para la generación de este modelo es el de menor volumen de todos los empleados, lo cual disminuye la complejidad del mismo. Sin embargo, el modelo generado por el árbol J48 permite una visualización más clara y detallada de qué variables influyen en la toma de decisiones y en qué medida lo hacen, obteniendo en este caso también una diferencia estadísticamente significativa con $p<0,01$ para el AUC y con $p<0,05$ para la precisión. Desde el punto de vista clínico seguramente este último sería la opción elegida.

5. Conclusiones

Este trabajo obtiene resultados prometedores a la hora de apoyar el tratamiento de adicción al alcohol y ampliar el conocimiento existente mediante técnicas de aprendizaje de máquina. Los datos provenientes de los cuestionarios psicológicos han demostrado ser los de mayor utilidad en ambos análisis realizados, quedando abierta la posible necesidad de un tratamiento previo de los datos de registros cerebrales para obtener un mayor rendimiento en los resultados.

Referencias

[1] Organización Mundial de la Salud. (2018). Informe sobre la situación mundial del alcohol y la salud.

[2] Página de MedlinePlus https://medlineplus.gov/spanish/ency/patientinstructions/000494.htm (Consultada: mayo 2020).

[3] Blas, L., Nro, T. Y., & Dni, D. Informe 2019, *Observatorio español de las drogas y las adicciones*, 2019, pp 34–36.

[4] Página web del National Institute on Alcohol Abuse and Alcoholism. https://pubs.niaaa.nih.gov/publications/treatmentSpanish/treatmentSpanish.pdf%0Ahttp://pubs.niaaa.nih.gov/publications/treatmentSpanish/treatmentSpanish.pdf (Consultada: mayo 2020).

[5] Logan, G. D. On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory and language*, 1994, pp 189–239 (ISBN 9780122004100)

[6] Sion. A, Jurado-Barba, R, Alonso, M. J. Evaluación de la capacidad de inhibición en dependientes de alcohol: traducción de la tarea de la señal de stop modificada. *Actas españolas de psiquiatría*, vol 45, sup 1, 2017, pp 21-31.

[7] Martínez-Maldonado, A., Jurado-Barba, R., Sion. Brain functional connectivity after cognitive-bias modification and behavioral changes in abstinent alcohol-use disorder patients. *International Journal of Psychophysiology*, vol 154, 2020, pp 46-58. (ISSN: 0167-8760).

[8] Eberl, C., Rinck, M., Becker, E.S., Lindenmeyer, J. Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychol. Sci.* vol 22, iss 4 2011, pp 490-497. (ISSN: 0956-7976)

[9] Pascual-Marqui, R. D., Esslen. (2002). Functional imaging with low-resolution brain electromagnetic tomography (LORETA): A review. *Methods and Findings in Experimental and Clinical Pharmacology*, vol 24, sup C, 2002, pp 91–95. (ISSN: 0379-0355).

[10] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter*, vol 11, sup 1, 2009, pp 10–18. (ISSN: 1931-0145).

[11] O'Halloran, L., Rueda-Delgado. Inhibitory-control event-related potentials correlate with individual differences in alcohol use. *Addiction Biology*, vol 25, sup 2, 2020. (ISSN: 1355-6215).

Clasificación de Bloqueadores de Ikr Basada en Protocolos de Voltage Clamp y Técnicas de Machine Learning.

Fernando Escobar[1], Julio Gomis-Tena[1], Javier Saiz[1], Lucía Romero[1]

[1]Centro de Investigación e Innovación en Bioingeniería, Universitat Politècnica de València, Valencia, Spain

Resumen

La evaluación de la cardiotoxicidad es clave en el desarrollo de nuevos compuestos y se suele realizar analizando el grado de bloqueo que provocan en los canales de potasio relacionados con el gen humano ether-à-go-go (hERG). El objetivo de este trabajo es desarrollar un clasificador que determine la preferencia de un fármaco para unirse a un estado determinado del canal.

Se ha creado un conjunto de 2600, divididos en 13 clases, bloqueadores virtuales con diferentes afinidades y cinéticas por los estados conformacionales del canal. Se realizaron simulaciones utilizando tres protocolos de estimulación que aumentan la probabilidad del canal de ocupar ciertos estados. Para cada simulación se han tomado tres medidas: IC_{50}, la constante de recuperación de la corriente de potasio I_{Kr} y una estimación del tiempo necesario para alcanzar el estado estacionario. Por lo tanto, se obtuvieron 9 variables por cada fármaco estudiado. Se desarrolló, entrenó y evaluó un clasificador en dos pasos. En primer lugar, se usaron máquinas de soporte vectorial sobre las IC_{50} para separar las 13 clases en tres grupos con 4, 5 y 4 clases respectivamente. En segundo lugar, se usaron redes neuronales en cada grupo para terminar de clasificar los bloqueadores.

Los tres clasificadores obtuvieron una precisión global en los grupos de test del 90.83, 88.66 y 89.16% respectivamente.

1. Introducción

La evaluación preclínica de la cardiotoxicidad de nuevos fármacos se ha convertido en una prioridad para las compañías farmacéuticas debido a potenciales efectos adversos como arritmias ventriculares [1]. Hoy en día, la forma más común de evaluar la seguridad de un fármaco se basa en medir el bloqueo in vitro que producen en el gen humano ehter-à-go-go (hERG) y la prolongación in vivo del intervalo QT. Ambos fenómenos han sido relacionados con la aparición de Torsades de Pointes (TdP).

Aunque estos test de seguridad han tenido éxito en evitar que fármacos potencialmente peligrosos llegasen al mercado, también han evitado el desarrollo de compuestos con potencial de ser útiles. De hecho, existen fármacos como el verapamil, reconocido como un potente bloqueador de la corriente I_{Kr}, que no está asociado al desarrollo de TdP.

Estudios recientes han demostrado que la evaluación de la seguridad de fármacos mejora cuando se tienen en cuenta las dinámicas y cinéticas de los mismos [2]. Las diferencias en la unión de un compuesto al canal dependiendo del estado conformacional en el que se encuentre el canal pueden alterar significativamente los valores de la IC_{50}, que es la concentración de fármaco a la cual se produce la mitad del bloqueo de una corriente iónica. A su vez, dichos valores también pueden variar dependiendo del protocolo de estimulación utilizado [3]. Por lo tanto, considerar el comportamiento del fármaco y estandarizar los protocolos empleados es necesario para mejorar la evaluación de los mismos. Un estudio reciente ha propuesto un conjunto de tres protocolos que maximizan la probabilidad del canal I_{Kr} de encontrarse en cada uno de los estados posibles [4]. En este trabajo se utilizaran dichos protocolos con el objetivo de determinar el estado de unión preferente haciendo uso de técnicas de *machine learning*.

2. Materiales y métodos

2.1. Modelos

Se han utilizado seis variantes del modelo de Markov de I_{Kr} de ventrículo humano de Fink et al. [5] para simular la interacción fármaco-canal, modelos mostrados en la figura 1. El modelo sin fármaco original tiene cinco estados conformacionales, tres cerrados (C_1, C_2, y C_3), uno abierto (O) y uno inactivado (I). Se incorporaron otros cinco estados para representar los estados del canal a los que el fármaco está unido (C_{1d}, C_{2d}, C_{3d}, O_d and I_d). Las transiciones entre los estados con y sin fármaco están regulados por una constante denominada $k_{(C, O, I)}$ que se multiplica por la concentración de fármaco para obtener la velocidad de transición. La liberación de un compuesto desde un estado está marcada por la velocidad de transición denominada $r_{(C, O, I)}$.

2.2. Protocolos

Tres protocolos de *voltage clamp* (mostrados en la figura 2) previamente desarrollados se utilizaron para simular los efectos de los fármacos [4]. Dichos protocolos, denominados P40, P0 y P-80, maximizan el tiempo que el canal permanece a 40, 0 y -80 mV

ISBN: 978-84-09-25491-0, pags. 137- 140

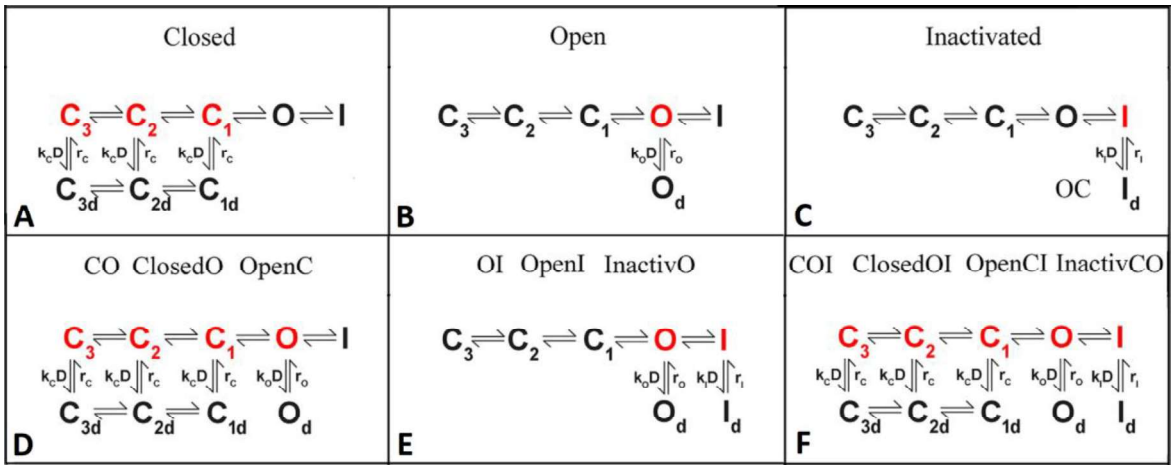


Figura 1: Modelos de Markov con estados libres (A, B y C) y con fármaco (D, E y F). D es la concentración de fármaco, k_C, ko y k_I se multiplican por D para calcular la velocidad de asociación desde los estados cerrados, abiertos e inactivados libres a los ocupados respectivamente. r_C, r_o y r_I son las velocidades de disociación del compuesto. Las letras rojas indican estados a los que se une el fármaco. Los nombres de las clases parecen con el modelo de Markov correspondiente.

respectivamente. P40 consiste en un paso de acondicionamiento de 5 s a 40 mV seguido de un paso de test de 0.2 s a -60 mV y 0.2 s de recuperación a -80 mV. P0 es similar a P40 pero el paso de acondicionamiento se produce a 0 mV. P-80 consiste en un test de 0.5 s a 20 mV, seguido de 0.2 s a -50 mV y un paso de acondicionamiento de 4.5 s a -80 mV. Los protocolos aumentan la posibilidad del canal de encontrarse en el estado inactivado, abierto y cerrado, respectivamente [6].

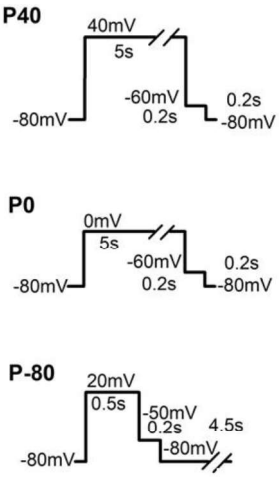


Figura 2: Protocolos de voltage clamp utilizados para similar los efectos de fármacos.

2.3. Fármacos

Se creó un conjunto de 2600 fármacos virtuales con diferentes dinámicas y cinéticas por lo estados conformacionales del canal. Los nombres de las diferentes clases de compuesto aparecen con el modelo de Markov correspondiente en la figura 1 y se definen como: Fármacos que se unen exclusivamente al estado cerrado (Closed), abierto (Open) e inactivado (Inactivated). Fármacos que se unen simultáneamente al estado cerrado y al abierto, con igual afinidad (CO), con preferencia por el cerrado (ClosedO) o por el abierto (OpenC). Compuestos que se unen con igual afinidad a los estados abierto e inactivado (OI), con preferencia por el abierto (OpenI) y con preferencia por el inactivado (InactivO). Finalmente, fármacos que se unen a los tres estados a la vez, de igual manera (COI), con mayor afinidad por el cerrado (ClosedOI), con preferencia por el abierto (OpenCI) y con preferencia por el inactivado (InactivCO). En todos los casos, la velocidad de asociación es igual para todos los estados bloqueados. La mayor afinidad por un determinado estado se simula multiplicando por 100 la constante de disociación del estado correspondiente. Se simularon 200 fármacos de cada clase, cuyas contantes k y r fueron generadas aleatoriamente dentro de un rango de valores fisiológicos usando Matlab (Mathoworks Inc.).

3. Resultados

Los efectos de los fármacos virtuales se simularon usando los tres protocolos mencionados anteriormente. Para cada simulación se obtuvieron tres variables por cada protocolo, 1) IC_{50}, el parámetro más utilizado para evaluar la seguridad cardíaca, 2) un estimación basada en derivadas del número de pulsos requeridos para alcanzar el

estado estacionario y 3) la constante de recuperación de la evolución de la inhibición de la corriente de potasio a la concentración de la IC_{50} estimada mediante un ajuste exponencial. Por lo tanto, se obtuvieron 9 variables por compuesto.

Tras obtener todas las variables se desarrolló un clasificador en dos pasos para intentar clasificar los fármacos en las trece clases objetivo (mostradas en la figura 1). Todo el análisis de datos se realizó con Matlab.

3.1. Máquinas de soporte vectorial

Se realizó un análisis exploratorio de los datos para visualizar las variables obtenidas. Se construyeron gráficos 3D donde se mostraron los valores de cada variable asignando cada protocolo a un eje. La representación 3D de la IC_{50} se muestra en la figura 3 (gráficos del resto de variables no mostrados).

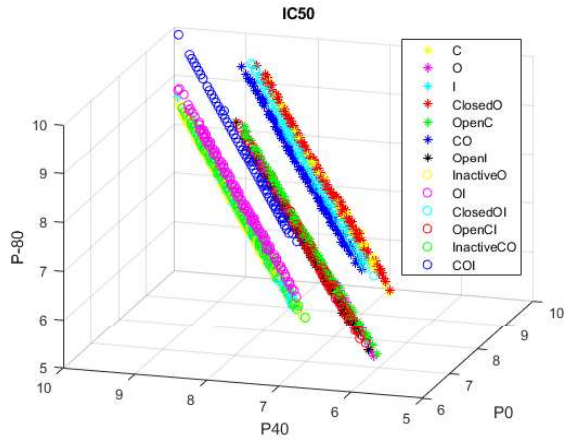


Figura 3: representación 3D de los valores de IC_{50} de los 2600 bloqueadores virtuales. Los ejes representan el –log([IC_{50}]) obtenido con los tres protocolos (P40, P0 y P-80).

Como se observa en la figura 3, existen tres grupos claramente diferenciados entre los fármacos, separables mediante un hiperplano. Por este motivo se decidió utilizar máquinas de soporte vectorial (SVM) para crear las separaciones entre grupos. Por lo tanto, se entrenaron dos SVMs utilizando un enfoque de uno contra todos. El primer SVM separaba el grupo superior, considerando todos los compuestos presentes como una única clase (Figura 4, panel superior). El segundo SVM separaba el grupo inferior del resto, de nuevo considerando como una todas las clases presentes en el mismo (Figura 4, panel inferior). Todos aquellos compuestos que no pertenecieran a ninguno de los dos grupos anteriores se asignaron al grupo central.

Este primer paso separaba los fármacos en tres grupos utilizando exclusivamente los valores de IC_{50}. El primer grupo se corresponde con fármacos con mayor afinidad por el estado cerrado y los que bloqueaban de igual manera el cerrado y el abierto (Closed, ClosedO, ClosedOI, CO). El grupo central contenía los compuestos con preferencia por el estado abierto y con igual afinidad por los tres estados a la vez (Open, OpenC, OpenI, OpenCI, COI). Finalmente, el grupo inferior agrupaba los compuestos con mayor afinidad por el estado inactivado y los que bloqueaban simultáneamente y de igual manera el abierto e inactivado (Inactivated, InactivO, InactivCO, OI)

Por lo tanto, haciendo uso solamente de la IC_{50} obtenida para cada protocolo se pueden separar los compuestos en base a su estado preferente.

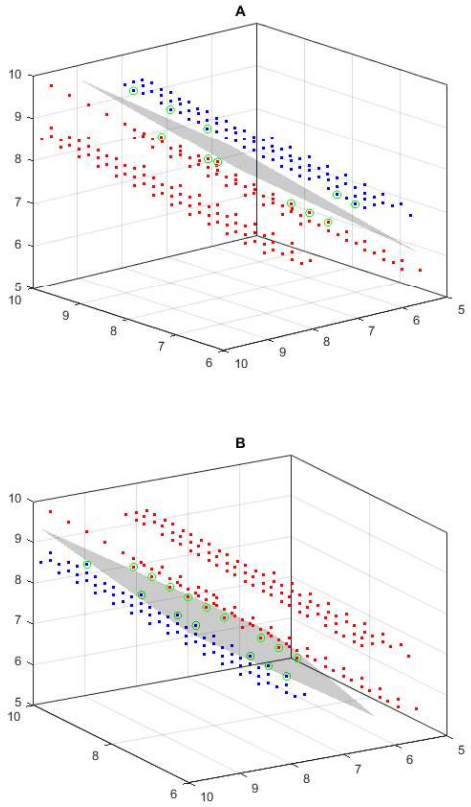


Figura 4: Hiperplanos creados mediante los SVMs para separar los compuestos en tres grupos por los valores de IC_{50} obtenidos con cada protocolo. (A) SVM que separa el grupo superior (preferencia por el estado cerrado). (B) SVM para separar el grupo inferior (preferencia por el estado inactivado). Los puntos azules representan fármacos que pertenecen al grupo separado y los rojos aquellos que no.

3.2. Redes Neuronales

Tras el primer paso del proceso de clasificación, se entrenó y evaluó un conjunto de tres redes neuronales utilizando la *toolbox* de Matlab. Las redes se usaron para clasificar los fármacos de cada uno de los grupos en las trece clases objetivo. Los grupos de entrenamiento y test fueron seleccionados automáticamente por la *toolbox*. 120,

126 y 120 fármacos fueron seleccionado para cada uno de los grupos de test. Para este paso se consideraron las nueve variables obtenidas. Las redes tienen una capa de entrada con 9 nodos (tantos como variables usadas), una capa oculta y una capa de salida con 3, 4 y 3 nodos respectivamente. El número de neuronas en la capa oculta se seleccionó evaluando un rango de 1 a 18 (el doble de las neuronas de la capa de entrada) y seleccionando el mejor resultado [7]. Para cada grupo, el mejor resultado se obtenía con 11, 7 y 9 neuronas respectivamente. Los resultados para los grupos de test se muestran en la figura 5. La figura muestra una precisión total de 90.83, 88.66 y 89.16% para los grupos de test.

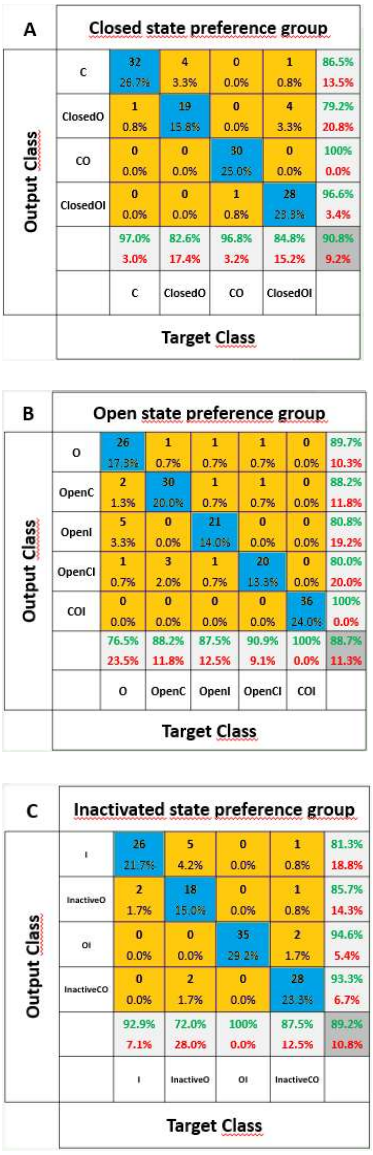


Figura 5: Matrices de confusión de los grupos de test para las redes neuronales. El número superior en las celdas representa el número absoluto de fármacos y el inferior el porcentaje relativo. (A) grupo de preferencia por el estado cerrado. (B) grupo de preferencia por el estado abierto. (C) grupo de preferencia por el estado inactivado. Las filas se corresponden con las clases predichas y las columnas con las clases reales. Celdas azules muestran los fármacos clasificados correctamente y las celdas naranjas los incorrectamente clasificados. La columna de la derecha muestra el valor predictivo positivo (verde) y ratio de falsos (rojo). La fila de abajo muestra el ratio de verdaderos positivos (verde) y el ratio de falsos negativos (rojo). La celda gris oscura muestra la precisión total.

4. Conclusión

Hemos desarrollado una herramienta útil para la clasificación de fármacos de acuerdo a sus afinidades y basándonos en medidas obtenidas de protocolos de *voltage clamp* y con elevada precisión

5. Referencias

[1] Hammond, T. G.; Carlsson, L.; Davis, A. S.; Lynch, W. G.;MacKenzie, I.; Redfern, W. S.; Sullivan, A. T.; Camm, A. J. Methods of Collecting and Evaluating Non-Clinical Cardiac Electrophysiology Data in the Pharmaceutical Industry: Results of an International Survey. Cardiovasc. Res. 2001; 49(4): 741−750.

[2] Noble, D. Computational Models of the Heart and Their Use in Assessing the Actions of Drugs. J. Pharmacol. Sci. (Tokyo, Jpn.). 2008; 107(2): 107−117

[3] Lee, W.; Windley, M. J.; Perry, M. D.; Vandenberg, J. I.; Hill, A. P. Protocol-Dependent Differences in IC50 Values Measured in Human Ether-Á-Go-Go-Related Gene Assays Occur in a Predictable Way and Can Be Used to Quantify State Preference of Drug Binding. Mol. Pharmacol. 2019; 95: 537−550.

[4] Gomis-Tena J. When Does the IC50 Accurately Assess the Blocking Potency of a Drug?. Journal of Chemical Information and Modeling. 2020; 60(3): 1779-1790.

[5] Fink M, Noble D, Virag L, Varro A, Giles W. Contributions of HERG K+ current to repolarization of the human ventricular action potential. Progress in Biophysics and Molecular Biology. 2008; 96(1-3): 357-376.

[6] Milnes, J. T.; Witchel, H. J.; Leaney, J. L.; Leishman, D. J.; Hancox, J. C. Investigating dynamic protocol-dependence of hERG potassium channel inhibition at 37 °C: Cisapride versus dofetilide. J. Pharmacol. Toxicol. Methods. 2010; 61: 178−191.

[7] Heaton J. The Number of Hidden Layers [Internet]. 2017 [cited 26 August 2020]. Available from: https://www.heatonresearch.com/2017/06/01/hidden-layers.html

Arrhythmogenicity quantification of two genetic defects affecting I_{Kr} channel in AF patients

R. Belletti[1], L. Martinez Mateu[2], L. Romero Perez[1], E. Cherry[3], F. H. Fenton[4], J. Saiz[1]

[1] Centro de Investigación e Innovación en Bioingeniería, Universitat Politecnica de Valencia, Valencia, España, rbelletti@ci2b.upv.es, jsaiz@ci2b.upv.es

[2] Departamento de Teoría de la Señal y las Comunicaciones y Sistemas Telemáticos y Computación, Universidad Rey Juan Carlos, Madrid, Spain, laura.martinez.mateu@urjc.es

[3] School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA, elizabeth.cherry@gatech.edu

[4] School of Physics, Georgia Institute of Technology, Atlanta, GA, USA, flavio.fenton@physics.gatech.edu

Abstract

Atrial fibrillation (AF) is the most common cardiac arrhythmia characterized by disorganized electrical activations of the upper chambers of the heart, leading to uncoordinated contraction and compromising the pumping action of the organ. AF risk factors include cardiovascular pathologies, endocrine disorders, advanced age, obesity, smoking and heritability. Genetic mutations affecting gene encoding for ion channel protein structures are in fact considered as cause of fibrillatory events in individuals who do not present any other co-morbidities.

In this work, two genetic mutations found in literature and affecting the alpha-subunit of the rapid delayed rectifier potassium channel are modelled, by reparametrizing the I_{Kr} current formulation and by fitting it to mutant experimental data. The modified potassium current was then incorporated into the Courtemanche-Ramirez-Nattel (CRN) model and single cell simulations have been performed to study the mutations' effects on action potential and current traces, as well as, restitution properties, in right and left atrium. Both mutations produced a shortening of the action potential duration at 90% of repolarization (APD_{90}), a higher current peak and lower APD values in the restitution curves. T895M yielded also to a reduction in the maximum slope of restitution curve.

Tissue patch simulations revealed that T895M and T436M provide a substrate to initiate and maintain re-entries during the 5 seconds of simulation. Rotor's meandering in T895M appears more stable with a less extended area and more regular pattern than in T436M. Investigation on 3D atria and torso models will be necessary to provide further insights in understanding the mechanisms behind these genetic mutations.

1. Introduction

Defects in genes encoding for ion channel protein structures have been studied extensively in the past few decades because of their potential role in arrhythmogenic events genesis in otherwise healthy individuals. Onset of atrial fibrillation in patient not affected by cardiovascular disorders or metabolic conditions has been in fact implicated to the presence of genetic mutations, which cause abnormal current conduction through the pore, thus impairing the delicate balance of inward and outward currents through the cell membrane. Genetic predispositions and heritability of AF have been thus added to the long list of common risk factors and in particular this conditions has been named 'Lone AF', accounting for 30% of current AF cases [1][2]. Even though it is not a life threatening disease, AF worldwide affects 1.5 - 2% of general population, leading to a poor quality of life and becoming a growing issue for the healthcare system [3]. In order to better understand the functioning of this pathology, in this study we focus on the modelling and simulation of the effects of two missense gain-of-function mutations, T895M and T436M, found in literature in a previous work by [4]. The two mutations under study were detected in two male patients affected by Lone AF, developed then into paroxysmal and chronic AF over years. The defects affected the KCNH2 gene, which encodes for the α-subunit of the rapid delayed rectifier potassium channels, Kv11.1, underlying the I_{Kr} current. Channels in presence of the mutations result in conducting larger amount of current with respect to the wild-type (WT) case and in a slower deactivation. The simulation study have been performed in order to assess the effects of the two mutations and their potential risk to become a substrate for initiation and maintenance of AF. Single cell and tissue patch simulations have been executed using the human atrial model by Courtemanche-Ramirez-Nattel (CRN) [5]. Results showed how apparently similar mutations affect the system in a different manner, pointing out the importance of a patient-specific approach, especially in the designing of a drug therapy, at treatment stage.

2. Material and Methods

The human atrial single cell model proposed by Courtemanche and colleagues in [5] has been used to reproduce both healthy and mutated behaviour of myocytes.

	RA	LA
g_{CaL}	1	0.9
g_{Kr}	1	2
g_{KACh}	1	1

Table 1. *Factors applied to currents' conductance to account for atrial heterogeneities*

Acetylcholine current (with [ACh] = 0.005 μM) has been introduced since its important role in atrial

ISBN: 978-84-09-25491-0, pags. 141- 144

electrophysiology, in particular in the myocytes' repolarization phase [6]. Atrial heterogeneities for right atrium (RA) and left atrium (LA) have been accounted by varying some currents' conductance as shown in ***Table 1***.

2.1. Potassium Current Parameterization

The rapid delayed rectifier potassium current formulation has been reparametrized to integrate mutant experimental data in the dynamics of the channel. Parameters to be applied in I_{Kr} equations were evaluated and chosen through a sensitivity analysis. This process allowed to reach a final parameters' configuration shown in (1) to (3), which resulted to be the one with the highest influence on the curves to be fitted.

$$I_{Kr} = \frac{P_1 \cdot g_{Kr} \cdot C_m \cdot x_R \cdot (V - E_K)}{1 + e^{\frac{V+15}{22.4}}} \quad (1)$$

$$\alpha_{x(r)} = 0.0003 \cdot \frac{(V+14.1)}{1 - e^{\frac{V+14.1}{-5}}} \quad (2)$$

$$\beta_{x(r)} = P_2 \cdot 7.3898 \cdot 10^{-5} \cdot \frac{(V - 3.3328 + P_3)}{e^{\frac{V-3.3328+P_3}{5.1237 \cdot P_4}} - 1} \quad (3)$$

Parameter tuning has been done accounting for relative changes within the WT and the experimental data of the mutations, which were digitalized along with their experimental ranges as reported in the original study.

2.2. Optimization Algorithm

Parameter estimation has been achieved with the use of a genetic algorithm, implemented in Python, with an initial population size of 100 individuals and a number of generations equal to 100. The most suitable method for selection, crossover and mutation has been chosen in order to find the best solution in this specific problem. Regarding the selection, the 'Roulette Select' method has been preferred because of its high searching power in the local parameter space, while as for regards the crossover, 'Mate' method has been employed. The mutation method used a Gaussian mutation operator with a mutation rate of 0.9. The cost function to be minimized has been defined as a weighted root-mean-square error, between simulated and experimental data. Elitism of individuals in every generation is introduced, allowing to keep unchanged the 10% of the most fitted individuals from generation to generation. The choice of such optimization algorithm is motivated by its high ability to spread in the entire parameter space, with a low probability of get stuck in a local minima.

2.3. Single Cell Simulation

CRN model incorporating the ACh current and mutations' dynamics in the I_{Kr} channel is then used to run single cell simulations applying a 20 pA/pF amplitude current stimulus for 2 ms at different basic cycle lengths (BCLs), both in RA and LA. After 60 pulses of stabilization, action potential (AP) and current traces properties at BCL 1000 ms have been studied and compared with WT case. Restitution curves (APDr) have been presented as APD_{90} respect to diastolic interval (DI), in order to study the dependency to pacing frequency, in presence of genetic mutations.

2.4. Tissue Simulation

Two-dimensional model of a tissue layer of dimension 5 x 5 cm^2 and 300 µm spatial resolution incorporated the modified version of CRN model reproducing the mutations' behaviour. The monodomain formalism has been solved using the Elvira software [7]. A 2 s stimulus current of 100 pA/pF amplitude has been applied on the tissue patch mesh, with a longitudinal conductivity of 0.0022 S/cm·pF and an anisotropy ratio equal to 0.35. Ten planar pulses at BCL 1000 ms were paced in order to stabilize the tissue model, then S1-S2 cross-field protocol has been used to study vulnerability of the tissue to re-entrant activity. Stimulus S2 has been applied at the left bottom corner of the tissue in a rectangular-like shape area in 5-s simulations, a sufficient time to get stable re entries in a simple tissue geometry. For those cases in which re-entries maintained up to the simulation time, spectral and phase analysis of the signals have been performed in order to compute dominant frequencies (DF) and to track rotors tips' trajectory. DF were computed as the highest peak in the power spectrum of the signal and tracking of rotors have been achieved computing singularity points from phase maps built on the simulated signal with the use of the Hilbert transform.

3. Results

3.1. Fitting of Experimental Data

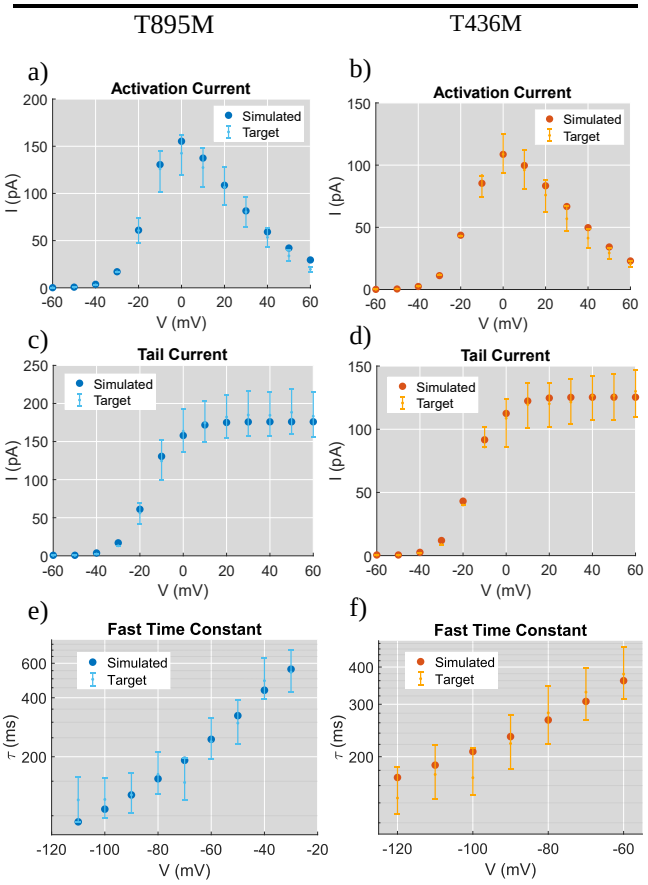


Figure 2. *Fitting procedure results. Comparison between experiments and simulations for* ***a)*** *and* ***b)*** *activation currents,* ***c)*** *and* ***d)*** *tail currents,* ***e)*** *and* ***f)*** *fast time constants, in T895M (blue) and T436M (orange) cases.*

The genetic algorithm successfully retrieve a set of parameter values realizing the adjustments to experimental data. Parameters that fit well the mutation datasets are shown in ***Table 2***.

T895M		T436M	
P_1	2.028	P_1	1.402
P_2	2.105	P_2	0.734
P_3	45.055	P_3	15.19
P_4	2.982	P_4	3.34

Table 2. *Parameters' values for T895M and T436M mutations*

Activation and tail currents, as well as time constants of the mutations, expressed with their experimental ranges, could be accurately adjusted to simulation data, as in ***Figure 2.***

3.2. Single Cell Simulation

Single cell simulations have been performed introducing the parameters shown in ***Table 2.*** in the formulation of the potassium current of the CRN model, along with factors accounting for atrial heterogeneities in ***Table 1.*** and ACh current formulation. Action potential and I_{Kr} current traces resulting from mutated models are shown and restitution properties are studied. Overall, the systems affected by the mutations presented a higher and earlier potassium current peak, which led to an action potential shortening. APDr were shifted to lower APD_{90} values in both cases, maintaining a rate-dependant behaviour.

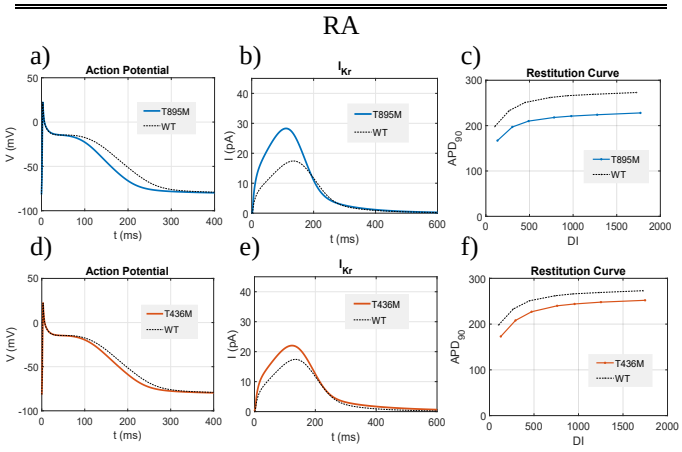


Figure 3. *Action potentials, I_{Kr} current and APDr in RA in presence of T895M (upper panel – solid blue lines) and T436M (lower panel – solid orange lines) compared to WT (dashed line).*

In right atrium, as shown in the upper panel of ***Figure 3.***, the mutation T895M yielded a reduction in the APD_{90} of 17% with respect to the control value, as a direct consequence to a 62% higher peak of the I_{Kr} current; moreover, potassium current time to peak is reduced by 24 ms. APDr in presence of T895M mutation appears flattened, with a 36% reduction in the maximum slope. In ***Figure 3.*** lower panel, results of the effects of T436M mutation in right atrium are depicted. APD_{90} shortens of about 8% the WT value, due to a current peak 26% higher and 12 ms earlier. Restitution curves maximum slope values did not show any difference respect to control case, but APD values were lowered. Results of the effects of the two mutations in left atrium single cell simulations are shown in ***Figure 4.***, where T895M and T436M are compared to control case. In case of T895M, APD is reduced by 23% the control value with a 60% higher current peak and a time to peak 25 ms earlier. APDr shifted to lower APD values and flattened its trend, reducing the maximum slope value of 31% with respect to WT case. The mutation T436M in LA yielded to an APD_{90} shortening of 12%, due to an increase in the potassium current peak of about 26%. As in RA case, APDr did not flattened and no changes in the maximum slope are reported.

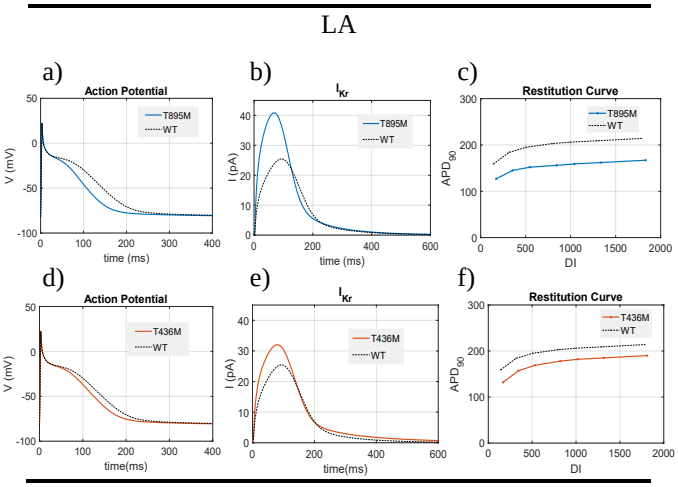


Figure 4. *Action potentials, I_{Kr} current and APDr in LA in presence of T895M (upper panel – solid blue lines) and T436M (lower panel – solid orange lines) compared to WT (dashed line).*

3.3. Tissue Patch Simulation

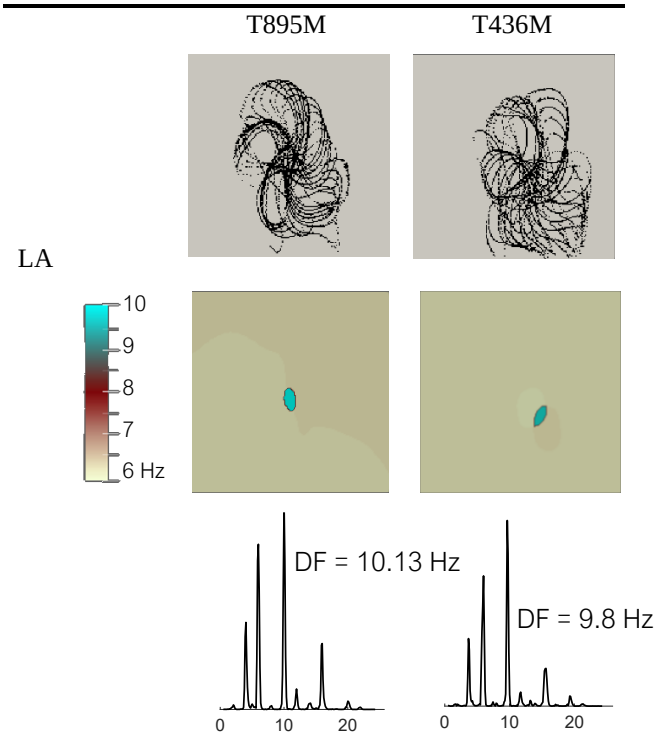


Table 3. *Upper panel: rotors' tip trajectory color-coded by ms of simulation. Middle panel: spatial distribution of high DF in the tissue. Lower panel: power spectrum of the two signals*

Temporal vulnerability of the tissue model to re entrant activities in presence of the two mutation studied revealed

that both of them were able to initiate a rotor, both in right and left atrium. However, not all the cases taken into consideration were able to maintain the spiral wave till the end of the simulation, which in this work has been set at 5 s. Re-entrant activities perpetuating till the end of the simulation have been observed in presence of both mutation in LA. Spectral analysis has been performed and dominant frequency values were equal to 9.8 Hz in T895M case and 10.13 Hz in T436M case. Information about the DF spatial distribution indicates that higher DF areas coincide with the central areas of rotors' meandering as shown in ***Table 3.***'s middle panel. Phase analysis provided information about temporal distribution of singularity points, allowing the tracking of the rotors' tip trajectory during the 5s-simulation. Rotor's meandering area is less extended in presence of T895M rather than in presence of T436M and, as it can be appreciated in the upper panel of ***Table 3.***, the pattern seems quite regular in T895M condition, if compared with T436M case.

4. Discussion and Conclusions

The simulation study presented in this paper proposed to characterize the effects of two gain-of-functions genetic mutations, T895M and T436M, found in patients affected by AF. With this aim, a parameter estimation procedure has been performed to incorporate mutant experimental datasets into a well-established human atrial myocyte model. Using an adaptive heuristic search method, the genetic algorithm, mutation data were successfully fitted and pathological channel dynamics were reproduced. Single cell and tissue patch simulations have been thus carried out in order to further understand the mechanisms behind these pathological conditions. 0D simulations revealed both genetic mutations led to an APD_{90} reduction, an increased I_{Kr} peak and to a shift to low APD values in the APDr. These effects were more pronounced in presence of the mutation T895M, rather than in presence of the mutation T436M; in particular, the former defects was also responsible for an APDr flattening, with a reduction in the maximum slope value, while the latter did not affect it at all. Both APD_{90} shortening and resistance to high frequency pacing excitations have been demonstrated to be related to the generation of a substrate more prone to re entrant activity initiation and progression [8]. Simulations on a two-dimensional tissue model confirmed this assumption. In presence of both mutations, spiral waves had been induced, but only in left atrium they could be maintained for the whole simulation time. The study of singularity points allowed to keep track of rotor's tip trajectory which were more regular in shape and less extended in area in T895M case, with respect to T436M case. Dominant frequencies are spatially organized in the areas of the meandering rotors' centres and DF values lied in the range of values reported in literature for subject suffering from paroxysmal and permanent AF [9]. Further investigation of the system affected by these pathological conditions will be necessary in order to better assess and characterize the effects on whole – atrial and torso geometries. The importance of an accurate modelling and study of these rare mutations are highlighted by the fact that even if they affect the same channel portion, different mutations can lead to different dynamical impairments, which will require different therapeutic strategies, including a patient-specific design of targeted drug therapies.

Acknowledgements

I am grateful to the teams at GaTech, who helped me in the realization of this work. This project has been funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie gran agreement No.766082 and by Generalitat Valenciana, Prometeo programme 2020/043.

References

[1] A. A. Y. Ragab, G. D. S. Sitorus, B. B. J. J. M. Brundel, and N. M. S. de Groot, "The Genetic Puzzle of Familial Atrial Fibrillation," *Front. Cardiovasc. Med.*, vol. 7, no. February, pp. 1–8, 2020, doi: 10.3389/fcvm.2020.00014.

[2] I. E. Christophersen and P. T. Ellinor, "Genetics of atrial fibrillation: From families to genomes," *J. Hum. Genet.*, vol. 61, no. 1, pp. 61–70, 2016, doi: 10.1038/jhg.2015.44.

[3] D. Fatkin, R. Otway, and J. I. Vandenberg, "Genes and atrial fibrillation: A new look at an old problem," *Circulation*, vol. 116, no. 7, pp. 782–792, 2007, doi: 10.1161/CIRCULATIONAHA.106.688889.

[4] K. Hayashi *et al.*, "Functional Characterization of Rare Variants Implicated in Susceptibility to Lone Atrial Fibrillation," *Circ. Arrhythmia Electrophysiol.*, vol. 8, no. 5, pp. 1095–1104, 2015, doi: 10.1161/CIRCEP.114.002519.

[5] M. Courtemanche, R. J. Ramirez, and S. Nattel, "Ionic mechanisms underlying human atrial action potential properties: insights from a mathematical model," *Am. J. Physiol. -Heart Circ. Physiol.*, vol. 275, pp. 301–321, 1998, doi: 10.1152/ajpheart.1998.275.1.H301.

[6] E. Grandi *et al.*, "Human atrial action potential and Ca 2+ model: Sinus rhythm and chronic atrial fibrillation," *Circ. Res.*, vol. 109, no. 9, pp. 1055–1066, 2011, doi: 10.1161/CIRCRESAHA.111.253955.

[7] E. A. Heidenreich, J. M. Ferrero, M. Doblaré, and J. F. Rodríguez, "Adaptive macro finite elements for the numerical solution of monodomain equations in cardiac electrophysiology," *Ann. Biomed. Eng.*, vol. 38, no. 7, pp. 2331–2345, 2010, doi: 10.1007/s10439-010-9997-2.

[8] U. Schotten, S. Verheule, P. Kirchhof, and A. Goette, "Pathophysiological mechanisms of atrial fibrillation: A translational appraisal," *Physiol. Rev.*, vol. 91, no. 1, pp. 265–325, 2011, doi: 10.1152/physrev.00031.2009.

[9] P. Sanders *et al.*, "Spectral analysis identifies sites of high-frequency activity maintaining atrial fibrillation in humans," *Circulation*, vol. 112, no. 6, pp. 789–797, 2005, doi: 10.1161/CIRCULATIONAHA.104.517011.

Feature Selection and Machine Learning for Predicting Multi-drug Resistance just After ICU Admission

J.Tarancón-Rey[1], I. Mora-Jiménez[1], J. Alvárez-Rodríguez[2], C. Soguero-Ruiz[1]

[1] Department of Signal Theory and Communications, Telematics and Computing Systems. Universidad Rey Juan Carlos, Madrid, España. j.tarancon.2016@alumnos.urjc.es, inmaculada.mora@urjc.es, cristina.soguero@urjc.es

[2] Intensive Care Department, Hospital Universitario de Fuenlabrada, Madrid, España. joaquin.alvarez@salud.madrid.org

Abstract

Multi-drug resistance (MDR) bacteria is one of the most current and greatest threats to the global health system assuming an increase in the mortality rate as well as health care costs. This situation is especially relevant in intensive care units (ICU) where the serious state of health of patients makes them more vulnerable. It is therefore necessary to identify MDR bacteria with the aim of taking appropriate action as soon as possible. Since cultures require 48-72 hours to obtain results, in this paper machine learning (ML) techniques are proposed to identify earlier whether a patient will develop MDR in the first 48 hours of admission. The use of the 48-hour threshold is due to the fact that it is the threshold used by clinical experts to discern between patients who have acquired the MDR in the ICU and those who have not. For this purpose, clinical and demographic features from the patient are considered (12 in total). The results of the feature selection identified a representative set of 24 features which can be used to identify MDR and non-MDR patients. That means that almost 75% of the initial features (95 after data preprocessing) are not informative for this task. We can conclude that the application of ML techniques contributes to the classification of patients according to the development of MDR during their first hours in ICU stay.

1. Introduction

Antimicrobials were one of the most outstanding inventions in the health sector since penicillin was discovered in 1928 [1]. An antimicrobial can be defined as a substance that kills or inhibits the growth of microorganisms such as bacteria, fungi, or parasites [2]. Depending on the microorganism they attack, antimicrobials are classified as: antibiotics, antivirals, fungal and antiprotozoal agents [3]. They constitute one of the most important tools in both the prevention and treatment of infectious diseases.

However, sometimes antimicrobials cannot be effective to treat infections. This happens when a microorganism goes from being sensitive to an antimicrobial to being unaffected by it, and it is called antimicrobial resistance [4]. The growing problem of antibiotic resistance, defined as the ability of bacteria to survive against chemical agents designed to kill them [5], poses a global threat in deteriorating the efficacy of these drugs [6].

There are several mechanisms by which a bacterium can resist the antibiotic action. Many of them are innate and these resistances do not suppose a major problem. On the other hand, the developed resistances can lead to complications in the treatment of diseases that are not a problem today, with a corresponding increase in healthcare costs [7]. Resistance is not necessarily limited to a single family of antibiotics. It can apply, simultaneously, to many chemically unrelated compounds to which the bacteria have never been exposed, this is known as multi-drug resistance (MDR) [8].

There are certain scenarios, where the appropriate conditions for the development of infections that generate antimicrobial resistance are presented. Health centers stand out among these scenarios since the use of antibiotics and the presence of germs are constant. It is in these places where nosocomial infections also originate. In addition, half of healthcare-related infections that occur in hospitals originate in the intensive care unit (ICU) [9] where the patient's health is more critical. In this sense, a MDR is considered to be acquired in the ICU if it was detected after the first 48 hours of the patient's admission in this unit. Therefore, there is a special interest in establishing a threshold in those 48 hours to know whether a patient will develop MDR during his/her admission in the ICU.

In this scenario, the application of data-driven tools currently available can help to control and mitigate the development of MDR. Machine learning (ML) allows the extraction of knowledge from data, to create models generalizing that knowledge to other cases, with the aim of anticipating and taking the appropriate measures.

ML techniques have been previously used to determine whether a bacterium will be resistant to different families of antimicrobials in the ICU [11]. In this work, we explore the potential of ML for early prediction of the MDR, since common antibiotics resistance tests require between 48 and 72 h after the culture is collected. Therefore, the objective of this work is to predict MDR earlier than the 48-72 hours usually required to determine the existence of antimicrobial resistance by a culture test.

The methodology followed in the work is presented in Figure 1. First, a data preprocessing stage is considered. Then, a FS strategy is carried out to get insights of possible risk factors. Finally, several ML techniques are applied to predict whether a patient will develop a MDR germ during the first 48 hours of the ICU stay.

The rest of the paper is organized as follows. In Section 2 the study dataset is described. Section 3 and Section 4 present the statistical approaches used for FS and learning process, respectively, while results are shown in Section 5. Finally, main conclusions and discussion are drawn in Section 6.

ISBN: 978-84-09-25491-0, pags. 145- 148

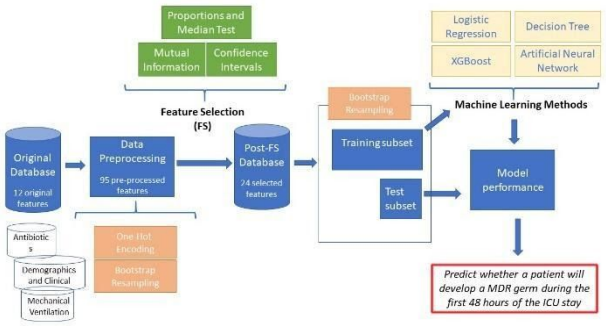


Figure 1. *Workflow diagram*

2. Database Description

For the development of this work, clinical data from the Hospital Universitario de Fuenlabrada, Madrid, Spain, have been used. More specifically, we analyze a total of 12 variables related to demographic data, the time the patient was with mechanical ventilation, the antibiotics administered as well as data related to cultures performed by the microbiology laboratory. This data was compiled for a total of 2600 patients admitted to the ICU during a period of 13 years between 2004 and 2016 (both inclusive). A patient can have more than one clinical episode, and therefore, there are a total of 3013 episodes (2743 episodes in which there are no multi-resistant germs in the first 48 hours of stay against 270 in which they are). We consider as a single sample all the events (admissions, cultures...) associated with a clinical episode while, in previous related studies, the study focus was at the culture level [11].

The exploratory analysis is carried out for 12 static variables. Since some of them are categorical, a one hot encoding strategy is considered, ending up in a total of 95 variables after considering this preprocessing stage. Additionally, the number of patients in the UCI and the number of MDR in the ICU the day the patient was admitted in the UCI are added as variables. They characterize the state of the ICU at a given time, providing information about the patients who shared the clinical unit. After doing this exploratory analysis, it is important to highlight the increase in the number of multi-resistant germs per year from 2004 to 2016.

3. Feature Selection Strategies

The next stage consists of a feature selection (FS) process using proportions test [12], median test [13], mutual information (MI) [14] and confidence intervals (CI) [15]. The FS is necessary in this type of study to identify the most informative variables to characterize the target variable. Thus, it allows us to eliminate those variables that may be noisy, irrelevant or redundant as they may adversely affect the proper design of the model as well as its interpretation. All these techniques are performed in combination with Bootstrapping [16] to reduce the risk of bias.

3. Machine Learning Methods

The last stage of the study focuses on designing statistical models to solve an automatic learning classification task. As mentioned previously, we deal with supervised learning since for each sample we know the associated label. Moreover, it is a binary classification task since the set of labels comprises only two values identifying each of the classes already explained. For this purpose, each sample in the original set $\mathcal{D}$ composed by N samples is composed by a vector of d features $\mathbf{x}$ (input variables) and the corresponding target output (denoted by a one-dimensional variable y), i.e.,

$$\mathcal{D} = \left\{\left(x^{(n)}, y^{(n)}\right)\right\}_{n=1}^{N}$$

The objective of ML techniques is to learn a model from the set of samples defined by a series of explanatory variables (determined after the FS process). ML methods are mainly based on optimizing a cost function [17]. To build the models, the dataset $\mathcal{D}$ is divided into two different subsets: training and test. The training set is used to design the model whereas the test set is used to evaluate it. We consider cross-validation strategy to tune the parameters of each classifier and to assess the generalization capacity.

In this paper, our goal is to learn the relationship between the feature vector (input variables to the model) and the labels associated with each vector (model output) by using FS and different ML strategies. We model this problem as a binary classification, where the target ($y^{(n)}$) is coded by values "1" and "0". The code "1" is associated with episodes with MDR during the first 48 hours after ICU admission. A label of "0" means that the episode has not been identified as MDR during the first 48 hours of the ICU stay (regardless of whether the patient develops MDR after the first 48 hours). More specifically, the following ML methods have been used: logistic regression (LR), decision tree (DT), xgboost (XGB) and artificial neural networks (ANN) (single-layer perceptron or SLP and multi-layer perceptron or MLP).

Logistic regression is a linear and parametric method of classification that allows obtaining the probability that each sample belongs to each of the classes considered. That is, it is a classifier with "soft" output [18]. Therefore, it is necessary to apply a threshold to assign the sample to one or another class depending on whether the probability obtained is greater or less than the established threshold.

On the other hand, a **decision tree** is a non-linear, non-parametric method that, intuitively, consists of a set of conditions organized in a hierarchical structure. In this way, the decision on the class to which each sample belongs can be determined by following the conditions that are met through the tree structure [19]. To this end, the parameters considered are the minimum number of samples to split a node, maximum number of samples for a terminal node and maximum depth.

Xgboost is a boosting method based on a set of decision trees so that the new trees that are incorporated into the model, correct the errors of those already in the model [20]. Thus, each new model is weighted by the results of the

previous model. In Xgboost, the parameters mentioned for decision trees together with the number of trees are tund.

Two approaches have been considered for **ANN**: a linear one by means of the SLP [21] and a nonlinear one as it is the MLP, which is considered a universal approximator of functions [22]. The ANN tries to imitate the structure of the nervous system. To do so, they interconnect neurons organized in layers, weighting the relationship between each pair of neurons by a weight that is necessary to learn during the design of the model. Thus, it is necessary to determine the number of neurons in hidden layers.

In addition, the dataset analyzed is imbalance, this is, there are 2743 episodes in which there are no multi-resistant germs in the first 48 hours of stay, and 270 in which they are. This can produce a bias that leads to a poor performance of the learning models because the minority class is not well represented. To solve this problem, two alternatives are considered: first, a balancing strategy has been developed, aimed at undersampling the majority class until the number of samples from both classes is equal [23]. This was done in combination with bootstrapping. It has also been considered a strategy based on changing the cost function used when designing the models [24]. In this strategy, models weigh more errors associated with the minority class.

4. Experiments and Results

This section includes the results of the proposed FS methods as well as the classification of MDR and non-MDR patients using the ML methods presented in Section 4.

The FS results when considering the proportions and medians tests discard 81% of the initial variables, ending up with 18 features as statistically significant (considering a significance level of 0.05), among them: *SAPS III Score*, the *age* of the patient or the reason of admission referred to *acute chronic respiratory failure*.

The results of FS with MI are very similar results to the ones obtained with the previous methods. Indeed, the features selected as statistically significant with proportions and median test are the ones with the highest mutual information values. There is an exception for the variable referring to the reason of admission: *acute chronic respiratory failure* that appears with a lower MI value than three features not considered significant with the previous experiment.

Considering the hypothesis test with CI, a total of 23 features are obtained as significant, while the median and proportions tests and MI select 18 attributes.

Finally, for the construction of the final subset of features, all the features obtained as significant with the proportions and median test and MI, adding those that have been significant with CI are considered. That is, the final subset consists of 24 features.

Once the relevant variables have been determined, we compute the correlation between the selected variables. We want to identify if there are redundant features, which could worse the performance of the ML models. Towards that end, Pearson correlation coefficient was used to determine the correlation between continuous features while the Spearman correlation coefficient was used for binary ones.

We check that there are not redundant features and, therefore, a total of 24 features are considered for the learning process. The performance of the ML schemes are evaluated with different figures of merit: *accuracy*, *sensitivity*, *specificity* and *Area Under the Curve* (*AUC) Score*.

Table 1 presents the results obtained with LR, DT, XGB, SLP and MLP, each one divided according to whether an undersampling strategy was used to balance the data, or the cost function was modified to work with the unbalanced classes. In general terms, the classifiers that have obtained the best results have been linear methods. Thus, LR has obtained the highest accuracy using the cost function modification, obtaining an accuracy of 66.1%. LR has also obtained the best results in terms of specificity and AUC, both by modifying the cost function to work with unbalanced classes. SLP achieves very similar results to LR, slightly improving sensitivity. Regarding sensitivity, the DT model is the one that has obtained the best results for this figure of merit, reaching a value of 62.8%.

Model	Strategy	Accuracy	Sensitivity	Specificity	AUC
LR	Undersampling	0.62±0.05	0.56±0.08	0.65±0.07	0.62±0.05
	Cost Function	**0.66±0.02**	0.61±0.07	**0.67±0.02**	**0.64±0.03**
DT	Undersampling	0.57±0.05	0.56±0.13	0.58±0.13	0.57±0.05
	Cost Function	0.56±0.10	**0.63±0.13**	0.55±0.12	0.59±0.03
XGB	Undersampling	0.59±0.05	0.54±0.08	0.61±0.08	0.59±0.05
	Cost Function	0.58±0.22	0.62±0.20	0.57±0.26	0.59±0.05
MLP	Undersampling	0.58±0.05	0.58±0.10	0.60±0.01	0.59±0.05
	Cost Function	0.64±0.05	0.61±0.09	0.64±0.05	0.63±0.04
SLP	Undersampling	0.62±0.05	0.60±0.07	0.65±0.07	0.64±0.04
	Cost Function	0.66±0.02	0.62±0.07	0.66±0.02	**0.64±0.03**

Table 1. *Mean ± standard deviation of the performance (Accuracy, Sensitivity, Specificity, AUC) provided by the models on the 50 test sets. Results are shown for the set of 24 selected features.*

In addition, a learning process is carried out considering also the number of patients admitted at the time of admission of a new patient and the number of them who present MDR. Despite the intuitive idea that these new variables could be relevant and enhance the obtained performance since cross transmission is one of the main risk factors of MDR germs, the results obtained are similar to those shown in Table 1.

5. Conclusions and Future Work

There are several studies that support the importance of the problem of antibiotic resistance, a problem that continues to grow worldwide [25]. MDR can be one of the major obstacles in the treatment of infectious diseases and considerably increasing hospital stays, mortality rates and the cost of health treatments, by delaying the use of appropriate therapy according to the effect it will have on the host. This is a particularly prevalent and serious

problem in the ICU [26], where due to the health conditions of the patients, the use of antibiotics is more frequent and accompanied by longer treatments. For all these reasons, it is essential to anticipate the development of MDR with the aim of isolating the patient as soon as possible, thus avoiding the spread of such resistance.

In this paper, it is proposed a strategy based on FS and ML techniques to identify MDR. On the one hand, the FS based on proportions and median test, MI and CI has demonstrated that a great number of the features considered in this work were not relevant. It has also been shown how ML methods (i.e, LR, DT, XGB and MLP) could be useful to predict MDR in the first 48 hours after ICU admission. Furthermore, they may accelerate the workflow of clinical centres despite the fact that there is still work to be done to obtain applicable results in the clinical practice. Among the results obtained from the experiments carried out, it should be noted that variables corresponding to patient severity classification systems such as *SAPS* and *Apache*, as well as others like *Age* and the medical service from which the patient comes are important features to distinguish the MDR patients from non-MDR patients.

Future work concerns deeper analysis of other clinical features to extract more knowledge about the bacterial transmission in the ICUs. Furthermore, other methodological approaches such as different strategies to balance classes aimed at generating new synthetic patients could enhance the obtained performance. Among these features, the patient's bed may be of interest because MDR could be transmitted to the closest patients, or even variables at ICU level as the total number of nurses working in the unit.

References

[1] De la Bédoyère, G. The discovery of penicillin. Evans Brothers Ltd, 2005.

[2] Franklin, T. J., and Snow, G. A. Biochemistry of antimicrobial action. Springer, 2013.

[3] Béahdy, J. Recent developments of antibiotic research and classification of antibiotics according to chemical structure. *Advances in applied microbiology*, vol 18, 1974, pp 309-406.

[4] Mendelson, M., & Matsoso, M. P. The World Health Organization global action plan for antimicrobial resistance. *SAMJ: South African Medical Journal*, vol 105, sup 5, 2015, pp 325-325.

[5] Siegel, J. D., Rhinehart, E., Jackson, M., & Chiarello, L. Management of multidrug-resistant organisms in health care settings, 2006. *American journal of infection control*, vol 35, sup 10, 2007, pp S165-S193.

[6] Alós, J. I. Resistencia bacteriana a los antibióticos: una crisis global. *Enfermedades infecciosas y microbiología clínica*, vol 33, sup 10, 2015, pp 692-699.

[7] Smith, R., & Coast, J. The true cost of antimicrobial resistance. *Bmj*, vol 346, 2013, pp f1493.

[8] Depardieu, F., Podglajen, I., Leclercq, R., Collatz, E., & Courvalin, P. Modes and modulations of antibiotic resistance gene expression. *Clinical microbiology reviews*, vol 20, sup 1, 2007, pp 79-114.

[9] Magill, S. S., Edwards, J. R., Bamberg, W., Beldavs, Z. G., Dumyati, G., Kainer, M. A., ... & Ray, S. M. Multistate point-prevalence survey of health care–associated infections. *New England Journal of Medicine*, vol 370, sup 13, 2014, pp 1198-1208.

[10] Witten, I. H., Frank, E., & Hall, M. A. Practical machine learning tools and techniques. Morgan Kaufmann, 2005.

[11] Martínez-Agüero, S., Mora-Jiménez, I., Lérida-García, J., Álvarez-Rodríguez, J., & Soguero-Ruiz, C. Machine Learning Techniques to Identify Antimicrobial Resistance in the Intensive Care Unit. Entropy, vol 21, sup 6, 2019,pp 603.

[12] Rosner, B. Fundamentals of Biostatistics. 7th ed, 2011.

[13] Tang, S., & Jeong, J. H. Median tests for censored survival data; a contingency table approach. Biometrics, vol 68, sup 3, 2012, pp 983-989.

[14] Shannon, C. E. A mathematical theory of communication. *The Bell system technical journal*, vol 27, sup 3, 1948, pp 379-423.

[15] Carlin, J. B., & Doyle, L. W. Basic concepts of statistical reasoning: standard errors and confidence intervals. Journal of Paediatrics and Child Health, vol 37, 2001, pp 72--77.

[16] James, G., Witten, D., Hastie, T., & Tibshirani, R. An introduction to statistical learning with Applications in R, 2013.

[17] Bishop, C. M. Pattern recognition and machine learning, 2006.

[18] Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. Applied logistic regression. Vol. 398, 2013.

[19] Kotsiantis, S. B. Decision trees: a recent overview. *Artificial Intelligence Review*, vol 39, sup 4, 2013, pp 261-283.

[20] Chen, T., & He, T. Higgs boson discovery with boosted trees. *NIPS 2014 workshop on high-energy physics and machine learning*, Montreal, 2015, pp 69-80.

[21] Rosenblatt, F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, vol 65, sup 6, 1958, pp 386.

[22] Ripley, B. D. Pattern recognition and neural networks. Cambridge university press, 2008.

[23] Liu, X. Y., Wu, J., & Zhou, Z. H. Exploratory undersampling for class-imbalance learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),* vol 39, sup 2, 2008, pp 539-550.

[24] He, H., & Garcia, E. A. Learning from imbalanced data. *IEEE Transactions on knowledge and data engineering*, vol 21, sup 9, 2009, pp 1263-1284.

[25] French, G. L. Clinical impact and relevance of antibiotic resistance. *Advanced drug delivery reviews*, vol 57, sup 10, 2005, pp 1514-1527.

[26] Zilahi, G., McMahon, M. A., Povoa, P., & Martin-Loeches, I. Duration of antibiotic therapy in the intensive care unit. *Journal of thoracic disease*, vol 8, sup 12, 2016, pp 3774.

Construcción de pupilómetro binocular para estudio de lentes intraoculares

J.C. Gómez Blanco[1], P. Serra[2], J.L. Moyano-García[1], A. Sánchez-Trancón[2], O. Torrado[2], F.M. Sánchez Margallo[1], J.B. Pagador Carrasco[1]

[1] Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España
{jcgomez,jmoyano,msanchez,jbpagador}@ccmijesususon.com

[2] Clínica Oftalmológica Sánchez Trancón, Badajoz, España, {pedro.serra, sancheztrancon, direccionmedica}@vistasancheztrancon.com

Resumen

Las cataratas son una patología muy extendida provocada por la descomposición de las proteínas del cristalino que provoca en última instancia la ceguera del paciente. Esta patología se trata mediante la extirpación del cristalino y la inserción de una lente intraocular. El cálculo de las lentes es un reto para el clínico que necesita datos biométricos fiables para conseguir el efecto que el paciente necesita. Uno de los datos necesarios es el diámetro pupilar que usualmente se mide mediante pupilómetros monoculares. Hay estudios que sugieren que el diámetro de la pupila cambia si se mide solo una o las dos a la vez. Por eso el fin de este trabajo es construir un pupilómetro binocular capaz de medir el diámetro de las dos pupilas a la vez. Los resultados demuestran que el prototipo construido es capaz de medir los dos diámetros pupilares con errores de 2 píxeles. Esto implica que el dispositivo desarrollado puede medir las pupilas con errores menores a 0.3 mm.

1. Introducción

Las cataratas son una alteración óptica muy extendida provocada en su mayoría por la edad. Se trata del cambio de transparencia del cristalino por la descomposición de sus proteínas hasta volverse totalmente opaco. Esta patología es causa de incapacidad visual, con impacto físico y social [1]. Su diagnóstico necesita un examen completo del ojo para detectar todas las posibles anomalías en todos los niveles del ojo. Una vez diagnosticada las cataratas el oftalmólogo decidirá el tratamiento en función de su tipo y del grado de opacidad. No obstante, el único tratamiento realmente efectivo es la extracción del cristalino y su reemplazo por una lente intraocular (LIO) [2]. Este procedimiento se trata de una intervención ambulatoria de unos 20 minutos con anestesia tópica.

Existe una diferenciación entre las diferentes LIO, las más utilizadas o estándar y las conocidas como “Premium” que tienen características específicas. La evolución de las estas últimas LIO permiten al paciente alcanzar un determinado desempeño visual que las lentes estándar no son capaces de conseguir [3]. La utilización de LIO premium supone una serie de desafíos en cuanto a su correcta elección y el mejor cálculo para un paciente determinado [4].

El cálculo de la potencia de una determinada LIO supone considerar el ojo como un sistema óptico compuesto por varias superficies ópticas para conseguir la generación de una determinada imagen en la retina [5]. Actualmente este cálculo se realiza mediante óptica Gaussiana y trazado de rayos. Para poder obtener buenos resultados es necesario incorporar en el cálculo de las LIO medidas biométricas precisas y parámetros clínicos optimizados para llegar hasta errores refractivos postoperatorios de 0.5 Dioptrías [6]. De entre todos los parámetros necesarios uno de los primeros parámetros a medir es el diámetro pupilar. Para ello existen diferentes modelos de pupilómetros en condiciones de baja iluminación y de forma monocular. Estos pupilómetros tienen un problema ya que según Kawamorita et al. [7] el diámetro pupilar obtenido monocularmente no se corresponde con la dimensión del medido binocularmente.

Por lo tanto, el objetivo de este trabajo es diseñar y testear un pupilómetro binocular que permita la medición de la pupila en condiciones de iluminacion variables y diferentes distancias de observación.

2. Materiales y métodos

Los requerimientos mínimos que el sistema debía tener se establecieron de la experiencia de los clínicos. Los puntos más característicos que señalaron fueron:1) el paciente no puede mirar al infinito para realizar ciertas pruebas y 2) no se podía controlar las condiciones lumínicas. Con eso en mente se les presentó el boceto mostrado en la figura 1.

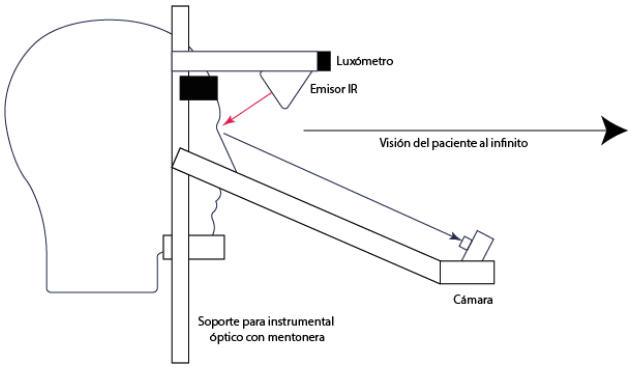


Figura 1.*Boceto de diseño de pupilómetro*

Estos dos requerimientos son necesarios para asegurar una medida estándar de la pupila en condiciones establecidas. El diámetro pupilar varía en función de la fijación de la imagen. Por lo tanto, en el seguimiento de determinados pacientes es necesario provocar que fijen la mirada a

ISBN: 978-84-09-25491-0, pags. 149- 151

diferentes distancias. Además, es necesario fijar la cantidad lumínica incidente para estandarizar todas las medidas realizadas.

En este diseño se propone un sistema abierto que permite al paciente mirar al infinito. Además, cuenta con un luxómetro para medir la intensidad lumínica incidente en los ojos del paciente. El prototipo cuenta con una serie de leds IR para iluminar los ojos del paciente y facilitar al sistema de captura una imagen clara (figura 2). Una vez que los clínicos validaron la idea inicial de pupilómetro, se diseñó mediante el software CAD Autodesk Inventor.



Figura 2.*Captura de imagen de cámara*

El prototipo final se construyó mediante impresión 3D toda la electrónica se controla mediante un Arduino Uno como se muestra en la figura 3. El sistema de captura está compuesto por una cámara infrarroja Genie Nano M1930-NIR (Sensor 1″ CMOS On-Semi Python 2000, tamaño de pixel 4.8 µm y 1920x1200 px) dispuesta con una angulación de 20° a 20 cm de las pupilas. La colocación estratégica de la cámara asegura que el error en la segmentación sea estable y aproximadamente 1 pixel en todo el borde, lo que implica un error máximo de 2 pixeles en la medida final que equivale a 260 µm.

Junto con el hardware se diseñó el software de adquisición de imágenes y segmentación de las pupilas. El software está compuesto por el módulo firmware de control de la cámara y luces IR y el software de medición biométrico.

El funcionamiento del firmware tiene como objetivo la comunicación entre los dispositivos hardware y el software de cálculo (figura 4). Todo el firmware se ha desarrollado en Python 2.7 y conecta la cámara con los leds IR para que sólo se enciendan cuando se van a tomar las imágenes. De esta manera se garantiza la seguridad del paciente minimizando la radiación IR directa en los ojos.

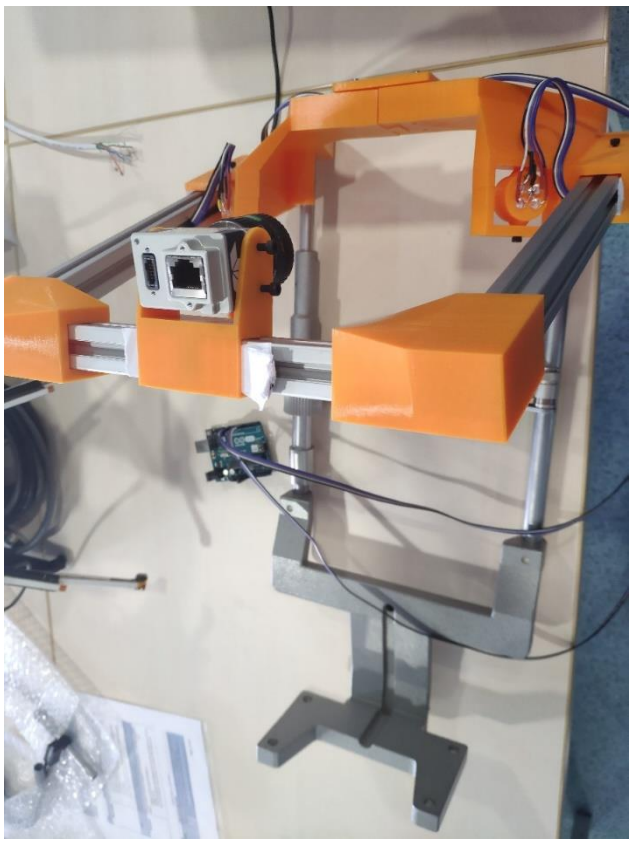


Figura 3.*Primer prototipo de pupilómetro*

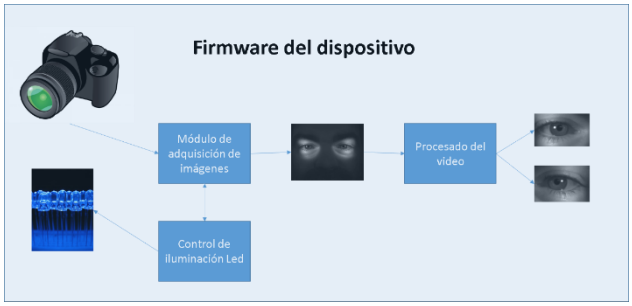


Figura 4.*Esquema de funcionamiento del firmware.*

El módulo de procesado del video en el firmware pide al usuario que establezca dos ROI – Region Of Interest (cada uno de los ojos) y genera un vídeo de 10 frames.

Estos vídeos son los datos de entrada del software biométrico donde se calcula el diámetro pupilar y la distancia entre pupilas. Los cálculos se realizan asimilando la pupila a una elipse (por la distorsión angular provocada por la posición de la cámara) y se obtiene el radio mayor (horizontal) y el centroide de la elipse. Con esos datos se calcula el diámetro de la pupila y la distancia entre ellas.

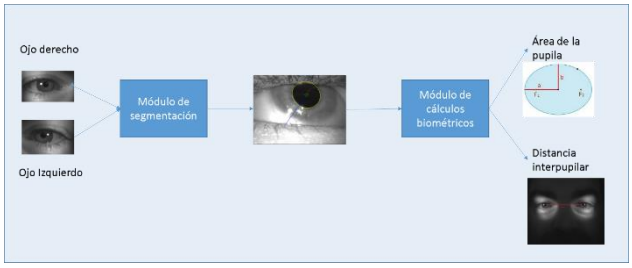


Figura 5.*Esquema de funcionamiento del software biométrico.*

Se realizó una prueba inicial de funcionamiento en la que se midió el diámetro pupilar de un paciente completamente sano, sin anomalías oculares. La medición consistió en obtener 2 segundos de adquisición (19 imágenes) de cada ojo y medir el diámetro pupilar. El valor del diámetro se calculó como la media entre el eje mayor y el menor de la elipse ajustada.

3. Resultados y discusión

De las imágenes obtenidas se calcula que cada pixel representa 130 µm de la imagen capturada. Los resultados de las mediciones del diámetro de la pupila en los dos ojos del paciente se muestran en la figura 6.

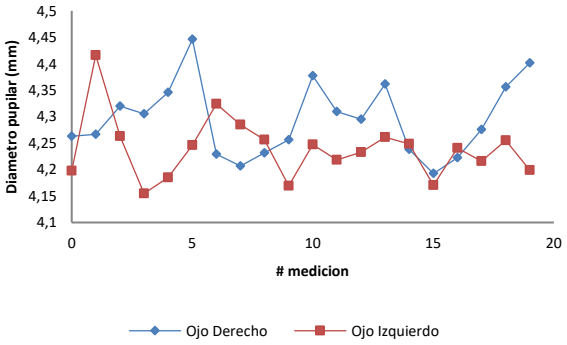


Figura 6.*Tamaño de la pupila (mm) obtenido por el prototipo de pupilómetro binocular en el paciente estudiado.*

De los resultados se obtiene que existe una fluctuación de 2 pixeles (0.26 mm) tanto en el ojo izquierdo como en el derecho. Esta oscilación coincide con el movimiento involuntario y periódico de las pupilas conocido como Hippus Pupilar (<0.5 mm). Junto con este movimiento de la pupila, el error medido es el esperado en el prototipo tal y como se ha mencionado en la metodología. Los valores de las pupilas son 4.295±0.069 mm y 4.239±0.0.059 mm para el ojo izquierdo y derecho respectivamente.

Dado que la prueba realizada es una prueba inicial de funcionamiento del prototipo no se cuenta con un tamaño pupilar del paciente para contrastar la medida. Se tiene previsto en trabajos futuros hacer una batería de pruebas para medir la bondad de las mediciones obtenidas utilizando ojos artificiales. De estas pruebas se obtendrán los errores reales del prototipo y las calibraciones que son necesarias para que funcione correctamente.

No obstante, con estos resultados se puede afirmar que el pupilómetro construido es capaz de medir el tamaño pupilar de los dos ojos simultáneamente con un error teórico de aproximadamente 260 µm teniendo en cuenta el Hippus Pupilar.

Agradecimientos

Los autores desean agradecer el apoyo de este trabajo de investigación al VI Plan Regional de Investigación y a la Junta de Extremadura y al Fondo Europeo de Desarrollo Regional (FEDER) vinculado a la financiación de los grupos de investigación del Centro de Cirugía Mínimamente Invasiva Jesús Usón (GR18199).

Referencias

[1] S. Resnikoff, D. Pascolini, D. Etya'ale, I. Kocur, R. Pararajasegaram, G.P. Pokharel, S.P. Mariotti, Global data on visual impairment in the year 2002., Bull. World Health Organ. 82 (2004) 844–851.

[2] A. Foster, Cataract and "Vision 2020-the right to sight" initiative., Br. J. Ophthalmol. 85 (2001) 635–637. https://doi.org/10.1136/bjo.85.6.635.

[3] W. Haigis, Challenges and approaches in modern biometry and IOL calculation., Saudi J. Ophthalmol. Off. J. Saudi Ophthalmol. Soc. 26 (2012) 7–12. https://doi.org/10.1016/j.sjopt.2011.11.007.

[4] E. Rosen, J.L. Alió, H.B. Dick, S. Dell, S. Slade, Efficacy and safety of multifocal intraocular lenses following cataract and refractive lens exchange: Metaanalysis of peer-reviewed publications., J. Cataract Refract. Surg. 42 (2016) 310–328. https://doi.org/10.1016/j.jcrs.2016.01.014.

[5] T. Olsen, Calculation of intraocular lens power: a review., Acta Ophthalmol. Scand. 85 (2007) 472–485. https://doi.org/10.1111/j.1600-0420.2007.00879.x.

[6] S. Norrby, Sources of error in intraocular lens power calculation., J. Cataract Refract. Surg. 34 (2008) 368–376. https://doi.org/10.1016/j.jcrs.2007.10.031.

[7] T. Kawamorita, Natural Pupil Size and Ocular Aberration under Binocular and Monocular Conditions, J. Comput. Sci. Syst. Biol. 7 (2014). https://doi.org/10.4172/jcsb.1000133.

Competición de Estudiantes de Grado en Ingeniería Biomédica

Miércoles, 25 de noviembre de 2020

Caracterización de una plataforma electroactiva de hidrogel-microesferas para la diferenciación osteogénica de células madre mesenquimales

M.I. García-Briega[1], M. Guillot-Ferriols[1,2], S. Lanceros-Méndez[3,4,5], J.L. Gómez-Ribelles[1,2], G. Gallego-Ferrer[1,2]

[1] Centre for Biomaterials and Tissue Engineering (CBIT), Universitat Politècnica de València, 46022, Valencia, Spain

[2] Centro de Investigación Biomédica en Red en Biomateriales, Bioingeniería y Nanomedicina (CIBER-BBN), Valencia, Spain

[3] Centro/Departmento de Física, Universidade Do Minho, 4710-057 Braga, Portugal

[4] BCMaterials UPV/EHU Parque Científico, 48940 Leioa, Spain

[5] IKERBASQUE, Basque Foundation for Science, 48013 Bilbao, Spain

Resumen

En el hueso existen patologías que el propio tejido es incapaz de reparar. Para tratarlas se emplean diversas técnicas, entre las que se encuentra la terapia celular. Esta terapia emplea habitualmente células madre mesenquimales de médula ósea (hMSC) autólogas, debido a su gran potencial regenerador. En la terapia, las hMSCs se diferencian in situ tras la inyección. Nuestro trabajo se sustenta en la hipótesis de que el uso de células prediferenciadas a linaje óseo mejoraría los resultados obtenidos. Para prediferenciar las hMSCs, se ha puesto a punto una plataforma de cultivo 3D basada en un hidrogel de gelatina y microesferas, capaz de convertir estímulos magnéticos externos en estímulos eléctricos, simulando el entorno piezoeléctrico del hueso para favorecer la diferenciación a linaje óseo. Las microesferas se han fabricado mediante electrospray y se componen del polímero piezoeléctrico fluoruro de polivinilideno (PVDF) y nanopartículas magnetoestrictivas de ferrita de cobalto (CFO). Se han fabricado y caracterizado físico-químicamente las microesferas, se ha observado la distribución de estas en el hidrogel y se ha probado la no citotoxicidad de la plataforma con y sin estimulación magnética externa mediante ensayos de proliferación.

1. Introducción

El tejido óseo se compone de una matriz extracelular basada en hidroxiapatita y colágeno I y distintos tipos celulares. Una de las propiedades más relevantes de los huesos es su capacidad de regenerarse y repararse a lo largo de toda la vida del individuo. Esta capacidad de remodelación ósea se ve promovida por los campos eléctricos inducidos por las cargas cíclicas a las que se ve sometido el hueso, gracias a la piezoelectricidad que el colágeno I le confiere al tejido [1]. Estos estímulos eléctricos también podrían estimular la diferenciación de las células madre mesenquimales de médula ósea (hMSCs), ya que afectarían a las proteínas transmembrana dependientes de voltaje. Esto desencadenaría procesos en el interior celular que acabarían activando la expresión de genes relacionados con la diferenciación celular [2].

La capacidad de auto-reparación ósea mengua con la edad de los individuos. Debido a que la esperanza de vida ha aumentado en los últimos tiempos, la incidencia de patologías óseas difíciles de curar también ha aumentado. Por ello, el interés por desarrollar terapias que contribuyan a la regeneración de daños asociados al tejido óseo ha crecido. De las terapias empleadas actualmente, destaca la terapia celular, que busca regenerar el tejido inyectando células en el lugar del defecto con o sin andamiajes porosos. En el tratamiento de patologías óseas, las células escogidas suelen ser las hMSCs por ser células madre autólogas. Sin embargo, su uso conlleva desventajas, como la diferenciación a otros tejidos o la pérdida del fenotipo durante la amplificación *in-vitro*. No obstante, estas desventajas desaparecerían si las células empleadas ya estuvieran prediferenciadas al linaje óseo, y esta es la motivación para el trabajo realizado.

Para sintetizar la plataforma de cultivo, se seleccionan la gelatina, el polímero piezoeléctrico fluoruro de polivinilideno (PVDF) y el material magnetostrictivo ferrita de cobalto (CFO). Esta elección se ha realizado teniendo en cuenta que se busca generar estímulos eléctricos dentro de la plataforma a partir de estímulos magnéticos externos y buscando mimetizar al máximo el entorno óseo. El hidrogel de gelatina proporciona un entorno 3D donde distribuir las microesferas y las células logrando así una correcta estimulación.

El PVDF es un polímero piezoeléctrico y semicristalino cuyas fases cristalinas son la α, la β y la γ. Cada fase confiere al PVDF distintas propiedades electroactivas, siendo la fase β la más piezoeléctrica de las tres [3]. Para maximizar la formación de cristales en fase β en las microesferas, se sintetizaron mediante electrospray.

En el diseño de las microesferas se hace valer el efecto magnetoeléctrico, que consiste en el trabajo combinado de un material piezoeléctrico y uno magnetostrictivo para convertir estímulos magnéticos en eléctricos y viceversa. Estas microesferas contendrán nanopartículas de CFO en la matriz de PVDF, por lo que al aplicar un campo magnético externo al constructo 3D, el PVDF será capaz de generar estímulos eléctricos, ya que se verá deformado gracias a la deformación del CFO.

2. Materiales y métodos

2.1. Preparación y caracterización de microesferas

Se han obtenido dos tipos de microesferas, de PVDF (PVDF) y de PVDF con nanopartículas de CFO (PVDF-CFO). Ambos tipos se han sintetizado mediante electrospray modificando el protocolo descrito por Gonçalves *et al*. [4]. La solución polimérica empleada para

ISBN: 978-84-09-25491-0, pags. 153- 156

las microesferas de PVDF se compone en un 9% w/v de PVDF, un 85% v/v de N,N-dimetilformamida (DMF) y un 15% v/v de tetrahidrofurano (THF). La solución polimérica empleada para las microesferas de PVDF-CFO se compone en un 4% w/v de PVDF, un 85% v/v de DMF y un 15% v/v de THF, un 20% CFO en relación al peso de PVDF y un 1% v/v de tritón x-100.

La caracterización morfológica de las microesferas se realizó empleando el Microscopio Electrónico de Barrido de Emisión de campo (FESEM) a 2 kV. Para conocer los porcentajes de las fases cristalinas en ellas se usó la espectroscopía infrarroja por transformada de Fourier (FTIR). Se empleó el espectrofotómetro en modo ATR, con un barrido desde 4000 a 400 cm^{-1}, una resolución de 4 cm^{-1}, y las ecuaciones propuestas por Gregorio *et al.* [5] (Ec.1 y Ec.2). El grado de cristalinidad del PVDF se midió por calorimetría diferencial de barrido (DSC), realizando un barrido de calentamiento de 0 a 200 ºC a 20 ºC/min. Se obtuvo el porcentaje de cristalinidad con la ecuación propuesta por Gonçalves *et al.* [4] (Ec.3), teniendo en cuenta fracción en peso de PVDF en las microesferas analizadas, w_{PVDF}. Mediante el magnetómetro de muestra vibrante (VSM) se determinó la magnetización de las microesferas de PVDF-CFO para calcular el porcentaje real en masa de CFO en ellas, valiéndose de la Ec. 4 descrita por Gonçalves *et al.* [4]. Para conocer la distribución de las nanopartículas de CFO dentro de la matriz de PVDF se empleó el cañón de iones focalizado (FIB) para cortar las microesferas, observando la sección mediante FESEM.

$$F(\alpha) = \frac{\Delta_\alpha}{\left(\frac{K_\alpha}{K_\beta}\right)\Delta_\beta + \Delta_\alpha} \qquad \text{(Ec.1)}$$

$$F(\beta) = \frac{\Delta_\beta}{\left(\frac{K_\beta}{K_\alpha}\right)\Delta_\alpha + \Delta_\beta} \qquad \text{(Ec.2)}$$

Δ_α y Δ_β absorbancia a 760 cm^{-1} y 840 cm^{-1} y K_α y K_β coeficientes de absorción característicos [5]

$$X_C = \frac{\Delta H}{w_{PVDF}(F(\alpha)\Delta H_\alpha + F(\beta)\Delta H_\beta)} \qquad \text{(Ec.3)}$$

ΔH entalpía de fusión de las microesferas, ΔH_α y ΔH_β entalpías de fusión de PVDF totalmente en fase α y en fase β, respectivamente[3].

$$\%\ CFO = \frac{Magnetización\ de\ saturación\ microesferas}{Magnetización\ de\ saturación\ CFO\ puro} x100 \qquad \text{(Ec.4)}$$

2.2. Hidrogel de gelatina

Se injertó tiramina (Tyr) a la gelatina (de piel porcina, gel strength 300, Tipo A) para entrecruzarla enzimáticamente mediante reacción no citotóxica y poder contener las micropartículas y las células. Se siguió el protocolo de Poveda-Reyes *et al.* [6], tanto para el injerto como para su posterior reticulación mediante peroxidasa (HRP) y peróxido de hidrógeno (H_2O_2). El hidrogel se preparó al 2% w/v en el buffer libre de calcio de Krebs Ringer. Cada hidrogel se obtuvo empleando las proporciones del 80% v/v de Gel al 2% w/v, 10% v/v de HRP (12.5 U/mL) y 10% v/v de H_2O_2 (20 mM). Para gelificarlo, se mezcló la disolución de gelatina y de HRP, se añadió un 0.6% w/v de microesferas y se reticuló añadiendo el volumen de H_2O_2.

Para comprobar la distribución de las microesferas en el hidrogel, se realizaron crio-secciones que permitieron observar su interior al microscopio óptico. Para los crio-cortes, se incluyeron los hidrogeles gelificados en O.C.T, se congelaron a -80 ºC y se cortaron empleando un criostato.

2.3. Cultivos celulares

Se realizaron dos tipos de cultivos *in vitro* a corto plazo de las hMSCs. Exceptuando la existencia de estimulación, las condiciones y la evaluación fueron iguales en ambos cultivos. Los hidrogeles fueron de 50 µL, con una densidad celular de 1 x 10^6 células/mL y con tres condiciones diferentes: gelatina sin microesferas (Gel), gelatina con microesferas de PVDF (Gel-PVDF) y gelatina con microesferas de PVDF-CFO (Gel-PVDF-CFO).

Siembra en hidrogel

Todos los materiales fueron esterilizados y filtrados. Los hidrogeles mantuvieron las mismas proporciones que para los crio-cortes, pero se añadieron hMSCs antes de añadir el H_2O_2 para reticularlos.

Para el cultivo estático, se empleó un medio basal compuesto por DMEM alto en glucosa (4.5 g/L) suplementado con un 10 % v/v de suero fetal bovino (FBS), 1% v/v de aminoácidos no esenciales (100X), 1% v/v de piruvato de sodio (100 mM) y 1% v/v de penicilina-estreptomicina (100 U/mL), en atmósfera controlada a 37 ºC y 5% de CO_2 y realizando un cambio de medio cada 72 h.

En el ensayo dinámico, las hMSCs se cultivaron en las mismas condiciones, pero en un biorreactor magnético. Éste aplica un campo magnético externo alterno a baja frecuencia (0.3 Hz), que llega a los 0.023 T y que mimetiza el comportamiento humano simulando ciclos de estimulación (8 horas de reposo, 16 de actividad) [7].

Proliferación celular (MTS)

Las pruebas de proliferación en los cultivos estático y dinámico se realizaron para los tiempos 3 h, 1, 3 y 7 días y 3 h, 1 y 3 días, respectivamente. Se empleó 3 réplicas por tiempo y material.

Para cada tiempo ensayado se incubaron los hidrogeles con el reactivo MTS (1:6 en DMEM sin rojo fenol) durante 2 horas a 37 ºC en oscuridad. Luego se midió la absorbancia del medio MTS a 490 y 690 nm en el espectrofotómetro de placas. Los resultados se normalizaron respecto a la absorbancia a tiempo 0 (3 h) obteniendo así la proliferación relativa para cada tiempo. Se llevó a cabo un análisis two way-ANOVA mediante el software GraphPad Prism 7 en el que se consideró que existían diferencias significativas a partir de un p-valor < 0.05.

3. Resultados y discusión

3.1. Preparación y caracterización de microesferas

El uso del FESEM permitió confirmar la obtención de las microesferas tanto de PVDF (Fig1.a) como de PVDF-CFO (Fig1.b) a través del electrospray. Esto se debe al porcentaje de PVDF en disolución, menor al 10 % w/v en

ambos casos, que asegura la formación de microesferas y no fibras [8]. La medición de los diámetros de las microesferas muestra un menor diámetro en las microesferas PVDF-CFO, atribuible a la reducción del porcentaje de PVDF en la disolución para evitar que el CFO dificultara la dispersión y formación de microesferas. Los diámetros son de 3.2 ± 1.6 µm en las microesferas de PVDF y de 2 ± 0.6 µm en las de PVDF-CFO.

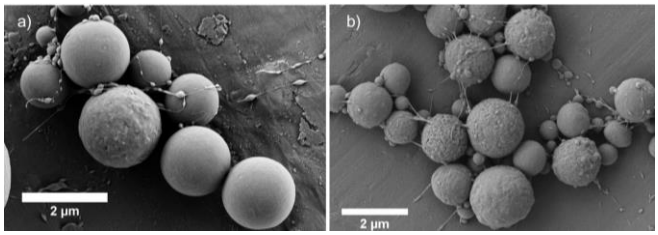


Figura 1. *a) Imagen FESEM de microesferas de PVDF. b) Imagen FESEM de microesferas de PVDF-CFO*

El espectro resultante del ensayo FTIR (Fig2) permite el cálculo de los porcentajes de las fases cristalinas a partir de los picos característicos de cada número de onda. Con estos resultados, se comprueba que se ha maximizado la presencia de fase β (Tabla 1).

Tabla 1. *Porcentajes de las fases en cada tipo de microesferas*

	PVDF	PVDF-CFO
Fase α (%)	15.2	13.4
Fase β (%)	84.8	86.6

Esta maximización de la fase β puede atribuirse a diversos sucesos durante el proceso de síntesis que incrementan su formación. Primero, la temperatura del proceso (menor a 70 ºC), segundo, el uso de DMF como solvente, y tercero, el uso de altos voltajes que generar fuerzas de tracción en el *jet* del electrospray [3].

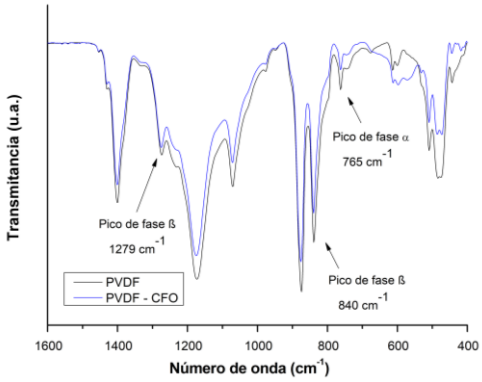


Figura 2. *Espectro FTIR de las dos clases de microesferas*

Los termogramas obtenidos del DSC demostraron que el electrospray puede ofrecer cristalinidades del PVDF superiores al 50% en ambos tipos de microesferas, de 57.9% en las de PVDF y 55.8% en las de PVDF-CFO.

Con el VSM obtenemos la magnetización de saturación de las microesferas PVDF-CFO y con ello, calculamos el porcentaje real de CFO en masa que éstas poseen. El porcentaje resultante es de 9.2 % de masa de CFO dentro de las microesferas, que comparado con el 20% de masa de CFO inicial, evidencia pérdidas superiores a la mitad de la masa inicial al sintetizar con electrospray. Pese a la disminución de CFO, las nanopartículas son capaces de distribuirse por la matriz polimérica, tal y como se observa en el corte del FIB (Fig3). Esta distribución genera superficies de contacto entre el PVDF y el CFO que permiten presuponer el correcto desarrollo del efecto magnetoeléctrico.

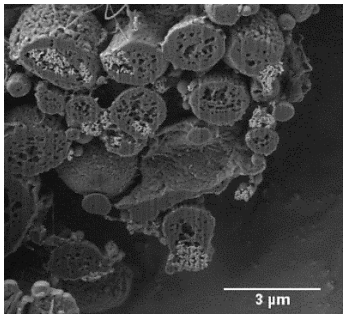


Figura 3. *Imagen de un corte transversal de las microesferas de PVDF-CFO mediante FIB obtenida empleando el FESEM.*

3.2. Síntesis de hidrogeles inyectables de gelatina

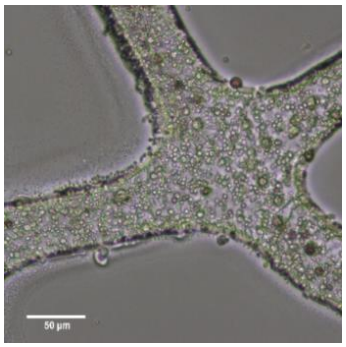


Figura 4. *Imagen representativa al microscopio óptico de un crio-corte a un hidrogel con microesferas de PVDF.*

Las imágenes obtenidas de las crio-secciones (Fig4) permiten observar trabéculas de poco grosor y poros de gran tamaño, atribuibles al bajo porcentaje de gelatina presente en el hidrogel (2% w/v). También se observó una distribución similar y homogénea de las microesferas de los dos tipos en las trabéculas, lo que genera un entorno en el que las hMSCs pueden ser correctamente estimuladas.

3.3. Cultivos celulares

Cultivo Estático

El objetivo del ensayo era comprobar la no-citotoxicidad de los materiales empleados en la plataforma 3D.

En la Figura 5 se puede apreciar que, mientras que en los geles de gelatina no hay diferencia significativa en la proliferación a lo largo del tiempo, en los que contienen microesferas sí la hay.

Existe una diferencia significativa en la proliferación entre los hidrogeles Gel-PVDF-CFO con respecto al control (Gel) a 1 y 3 días (no indicada en el gráfico por claridad), que se atribuye a la menor proliferación celular del hidrogel Gel-PVDF-CFO. Esta baja proliferación, probablemente, sea debida a una mayor muerte inicial por presencia de la ferrita en el constructo 3D. Tras 7 días de cultivo, la proliferación no es diferente si comparamos los

3 tipos de hidrogeles, lo que indica la no citotoxicidad de las microesferas.

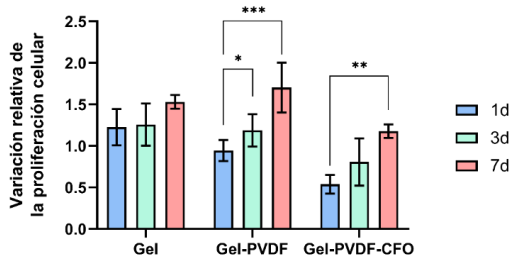


Figura 5. Variación relativa de la *proliferación celular del cultivo estático medida por MTS. * p-valor < 0.05, ** p-valor < 0.01, *** p-valor <0.005*

Cultivo dinámico

El objetivo del ensayo era determinar si el campo magnético externo afecta negativamente a la proliferación de las hMSCs y comprobar si la interacción de este campo con los biomateriales resultaba nociva para el cultivo.

En la Figura 6 vemos que, pese a que no hay diferencias de proliferación al comparar entre materiales, sí las hay en todos los constructos entre 1 y 3 días. Esto confirma de nuevo la no-citotoxicidad de los constructos empleados y del campo magnético externo. Si comparamos con los resultados del cultivo en estático, vemos que en este se ha incrementado significativamente la proliferación en los hidrogeles de gelatina, cosa que no pasaba en el anterior. Esto implica que el campo magnético tiene algún tipo de efecto por sí mismo en la proliferación. En trabajos futuros se debería comprobar si el campo magnético estimula la diferenciación osteogénica en los materiales piezoeléctricos y no solo la proliferación como se ha visto en este trabajo.

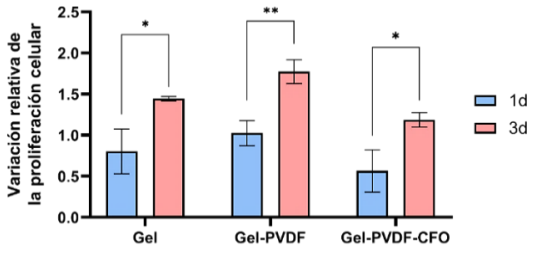


Figura 6. Variación relativa de la *proliferación celular del cultivo dinámico medida por MTS. * p-valor < 0.05, ** p-valor < 0.01*

4. Conclusiones

Se ha comprobado que el método del electrospray es capaz de producir microesferas de PVDF con y sin nanopartículas de CFO con un alto porcentaje de fase β. Además, estas nanopartículas están distribuidas por toda la microesfera y pueden desarrollar su función, a pesar de que la mitad de la masa de CFO se retiene en la jeringa. También se ha probado la homogénea distribución de las microesferas dentro del hidrogel, permitiendo una correcta estimulación de las células cultivadas.

A través de los cultivos celulares, se concluye que los constructos 3D con microesferas de PVDF y PVDF-CFO no son citotóxicos ya que no afectan negativamente a la proliferación celular de las hMSC. De igual forma, la aplicación de un campo magnético junto al constructo, tampoco perjudica al cultivo celular. Sin embargo, el propio campo magnético parece tener un efecto en sí mismo sobre la proliferación celular, independientemente de la plataforma 3D empleada.

Agradecimientos

Agradecemos la financiación de la Agencia Española de Investigación por los proyectos PID2019-106099RB-C41 y C43 /AEI/10.13039/501100011033, incluidos los fondos FEDER. MGF agradece la financiación FPI BES-2017-080398.

Referencias

[1] C. A. Bassett, “Electrical effects in bone.,” *Scientific American*, vol. 213, no. 4, pp. 18–25, 1965, doi: 10.1038/scientificamerican1065-18.

[2] J. Jacob, N. More, K. Kalia, and G. Kapusetti, “Piezoelectric smart biomaterials for bone and cartilage tissue engineering,” *Inflammation and Regeneration*, vol. 38, no. 1, pp. 1–11, 2018, doi: 10.1186/s41232-018-0059-8.

[3] P. Martins, A. C. Lopes, and S. Lanceros-Méndez, “Electroactive phases of poly(vinylidene fluoride): Determination, processing and applications,” *Progress in Polymer Science*, vol. 39, no. 4, pp. 683–706, 2014, doi: 10.1016/j.progpolymsci.2013.07.006.

[4] R. Gonçalves *et al.*, “Development of magnetoelectric CoFe2O4 /poly(vinylidene fluoride) microspheres,” *RSC Advances*, vol. 5, no. 45, pp. 35852–35857, 2015, doi: 10.1039/c5ra04409j.

[5] R. Gregorio, and M. Cestari, “Effect of crystallization temperature on the crystalline phase content and morphology of poly(vinylidene fluoride),” *Journal of Polymer Science Part B: Polymer Physics*, vol. 32, no. 5, pp. 859–870, 1994, doi: 10.1002/polb.1994.090320509.

[6] S. Poveda-Reyes *et al.*, “Injectable composites of loose microfibers and gelatin with improved interfacial interaction for soft tissue engineering,” *Polymer*, vol. 74, pp. 224–234, Aug. 2015, doi: 10.1016/j.polymer.2015.08.018.

[7] M. M. Fernandes, D. M. Correia, C. Ribeiro, N. Castro, V. Correia, and S. Lanceros-Mendez, “Bioinspired Three-Dimensional Magnetoactive Scaffolds for Bone Tissue Engineering,” *ACS Applied Materials and Interfaces*, vol. 11, no. 48, pp. 45265–45275, Dec. 2019, doi: 10.1021/acsami.9b14001.

[8] D. M. Correia *et al.*, “Electrosprayed poly(vinylidene fluoride) microparticles for tissue engineering applications,” *RSC Advances*, vol. 4, no. 62, pp. 33013–33021, 2014, doi: 10.1039/c4ra04581e.

Generative Adversarial Networks for Peripheral Cell Blood Images Standardization

A. Pérez[1], S. Alférez[2], K. Barrera[1], A. Merino[3] and J. Rodellar[1]

[1] Department of Mathematics, Barcelona Est Engineering School, Technical University of Catalonia, Barcelona, Spain, alejo.perez@estudiant.upc.edu, kevinivancho@hotmail.com, jose.rodellar@upc.edu

[2] Faculty of Natural Sciences, Universidad del Rosario, Bogotá, Colombia, edwin.alferez@urosario.edu.co

3 Biomedical Diagnosis Center, Hospital Clínic of Barcelona, amerino@clinic.cat

Abstract

Our multidisciplinary research group has been developing models based on Convolutional Neural Networks for the automatic morphological recognition of normal, reactive and abnormal blood cells circulating in blood, lymphoma and acute leukaemia.

The problem addressed in this study is related to the effect of different staining protocols used to prepare the blood smear from which microscope images are analyzed. Specifically, the automatic blood cell classification with models trained using an image database collected in a single hospital may be affected when images are coming from another hospital. Thus resulting in a certain number of misclassified images. Such accuracy decrease is because some color, staining and texture characteristics may suffer variations in the images depending on the staining protocol.

Generative Adversarial Networks (GANs) were leveraged to transform the images of the X domain (Hospital Germans Trias i Pujol) into the Y domain (Hospital Clínic). Characteristics of color, staining and texture of the latter were emulated. After this transformation, the classification accuracy of the images taken at Hospital Germans Trias i Pujol improved. CycleGAN and ColorizationGAN architectures were used for the transformations. ResNet 18, ResNet 34, ResNet 18 with Focal Loss and ResNet 34 with Focal Loss were used for the classifications. The classification accuracy improved from 34% to 81%.

1. Introduction

Morphological analysis of cells circulating in peripheral blood is a useful tool for the initial and quick detection of haematological and non haematological diseases. Visual inspection of blood cells on the stained smear using the microscope requires skills and expertise, being prone to subjectivity and inter-observer variability. Specialized automated microscopes are currently in use in specialized laboratories to help with such analysis by pre-classifying cells, such as CellaVision DM96 [1]. There exist a large number of pathologies, like leukemia, lymphoma and others, which are associated with the presence of specific classes of cells in peripheral blood. Morphological differences exist among those cell types but they are sometimes difficult to discern. In the last decade, researchers and practitioners have carried out a significant research effort to develop image-based machine learning algorithms to achieve automatic recognition in a more objective way and assist the clinical pathologists in their initial diagnosis [2, 3]. Deep learning has seen a recent tremendous interest in being applied to different problems involving automatic blood cell classifications [4, 5].

Generative Adversarial Networks (GAN) have been recently proved successful in allowing transformation procedures to improve classification models in biomedical problems. For instance Hu et al. [6] propose the use of these architectures as a complement to the unsupervised learning for the visual representation of histopathologies.

This work responds to the need to standardize the staining of the images obtained in different laboratories when developing and implementing machine or deep learning models. Our group has been working mainly with images obtained in the Core Laboratory of the Hospital Clínic of Barcelona (Clinic from now on). Therefore, training and validation of algorithms have been performed with data sets of images acquired with specific colorization depending on the staining used to prepare blood smears. Although smears are routinely performed by automatic stainers, they are prone to variability depending on the proportion of certain chemicals, times and local protocols. Therefore, classifiers based on convolutional neural networks (CNN) may decrease their effectiveness when tested with images from other laboratories.

This work explores whether using GAN can help to address this problem. The starting point is that CNN models have been trained with a data set of images from Clinic and we plan to test that models using images from a different center, concretely Hospital Germans Trias i Pujol (called HGTP from now on).

There are three specific objectives proposed in this work, which are directly related to three different methodology blocks:

1. Development of a preprocessing algorithm to transform the images of HGTP emulating the staining characteristics of the Clinic images with CycleGAN.
2. Assessment of the previous process through classification of the HGTP images (original and transformed ones) by CNNs pretrained with Clinic images. Moreover, qualitative evaluations of the transformation by clinical pathologists.

ISBN: 978-84-09-25491-0, pags. 157- 160

3. Artificial generation of abnormal peripheral lymphocytes out of normal lymphoid blood images (and vice versa) by means of CycleGAN. Those images belong to the Hospital Clínic.

2. Materials

As for Block 1, three different datasets were used to train the CycleGAN to transform the HGTP images: The Clinic training set (3,000 images), the Clinic testing set (184 images), and the HGTP dataset, used for both training and testing (due to the scarcity of the set) containing 184 images. The Clinic training and testing sets consisted of a random sampling of the Clinic database containing all reactive, abnormal lymphocytes and leukaemic cells (blasts). To train the CycleGAN, the specific number of each class was not taken into account because just changes in color and style were intended. The images were labeled with the categories mentioned before but just the general labels of HGTP and Clinic were relevant in the transformation process. Then, they were taken into consideration in the classification stage.

In the Second Block (classification) a sampling from the Clinic database was used to fine-tune the classifiers (CNN) prior to the classification of the HGTP images (both original and transformed). This data distribution is outlined in Table 1. We excluded the normal lymphocyte class as the HGTP dataset lacks this category. The HGTP dataset was split in training and testing to ensure the network does not see the same images both in training and testing, see Table 2.

Dataset	Validation	Testing	Training	Total per category
Abnormal Lymphocyte	243	69	1026	1338
Reactive Lymphocyte	112	102	439	653
Blast	227	93	866	1186
Total per set	582	264	2331	3177

Table 1. *Clínic Validation, Training and Testing Set*

Dataset	Training	Testing	Total per category
Abnormal Lymphocyte	22	27	47
Reactive Lymphocyte	27	30	57
Blast	43	37	80
Total per set	92	92	184

Table 2. *HGTP Training and Testing Set for Classification*

For Block 3, aiming to transform normal lymphocytes to abnormal lymphocytes and vice versa, a sampling dataset as shown in Table 3 was used. It was split in 3 subsets for each training and testing experiment. This was intended for performing three different experiments: confronting normal lymphocytes against lymphocytes circulating in blood in splenic marginal zone lymphoma (SMZL), then against mantle cell lymphoma cells (MCL) and finally against blasts.

	Abnormal Lymphocyte		Normal Lymphocytes	
Dataset	Training	Testing	Training	Testing
Normal Lymphocyte vs SMZL	1,000	583	799	287
Normal Lymphocyte vs MCL	800	381	800	286
Normal Lymphocyte vs Blast	1,040	200	987	100

Table 3. *Training and Testing sets for Block 3*

Regarding the network architectures, they were designed, trained, validated and tested using the *fast.ai* library (built on top of *Pytorch*) and *Python* programming language. It was also used with a DeepBox with a *GPU Nvidia Titan XP* (12 GB).

3. Methodology

3.1. Block 1 - Transformation of HGTP images by CycleGAN

Firstly, the datasets referent to the First Block were arranged to be introduced to the CycleGAN. As shown in Figure 1, both Clinic and HGTP training sets are confronted one against the other to train the CycleGAN during 200 epochs. We chose the parameter values intended for training similar dimensions datasets recommended by Jun-Yan Zhu, main developer of the architecture. These can be seen in its Github repository [7]. Mini-batches of 1 image were selected as well as a load size of 286 pixels and a crop size of 256 pixels. A number of 64 filters was chosen for both the last discriminator layer and the first generator layer.

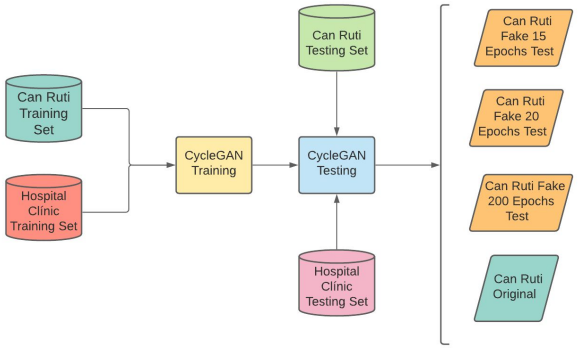


Figure 1. *Diagram of Transformation of HGTP images by CycleGAN*

After that, the testing sets were introduced to perform the testing step (transformation). The transformation was carried out bidirectionally although just the HGTP resulting datasets were kept. The CycleGAN model trained was tested at 10, 15, 20, 25 and 200 epochs. Every epoch represents one time for the model to have seen the whole training set and update its parameters. This multiple epoch testing was done because we noticed that the more epochs

the network was trained, the transformations involved more cell morphology structural changes. Therefore, it was necessary to select the epoch testing that emulates best the targeted style features. As visible in Figure 1, just 15 and 20 epoch generated datasets were selected because they complied with the Clinic standards. The transformation at 200 epochs was kept to show the resulting images when they underwent a transformation of morphology due to too many epochs in the training stage.

3.2. Block 2 - Transformed Cell Images Classification

The Block 2 procedure is outlined in Figure 2. We used ResNets with 18 and 34 layers pre-trained on the Imagenet Dataset [8] as well as a model provided by Clinic pre-trained with a bigger dataset than the used in this study (called ResNet34_Resample_NoSIND, from now on ResNet34_NoSIND for simplicity).

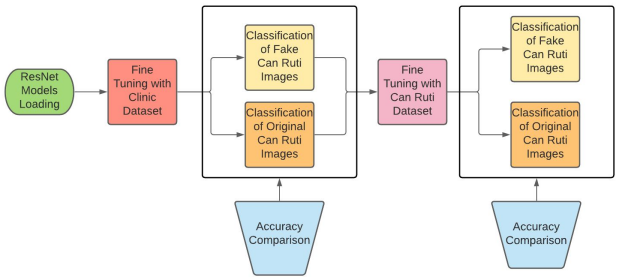


Figure 2. *Diagram of Transformed Cell Classification*

A Focal Loss function was added to the first two architectures, proven to improve the training convergence in unbalanced datasets [9]. Therefore, two new custom architectures were produced: Resnet18 Focal and Resnet34 Focal. The models (except ResNet34_NoSIND) were fine-tuned with the Clinic dataset distribution. We used the 1 cycle policy [10] to select the learning rate at the fine-tuning process (for the "Fine-tuning with Clinic Dataset" and "Fine Tuning with HGTP Dataset" stages). The HGTP Dataset was split in two (training and testing). After that, a test classification of the Fake HGTP images and the Original HGTP images was carried out (in the testing set). Then, a fine-tuning with this data distribution was performed. A posterior classification was done again with the aforementioned HGTP testing dataset after this fine-tuning. Finally, accuracy metrics were registered in both classification stages.

3.2. Block 3 - Image Cell Transformation by CycleGAN (Normal to Abnormal Lymphocyte and vice versa)

The Block 3 procedure is outlined in Figure 3. Three training experiments were done in parallel following the Block 1 same procedure for the CycleGAN training (200 epochs training).

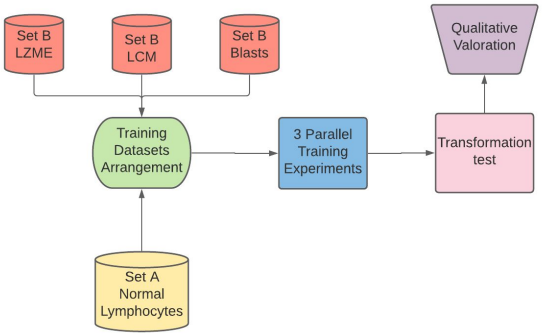


Figure 3. *Diagram of the Third Block, Transformation by CycleGAN between classes*

4. Results and Discussion

4.1. Results of Block 1 - Transformation by CycleGAN

As shown in Figure 4, after a qualitative inspection the best result was the transformation at 20 epochs of CycleGAN training. In the case of the 200 epoch transformation, it is visible that the transformations at 200 epochs of training suffered undesired morphological transformations, please see arrows in Figure 4.

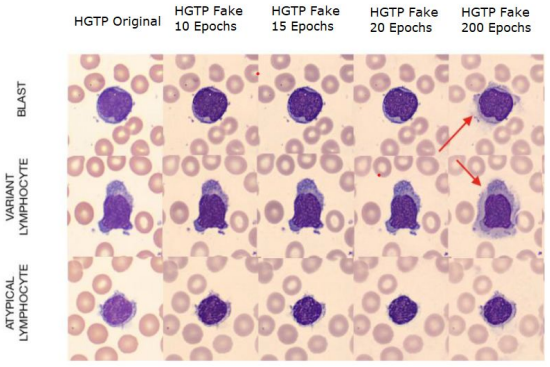


Figure 4. *HGTP Fake Dataset generation result by CycleGAN*

4.2. Results of Block 2 - Transformed Cell Classification

As shown in Table 4, there are classifications whose result did not exceed 0.5 of accuracy, representing a poor performance of the classifier model (executions in the original HGTP dataset). The accuracies in bold are those that have exceeded 0.6.

Dataset	HGTP Original	HGTP Fake 15 Epochs	HGTP Fake 20 Epochs
ResNet_18	0.3152	0.5869	0.5326
ResNet_18_Focal	0.4565	**0.6847**	0.5869
ResNet_34	0.2826	0.5543	0.5326
ResNet_34_Focal	0.3043	**0.6739**	0.5326
ResNet_34_No_Sind	0.3478	**0.8152**	**0.75**

Table 4. *Accuracy results of automatic classification of the* HGTP *Datasets (Fake ones and the Original)*

The transformation at 15 and 20 epochs showed a significant improvement in the performance of the model with respect to the classification in the Original Dataset. It is also shown in Figure 5 the normalized confusion matrix of the best test result without evident morphological changes (test with ResNet_34_No_Sind with the HGTP Fake CycleGAN dataset at 15 epochs).

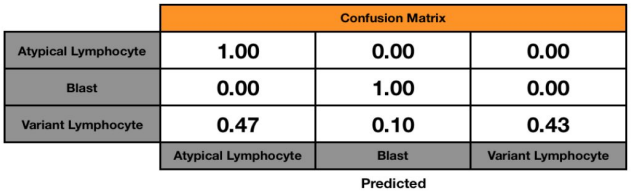


Figure 5. *Normalized Confusion Matrix of the best classification test result. Tested with the ResNet_34_No_Sind with the dataset Can Ruti Fake (15 epochs of training with CycleGAN)*

After having fine-tuned again the CNN models with a training set of HGTP dataset (for the three Fake HGTP ones and the Original) all accuracy results reached at least 0.9 for all HGTP datasets (Originals and Fake).

4.3. Results of Block 3 - Cell Transformation between Classes

In the Third Block, the image transformation from normal cells to other types of malignant cells has been possible thanks to the CycleGAN architecture. The conversion has also been performed in reverse and has been generally successful in both directions.

5. Conclusions

Regarding Blocks 1 and 2, the pre-processing algorithm to transform the HGTP into the Clinic standards has been successfully created. Classification accuracy has been used to determine the most effective image transformation technique is the CycleGAN usage with 15 epochs of training. The ResNet34_NoSIND model has achieved an accuracy of 81.52% in the dataset resulting from that transformation, increasing approximately 60 percentage points of accuracy with respect to the classification of the Original HGTP dataset. In the next stage, the high effectiveness of the fine-tuning has been proven being carried out in the training sets of the HGTP datasets, both in the Fake ones and in the Original. This is an alternative to preprocessing using CycleGAN in case the latter is not available. By means of this method, performances exceeding 90% have been achieved with all the models in the original dataset, even reaching 98.91%.

In matters of the Third Block, it has been possible to make a comparison and qualitative assessment of the cell transformations from Normal to Abnormal Lymphocytes and vice versa. The possibility of transforming blood cell images from one type to another opens up the possibility of training classifiers with images of malignant cells corresponding to rare diseases. Obtaining an adequate number of images containing this type of cells from low prevalence diseases is very difficult, so the use of CycleGAN architecture for their acquisition can be very useful.

It should be pointed out that an increase in the size of the training datasets of this study would lead to greater precision and quality of the training-testing performance of both CycleGAN and CNNs.

Acknowledgments

We are grateful to Cristian Morales-Indiano for providing the HGTP images. This work is part of a research project funded by the Ministry of Science and Innovation of Spain, with reference PID2019-104087RB-I00.

References

[1] A. Merino, R. Brugués, R. García, M. Kinder, F. Torres, and G. Escolar, "Estudio comparativo de la morfología de sangre periférica analizada mediante el microscopio y el CellaVision DM96 en enfermedades hematológicas y no hematológicas", Revista del Laboratorio Clínico, vol. 4, no. 1, pp. 3–14, Jan. 2011, doi: 10.1016/j.labcli.2010.10.002.

[2] J. Rodellar, et al. "Image processing and machine learning in the morphological analysis of blood cells", Int. Journal of Laboratory Hematology, Vol. 40, pp. 46-53, 2018.

[3] S. Alférez, et. al. "Automatic recognition of atypical lymphoid cells from peripheral blood by digital image analysis", American Journal of Clinical Pathology, Vol. 143, pp. 168-176, 2015.

[4] J. Zhao, M. Zhang, Z. Zhou, J. Chu, and F. Cao, "Automatic detection and classification of leukocytes using convolutional neural networks," Med Biol Eng Comput, vol. 55, no. 8, pp. 1287–1301, 2017, doi: 10.1007/s11517-016-1590-x. [Online]. Available: http://link.springer.com/10.1007/s11517-016-1590-x.

[5] A. Acevedo et al. "Recognition of peripheral blood cell images using convolutional neural networks", Comp Methods and Programs in Medicine, Vol. 180, 2019, https://doi.org/10.1016/j.cmpb.2019.105020.

[6] B. Hu, Y. Tang, E. I.-C. Chang, Y. Fan, M. Lai, and Y. Xu, "Unsupervised Learning for Cell-Level Visual Representation in Histopathology Images With Generative Adversarial Networks," IEEE J. Biomed. Health Inform., vol. 23, no. 3, pp. 1316–1328, May 2019, doi: 10.1109/JBHI.2018.2852639. [https://ieeexplore.ieee.org/document/8402089/.

[7] J. P. Zhu, T. Isola, and P. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," Unpaired image-to-image translation using cycle-consistent adversarial networks.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Commun. ACM, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386. [Online]. Available: https://dl.acm.org/doi/10.1145/3065386.

[9] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection," arXiv: 1708.02002, Feb.2018.http://arxiv.org/abs/1708.02002.

[10] G. Sylvain, "The 1cycle policy," Github. Apr2018 https://sgugger.github.io/the-1cycle-policy.html

Effect of cardiac stem cell derived extracellular vesicles on cardiomyocytes

M. Moro López[1,2*], L. Gómez Cid[1,2,3], A. S. de la Nava[2,3], M. E. Fernández Santos[2,3], L. Grigorian Shamagian[2,3] and F. Fernández Avilés[2,3]

[1] Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Leganés, Spain
[2] Instituto de Investigación Sanitaria Gregorio Marañón, Hospital Gregorio Marañón, Madrid, Spain
[3] CIBERCV, ISCIII, Madrid, Spain
* Correspondence: marina.moro@alumnos.uc3m.es

Abstract

Cell and extracellular vesicle-based therapies have shown antiarrhythmic potential but have not been extensively studied. Hence, in this work it is investigated the influence of extracellular vesicles secreted by atrial cardiosphere-derived cells (CDC-EVs) on cardiomyocytes. The HL-1 cell line was used, adding CDC-EVs after 48h of culture. When completely confluent, cell cultures were analyzed with optical mapping and posteriorly with MATLAB software for the measurement of the main electrophysiological properties: dominant frequency, complexity, curvature, rotor meandering and conduction velocity (CV). Then, a two-step RT-qPCR was performed to quantify the expression of genes involved in the cardiac impulse propagation (SCN5A, CACNA1C and GJA1, coding for the sodium and calcium channels and connexin 43, respectively). As significant results, CV and meandering increased and complexity and curvature decreased after the CDC-EVs administration, supported by an increase in the expression of the three genes under study. This suggests that the CDC-EVs treatment induces a less arrhythmogenic profile driven by changes in the gene expression, increasing CV and modifying rotor dynamics so complexity is decreased. In spite of having an in vitro nature and working with non-human cells, this study opens new insights into the antiarrhythmic potential of CDC-EVs.

1. Introduction

Atrial fibrillation (AF) is the most common sustained arrhythmia with a global prevalence of 2%, entailing a five-fold higher risk of stroke and a two-fold higher risk of death [1]. AF implies rapid and anarchical atrial activity. Once it is initiated, the atrial tissue suffers remodeling at the electrical, anatomical and autonomic levels, causing conduction disturbances that lead to the emergence of arrhythmic activation patterns; these are called rotors, which are spiral wave reentries that rotate around a singularity point [2]. They are associated with low conduction velocity values and short action potential (AP) duration, as cardiomyocytes can recover the excitability during the AP propagation to nearby cells [3]. Conduction velocity is dictated by gap junctions (made of connexins), as they allow the ion passage between cells, and by the sodium channels (I_{Na} current), which control the cellular membrane depolarization; on the other hand, the duration is driven by the calcium channels (I_{CaL} current), which govern the plateau phase of the AP [4].

Present treatments for AF aim to control the atrial and ventricular rhythms, either pharmacologically or ablating the area that triggers the fibrillatory activity, but these techniques lack efficiency and imply serious side effects, both cardiac and non-cardiac [5]. There are no clinical treatments directed to the maintenance mechanisms of AF themselves [6]. In terms of biological therapies, in particular cell and extracellular vesicle-based approaches, their use for regenerative purposes is increasingly being investigated due to their great potential [7].

Following this line of research, the present study aims to study the effect of extracellular vesicles (EVs) secreted by cardiosphere-derived cells (CDCs), a type of cardiac stem cells which have shown cardioprotective function [8], on HL-1 cardiomyocytes, which present fibrillatory activity *in vitro* [9]. The purpose is to test if human CDC-EVs have an effect on these cells, leading to a potential approach for antiarrhythmic therapy.

2. Materials and Methods

2.1. CDCs and extracellular vesicles isolation

The used EVs were extracted from CDCs from pediatric donors following the protocol described by Smith *et al*. [8]. As a summary, atrial biopsies from patients undergoing cardiac surgery were digested as millimetric explants and then cultured in Iscove's Modified Dulbecco's Medium (IMDM) to obtain explant-derived cells (EDCs). These EDCs were transferred to ultra-low attachment Nunclon Plates so they formed 3D spheroids (cardiospheres), which were passaged to produce CDCs. CDCs' identity was confirmed by flow cytometry and they were changed to Fetal Bovine Serum (FBS)-free IMDM for 15 days so EVs were secreted. The EVs content in the culture media was quantified using Bradford assay and precipitated with polyethylene glycol (4% w/v in PBS) at 4º and further centrifugation for 30 minutes at 4º and 1,500xg. EVs were finally resuspended in FBS-free Claycomb Medium (final concentration 100 µg/ml) for the experiments.

2.2. Experimental protocol

The overall experimental protocol is illustrated in **Figure 1**. 26,000 HL-1 cells/cm^2 were seeded at day 0 on 35-mm-diameter Petri dishes and cultured according to the

ISBN: 978-84-09-25491-0, pags. 161- 164

standard HL-1 protocol [9]. Cell cultures were changed to FBS-free media after 48h, resulting in a control group and a CDC-EVs group (with EVs resuspended in the culture media, 20 µg/cm^2 doses). At day 4, optimal confluency was reached so cultures were analyzed with optical mapping (N = 12 in control and N = 12 in CDC-EVs group). Then, the gene expression analysis was performed to contrast the electrophysiological results with cardiac remodeling at the molecular level.

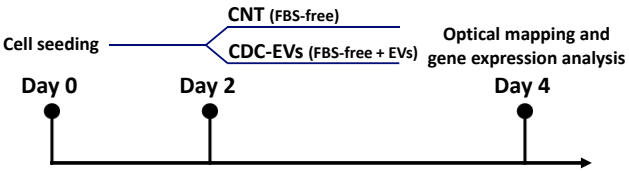


Figure 1. *Experimental protocol. HL-1 cells were seeded and changed to FBS-free media with or without CDC-EVs after 48h. At day 4, cell cultures were analyzed.*

2.3. Optical mapping

Cell cultures were treated with a solution with the calcium-sensitive dye Rhod-2 AM (AAT Bioquest; final concentration 3.3 µM, from 1 mM aliquots in dimethyl sulfoxide) and Probenecid (AAT Bioquest; final concentration 1.75 mM, from 265 mM aliquots in Tyrode) diluted in FBS-free media for 30 minutes at incubation conditions so the propagation patterns could be recorded by means of the calcium transients. Prior to the optical mapping, the media was replaced with pre-warmed Tyrode (containing, in mM, NaCl 130, $NaHCO_3$ 24, NaH_2PO_4-H_2O 1.2, $MgCl_2$ 1, KCl 4, glucose 5.6 and $CaCl_2$ 2.3).

Cell cultures were illuminated with a green light-emitting diode with a plano-convex lens and a green excitation filter, and fluorescence was recorded with an electron-multiplying charge-coupled device with a custom emission filter and a high-speed camera lens, as defined by Climent *et al*. [10]. Movies with a duration of 20 seconds at 50 frames per second and a field of view of ~2x2 cm^2 were acquired. For the velocity measurement, cell cultures were delimited in rectangular areas and stimulated with pulses of 37V amplitude, 4 ms width and 1 Hz pacing period.

2.4. Electrophysiological analysis

The optical mapping recordings were processed as illustrated in **Figure 2** with custom MATLAB software to quantify dominant frequency (DF), complexity, curvature, rotor meandering and conduction velocity (CV) as described by Climent *et al*. [10].

Power spectra of the samples were obtained applying a Welch periodogram (2-second Hamming window overlap) to the recordings, considering the DF for each pixel as the largest peak in its power spectrum and the DF for each culture as the maximum DF value.

Phase maps were acquired using the Hilbert transform. Phase singularities (PS) were defined as points surrounded by phases from 0 to 2π monotonically changing. Complexity was defined as the mean number of PS in a defined area (i.e. PS/cm^2). From the PS map, lines connecting phase transitions from 0 to 2π were traced from the rotor tip to the periphery. The curvature of the rotor was the mean of the curvatures at each point, which was estimated as the derivative of the relative angle (α) over the relative distance (δ) (i.e. rad/cm). Rotor meandering was defined as the distance covered by the tip of the rotor over PS duration (i.e. cm/s). From the phase map, the path followed by the tip of the PS was traced and quantified, determining the meandering value for each culture as the mean value of the meandering of all the PS.

CV was defined as the covered distance by the propagation front (i.e. cm/s) in unidirectional propagation areas. The signals of two pixels at a known spatial distance were correlated to obtain the time difference between their activation instants (Δt). CV was calculated as the distance between the pixels over the elapsed time, being the CV value for each culture the mean of the CV of at least eight delimitated areas.

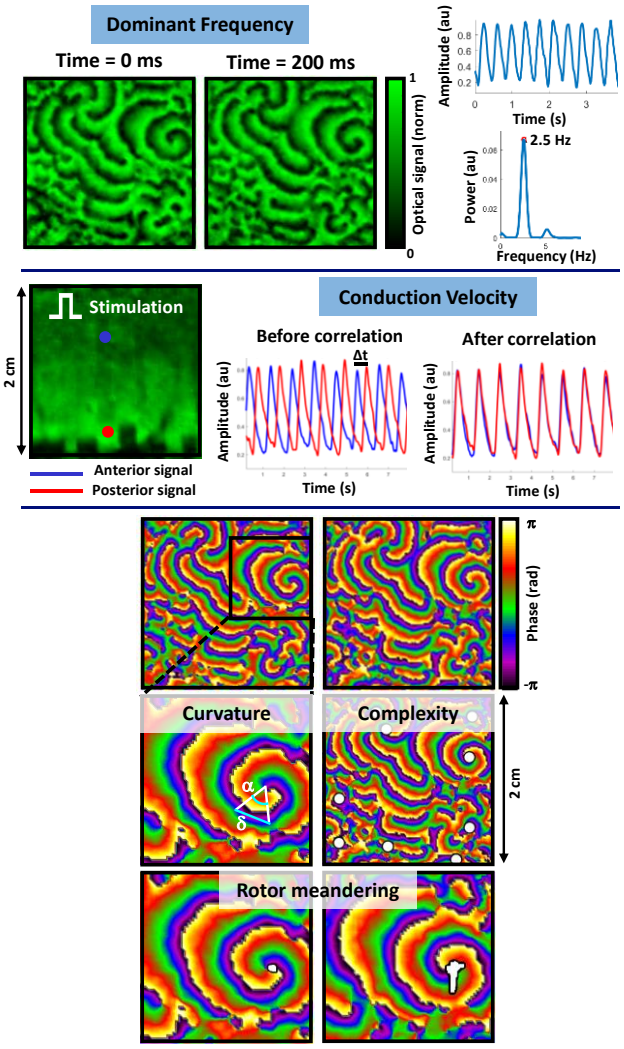


Figure 2. *Electrophysiological analysis. Two snapshots of a recording are shown at different time instants with an example of signal for a pixel and its power spectrum with the DF value in red (upper). For CV two pixels were selected and their signals were time-correlated (middle). Phase maps were obtained to calculate complexity, curvature and rotor meandering (lower).*

2.5. Gene expression analysis

RNA was extracted from the samples using QIAzol Lysis Reagent (QIAGEN, Venlo, The Netherlands) and relative gene expression quantification was performed with a two-

step RT-qPCR. RNA was reverse transcribed into complementary DNA (cDNA) using the iScript™ cDNA Synthesis Kit following the manufacturer's instructions. The quantification was performed in a CFX Real-Time PCR Detection System (Bio-Rad Laboratories) using 96-well plates with cDNA, specific primers (**Table 1**) and SYBR Green chemistry (final volume 20 µl), analyzing each sample per duplicate. Ct data was analyzed applying the $2^{-\Delta\Delta Ct}$ method, using *GUSB* and *36B4* as reference genes. The genes under study were *SCN5A*, *CACNA1C* and *GJA1*, encoding for proteins related to the cardiac impulse propagation and involved in the voltage-activated sodium current (I_{Na}), L-type calcium current (I_{CaL}) and connexin 43 (Cx43) channels, respectively.

Gene	Forward and reverse primers (5'-3')
SCN5A	CACCTTCACCGCCATCTACA
	AAGGTGCGTAAGGCTGAGAC
CACNA1C	CCTCGAAGCTGGGAGAACAG
	TGTGTGGGAGTCAATGGAGC
GJA1	CCGGTGTGAAAATGTCTGCT
	CGTGAGCCAAGTACAGGAGT
GUSB	TGCTCTAGACTTCGAGCAG
	AAGAGCCTCAGGGCATCGG
36B4	GCGACCTGGAAGTCCAACTA
	ATCTGCTGCATCTGCTTGG

Table 1. *RT-qPCR primers*

2.6. Statistical analysis

Results are presented as mean ± standard deviation in the text and figures. Continuous variables were compared using Student's t-tests, considering a p-value of 0.05 or less as statistically significant.

3. Results

3.1. Electrophysiological analysis

Figure 3 shows the comparison between the electrophysiological results of the control and CDC-EVs groups. No significant differences were observed during microscopic observation of the samples after the CDC-EVs administration and spontaneous beating was appreciated for both groups during cell expansion.

Complexity decreased in ~40% after the EVs treatment (1.54 ± 0.20 vs. 2.60 ± 0.57 PS/cm^2, $p < 0.01$), finding an evidently greater presence of depolarization fronts and wavebreaks in control samples, while for CDC-EVs samples rotors are wider, more disperse and not so fragmented. Likewise, curvature was decreased as well for CDC-EVs samples (1.95 ± 0.45 vs. 0.92 ± 0.23 rad/cm, $p < 0.01$), being the rotors after the CDC-EVs administration less tight and covering greater areas when dispersing. On the other hand, rotor meandering increased in CDC-EVs samples (1.45 ± 0.70 vs. 4.48 ± 1.95 cm/s, $p < 0.01$), as the trajectories followed by the rotor tips of treated cultures were wider and more irregular. DF was similar for both groups although slightly higher for CDC-EVs samples (3.61 ± 1.44 vs. 4.69 ± 2.15 Hz, not significant). Finally, CV was 80% greater after the CDC-EVs administration (1.43 ± 0.53 vs. 2.62 ± 0.75 cm/s, $p < 0.01$). This difference can be appreciated in the isochrone maps, as CDC-EVs samples covered 2 cm within ~600 ms, while control ones took ~1200 ms.

3.2. Gene expression analysis

A general increase was observed in the expression of the three genes under study after the CDC-EVs administration (**Figure 3**) (*SCN5A* 1 ± 0.77 vs. 3.27 ± 2.36, $p < 0.05$; *GJA1* 1 ± 0.65 vs. 2.63 ± 1.19, $p < 0.05$; *CACNA1C* 1 ± 0.39 vs. 1.73 ± 1.37, not significant).

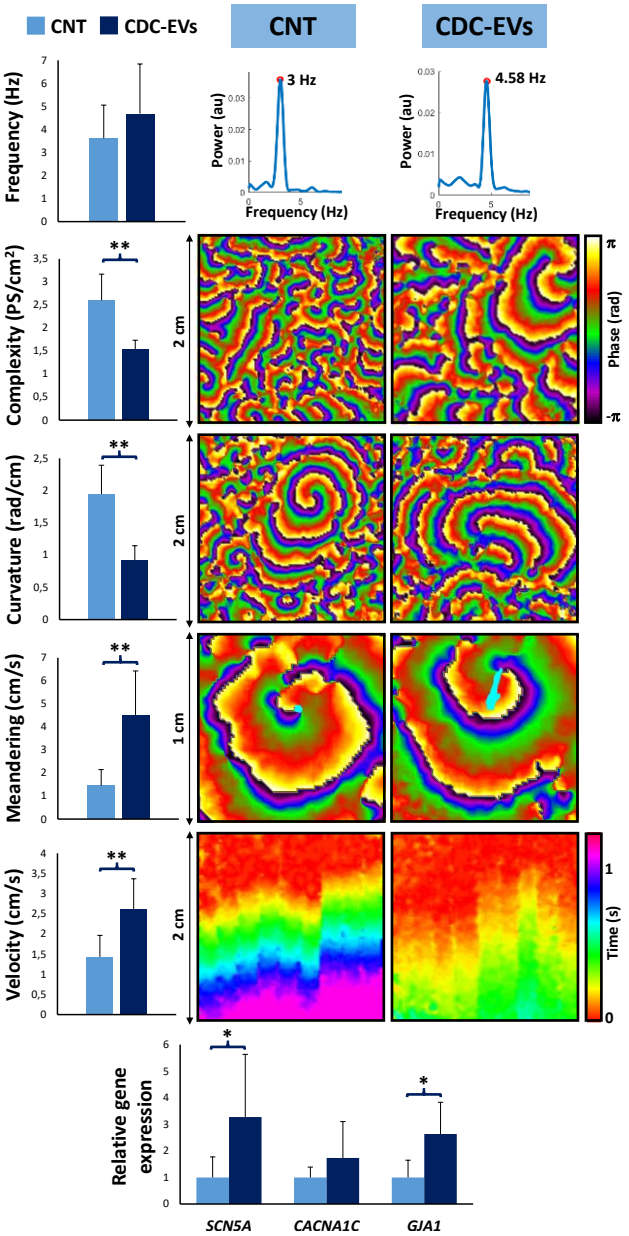


Figure 3. *Results and representative examples of CNT and CDC-EVs samples.*

4. Discussion

The major finding of this study is that CDC-EVs reduce the activation complexity of HL-1 cells, driven by modifications in the CV and rotor dynamics supported by changes in the gene expression.

CV increased after the CDC-EVs administration. Low CV values have been associated with AF maintenance, as cardiomyocytes are capable of recover the excitability

during the AP propagation [11], and with a low and heterogeneous expression of Cx43 (*GJA1* gene) [12]. On the other hand, high sodium channels expression (*SCN5A* gene) has been linked to faster depolarization and therefore higher CV [13]. Hence, the obtained increase in the CV for CDC-EVs samples seems to be driven by the higher sodium channels and Cx43 expressions and supposes a conduction substrate with lower reentry recurrence. In addition, the not significant increase in DF could be a collateral effect of this higher CV.

Regarding rotor meandering and curvature, CDC-EVs samples showed wider rotors that covered larger areas accounting for less depolarization front breaks and with more extended tip movements, implying an activation pattern with higher instability and that needs a larger area for rotor self-sustaining. For these reasons, the increase in meandering and decrease in curvature have been associated with AF termination [14]. Besides, they have been related to the AP duration (APD), which is shorter under AF conditions and governed by a low expression of the calcium channels (*CACNA1C* gene) [15], so the rotor dynamics results seem to be driven by the increase in the calcium expression and may be caused by a longer APD, implying a less probable AF maintenance because of rotor instability and giving place to lower complexity values.

CDCs have been tested under ventricular arrhythmia and heart failure scenarios, resulting in a decrease in fibrosis and inflammation and an increase in the heart function by upregulating the potassium channels and decreasing and homogenizing the APD [16]. The results of this study are complementary to these findings and illustrate that the antiarrhythmic effect of CDC-EVs is mediated by an increase in CV and modifications of rotor dynamics, offering an innovative knowledge about the electrophysiological response after a CDC-EVs treatment.

Although the HL-1 cell line is widely used as a study model of cardiac electrophysiology and arrhythmias because of its spontaneous fibrillatory activation [17], it has several drawbacks. HL-1 cells are mature but of murine origin and present differences with adult human cardiomyocytes. In addition, the *in vitro* conditions do not resemble the real environment of a human heart, as the properties derived from the 3D arrangement are missing. Considering these limitations, the effect of CDC-EVs should be verified in other cell models and human samples as future work.

5. Conclusions

The use of CDC-EVs on HL-1 cardiomyocytes reduce their spontaneous activation complexity by increasing the CV and modifying the rotor dynamics, leading to a less arrhythmogenic profile. If validated in other cell models, this approach presents great potential and may open new insights into a promising therapeutic alternative for tachyarrhythmias as AF.

6. Acknowledgments

Very special thanks to Lidia Gómez for this opportunity and her unconditional support. Thanks also to my team at the Instituto de Investigación Sanitaria Gregorio Marañón.

References

[1] Smid J, Rokyta R. Atrial fibrillation and its relation to cardiac diseases and sudden cardiac death. *Cor et Vasa*, vol 59, sup 4, 2017, pp e325-e331 (ISSN: 1803-7712).

[2] Schotten U, Verheule S, Kirchhof P, Goette A. Pathophysiological mechanisms of atrial fibrillation: A translational appraisal. *Physiological Reviews*, vol 91, sup 1, 2011, pp 265-325 (ISSN: 1522-1210).

[3] Antzelevitch C, Burashnikov A. Overview of Basic Mechanisms of Cardiac Arrythmia. *Cardiac Electrophysiology Clinics*, vol 3, sup 1, 2011, pp 23-45 (ISSN: 1877-9182).

[4] Nattel S, Burstein B, Dobrev D. Atrial remodeling and atrial fibrillation: mechanisms and implications. *Circulation: Arrhythmia and Electrophysiology*, vol 1, sup 1, 2008, pp 62-73 (ISSN: 1941-3084).

[5] Forleo GB. Present concepts in management of atrial fibrillation: From drug therapy to ablation. *World Journal of Cardiology*, vol 1, sup 1, 2009, p 11 (ISSN: 1949-8462).

[6] Haugan K *et al*. Rotigaptide improves atrial conduction slowing in chronic volume overload-induced dilated atria. *Basic and Clinical Pharmacology and Toxicology*, vol 99, sup 1, 2006, pp 71-79 (ISSN: 1742-7835).

[7] Bellin G *et al*. Exosome in Cardiovascular Diseases: A Complex World Full of Hope. *Cells*, vol 8, sup 2, 2019, p 166 (ISSN: 2073-4409).

[8] Smith RR *et al*. Regenerative potential of cardiosphere-derived cells expanded from percutaneous endomyocardial biopsy specimens. *Circulation*, vol 115, sup 7, 2007, pp 896-908 (ISSN: 0009-7322).

[9] Claycomb WC *et al*. HL-1 cells: A cardiac muscle cell line that contracts and retains phenotypic characteristics of the adult cardiomyocyte. *Proc. National Academy of Sciences*, vol 95, sup 6, 1998, pp 2979-2984 (ISSN: 0027-8424).

[10] Climent AM *et al*. Role of atrial tissue remodeling on rotor dynamics: an in vitro study. *American Journal of Physiology – Heart and Circulatory Physiology*, vol 309, sup 11, 2015, pp H1964-H1973 (ISSN: 1522-1539).

[11] Pellman J, Sheikh F. Atrial fibrillation: Mechanisms, therapeutics and future directions. *Compr. Physiology*, vol 5, sup 2, 2015, pp 649-665 (ISSN: 2040-4603).

[12] Thibodeau IL *et al*. Paradigm of genetic mosaicism and atrial fibrillation: Physiological characterization of a connexin 43-deletion mutant identified from atrial tissue. *Circulation*, vol 122, sup 3, 2010, pp 236-244 (ISSN: 0009-7322).

[13] King J *et al*. Determinants of myocardial conduction velocity: Implications for arrhythmogenesis. *Frontiers in Physiology*, vol 4, 2013, p 154 (ISSN: 1664-042X).

[14] Comtois P, Kneller J, Nattel S. Bridging the gap between the leading circle and spiral wave concepts of cardiac reentry. *Europace*, vol 7, sup 2, 2005 (ISSN: 1099-5129).

[15] Van Wagoner DR, Pond AL, Lamorgese M, Rossie SS, McCarthy PM, Nerbonne JM. Atrial L-Type Ca^{2+} Currents and Human Atrial Fibrillation. *Circulation Research*, vol 85, sup 5, 1999, pp 428-436 (ISSN: 0009-7330).

[16] Cho JH *et al*. Reverse electrical remodeling in rats with heart failure and preserved ejection fraction. *JCI insight*, vol 3, sup 19, 2018 (ISSN: 2379-3708).

[17] Van Gorp PRR *et al*. Multicellular In vitro Models of Cardiac Arrhythmias. *Frontiers in Cardiovascular Medicine*, vol 7, 2020, p 43 (ISSN: 2297-055X).

Design, construction and validation of a low-cost device to measure maximal inspiratory and expiratory pressures

C. Aymerich Moreno, M. Rodríguez Lázaro, J. Otero Díaz, R. Farré Ventura

Biophysics and Bioengineering Unity, School of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain, claudiaaymerich@gmail.com, m.rodriguez@ub.edu, jorge.otero@ub.edu, rfarre@ub.edu

Abstract

Measurement of maximal inspiratory (MIP) and expiratory pressures (MEP) is the most widely used technique to non-invasively evaluate respiratory muscle strength in clinical practice. Moreover, it is thought that MIP and MEP could serve as accurate markers for respiratory disease prognostics, but more evidence is needed. Commercial devices that perform this test range from whole body plethysmographs to portable spirometers, both expensive and covering a wide range of respiratory tests. Given that a portable, low-cost and specific option for MIP and MEP measuring device is not available right now in the market, a high-performance and easy-to-build prototype has been developed to fill the gap. The novel device has been based on an Arduino microcontroller with a digital display, an integrated pressure transducer, and a 3D printed case (total retail cost 80 €). The validation of the device was performed by comparison with a laboratory reference setting, and results showed an accuracy within ±1%. As the device design will be published as open source, free replication and use in low-resources countries will be facilitated. Moreover, the availability of the device can contribute to obtain further evidence on respiratory disease prognostic markers.

1. Introduction

1.1. MIP and MEP test

MIP/MEP measurement is an easy, non-invasive and rapid test to measure the strength of the respiratory muscles [1]. MIP is the maximum negative pressure that can be generated by forced inspiration, and it is generated by the maximum contraction of diaphragm and intercostal muscles which tend to increase the volume of the rib cage and consequently lung volume. MEP is the maximum positive pressure that can be generated on forced expiration [2], when the abdominal muscles push the diaphragm and the internal intercostals up, thus tending to reduce the thorax and lung volumes. Therefore, MIP is an indicator of the strength of the diaphragm and the intercostal muscles and MEP measurement assess the strength of abdominal and internal intercostal muscles.

This test is a routine procedure in the diagnosis of certain pulmonary diseases. It is mainly performed in patients with suspected respiratory muscle weakness and is of great value for excluding important respiratory muscle weakness. Some examples of prevalent diseases which alter MIP and MEP values are Chronic Obstructive Pulmonary Disease (lower MIP), neuromuscular diseases such as multiple sclerosis (lower MEP), or chronic heart failure (significantly decreased MIP) [2,3].

The standard test is a volitional assessment that is highly dependent on the patient's performance. It also depends on the methodology, the device used, and the training of the professionals who supervise it. Therefore, the American Thoracic Society (ATS) and the European Respiratory Society (ERS) agreed to establish a standard protocol, which was published in 2002 [4] and was updated by the ERS in 2019 [5].

MIP/MEP measurements are obtained in a patient's seated position. Maximum inspiratory (Mueller manoeuvre) and maximal expiratory (Valsalva manoeuvre) measurements should be supervised by an experienced technician, who must urge the patient to exert his/her maximal effort. Both forced inspiratory and expiratory efforts should ideally be maintained for at least 1.5 seconds, allowing that maximum pressure can be measured continuously for 1 second. The MIP/MEP device should give a visual feedback of the patient manoeuvre by displaying of the pressure-time curves and the 1-sec measured pressure value. The technician determines whether the manoeuvre has been satisfactorily performed and the maximum value of three correct performances that vary by less than 10% is recorded. It is also stated that, to perform a reliable test, at least 5 efforts should be done for achieving peak values for MEP and 9 for MIP. Regarding test results, the ATS and the ERS indicate in their protocol of 2002 how a typical standard pressure trace of MIP and MEP looks like (Figure 1).

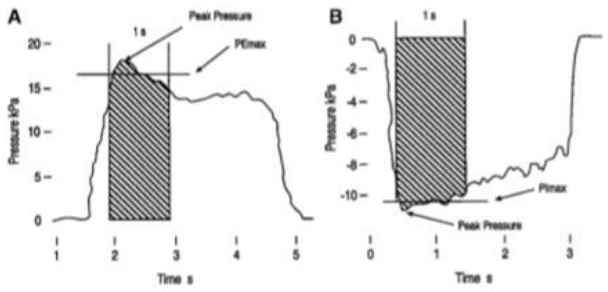


Figure 1. *(A) Typical pressure tracing for MEP. (B) Typical pressure tracing for MIP*

Regarding reference normal values, ATS/ERS agree that MIP values below -80 cmH_2O usually exclude the possibility of weakness and that values less negative than -80 cmH_2O (hence lower MIP) are difficult to interpret.

ISBN: 978-84-09-25491-0, pags. 165- 168

Therefore, according to ATS and ERS, there is a range of standard normal values but the LLN (Lower Limit Normal), which is the parameter of interest in deciding if a patient has a weak respiratory muscle, is not completely established. Other studies have analysed the normal values of these parameters. As an example, the book "Clinical Respiratory Medicine" establishes that normal values are less than -80 cmH_2O in men and -70 cmH_2O in women for MIP test and +130 cmH_2O in men and +100 cmH_2O in women for MEP [6].

In fact, these values considerably depend on several factors. The most important ones are age and gender. Men have higher values than women and these values decrease with age. For elderly patients (>70 years), it is really complicated to assess which is a good value because there is a lot of variation. Height, weight, body-mass index and spirometry indices are also parameters that must be taken into consideration. Accordingly, in the recent years, it has been highly recommended to combine the MIP and MEP technique with a novel volitional technique, the SNIP (Sniff nasal inspiratory pressure).

1.2. State of the art

Traditionally, the MIP/MEP test has been performed with whole-body plethysmography equipment, which has a cost >100 k€. In recent years, several companies have invested on the development of portable solutions, mainly regarding spirometry tests. Although the cost of these portable spirometers is significantly lower than the whole-body plethysmograph equipment, the devices are still too expensive (~2 k€) for low and middle income countries (LMICs).

Affordable and easy-to-use open source hardware electronics, such as Arduino, or distributed digital manufacturing strategies, like 3D printing, have become disruptive tools to design new medical devices in a cost-effective way without compromising the quality of the performance [7]. The development and commercialization of very accurate, easy-to-install, compensated and /or amplified low-cost sensors has also been a key fact for the expansion of this type of low-cost applications.

Accordingly, the **objective of this project** was to develop and test a portable low-cost easy-to-build and specific device to measure MIP and MEP, by using the technologies mentioned above. The new device is intended to expand the accessibility of medical tests that otherwise require much more expensive equipment.

2. Methods

The employed components and materials have been chosen according to the rationale of developing a device very simple to replicate with ease-to-find components, mostly through e-commerce. The device consists of a development board with a microcontroller, an LCD screen, a pressure transducer, a rechargeable 9V battery block, a switch, a power supply base, and two resistors. A PCB copper sheet has also been used for learning reasons but it is actually not needed. The structure has been produced in plastic by using 3D printing technology.

2.1. Electronics

The board chosen for this device is the Arduino Mega 2560 due to higher memory capacity to run the developed program than other Arduino boards (256kB of FLASH memory and 8kB of SRAM instead of 32kB and 2kB respectively of the well-known Arduino UNO). The chosen LCD touch screen (Open Smart 3.2-inch touch screen TFT LCD Shield) is compatible with the Arduino Mega 2560, with a resolution of 240x400 pixels. The use of a touch screen avoids the need of buttons or any other external component to select the parameters for the measurement. Regarding the pressure transducer, a piezoresistive strain gauge has been used. Considering that MIP and MEP values go easily from -100 cmH_2O to more than +140 cmH_2O, the SSCDRNN160MDAA5, from Honeywell, with a differential pressure of ±163 cmH_2O has been chosen.

2.2. 3D model of the prototype

The device has been designed to have two independent blocks. The first one (Figure 2) is the hand-held framework (the one used by the technician in the MIP/MEP test) containing the electronics and digital display of the measurement. The second one (Figure 3) is a mouthpiece support to contain a disposable card-board mouthpiece (for the patient). Both blocks are connected through a 1-m length silicone tube.

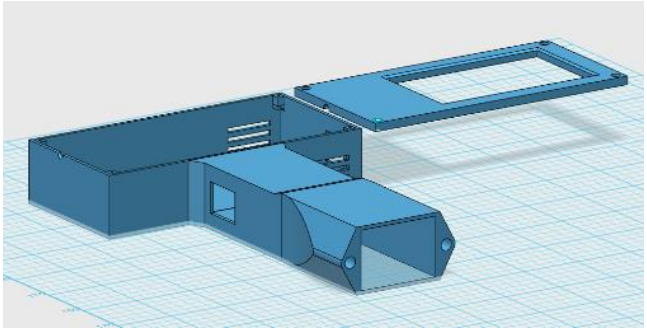


Figure 2. *Final design of the framework*

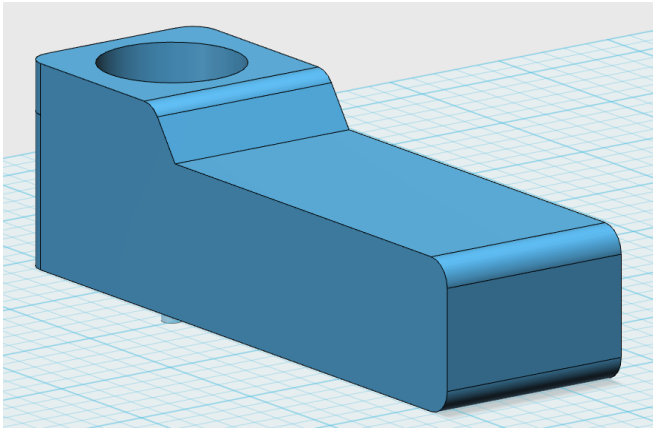


Figure 3. *Final design of the mouthpiece support*

The mouthpiece incorporates two small holes. The hole in the centre allows the blown air to travel from the mouthpiece to the transducer along the silicon tube. The second one, on its right, allows a small air-leak required to prevent closure of the glottis during forced inspiration

and decreased use of the oral muscles during forced expiration. Its implementation is mandatory as states the standard protocol of the American Thoracic Society (ATS) and the European Respiratory Society (ERS) [4,5].

2.3. Code

The code has been developed with the Arduino Integrated Development Environment (Arduino IDE), which supports the C and C++ languages. The diagram of the developed code can be seen in Figure 4.

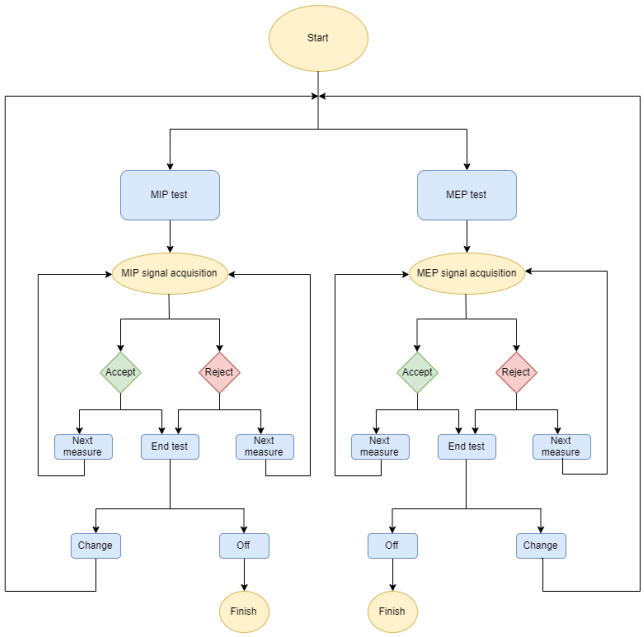


Figure 4. *Flow diagram of the MIP and MEP device*

When selecting to run the MIP or MEP measurement, the acquisition of the data starts immediately. The acquisition lasts 5 s and is carried out with a sampling frequency of 70 Hz. The gain factor of the transducer has been calibrated by using a water column as reference to maximize the accuracy of the device. However, as in-factory calibration is provided with the transducer, we recommend using this latter value for device replication.

2.4. Device testing

To assess the accuracy of the device, it has been validated by comparing its performance with a reference transducer (Honeywell 26PC Series). Four series of eight measurements were carried out. Two different adult subjects performed the experiment. The data coming from the reference transducer was acquired with a DAQ device and LabVIEW software. The MIP and MEP computation were performed with a Python 3 script.

3. Results and discussion

The final aspect of the prototype after its assembly is shown in Figures 5 and 6, and a view of the screen appearing during the acquisition of a MEP signal is shown in Figure 7.

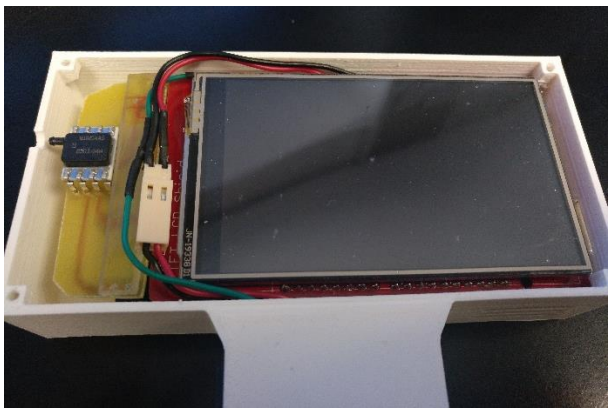


Figure 5. *Final aspect of the MIP and MEP device prototype inside view*

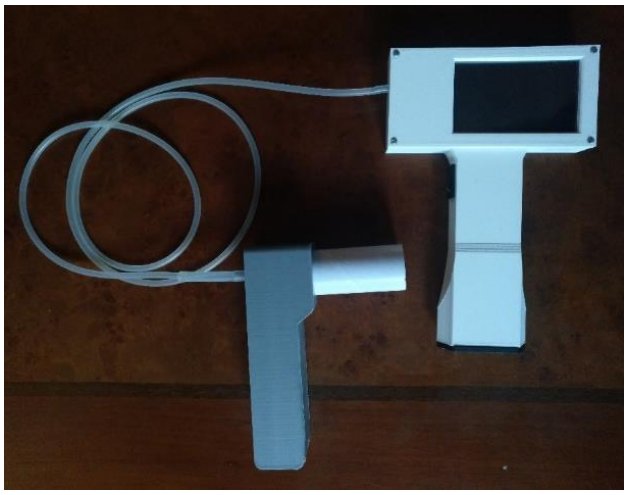


Figure 6. *Final external aspect of the MIP and MEP device prototype*

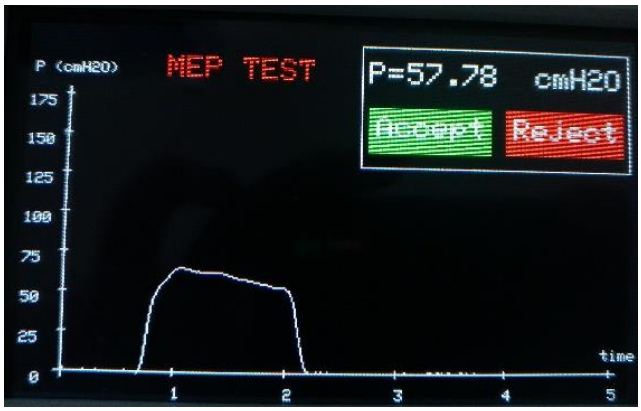


Figure 7. *Example of one of the screens during device operation*

The detail of the prototype's components cost is summarized in Table 1.

Concept	Price	Units	Total
Components			
Arduino Mega 2560	35€	1	35€
LCD	20€	1	20€
Transducer	10.36€	1	10.36€
PCB copper sheet	4.57€	1	4.57€

PLA	20€/kg	0.25 kg	5€
Silicone tube	0.47€/m	1m	0.47€
Switch	1.29€	1	1.29€
Power supply base	0.83€	1	0.83€
Total			77.52 €

Table 1. *Prototype components cost*

The results obtained when the device was evaluated by comparison with a laboratory reference setting are presented in Figures 8 and 9 following a Bland-Altman plot and a Linear Regression plot, respectively.

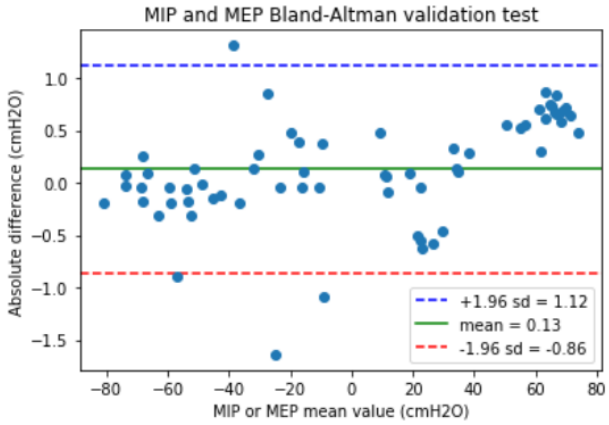


Figure 8. *Bland-Altman plot. Difference between values measured by the prototype and the reference equipment as a function of the measured values for both MIP (negative values) and MEP (positive values). Green line is the prototype bias and blue-red lines indicate the limits of agreement.*

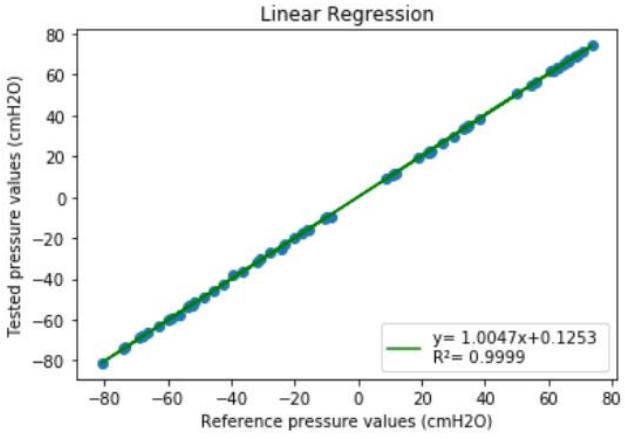


Figure 9. *Linear regression of the values obtained with the developed device and the laboratory reference.*

The obtained average difference between both tests was 0.13 cmH_2O, which corresponds to ±1% accuracy. Moreover, the coefficient of regression was 1.00, which confirms that the developed device is accurate to perform maximal inspiration and expiration pressure measurements within realistic ranges.

4. Conclusion

The designed, constructed and evaluated novel device showed that the combination of open-source technologies, like Arduino or 3D printing, with simple and accurate low-cost sensors allows the development of accurate easy-to-build, low-cost healthcare devices in the field of respiratory diseases assessment. Along with other easy-to-build medical devices in open source [7-9], an approach as the one described here opens the door to a wide range of possibilities, like the ability for LMICs to offer a greater range of diagnostic or prognostic tools in their service book or the possibility to enlarge the number of respiratory function tests in developed countries.

References

[1] P. Caruso *et al.*, "Diagnostic methods to assess inspiratory and expiratory muscle strength," *J. Bras. Pneumol.*, vol. 41, no. 2, pp. 110–123, Apr. 2015.

[2] V. Nambiar, S. Ravindra, and B. N. Kumar, "Maximal inspiratory and expiratory pressures in men with chronic obstructive pulmonary disease: A cross-sectional study," *Indian J. Respir. Care*, vol. 7, no. 2, p. 88, 2018.

[3] R. C. Kelley and L. F. Ferreira, "Diaphragm abnormalities in heart failure and aging: mechanisms and integration of cardiovascular and respiratory pathophysiology.," *Heart Fail. Rev.*, vol. 22, no. 2, pp. 191–207, 2017.

[4] "American Thoracic Society/European Respiratory Society ATS/ERS Statement on Respiratory Muscle Testing this joint statement of the american thoracic society (ats), and the european respiratory society (ers) was adopted by the ats board of directors, march 2001 and by the ers executive committee, june 2001."

[5] P. Laveneziana et al., "Early View Task Force Report ERS Statement on Respiratory Muscle Testing at Rest and during Exercise."

[6] "Diseases of the Thoracic Cage and Respiratory Muscles," pp. 763–775, Jan. 2012.

[7] Farré R, Montserrat JM, Solana G, Gozal D, Navajas D. Easy-to-build and affordable continuous positive airway pressure CPAP device for adult patients in low-income countries. Eur Respir J. 2019;53(5):1802290

[8] Garmendia O, Rodríguez-Lazaro MA, Otero J, et al. Low-cost, easy-to-build noninvasive pressure support ventilator for under-resourced regions: open source hardware description, performance and feasibility testing. Eur Respir J. 2020;55(6):2000846.

[9] Farré R, Trias G, Solana G, Ginovart G, Gozal D, Navajas D. Novel Approach for Providing Pediatric Continuous Positive Airway Pressure Devices in Low-Income, Underresourced Regions. Am J Respir Crit Care Med. 2019;199(1):118-120.

Development of a Virtual Reality Application for Wrist, Elbow and Shoulder Rehabilitation Using Oculus Quest

Rodrigo Enríquez Medina, José Ángel Santos Paz, Rodrigo García Carmona

Escuela Politécnica Superior, Universidad CEU San Pablo, Madrid, Spain

{rm.enriquez@usp.ceu.es, ja.santos3@usp.ceu.es, rodrigo.garciacarmona@ceu.es}

Abstract

In recent years, video games have gained more traction in fields other than pure entertainment. This has led to a diverse collection of serious games, designed to provide a specific benefit through gamification. In parallel, the latest development in Virtual Reality (VR) devices, initially designed for videogames, has provided an enticing and attractive platform that can be used for the development of healthcare applications. Recent studies have proven that specially tailored games are a viable rehabilitation option, and the intensive development of VR devices during the last 2 years, suggests that the technology has reached a point where it can be used to develop rehabilitation scenarios. with this in mind, this work has explored the viability of developing a VR upper-limb rehabilitation application for the Oculus Quest device that bridges the gap between commercially available gear and medical needs and can be used in a practical way in a rehabilitation clinic or at the patient's home.

1. Introduction

The Oculus Quest is a three-part device comprised of a headset and two *touch controllers* [1]. Each of the three components register movement with up to 6 degrees of freedom [2], three more than other, more limited devices. However, unlike more expensive and cumbersome solutions, like the HTC Vive, Oculus Quest doesn't need external tracking beacons (features inside-out tracking through a set of infrared cameras) nor a high-performance computer. This makes the device very convenient for upper limb rehabilitation as we can measure the pitch, yaw and roll while at the same time accessing translational movements data, all in an easy to setup and inexpensive package, ideal for rehabilitation clinics and home use, where time and convenience is paramount.

We wanted to create a VR upper-limb rehabilitation application using this device by focusing on two tasks: a) combining these Oculus Quest measurements with the data provided by an IMU (Inertial Measurement Unit) sensor developed by Werium Solutions[3], and b) developing a sample VR application that features several possible upper-limb rehabilitation scenarios.

2. Methods

During the development of this work we encountered several problems and difficulties, some of them related to the early development stage of the technologies used. Therefore, we find value in detailing the steps taken so future developers can learn from them; they will be thoroughly explained in the following sections.

2.1. Integrating the Device with the Unity Engine

Unity is a game engine that provides a platform in which a vast array of games can be developed, from 2D platformers to 3D MMORPGs[4], with the possibility to develop other genres such as rehabilitation games, which in the last few years have gained more attention [5,6]. In order to develop VR games, a 3D project must be first made and configured appropriately. This requires a Unity version higher or equal to the 2019.2.0 (though preferably the latest version). Then, once the new 3D project has been created, we must access the asset store, import the "Oculus Integration" pack [7] and install it together with the "Oculus Utilities". This package will provide the necessary tools to work with all Oculus devices. Unfortunately, this programming interface is cumbersome, unreliable and difficult to use on its own. To ease up the development process, we will complement it with the VRTK package, which will provide some convenient scripts [8]. To do that, we must install "Oculus Desktop" in the package manager. Then, by opening the "Packages" folder in the project directory we must open the "manifest" file and modify it by hand. We must include the following lines into the manifest file:

```
"scopedRegistries":[{"name":"npmjs","url":"https://regist
ry.npmjs.org/","scopes": ["io.extendreality" ]}]

"io.extendreality.vrtk.prefabs": "1.1.3"
```

The latter snippet fits in the dependencies section of the file.

Once all the necessary packages have been installed, we must configure the project settings. In the "Player" tab, the "Virtual Reality Supported" checkbox must be selected, and the "Shared Depth Buffer" and "Dash Support" checkboxes disabled.

VR applications can now be developed using the prefabs provided by VRTK and Oculus Integration. However, such applications lack the necessary bridge to be installed into the Quest headset. Therefore, an ADB library needs to be installed [9]. Unfortunately, Oculus divides the Oculus Quest device into two privilege modes, normal and developer mode and only the developer mode enables debugging through ADB [10]. This means that the headset needs to be configured too. When acquiring a headset, it first requires an Oculus account and a mobile device. It is from the latter, using the Oculus app and connecting to the headset, that the developer mode can be toggled. Finally, we need to activate the debug mode by using the Windows command prompt, entering into the ADB directory and listing the devices by typing: adb devices [11].

Should the device be listed as unauthorized, the server ought to be restarted by typing "adb kill-server" [11] and then relisting the devices using "adb devices" [11] once again.

ISBN: 978-84-09-25491-0, pags. 169- 172

This will trigger the headset to ask for permission and authorize the given computer to debug and install applications.

Once the development process is concluded, the final result should be an application ".apk" file. To install it into the headset, it should be moved into the ADB directory(for convenience) and, from the command prompt, run the "adb install -r application.apk" [11] command.

2.2. Building the environment

Designing a suitable rehabilitation environment proved to be challenging, it took two test environments to establish a delicate balance between what was a welcoming playable experience and a feasible rehabilitation task. We realized that certain issues, such as the physics engine, were going to limit the development of the desired experience.

One of the most important aspects of an application like this is giving a proper positive feedback. This feedback not only makes the exercise enjoyable but also improves the player's performance [12]. After pondering the key aspects of the application, a list of priorities containing the elements the scene had to feature was produced.

The most important aspect was to keep the focus on the rehabilitation exercises; they are the core of the app and not just a game mechanic. It was also important to create an environment as immersive as possible, to maximize the player's attention. The final element of the application was the ability to produce a feeling of accomplishment once the user managed to perform an exercise.

In the end, the definitive environment was based upon the Star Wars saga Jedi training scenes. This scenario is conductive to several rehab exercises disguised as several uses of the fictional "force". The scene needed to be immersive but, at the same time, keep the player calm so that he could perform the exercises without stress or urgency. A temple room was chosen as the fictional location, as can be seen in Figure 1.

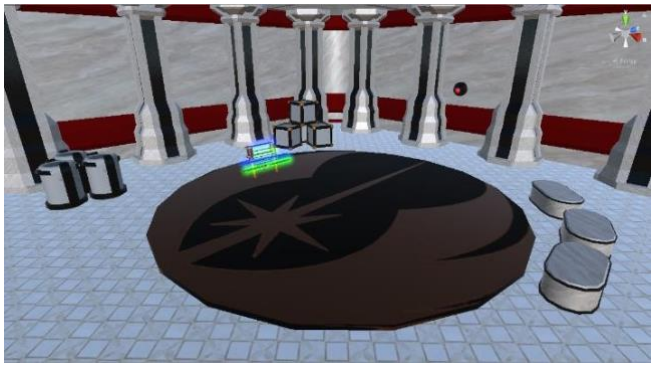


Figure 1. *The scene*

Regarding the exercises, the idea was to provide an example for wrist, elbow and shoulder exercises. These were coded in the following way:

- *Wrist movements [13]:* Several lightsabers were placed in the scene, so the player could train the forearm's pronation and supination. Then, one of the force powers (force slam) was implemented using the active range of motion of the wrist (Figure 2).

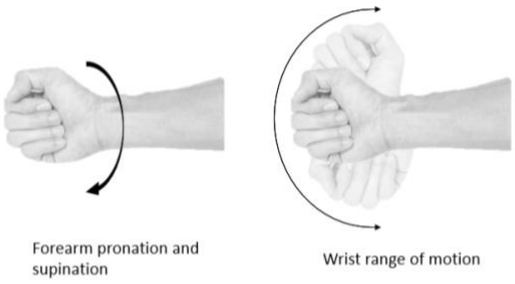


Figure 2. *Wrist exercises [13]*

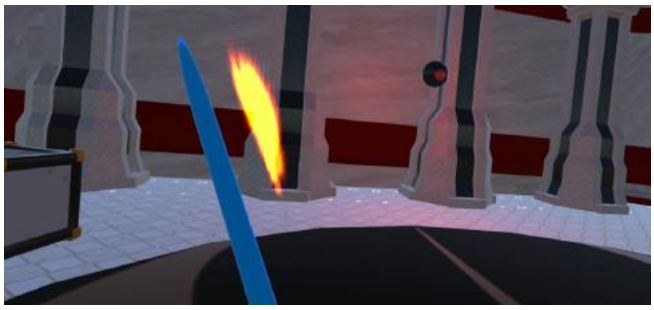


Figure 3. *Lightsabre usage exercising wrists*

- *Elbow movements [13]:* One of the force powers (force pull) was implemented through the elbow's flexion and extension (Figure 5). Another force power (force push) was implemented through the eccentric wrist extension exercise (Figure 6).

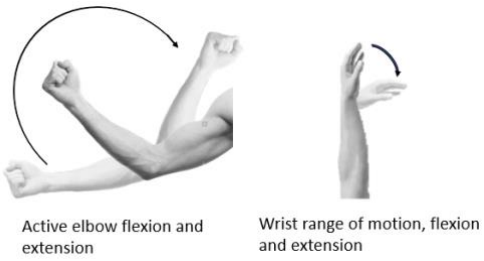


Figure 4. *Elbow exercises [13]*

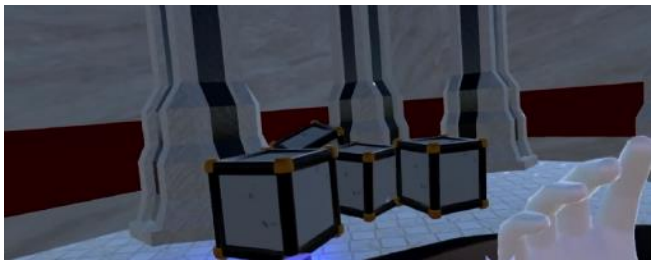


Figure 5. *Force pull instants before the cubes are flung over*



Figure 6. *Force push instants before the benches are flung around.*

- *Shoulder movements [14]:* The final force power (force lift) was implemented using the shoulder flexion mechanic (Figure 4).



Figure 7. *Shoulder Exercise [14]*



Figure 8. *barrels lifted moments after using force lift*

To determine if the player is performing these exercises, the rotation and position of the controller is analysed for all exercises except the forearm supination and pronation. This is done through a series of scripts. The way these work is by analysing the position of the controller in every update; if the controller reaches a specific location, then the movement is triggered. When this happens, the script actively monitors the movement and tries to ensure that the player continues to do so. If the player successfully performs the movement, then he is rewarded with some power which he can use somewhere in the scene. If the player fails to perform the exercise, then the force power will remain untriggered.

In the case of the forearm pronation and supination exercises, the mechanism works differently. For this exercises some lightsabers are provided in the scene which the player can use. As soon as they are picked up a sphere will start attacking the player with some blaster shots which the player must deflect using the lightsabers. This exercise is the most versatile of all and, by its own nature, the most unreliable. There is no control mechanism: the player is expected to perform the task correctly without any "rails" or other guidance.

Regarding the design of the sphere which attacks the player, it is an active element which at the beginning of the scene starts rotating around the room to catch the player's attention. As soon as the player picks one of the lightsabers, the sphere is programmed to move into its original position and start attacking the player. The way it does that is by targeting the position of the right controller and then calculating the necessary vector to translate the blaster shot. Once the optimal trajectory has been calculated, a small random error is added in the positive Y and X axes so it lands in a circle reachable by moving the wrist to align the lightsaber's position with the shot. This way an illusion is created in which the sphere seems to target the player, but all the shots can be parried with the lightsaber effortlessly.

An important detail regarding the blaster's design is the orientation it takes when heading towards the player. The player might notice that the blaster does not fully face them and that it is slightly rotated. This is done on purpose as a perfectly perpendicular blaster shot increases the parrying difficulty. As the exercises do not revolve around the cognitive ability of the patient but rather the motor skills, sacrificing some realism for the sake of playability seems reasonable. To ensure the sphere works correctly there are three main scripts. Two of them are overviewing the movement of the sphere so that the player notices it when entering the scene and the third one controls the blaster shot. This is half of the overall process, there are more scripts controlling the lightsabers and the behaviour of the shot.

The source code of this project can be found online [15], together with several videos illustrating its usage.

3. Results and Discussion

In this work we managed to set up an Oculus Quest development environment using Unity and use it to develop a viable rehabilitation application for the upper limb disguised as an engaging game.

The development of VR applications for this platform remains a novelty and thus, it has the benefits and the inconveniences of all cutting-edge technology. The opportunities are vast but the difficulties to overcome made the development an arduous and complicated task. Errors were frequent, glitches ever present, and the use of third-party software required. As will be explained in the next section, one of our future ideas requires the integration of external sensors that complement the on-board Quest sensors. The difficulty of connecting such sensors might be reduced as more developers flock to work on applications especially designed for the Oculus Quest, since this headset has been particularly successful [16] and a second generation version has been announced [17]. Another positive impact of a more mature platform is the introduction of permanent and reliable documentation sources. As of now, documentation on the issues and implementation (if covered at all) is spread across blogs, repositories, official sites and amateur pages. It might also be contradictory, deprecated or directly useless, so it might be better to wait out until there is a more fleshed out development landscape.

A final note regarding the design choices of the project: When defining the exercises, they were based upon angles and distances. This means that they are calibrated for a certain height range and angles. In order to translate this project into any commercial application of some sort, these parameters should be calibrated to the specific needs of each player. In addition, the application could use some more polish and advanced game mechanics, but always keeping the focus on the exercises at hand.

4. Conclusions and Future Work

Developing applications through Unity directed towards the Oculus Quest is possible, though cumbersome. The developer ought to focus on the critical points he wishes to emphasize in their application and work on them as thoroughly as possible as it is more than likely that they will be problematic to some extent. Once the critical aspects of the application have been settled on, developing the

environment ought to be easier, as long as the game mechanics are dissociated from the scene. Lastly, the lack of support regarding the experimental functions such as the Bluetooth functionality, make the integration of external devices theoretically possible but complicated.

A future line of work is the integration of an IMU (Inertial Measurement Unit) sensor that can provide the rotation of a particular segment of the patient's body, other than the head and hands (already provided by Oculus Quest). It remains to be seen if such a sensor could be integrated, since the Quest headset does not officially support Bluetooth so accessing this feature requires the developer mode (see section 2.1). As a follow up of the project, the development of a library which tests this theory could be a viable goal. Another alternative could be using other third-party software which could bridge this gap. However, of all the asset store candidates, none fulfilled the requirements needed [18-20].

The idea behind the sensor integration is to provide one or several additional measurements with which to confirm the legitimacy of the movement. By providing a two-point reference, the arm / forearm (depending on the sensor's location) could be modelled and thus confirm that the user is indeed moving the elbow and not using the back muscles.

We also plan to properly validate the developed application, to ensure that it can fulfil its design goals with a wide range of users, featuring different heights and body types. This validation will be performed with people without disabilities before proceeding with a clinical validation with therapists.

Acknowledgements

I would like to thank Sara Fernández Pomar for her contribution both in the design and in the debugging of the project.

References:

[1] *"Oculus Quest: Visor De Realidad Virtual Todo En Uno | Oculus". 2020. Oculus.Com. Accessed January 19 2020. https://www.oculus.com/quest/?locale=es_ES.*

[2] *"Degrees Of Freedom (Dof): 3-Dof Vs 6-Dof For VR Headset Selection". 2020. Virtualspeech.Com. Accessed January 19 2020. https://virtualspeech.com/blog/degreesof-freedom-vr.*

[3] *Raya, Rafael, Rodrigo Garcia-Carmona, Cristina Sanchez, Eloy Urendes, Oscar Ramirez, Alvaro Martin, and Abraham Otero. 2018. "An Inexpensive And Easy To Use Cervical Range Of Motion Measurement Solution Using Inertial Sensors". Sensors 18 (8): 2582. doi:10.3390/s18082582.*

[4] *"Unity Game Development Essentials". 2020. Google Books. Accessed January 19 2020. https://books.google.es/books?hl=es&lr=&id=WfAWzVW9IK0C&oi=fnd&pg=PT16&dq=unity+engine&ots=AU9VCyNLFi&sig=8o95HrJHDrNpYVv_WNUo-21s3ZQ#v=onepage&q=unity%20engine&f=false.*

[5] *"Serious Games For Rehabilitation: A Survey And A Classification Towards A Taxonomy - IEEE Conference Publication". 2020. Ieeexplore.Ieee.Org. https://ieeexplore.ieee.org/abstract/document/5556674.*

[6] *Lohse, Keith, Navid Shirzad, Alida Verster, Nicola Hodges, and H. F. Machiel Van der Loos. 2013. "Video Games And Rehabilitation". Journal Of Neurologic Physical Therapy 37 (4): 166-175. doi:10.1097/npt.0000000000000017.*

[7] *"Oculus Integration - Asset Store". 2020. Assetstore.Unity.Com. https://assetstore.unity.com/packages/tools/integration/oculus-integration-82022*

[8] *"Design, Develop, And Deploy For VR - Unity Learn". 2020. Unity Learn. Accessed January 19 2020. https://learn.unity.com/course/oculus-vr?language=en*

[9] *"Download 15 Seconds ADB Installer (All Versions)". 2016. Android MTK. Accessed January 19 2020. https://androidmtk.com/download-15-seconds-adbinstaller.*

[10] *Base, Knowledge, and How Go/Quest). 2020. "How To Enable Developer Mode (Oculus Go/Quest)". EZ360. Accessed January 19 2020. https://ez-360.com/knowledgebase/how-to-enable-developer-mode-oculus-go-quest/.*

[11] *"Android Debug Bridge (Adb) | Android Developers". 2020. Android Developers. Accessed January 19 2020. https://developer.android.com/studio/commandline/adb?hl=es-419.*

[12] *"Virtual Reality Games For Rehabilitation Of People With Stroke: Perspectives From The Users". 2020. Disability And Rehabilitation: Assistive Technology. https://www.tandfonline.com/doi/abs/10.3109/17483107.2011.574310.*

[13] *"Summit Medical Group". 2020. Summitmedicalgroup.Com. Accessed January 19 2020. https://www.summitmedicalgroup.com/library/pediatric_health/sma_tennis_elbow_exercises/.*

[14] *"Summit Medical Group". 2020. Summitmedicalgroup.Com. Accessed January 19 2020. https://www.summitmedicalgroup.com/library/adult_health/sma_shoulder_bursitis_exercises/.*

[15] *Jedi Academy Git Repository: https://github.com/SlothThyLord/Proyectos-III*

[16] *"Oculus surpasses $100 million in Quest content sales". 2020. TechCrunch.Com. Accessed September 15 2020. https://techcrunch.com/2020/05/18/oculus-surpasses-100-million-in-quest-content-sales/*

[17] *"A new Oculus Quest headset is reportedly on the way". 2020. TechCrunch.Com. Accessed September 15 2020. https://techcrunch.com/2020/05/05/a-new-oculus-quest-headset-is-reportedly-on-the-way/*

[18] *"Simple Bluetooth For Android - Asset Store". 2020. Assetstore.Unity.Com. Accessed January 19 2020. https://assetstore.unity.com/packages/tools/network/simplebluetooth-for-android-140551.*

[19] *"Bluetooth Plugin For Android - Asset Store". 2020. Assetstore.Unity.Com. Accessed January 19 2020. https://assetstore.unity.com/packages/tools/network/bluetooth-plugin-for-android-119760.*

[20] *"Android Bluetooth Multiplayer (Pro) - Asset Store". 2020. Assetstore.Unity.Com. Accessed January 19 2020. https://assetstore.unity.com/packages/tools/network/android-bluetooth-multiplayer-pro-9643.*

On the use of texturized electrospun wound dressings for absorption maximization

Aitor Diaz Zelaia[1], Lorea Buruaga Lamarain[1,*], Judith Zubia Aranburu[1], Cristina Eguizabal Argaiz[2,3], Silvia Santos Cabrera[2,3], Clara Rodriguez Aierbe[2,3], Alaitz Zabala Eguren[2,**]

[1] Mechanical and Industrial Manufacturing Department, Mondragón, Spain, * lburuaga@mondragon.edu, **azabalae@mondragon.edu

[2] Cell Therapy, Stem Cells and Tissues Group, Biocruces Bizkaia Health Research Institute, Barakaldo, Spain

[3] Osakidetza, Basque Centre for Blood Transfusion and Human Tissues, Galdakao, Spain

Abstract

We have studied a new approach for maximizing the absorption capabilities of a wound dressing by using textured collectors in the electrospinning process. Thus, we aim to increase the surface to volume ratio of the mesh, obtaining a larger area to interact with the environment and to improve the healing process of a wound by a larger absorption rate of the dressing. To do so, a non-textured reference collector (2R BA Finish) and a textured collector (2M Linen Finish) from *ACERINOX* have been used in the electrospinning process. Then, the characterization of the fibers has been done using a scanning electron microscopy, followed by a surface analysis with an optical perfilometer, to ensure the geometries are correctly printed in the final mesh. An absorption test has been performed to see the improvement in the fluid uptake capabilities of the resultant dressings. The absorption capabilities of the textured dressings resulted to be a 90 percentage points bigger in low relative humidity manufacturing condition samples compared to the non-textured ones, whereas the fibers manufactured at high relative humidity conditions reported decreased absorption capabilities. Thus, a promising approach to increase the liquid uptake has been demonstrated that could improve the performance of wound dressings in wound healing processes.

1. Introduction

Burden of treating wounds in health care is rapidly increasing due to an aging population, a sharp rise of the incidence of diabetes and obesity worldwide, and other issues such as amputations (one every five seconds) that may often derive in wounds [1]. Thus, it is estimated that at least 2 % of the global population will experience a chronic wound during lifetime [2].

Moreover, the lack of specialized training in the diagnosis and treatment of wound and the lack of priority of resources and capabilities in this area, has led to no innovation in therapies, diagnosis and procedures, leading to a world spreading “silent epidemic” . So far, wound care is already a multibillion-dollar world-wide problem, which is affecting just in the USA to more than 5,7 million people, with an annual cost of US$20 billion [3]. Besides, the treatment and care of wounds account for around 3-6 % of the total health-care expenditure in developed countries [4].

This generates a huge burden to society, not just economically, but also in productivity, environmental issues, and in both physical and psychological distress on patients and their relatives [1]. Nevertheless, the implementation of a strategy based on best practice guidelines, may help to reduce the cost of wound management with up to a 30 % [5].

To generate an appropriate wound dressing, several requirements should be fulfilled to ensure a successful healing process. The dressing must provide a physical barrier between the wound and the environment, be permeable to moisture and oxygen, to be non-adherent, and to have enough absorption capacity for maintaining the moist environment required for appropriate wound healing[6].

Electrospinning technique has attracted much attention in wound dressing applications due to the large surface area to volume ratio and high porosity of the nanofibers. Due to the importance of the ability to swell and absorb exudates from the wound, several strategies have been studied in order to improve the absorption capabilities of electrospun mats. The use of a combination of materials [1], core-shell structures [7] modified hydrophilic properties [8], and nanopore structures [9] have proven to be a successful strategies to promote exudate absorption. More recently, the use of a textured collector to enhance the volume to surface ratio of the electrospun mat surfaces has reported improvements in oil absorption capacities [10]. This approach presents high potential for the manufacturing of wound dressings since the absorption improvement might be obtained without the need of any further treatment. In this context, 3D printed collectors presenting desired geometries on the surface have been used to emboss textured patterns on electrospun mats, reporting improved cell attachment, proliferation, differentiation, and growth [11]. However, to the best of the author’s knowledge, the use of textured collector as absorption enhancement strategy has not been explored yet.

The main objective of the following study is to maximize the absorbance capacity of electrospun wound dressings by improving the volume to surface ratio using textured collectors.

ISBN: 978-84-09-25491-0, pags. 173- 176

2. Methodology

2.1. Manufacturing of electrospun mats

High molecular weight polycaprolactone (PCL) (80,000 g/mol) and chloroform was purchased from Sigma Aldrich. The peening solution was prepared by dissolving a 2,4 g PCL in 11,8 ml chloroform (12 % wt). In order to analyse the effect of the pattern textures on the absorption capabilities, two collectors were used, a non-patterned one as reference (2R BA Finish) and the textured (2M Linen Finish) AISI 304 references from Acerinox (see Figure 1). Moreover, the effect of the relative humidity (RH) in the swelling capabilities was also analysed by the manufacturing of the mats in both, 40 and 80 % of RH.



Figure 1. *Images of the non-textured reference (2R BA Finish) and textured (2M Linen Finish) collectors used for the comparative study.*

The process parameters based in previous studies are indicated in Table 1 for all dressings.

Collector	Flow rate (ml/h)	Distance (cm)	Voltage (kV)
Reference / Textured	0.5	15	13.3

Table 1. *Parameters used for the manufacturing of the electrospun dressings*

2.2. Fiber diameter characterization

The morphology of the textured and reference mats was characterized by scanning electron microscopy (FEI Nova NanoSEM) at accelerating voltage of 3 kV and at low vacuum condition (80 Pa). The diameters of the fibers were measured by image J software.

2.3. Surface analysis

The surface of both textured and reference collectors was analysed by an optical profilometer SensoFar S-NEOX using the intereferometry technique (20xDI). The measurements were post-processed by means of the metrological software SensoMap Premium. The form was removed by means of plane fitting and the root mean square roughness (Sq) and the developed area (Sdr) was computed in the primary surface according to ISO 25178. The surfaces of the electrospun mats were analysed by means of the optical Microscope Leica DM500.

2.4. Assessment of fluid absorption properties

The weight of the dry and wet samples was measured 5 times by means of the Mettler-Toledo XPE205 microbalance (resolution 10^{-5} gr), and the average value and standard deviation were computed. The procedure for the measurements was as follows: the dried samples were weighted and recorded as m_0, then added into a recipient containing 4 ml of HANKS solution. After immersing of 1h, 24 h and 3 days at room temperature, the samples were picked up by means of tweezers, removed excess water from the surface using filter papers and weighted as m_1. The fluid absorption (Q) was then calculated by the following equation 1:

$$Q_1 = \frac{m_1 - m_0}{m_0} \, x \, 100 \, \% \qquad (1)$$

3. Results and discussion

3.1. SEM analysis and fiber diameter characterization

The morphology of the electrospun dressings was studied by FE-SEM image analysis, as observed in the Figure 2. Moreover, the effect of high relative humidity in the fiber production has also been considered.

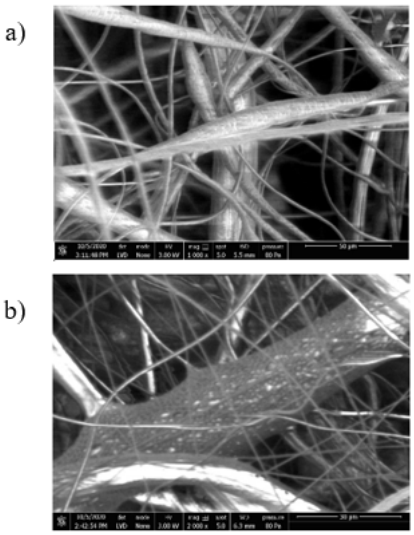


Figure 2. *Representative FE-SEM image of electrospun fibers for diameter characterization. A) 40 % of RH. B) 80 % of RH*

Both the samples manufactured with reference collector and the textured ones presented similar fibre diameters of 10 ± 5 µm when produced under low RH condition. However, the difference in the diameters was noticeable when comparing the samples generated at high RH (see Figure 2). The following Figure 3 presents the measured diameters of the reference (R) and textured (T) collectors produced at low RH (RT) and high RH (RH) condition.

A great increment in the fiber diameter can be observed when comparing the dressings manufactured at different RH, being the values at least twice bigger in high RH conditions. Moreover, these fibers look less homogeneous and ramifications tend to appear from the largest fibers, increasing the deviation values and showing a greater dispersion.

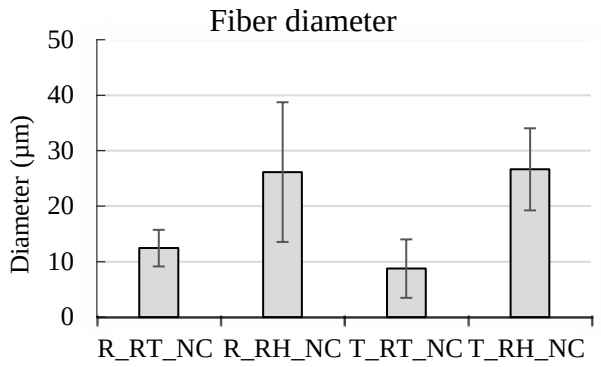


Figure 3. *Graphical representation of the average diameter of the fibers within the texturization and processing relative humidity*

3.2. Surface analysis

Figure 4 shows the 3D axonometric representations of the reference and textured collectors at imposed scale, at which the enhanced texture of the textured collector can be observed. The 2D images shows the morphology of each surface, presenting a very different shapes on the surface. The 3D parameter results are shown in Table 2.

Collector	Sq [µm]	Sdr [%]
Reference	0.4±0.03	0.03±0.004
Textured	25±0.15	14.4±0.03

Table 2. *3D topographical parameters for the reference and textured collectors, computed according to ISO 256178.*

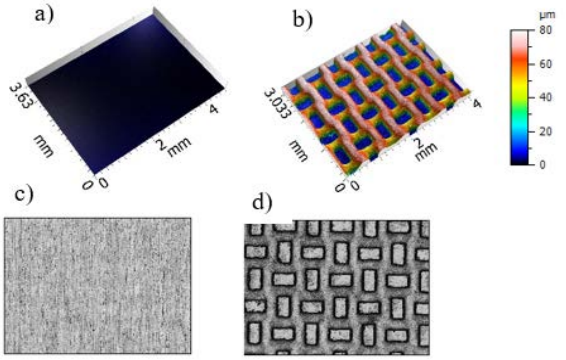


Figure 4. *Axonometric projections of the reference and textured collectors (a, b) at imposed scale, and the corresponding free scale 2D images for morphological observation (c, d).*

The Sq parameter indicating height properties of the surfaces shows that the textured samples present deeper motives in the surfaces. The developed area ratio parameter Sdr is expressed as the percentage of additional surface area contributed by the texture as compared to an ideal plane. The developed area of the texture sample (Sdr) is 14 times bigger than the reference collector. Looking at the cross section of the textured collector surface in Figure 5, it can be observed that the movies presents a 63 µm hight (h), 240 µm width (w), and 563 and 670 µm spacing (s1, s2).

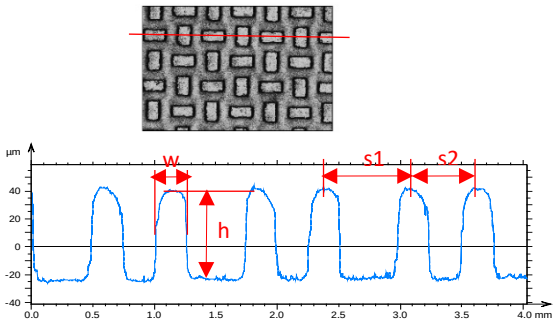


Figure 5. *Cross section of the textured collector surface for pattern size calculation.*

Figure 6 shows the microscope images of the textured and reference electrospun mats. It can be observed that the micro texture of the textured collector was successfully printed in the electrospun mat surface, presenting the negative shapes on it.

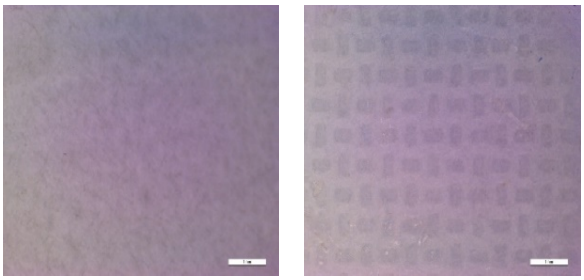


Figure 6. *On the left, the image of the reference dressing. On the right, the point-based dressing*

3.3. Absorption test

Due to the need of absorbing the excess exudates and handling with moist environment, high liquid uptake capacity is an important parameter to take into account in wound dressing materials. After the immersion of the samples in Hanks solution for a period of time of 24 days, the absorbance capabilities of the dressings has been observed in the following Figure 7.

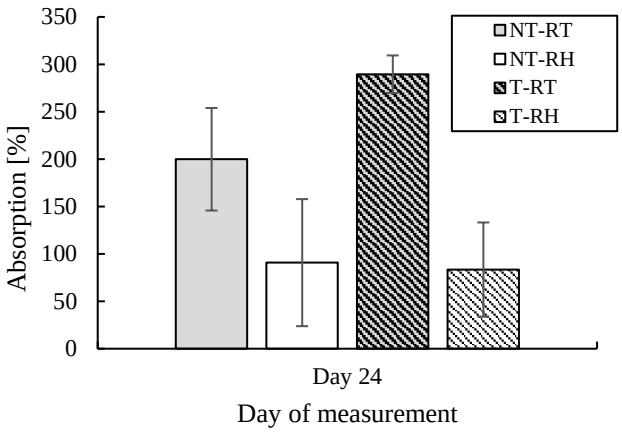


Figure 7. *Absorption (%) of the reference (NT) and textured (T) dressings according to the RH after 24 days of immersion, where RT refers to a value of 40 % and RH 60 %.*

By analysing the results, an important improvement in the swelling capabilities can be seen among the reference and textured mesh in RH of 40 %, due to the enhancement of

the volume to surface relation. An increase of almost a 90 percentage points is obtained when comparing the liquid uptake capabilities of both dressings.

Nevertheless, if we analyse the uptake within the RH of 80 %, we cannot observe any significant difference among the reference and textured dressings, and there is a significant worsening in comparison to the low RH. This can be explained attending the lack of stability in the electrospinning process and increment of the diameter values caused by too high relative humidity [12]. This tends to generate pores that should help in the swelling capabilities, but also causes larger fiber diameters and lower fiber densities that avoids an appropriate absorption process, leading to an overall decrease in the uptake capabilities.

Despite the clear gain in uptake capabilities observed in the textured dressings, the resultant deviations are bigger than the desired ones. Thus, these deviations need to be reduced by manufacturing more compact and dense dressings that may offer more consistent measures, without so high deviations. However, the difference among the average values of the swelling capabilities are still high enough in favour of the textured dressings over the non-textured ones, showing a potential use of this approach to improve uptake in wound dressing processes.

4. Conclusions

Textured polycaprolactone (PCL) electrospun dressings for wound healing were successfully obtained by applying a textured collector in the manufacturing process. The use of commercially available non-patterned reference collector (2R BA Finish) and the textured collector (2M Linen Finish) from Acerinox showed the potential to print the geometries in the dressings, as observed in the optical images and the negative shapes in the meshes. Both dressings showed an average fiber diameter of 10 ± 5 µm which is increased more than twice the value when changing the processing relative humidity from the environmental 40 % to a higher 80 %. Moreover, this RH has also a negative effect on the swelling capability of the dressings, as the increase in the diameter values and the instability in the manufacturing process generates non favourable conditions to handle appropriate absorption rates. Nevertheless, the texturization of the dressing has shown a significant increment on the liquid uptake capabilities of the dressing, offering a 90 percentage points greater absorption rate than the non-patterned reference dressing. Indeed, a promising and simple method is being demonstrated to maximize the swelling capabilities of a dressing, offering a great opportunity to improve the wound healing processes.

Acknowledgements

This project has received funding from the Basque Government Health Department, under “Ayudas a proyectos de investigación y desarrollo en salud” program (project “ALOPRP”).

References

[1] K. Järbrink et al., “The humanistic and economic burden of chronic wounds: A protocol for a systematic review,” Systematic Reviews, vol. 6, no. 1, Jan. 2017, doi: 10.1186/s13643-016-0400-8.

[2] H. Brem et al., “Molecular markers in patients with chronic wounds to guide surgical debridement,” Molecular Medicine, vol. 13, no. 1–2, pp. 30–39, Jan. 2007, doi: 10.2119/2006-00054.Brem.

[3] R. G. Frykberg and J. Banks, “Challenges in the Treatment of Chronic Wounds,” Advances in Wound Care, vol. 4, no. 9, pp. 560–582, Sep. 2015, doi: 10.1089/wound.2015.0635.

[4] J. Posnett and P. J. Franks, “The burden of chronic wounds in the UK | Nursing Times,” 2008. https://www.nursingtimes.net/clinical-archive/tissue-viability/the-burden-of-chronic-wounds-in-the-uk-23-01-2008/ (accessed Oct. 15, 2020).

[5] A. Hjort and F. Gottrup, “Cost of wound treatment to increase signifi cantly in Denmark over the next decade,” 2010.

[6] M. R. MacEwan, S. MacEwan, T. R. Kovacs, and J. Batts, “What Makes the Optimal Wound Healing Material? A Review of Current Science and Introduction of a Synthetic Nanofabricated Wound Care Scaffold,” Cureus, Oct. 2017, doi: 10.7759/cureus.1736.

[7] R. Li et al., “Novel SA@Ca2+/RCSPs core-shell structure nanofibers by electrospinning for wound dressings,” RSC Advances, vol. 8, no. 28, pp. 15558–15566, 2018, doi: 10.1039/c8ra00784e.

[8] D. Simões, S. P. Miguel, and I. J. Correia, “Biofunctionalization of electrospun poly(caprolactone) fibers with Maillard reaction products for wound dressing applications,” Reactive and Functional Polymers, vol. 131, pp. 191–202, Oct. 2018, doi: 10.1016/j.reactfunctpolym.2018.07.021.

[9] A. D. Juncos Bombin, N. J. Dunne, and H. O. McCarthy, “Electrospinning of natural polymers for the production of nanofibres for wound healing applications,” Materials Science and Engineering C, vol. 114. Elsevier Ltd, Sep. 01, 2020, doi: 10.1016/j.msec.2020.110994.

[10] J. Lin, B. Ding, J. Yang, J. Yu, and G. Sun, “Subtle regulation of the micro- and nanostructures of electrospun polystyrene fibers and their application in oil absorption,” Nanoscale, vol. 4, no. 1, pp. 176–182, Jan. 2012, doi: 10.1039/c1nr10895f.

[11] M. M. Saadatmand, M. E. Yazdanshenas, R. Khajavi, F. Mighani, and T. Toliyat, “Patterning the surface roughness of a nano fibrous scaffold for transdermal drug release,” 2019.

[12] B. Zaarour, L. Zhu, C. Huang, and X. Jin, “Controlling the Secondary Surface Morphology of Electrospun PVDF Nanofibers by Regulating the Solvent and Relative Humidity,” Nanoscale Research Letters, vol. 13, 2018, doi: 10.1186/s11671-018-2705-0.

Análisis de la importancia de las características de la esteatohepatitis no alcohólica aplicando modelos Random Forest

R. Holgado-Cuadrado[1], R. García-Carretero[2], O. Barquero-Pérez[1]

1 Departamento Teoría de la Señal y Comunicaciones y Sistemas Telemáticos y Computación, Universidad Rey Juan Carlos, Fuenlabrada, España, r.holgado.2016@alumnos.urjc.es,oscar.barquero@urjc.es

2 Servicio Medicina Interna, Hospital Universitario de Móstoles, Móstoles, España, rgcarretero@salud.madrid.org

Resumen

La enfermedad de hígado graso no alcohólico (EHGNA) es, en la actualidad, la enfermedad hepática crónica más prevalente. Una de las manifestaciones más típicas es la esteatohepatitis no alcohólica (EHNA), que puede derivar en cirrosis. Existen, no obstante, dos problemas por resolver, a saber, (1) crear diagnósticos precoces y no invasivos, y (2) entender mejor la base fisiopatológica del desarrollo de la enfermedad para conseguir tratamientos eficaces. El objetivo del trabajo fue diseñar un modelo predictivo para estimar la probabilidad de EHNA, e identificar las variables más relevantes en el desarrollo de la enfermedad. Para ello, se han analizado los datos de 2.828 pacientes hipertensos utilizando técnicas de machine learning, incluyendo modelos Random Forest, estrategias de desbalanceo de datos y herramientas de interpretación que permiten identificar los factores más relevantes para determinar la probabilidad de EHNA. Los resultados sugieren que los niveles sanguíneos de gamma glutamil transferasa y ferritina, así como la función renal (ckd-epi-creatinina) fueron los factores más relevantes a la hora de predecir la probabilidad de EHNA. Los modelos propuestos consiguieron una sensibilidad y especificidad del 70%, y un área bajo la curva ROC de 0,70. El modelo propuesto permite identificar las contribuciones de cada una de las variables identificadas para estimar la probabilidad final de EHNA, permitiendo caracterizar el impacto en cada paciente individual de sus variables fisiológicas.

1. Introducción

La enfermedad de hígado graso no alcohólico (EHGNA) se ha convertido en la enfermedad hepática crónica más frecuente [1]. La prevalencia mundial de EHGNA se estima en un 25%. La principal característica es la acumulación lipídica en el hígado, definida por la presencia de esteatosis en más de un 5% de las células hepáticas, sin que existan factores habituales como consumo alcohólico o enfermedad hepática [1], [2]. Una de las manifestaciones más típicas es la esteatohepatitis no alcohólica (EHNA) caracterizada por el daño hepatocelular en forma de balonización e inflamación. Esta forma de EHGNA es especialmente peligrosa, pues puede derivar en cirrosis y la necesidad de trasplante hepático [3]. Desde un punto de vista cardiovascular, EHGNA se puede considerar como una manifestación hepática del síndrome metabólico, y tienden a ser pacientes obesos y presentar hipertensión [1], [4]. No cabe duda, por lo tanto, de que EHGNA y EHNA son responsables de un aumento en los eventos cardiovasculares [5]. Sin embargo, persisten, al menos, dos importantes desafíos. En primer lugar, conseguir diagnósticos precoces y no invasivos. Típicamente, EHGNA cursa sin síntomas aparentes en los primeros estadios y la confirmación diagnóstica implica biopsia de hígado [6]. Se han propuesto algunos biomarcadores para la detección temprana, como los niveles de transaminasas, pero se ha comprobado que no existe una correlación significativa con la evolución de la enfermedad [7]. En segundo lugar, conseguir desarrollar tratamientos eficaces y específicos. Es necesario aumentar el conocimiento sobre la base fisiopatológica de la enfermedad con el objetivo de obtener tratamientos eficaces. En particular, y dado que la EHGNA va acompañada de otras enfermedades como obesidad, aterosclerosis y diabetes, es necesario ajustar los tratamientos a los diferentes pacientes [6], [7].

En este trabajo proponemos utilizar métodos de machine learning de clasificación basados en agregado de árboles de decisión binaria (Random Forest, RF) para crear modelos de predicción de EHNA [4], [8]. La base de datos etiquetada corresponde a un conjunto de pacientes hipertensos del Hospital Universitario de Móstoles, de los que se tiene información tanto demográfica como analítica. El objetivo del trabajo es, por un lado, obtener un modelo predictivo que permita estimar la probabilidad de desarrollar EHNA a partir de los datos clínicos. Por otro lado, y quizá más importante, es identificar las variables más relevantes para el desarrollo de EHNA, con el objetivo de aumentar el conocimiento patofisiológico del desarrollo de la enfermedad.

La estructura del artículo es la siguiente. En la Sección 2 se describe la base de datos y análisis exploratorio. En la Sección 3 se introduce brevemente RF. En la Sección 4 se presentan los resultados. Por último, en la Sección 5 se presenta la discusión y conclusiones del trabajo.

2. Base de datos y análisis exploratorio

Se realizó un estudio retrospectivo, transversal con pacientes referidos a la unidad de hipertensión del Hospital Universitario de Móstoles (España) entre 2005 y 2017. El total de pacientes incluidos en el estudio fue de 2.828, de los que 44 fueron excluidos por datos incompletos. A partir de sus registros electrónicos e historias clínicas se recogieron los datos demográficos (edad, sexo), antropométricos (talla, peso), clínicos (presión arterial sistólica y diastólica) y analíticos (perfil

ISBN: 978-84-09-25491-0, pags. 177- 180

férrico y lipídico, glucemia basal, hemoglobina glicosilada e insulinemia plasmática basal, colesterol). La base de datos final incluye un total de 2.784 pacientes, de los que 646 tenían EHNA (23.2%) [4].

Un paso fundamental para el buen desempeño de modelos de machine learning es el de análisis exploratorio, feature-engineering y el tratamiento de valores perdidos y atípicos. Para cada una de las características, se establecieron unos límites fisiológicos, de forma que se considera un valor atípico cualquier valor fuera estos. En primer lugar, se eliminan de la base de datos todos los pacientes que contienen más de la mitad de las características con valores perdidos y/o atípicos. Antes de realizar el método de imputación, se realiza previamente una partición de los datos en un conjunto de training y de test. Los valores perdidos y/o atípicos de ambos conjuntos se sustituyen por la mediana de cada variable del conjunto de training. Se realiza una transformación logarítmica para variables con histograma tipo exponencial. En concreto: GGT, HbA1C, PCR, albúmina, albuminuria, cpk, ferritina, glucemia, índice de homa, insulina, IST, triglicéridos, vitamina D.

3. Modelo de clasificación basado en Random Forest

3.1. Random Forest

Random Forest (RF) es un algoritmo de predicción basado en el agregado de distintos árboles de decisión, en el que cada árbol se construye de forma independiente utilizando una remuestra con reemplazamiento a partir del conjunto de datos inicial (bootstrap). De esta forma, cada árbol individual realiza una predicción, siendo el resultado final elegido por votación, en el caso de clasificación. El objetivo de esta aproximación es construir predictores incorrelacionados, con el objetivo de reducir el sobreajuste, mejorando las prestaciones de generalización [7]. RF posee una serie de parámetros libres, hiperparámetros, de los cuales seleccionamos el mínimo de muestras en cada hoja del árbol y el número de variables evaluables en cada ramificación del árbol para la búsqueda de los hiperparámetros óptimos.

Con el objetivo de utilizar las mejoras prácticas en el machine learning, el primer paso consiste en separar el conjunto de datos en un conjunto de training (80%) y un conjunto de test (20%). El análisis exploratorio y cálculo de valores para imputación de valores perdidos o atípicos, se realiza con el conjunto de training. Asimismo, la búsqueda de los hiperparámetros óptimos, se realizará sobre el conjunto de training utilizando la técnica de 10-fold cross validation (10-fold CV), utilizando como figura de mérito para la comparación el accuracy.

3.2. Balanceado de datos

Los datos biomédicos reales suelen adolecer de desbalanceo, de forma que los pacientes que presentan enfermedad suelen estar subrepresentados. Este problema está presente en nuestra base de datos de EHNA, en la que los pacientes sin enfermedad son 4 veces más numerosos que con enfermedad, aproximadamente. Los conjuntos de datos desbalanceados suponen un problema para los algoritmos de clasificación, dado que tienden a proporcionar soluciones que favorecen la clase más representada, optimizando el accuracy (porcentaje de acierto) [4]. De hecho, se recomienda trabajar con otro tipo de métricas, tales como *sensibilidad, especificidad* y F1-Score. Por ello, se decide aplicar dos diferentes estrategias de desbalanceo sobre el conjunto de training del modelo: (1) ponderación de clases, en el cual se modifican los pesos de las clases en la división de cada árbol, de forma que, la penalización de los errores de la clase minoritaria sea mayor; y (2) undersampling, que reduce aleatoriamente el número de pacientes de la clase mayoritaria hasta igualarlo al número de pacientes de la clase minoritaria, de forma que se tienen las mismas proporciones de las diferentes clases [9].

3.3. Interpretabilidad fisiológica de los modelos

Los árboles de decisión tienen la gran ventaja de ser muy interpretables y, por tanto, de ayudar en la comprensión del desarrollo de la predicción y de los resultados; ofrece toda la información de la toma de decisiones en la clasificación de una muestra específica. Los modelos RF mejoran las prestaciones de desempeño agregando árboles, a costa de perder algo de la interpretabilidad original. Sin embargo, existen diferentes herramientas, en algunos casos exclusivas de RF, para la interpretación y análisis de los resultados. En este trabajo utilizamos: (1) *feature importance:* basada en el cálculo de la contribución relativa de cada variable en el modelo diseñado. Estos valores se calculan siguiendo un método basado en el incremento de pureza del nodo. (2) *partial dependence plots (PDP)* las cuales aportan información acerca del efecto marginal que tiene cada característica en la predicción de EHNA. En concreto, se representa la probabilidad de padecer EHNA según varían valores de una variable, para el valor medio de las restantes. (3) *tree interpreter:* cada predicción (en forma de probabilidad) se puede descomponer en la suma de las contribuciones de cada característica. En RF, la predicción final es la media de las predicciones de todos los árboles [8].

Estas técnicas permiten ordenar las variables en función de la importancia a la hora de predecir la probabilidad de EHNA. Adicionalmente, utilizaremos esta característica para realizar selección de características en los modelos, reteniendo las variables más importantes, utilizando 10-fold CV. Se utilizarán los PDPs para entender la relación entre las variables y la probabilidad de EHNA. Por último, se utilizará el *tree interpreter* para entender cuál es la contribución de cada variable fisiológica a la hora de aumentar la probabilidad de padecer EHNA.

4. Resultados

El proceso de training de cada modelo fue: (1) selección de hiperparámetros utilizando 10-CV; (2) selección de características ordenando las características por *feature importance,* a continuación, se eliminan las características una a una desde las menos importantes a más, y midiendo el accuracy con 10-CV. La Tabla 1 muestra la

comparación de los modelos con variables transformadas sin y con las dos técnicas de desbalanceado, para las diferentes métricas utilizadas, y después de seleccionar las características más relevantes. Analizando los resultados de los diferentes modelos RF construidos, el que ofrece el mejor conjunto de prestaciones es el RF con variables transformadas, undersampling, mínimo número de muestras en cada hoja: 10 y número de variables en cada ramificación: 0.5. Este modelo tiene una mayor tasa de detección de pacientes con EHNA, y una menor probabilidad de pérdida.

	Modelo sin técnicas balanceo	Modelo usando class weights	Modelo usando undersampling
AUROC	0,640	0,675	**0,704**
Sensibilidad	0,336	0,425	**0,709**
Especificidad	**0,943**	0,924	0,700
F1-score	0,443	0,511	**0,534**
Tasa de acierto	0,797	**0,804**	0,702

Tabla 1. *Prestaciones de los diferentes modelos RF, con los datos desbalanceados (primera columna), y con dos métodos de balanceo de datos (segunda y tercera columna).*

El modelo RF Undersampling que presenta una mayor tasa de acierto es el creado con las 3 variables más relevantes en la predicción: GGT_log (niveles sanguíneos de gamma glutamil transferasa), ferritina_log (niveles sanguíneos de ferritina) y Epi.creat (función renal estimada según ckd-epi-creatinina).

La importancia relativa de estas variables se muestra en la Figura 1. Se puede observar que la variable GGT_log tiene una destacable contribución en el modelo, siendo la más influyente en la predicción de EHNA (53% aproximadamente).

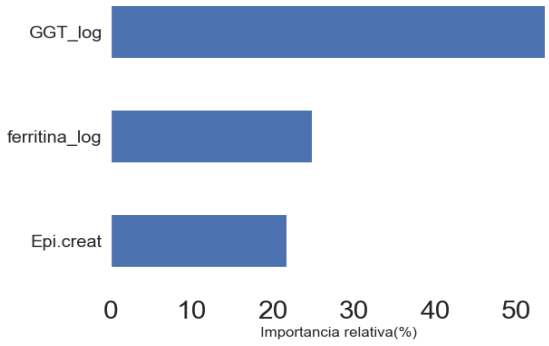


Figura 1. *Importancia de cada variable en el modelo RF variables transformadas undersampling.*

Una vez identificadas las variables más influyentes, analizamos su efecto marginal en la predicción de EHNA. La Figura 2 muestra los PDPs de las dos variables más importantes (GGT_log y ferritina_log). Se puede observar que, para ambas variables, cuanto mayor son, mayor es la probabilidad de padecer EHNA. La influencia es mayor para el caso de GGT_log. existiendo para niveles próximos de 28,5 GGT (3,35 GGT_log), una probabilidad estimada del 50%, aproximadamente. En ambos casos existe un umbral a partir del cual aumenta notablemente la probabilidad de padecer EHNA.

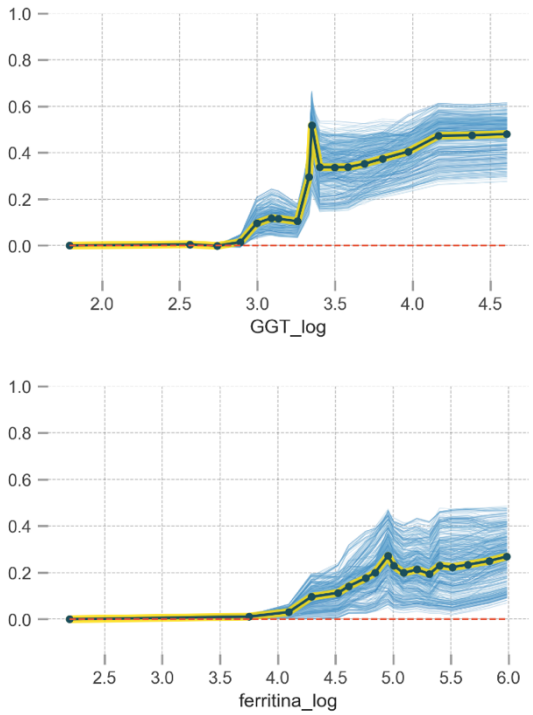


Figura 2. *PDPs de las variable GGT_log (panel superior) y ferritina_log. (panel inferior).*

Por último, calculamos las contribuciones de las variables en la predicción de EHNA (tree interpreter). Este análisis se realiza sobre los subconjuntos de pacientes que fueron diagnosticados con y sin esta enfermedad (Tabla 2). Como se puede observar, de nuevo, la variable que más contribuye a la probabilidad final es GGT_log.

	Pacientes con EHNA	Pacientes sin EHNA
GGT_log	0,070	-0,090
ferritina_log	0,026	-0,019
Epi.creat	0,012	-0,016

Tabla 2. *Contribuciones de las variables del modelo RF variables transformadas undersampling.*

A modo de ejemplo, mostramos las contribuciones de las variables en la predicción de EHNA sobre dos pacientes correctamente clasificados con y sin la enfermedad. Se puede observar en la Figura 3 que, para el paciente con EHNA, estas tres variables influyen positivamente en la predicción de EHNA, destacándose considerablemente la variable GGT. Por el contrario, en el paciente sin EHNA, destaca la variable Epi.creat con un notable efecto negativo en la predicción de EHNA.

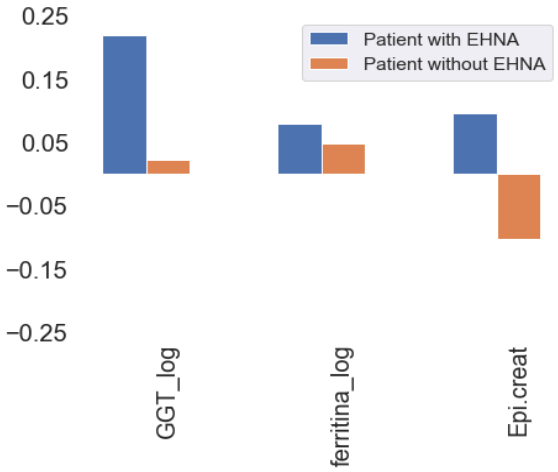


Figura 3. *Contribuciones de las variables de un paciente, con y sin EHNA, que ha sido correctamente clasificado.*

1. Discusión y conclusiones

El objetivo del trabajo fue diseñar un modelo de predicción que permita estimar la probabilidad de desarrollar EHNA a partir de datos clínicos, así como identificar las variables más relevantes en el desarrollo de EHNA, con el fin último de aumentar el conocimiento patofisiológico de la enfermedad. Proponemos utilizar el algoritmo RF porque, además de considerarse como uno de los principales esquemas de machine learning, permite realizar un exhaustivo análisis de interpretación. Sin embargo, hemos comprobado que, a pesar de su alta capacidad de generalización, es necesario aplicar alguna estrategia de balanceo de datos para garantizar una mayor fiabilidad no sólo en los resultados del modelo, sino también en su interpretación. Para cualquier tipo de problema, el diseño de modelos a partir de datos desbalanceados puede ofrecer prestaciones poco relevantes y, por consiguiente, interpretaciones erróneas.

Analizando los diferentes modelos, las mejores prestaciones se obtienen con el modelo RF con variables transformadas, undersampling, mínimo número de muestras en cada hoja: 10 y número de variables en cada ramificación: 0,5. Usando sólo GGT_log, ferritina_log y Epi.creat, este modelo alcanza un área bajo la curva ROC de 0,70 y una sensibilidad y especificidad del 70%.

En la predicción de EHNA, destacan las variables GGT_log con una importancia relativa muy superior (más del doble, aproximadamente) al resto de variables y, ferritina_log, como la segunda variable más relevante. Los niveles elevados de GGT en la EHNA se explican por la filtración de la enzima debida a la existencia de lesiones en el hígado. La ferritina puede estar elevada en la EHNA debido a una inflamación sistémica o al aumento de las reservas de hierro [4]. Varios estudios recientes han identificado niveles elevados de ferritina como un predictor independiente de EHNA y fibrosis hepática entre pacientes con EHGNA, porque la ferritina puede predecir la gravedad histológica de lesión hepática en pacientes con EHGNA [10]. Por lo tanto, ambas variables fisiológicas pueden considerarse un biomarcador no invasivo para identificar pacientes hipertensos con alto riesgo de EHNA. Una de las limitaciones del estudio es el desempeño predictivo, por lo que una de las líneas futuras del trabajo consiste en mejorar esta capacidad, de forma que puede realmente ser utilizado en la práctica clínica.

Los algoritmos de machine learning pueden servir de apoyo en el diagnóstico y conocimiento de enfermedades. Este trabajo identifica a las variables fisiológicas GGT y ferritina como mejores biomarcadores para predecir el desarrollo de EHNA. Actualmente, no existe una modalidad no invasiva para predecir el desarrollo de esta enfermedad, pero estos biomarcadores se pueden utilizar en la práctica clínica como prueba de detección de pacientes hipertensos para una biopsia de hígado y un tratamiento más especializado.

Agradecimientos

Este trabajo ha sido parcialmente subvencionado por los proyectos de investigación: TEC2016-75361-R y PID2019-105032GB-I00 del Ministerio de Ciencia e Innovación.

Referencias

[1] N. Chalasani et al., “The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases: Hepatology, 2017,” Hepatology, vol. 67, no. 1, pp. 328–357, 2018.

[2] VV.AA., “EASL–EASD–EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease,” Journal of Hepatology, vol. 64, no. 6, pp. 1388–1402 2016.

[3] D. L. White et al., “Association Between Nonalcoholic Fatty Liver Disease and Risk for Hepatocellular Cancer, Based on Systematic Review,” Clinical Gastroenterology and Hepatology, vol. 10, no. 12, pp. 1342-1359.e2, 2012.

[4] R. Garcia-Carretero, et al., “Relevant Features in Nonalcoholic Steatohepatitis Determined Using Machine Learning for Feature Selection,” Metabolic Syndrome and Related Disorders, vol. 17, no. 9, pp. 444–451, 2019.

[5] M. R. Pagadala and A. J. McCullough, “The Relevance of Liver Histology to Predicting Clinically Meaningful Outcomes in Nonalcoholic Steatohepatitis,” Clinics in Liver Disease, vol. 16, no. 3, pp. 487–504, 2012.

[6] N. Tanaka, et al., “Current status, problems, and perspectives of non-alcoholic fatty liver disease research,” WJG, vol. 25, no. 2, pp. 163–177, 2019.

[7] T. Hastie, R. Tibshirani, J.Friedman, “The elements of statistical learning”, Springer, 2009.

[8] C.C. Mark, et al. Current modalities of fibrosis assessment in nonalcoholic fatty liver disease. Journal of clinical and translational hepatology, 5(3):261, 2017.

[9] A. Fernández, S. García, M. Galar, Ronaldo C. Prati, B. Krawczyk, and F. Herrera. “Learning from Imbalanced Data Sets”, Springer, 2018.

[10] K.V. Kowdley, P. Belt, L.A. Wilson, et al. Serum ferritin is an independent predictor of histologic severity and advanced fibrosis in patients with nonalcoholic fatty liver disease. Hepatology 2012;55:77–8.

Design and implementation of predictive models based on radiomics to assess response to immunotherapy in non-small-cell lung cancer

M. Corral Bolaños[1,2], B. Farina[2], A.D. Ramos Guerra[2], C. Palacios Miras[3], G. Gallardo Madueño[4], A. Muñoz-Barrutia[5], G. R. Peces-Barba[3], L. M. Seijo[4], J. Corral[4], I. Gil-Bazo[4,6], M. Dómine Gómez[3], M. J. Ledesma-Carbayo[2]

[1] ETSI de Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain, m.corral@alumnos.upm.es

[2] Biomedical Image Technologies, Universidad Politécnica de Madrid & CIBER-BBN, Madrid, Spain

[3]Hospital Universitario Fundación Jiménez Díaz & CIBERES, Madrid, Spain

[4]Clínica Universidad de Navarra & CIBERES & CIBERONC, Madrid, Spain

[5]Dpto. Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid & IiSGM, Spain

[6]Instituto de Investigación Sanitaria de Navarra (IDISNA), Pamplona, Spain

Summary

Lung cancer is the leading cause of cancer-related deaths in Europe. Immunotherapy treatments have been proved as the new standard of care for stage III-IV non-small cell lung cancer patients. However, the treatments vary in success, and there is not a reliable biomarker. This retrospective project aimed to develop a predictive model based on radiomics through machine learning or deep learning techniques to assess the response to the treatment, understood as the progression (or not) of the disease. Then, the study was complemented with an analysis of the progression-free survival time and an attempt of association with biological data.
We used the basal computed tomography images of the primary tumour lesions from a cohort with 84 patients with IV stage non-small-cell lung cancer. The best performance model reached an AUC of 0.80 – 90 % CI [0.62, 0.99]. Our results suggest that the radiomics models may be useful for patient classification.

1. Introduction

Lung cancer is the leading cause of cancer-related deaths in Europe. In 2016, 239 000 people died from lung cancer, more than one fifth (20.5 %) of all deaths from cancer [1]. In most patients, lung cancer is diagnosed once it has reached an advanced stage, with more than 60 % of patients presenting advanced or metastatic stage, with a typical survival time of less than a year [2]. At this point, new treatments, such as immunotherapy and targeted therapies have emerged as new standard of care treatments for non-small-cell lung cancer (NSCLC) patients, the most common subtype.

Despite their success, clinical benefits are only observed in a subset of patients. Hence, there is an important need of development of predictive markers to assess which treatment is more likely to be appropriate for each patient and to allow the tumour surveillance during the treatment, since immunotherapy may present important side effects and is expensive. Different biomarkers have been studied with limited success for patient stratification according to their potential benefit before the treatment, such as gene alterations or the expression of immune checkpoints proteins, like the PD-L1 [3]. The current standard to determine the treatment response is based mainly on tumour size evolution according to the RECIST and iRECIST criteria. Unfortunately, targeted treatments can cause morphological changes without changing size, so it seems insufficient.

The solution to this challenge could be found in radiomics. Radiomics is the high-throughput extraction and analysis of quantitative features from medical imaging to transform images into mineable high-dimensional data. This process is motivated by the idea that images reflect the underlying pathophysiology, and the quantitative features offer information on the tumour phenotype and its microenvironment.

In this study, we hypothesize that the basal computed tomography images and the radiomic features of the main tumour can be used to develop predictive biomarkers with prognostic value about the progression of the disease.

2. Methods

2.1. Patients cohort

Patients with confirmed stage IV NSCLC receiving immunotherapy treatment from January 2013 to December 2019 at Hospital Universitario Fundación Jiménez Diaz (FJD) and Clínica Universidad de Navarra (CUN) were analysed retrospectively after the approval of the corresponding institutional review boards. Immunotherapy treatment could be monotherapy, a combination of immune-based treatments or immunotherapy combined with chemotherapy or radiotherapy. A patient was excluded from the study if the clinical data were not available, primary tumour boundaries were not clear or there was not basal CT. Finally, 84 patients were included in the study, 60 with progression and 24 with not progression. Our models aimed to predict the progression of the disease (labelled as 1), or the non-progression (labelled as 0). The progression was defined by an oncologist based on radiological or clinical evidence on patient status. The whole dataset was divided into a training set with 58 patients (70 %) and a validation set with

ISBN: 978-84-09-25491-0, pags. 181- 184

26 patients (30 %), they were balanced to get a similar proportion of progression/non-progression response to the treatment. From the whole dataset we had information of the PD-L1 expression levels in 34 patients.

2.2. Imaging and tumour segmentation process

In this project, the basal computed tomography lung images were employed. The acquisition and reconstruction parameters are summarized in Table 1.

The segmentation of the main lesions was performed using syngo®.via software. This tool provides a semiautomatic seed-based method for tumour segmentation, although manual correction was commonly required. After segmentation, images were exported to DICOM format and transformed into Nearly Raw Raster Data (NRRD) [4] format through a MATLAB script to work with them in other environments.

Acquisition		
Tube voltage	80 KVP	1 image
	100 KVP	22 images
	120 KVP	58 images
	140 KVP	3 images
Tube current	< 300 mA	34 images
	300 – 500 mA	21 images
	> 500 mA	28 images
Exposure time	285 ms	1 image
	426 – 493 ms	12 images
	500 – 562 ms	71 images
Image reconstruction		
Slice Thickness	1 mm	38 images
	1.5 mm	22 images
	2 mm	21 images
	3 mm	2 images
Convolutional kernel	B, B26f, B30f, B31f	59 images
	C	16 images
	FC01	9 images

Table 1: *Image acquisition and* reconstruction *parameters.*

2.3. Radiomic features extraction and test-retest

The feature extraction was performed through *PyRadiomics*, an open-source Python package [5]. The radiomic features were extracted from both intranodular and perinodular regions, as well as from both regions merged. The perinodular region mask was generated through a morphological dilation of the tumour mask. It was used a 4 mm, 5 mm, or 6 mm radius 3-D spherical structural element for tumours with major axis length (2D) <25 mm, ≥ 25 mm and 50 mm, respectively. After dilation, the tumour mask was subtracted from the dilated area which produced the border mask.

In this project, first order features, shape-based features and texture features (obtained from the Gray Level Cooccurrence Matrix, GLCM; Gray Level Run Length Matrix, GLRLM; Gray Level Size Zone Matrix, GLSZM; Neighbouring Gray Tone Difference Matrix, NGTDM; and Gray Level Dependence Matrix, GLDM) were extracted from original images as well as from different filtered images: the Laplacian of gaussian, wavelet decomposition, square root, and local 3D binary pattern. Shape features were not extracted for the perinodular mask. Moreover, we performed the feature extraction with four different pairs of hyperparameters associated with the resampling resolution and bin width discretization to choose the best performance radiomic features subset.

Then, non-reproducible features were discarded using test/re-test scans from the Reference Image Database to Evaluate Therapy Response (RIDER) dataset [6]. This data set is composed of 31 NSCLC patients who underwent two chest CT scanners by using the same imaging protocol and the same scanner with a difference of 15 minutes. All those features whose Lin's concordance correlation coefficient value was lower than 0.9 were discarded.

2.4. Machine learning model building

We developed four types of models attending to the different feature extraction regions of interest: a model for the intranodular region, the perinodular region, the model which merged the regions and a model which uses both intranodular and perinodular features (Figure 1). For the feature selection and training strategy, we used the Python open-source package *Scikit-learn* [7].

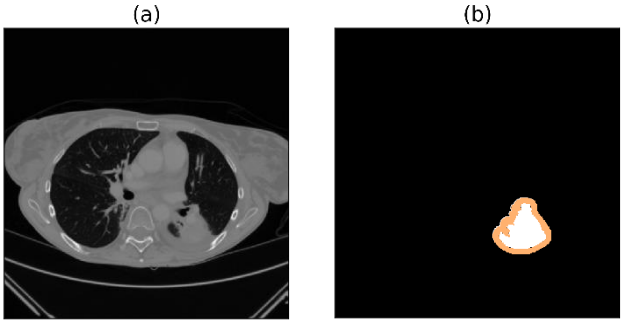


Figure 1: *a) Original CT slide, b) In orange, perinodular mask. In white, intranodular mask. The union of both masks composes the merged mask.*

To avoid the *curse of dimensionality* several steps for dimensionality reduction were performed. Firstly, Pearson's correlation coefficient was used to remove one feature from each pair of highly correlated features. Afterwards, we used two additional methods for feature selection based on a logistic regression algorithm: recursive feature selection (a wrapper method) and LASSO feature selection (an embedded method). Then, each subset of selected features was used to train three different types of machine learning classical algorithms: logistic regression, support vector machines (SVM) and random forest, and their performance was evaluated on the validation test.

For both feature selection and training stages, each classifier was trained on the training set using 3-fold cross-validation. Moreover, we used Bayesian sequential model-based optimization (SMBO) for hyperparameter tuning with the Python package *Scikit-optimize* [8]. In all the steps, the classifier which achieved the highest AUC score was selected as the candidate solution.

2.5. Deep learning model building

Due to the limited number of available samples, we opted to develop a convolutional neural network (CNN) model through transfer learning on an already developed CNN model for prediction of lung nodule malignancy [9], whose architecture followed the one published by Causey et al. [10], *NoduleX*. This task was done with *Keras* software package (a Python deep learning API) [11].

The input of the CNN was a small 3D volume of dimensions 5x47x47 pixels. These volumes have been extracted selecting the centroid of the tumour and getting the above and below slices through a Python implementation. The pixel values were rescaled by dividing by the maximum pixel intensity in the cropped volume. To deal with overfitting, real-time data augmentation has been performed. Since the model is developed through transfer learning, the weights of the first layers are exported from the previous model, while only the three top layers are trained to adjust the output to the new task during 200 epochs. Moreover, we tested the effect of *early stopping* on the models. In this case, we use the 20 % of samples of the training set for validation (i.e., 46 samples for training, 12 for validation, 26 for testing). After, we also tested to *unfreeze* all layers and adjust their weights with a small learning rate and during just 30 epochs.

2.6. Model selection and statistical analysis

We considered various metrics to choose the candidate solutions for each region of interest. Firstly, all classifiers whose AUC differed more than 0.05 between training and validation were discarded, and when there were two similar performance models, the one with the lowest standard deviation during training was preferred. We computed the 90 % confidence interval (CI) of the validation AUC using the Delong method. To check whether the classifier achieved a significant stratification for each kind of patient (i.e., if the classifier probability scores are significantly different between the patients in which the disease progresses or not), we calculated the p-value using the Mann-Whitney U test. We used the Youden Index to select the optimal cut-off point of the classifiers. Then, we compute other metrics such as the accuracy, balanced accuracy, sensitivity, and specificity to get a better insight into the behaviour of the models. All metrics were compared with a baseline classifier which always classified the most common class.

To correlate the classifier behaviour with biological meaning, we used the spearman rank-order correlation to analyse the relevance between the PD-L1 expression levels and the selected features of the best performance classifier.

Finally, we evaluated the progression-free survival (PFS) (i.e., the time between the first immunotherapy cycle and the progression/death date) with the Kaplan-Meier method for each one of the stratified groups by the selected classifier. Endpoints were the progression of the disease, death from any cause or any recurrence. The log-rank test was used to compare the survival distributions of the two groups. Moreover, we used Cox's proportional hazard model to evaluate how the different features (covariates) of each sample affect the survival time of the subject. The Kaplan-Meier curves and Cox's models have been developed through the Python open-source package *lifelines* [12].

3. Results

The main issue of most of the developed models was their high variance between the training and validation performance. Indeed, no perinodular model achieved the requirements to be considered. Anyway, some models of the other regions of interest reached acceptable stability (Figure 2). In the intranodular region of the tumour, an SVM model achieved an AUC of 0.70 – 90% CI [0.52, 0.88], however, its accuracy was worse than the baseline classifier, so it had no practical purpose. For the features in the merged region, the best classifier reached an AUC of 0.76 – 90 % CI [0.58, 0.93]. Anyway, the best performance was achieved by a logistic regression method which uses 20 selected features by recursive feature elimination among the intra- and perinodular features, AUC of 0.80 – 90 % CI [0.62, 0.99].

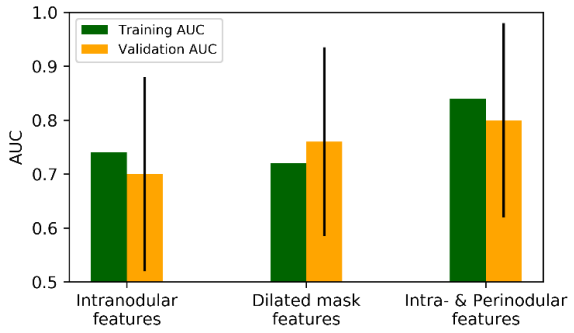


Figure 2: *Performance of the best classifiers in the training and validation set. The validation set shows the 90 % CI.*

The 20 selected features were extracted with a voxel resampling of 3 mm isotropic resolution and 25 bin width discretization. Five of them correspond to the intranodular region, and 15 to the perinodular region. Moreover, just one was computed from the original image, and the other 19 were computed from filtered images. The spearman rank-order correlation revealed three significant correlations between the features and the PD-L1 expression: the intranodular wavelet-HHH GLSZM Small Area Emphasis, perinodular wavelet-LHH first order Root Mean Squared and the perinodular square root NGTDM Busyness (p-value < 0.05). Nonetheless, the correlation coefficient value of these features is below 0.5, which means soft correlation.

Regarding the CNN models, all of them obtained similar results with AUCs between 0.643 and 0.703 in the validation phase. The model with slightly better performance was the one with early stopping after 153 epochs and only the training of the three last layers. This model reached an AUC of 0.703 - 90% CI [0.48, 0.92]. To compare the machine learning and the deep learning model, we used additional metrics which are presented in Table 2 and the ROC curves (Figure 3)

	p-value	AUC [90% CI]	Acc.	Sens.	Spec.
ML	**0.01**	**0.80 [0.62, 0.99]**	**0.85**	0.84	**0.86**
DL	0.06	0.70 [0.48, 0.92]	0.81	0.84	0.71

Table 2: *Different metrics for the ML and DL performance. Acc: Accuracy, Sens: Sensitivity, Spec: Specificity*

We have used the selected machine learning method for the PFS analysis. There is not a significant difference between the Kaplan-Meier curves for each of the classified groups by the machine learning algorithm, log-rank test p-value = 0.55.

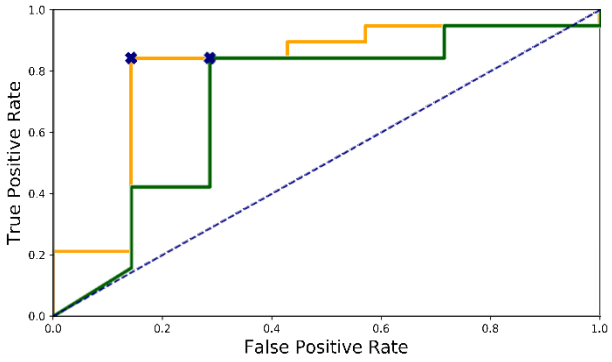


Figure 3: *ROC curve of the ML model (in yellow) and the DL model (green). The blue cross represents the optimum threshold.*

We used the Cox's models to analyse the effect of the 20 features in the PFS time of the patients. Two of them obtained significant results (p-value < 0.05): the perinodular wavelet-HHL NGTDM Strength (the higher the value, the shorter PFS) and perinodular wavelet-LLL NGTDM Contrast (the higher the value, the longer PFS). However, there was no significant association between the classifier scores and the PFS time of the validation patients (p-value = 0.26). But if we only consider the group of the progression patients, results were borderline significant (p-value = 0.05).

4. Discussion

This project aimed to analyse the suitability of basal CT images of advanced NSCLC to build radiomic biomarkers.

The development of different models considering different regions of interest allowed us to confirm the relevance of the surrounding peritumoral area. By its own, the perinodular regions did not obtain stable results but combined with the intranodular features the model achieved good classifications results with high sensitivity and specificity despite the imbalance of our dataset. These results support the current studies which assert the biological relevance of the surrounding regions due to different reasons such as the tumour vascularization or the lymphocyte infiltrations. We attempted to integrate biological evidence into our model with the PD-L1 correlation since it is a key step to be able to develop clinically applicable models. Three features showed significant results, although the limited number of patients with this information does not allow us to consider them definitive findings.

One of the main limitations of our work is the limited number of patients in the study. Hence, we were not able to develop a complete deep learning model with our data. Nonetheless, this first transfer learning approach obtained comparable results with respect to the machine learning model in most of the metrics (despite that the CNN classification was not significant, p-value = 0.06).

The PFS time analysis also suffered the lack of non-progression patients, not showing significant difference in the Kaplan-Meier curves. This analysis was deepened through the Cox's models, in which we observed the same effect since we can associate the classifier scoring to the PFS in only progression patients, but not in non-progression patients. Besides, two radiomic features revealed some association between our radiomic biomarker and the PFS.

5. Conclusion

Despite the limitations of this work, mainly associated to the limited number of patients, this analysis succeeds to demonstrate the predictive value of the intranodular and perinodular radiomic features as a feasible way to develop cheap, non-invasive and easy-obtainable biomarkers for immunotherapy effectiveness prediction at a patient-level. The current use of CT images in the follow-up of NSCLC patients motivates to continue in this direction, increasing the number of available patients and associate the model's behaviour to biological tumour biomarkers to be able to develop a clinically applicable radiomics model to support the oncological decision-making process.

6. Acknowledgements

This work was partially funded by projects RTI2018-098682-B-I00 and PID2019-109820RB-I00 (MCIU/ AEI/ FEDER, UE) by the Spanish Ministry of Science and Innovation, and the Leonardo grant 2019 from Fundación BBVA. MCB, BF and ADRG were supported by the Spanish Ministry of Education and Comunidad de Madrid.

References

[1] "Cancer statistics - specific cancers," *Eurostat*, Aug. 2020. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Cancer_statistics_-_specific_cancers#Lung_cancer (Oct. 18, 2020).

[2] C. Gérard and C. Debruyne, "Immunotherapy in the landscape of new targeted treatments for non-small cell lung cancer," *Molecular Oncology*, vol. 3, no. 5–6. John Wiley and Sons Ltd, pp. 409–424, Dec. 01, 2009.

[3] P. Villalobos and I. I. Wistuba, "Lung Cancer Biomarkers," *Hematology/Oncology Clinics of North America*, vol. 31, no. 1. W.B. Saunders, pp. 13–29, Feb. 01, 2017.

[4] "nrrd: Definition of NRRD File Format." http://teem.sourceforge.net/nrrd/format.html (Jun. 18, 20).

[5] J. J. M. van Griethuysen *et al.*, "Computational radiomics system to decode the radiographic phenotype," *Cancer Research*, vol. 77, no. 21, pp. e104–e107, Nov. 2017

[6] B. Zhao, L. H. Schwartz, and M. G. Kris, "Data From RIDER_Lung CT." The Cancer Imaging Archive., 2015.

[7] F. Pedregosa, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[8] T. Head, "scikit-optimize/scikit-optimize: v0.5.2." [Sw package] 2018.

[9] B. Farina, "Deep Learning-Based Models for Prediction of Lung Nodule Malignancy with CT scans," Master Thesis, University Federico II, 2018.

[10] J. L. Causey *et al.*, "Highly accurate model for prediction of lung nodule malignancy with CT scans," *Scientific Reports*, vol. 8, no. 1, 2018, doi: 10.1038/s41598-018-27569-w.

[11] F. Chollet and others, "Keras." [Sw package] 2015.

[12] C. Davidson-Pilon *et al.*, "CamDavidsonPilon/lifelines: v0.24.9.." [Sw package] 2020,

Diseño y fabricación de un sistema para el cultivo celular rotatorio en paredes de dispositivos microfluídicos con canales circulares embebidos

X. Bujanda Cundin[1] O. Mitxelena-Iribarren [1,2]

[1] Universidad de Navarra, Tecnun, Manuel Lardizabal 13, 20018 Donostia / San Sebastian, España, a904579@alumni.unav.es

[2] Ceit, Manuel Lardizabal 15, 20018 Donostia / San Sebastian, España, omitxelena@ceit.es

Resumen

El desarrollo de vasos sanguíneos funcionales es hoy en día un pilar fundamental en la evaluación de nuevas terapias y agentes de diagnóstico. En este artículo se describe la fabricación y posterior funcionalización, mediante cultivo celular, de un dispositivo microfluídico con sección circular. Su finalidad es simular un vaso sanguíneo con el objetivo de probar nuevos tratamientos para la hipertensión arterial pulmonar. La fabricación fue llevada a cabo mediante un proceso en el que un hilo de sección circular determina las dimensiones del canal. Para simular la histología de los vasos sanguíneos, se realizó un cultivo celular rotatorio con el fin de adherir las células a la pared de los dispositivos. Este es un método simple y reproducible que permite generar modelos de vasos sanguíneos in vitro.

1. Introducción

En las últimas décadas, la microfluídica se ha convertido en una herramienta muy útil para simular vasos sanguíneos in vitro y realizar estudios farmacológicos o terapéuticos en estos. El uso de técnicas como *soft-lithography* para fabricar dispositivos microfluídicos con canales circulares [1] ha permitido simular con mayor precisión el sistema cardiovascular en comparación con los dispositivos tradicionales que presentaban canales rectangulares [2]. Esta no es la única técnica empleada para obtener este tipo de canales. Métodos como *microwire-molding* [3] o la inyección de PDMS líquido en canales rectangulares y a continuación aire, permiten obtener este tipo de dispositivos [4].

La principal ventaja de trabajar con canales circulares es que permiten reproducir las fuerzas que sufre la sangre al circular por los vasos; esto no se podría lograr mediante canales rectangulares. [5]. Otras ventajas de estos es que al trabajar a escala microscópica el flujo es laminar, las fuerzas de tensión superficial dominan respecto de la gravedad en escala microscópica y son utilizadas para dirigir el flujo sin necesidad de control externo y, además, los tiempos de reacción son mucho menores que a escala macroscópica debido a que la difusión se da de forma más rápida [6].

La funcionalización de los dispositivos para simular vasos sanguíneos se lleva a cabo mediante el co-cultivo de dos tipos celulares. Es un cultivo dinámico en el que el dispositivo sufre una rotación que puede ser de diferente naturaleza. Los cultivos celulares rotatorios automáticos, se caracterizan por utilizar un eje rotatorio que permite dar vueltas al dispositivo tras haber insertado las células en su interior [7]. Los cultivos celulares rotatorios manuales, se basan en la inserción de células en repetidas ocasiones: tras cada inserción, se gira el dispositivo manualmente una cantidad determinada de grados. Tras completar una vuelta, el dispositivo se incubará estáticamente [8]. Por último, el cultivo orbital se distingue por el uso de un agitador orbital de laboratorio que permite los desplazamientos verticales y horizontales de los dispositivos [9].

2. Materiales y métodos

2.1. Fabricación y caracterización de los dispositivos microfluídicos

La técnica empleada para fabricar los dispositivos microfluídicos con canales circulares fue una variante de la técnica conocida como *microwire*. En primer lugar, se elaboró un molde mediante impresión 3D con el fin de estandarizar la fabricación de dispositivos. El molde se perforó a los lados con una broca por donde se introdujo el hilo que proporcionaría la estructura a los canales y en la parte inferior se dispuso un cristal para contener el polímero.

El polímero empleado fue PDMS, Sylgard 184, Dow Corning, con una ratio 10:1 entre el polímero y el agente de curado. En una mezcladora se agregaron ambos componentes durante 3 minutos y 30 segundos y, a continuación, no se realizó desgasificado. Tras introducir el hilo por los agujeros, se selló el molde con *parafilm* con el objetivo de evitar fugas. A continuación, se vertió el polímero y se dejó curar durante 48 horas a temperatura ambiente. Finalmente, se procedió a la extracción del hilo y cada canal fue cortado en canales más pequeños. Se emplearon 3 diámetros de hilo: 0.29mm, 0.20mm y 0.12mm.

En la caracterización de los dispositivos fabricados se utilizaron 4 métodos para analizar fugas, problemas en la continuidad de los canales y la topografía de la pared interna de los mismos. El primer método se basó en la circulación de un colorante por los canales empleando un tinte, una jeringuilla y un conector. El segundo método empleado fue la microscopía en el que se usaron tres microscopios: Leica M60, Nikon Eclipse Ti y un microscopio electrónico de barrido Phenon PRO-SEM de la marca Phenon World. En este último, la muestra se tuvo que recubrir mediante un material

ISBN: 978-84-09-25491-0, pags. 185- 188

conductor para poder ser analizada en el microscopio. El proceso de recubrimiento se llevó a cabo en una cámara de vacío de la marca Emitech, donde mediante *sputtering* la superficie de la muestra quedó cubierta de oro. En el último método de caracterización de dispositivos y, siguiendo el protocolo empleado por el grupo de investigación en el que se llevó a cabo este proyecto, se lograron fijar esferas fluorescentes en la pared de los dispositivos [10]. Mediante un microscopio de fluorescencia Nikon Eclipse Ti, se analizó la superficie interna de los canales.

2.2. Fabricación de un sistema rotatorio para la fijación celular en canales circulares

Para conseguir la fijación de células alrededor de todo el canal, es necesario una rotación o movimiento del dispositivo. Para ello, se llevó a cabo la caracterización de un motor, ya que permitiría conocer la relación entre los voltios con los que se alimentaría el motor y las revoluciones por minuto del eje de rotación. Se varió el voltaje con el que se alimentaba el motor, empezando con 0,5 V y llegando hasta 10 V, incrementando en cada paso 0,5 V la alimentación y midiendo el valor de las revoluciones por minuto del eje.

Paralelamente al desarrollo del sistema rotatorio y debido a sus dimensiones, surgió la necesidad de fabricar una incubadora para dar cabida tanto al motor como al dispositivo microfluídico. Para ello, con una nevera portátil y mediante un sensor de temperatura y otro de humedad, se desarrolló una incubadora celular. El objetivo del control de temperatura era alcanzar los 37 °C constantes dentro de la incubadora y, para ello, se dispuso de tres alfombrillas calefactoras rectangulares de 12 V-5 W, de un termostato digital y de una PT-100 conectada a Arduino. El procedimiento consistió en aumentar el voltaje que se aplicaba a las alfombrillas calefactoras, desde 1 V hasta 12 V, con un paso de 1 V y se registraba la temperatura medida tanto por la PT-100 como por el termostato digital. Por último, con la misma configuración que en el caso anterior, se realizó un calentamiento constante a 12 V en las 3 alfombrillas y se calculó el tiempo que era necesario para alcanzar los 37 °C. En cuanto al control de humedad, se emplearon 2 placas de Petri con agua en la base de la incubadora, un sensor de humedad DHT11 y un Arduino. Se realizaron dos pruebas: la primera consistió en analizar la variación de la humedad al variar los voltios con los que se alimentaba a las alfombrillas calefactoras y en la segunda, se calentó la incubadora hasta 37°C constantes y se midió la evolución de la humedad a esta temperatura durante 150 minutos.

2.3. Cultivo celular en dispositivos con canales circulares

Para realizar el cultivo celular, y siguiendo las fuentes consultadas, se emplearon 3 métodos: rotación manual, cultivo orbital y rotación mediante motor. La línea celular empleada fue MCF-7, una línea celular de adenocarcinoma humano de mama. Debido a la falta de control del ambiente de gases dentro de la incubadora creada en este proyecto, se optó por utilizar un medio alternativo al DMEM para el cultivo de células. Para ello se utilizó el buffer HEPES (Life Technologies) como controlador del pH del medio. Las células introducidas en los canales procedían de cultivos en confluencia, tanto en medio HEPES como DMEM, asegurando de esta forma la inserción de una cantidad suficiente de células para recubrir el canal. Tras la tripsinización, centrifugado y recuento celular en la cámara Neubauer, las células fueron introducidas en los canales mediante conectores y jeringuillas. La cantidad de células introducidas en cada canal quedó registrada para facilitar el análisis de los resultados. Únicamente se emplearon los dispositivos de diámetro 0.20 mm y 0.29mm.

En el cultivo celular llevado a cabo mediante la técnica de rotación manual se realizaron 4 giros de 90º y el tiempo de incubación fue de 2 horas en cada posición. Tras completar una vuelta entera los dispositivos se cultivaron estáticamente. En el cultivo celular mediante orbital, se realizaron dos pruebas. En la primera el tiempo de incubación con orbital fue de 30 minutos y en la segunda de 2 horas. Las revoluciones por minuto del orbital fueron constantes en ambas pruebas, 6 rev/min. Además, se varió la concentración y el diámetro de los canales en las pruebas realizadas. Por último, en el cultivo mediante motor, se realizaron 3 pruebas en las que se varió la concentración, el diámetro de los canales y el tiempo de incubación rotacional.

3. Resultados y discusión

3.1. Fabricación y caracterización de los dispositivos microfluídicos

La fabricación de los dispositivos a partir del molde facilitó el proceso de fabricación y aseguró la reproducibilidad de los canales. Las técnicas de caracterización proporcionaron información sobre la funcionalidad y geometría de los canales. En primer lugar, el análisis mediante tintes permitió comprobar lo siguiente: 1) que ningún canal presentaba fugas ni obstrucciones del flujo; 2) que la introducción de los conectores y del líquido era muy complicada en aquellos dispositivos en los que uno de sus extremos estuvo en contacto con la pared del molde en el proceso de fabricación. Una vez comprobada la continuidad de los canales mediante tintes, se procedió a la caracterización de la superficie interna del canal mediante microscopía. Esta técnica permitió comprobar que todos los canales eran totalmente rectos y cilíndricos. A continuación, se cortaron los dispositivos por la mitad de modo que se obtuviesen dos piezas idénticas que facilitasen la caracterización de la superficie interna del canal y se comprobó que: 1) uno de los canales presentaba burbujas, debido a que no se hizo el proceso de desgasificado; 2) todos los canales presentaban rugosidades por la tensión a la que se expone el polímero en el desmolde. La siguiente caracterización consistió en el uso del microscopio SEM para estudiar las dimensiones reales de los canales. A los dispositivos se les practicó 2 cortes: uno, horizontal que permitía ver el orificio circular del canal y otro, que dividía el canal en 2 partes iguales para analizar su interior. Se pudo observar que las dimensiones de los canales eran muy similares al diámetro del hilo con el que se habían fabricado (Figura 2).

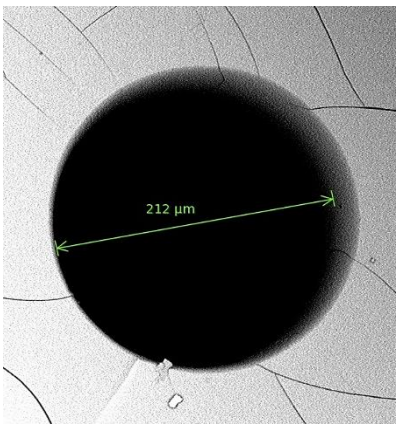


Figura 2. *Imagen obtenida mediante SEM de un dispositivo fabricado mediante un hilo de diámetro 0,2mm.*

Por último, se emplearon esferas fluorescentes para caracterizar la superficie interna de los canales sin seccionarlos. Los resultados de las imágenes de la pared del canal no fueron definitivos porque, al estar trabajando con un objeto de un grosor considerable, las imágenes tenían ruido procedente de las distintas capas. Para solventar este inconveniente, se empleó una de las opciones del microscopio consistente en establecer una profundidad fija y evitar que las imágenes se vieran afectadas por la luz proveniente de otros planos. Sin embargo, las imágenes tomadas mediante esta herramienta no proporcionaron información diferente a la obtenida anteriormente. De las imágenes obtenidas no se puedo deducir la geometría de los canales ni la disposición espacial de las paredes, por lo que este método no aportó ninguna información sobre los mismos.

3.2. Fabricación de un sistema rotatorio para la fijación celular en canales circulares

La relación entre los voltios con los que se alimentó el motor y las revoluciones por minuto de salida del eje rotatorio fue obtenida mediante la representación gráfica de los resultados obtenidos en la medición. De esta forma se obtuvo una ecuación que permitía controlar el giro del eje variando el voltaje.

En cuanto al control de temperatura, se consiguió alcanzar los 37 °C constantes dentro del incubador y el tiempo necesario para llegar a este valor fue de 46 minutos. Además, para alcanzar esta temperatura se comprobó que las alfombrillas calefactoras se debían de alimentar con 12 voltios constantes. Respecto al control de humedad, este no proporcionó resultados concluyentes. Los registros de humedad variaban constantemente sin un patrón determinado y con incongruencias. Por ejemplo, cuando se emplearon los calefactores y la temperatura del incubador pasó de los 24 °C a los 32 °C, la humedad descendió del 27% al 11%. Este resultado fue llamativo, pues, según se puede leer en la ficha técnica del sensor DHT11, el valor mínimo de humedad capaz de medir el sensor es del 20%. Dados los resultados y que el control de la humedad no se consideró prioritario, se decidió desistir de su implementación y únicamente se dispusieron las placas de Petri con agua al inicio de cada incubación con el motor.

3.3. Cultivo celular en dispositivos con canales circulares

Los resultados obtenidos en el cultivo celular en los canales circulares de los dispositivos microfluídicos se basan en las fotos obtenidas mediante un microscopio óptico. Son resultados cualitativos y, debido al uso de este tipo de microscopio, no se pudo localizar la posición exacta de las células, por tanto, no se determinó de forma precisa si los canales estaban completamente recubiertos de estas. Así pues, los resultados expuestos se basan en la presencia o ausencia de células en los canales tras una cantidad determinada de horas de cultivo. Los resultados obtenidos mediante la técnica de rotación manual muestran que, a mayor concentración de células insertadas, mayor probabilidad de que estas se mantengan vivas en los canales durante el tiempo (Figura 3).

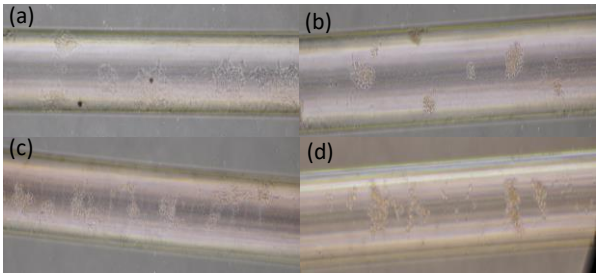


Figura 3. *Resultados del cultivo celular mediante rotación manual. (a) Medio DMEM – Concentración celular: 2.5×10^{6}células/ml – Diámetro del canal: 0,2mm – Tiempo de incubación: 24 h. (b) HEPES – 7,8×10^{5}células/ml– 0,2mm – 48 h. (c) DMEM - 2.5×10^{6}células/ml – 0,29mm – 96 h. (d) HEPES - 7,8×10^{5}células/ml– 0,29mm – 96 h.*

En cuanto a los resultados obtenidos mediante la técnica de cultivo con orbital, los resultados muestran una tendencia similar a la analizada en la prueba de cultivo mediante rotación manual. Además, se observó que las células se disponían de forma agrupada y en pequeños grupos a lo largo de los canales. Por último, los resultados obtenidos con el cultivo rotacional mediante motor mostraron lo mencionado en los resultados del método de rotación manual. A mayor concentración celular introducida en los dispositivos, estas se mantienen durante más tiempo vivas. Además, el uso de canales con mayor diámetro favorece la supervivencia celular y son más manejables para la inserción de las células (Figura 4).

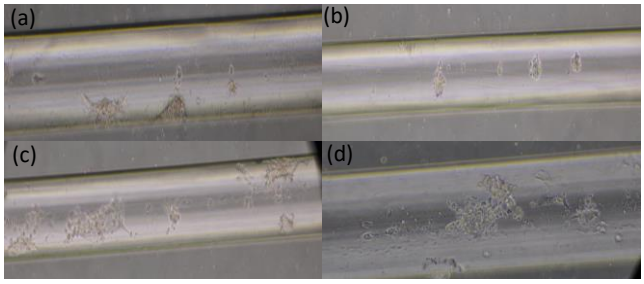


Figura 4. *Resultados del cultivo celular mediante motor. (a) Medio: DMEM – Concentración celular: 1,5×10^{6}células/ml – Diámetro del canal: 0,2mm – Horas de incubación: 72 h. (b) HEPES – 1,25×10^{6}células/ml – 0,2mm – 72 h. (c) HEPES – 1,25×10^{6}células/ml – 0,2mm – 96 h. (d) HEPES – 1,25×10^{6}células/ml – 0,29mm – 96 h.*

Tras analizar los resultados se concluye que: en primer lugar, no se ha podido determinar la posición exacta de las células en los

canales debido a que el microscopio empleado no era confocal. Por tanto, no se pudo garantizar que las células observadas estuvieran pegadas en las paredes de los canales o flotando en el medio. En segundo lugar, el método que mejores resultados proporcionó fue el cultivo mediante orbital. Los resultados no fueron satisfactorios en las primeras pruebas debido a que la presencia de células en tiempos avanzados era mínima. En las siguientes pruebas hubo presencia de células en todos los dispositivos y en todos los tiempos de observación. Este resultado está en concordancia con los obtenidos en una investigación previa en donde se consiguieron cultivar células empleando un orbital [9]. En tercer lugar, el uso del motor acoplado a la incubadora fabricada permitió cultivar células de forma rotatoria. Por último, la disposición de las células en los canales fue en grupos y no de forma extendida ocupando toda la pared del canal como se pretendía en esta investigación (Figura 5). Este resultado fue llamativo ya que lo esperado era que la disposición de las células en el canal fuese de manera que estas cubrieran toda la superficie interna del mismo. Este comportamiento puede derivarse de las características propias de la línea celular MCF-7. Aunque este último resultado no era el esperado a priori, una investigación llevada a cabo con la misma línea celular, dispositivos con canales circulares y condiciones similares, obtuvo estos mismos resultados [11]

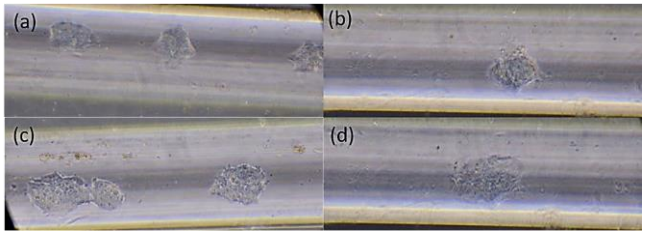


Figura 5. *Imágenes obtenidas tras la incubación mediante orbital. (a)2 h de incubación orbital – imagen obtenida 24 h después. (b) 30 min de incubación orbital – imagen obtenida 24 h después. (c) 2 h de incubación orbital – imagen obtenida 48 h después. (d) 30 min de incubación orbital – imagen obtenida 48 h después.*

4. Conclusiones

La fabricación de dispositivos microfluídicos con canales circulares mediante la técnica *micro-wire* permite generar dispositivos con canales rectos, circulares y sin daños, además de ser una técnica reproducible. La estabilidad dimensional de este método es muy alta ya que los diámetros medidos en el SEM eran iguales a los diámetros de los hilos empleados en la fabricación. El complejo motor-incubadora fue desarrollado correctamente y permitió cultivar células de manera rotatoria con buenos resultados. En cuanto al cultivo celular mediante las 3 técnicas empleadas, se puede concluir lo siguiente: 1) la concentración de células que hay que introducir en los dispositivos tiene que ser la mayor posible, entre $4{,}5\times10^6$ células/ml y 5×10^6 células/ml; 2) el uso de canales con mayor diámetro favorece el cultivo y adherencia de células en este tipo de canales; 3) las técnicas de cultivo que involucran al motor y el orbital favorecen la adhesión de las células; 4) las células de la línea MCF-7 se distribuyen en grupos de forma agrupada a lo largo de los canales.

Referencias

[1] Chen YC, Chen GY, Lin YC, Wang GJ. A lab-on-a-chip capillary network for red blood cell hydrodynamics. *Microfluidics and Nanofluidics*, vol. 9, sup 2, 2010, pp 585-591 (ISSN: 1613-4982).

[2] Sato K, Sato K. Recent Progress in the Development of Microfluidic Vascular Models. *Analytical Sciences*, vol 34, sup 7, 2018, pp 755–764 (ISSN: 0910-6340).

[3] Li G, Xu S. Small diameter microchannel of PDMS and complex three-dimensional microchannel network. *Materials and Design.*, vol. 81, 2015, pp 82–86 (ISSN: 0264-1275).

[4] Abdelgawad M, Wu C, Chien WY, Geddie WR, Jewett MAS, Sun Y. A fast and simple method to fabricate circular microchannels in polydimethylsiloxane (PDMS). *Lab on a Chip*, vol. 11, sup 3, 2011, pp 545–551 (ISSN: 1473-0189).

[5] Truskey G. Transport phenomena in biological systems. Pearson/Prentice Hall, 2004 (ISBN: 0130422045).

[6] Convery N, Gadegaard N. 30 years of microfluidics. *Micro and Nano Engineering*, vol. 2, 2019, pp 76–91 (ISSN: 2590-0072).

[7] Radtke AL, Herbst-Kralovetz MM. Culturing and Applications of Rotating Wall Vessel Bioreactor Derived 3D Epithelial Cell Models. *Journal of visualized experiments*, no 62, 2012 (ISSN: 1940-087X).

[8] Van Engeland NCA, Pollet AMAO, Den Toonder JMJ, Bouten CVC, Stassen OMJA, Sahlgren CM. A biomimetic microfluidic model to study signalling between endothelial and vascular smooth muscle cells under hemodynamic conditions. *Lab on a Chip*, vol 18, sup 11, 2018, pp 1607–1620 (ISSN: 1473-0189).

[9] Costa PF, Albers HJ, Linssen JEA, Middelkamp HHT, van der Hout L, Passier R, van den Berg A, Malda J, van der Meer AD. Mimicking arterial thrombosis in a 3D-printed microfluidic in vitro vascular model based on computed tomography angiography data. *Lab on a Chip*, vol 17, sup 16, 2017, pp 2785-2792 (ISSN: 1473-0189).

[10] Mitxelena-Iribarren O, Zabalo J, Arana S, Mujika M. Improved microfluidic platform for simultaneous multiple drug screening towards personalized treatment. *Biosensors and Bioelectronics*, vol 123, 2019, pp 237–243 (ISSN: 0956-5663).

[11] Morley ST, Walsh MT, Newport DT. The advection of microparticles, MCF-7 and MDA-MB-231 breast cancer cells in response to very low Reynolds numbers. *Biomicrofluidics*, vol 11, sup 3, 2017 (ISSN: 1932-1058).

Ingeniería de Rehabilitación

Miércoles, 25 de noviembre de 2020

Virtualización del test clínico Jebsen-Taylor Hand Function basado en Leap Motion Controller

V. Lozano Berrio[1], M. Álvarez Rodríguez[1], S. Ceruelo Abajo[2], E. López Dolado[2], F. Talavera Díaz[2], A. Gil Agudo[2], A. de los Reyes Guzmán[1]

[1] Unidad de Biomecánica y Ayudas Técnicas, Hospital Nacional de Parapléjicos (SESCAM), Toledo, España, {vlozanob,adlos}@sescam.jccm.es; maria.alvarez.rodriguez93@gmail.com

[2] Servicio de Rehabilitación, Hospital Nacional de Parapléjicos (SESCAM), Toledo, España, {sceruelo,elopez,ftalavera,amgila}@sescam.jccm.es

Resumen

El test clínico Jebsen Taylor Hand Function es una escala validada y estandarizada para su uso en el entorno clínico que permite la valoración de las capacidades funcionales del paciente. Consiste en 7 tareas, diseñadas como representativas de varias actividades funcionales centradas en la mano y destinada a la población general. Valora en cada tarea el tiempo de ejecución de la misma con la mano no dominante y dominante. No obstante, se ha aplicado para la valoración de las capacidades funcionales en pacientes con patologías neurológicas. El objetivo de este trabajo es presentar la aplicación virtual relativa al subtest de la escritura del test clínico Jebsen Taylor Hand Function, para su manipulación a través de Leap Motion Controller. Se analiza el desempeño en el test virtual en personas sanas estudiando la relación existente entre el tiempo de ejecución de la tarea y la precisión obtenida en la grafía de los caracteres medida en porcentaje de contacto. Se ha obtenido una correlación elevada entre ambas variables en personas neurológicamente sanas. Este análisis se está ampliando a una muestra de personas con afectación de la función del miembro superior tras enfermedad neurológica, más concretamente, lesión medular. En un trabajo posterior se analizará también la correlación entre el tiempo de ejecución del test virtual con el desempeño obtenido en el test real.

1. Introducción

La afectación en la destreza manipulativa y pérdida de coordinación motora son algunas de las consecuencias derivadas de trastornos y enfermedades neurológicas, condicionando en gran medida la autonomía y nivel de independencia y participación social de las personas que las padecen. Así, la calidad del movimiento y desempeño de actividades de la vida diaria está muy influenciada por la función de la mano y la destreza manipulativa [1].

En este contexto, la capacidad funcional de la mano suele valorarse mediante la administración de test clínicos. Uno de ellos es el Jebsen Taylor Hand Function (JTHF) [2]. Se trata de un test estandarizado para evaluar la función completa de la mano. Así, la mano dominante y no-dominante se evalúa mediante 7 subtests que simulan actividades de la vida diaria.

Se tiene evidencia de un estudio que ha aplicado el JTHF como uno de los métodos para evaluar la capacidad funcional de la mano en pacientes con patologías neurológicas como el Accidente Cerebrovascular (ACV). Este estudio detecta una estructura de factores detrás de las deficiencias motoras finas en pacientes con ACV que quedan reflejadas en gran medida por los resultados del test clínico JTHF [3]. También ha sido aplicado en pacientes con enfermedad de Parkinson [4]. Otro estudio ha validado la aplicación del JTHF para predecir el resultado en actividades de la vida diaria (AVD) [5].

Actualmente, hay un interés creciente en el uso de tecnología en el campo de la Neurorrehabilitación con el objetivo de cuantificar y generar conocimiento acerca de los trastornos sensorimotores tras enfermedades neurológicas, entendiendo que la tecnología posee un elevado potencial para su uso como herramientas terapéuticas [6]. Así, en la actualidad, son varios los trabajos que proponen versiones virtuales de test clínicos como, por ejemplo, el Box and Block en principio pensada para pacientes con daño cerebral y que utiliza tecnología de bajo coste para permitir la interacción del paciente con la aplicación, como Microsoft Kinect [7] y Leap Motion Controller (LMC) [8]. Sin embargo, hasta la fecha no se tiene evidencia de estudios que propongan una virtualización del test JTHF. Solamente se tiene evidencia del desarrollo de un sistema para el reconocimiento preciso y rápido de caracteres escritos a mano en el aire utilizando LMC. Para ello, el sistema previamente ha de hacer un registro de los caracteres como patrón, para su posterior reconocimiento [9].

Por tanto, el objetivo de este trabajo es presentar la aplicación virtual relativa al subtest de escritura del test clínico JTHF. Se analiza en personas sanas la relación existente entre el tiempo de desempeño de la tarea y el porcentaje real de contacto con la frase cuyos caracteres hay que trazar. La aplicación virtual presentada forma parte del prototipo RehabHand basado en LMC, consistente en 7 aplicaciones virtuales dotadas de un sentido terapéutico para rehabilitación de los miembros superiores. RehabHand ha sido desarrollado íntegramente en el Hospital Nacional de Parapléjicos mediante la colaboración de un equipo de trabajo interdisciplinar.

ISBN: 978-84-09-25491-0, pags. 190- 193

2. Test Jebsen Taylor virtual

2.1. Descripción e instrucciones del JTHF

El JTHF es una escala específica para valorar los patrones de funcionalidad de la mano centrada en los patrones estáticos y dinámicos, ya que tiene en cuenta la participación en determinadas actividades de la vida diaria, permitiendo conocer las capacidades y limitaciones reales en la función de la persona que lo ejecuta.

Pretende evaluar funcionalidad, habilidad y velocidad, valorando el tiempo de ejecución de los siguientes subtests:

- Escritura
- Pasar tarjetas de 3x5 pulgadas (simulando pasar páginas)
- Coger determinados objetos pequeños comunes
- Mover fichas apilando 4 damas en forma de torre
- Simular la alimentación usando una cuchara para coger una a una cinco alubias
- Alcanzar, coger y transportar de forma unilateral cinco objetos grandes ligeros
- Alcanzar, coger y transportar de forma unilateral cinco objetos grandes pesados

Cada uno de ellos se ejecuta sin centrarse en la calidad del movimiento, mediante un proceso estandarizado y evaluando de forma objetiva y unilateral cada mano, comenzando por la no dominante. En este contexto, es necesario estandarizar y homogeneizar las condiciones ambientales de la prueba, los materiales a utilizar, las instrucciones que se deben seguir y mantener el mismo orden de ejecución de los subtests, así como los posibles límites de tiempo para garantizar la repetibilidad de las pruebas.

Son varios los objetivos que perseguían los autores con el diseño de esta escala [2]:

- Obtener medidas objetivas a partir de tareas funcionales con patrones de normalidad con los que comparar los resultados del paciente.
- Evaluar aspectos funcionales de la mano frecuentemente usados en AVD
- Ser capaz de documentar aspectos de habilidad dentro de cada categoría de la función de la mano.
- Ser administrada fácilmente en un corto período de tiempo.
- Utilizar materiales que están fácilmente disponibles y accesibles.

2.2. Desarrollo de la tarea de escritura virtual

En muchos casos, el resultado de los test clínicos está condicionado a la persona que administra y puntúa el test. Por ello, surgen intentos de diseñar test clínicos que de forma virtual y con el uso de tecnología aporten la objetividad necesaria a las valoraciones. En este caso la tecnología utilizada es el controlador de movimiento Leap Motion Controller (LMC). Se trata de un dispositivo diseñado para controlar aplicaciones mediante gestos de la mano. Consta de tres sensores emisores de luz infrarroja y dos cámaras. Es un dispositivo pequeño y rectangular (13 mm x 13 mm x 76 mm) que pesa 45 g. Se trata de un sistema de captura óptico que, a través de sus algoritmos internos, ofrece la posición en el espacio cartesiano de objetos predefinidos como las manos y los dedos. Las posiciones medidas son relativas respecto al origen de LMC, situado en la posición del segundo sensor infrarrojo. La precisión de LMC es de 0,2 mm en condición estática [10]. El dispositivo se conecta al ordenador vía USB y es compatible con Mac y Windows. En este trabajo se utilizó un ordenador con procesador i7, 12 GB de RAM y tarjeta gráfica dedicada, trabajando bajo Windows 10.

Algunos dispositivos, como Microsoft Kinect, pueden capturar amplios rangos de movimiento del cuerpo. Sin embargo, LMC captura el movimiento de las manos y, por tanto, es más viable para trabajar movimientos finos del miembro superior y la función de la mano.

En el contexto de este trabajo, LMC se ha utilizado para simular la tarea de la escritura. Se ha utilizado la SDK versión 2.3.1., puesto que aunque hay otras versiones posteriores, ésta distingue el objeto *tool* como un elemento independiente de la mano y lo caracteriza como el elemento asemejando al lápiz o bolígrafo. Esta funcionalidad no está permitida en versiones superiores de la SDK de LMC.

La aplicación virtual que simula la tarea de escritura se ha desarrollado en Unity3D. Previamente se han modelado todas las letras y caracteres más importantes de la fuente Arial. Siguiendo la terminología y métodos de Unity3D, cada una de esas letras está definida por un *prefab*, cuyos componentes son *Transform*, el script *Letter*, y un conjunto de *colliders*. El script se encarga de determinar el porcentaje de contacto del puntero con cada uno de los puntos que definen la letra.

Cuando se inicia la aplicación desarrollada, ésta ofrece de forma aleatoria una de las cuatro frases disponibles en el test JTHF real para trabajar la función de escritura:

- Las ballenas nadan en ese mar.
- Ese anciano parecía cansado.
- John vio venir un camión rojo
- El pez aspira dentro del agua.

Todas ellas tienen 24 caracteres y son iguales en contenido semántico. Se trata de frases de dificultad lectora media. En la Figura 1 se puede ver parte de la aplicación virtual.



Figura 1.*Representación parcial de una de las frases correspondientes a la tarea de escritura del JTHF virtual*

El objetivo es ir repasando la caligrafía punteada, de forma que cuanto más contacto haya con la frase tipografiada mayor será el resultado obtenido en el test.

El marcado de la trayectoria se puede hacer tanto con el dedo índice como con un objeto que simule un lápiz, reconocido por LMC como un elemento *tool* independiente de la mano, proyectando en un plano las coordenadas XY, de forma que la trayectoria sea bidimensional.

La cámara puede desplazarse por el texto siguiendo las teclas de dirección. Los elementos del LMC están asociados a la cámara en la jerarquía de la escena, de forma que se desplazan con la misma. Las proyecciones se representan en cada momento con la posición de un puntero en forma de estrella y se irán dibujando trayectorias como objetos individuales cada vez que se deje de pintar.

Todos los sujetos realizaron la prueba en sedestación frente a una mesa regulable en altura. La altura de la mesa se adaptó en cada caso de forma que el codo, apoyado sobre la mesa, estuviera flexionado 90º. El sensor se ubicó entre el sujeto y el ordenador, a una distancia que permitiese que la mano quedase sobre el sensor con el antebrazo apoyado sobre el borde de la mesa y evitando la realización de movimientos de compensación de tronco. Se tarea se realizó con la mano dominante (Figura 2).



Figura 2.*Imagen del set-up experimental mientras un participante del estudio realiza la tarea*

El test virtual requiere aclimatación al entorno, por lo que este test se realizó previamente una vez a modo de aclimatación.

2.3. Variables analizadas

El análisis consistió en la ejecución del test virtual JTHF en la tarea de escritura por un grupo de 15 adultos sanos (12 mujeres y 3 hombres), con una edad de 30,27± 8,28 años. Se obtuvo el tiempo total de desempeño de la tarea y la precisión en la misma a través del porcentaje de contacto con la grafía punteada. Se analizó mediante un modelo de regresión lineal (índice de correlación de Pearson) la relación existente entre las dos variables propuestas.

3. Resultados

En la comparación de la tendencia entre el tiempo total de desempeño de la tarea y la precisión en la misma a través del porcentaje de contacto con la grafía punteada, se ha obtenido un coeficiente de correlación elevado y positivo ($r=0.857$, $p<0.01$), indicando que un mayor tiempo de ejecución de la tarea se corresponde también con desempeño mejor en tarea virtual ofreciendo un porcentaje mayor de contacto con la grafía punteada disponible cuando se lanza la aplicación (Figura 3).

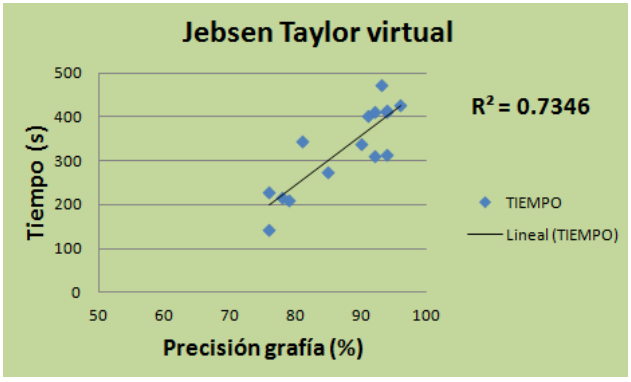


Figura 3.*Diagrama de dispersión entre el tiempo total de desempeño de la tarea y la precisión en la misma a través del porcentaje de contacto con la grafía punteada.*

4. Discusión

En este trabajo se ha presentado una aplicación virtual del test JTHF para la tarea funcional de la escritura, pudiéndose analizar la relación existente entre el tiempo total de desempeño de la tarea y la precisión en la misma a través del porcentaje de contacto con la grafía punteada.

Los resultados obtenidos sugieren que se mantiene la tendencia entre las dos variables analizadas de forma que un menor tiempo de ejecución del test virtual corresponde con una menor precisión en cuanto al porcentaje de contacto con la grafía punteada, que correspondería a la ejecución ideal. Este resultado refleja que el desempeño de una tarea funcional a nivel distal es un compromiso entre la precisión en la tarea y la velocidad con que ésta se realiza. Sin embargo, estos resultados han sido los obtenidos en una muestra de 15 personas sanas adultas. En este contexto, es necesario ampliar este estudio a una muestra de pacientes con patología neurológica, en los que hay evidencia de que el desempeño motor dentro de una tarea funcional es menor para un mayor tiempo de ejecución de la misma. Así, resulta de interés estudiar si

se conservaría esta tendencia o por el contrario se produciría una inversión de la tendencia para las dos variables analizadas. De la misma forma, igual que en estudios previos, es necesario analizar la relación entre los resultados en el test virtual con los obtenidos en el test real.

No se tiene evidencia de estudios previos que hayan propuesto una virtualización del JTHF para la tarea de escritura. Como consecuencia, los resultados obtenidos no se han podido contrastar con los de otros autores. Solamente se tiene evidencia de un trabajo en el que se propone un sistema para el reconocimiento de caracteres basado en LMC. Para su correcto funcionamiento, previamente es necesario hacer un registro de patrones. La tasa de aciertos que han obtenido en el reconocimiento de caracteres ha sido del 86.7%, sugiriendo ampliar el estudio y continuar trabajando en la mejora del sistema [9].

La aplicación objetivo de este trabajo incorpora fuentes de *feedback* visual: un contador que informa sobre el tiempo de ejecución y la frase completa a escribir en la parte inferior de la pantalla. De este modo, la persona que realiza el test es consciente de los caracteres restantes para completar la tarea. La tarea propuesta se puede realizar con el dedo índice o bien con un objeto adicional simulando al bolígrafo o lápiz. Para hacerlo posible, la aplicación se ha desarrollado usando la SDK del LMC (versión 2.3.1) que permite registrar y diferenciar el objeto *tool* como un elemento diferente e independiente de la mano. Obviamente, en una aplicación de esta metodología en pacientes con afectación de la función de los miembros superiores esto no es tan trivial. El paciente normalmente no posee la capacidad ni fuerza suficiente para sostener el lápiz en la mano. En este caso, se cuenta con una adaptación obtenida mediante impresión 3D que permitiría la ejecución de esta aplicación por parte de los pacientes.

Además, resultaría muy útil combinar esta aplicación con otras que componen el prototipo RehabHand y que entre sus objetivos, por ejemplo se valore la suavidad de la trayectoria realizada durante la ejecución de la tarea. Esto permitiría la utilización de este prototipo para realizar una valoración objetiva de la función del miembro superior, teniendo en cuenta que la falta de suavidad durante la ejecución de un movimiento es una de las principales características en presencia de afectación neurológica.

No obstante, el presente trabajo también presenta limitaciones. Algunas de ellas son inherentes a la tecnología utilizada, el LMC, cuyo rendimiento está fuertemente condicionado por las condiciones del entorno y por las prestaciones del ordenador en el que se ejecuta. Las demás limitaciones se deben al desarrollo del entorno virtual en sí. La aplicación desarrollada posee una elevada complejidad en cuanto a desarrollo software. Se han modelado todos los caracteres del alfabeto y los principales signos de puntuación. Lo ideal sería poder cambiar el tipo de fuente, pero este hecho añadiría a la aplicación una complejidad de desarrollo muy importante. Por eso se optó por modelar la fuente tipo Arial.

5. Conclusión

En este trabajo se ha mostrado que la versión virtual del JTHF desarrollada para trabajar la función de la escritura es viable aplicada a población adulta sana. Este mismo estudio ha de ser ampliado a una muestra de pacientes con patología neurológica, recogiendo también los resultados del test clínico real para esta tarea.

Agradecimientos

Este trabajo forma parte del Proyecto “Plataforma de bajo coste para rehabilitación del miembro superior basado en Realidad Virtual” con referencia DPI2016-77167-R financiado por MINECO a través del Programa Estatal I+D+i orientado a los Retos de la Sociedad 2016 y cofinanciado con fondos FEDER.

Referencias

[1] Oliveira CS, Almeida CS, Freias LC, Santana R, Fernandes G, Junior PRF, Moura RCF. Use of the Box and Block Test for the evaluation of manual dexterity in individuals with central nervous system disorders: a systematic review. *Manual Therapy, Posturology & Rehabilitation Journal Revista Manual Therapy*, *14*, (2016).

[2] Jebsen RH. An objective and standardized test of hand function. *Arch. Phys. Med. Rehabil.*, *50*(6) (1969), 311-319.

[3] Allgöwer K, Hermsdörfer J. Fine motor skills predict performance in the Jebsen Taylor Hand Function Test after stroke." *Clinical Neurophysiology* 128.10 (2017): 1858-1871.

[4] Mak MKY, Lau ETL, Tam VWK, Woo CWY, & Yuen S KY. Use of Jebsen Taylor Hand Function Test in evaluating the hand dexterity in people with Parkinson's disease. *Journal of hand therapy*, *28*(4), 2015, 389-395.

[5] Lynch KB, Bridle MJ. Validity of the Jebsen-Taylor Hand Function Test in predicting activities of daily living. (1989): 316-318.

[6] Shirota C, Balasubramanian S, Melendez-Calderon A. Technology-aided assessments of sensorimotor function: current use, barriers and future directions in the view of different stakeholders. *Journal of neuroengineering and rehabilitation*, *16*(1), 2019, 53.

[7] Hsiao CP, Zhao C, Do EYL. The Digital Box and Block Test Automating traditional post-stroke rehabilitation assessment. In *2013 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)* (2013, March) (pp. 360-363). IEEE.

[8] Dall’Alba D, Dimbwadyo I, Piazza S, Magnabosco E, Menegozzo G, Fiorini P. ViTAS gaming suite: virtual therapy against stroke. In *Converging Clinical and Engineering Research on Neurorehabilitation II*, 2017 (pp. 407-411). Springer, Cham.

[9] Tsuchida K., Miyao H, Maruyama M. Handwritten character recognition in the air by using leap motion controller. In *International Conference on Human-Computer Interaction* (pp. 534-538). (2015, August). Springer, Cham.

[10] Weichert F, Bachmann D, Rudak B, Fisseler D. Analysis of the accuracy and robustness of the leap motion controller. *Sensors*, *13*(5), 2013, 6380-6393.

Development of a motion measurement system of a long cane for Visually Impaired People in rehabilitation.

K. M. Reyes Leiva[1,2], S. Sanchez Lara[1,3], M. Jaén-Vargas[1],

J. J. Serrano Olmedo[1,3,4]

1. Bioinstrumentation and Nanomedicine laboratory (LBN), Center for Biomedical Technology (CTB), Universidad Politécnica de Madrid, Madrid, Spain.

2. Engineering Faculty, Universidad Tecnológica Centroamericana UNITEC, San Pedro Sula, Honduras

3. Escuela Técnica Superior de Ingenieros de Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain.

4. Networking Center of Biomedical Research for Bioengineering, Biomaterials and Nanomedicine, Instituto de Salud Carlos III, Madrid, Spain.

Karla.reyes@ctb.upm.es, Sergio.sanchez.lara@alumnos.upm.es, milagros.jaen@ctb.upm.es, josejavier.serrano@ctb.upm.es

Resume

The rehabilitation of a blind person is a systematic process where the person is dotted with tools that allow them to deal with the impairment, including the correct use of the long cane for navigation in the Orientation and Mobility training. In this process, technology as inertial measurement sensors could be included to track the motion of the cane in order to evaluate the correct use. One sensor added to the cane can provide valuable information about orientation, position, velocity and trajectory of the tip. In this research, a system was developed by adding a BNO055 inertial sensor to the lower part of the long cane in order to analyze the output with the help of Arduino, Simulink and MATLAB. A method was developed to obtain qualitative information about the movements of the long cane during the training of mobility and the reliability was tested. The tests were carried out with a real cane, in isolated situations and a controlled environment, corresponding to real techniques that will be trained during rehabilitation. The experimental results indicate that the parameters obtained in the orientation test and in the short displacement test presented a standard error ~1% in both cases and an accuracy of 98 and 97% respectively. However, for the sliding test for simulation of the diagonal technique, the difference in the value of the mean obtained with respect to the real value was 39%. Despite presenting reproducibility in tests with low standard errors, to obtain precise displacement or position value, a more complex census is required, or a new method should be developed for evaluation. The system requires further development in order to be used in the assistant program for the rehabilitation process.

Introduction

According to the OMS, 28% of the worldwide population suffers from a visual impairment [1]. In Spain, in 2020 there are 71,331 affiliated people to ONCE, where 14,15% of this suffers total blindness and in high percentages, other affiliates are registered with different pathologies that can lead to a total loss of vision, such as myopia, macular degeneration, hereditary retinal dystrophies and optic nerve pathologies [2]. When a person is born with a visual disability or suffers from a trauma or illness that leads to a total loss of vision, they must be assisted through a Rehabilitation process. According to the last report of the study of Rehabilitation of the blind [3], the services of rehabilitation that are provided with the highest incidence throughout the worldwide rehabilitation programs are the Training of Orientation and Mobility (O&M) and the Activities of daily living / Self-care or life skills (ADL) training.

The long cane is a mobility device, its carried in order to bring information about the visually impaired, it is used for safe displacement during navigation and mobility [4]. A fundamental part of the mobility training is the use of the long cane, the user should learn how to hold it correctly, how to grip, how to walk and sweep in order to detect obstacles, different techniques of exploration, and other parameters according to the complexity of the environment in which the user will be navigating [5]. This training is usually done in person with an O&M specialist, which leads to an accessibility problem in rural communities, also it compromises the rehabilitation duration, as well as the number of users that can be rehabilitated at the same time. In this training, depending on the scenario there is a recommended technique and according to the complexity and advances of the training, the scenarios will change [6]. However, the parameters for evaluation of the correct use, regardless of the change of scenario, will keep the same; these allow the possibility to register parameters and quantitative values of the movement [7], in order to support the O&M training in the rehabilitation processes.

On the other hand, the inertial measurement unit sensors have been widely used for navigation, robotics, body motion, processing and control, autonomous vehicles, and others [8]–[11] These sensors can measure the acceleration (accelerometer), angular velocity (gyroscope) and magnetic fields (magnetometer), together they can bring information about the orientation (Euler angles and Quaternions) and trajectory of an object (position and velocity) [12] and, there is a large and ever-growing number of applications for this sensors.

According to the literature, a diversity of technological proposals have been designed for orientation and mobility, such as ETAS (Electronic Travel Aid Systems) [13], focused on obtaining information from the environment and providing it to the visually impaired in order to assist them in autonomous navigation. There have been many attempts to dotted the long cane with technology, these systems are developed from technologies such as Global Positioning System, BLE Beacons, RFID or radio frequency identification, to obtain information on position and displacement and optical sensors (RGB-D cameras, Laser), inertial sensors, speed sensors among others for obtaining information regarding object detection [14]–[18] . However, the use of any of these ETAS requires previous O&M training [19], leading to an existing gap, which is the development of assistive technologies specifically focused on evaluation and assistance of the training process, so it can be more accessible for users.

ISBN: 978-84-09-25491-0, pags. 194- 197

The aim of this research was to develop a simple-architecture hardware system of data acquisition, and test its reliability in the quantitative analysis of the parameters evaluated in the rehabilitation process of blind people by obtaining metrics that the O&M trainers personally examine, such as orientation and grip rotation, trajectory and correct displacement of the long cane movement. For this objective, three tests were designed as described below.

Methodology

A BNO055 IMU Bosch sensor [20] was placed in the lower part of a 117 cm long cane to estimate the measure of the movement of the tip as shown in Fig. 1. Serial communication was used to transfer the sensor data to a microprocessor Arduino MKR1010 at a sample rate of 0.01 s, the Arduino served as an interface within Simulink, and was processed through a developed program in MATLAB. Also, in order to remove noise components from the signal, low pass filtering was performed, with a cutoff frequency of 20 Hz.

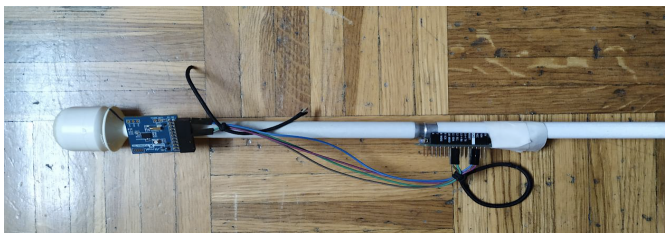


Fig 1. Long cane (117 cm) with the incorporation of the IMU sensor and Arduino MKR1010 for testing.

A methodology was developed to evaluate the orientation and displacement in controlled environments. To each of the three exercises belongs a different scenario, specifically developed to control the evaluated paraments and minimize human error. The three tests were created to evaluate the situation corresponding to real exercises of the O&M process of the blind, based on the O&M evaluation theory [21]. For each test, an objective measure was settled in order to analyze the validation of the developed model during the sampling time. Statistics parameters as media, standard deviation, and standard error were obtained after the testing.

Results and Discussion

3.1 Orientation, Velocity and Position

The sensor can provide orientation-related data in different formats so that it is only necessary to adapt the data to our reference system. The different formats in which orientation is presented are useful depending on their application. With them, we can rotate acceleration matrices, represent orientation variations in relative reference systems, or absolute with the help of the magnetometer [11].

With respect to obtaining velocity and position, several challenges were faced due to the use of discrete time data, cumulative errors in time due to integration, and accelerometer readings that do not correspond to the desired ones, due to the influence of the sensor orientation. First, it must be considered that when integrating information related to acceleration, with the intention of obtaining velocity, and then acceleration, we find an error known as time drift. The temporal drift can be treated by filtering the low-frequency components, which is possible thanks to the fact that the movement to be studied is a periodic and oscillating movement, with high-frequency components, compared to those of the temporal drift [11]. In addition, we rotated the acceleration reading matrix, depending on the sensor orientation at the sampling moments, in this way we keep three spatial axes (X, Y, Z) constant in time undoing the changes produced in the readings due to the variation in the sensor orientation.

In this way, we can obtain relatively reliable data on the velocity and position of the sensor, even though it requires a finer adjustment to obtain valid information. In addition, the use of data fusion from different sources is recommended, to reduce the error margin in obtaining the position.

3.2 Displacement of the long cane

The capacity to measure the displacement of the long cane was tested. As described by Blasch & LaGrow [21], the performance of the training can be evaluated in terms of "coverage" provided by the long cane, a full coverage includes, for instance, object preview: the capacity to identify objects in the path of travel. This path can be established by the measure from shoulder to shoulder of a user (~50 cm). since is assumed that the user will walk within a straight line, however side view should be added in the amplitude of the sweeping technique, because it will afford additional information of the travel path (~100 cm).

For this, the oscillating movements in one axis of the sensor were evaluated in order to validate measure the displacement needed for object preview in the Diagonal Technique (DT), this technique as described by the Rehabilitation Manual of ONCE, is the first exploration technique to be taught in rehabilitation [6]. A signaling system was created on the floor with a length of one meter and the help of a rail, with the intention of containing the movement in a single axis, avoiding human error. The movement carried out, consisted of a sweeping oscillatory movement within the rail so that the readings obtained were sinusoidal signals of one-meter peak-peak amplitude. The data for this test was extracted from graphical bidimensional representations of x and y position using MATLAB, and the displacement (D) measures were obtained with the Pythagoras equation $D = \sqrt{X^2 + Y^2}$. The displacement procedure of the cane in the rail was repeated 10 times in each test. So, the measured values were a statistical reference of the oscillation period, this test was repeated ten times by the same sighted volunteer and the statistical values are presented in Table 1. When measuring the motion of one axis, there are relative errors to the displacement axis and, a perpendicular axis to this one, which is denominated longitudinal and transversal, respectively.

The standard error of the measurements (0.02 and 0.01) is low since it represents approximately 4% for the media of the longitudinal displacement measure and 7% for the transversal, this result suggests that the data processing with the developed system is trustworthy for short displacements. When estimating the real path, the main measure has a difference of 0.39 m. This difference in displacement is mainly attributed to the processing of the data when

obtaining changes in position due to the drift error, as mentioned before, and could improve with a sensor of higher precision, with sensor fusion, with more references from the environment, or with a different method of measurement. Some authors reported displacement errors in navigation systems using smart canes, these errors are of about 0.2 m when using a VN-100 IMU / AHRS VectorNav inertial sensor fused with an RGB-D camera [17], or when fusing inertial sensors with RGB-D cameras and line laser [16]. Although these reported systems present evaluation conditions of displacements schemes greater than in this displacement test, they present us a notion of the difference of estimated error when there is sensor fusion in more complex systems, however, in order to evaluate that the user is performing a correct object preview in this technique, the accuracy of the measurement system should increase.

Table 1. Statistical values obtained in the tests.

3.2 Displacement test	**Longitudinal**	**Transversal**
Real value	1.0 m	0.01 m
Average Measured	0.61 m	0.09 m
Mean Standard deviation	0.07 m	0.02 m
3.3 Linear Slip Test	**Values**	
Real Value	50 cm	
Average Measured	48.54 cm	
Mean Standard deviation	1.80 cm	
3.4 Orientation test	**Values**	
Real Value	180°	
Average Measured	176.33°	
Mean Standard deviation	5.29 °	

3.3 Long cane linear slip

This test was developed as a validation of a measurement method focused on the modified diagonal technique (MDT). This technique is used to obtain information about the dimension of an unknown object, the cane must be placed vertically and a displacement in one or more axis must be done in order to identify the dimensions [6]. Due to the low accuracy of the measurement compared to the real value presented in the last test, a new method of measurement was proposed. For this test, the orientation angles were used, taking into count the inclination with respect to the vertical and longitudinal, obtaining the final distance D with oscillatory movements, following the formula according to (Fig 2.B). The Y-axis shows the inclination of the cane with respect to the vertical and the values of β are the variation in the degrees from peak to peak in the Y-axis. This data could be obtained by the conversion of quaternion to Euler angles, in order to obtain a system of Absolut orientation (Fig. 2.C). A box of 50x50 cm was created to measure the dimensions with the developed program (Fig. 2.A).

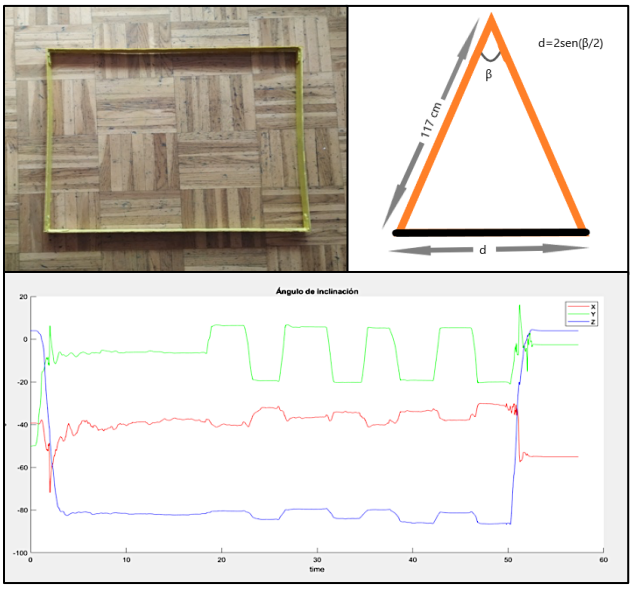


Fig 2. A – 50 cm square box representation for testing. B – Representation of the long cane in an oscillatory movement. C- Orientation of the sensor in time function.

The media of the diameter obtained was 48.54 cm as shown in Table 1. The standard deviation was 1.80 cm with a mean standard error of 0.57 cm. In this case, for a smaller displacement (50 cm instead of 1 cm), the difference from the measured distance to the real distance is 1.47 cm and the standard error represents only 1% of the obtained measure, which suggest that there is more precision in the reproducibility of smaller displacements and smaller difference of the values obtained to the real measure when using absolute orientation angles to measure a displacement.

This method should be used to obtain a reference of the long cane coverage evaluation in any technique and is recommended to be used in other parameters such as “hand height”, “safety zone” verification, and “surface preview”.

3.4 Long cane orientation

The orientation of the long cane is an important factor to considerer in order to identify traits when training the use of the long cane, since a bad grip when the wrist rotates can be early identified and corrected. According to the O&M specialist [22][23], the hand should produce a translational movement within the long cane and should not rotate with the movements of the wrist.

For the orientation test, 180-degree turns were made on the vertical axis, which indicates where the flat part of the handle is pointing. The cane was rotated within two points of reference. In this test, the measures were taken from the Euler angles output within the roll angle. A media of 176.33° was obtained, within a standard deviation of 5.29° with a standard error of 1.67° which represents the 0.94% of the measured angle and the difference of the mean measured to the real value, represents only 3.67% of the real orientation value 180°, as shown in Table 1.

According to the literature [7], this obtained results indicate that the system is very effective for the orientation measurement and within a low standard deviation. This suggests that this system could allow us

to clearly analyze parameters about the correct grip of the long cane and identify with reliability if there are wrist rotations during the training of the techniques.

In order to determine if the system is capable to verify the full coverage of the long cane while the user is in training, more tests should be developed. Besides object preview, surface preview and foot placement preview should also be accomplished [23][21]. Future work will be done regarding the evaluation of both surface and foot placement preview, the future work is focused on evaluating the capability of the system to measure the height of the hand and the safety zone, also, by adding one more inertial sensor to the body, the system will be capable to measure the stride, the posture, and the body rotation.

Conclusions

A quantitative study was done in order to evaluate representative parameters of O&M training in the rehabilitation process of the blind in experimental tests while using the long cane. The results obtained indicate that the developed system allows obtaining information on the orientation, velocity, position, and displacement of the movement of the long cane that contains an inertial sensor in its lower tip and a microprocessor in its mount. The carried-out tests, determinate the quality of the information provided by the system, indicating that to obtaining an exact displacement measurement with one inertial sensor is a complex task is we use traditional dead reckonin. This, because the raw data requires a large post-processing load to obtain accurate readings. This displacement must be adapted to the conditions of the movement to be studied. As shown in the experimental tests, the parameters obtained in the orientation test and in the short displacement test presented a standard error of around 1% in both cases and an accuracy of 98 and 97% respectively. These results suggest a reasonable estimate of the distance traveled over short distances and in the grip orientation measurements using absolute orientation angles. However, for the sliding test for simulation of the diagonal technique, the difference of the mean value obtained with respect to the real value was 39%. Despite presenting reproducibility in the tests with low standard error, as mentioned above, to obtain precise displacement for position values a new method should be developed for obtaining positioning parameters with this sensor. Further work should be done in order to obtain more parameters to evaluate the long cane full coverage during O&M rehabilitation training.

References

[1] World Health Organisation, *World report on vision*, vol. 214, no. 14. 2019.

[2] ONCE, "DATOS visuales y sociodemográficos de los afiliados a la once 2020." .

[3] American Foundation for the Blind, "International Approaches to Rehabilitation Programs for Adults who are Blind or Visually Impaired: Delivery Models, Services, Challenges, and Trends." 2016.

[4] L. Deverell, "La O & M Environmental Complexity Scale , una escala para medir la complejidad del entorno desde la perspectiva de la Orientación y Movilidad I," *Integr. Rev. Digit. sobre Discapac. Vis.*, vol. 62, pp. 141–163, 2013.

[5] B. B. Blasch, S. La Grow, and W. Penrod, "Environmental Rating Scale for Orientation and Mobility," *Int. J. Orientat. Mobil.*, vol. 1, no. 1, pp. 9–16, 2008.

[6] O. N. de C. Españoles, *Discapacidad Visual y autonomía personal. Enfoque práctico de la rehabilitación*. 2011.

[7] Y. Kim, A. Moncada-Torres, J. Furrer, M. Riesch, and R. Gassert, "Quantification of long cane usage characteristics with the constant contact technique," *Appl. Ergon.*, vol. 55, pp. 216–225, 2016.

[8] D. O. F. Imu, *Miniature Inertial Measurement Unit*. 2018.

[9] A. I. Faisal, S. Majumder, T. Mondal, D. Cowan, S. Naseh, and M. J. Deen, "Monitoring methods of human body joints: State-of-the-art and research challenges," *Sensors (Switzerland)*, vol. 19, no. 11, pp. 1–40, 2019.

[10] D. K. Shaeffer, "MEMS inertial sensors: A tutorial overview," *IEEE Communications Magazine*. 2013.

[11] M. Kok, J. D. Hol, and T. B. Schön, "Using Inertial Sensors for Position and Orientation Estimation," *Found. Trends Signal Process.*, vol. 11, no. 1–2, pp. 1–153, 2017.

[12] T. Tamura, *Wearable Inertial Sensors and Their Applications*, no. 1. Elsevier Inc., 2014.

[13] R. Tapu, B. Mocanu, and T. Zaharia, "Wearable assistive devices for visually impaired: A state of the art survey," *Pattern Recognit. Lett.*, 2018.

[14] K. Fan *et al.*, "Hardware implementation of a virtual blind cane on FPGA," *2017 IEEE Int. Conf. Real-Time Comput. Robot. RCAR 2017*, vol. 2017-July, pp. 344–348, 2018.

[15] A. Dastider, B. Basak, M. Safayatullah, C. Shahnaz, and S. A. Fattah, "Cost efficient autonomous navigation system (e-cane) for visually impaired human beings," *5th IEEE Reg. 10 Humanit. Technol. Conf. 2017, R10-HTC 2017*, vol. 2018-Janua, pp. 650–653, 2018.

[16] Q. K. Dang, Y. Chee, D. D. Pham, and Y. S. Suh, "A virtual blind cane using a line laser-based vision system and an inertial measurement unit," *Sensors (Switzerland)*, vol. 16, no. 1, 2016.

[17] H. Zhang and C. Ye, "Human-Robot Interaction for Assisted Wayfinding of a Robotic Navigation Aid for the Blind," *Int. Conf. Hum. Syst. Interact. HSI*, vol. 2019-June, no. January, pp. 137–142, 2019.

[18] V. V. Meshram, K. Patil, V. A. Meshram, and F. C. Shu, "An Astute Assistive Device for Mobility and Object Recognition for Visually Impaired People," *IEEE Trans. Human-Machine Syst.*, vol. 49, no. 5, pp. 449–460, 2019.

[19] T. C. Davies, C. M. Burns, and S. D. Pinder, "Mobility interfaces for the visually impaired: What's missing?," *ACM Int. Conf. Proceeding Ser.*, vol. 254, pp. 41–47, 2007.

[20] Bosch Sensortec, "BNO055 Intelligent 9-axis absolute orientation sensor, https://cdn-shop.adafruit.com/datasheets/BST_BNO055_DS000_12.pdf," no. November, p. 105, 2014.

[21] B. B. Blasch, S. J. LaGrow, and W. R. De L'Aune, "Three aspects of coverage provided by the long cane: Object, surface, and foot-placement preview," *J. Vis. Impair. Blind.*, vol. 90, no. 4, pp. 295–301, 1996.

[22] V. K. Ramsey, B. B. Blasch, and A. Kita, "Effects of Mobility Training on Gait and Balance," *J. Vis. Impair. Blind.*, vol. 97, no. 11, pp. 720–726, 2003.

[23] S. J. LaGrow, B. B. Blasch, and W. De L'Aune, "Efficacy of the touch technique for surface and foot-placement preview," *J. Vis. Impair. Blind.*, vol. 91, no. 1, pp. 47–52, 1997.

Optimización de la focalidad espacial en tDCS mediante modelos computacionales bioelectromagnéticos: un enfoque alternativo

Pablo Franco-Rosado[1,2,3], M. Amparo Callejón[3], Javier Reina-Tosina[3], Laura M. Roa[3], Juan F Martín-Rodríguez[1,2,4], Pablo Mir[1,2]
pfranco.eng@gmail.com, mcallejon@us.es, jreina@us.es, lroa@us.es , jf.martin.rodriguez@gmail.com, pmir@us.es

(1) Unidad de Trastornos del Movimiento, Servicio de Neurología y Neurofisiología Clínica, Instituto de Biomedicina de Sevilla, Hospital Virgen del Rocío/CSIC/Universidad de Sevilla.

(2)Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED)

(3) Grupo de Ingeniería Biomédica, Universidad de Sevilla.

(4) Departamento de Psicología Experimental, Universidad de Sevilla.

Resumen

El empleo de modelos computacionales bioelectromagnéticos para la simulación y evaluación de las técnicas de neuroestimulación no invasiva se ha asentado en los últimos años como una potente estrategia de para la predicción de la distribución y magnitud de campo eléctrico inducido. En este trabajo se propone un estudio de la estimulación trascraneal de corriente directa (tDCS), en términos de focalidad volumétrica, sobre la región sensomotora, considerándose dos montajes de electrodos: C3Fp2, montaje clásico para este objeto, y P3F3, un montaje alternativo. Los resultados presentan una ventaja concluyente del montaje alternativo para un amplio rango de intensidades de estimulación, en el que se incluyen los más prevalentes en el ámbito clínico. Estos resultados deben complementarse con estudios más extensos y validarse mediante estudios neurofisiológicos.

1. Introducción

Durante los últimos 30 años las técnicas de neuroestimulación no invasiva se han consolidado como instrumentos de gran potencial terapéutico en un amplio espectro de dolencias del sistema nervioso [1]. Entre ellas destaca la estimulación transcraneal de corriente directa (tDCS) por su capacidad para inducir plasticidad neuronal y mejorar la respuesta motora de los pacientes con patologías del ámbito de los trastornos del movimiento [2-3].

La estimulación es vehiculada por una corriente continua que circula entre dos o más electrodos emplazados sobre el cuero cabelludo, actuando sobre las regiones corticales que intersecan la dirección de la corriente. La complejidad del fenómeno de estimulación por tDCS genera una dependencia multifactorial en la excitación cortical inducida, por lo que variables como el espesor del cráneo, la profundidad de los surcos y la geometría de los giros corticales, citando alguna de las de mayor peso específico, provocan una elevada variabilidad interindividual en la respuesta [4]. Con el fin de optimizar y personalizar las terapias se hace necesario ahondar en el alcance de los factores implicados en esta respuesta excitatorial, permitiendo personalizar los parámetros de estimulación a sus características morfológicas y fisiológicas concretas. Es aquí donde los modelos computacionales electromagnéticos realistas de la cabeza se tornan como una potente herramienta para el estudio de la variabilidad paramétrica en la distribución de corriente inducida en la corteza cerebral [5]. Asimismo, una cuestión fundamental es esclarecer el efecto de las diferentes componentes tangencial y normal del campo eléctrico sobre la superficie cortical, y las posibles repercusiones que a nivel neurofisiólogico puede tener cada una de estas en el proceso de estimulación [6]. Existen estudios computacionales previos que ya han abordado la cuestión [7], pero la validación empírica de los resultados obtenidos por los mismos está aún por consolidarse, siendo cuestión de debate.

En este contexto, este estudio propone un flujo computacional que permite simular y analizar la focalidad del campo eléctrico y sus respectivas componentes direccionales sobre determinadas estructuras cerebrales. Para ello, se proponen dos montajes de electrodos distintos: el montaje clásico para la estimulación de la región sensomotora, C3Fp2, y un montaje alternativo, P3F3, que pretende optimizar la focalidad del campo eléctrico, estando las posiciones referenciadas al sistema estándar de colocación de electrodos en EEG. Nuestra motivación reside en optimizar la estimulación sobre esta zona, *target* de especial relevancia el ámbito de los trastornos del movimiento [8] y continuar con la línea de estudio inciciada en [9]. Mediante modelos computacionales simularemos las distribuciones de campo eléctrico inducido sobre el cerebro, diferenciando los resultados obtenidos en la región de interés y realizando un estudio paramétrico de la evolución de la focalidad frente a la intensidad de corriente dispensada.

2. Materiales y métodos

2.1. Modelos computacionales

ISBN: 978-84-09-25491-0, pags. 198- 200

Para la construcción de los modelos se dispone de resonancias magnéticas estructurales T1 y T2 de alta resolución, adquiridas con un escáner PHILIPS Ingenia 3T, de 12 sujetos (7 hombres) de edad promedio de 28±3 años, que dieron su consentimiento para participar en el estudio, siendo este aprobado previamente por un Comité de Ética independiente.

En primer lugar, estas imágenes sirven como entrada al flujo de trabajo del software de modelado y simulación SimNIBS 3.1.1(https://simnibs.github.io), que tras segmentar los tejidos y optimizar la malla volumétrica resultante, proporciona un modelo realista de la cabeza como conjunto de mallas tetraédricas de los diferentes tejidos: materia blanca (WM), materia gris (GM), líquido cefalorraquídeo (CSF), hueso y cuero cabelludo. Se ha optado por la máxima resolución posible para el mallado tetraédrico de cara a conseguir la máxima fidelidad anatómica de los dominios, siendo el número promedio de tetraedros para los 12 modelos de 19,34 millones de tetraedros.

Posteriormente, extraemos la región sensomotora realizando una intersección del modelo con una malla superficial de geometría semejante a la región de interés (ROI). Partimos de una máscara binaria obtenida mediante la edición y combinación de las máscaras correspondientes al giro precentral y postcentral del hemisferio izquierdo del "MNI152 Harvard-Oxford Cortical Structural Atlas", atlas anatómico incluido en el entorno de neuroimagen FSL (https://fsl.fmrib.ox.ac.uk). Esta máscara será transformada desde espacio MNI a espacio sujeto y posteriormente traducida a la deseada malla triangular mediante 3D Slicer (https://www.slicer.org/). La operación morfológica se realiza mediante un *script* de MATLAB diseñado *ad hoc* y

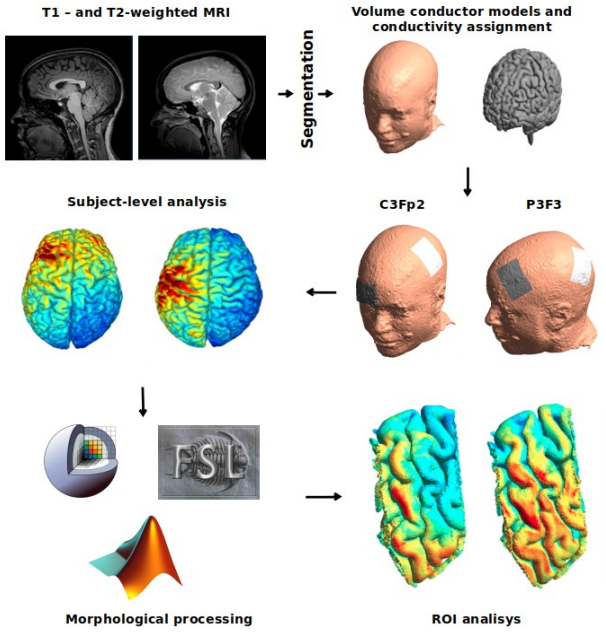


Figura 1. *Flujo de trabajo propuesto. Desde las resonancias de alta definición se generan modelos realistas de la cabeza sobre los que se calcula la distribución de campo eléctrico inducida. Tras las operaciones morfológicas necesarias, obtenemos la GM de la región de interés para cada sujeto con sus correspondientes valores de campo.*

que nos proporciona una malla que contiene únicamente los tetraedros pertenecientes a la región de GM confinada en el interior de la malla superficial (figura 1).

2.2. Configuraciones de electrodos

Se definen electrodos de goma rectangulares de 4.5 x 6.5 cm² de 1 mm de espesor que se sustentan sobre una capa de gel conductor de tres milímetros de espesor. Su conector estará posicionado un milímetro por debajo del centro del eje longitudinal, siendo sus dimensiones de 1.6x2.3 cm². La importancia del detalle en la caracterización de los electrodos en la distribución de campo inducida se recoge en [9], optando para este estudio por unas dimensiones típicas de los electrodos rectangulares de tDCS.

Se emplea el sistema internacional 10/20 para determinar la posición de los electrodos. Para el montaje clásico, ánodo ubicado en C3 y cátodo en Fp2. El montaje alternativo dispondrá el ánodo en P3 y el cátodo en F3.

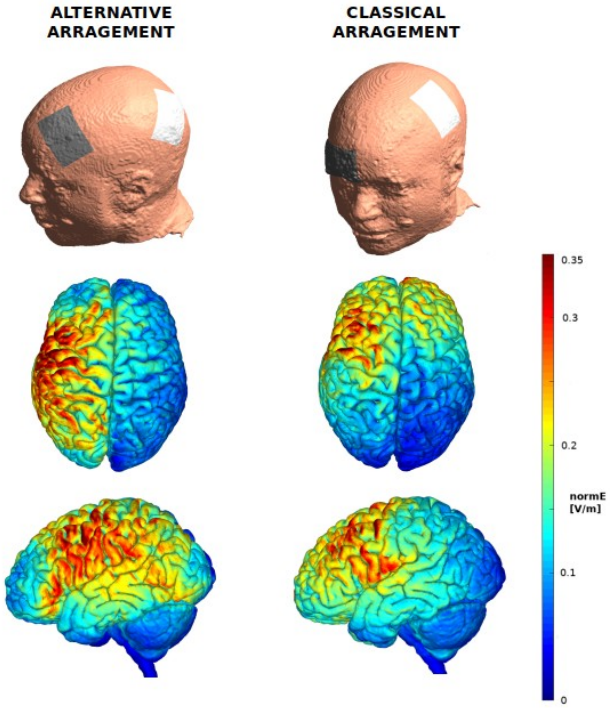


Figura 2. *Magnitud del campo eléctrico para ambos montajes sobre uno de los sujetos para una intensidad de corriente de 1 mA.*

2.3. Simulaciones mediante elementos finitos

Mediante SimNIBS 3.1.1 simulamos la física de la estimulación con corriente continua. Para cada elemento del dominio se resuelve la ecuación de Laplace en condiciones cuasi-estacionarias. Se considera aplicable la conservación de la componente normal de la densidad de corriente en la interfaz entre los distintos tejidos (condición de contorno de Neumann), quedando los límites ulteriores aislados eléctricamente. Para caracterizar dieléctricamente los tejidos se emplearon las conductividades recogidas en la literatura: 0.126 S/m para WM, 0.275 S/m para GM, 1.654 S/m para CFS, 0.010 S/m para el hueso y 0.465 S/m para el cuero cabelludo. A los electrodos se les asignó una conductividad de 1 S/m y al gel conductor 29.4 S/m [10].

Con el objetivo de determinar el efecto que sobre la focalidad del campo tiene el incremento de la intensidad de corriente introducida, se realizaron 34 simulaciones por sujeto (17 por montaje), donde la intensidad varía con un paso de 0.25 mA en el intervalo de 0.5 a 4.5 mA. Así, con un total de 374 simulaciones, obtenemos el campo eléctrico inducido como solución del problema bioelectromagnético, el cual se reformula en como focalidad volumétrica (figura 2). Definimos la focalidad volumétrica como el ratio entre el volumen que supera cierto umbral de estímulo frente al volumen total la región. Tomamos este umbral como un percentil del máximo valor de campo registrado en todo el cerebro. Se estudiarán los percentiles 70, 75 y 80 y se compararan las medias de los resultados de ambos montajes mediante el test t-Student.

3. Resultados

En la figura 3 podemos observar la evolución del valor-p con el aumento de la intensidad, pudiéndose apreciar como en el rango comprendido entre los 1.25 y 4.25 mA existe una diferencia significativa entre ambos montajes ($p<0.05$).

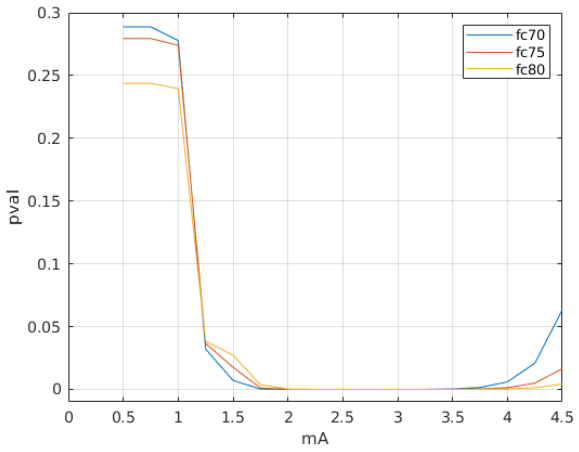


Figura 3. *Evolución del p-valor con la intensidad de estimulación para distintos umbrales de campo eléctrico.*

La figura 4 muestra la focalidad en la región sensomotora frente a la intensidad para un sujeto representativo. La tendencia dominante del montaje P3F3 se mantiene hasta superados los 4 mA.

4. Discusión y conclusiones

Podemos observar una tendencia a la saturación del volumen de interés según aumenta la intensidad de estimulación. Si bien los resultados confirman el ineludible compromiso existente entre focalidad e intensidad, es de especial relevancia la consistente superioridad del montaje alternativo frente al clásico en términos de focalidad en el intervalo de 1.5 a 3 mA, que abarca la mayor parte de los protocolos de actuación clínica en esta técnica. Nuestros resultados sugieren que el montaje alternativo es capaz de estimular mayores volúmenes de corteza en la ROI con menores intensidades. De esta forma, la propuesta presentada permite optimizar el campo eléctrico en la región sensomotora, presentando un robusto comportamiento ante la variabilidad interindividual de la geometría de zona, factor clave para determinar estrategias de estimulación adecuadas en el dominio clínico. Confirmar esta tendencia con una mayor muestra, así como extender el estudio a las distribuciones de las componentes de campo normal y tangencial inducidas, proporcionará un completo estudio *in-silico* que nos permitiría considerar extender estas hipótesis al ámbito experimental.

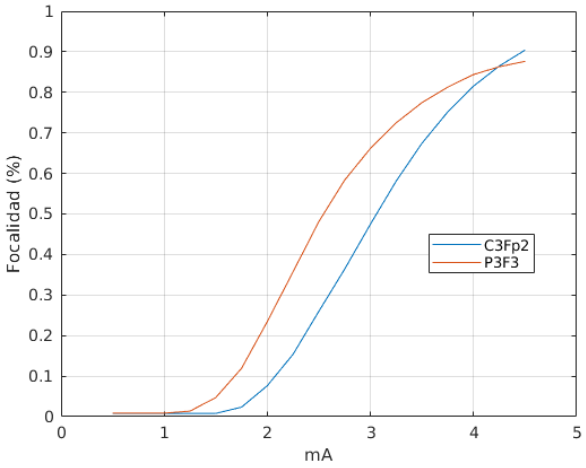


Figura 4. *Focalidad (%) frente a intensidad de estimulación (mA) para uno de los sujetos de la muestra y los dos montajes simulados.*

Referencias

[1] S. Luan, I. Williams, and K. Nikolic, "Neuromodulation: present and emerging methods," Front., 2014, Accessed: Dec. 18, 2016.

[2] F. Yavari, A. Jamil, M. Mosayebi Samani, L. P. Vidor, and M.A. Nitsche, "Basic and functional effects of transcranial Electrical Stimulation (tES)—An introduction," Neurosci. Biobehav. Rev., vol. 85, no. July 2017, pp. 81–92, 2018.

[3] D. J. Kidgell, A. M. Goodwill, A. K. Frazer, and R. M. Daly, "Induction of cortical plasticity and improved motor performance following unilateral and bilateral transcranial direct current stimulation of the primary motor cortex," BMC Neurosci., vol. 14, 2013.

[4] A. Opitz, W. Paulus, S. Will, A. Antunes, and A. Thielscher, "Determinants of the electric field during transcranial direct current stimulation," Neuroimage, vol. 109, pp. 2015.

[5] P. C. Miranda, M. A. Callejón-Leblic, R. Salvador, and G. Ruffini, "Realistic modeling of transcranial current stimulation: The electric field in the brain," Current Opinion in Biomedical Engineering, vol. 8. 2018.

[6] P. T. Fox et al., "Column-Based Model of Electric Field Excitation of Cerebral Cortex," Hum. Brain Mapp., vol. 22, no. 1, pp. 1–14, 2004.

[7] M. A. Callejón-Leblic and P. C. Miranda, "A Computational Parcellated Brain Model for Electric Field Analysis in Transcranial Direct Current Stimulation," in Brain and Human Body Modeling 2020, Springer International P ublishing, 2021, pp. 81–99.

[8] D. Doruk, Z. Gray, G. L. Bravo, A. Pascual-Leone, and F. Fregni, "Effects of tDCS on executive function in Parkinson's disease," Neurosci. Lett., vol. 582, 2014.

[9] Franco et.al, Modelos computacionales bioelectromagnéticos para optimizar la focalidad espacial en las técnicas tDCS. Actas del URSI 2020. Por publicar.

[10] G. B. Saturnino, A. Antunes, and A. Thielscher, "On the importance of electrode parameters for shaping electric field patterns generated by tDCS," Neuroimage, vol. 120, pp. 25–35, 2015.

Comparison of electrode positions for a multi-field FES device for compensation of foot drop

A. Martín Odriozola[1], L. Francés Morcillo[1], C. Rodríguez-de-Pablo[1], H. Zabaleta Rekondo[1]

[1] Fesia Technology, Donostia, España {amartin, lfrances, crodriguez, hzabaleta}@fesia.net

Abstract

Neurologic diseases often cause "foot drop", a gait impairment in which the patient is not able to lift the foot due to the weakness or absence of ankle dorsiflexors. Functional electrical stimulation (FES) is one of the options for treating foot drop. Surface multi-field electrodes have emerged in the last decade, showing great potential on improving selectivity of stimulation and providing easier donning and doffing. So far, usability and clinical results of FES have been researched. This preliminary study attempts to analyse and compare the movements created using different electrode positions with a commercial multi-field technology-based FES device. 6 healthy subjects participated in the study, and one configuration was held with 2 electrode positions within 3 days. Results show that when the electrode is located more caudally, the capability of the system to produce clear dorsiflexion is increased, while when it is located more cranially, it increases the capacity to induce ankle eversion and plantar flexion. No significant differences were observed between both electrode positions. This study suggests that electrode position influences the elicited ankle movements and that therefore, it should be considered in the configuration process.

1. Introduction

Neurologic patients often show "foot drop", a gait impairment consisting of the inability to lift the foot due weakness or absence of ankle dorsiflexors [1]. These patients usually adopt a steppage or equine gait, resulting in the inability to perform meaningful daily activities [2]. Promoting motor recovery to correct abnormalities of hemiplegic gait is essential for these patients [3].

Functional electrical stimulation (FES) consists of providing short bursts of electrical pulses to induce action potentials in the motor neurons that innervate a muscle, eliciting muscle contractions and thus, restoring motor function [4]. It is one of the options for treating foot drop, and its use is widely supported by the scientific evidence [5]. Most commercial FES-based devices stimulate the common peroneal nerve, producing an ankle dorsiflexion during gait. Some stimulators also produce plantar flexion, resulting in better knee flexion and better swing phase [6].

The configuration of FES devices based on conventional 1 -and-1 electrodes is done manually, being very time-consuming and requiring a big clinical experience [7]. However, surface multi-field electrodes have emerged in the last decade, showing great potential on improving selectivity of stimulation and providing easier donning and doffing [8]. Another remarkable feature is the possibility of doing the motor-point search automatically by means of a configuration process involving sensors [9].

In this pilot study we present an analysis that compares the movements created using different electrode positions with a commercial multi-field technology-based FES device.

2. Methods

2.1. Material

For this study, the Fesia Walk system (Fesia Technology, Donostia – San Sebastián, Spain) was used. It is a commercial multi-field technology-based FES device, used for the treatment of foot drop. The system, shown in Figure 1, consists of a stimulator, garment with integrated multi-field electrode (comprises 16 cathodes and 4 anodes), wireless inertial sensor and tablet PC with software application. It works delivering a train of biphasic pulses of different widths and amplitudes to different fields of the electrodes in an asynchronous manner.



Figure 1. *The Fesia Walk system (Fesia Technology, Donostia – San Sebastián, Spain). a Stimulation unit, garment, multi-field electrode and wireless sensor. b Tablet PC with software application. c Neurorehabilitation with FES.*

The configuration process of the Fesia Walk system can be performed in an automatic manner, which consists on the activation of each of the 16 electrode fields individually with different stimulation amplitudes. The sensor, placed

ISBN: 978-84-09-25491-0, pags. 201- 204

on the foot, comprises a gyroscope and an accelerometer, and measures the ankle movements elicited in response to the stimuli applied to each field. This sensor includes a classification algorithm that classifies the movements in 3 movement groups: eversion, dorsiflexion or plantarflexion. It also selects the fields that produce these movements optimally [9].

2.2. Experimental protocol

6 healthy volunteers aged between 19 and 53, who signed an informed consent, were included in this study. The laterality was right for 3 of them and left for the other 3.

This study lasted 3 days. All the configurations were done with the subjects sitting on the top of a table or stretcher with knee at 90 degrees and foot not touching the floor. 2 different electrode positions were included in the study (Figure 2):

- Position 1: electrode positioned on the shank taking as a reference the midline between anodes 2 and 3, which was aligned with the tibia, 2-2,5 cm below the patella.
- Position 2: electrode positioned on the shank taking as a reference the midline between anodes 2 and 3, which was aligned with the tibia, 4-4,5 cm below the patella.

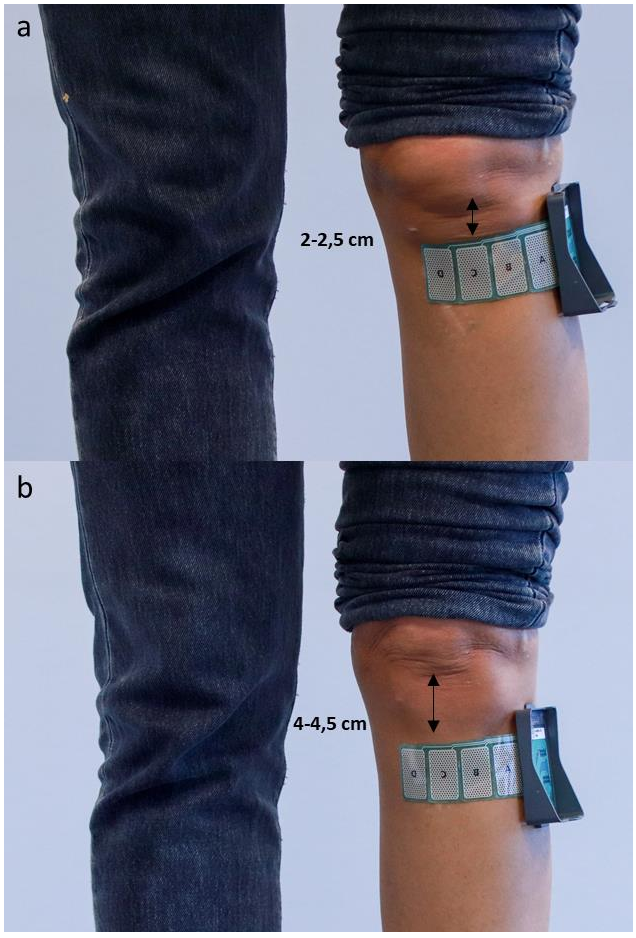


Figure 2. *Electrode positions included in the study. a Position 1. b Position 2.*

A minimum of 5 minutes of rest between electrode positions was ensured to avoid fatigue affecting results.

The stimulation waveform was biphasic compensated. Parameters were set to frequency 25 Hz and pulse width 250 µs. Anodes were fixed and set as B and C in Fig. 1.

The amplitude was set for each day and electrode position at 10 mA above motor threshold if tolerated by the subject, otherwise above motor threshold at subject's tolerance. The motor threshold was defined as the minimum amplitude at which foot movement was visually appreciated with at least one of the fields.

2.3. Analysis

The analysis focused on the movements assigned by the classification algorithm depending on the electrode position. The aim was to find differences in terms of generated movements between different electrode positions. To make statistical comparisons between the two groups, the unpaired Student's t-test was employed. The degree of significance was established at $p = 0.05$ (95% confidence limit).

3. Results

In Figure 3, the ankle movement classification results obtained with different electrode positions in different subjects are shown.

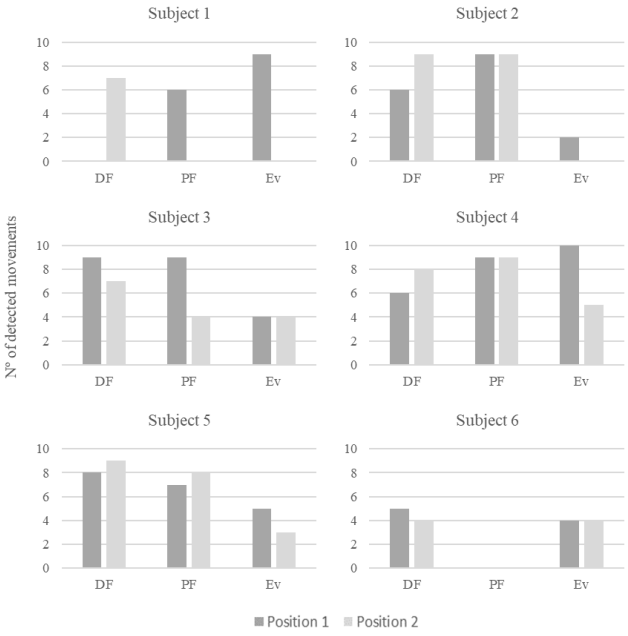


Figure 3. *Ankle movements selected by the classification algorithm for each electrode position and each subject. DF = dorsiflexion, PF = plantarflexion, EV = eversion.*

The figures show that the generated movements were very dependent on subjects. Subjects 3, 4 and 5 obtained all the movements with both electrode placement positions. In the rest of the subjects, however, not always all the movements were generated with both positions of the electrode. In subject 1, dorsiflexion was only generated with electrode position 2, while plantar flexion and eversion only occurred when the electrode was in position 1. In subject 2, eversion was not generated with electrode position 2. Finally, neither of the two positions was capable to induce plantar flexion in subject 6.

Figure 4 shows the number of times that the classification algorithm did select each of the three movements (dorsiflexion, plantar flexion, and eversion) with each electrode position.

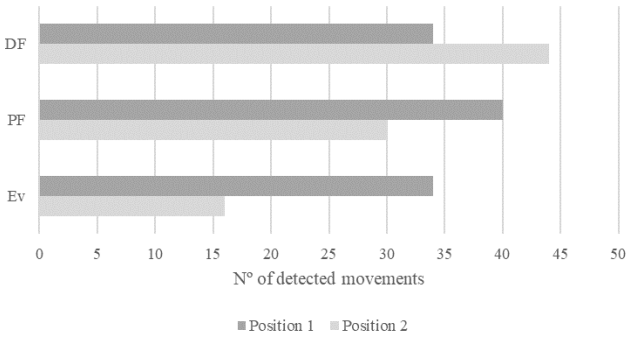


Figure 4. *Ankle movements selected by the classification algorithm por each electrode position. DF = dorsiflexion, PF = plantarflexion, EV = eversion.*

If we focus on dorsiflexion, it was detected 34 times with electrode position 1 and 44 times with electrode position 2. On the other hand, plantarflexion was produced 40 times with electrode position 1 and 30 times with electrode position 2. Finally, a similar behaviour occurred with the eversion, which was detected 34 times with electrode position 1 and 16 times with electrode position 2.

No statistically significant differences were found between groups in any of the three movements (dorsiflexion $p = 0.29$; plantarflexion $p = 0.48$; eversion $p = 0.08$).

4. Discussion

The aim of the study was to determine the optimal electrode position for the configuration of the multi-field FES system, continuing with the work developed by our research group to optimize the use of these devices for the rehabilitation of foot drop [10-12].

Although the differences were not statistically significant, this analysis proved that when the electrode was located more caudally (position 2), the capability of the system to produce clear dorsiflexion was increased. On the other hand, the most cranial position of the electrode (position 1) showed the capacity to induce ankle eversion and plantar flexion more often.

These differences on the movements generated by the different electrode positions could be due to their relative position with respect to the muscular and nervous structures of the leg. The height at which the electrode is placed in position 1, coincides with a level at which both the tibial nerve and the common peroneal nerve are more superficially located. This allows them to be stimulated more easily, generating plantar flexion and eversion, respectively. At this height, the tibial nerve passes under the tendinous arch formed between the tibial and peroneal heads of the soleus muscle, before continuing vertically through the deep region of the posterior compartment of the leg on the surface of the posterior tibial muscle. In addition, the common peroneal nerve, whose stimulation generates eversion and dorsal flexion, is in a very superficial location. Then, it is bifurcated into the deep peroneal nerve and the superficial peroneal nerve behind the head of the fibula. If the electrode is placed more caudally, however, the tibial nerve is in a deeper location, which makes it difficult to achieve plantar flexion. Also, the common peroneal nerve has already bifurcated, allowing dorsal flexion to appear a greater number of times, since the deep peroneal nerve only produces movements in that direction, and not towards eversion.

This finding is interesting, since it could be used in clinical practice to select different electrode positions, depending on the movement that is abolished in each patient. This could increase the selectivity of the stimulation, improve the patient's sensation, decrease the set-up times etc. All this could contribute to increase the therapy intensity, improving patient motivation, which would end up increasing the motor rehabilitation and improving the quality of life of the patient.

5. Conclusions

We conclude that position 2, where the electrode is placed 4-4,5 cm below the patella, is the most appropriate electrode position for achieving a clear ankle dorsal flexion with the Fesia Walk device. On the other hand, the eversion and the plantarflexion movements are elicited better with electrode position 1, in which it is located 2-2,5 cm caudal to the patella. This contribution is interesting, as it can be used in clinical practice to personalise electrode placement in each patient, obtaining improvements in clinical and usability terms. Although the study was carried out with healthy people, these results can be extrapolated to patients with neurological pathology, since it has been studied a characteristic that does not vary if there is a neurological disorder. Further research with people with neurological disorders is necessary to confirm these conclusions.

Acknowledgment

This work has been supported by Fesia Technology S.L.

References

[1] Perry J, Davids JR. Gait analysis: normal and pathological function. Journal of Pediatric Orthopaedics, vol 12, 1992, pp 815 (ISSN: 0271-6798).

[2] Graham J. Foot drop: explaining the causes, characteristics and treatment. British Journal of Neuroscience Nursing, vol 6, sup 4, 2010, pp 168–172 (ISSN: 2052-2800).

[3] Howlett OA, Lannin NA, Ada L, McKinstry C. Functional electrical stimulation improves activity after stroke: A systematic review with meta-analysis. Archieves of Physical Medicine and Rehabilitation, vol 96, sup 5, 2015, pp 934-943 (ISSN: 0003-9993).

[4] Renfrew L, Paul L, McFadyen A, Rafferty D, Moseley O, Lord AC, Bowers R, et al. The clinical- and costeffectiveness of functional electrical stimulation and anklefoot orthoses for foot drop in Multiple Sclerosis: a multicentre randomized trial. Clinical Rehabilitation, 2019, vol 33, sup 7, pp 1150-1162 (ISSN: 1477-0873).

[5] Sheffler LR, Chae J. Neuromuscular electrical stimulation in neurorehabilitation. Muscle & Nerve, vol 35, sup 5, 2007, pp 562-590 (ISSN: 1097-4598).

[6] Embrey DG, Holtz SL, Alon G, et al. Functional electrical stimulation to dorsiflexors and plantar flexors during gait to improve walking in adults with chronic hemiplegia. Archives of Physical Medicine and Rehabilitation, vol 91, sup 5, 2010, pp 687-696 (ISSN: 0003-9993).

[7] Koutsou AD, Moreno JC, Del-Ama AJ, Rocon E, Pons JL. Advances in selective activation of muscles for non-invasive motor neuroprostheses. Journal of NeuroEngineering and Rehabilitation, vol 13, sup, 1, 2016, pp 1-12 (ISSN: 1743-0003).

[8] Imatz-Ojanguren E, Sánchez-Márquez G, Asiain-Aristu JR, Cueto-Mendo J, Jaunarena-Goicoechea E, Zabaleta H, Keller T. A foot drop compensation device based on surface multi-field functional electrical stimulation-Usability study in a clinical environment. Journal of Rehabilitation and Assistive Technologies Engineering, vol 6, 2019 (ISSN: 2055-5583).

[9] Malešević J, Dedijer-Dujović S, Savić AM, Konstantinović L, Vidaković A, Bijelić, G, Malešević N, et al. A decisión support system for electrode shaping in multi-pad FES foot drop correction. Journal of NeuroEngineering and Rehabilitation, vol 14, sup 1, 2017, pp 1–14 (ISSN: 1743-0003).

[10] Martin A, Aranguren A, Imatz-Ojanguren E, Zabaleta H, Keller T. Repeatability of calibration process of a multi-field FES device for compensation of foot drop. Proceedings of the XXXVII Annual Congress of the Spanish Society of Biomedical Engineering (CASEIB 2019), Santander, 2019, pp 263-266 (ISBN: 978-84-09-16707-4).

[11] Martin A, Rodriguez-de-Pablo C, Zabaleta H, Imatz-Ojanguren E, Keller T. Comparison of Configuration Postures for a Foot Drop Multi-Field FES Device. 5th International Conference on NeuroRehabilitation (ICNR2020), virtual format, 2020.

[12] Martin A, Rodriguez-de-Pablo C, Zabaleta H, Imatz-Ojanguren E, Keller T. Effect of Gel Type and Anode Selection in Ankle Movements Elicited by a Multi-Field FES Device. 5th International Conference on NeuroRehabilitation (ICNR2020), virtual format, 2020.

Sistemas de Ayuda a la Decisión I

Miércoles, 25 de noviembre de 2020

Método de reconstrucción tomográfica con información a priori obtenida con aprendizaje profundo

A. Piol[1,2], C. F. del Cerro[1,2], J. García-Blas[3], M. Desco[1,2,4,5], M. Abella[1,2,5]

[1] *Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III, Madrid, Spain*

[2] *Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), Madrid, Spain*

[3] *Departamento de Informática, Universidad Carlos III de Madrid, Madrid, España*

[4] *Centro de Investigación en Red de Salud Mental (CIBERSAM, CIBER CB07/09/0031), 28007 Madrid, Spain*

[5] *Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, Spain*

Resumen

En sistemas de tomografía computarizada tradicionales, los datos se adquieren girando fuente y detector alrededor de la muestra completando un rango angular de 360 grados y obteniendo, generalmente, más de 360 proyecciones. Sin embargo, existen situaciones donde solo es posible obtener un número limitado de proyecciones, como en el caso del uso de sistemas de arco en C en un entorno intraoperatorio. Esta limitación de datos resulta en artefactos específicos en la reconstrucción y, consecuentemente, en pérdida de detalles importantes para el diagnóstico clínico. Una solución para compensar la falta de datos es utilizar información a priori en algoritmos iterativos. El algoritmo SCoLD permite la reconstrucción del volumen con un número limitado de proyecciones, incorporando como información a priori el soporte de la muestra, identificado con una máscara binaria.

En este trabajo se propone una metodología de reconstrucción de datos incompletos basado en la combinación de un algoritmo de reconstrucción que incorpora el soporte de la muestra como información a priori con un paso de obtención de este soporte usando metodología de aprendizaje profundo.

1. Introducción

En sistemas de tomografía computarizada, el par fuente-detector gira simultáneamente alrededor del paciente, cubriendo generalmente 360 grados, adquiriendo un conjunto de proyecciones (radiografías planas) desde diferentes posiciones angulares. Cuando el conjunto de proyecciones cumple determinados requisitos, de número y rango angular, la reconstrucción del volumen bajo análisis se puede hacer utilizando algoritmos tradicionales basados en retroproyección filtrada [1]. Sin embargo, existen diferentes situaciones en las que el número de proyecciones adquiridas no alcanza el mínimo para poder utilizar estos algoritmos. Un ejemplo es el uso de arcos en C para obtener imagen tomográfica en quirófano, que presentan limitaciones mecánicas que no permite cubrir todo el rango angular. Otro caso de limitación de datos es la selección de un paso angular grande como medio para reducir las dosis de radiación. En estos casos, el resultado de reconstrucciones con algoritmos convencionales basado en FDK está afectado por dos tipos de artefactos: rayas, en el caso de un paso angular demasiado grande, y distorsión de los bordes, en el caso de un rango angular limitado.

Una forma de compensar la falta de datos es la incorporación de información a priori en algoritmos iterativos. Ejemplos de información a priori son la positividad de la solución o la característica de homogeneidad de los tejidos. Esta última característica se puede conseguir con métodos de regularización basados en la función de Total Variation (TV) [2-4]. Sin embargo, regularizar con la función TV no soluciona los artefactos causados por un rango angular limitado. Para estos casos, se ha propuesto incorporar como información previa una reconstrucción con datos completos. En [5], se utiliza una reconstrucción con datos completos del mismo paciente en un estudio con *gating* respiratorio, obtenida a partir de las proyecciones de las diferentes fases respiratorias. Una alternativa que se ha propuesto para la reconstrucción de datos limitados en rayos X se basa en técnicas de aprendizaje profundo [6, 7]. Estos trabajos tienen limitaciones en resolución espacial, con un tamaño máximo de imagen de 256×256 pixeles. Por otra parte, para evitar sesgos que resulten en estructuras en la imagen final que realmente no están presentes en los datos, es necesario disponer de una base de datos reales grande. Otra información a priori que se ha propuesto en la literatura es la restricción del espacio de búsqueda de la solución en un determinado área del campo de visión en la que sabemos que se encuentra la muestra. En [8], se presenta *Support-Constraint Limited Data* (SCoLD), un algoritmo de reconstrucción que utiliza el contorno de la muestra para definir la región de búsqueda de la solución. La Figura 1 muestra en ejemplo de resultados aplicando el algoritmo SCoLD a un conjunto de datos de 70 proyecciones en un rango angular de 140 grados, que permite reducir los artefactos y alcanzar una calidad de imagen muy parecida a la de un TAC estándar (360 proyecciones con 360 grados de rango angular).

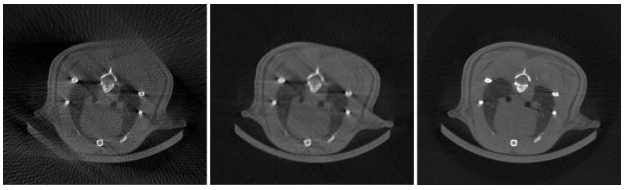


Figura 1: *Corte axial de un estudio de datos limitados de un pequeño animal reconstruido con FDK (izquierda) y SCoLD (centro). Reconstrucción con datos completos utilizando algoritmo FDK* (derecha).

En [8], el contorno de la muestra se obtiene mediante un escáner de superficie. La adquisición de esta superficie

ISBN: 978-84-09-25491-0, pags. 206- 209

puede complicar el protocolo clínico e incluso ser imposible por la existencia de tejido cubriendo al paciente. En este trabajo se propone eliminar los problemas derivados del uso de dispositivos externos obteniendo este contorno por medio de técnicas de aprendizaje profundo.

2. Método propuesto

El método propuesto consiste en tres pasos pasos: (1) reconstrucción preliminar con FDK, que será la entrada a la red, (2) obtención de la máscara binaria que define el soporte de la muestra con la red y (3) uso de esta máscara en SCoLD para obtener la reconstrucción final sin artefactos.

Para la obtención de la máscara, se utiliza U-net, red propuesta para segmentación en [9], con el Error Cuadrático Medio (MSE) como función de coste. Como parámetros de entrenamiento se utilizan una tasa de aprendizaje de 10^{-3} durante 40 épocas y el optimizador Adam [10]. Los datos de entrada de la red son cortes 2D de tamaño 256×256×1. Para el proceso de enteramiento y validación del modelo se ha utilizado Python (versión 3.8). La Figura 2 muestra un esquema de la arquitectura de la red utilizada.

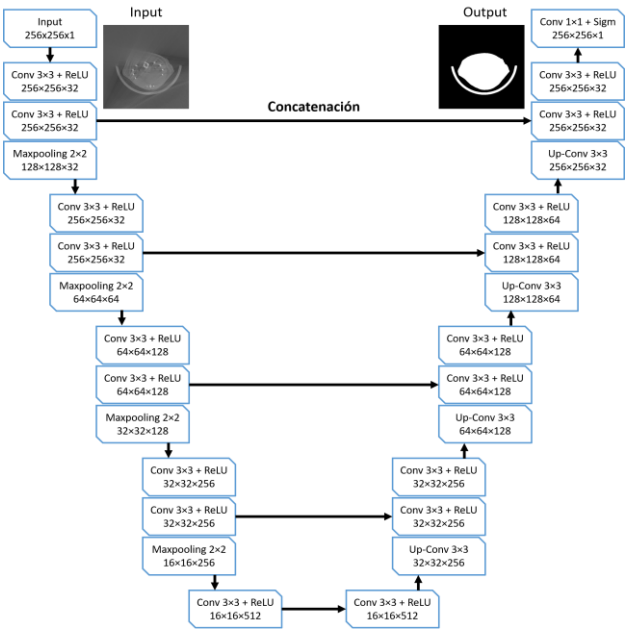


Figura 2: *Arquitectura de la red utilizada.*

Para poder utilizar la máscara binaria en el algoritmo SCoLD, las imágenes en salida a la red se redimensionan al tamaño de la reconstrucción.

3. Generación de la base de datos

La base de datos para el entrenamiento del modelo se ha creado a partir de cuatro estudios de roedor adquiridos con el sistema micro CT Add-on de SEDECAL [11]. Cada estudio está formado por 360 proyecciones de 768×486 pixeles obtenidas en un rango angular de 360 grados, distribuidas uniformemente. La máscara ideal se ha obtenido por umbralización en la reconstrucción de los datos completos con un método basado en FDK [12], un volumen de tamaño 768×768×486.

Se han simulado con FuxSim [13] cuatro estudios eliminando proyecciones de manera aleatoria para cada uno de los siguientes escenarios de datos limitados (Figura 3): (1) Rango angular limitado, (2) número de proyecciones limitado y (3) rango angular y número de proyecciones limitados.

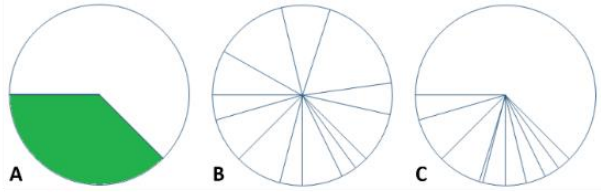


Figura 3: *Esquema de selección de rango angular y proyecciones para los casos de rango angular limitado (A), numero de proyecciones limitado (B) y rango angular y número de proyecciones limitados (C).*

El conjunto de datos está compuesto por los cortes centrales de cada estudio, obteniendo un total de 8904 imágenes, separados en 6410 para entrenamiento y 1781 para validación. Dado que lo que esperamos obtener de la red es información paso bajo (la máscara soporte del objeto), estas imágenes se han submuestreado al tamaño de 256×256.

La Figura 4 muestra ejemplos de cortes para cada escenario de datos limitados junto a la reconstrucción de referencia.

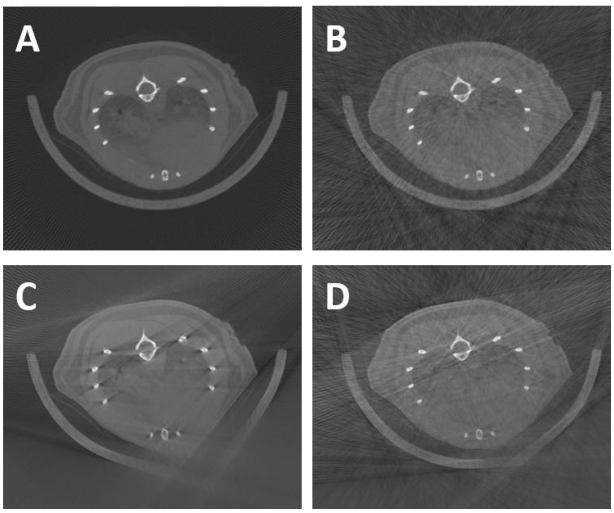


Figura 4: *Corte axial central de la reconstrucción FDK de datos completos (A), datos con un numero de proyecciones limitado (B), datos con rango angular limitado (C), datos con número de proyecciones y rango angular limitados (D).*

4. Evaluación y Resultados

Para evaluar la calidad de la máscara obtenida se ha utilizado el coeficiente de Dice-Sørensen (DSC) con respecto a la máscara ideal:

$$\mathrm{DSC}=2\cdot|X\cap Y||X|+|Y| \qquad (1)$$

donde X y Y son los dos volúmenes binarios de comparar.

Los datos limitados se han reconstruido con FDK, con SCoLD incorporando la máscara obtenida por umbralización, con SCoLD utilizando la máscara ideal y con el método propuesto.

Para evaluar el efecto en la reconstrucción final de la máscara utilizada como información a priori, se ha calculado la raíz del error cuadrático medio (RMSE) entre FDK con datos completos (referencia) y todas las reconstrucciones obtenidas con SCoLD, con las diferentes mascaras. Se presentan resultados en dos estudios representativos: (1) rango angular de 130 grados, con un paso angular de 1 grado y (2) rango angular de 140 grados, con un paso angular de 2 grados (70 proyecciones totales).

La Figura 5 muestra el caso de rango angular de 130 grados y paso angular de 1 grado. Se puede ver la distorsión de los bordes en la parte superior izquierda de la reconstrucción de datos limitados, que se reflejan en la máscara obtenida por umbralización. La máscara obtenida con la red propuesta elimina estos artefactos en los bordes y permite una cama está bien definida y separada de la muestra, como ocurre en la máscara ideal.

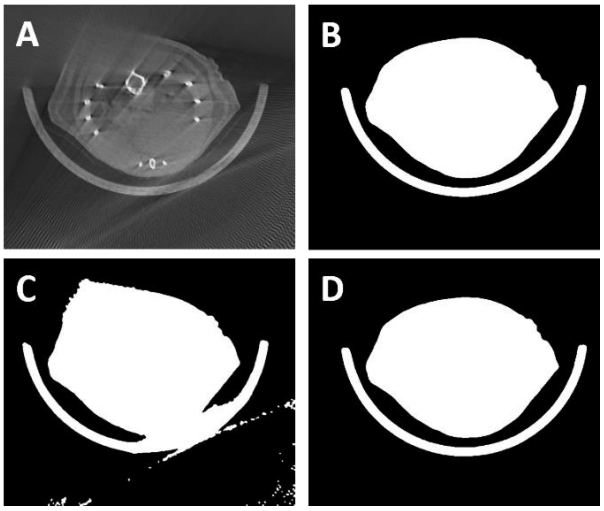


Figura 5: *Corte axial de la reconstrucción FDK de datos limitados (A), máscara ideal (B), máscara obtenida por umbralización (C) y máscara obtenida del método propuesto. Rango angular de 130 grados y paso angular de 1 grado.*

La Figura 6 muestra el caso de rango angular de 140 grados y paso angular de 2 grados (70 proyecciones).

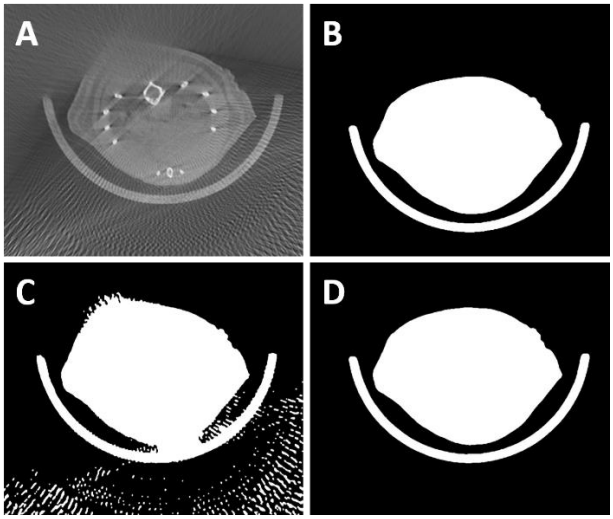


Figura 6: *Corte axial de la reconstrucción FDK de datos limitados (A), mascara ideal (B), máscara obtenida por umbralización (C) y máscara obtenida con el método propuesto. Rango angular de 140 grados y paso angular de 2 grados.*

La máscara obtenida por umbralización a partir de la reconstrucción con FDK muestra la distorsión de los bordes en la parte arriba-izquierda y en la inferior derecha y los artefactos de rayas de la reconstrucción FDK de datos limitados. Estos artefactos se eliminan con el método propuesto, quedando la cama bien definida y separada de la muestra, como ocurre en la máscara ideal.

La Figura 7 muestra los resultados de las reconstrucciones con SCoLD, donde se puede ver como el método propuesto elimina los artefactos de distorsión de bordes, obteniendo un resultado similar al de SCoLD + máscara ideal.

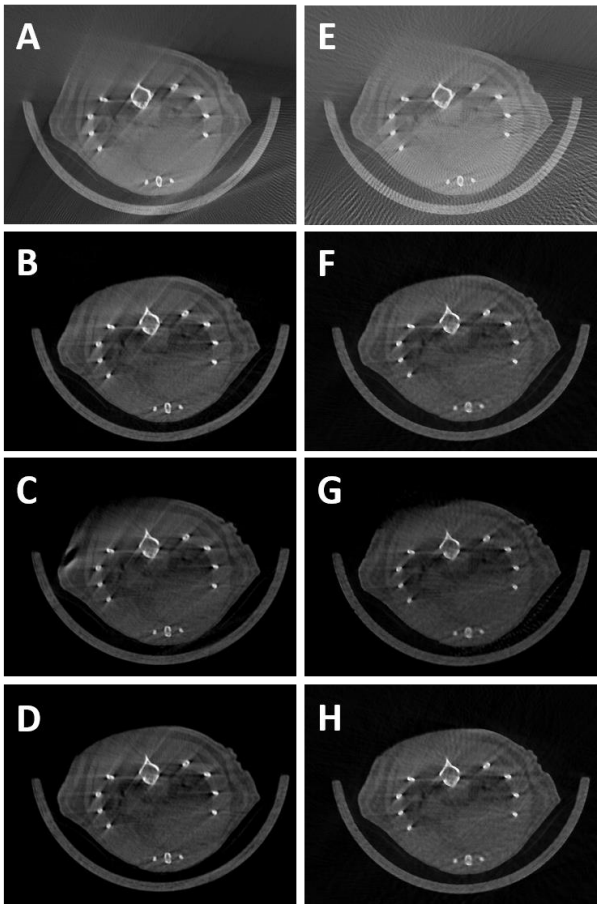


Figura 7*: Izquierda: estudio con 130 grados de rango angular, reconstruido con FDK (A), SCoLD + máscara ideal (B), SCoLD + máscara por umbralización (C) y método propuesto (D). Derecha: estudio con 140 grados de rango angular y 2 grados de paso angular reconstruida con FDK (E), SCoLD + máscara ideal (F), SCoLD + máscara datos limitados (G) y método propuesto (H).*

La Tabla I muestra coeficientes de DICE cercanos a 1 y un valor de RMSE inferior para el método propuesto.

TABLA I

		Rango 130˚, paso 1˚	Rango 140˚, paso 2˚
DICE		0.993	0.995
RMSE	FDK	267	332.40
	SCOLD + umbralización	164.37	165.33
	Método propuesto	156.56	159.15

5. Discusión y conclusiones

Este trabajo presenta una metodología de reconstrucción de datos incompletos basado en la combinación de SCoLD, un algoritmo de reconstrucción que incorpora el soporte de la muestra como información a priori, con un paso de obtención de este soporte usando metodología de aprendizaje profundo. En trabajos previos, este soporte se obtenía con un escáner de superficie, no siempre posible por la presencia de ropa o sábanas tapando al paciente y suponiendo una dificultad extra a proceso de adquisición. El método propuesto, además de eliminar la necesidad de este paso de adquisición de la superficie, resulta en una superficie ajustada a la información de rayos X, evitando diferencias debidas al uso de otros tipos de radiación para la adquisición de la superficie.

La evaluación preliminar sobre simulaciones a partir de proyecciones reales, cubriendo diferentes escenarios de datos limitados, muestra la viabilidad de la propuesta. En todos los casos analizados, la máscara obtenida se corresponde con la ideal obtenida por umbralización a partir del estudio de alta dosis. El método propuesto reduce el RMSE en 41.36% para el caso 130 de rango angular y 52.12%, para el caso 140 de rango angular con respecto a la reconstrucción FDK con datos limitados. En ambos estudios, la máscara obtenida por umbralización desde la reconstrucción con datos limitados, lleva a un peor resultado con respecto al método propuesto: el método propuesto reduce el RMSE en un 5.12% para el caso de 130 grados de rango angular y en 3.77% para el caso de 140 grados de rango angular.

Sin embargo, el uso de métodos híbridos como el propuesto, que incluye un paso de reconstrucción iterativa, comporta tiempos de ejecución altos.Como trabajo futuro, se estudiará una metodología que elimine el paso de reconstrucción iterativa, obteniendo de la red directamente la imagen procesada. Para ello, será necesario resolver los problemas de resolución espacial y optimizar la fase de entrenamiento para evitar el sobreentrenamiento.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia, Innovación y Universidades (Instituto de Salud Carlos III, proyecto DTS17/00122; Agencia Estatal de Investigación, proyecto DPI2016-79075-R-AEI/FEDER, UE), cofinanciado por Fondos de la Unión Europea (FEDER), “A way of making Europe”. Además, ha sido financiado por el Programa de apoyo a la realización de proyectos interdisciplinares de I+D para jóvenes investigadores de la Universidad Carlos III de Madrid 2019-2020 en el marco del Convenio Plurianual Comunidad de Madrid- Universidad Carlos III de Madrid (proyecto DEEPCT-CM-UC3M) y por CRUE Universidades, CSIC y el Banco Santander (Fondo Supera, proyecto RADCOV19). El CNIC está financiado por el Ministerio de Ciencia, Innovación y Universidades y la fundación PRO-CNIC y es un centro de excelencia Severo Ochoa (SEV-2015-0505).

Referencias

[1] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Amer.*, vol. 10, pp. 612–619, 1984.

[2] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys Med Biol*, vol. 53, pp. 4777–4807, 2008.

[3] E. Y. Sidky, C. M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *Journal of X-ray Science and Technology*, vol. 14, pp. 119-139, 2006.

[4] M. Chang, L. Li, Z. Chen, Y. Xiao, L. Zhang, and G. Wang, "A few-view reweighted sparsity hunting (FRESH) method for CT image reconstruction," *J Xray Sci Technol*, vol. 21, pp. 161-176, 2013.

[5] F. P. J. Abascal, M. Abella, E. Marinetto, J. Pascau, and M. Desco, "A novel prior- and motion-based com-pressed sensing method for small-animal respiratory gated CT," *PLoS One*, vol. 11, p. e0149841, 2016.

[6] H. Zhang, L. Li, K. Qiao, L. Wang, B. Yan, L. Li, *et al.*, "Image prediction for limited-angle tomography via deep learning with convolutional neural network," *arXiv preprint arXiv:1607.08707*, 2016.

[7] Y. Huang, S. Wang, Y. Guan, and A. Maier, "Limited angle tomography for transmission X-ray microscopy using deep learning," *Journal of Synchrotron Radiation*, vol. 27, 2020.

[8] M. Monica Abella, C. de Molina, N. Ballesteros, A. García-Santos, A. Martínez, I. García, *et al.*, "Enabling tomography with low-cost C-arm systems," *Plos One*, vol. 13, p. e0203817, 2018.

[9] M. Buda, A. Saha, and M. A. Mazurowski, "Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm," *Computers in biology and medicine*, vol. 109, pp. 218-225, 2019.

[10] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[11] J. J. Vaquero, S. Redondo, E. Lage, M. Abella, A. Sisniega, G. Tapias, *et al.*, "Assessment of a New High-Performance Small- Animal X-ray Tomograph," *IEEE Trans Nucl Sci*, vol. 55, pp. 898-905, 2008.

[12] M. Abella, J. J. Vaquero, A. Sisniega, J. Pascau, A. Udías, V. García, *et al.*, "Software architecture for multi-bed FDK-based reconstruction in X-ray CT scanners," *Computer methods and programs in biomedicine*, vol. 107, pp. 218-232, 2012.

[13] M. Abella, E. Serrano, J. Garcia-Blas, I. García, C. De Molina, J. Carretero, *et al.*, "FUX-Sim: Implementation of a fast universal simulation/reconstruction framework for X-ray systems," *PloS one*, vol. 12, p. e0180363, 2017.

Implementation of a neural-network-based Surgical Workflow Analysis algorithm for surgical educational environments

C. Jiménez-Torres[1], C. Guzmán-García[1], P. Sánchez-González[1, 2], I. Oropesa[1], E.J. Gómez[1, 2].

[1] Centro de Ingeniería Biomédica y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España, catherine.jimenezt@alumnos.upm.es, {c.guzman, p.sanchez, i.oropesa, e.gomez}@upm.es

[2] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Zaragoza, España

Abstract

The automatic analysis of surgical workflow plays an important role in the modeling of surgical processes. Although current automatic approaches to recognize surgical procedures' phases achieve accuracies varying from 0.8 and 0.9, they do not usually analyze their performance per phase, which reports relevant data on their strong and weak points. In this study, an innovative algorithm is implemented to classify the different phases of a laparoscopic cholecystectomy, following an approach based on neural networks and probabilistic methods, and analyzing the model performance for each phase. The best results were achieved using a pre-trained convolutional neural network, VGG-16, in combination with a Hidden Markov Model (accuracy = 0.91). The potential of this algorithm may further be explored for the definition of training and assessment systems of surgeons' cognitive and decision-making skills.

1. Motivation

Minimally invasive surgery (MIS) has evolved as the gold standard for highly accurate, sensitive, and less invasive surgical procedures [1]. These procedures have gained clinical acceptance due to their advantages as compared to open surgery, such as less pain, less scarring, less damage to healthy tissue and a faster recovery [1]. Although MIS can indeed be beneficial, surgical skills for this kind of surgeries are generally difficult to acquire [2].

The concept of surgical expertise can be very complex to define, being characterized by different types of skills, ranging from technical (e.g. dexterity, depth perception, efficiency, autonomy, etc.); to nontechnical (e.g. stress management, decision-making, etc.) and cognitive skills (e.g. conceptual and procedural knowledge) [2]. Thus, the appropriate training of surgeons is important and necessary for the safe and accurate performance of laparoscopic procedures.

Being able to recognize the phases of a surgical procedure is an important component to develop context-aware systems to monitor surgical processes and enhance coordination among surgical teams. Phase recognition is one of the most relevant aspects of surgical workflow analysis (SWA).

SWA can be defined as the set of techniques aimed at detecting the main phases or actions of a surgical procedure and is used for building surgical process models [3]. As a result, these process' models can be used in surgical educational environments and contribute to the training of surgeons and the evaluation of their skills. The algorithms related to SWA using laparoscopic videos found in the literature are usually based on supervised training [4]. The use of videos for SWA is a complex task since laparoscopic videos can highly vary between patients, surgeons, and hospital environments [4].

Several types of features have been used in the literature to recognize surgical phases such as visual cues and surgical instrument tracking [5,6]. Apart from feature selection, the most used approaches to this day for SWA are neural networks and Hidden Markov Models (HMM) [7-9]. The best results obtain accuracies varying from 0.8 to 0.9 approximately [4,8,10,11]. However, in current studies, an individual per-phase analysis is rarely performed [5,11].

The main goal of this study is to obtain an algorithm capable of classifying laparoscopic surgery phases using visual information from recorded procedures. In addition, we aim to analyze the performance of the model for each different phase, taking into consideration the particularities of each of them Specifically, the procedure selected for this work is the laparoscopic cholecystectomy, which consists on the removal of the gallbladder and was selected since it is a widespread and convenient procedure for SWA [5,12], which facilitates comparisons. The implemented algorithm is based on neural networks and HMM approaches, due to their extensively high reported performance. The final purpose of this study is oriented towards contributing to the training and evaluation of surgeons' cognitive and decision-making skills.

2. Materials and methods

The database used for the classification task was extracted from the sub-challenge "Surgical Workflow and Skill Analysis" of the Endoscopic Vision Challenge [13]. The dataset consists of endoscopic videos obtained during laparoscopic cholecystectomy procedures at the University Hospital of Heidelberg and its affiliate hospitals. Surgical phases for all interventions were annotated frame-wise by surgical experts. Annotated phases are shown in Figure 1.

ISBN: 978-84-09-25491-0, pags. 210- 213

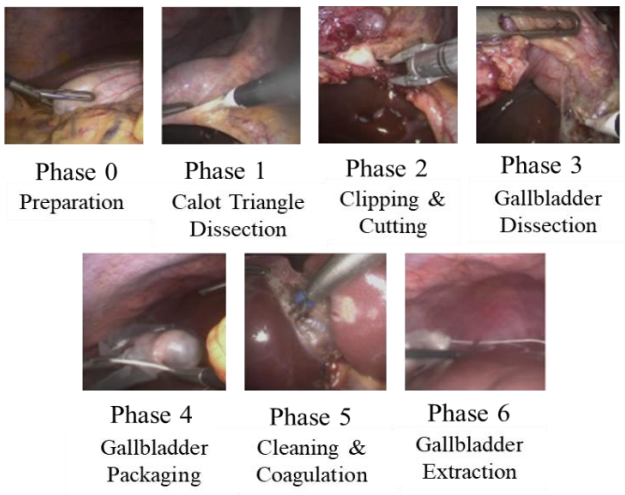


Figure 1. *Laparoscopic cholecystectomy surgical phases*

2.1. Data pre-processing

First of all, the videos of the database were downsized and resampled to cope with the limited hardware resources. To ensure the resampled videos kept the same phase distribution as the originals, the number of frames per second was analyzed in order to set a specific sampling frequency for each video.

Since the distribution of classes (phases) of the original videos was imbalanced, an adjustment to the samples had to be undertaken to obtain an even distribution of the phases and allow for a better training of the neural networks. The majority classes (phases 1 and 3) were randomly under-sampled and the minority classes (phases 0, 2, 4, 5 and 6) were randomly oversampled, resulting in a total of 300 frames per phase.

After balancing the classes, the videos were divided in three groups, 65% for training, 15% for holdout validation to use during training, and 20% for testing.

2.2. Model selection

CNNs have demonstrated outstanding accuracy and compelling optimization behaviors on large-scale image recognition tasks. After analyzing the literature's approaches we have decided to apply transfer learning methodologies and studied the different pre-trained CNNs existing in the literature, from which two were selected to evaluate their performance on the phase recognition tasks: ResNet-50 and VGG-16 [9,14]. Transfer learning consists of using the features learned on a specific problem, and refining them to be applied on a new, similar problem [15]. Both networks (i.e. ResNet-50 and VGG-16) were analyzed using ImageNet pre-defined weights and refined weights.

In this study, we have also taken the temporal information into account and combined the models with (1) s recurrent neural network (RNN), specifically a Long Short-Term Memory (LSTM) layer to the pre-trained networks, and (2) a HMM model.

The combination of RNNs with CNNs has gained popularity in surgical phase recognition, due to the fact that it enables better correlation between the spatial features extracted by the CNN and the temporal knowledge acquired by the RNN [16]. LSTM is an appropriate method to classify, process and predict time series given time lags of unknown duration [17]. These units help to preserve the time-related errors that are usually not taken into account by CNNs.

On the other hand, an HMM is a probabilistic system designed to model a sequence as a result of a markovian process that cannot be observed [18]. HMMs are especially useful when modelling a process in which there is not a direct knowledge about the state in which the system is, such as the case of using visual information from laparoscopic videos to classify phase [18].

While the combination of ResNet-50 with HMM has been used by several authors in the literature [10], the combination of VGG-16 with HMM is an approach not previously reported, thus becoming one of the main contributions of this study.

2.3. Model Validation

In order to assess the models, performance was analyzed based on the accuracy and the F1 score, which is the harmonic mean of precision and recall. In addition, confusion matrices were used with the aim of analyzing the behavior of the models for each different phase.

3. Results and discussion

The learning curves based on the F1-score of both pre-trained CNNs are shown in Figure 2.

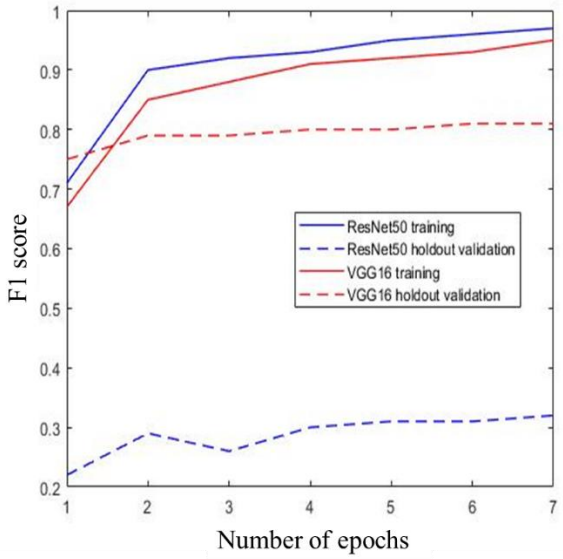


Figure 2. *Learning curve for the training and holdout validation dataset of ResNet-50 and VGG-16. An epoch is a training attempt.*

As observed, ResNet-50's behavior during training and holdout validation suggests that this model has bias problems and performs poorly when adjusting to new data (i.e. training data visibly increases to almost an F1 score of 1 whilst the holdout validation data used during the training process hardly increases, obtaining a low score).

On the other hand, when using VGG-16 neural network, the learning curve for holdout validation data appears to be

closer to state-of-the-art results. In the light of these results, and seeing VGG-16 has a better performance as compared to ResNet-50, the former was selected for the rest of the analysis.

Next, and to analyze how transfer learning methodologies affected the training, the F1 scores obtained using ImageNet predefined weights and refined weights are shown in Table 1. As expected, training with the refined weights seemed to yield better results, because the model adapts better to the specific type of data used for the training [15].

	ImageNet weights	Refined weights
Training data	0.95	0.96
Test data	0.77	0.83

Table 1. *F1 score for VGG-16 models using training and test data.*

Although using refined weights increases the accuracy of test data, these results could still be improved, regarding state-of-the-art models. For this reason, the next step was to evaluate the performance of the additional components of the VGG-16 network using refined weights with the objective of introducing the temporal information that is not extracted when using CNNs and selecting the best approach for the final model. To do that, the F1 scores when adding LSTM and HMM were calculated (Table 2). As observed in Table 2, applying HMM yields considerable improvements as compared to the F1 score obtained with the inclusion of LSTM. This suggests that HMMs are able to capture the temporal discriminative features of our database better than LSTM.

	LSTM		HMM	
	F1 score	Accuracy	F1 score	Accuracy
Training data	0.92	0.96	0.99	0.99
Test data	0.70	0.78	0.92	0.91

Table 2. *F1 score for VGG-16 models using training and test data.*

The confusion matrices of the VGG-16 coupled to LSTM units (Figure 3) and HMMs (Figure 4) suggest that the best predictions are obtained in phase 1.

In the specific case of LSTM, phases 0, 3 and 5 present satisfactory results; however, the model cannot generalize for the rest of phases. Additionally, phase 2 tends to be confused with phase 1, which might be due to the fact that those phases happen in the same working area, which means the same tissues and organs appear in the image. Moreover, another reason could be that test data for these phases may not be representative enough, after performing oversamplig and producing sintetic data.

Contrary to the low classification performance per class of the LSTM model for certain phases, the HMM seemed to generalize condsiderably better for all phases. As can be observed in Figure 4, phase 2 is confused with phase 1, but in this case with a considerably lower error. Both phases, as mentioned above, are performed in the same working area. A possible solution to improve the accuracy for this phase could be introducing instrument information, from feature extraction, since the instruments used for each phase are different (i.e. clipper for phase 2 and dissector for phase 1). Phase 0 is also confused with some other phases, being this phase the beginning of the procedures. This might be due to the high variability between procedures regarding phase 0 (i.e. the endoscope may be calibrated differently and the ports may not be made with exactly the same material, which makes them appear different onscreen). Counting with a higher number of frames for this phase may contribute to have a better representation of phase 0.

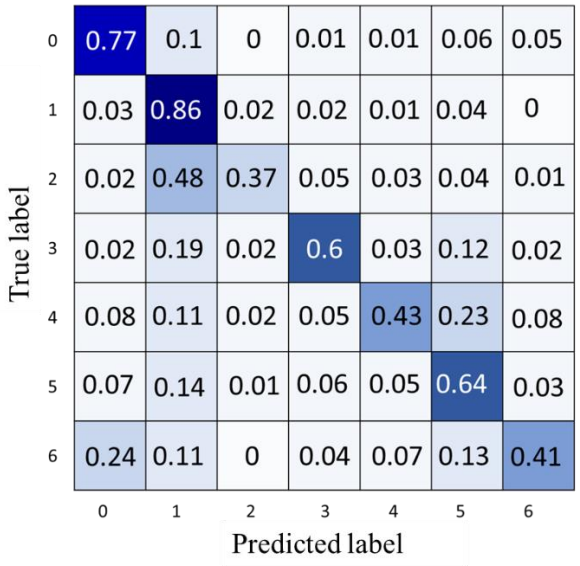


Figure 3. *Confusion matrix for test data, using VGG-16 trained with refined weights and adding LSTM.*

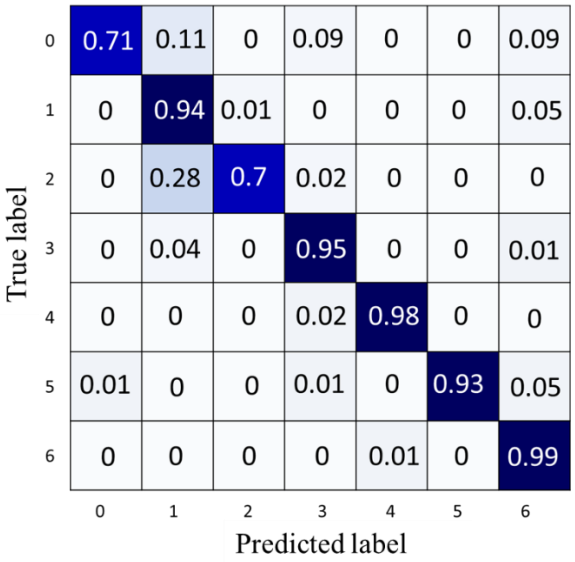


Figure 4. *Confusion matrix for test data, using VGG-16 trained with refined weights and adding HMM.*

Results reveal that the most adequate system is formed by the VGG-16 network trained using refined weights and coupled to an HMM. This model has an accuracy for test data of 0.91, which is indeed a state-of-the-art result. The accuracy would be even higher should the hardware limitations not have taken place (i.e. using full-resolution images). Additionally, the study of the confusion matrices provided further knowledge of which phases are easily recognised and which ones are not.

4. Conclusions

SWA can highly contribute to the characterization of surgical procedures. As a part of SWA, phase recognition plays a crucial role for this task. However, despite existing efforts, current solutions are not satisfactory enough, which led to the purpose of this work.

Accomplishing the main goal of this study, an algorithm for the classification of laparoscopic cholecystectomy procedures was designed. This innovative algorithm consisted on the combination of VGG-16 and HMM, reaching a state-of-the-art accuracy (91%). Moreover, we have conducted an analysis of the algorithm' performance for each phase. This has allowed us to identify which phases were better recognized and which ones were not, arriving to possible reasons of this behavior and also possible solutions. This exhaustive phase analysis may help set future lines in order to improve the recognition, such as introducing instrument information that could contribute to a better description and representation of each phase.

Although the algorithm proposed in this study has been focused on the recognition of phases for laparoscopic cholecystectomy, it can easily be adapted to any other surgical procedure. The main challenge when applying the algorithm to other procedures may be the number of phases to differentiate (i.e. a different number of phases may difficult the extraction of unique characteristics from each of them).

The implemented model can contribute to the training and assessment of surgical skills, helping to evaluate the progress of surgeons' cognitive and decision-making skills. Specifically, it can be useful to assess how residents follow the procedures and their behavior when encountering surgical complications. Thus, this study can be said to represent an attractive approach for further use in the assessment of non-technical surgical skills.

References

[1] J. Verdura, E. Maureen, S. Ek, and M. P. Callery, "Systems, methods, and instruments for minimally invasive surgery," *Relat. US Appl. Data*, vol. 09, no. 186, p. 722, 2000, doi: https://patents.google.com/patent/US6165184.

[2] M. Uemura, P. Jannin, and M. Yamashita, "Procedural surgical skill assessment in laparoscopic training environments," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 11, no. 4, pp. 543–552, 2016, doi: 10.1007/s11548-015-1274-2.

[3] C. Loukas, "Video content analysis of surgical procedures," *Surgical Endoscopy*. 2018, doi: 10.1007/s00464-017-5878-1.

[4] C. Lea, J. H. Choi, A. Reiter, and G. D. Hager, "Surgical Phase Recognition: from Instrumented ORs to Hospitals Around the World," *M2CAI - Satell. Workshop Int. Conf. Med. Image Comput. Comput. Assist. Interv. MICCAI*, 2016.

[5] O. Dergachyova, D. Bouget, A. Huaulmé, X. Morandi, and P. Jannin, "Automatic data-driven real-time segmentation and recognition of surgical workflow," *Int. J. Comput. Assist. Radiol. Surg.*, 2016, doi: 10.1007/s11548-016-1371-x.

[6] O. Weede, F. Dittrich, H. Worn, and H. Feussner, "Workflow analysis and surgical phase recognition in minimally invasive surgery," *2012 IEEE Int. Conf. Robot. Biomim. ROBIO 2012 - Conf. Dig.*, vol. 3, no. 12, pp. 1068–1074, 2012, doi: 10.1109/ROBIO.2012.6491111.

[7] V. E. Balas, S. S. Roy, D. Sharma, and P. Samui, *Handbook of deep learning applications*. Springer, 2019.

[8] A. P. Twinanda, S. Shehata, D. Mutter, J. Marescaux, M. De Mathelin, and N. Padoy, "EndoNet: A Deep Architecture for Recognition Tasks on Laparoscopic Videos," *IEEE Trans. Med. Imaging*, 2017, doi: 10.1109/TMI.2016.2593957.

[9] N. Padoy, "Workflow and Activity Modeling for Monitoring Surgical Procedures," Université Henri Poincaré - Nancy, 2010.

[10] Y. Jin *et al.*, "SV-RCNet: Workflow recognition from surgical videos using recurrent convolutional network," *IEEE Trans. Med. Imaging*, 2018, doi: 10.1109/TMI.2017.2787657.

[11] Y. Jin, Q. Dou, H. Chen, L. Yu, and P. A. Heng, "EndoRCN: Recurrent Convolutional Networks for Recognition of Surgical Workflow in Cholecystectomy Procedure Video," Hong Kong, 2016.

[12] I. M. Macintyre and R. G. Wilson, "Laparoscopic cholecystectomy," *Br. J. Surg.*, vol. 80, no. 5, pp. 552–559, 1993.

[13] S. Speidel, "EndoVis - Grand Challenge," 2015. http://www.endovis-grand-challenge.org.

[14] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015.

[15] L. Torrey, "Transfer learning," in *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods and Techniques*, 2010, p. 23.

[16] G. Yengera, D. Mutter, J. Marescaux, and N. Padoy, "Less is More: Surgical Phase Recognition with Less Annotations through Self-Supervised Pre-training of CNN-LSTM Networks," 2018.

[17] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. 2016.

[18] M. Franzese and A. Iuliano, "Hidden markov models," in *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, 2018.

Desarrollo de un Índice de Privación y un Algoritmo de Aprendizaje Automático para la Identificación de Pacientes Vulnerables

S. Sorbet Santiago[1], J. Díaz Carnicero[1], S. González de Julián[1], M.P. Botija Yagüe[2], R. Usó Talamantes[2]

[1] Centro de Investigación en Economía y Gestión de la Salud, Universitat Politècnica de València, Valencia, España, sorbetsantiago@gmail.com

[2] Departamento de Salud Clínico-Malvarrosa, Consellería de Sanitat Universal i Salut Pública, Valencia, España

Resumen

Además de las influencias biológicas, la salud se ve afectada por determinantes sociales. Esto ha incentivado en los últimos años el desarrollo de metodologías para identificar a las personas más vulnerables a partir de factores socioeconómicos, resultando de especial interés para la gestión de recursos sanitarios.

El primer objetivo de este trabajo consiste en desarrollar un índice de privación construido con variables socioeconómicas que permita clasificar las secciones censales del área geográfica de un departamento de salud de la Comunidad Valenciana. El segundo radica en elaborar un indicador de vulnerabilidad para clasificar individualmente a cada paciente utilizando marcadores socioeconómicos y clínicos.

La construcción del índice de privación por sección censal se llevó a cabo mediante un análisis multivariante y de componentes principales. Para la identificación de pacientes vulnerables, primero un comité de expertos clasificó una lista de potenciales perfiles de vulnerabilidad, que sirvió para entrenar tres algoritmos de clasificación supervisada: Regresión Logística, Random Forest y Red Neuronal, comparados mediante diferentes métricas evaluadoras.

Existen evidencias de una reducción en los niveles de salud y un mayor consumo de recursos sanitarios en aquellas áreas marcadas con mayor privación. En cuanto a los algoritmos probados para la detección de pacientes vulnerables, el Random Forest es con el que mejores resultados se ha obtenido, con valores superiores al 99% en los valores de sensibilidad y especificidad. Los índices de privación y vulnerabilidad obtenidos utilizando herramientas de análisis de Big Data y Machine Learning pueden ayudar a la gestión sanitaria de aquellos pacientes más vulnerables.

1. Introducción

La precariedad laboral, los procesos de exclusión social o la soledad entre las personas más mayores, son algunos de los motivos que se han demostrado como influyentes en la salud de aquellas personas que los sufren [1]. A estas diferencias en salud generadas por causas injustas y evitables se las denomina desigualdades en salud [2].

Pese a que, tradicionalmente, las causas de la enfermedad o de la salud se relegaban a un ámbito de estudio de las características biológicas del individuo, cada vez existen más evidencias que señalan que la biología podría ocupar un plano secundario [3]. En su lugar, la información contextual del individuo está empezando a ganar relevancia en el estudio de la salud. Pese a que estos grupos presentan, en general, una mayor necesidad de acceso a recursos sociales, se presenta la paradoja de que suelen ser el grupo que más dificultades tiene para disfrutarlos [4].

Esta situación de desigualdad resulta evidente en el ámbito sanitario, donde cada vez surgen más estrategias para compensar el efecto sobre la salud de ciertos factores contextuales del individuo. El foco de actuación actualmente se encuentra en Atención Primaria, ya que reúne características que la hacen única, como permitir el acceso a información relevante para generar una planificación adecuada de los recursos o a las herramientas necesarias para implantar planes de acción que tengan repercusiones reales sobre la ciudadanía [5]. Para poder llevar a cabo esta planificación de reparto de recursos o la predicción de prevalencias de ciertas patologías es indispensable el desarrollo de herramientas que tengan en cuenta la información social, económica, ambiental y biológica. No obstante, debido a que todavía es un área de estudio, estas herramientas no están implantadas en la clínica.

2. Metodología

Para este estudio se dispone de una base de datos anonimizada procedente de los sistemas de información de un departamento de salud de la Comunidad Valenciana, para el año 2015. Se incluye la información de todos los pacientes pertenecientes a los cupos de atención primaria de dicho departamento (320.956). En total existen 19 variables, que se han diferenciado en los siguientes tipos:

- Demográficas: edad (grupos etarios), sexo.
- Clínicas y de consumo de recursos sanitarios: índice de morbilidad, costes de pruebas de laboratorio, de farmacia ambulatoria, de pruebas de imagen y totales, número de urgencias hospitalarias, de visitas a atención primaria, de ingresos hospitalarios y de hospitalizaciones evitables.
- Socioeconómicas: zona básica de salud, nacionalidad, coordenadas de residencia, actividad laboral, conjunto geopolítico de origen, composición familiar, tamaño familiar y nivel de renta.

2.1. Índice de privación material

Para el desarrollo del índice de privación material, en primer lugar, se procedió a la delimitación de las secciones

ISBN: 978-84-09-25491-0, pags. 214- 217

censales. Las coordenadas de las secciones censales de las zonas urbanas, pertenecientes al ayuntamiento de Valencia, fueron extraídas de un fichero disponible en la web del propio ayuntamiento [6]. Las coordenadas de las secciones censales de las zonas rurales no se encontraban disponibles en ningún fichero abierto, por lo que hubo que delimitarlas manualmente a partir del visor de mapas del Censo de 2011 disponible en la página web del INE [7].

Una vez se consiguió delimitar cada una de las secciones censales del departamento, se procedió a la asignación de cada persona a su correspondiente sección censal, en función de sus coordenadas individuales de la base de datos de estudio.

El siguiente paso radicó en convertir cada una de las variables en un indicador. Para ello, se seleccionó, para cada variable socioeconómica, una o varias categorías de interés y se calculó el porcentaje de personas que presentaban dicha categoría por cada sección censal. Posteriormente, se calcularon las tasas por edad y sexo, utilizando los grupos de edad: < 15 años, 15 a 44 años, 45 a 64 años, 65 a 74 años, > 74 años. Como medida de desigualdad en salud se utilizaron los indicadores de las variables clínicas y de recursos sanitarios.

A continuación, se realizó un análisis descriptivo y estadístico sobre dichos indicadores para evitar colinealidades entre ellos y para corregir aquellos con altas desviaciones de la normalidad.

Se calculó la PCA (en inglés, *Principal Component Analysis*) sobre los indicadores socioeconómicos y se seleccionaron aquellos con mayor puntuación en la primera componente. Tras esto, se calculó la PCA sobre el subconjunto seleccionado, empleando las coordenadas transformadas asociadas a la nueva primera componente como índice de privación. Para lograr una mayor interpretabilidad, se dividió el histograma de este índice en 7 partes, obteniendo los que se denominaron septiles de privación, donde, a mayor valor, mayor privación material presentaba esa zona.

Por último, se estudió la relación entre el índice de privación y los indicadores de morbilidad y de recursos sanitarios y se estudió si la correlación había mejorado al agregar los indicadores en el índice de privación con respecto a los indicadores socioeconómicos individuales.

En el caso del índice de privación y la morbilidad se llevó a cabo un test estadístico Chi-cuadrado para determinar si existían diferencias significativas de morbilidad entre los distintos septiles de privación. Para el estudio de la relación entre el índice de privación y los indicadores de recursos sanitarios se llevó a cabo un test de Kruskal-Wallis para determinar si existían diferencias significativas de consumo de recursos sanitarios entre los distintos septiles de privación.

2.2. Algoritmo de identificación de pacientes vulnerables

Para esta segunda parte, lo primero que se necesitó fue la identificación de los distintos perfiles de vulnerabilidad que contenía la base de datos inicial, donde se tuvieron en cuenta las variables cualitativas, incluyendo para cada paciente tanto el septil al que pertenecía en base a su situación de residencia, como las variables clínicas. Para ello, se identificaron los distintos tipos de pacientes existentes; es decir, se hallaron todas las combinaciones de características distintas existentes en la base de datos. A las distintas combinaciones existentes se les llamó perfiles.

Para realizar una preselección de perfiles a enviar al comité de expertos, debido al elevado número de combinaciones existentes en la base de datos, se decidió detectar de forma automática a dichos pacientes utilizando la entropía de Shannon, calculándose mediante las frecuencias relativas de cada categoría de cada variable de la base de datos.

Para cada tramo de edad (< 15 años, 15 a 44 años, 45 a 64 años, 65 a 74 años, > 74 años), se tomaron únicamente aquellos perfiles que presentaban una menor entropía. Como criterio de inclusión para esta selección, para cada rango de edad, se seleccionaron todos aquellos pacientes con entropía menor del percentil 50. La menor entropía se asocia con aquellos perfiles que más información presentan, lo cual ha parecido de interés para obtener perfiles que resuman la información contenida en la base de datos total. Con estos perfiles se preparó una encuesta para pasar a un comité de expertos que clasificaría cada paciente como vulnerable o no vulnerable. Una vez se había marcado toda la base de datos, se procedió a analizar tres métodos de clasificación supervisados distintos.

El primero de ellos fue *Random Forest*, con 100 árboles de decisión, con el método RUSBoost. Este método es especialmente bueno cuando el desbalance entre las dos categorías de clasificación es muy grande. El segundo fue una Regresión Logística. Por último, la Red Neuronal, algoritmo ADAM, que contenía un máximo de interacciones de 500, con una sola capa oculta, con un número de neuronas en la capa oculta igual a 8. La función de activación utilizada fue una sigmoide. Debido al número elevado de datos con los que se contaba, se decidió dividir la base de datos en un 70% entrenamiento, 30% test, manteniendo las proporciones de las categorías de las variables.

Se evaluaron los modelos utilizando los siguientes evaluadores: exactitud, precisión, sensibilidad, especificidad, estadístico-F, coeficiente Phi de correlación de Matthews (MCC) y área bajo la curva ROC (AUROC).

Una vez seleccionado el método que mejor resultados reportaba, se trató de hacer una reducción de variables que participarían en el modelo final. Para seleccionar el número de variables adecuado, se fueron eliminando una por una las variables del modelo de clasificación, de menos importante a más importante y se analizó su influencia en las métricas de evaluación de los modelos.

3. Resultados

3.1. Índice de privación material

Analizando la literatura de referencia [8], [9], y teniendo en cuenta las variables disponibles en la base de datos, se decidió incluir los indicadores siguientes, calculados para cada una de las 228 secciones censales: extranjeros sin

nacionalidad española, población con renta anual inferior a 18.000€, población con renta anual superior a 100.000€, desempleados, extranjeros procedentes de continentes con bajo PIB, familias monoparentales, mayores de 65 años que viven solos, y familias de más de 4 miembros.

Tras el cálculo de la primera PCA, los indicadores con mayor peso en la primera componente, que fueron seleccionados para el índice final, fueron: extranjeros sin nacionalidad española, población con renta anual inferior a 18.000€, población con renta anual superior a 100.000€, desempleados, extranjeros procedentes de continentes con bajo PIB y mayores de 65 años que viven solos. La primera componente de la segunda PCA calculada sobre estas variables representó un porcentaje de varianza explicado de un 45%.

Tanto para el test Kruskal-Wallis como para el test Chi-Cuadrado se obtuvo un p-valor menor de 0,01. Esto permite rechazar la hipótesis nula de que los datos provienen de la misma distribución. Es decir, existe una estructura asociada a los septiles respecto al consumo de recursos y morbilidad. La correlación entre el indicador de privación y los indicadores de morbilidad y de recursos sanitarios se presenta en la Tabla 1.

Nombre indicador	**Valor de correlación**
Indicador de morbilidad	
Sin morbilidad	-0,455[a]
Multimorbilidad crónica	0,489 [a]
Indicador de recursos sanitarios	
Costes en laboratorio	0,338 [a]
Costes en farmacia	0,444 [a]
Costes totales	0,574 [a]
Nº consultas a urgencias hospitalarias	0,502 [a]
Nº consultas a facultativos en A.P.	0,572 [a]
Nº ingresos hospitalarios	0,233 [a]
Nº hospitalizaciones evitables totales	0,255 [a]

Tabla 1. *Valor de correlación entre el índice de privación y los indicadores de morbilidad y de recursos sanitario. Donde: p<0,01[a]*

3.2. Algoritmo de identificación de pacientes vulnerables

En primer lugar, se obtuvieron todos los perfiles poblacionales distintos, que ascienden a 8.659, introduciendo el septil como una variable, además de las variables socioeconómicas y de las variables clínicas. Tras la selección de los potenciales perfiles de vulnerabilidad utilizando la entropía, y para simplificar los resultados, se eliminaron aquellos perfiles con un tramo de renta superior a los 100.000 € (al considerarlos no vulnerables socioeconómicamente). Con esto se obtuvieron un total de 154 perfiles diferentes. El comité de expertos identificó un total de 10.452 pacientes como vulnerables, lo cual supone un 4,54% del total.

Respecto al análisis de los resultados obtenidos por las tres técnicas de clasificación: *Random Forest*, Red Neuronal y Regresión Logística, los resultados obtenidos se evaluaron respecto al *ground-truth* de los expertos, y pueden visualizarse en la Figura 1.

En base a esto, se seleccionó el *Random Forest* como modelo que mejores resultados aportaba a nivel predictivo. Finalmente se decidió que no se descartaría ningún predictor del modelo, ya que descartar tres predictores suponía un decremento considerable en la precisión, aunque el resto de las métricas se mantuvieran estables.

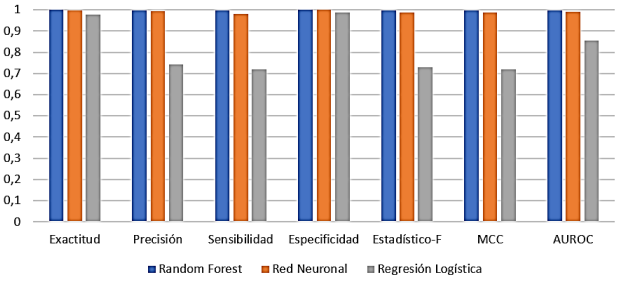


Figura 1: *Gráfico comparativo de las métricas de evaluación entre los tres modelos de clasificación.*

4. Discusión

Los resultados finales del índice de privación fueron validados por el mismo comité de expertos comentado anteriormente. Se puede observar que la varianza explicada es menor al 45%, lo cual es un resultado menor que el de los artículos que se habían tomado como referencia [8], [9]. Así mismo, se debe tener en cuenta a la hora de analizar los resultados que la cantidad de población residente en secciones censales marcadas con los septiles 1, 2 y 7 es muy pequeña en comparación con la que vive en el resto de septiles, lo cual genera un desbalance que afecta a los resultados estadísticos, traduciéndose en unos mayores intervalos de confianza para ambos.

El hecho de agregar los indicadores socioeconómicos en el índice de privación consigue mejorar, en general, el valor de correlación con los indicadores de morbilidad y de recursos sanitarios. Tanto es así, que en todos los casos se obtienen valores estadísticamente significativos de correlación.

Se ha podido observar que, en general, a mayor nivel de privación, mayor consumo de recursos sanitarios. Más en concreto, se podrían apreciar tres agrupaciones básicas: una formada por aquellas secciones censales con un índice de privación de 1, 2 y 3, otra con un índice de privación de 4 y, por último, la formada por aquellas con un índice de privación de 5, 6 y 7. Esto sugiere que existe un gradiente de consumo de recursos en función de las diferencias socioeconómicas, confirmando esta hipótesis de partida. No obstante, hay otras variables en las que no se cumple esta tendencia, como en las hospitalizaciones evitables o el número de ingresos hospitalarios. Igualmente reseñable es el hecho de que se hayan encontrado interacciones entre este índice y el estado de salud, mostrándose una tendencia clara de empeoramiento de salud ante un aumento del índice de privación.

Sobre la interpretación de este índice, resulta especialmente importante evitar caer en una falacia ecológica [10]. De hacerlo, se trasladarían resultados agregados a nivel de sección censal a nivel individual, lo cual supondría graves errores en las posibles interpretaciones del mismo. Se debe entender como la aportación de información sobre el contexto socioeconómico de cada individuo.

En lo referente a la detección de pacientes vulnerables, lo primero a destacar es que, dado el elevado número de combinaciones entre las variables, se debió preseleccionar un conjunto de perfiles para enviar al comité de expertos. Existe la posibilidad de que algún perfil descartado en este primer paso pueda ser también de especial interés, y debiera ser marcado posteriormente.

Otro punto a tener en cuenta es que únicamente un 4,54% del total de la base de datos fue marcada como vulnerable, lo cual presenta un desbalance que puede tener grandes consecuencias en la aplicación del algoritmo. No obstante, al analizar los resultados obtenidos por los métodos de evaluación de clasificación, se observa que todos los resultados son, en general, muy buenos.

Tanto el *Random Forest* como la Red Neuronal presentan excelentes resultados en todas las métricas evaluadoras, siendo ligeramente mejores las del *Random Forest*. Gracias a los valores de sensibilidad y MCC se ve que la Red Neuronal ha cometido más fallos que el *Random Forest* en la clasificación de las personas vulnerables (verdaderos positivos), siendo muy superior el error cometido con la Regresión Logística. La sensibilidad, la precisión y el MCC son de gran relevancia para la evaluación de los modelos probados, ya que se pretende que el valor de verdaderos positivos sea lo más elevado posible. Una posible explicación para unos valores tan altos de todas las métricas de evaluación se basa en cómo el comité de expertos seleccionó qué pacientes eran vulnerables y cuáles no. Los criterios utilizados por ellos podrían ser fácilmente reproducibles por un algoritmo de aprendizaje supervisado.

5. Conclusiones

La información socioeconómica se ha demostrado relevante en el estudio de las desigualdades en salud entre población del mismo grupo de edad. Se ha comprobado que las áreas con un índice de privación mayor tienen asociado un mayor consumo de recursos sanitarios, pudiendo establecer un gradiente de consumo y gasto donde los septiles 1, 2 y 3 presentan valores inferiores, el 4 valores intermedios y el 5, 6 y 7 valores superiores.

Los sistemas de aprendizaje automático permiten la detección de pacientes vulnerables de forma supervisada. En particular, el clasificador *Random Forest* es el que mejores parámetros ha obtenido de clasificación respecto al *ground-truth* marcado por los expertos, demostrando que la variable más relevante para la clasificación, con diferencia, es la actividad laboral. La Red Neuronal, así mismo, obtiene unos valores muy elevados en todas las métricas evaluadoras.

La detección adecuada de pacientes vulnerables puede servir tanto desde un punto de vista de gestión hospitalaria para la gestión de los recursos sanitarios como para llevar a cabo políticas que palíen la situación de este grupo poblacional.

Poder añadir otras variables que indiquen la situación del paciente a lo largo del tiempo y también del espacio, así como información relevante como el nivel de educación, sería una posibilidad de mejora para la predicción de pacientes vulnerables y para la planificación de recursos más a largo plazo.

Referencias

[1] M. Machón, I. Mosquera, I. Larrañaga, U. Martín, and I. Vergara, "Socioeconomic inequalities in health among the elderly population in Spain," *Gac. Sanit.*, vol. 34, no. 3, pp. 276–288, 2020, doi: 10.1016/j.gaceta.2019.06.008.

[2] O. Solar and A. Irwin, "A Conceptual Framework for Action on the Social Determinants of Health," *Soc. Determ. Heal. Discuss. Pap. 2 (Policy Pract.*, p. 79, 2010, doi: ISBN 978 92 4 150085 2.

[3] P. Slade-Sawyer, "Is health determined by genetic code or zip code? Measuring the health of groups and improving population health," *N. C. Med. J.*, vol. 75, no. 6, pp. 394–397, 2014, doi: 10.18043/ncm.75.6.394.

[4] J. Benach, M. Vergara, and Y. C. Muntaner, "Desigualdad en salud: la mayor epidemia del siglo XXI," *Papeles*, vol. 103, pp. 29–40, 2008.

[5] C. Hernández Quevedo and D. Jiménez Rubio, "Las diferencias socioeconómicas en salud entre la población española y extranjera en España: evidencia de la Encuesta Nacional de Salud," *Gac. Sanit.*, vol. 23, no. SUPPL. 1, pp. 47–52, 2009, doi: 10.1016/j.gaceta.2009.07.009.

[6] Ajuntament de València, "División de las secciones censales de la ciudad," 2019. https://www.valencia.es/ayuntamiento/datosabiertos.nsf/resultadoCapas/6FB9A45199E5FEEBC1257C70003E4F80?OpenDocument&lang=1&nivel=2&seccion=1&bdorigen=&idapoyo=22ADF97C1FD223B5C1257C55003BD01F (accessed Aug. 14, 2020).

[7] INE, "Visor de mapas. CENSO 2011," 2011. https://www.ine.es/censos2011/visor/ (accessed Aug. 25, 2020).

[8] M. F. Domínguez-Berjón *et al.*, "Construcción de un índice de privación a partir de datos censales en grandes ciudades españolas (Proyecto MEDEA)," *Gac. Sanit.*, vol. 22, no. 3, pp. 179–187, 2008.

[9] C. Colls, M. Mias, and A. García-altés, "Un índice de privación para reformar el modelo de financiación de la atención primaria en Cataluña," vol. 34, no. 1, pp. 44–50, 2020.

[10] M. C. Schmidtlein, R. C. Deutsch, W. W. Piegorsch, and S. L. Cutter, "A sensitivity analysis of the social vulnerability index," *Risk Anal.*, vol. 28, no. 4, pp. 1099–1114, 2008, doi: 10.1111/j.1539-6924.2008.01072.x.

Detección Automática de Tos para la Monitorización de Pacientes con Enfermedad Pulmonar Obstructiva Crónica

F.Piras[1], A. Lara-Doña[2], M. Morales-González[3], A. León-Jimenez[3], D. Sánchez-Morillo[2]

[1] Control Department, University of Mons, Mons, Belgium, florian.piras@alumni.umons.ac.be

[2] Departamento de Ingeniería Automática, Electrónica y Arquitectura y Redes de Computadores, Grupo de Investigación de Ingeniería Biomédica y Telemedicina, Universidad de Cádiz, Cádiz, España, {alejandro.lara@uca.es, daniel.morillo}@uca.es

[3] Servicio de Neumología, Alergia y Cirugía Torácica, Hospital Universitario Puerta del Mar, Cádiz, España, {moralesgonzalez.maria@gmail.com, antonio.leon.sspa@juntadeandalucia.es }

Resumen

Los sonidos pulmonares han recibido en los últimos años una creciente atención por su potencial para la evaluación de ciertos trastornos respiratorios. Más concretamente, de entre los sonidos adventicios que se manifiestan en el curso de la enfermedad, la tos se ha mostrado como el síntoma más común en algunas enfermedades respiratorias. Sin embargo, y a pesar de su importancia, el proceso convencional de recogida de información para este síntoma ha estado sustentado en el uso de encuestas, con la consiguiente limitación de la subjetividad y el error humano. El objetivo de este trabajo era el desarrollo y prototipado de un sistema automático para la detección de episodios de tos, para su uso en pacientes con EPOC.

Se implementó un sistema de bajo coste basado en la plataforma Raspberry Pi Zero y en el empleo de técnicas de aprendizaje automático aplicadas a la señal registrada por un sensor de aceleración digital colocado en el hueco supraesternal. Contando con la participación de 21 sujetos, se extrajeron características en el dominio tiempo-frecuencia y se evaluaron diversos clasificadores. Los mejores resultados se obtuvieron con un clasificador basado en el ensamblado de árboles de decisión (RUSBoost Tree), para el que se alcanzó una precisión en la detección del 89.8%. Los resultados demuestran que el sistema desarrollado tiene potencial para ser usado con pacientes reales en registros ambulatorios de larga duración.

1. Introducción

Actualmente, las enfermedades respiratorias crónicas son una de las principales causas de mortalidad en el mundo. Desde 2016, la enfermedad pulmonar obstructiva crónica (EPOC) es la tercera causa de muerte en el mundo [1], y se espera que en 15 años pase a ser la primera [2].

La Unión Europea dedica alrededor del 6% de su presupuesto de salud a las enfermedades respiratorias, estando el 56% de esta partida, unos 38.600 millones de euros, asignado a la EPOC. Un factor crítico en la mortalidad, la morbilidad y en los costes atribuibles a la gestión de esta enfermedad están relacionados con las hospitalizaciones, causadas mayoritariamente por episodios recurrentes de agudización de los síntomas denominados exacerbaciones [3].

Debido a la relevancia socioeconómica de la EPOC, en los últimos años han sido reportados en la literatura numerosos estudios sobre telemonitorización domiciliaria de pacientes [4]–[7]. El objetivo principal ha sido, en la mayoría de los casos, el de reducir la tasa de hospitalización a través de la detección automática de las exacerbaciones [7]–[10].

Sin embargo, la identificación temprana de las exacerbaciones se ha convertido en un difícil reto para la comunidad respiratoria, dadas las divergencias en las condiciones basales, causas y manifestaciones clínicas de estos episodios, que presentan una gran variabilidad intra e inter individuo [11].

Diferentes aproximaciones han sido empleadas para la telemonitorización de la EPOC. Las variables monitorizadas con más frecuencia han sido los síntomas diarios reportados por el propio paciente, la saturación arterial de oxígeno (SpO_2), la función pulmonar, la presión sanguínea, la tasa cardiaca, el peso, la temperatura, la tasa respiratoria, los sonidos pulmonares y la actividad física. Estas variables han sido monitorizadas de forma individual aunque con frecuencia se han combinado. Además, el uso de técnicas de aprendizaje automático se ha extendido en los estudios más recientes para proporcionar un enfoque personalizado para la gestión de la EPOC [12].

En particular, los sonidos pulmonares han recibido, en la última década, una creciente atención como marcador con potencial para la evaluación de la enfermedad y para la detección precoz de las exacerbaciones [13]–[15]. Más concretamente, de entre los sonidos adventicios que se manifiestan en el curso de la enfermedad, la tos frecuente se ha mostrado como un signo constante desde el inicio, que empeora de manera lenta pero progresiva a lo largo de los años. Como ejemplo, la tos nocturna ha despertado especial interés y ha sido analizada muy recientemente [16]. La recogida de información relativa a este signo se ha llevado a cabo convencionalmente mediante el uso de encuestas, con la consiguiente limitación de la subjetividad y el error humano. Por ello, la detección automática y la caracterización de los episodios de tos ha sido referida como una estrategia de gran interés que puede mejorar las monitorización de la enfermedad [17].

El objetivo de este trabajo es el desarrollo y prototipado de un sistema automático de bajo coste para la detección y caracterización mediante técnicas de aprendizaje automático de episodios de tos, con el objeto de ser empleado para monitorizar el curso de la enfermedad en pacientes con EPOC.

ISBN: 978-84-09-25491-0, pags. 218- 221

2. Materiales y métodos

2.1. Participantes y metodología

Considerando que se emplearán técnicas de aprendizaje automático supervisado, se hace necesario disponer de datos anotados para en entrenamiento y validación de los modelos. Se planteó como objetivo principal la detección automática de tos diferenciándola de otros eventos acústicos como el habla, la risa, el carraspeo o el ruido de fondo. Con objeto de disponer de muestras etiquetadas de todos estos patrones, se incluyeron 21 sujetos (9 mujeres y 12 hombres), con edades comprendidas entre los 21 y los 50 años, que firmaron un consentimiento informado aceptando su participación en el estudio. Se registró el audio de cada participante en un único fichero empleando una estructura común para todos los sujetos que facilitaba la posterior separación y etiquetado de cada segmento de audio. Cada grabación, realizada con el sujeto en posición sentada, incluyó 10 eventos de tos voluntaria, 5 segundos de habla, dos sonidos de risa, dos de carraspeo y dos pausas de silencio en las que se registró exclusivamente la respiración del sujeto. Cada episodio estuvo separado del siguiente por dos segundos de silencio. La interfaz de usuario fue diseñada para guiar al participante en la simulación de los sonidos, estableciendo pautas temporales en cada caso.

2.2. Hardware

La Figura 1 detalla el sistema de adquisición y almacenamiento de los datos captados por el sensor. Las señales no son procesadas y analizadas en este proceso inicial de registro. Este procesamiento y análisis se realiza posteriormente fuera de línea. Consta de dos partes principales: un sistema embebido y un sensor digital. El proceso de adquisición está controlado desde una interfaz de usuario ejecutada en una tableta con sistema operativo Android.

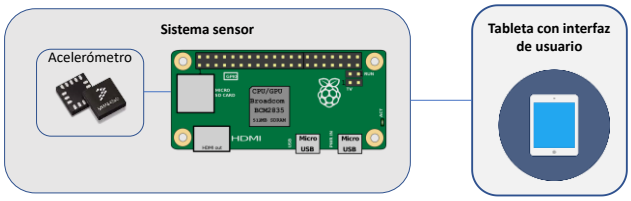


Figura 1.*Arquitectura general del sistema.*

Como microprocesador se seleccionó la unidad Raspberry Pi Zero W (procesador Broadcom BCM2835 con núcleo ARM11 de 1GHz, 512MB de memoria RAM y sistema operativo Linux en memoria microSD), por su bajo consumo, su conectividad inalámbrica y su reducido tamaño. La tarjeta SD permitía el almacenamiento de los sonidos respiratorios registrados.

El sistema fue alimentado mediante una batería recargable LiPo, conectada al microprocesador a través de un convertidor DC/DC cargador. El conjunto fue instalado en una envolvente plástica diseñada ad-hoc, como ilustra la Figura 2. En el estudio realizado por Drugman et al. [18] se comparó la eficacia de distintos sensores para la detección de la tos. Entre ellos, electrodos de electrocardiografía (ECG), termistores, galga extensiométricas, acelerómetros, y micrófonos.

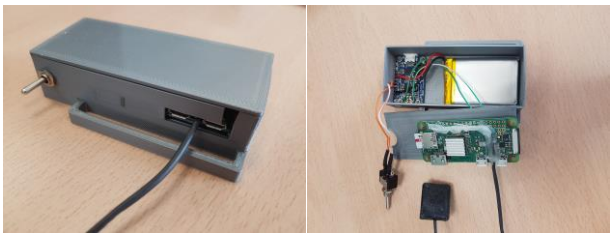


Figura 2.*Detalle del prototipo para el registro de las señales.*

Considerando el ruido de fondo, la espiración forzada, el carraspeo, y el habla y la risa se determinó que pese a que los micrófonos son más resistentes al ruido ambiental en una habitación cerrada, tienen el inconveniente de registrar sonidos procedentes del exterior o de otras personas presentes en la estancia. Por ello, para este estudio se seleccionó como sensor un acelerómetro triaxial digital MMA8452Q que proporciona la mejor relación de compromiso entre la robustez al ruido y la precisión en la detección.

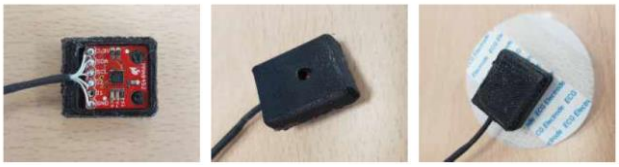


Figura 3.*Sensor de aceleración y sistema de fijación mediante adhesivo reusable.*

La posición del acelerómetro se determinó de acuerdo con las observaciones realizadas [19]. En este estudio, se evaluaron 15 posiciones diferentes en tórax y tráquea con para la detección de la señales de mecanomiografía relacionadas con la respiración, los ronquidos, la tos o los latidos del corazón. Para la tos y los ronquidos, se estimó que la ubicación óptima está alrededor del cartílago tiroideo en la parte delantera del cuello. Se colocó por tanto el sensor en esta posición, y se adhirió a la piel del participante empleando como base un electrodo adhesivo, como ilustra la Figura 3. La conexión entre el sensor y la unidad microprocesadora se realizó mediante el protocolo de comunicaciones I2C. La comunicación entre el sistema de detección y la tableta se realizó vía Bluetooth.

2.3. Adquisición y preprocesado de la señal

Para el acelerómetro, se seleccionó la máxima frecuencia de muestreo disponible para el sensor (800 Hz). La lectura se realizó mediante interrupciones. El proceso de adquisición y registro se controló empleando el lenguaje de programación Python 3. Para eliminar el efecto por la respiración, se aplicó un preprocesado consistente en un filtro FIR (Finite Impulse Response) de paso-banda de orden 10, con frecuencias de corte de 80 y 250Hz.

2.4. Extracción y selección de características

Para superar la limitación dada por la no estacionariedad de la señal de audio, se hizo uso de la transformada de

Fourier enventanada o STFT (Short Time Fourier Transform). Se empleó la ventana de Hanning y se evaluaron diferentes tamaños (300ms, 500ms y 700ms. Para cada ventana, se extrajeron 24 características: energía, frecuencia media, frecuencia mediana, momento de segundo orden y su raíz cuadrada, curtosis, coeficiente de asimetría, entropías de Shannon, Renyi y Tsallis, factor de cresta y coeficiente de aplanamiento espectral. Además, se calcularon 12 coeficientes cepstrales en las frecuencias de Mel (MFCC), frecuentemente usados en el procesado de señales de audio [20]. Para la selección de características se aplicó el método de selección secuencial hacia atrás o SBFS (Sequential Backward Feature Selection), que redujo el conjunto inicial a 12 características.

2.5. Clasificadores, validación y métricas de rendimiento

Se planteó la resolución del problema a través de un clasificador binario capaz de clasificar un evento como tos o como otro tipo de sonido. Se evaluaron varios modelos de clasificadores: máquina de vectores de soporte (SVM) cúbica y cuadrática [21], [22], y ensamblado de árboles de decisión RUS Boost Tree [23]. Para asegurar la consistencia y generalización de cada modelo resultante, se empleó validación cruzada de orden 10. Se realizó un análisis ROC (acrónimo de Receiver Operating Characteristic, o Característica Operativa del Receptor) para evaluar el rendimiento de cada modelo, estimando la sensibilidad, especificidad y precisión en cada caso. El proceso completo se ilustra en la Figura 4.

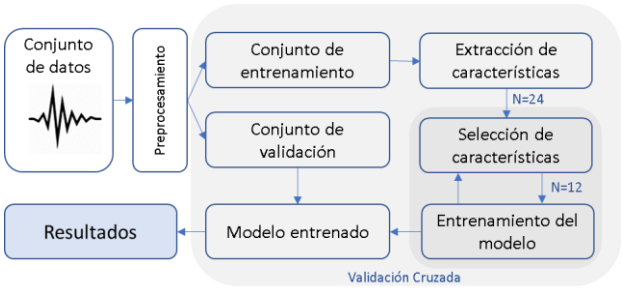


Figura 4.*Extracción, selección de características, entrenamiento y validación del modelo.*

3. Resultados

La Figura 5 representa un ejemplo de la señal registrada para los episodios de tos simulados por uno de los participantes.

Ventana (ms)	*Modelo*	*Pr*	*Se*	*Es*
300	SVM cúbica	83.2	88	78
500	SVM cuadrática	87.0	88	84
700	RUS Boost Tree	89.8	89	91

Tabla 1. *Síntesis de los parámetros de rendimiento para los modelos con mejores resultados para cada tamaño de ventana . Pr: Precisión (%). Se: Sensibilidad (%). Es: Especificidad (%).*

La Tabla 1 sintetiza los mejores resultados obtenidos tras el entrenamiento y validación de los diferentes modelos evaluados para cada uno de los tamaños de ventana considerados. Como se aprecia, la precisión del modelo crece con el tamaño de la ventana temporal seleccionada para la STFT. El mejor resultado se obtuvo para un tamaño de ventana de 700ms y un clasificador basado en el ensamblado de árboles de decisión o RUS Boost Tree. En este caso se alcanzó una precisión del 89.8%.

Empleando este modelo, el 9% de los eventos de tos fue incorrectamente clasificado mientras el 11% de los sonidos de otro tipo fue considerado erróneamente por el sistema como tos.

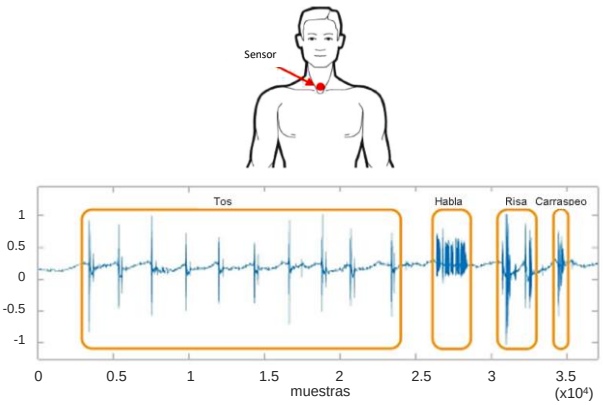


Figura 5.*Señal registrada por el sensor de aceleración para eventos de tos simulados por uno de los sujetos participantes en el estudio.*

4. Discusión y Conclusiones

En este trabajo se ha presentado una aproximación a la clasificación automática de episodios de tos empleando un sistema de bajo coste, basado en el empleo de técnicas de aprendizaje automático aplicadas a la señal registrada por un sensor de aceleración digital ubicado en el hueco supraesternal. El empleo de acelerómetros para la cuantificación de señales biomédicas como la tasa y variabilidad cardiaca, la tasa respiratoria o el ronquido está extendido en la literatura, y presenta para esta aplicación ciertos beneficios sobre los micrófonos convencionales [18].

Los modelos validados se comportan consistentemente bien, con rendimientos altos similares a los alcanzados en otros estudios recientes [16], [24], [25]. Además, desde un punto de vista práctico, el monitor de tos ha sido concebido como un sistema extremadamente simple que puede ser empleado para un uso ambulatorio de 24h.

Un trabajo futuro supondrá la evaluación del ensamblado de diferentes clasificadores y de nuevas características para mejorar la precisión, además del entrenamiento y validación de los modelos empleando datos procedentes de pacientes de edad avanzada con patologías respiratorias crónicas, dado que entre las limitaciones del estudio figura la naturaleza voluntaria de los sonidos registrados y la condición no patológica de los participantes.

Finalmente, la naturaleza portable del hardware embebido (Raspberry Pi Zero) y del software desarrollado y la experiencia adquirida permitirá abordar el problema real de diferenciar, en pacientes de EPOC, entre tos seca y

húmeda, detectar secuencias de tos [26] y analizar las propiedades específicas de cada tipología que resulten de interés para el seguimiento de la enfermedad.

Financiación

Este estudio ha sido financiado con fondos del proyecto "Soporte Inteligente a la Autogestión de la EPOC", cofinanciado por la Consejería de Salud de la Junta de Andalucía, en el marco de la "Iniciativa Territorial Integrada" con número de expediente PI-0006-2017, y con cargo al Fondo Europeo de Desarrollo Regional (FEDER), dentro del Programa Operativo de Andalucía FEDER 2014-2020.

Referencias

[1] J. Lortet-Tieulent, I. Soerjomataram, J. L. López-Campos, et al., "International trends in COPD mortality, 1995-2017," *European Respiratory Journal*, vol. 54, no. 6. European Respiratory Society, 01-Dec-2019.

[2] A. Lwan, "WHO | Global status report on noncommunicable diseases 2010," *Who*, 2015.

[3] S. M. May and J. T. C. Li, "Burden of chronic obstructive pulmonary disease: Healthcare costs and beyond," *Allergy Asthma Proc.*, vol. 36, no. 1, pp. 4–10, Jan. 2015.

[4] C. Kruse, B. Pesek, M. Anderson, K. Brennan, and H. Comfort, "Telemonitoring to manage chronic obstructive pulmonary isease: Systematic literature review," *J. Med. Internet Res.*, vol. 21, no. 3, p. e11496, Mar. 2019.

[5] J. Cruz, D. Brooks, and A. Marques, "Home telemonitoring effectiveness in COPD: A systematic review," *Int. J. Clin. Pract.*, vol. 68, no. 3, pp. 369–378, 2014.

[6] C. L. Bentley *et al.*, "A pilot randomised controlled trial of a Telehealth intervention in patients with chronic obstructive pulmonary disease: challenges of clinician-led data collection," *Trials*, vol. 15, no. 1, p. 313, Jan. 2014.

[7] M. A. Fernandez-Granero, D. Sanchez-Morillo, and A. Leon-Jimenez, "An artificial intelligence approach to early predict symptom-based exacerbations of COPD," *Biotechnol. Biotechnol. Equip.*, vol. 32, no. 3, pp. 778–784, May 2018.

[8] E. Iadanza, V. Mudura, P. Melillo, and M. Gherardelli, "An automatic system supporting clinical decision for chronic obstructive pulmonary disease," *Health Technol. (Berl).*, vol. 10, no. 2, pp. 487–498, Mar. 2020.

[9] C. Wang, X. Chen, L. Du, Q. Zhan, T. Yang, and Z. Fang, "Comparison of machine learning algorithms for the identification of acute exacerbations in chronic obstructive pulmonary disease," *Comput. Methods Programs Biomed.*, vol. 188, p. 105267, May 2020.

[10] L. Jiang and A. S. Gershon, "Using health administrative data to predict chronic obstructive pulmonary disease exacerbations," *Annals of the American Thoracic Society*, vol. 17, no. 9. American Thoracic Society, pp. 1056–1057, 01-Sep-2020.

[11] C. Burton, H. Pinnock, and B. McKinstry, "Changes in telemonitored physiological variables and symptoms prior to exacerbations of chronic obstructive pulmonary disease.," *J. Telemed. Telecare*, vol. 21, no. 1, pp. 29–36, Jan. 2015.

[12] D. Sanchez-Morillo, M. A. Fernandez-Granero, and A. Leon-Jimenez, "Use of predictive algorithms in-home monitoring of chronic obstructive pulmonary disease and asthma: A systematic review," *Chron. Respir. Dis.*, vol. 13, no. 3, pp. 264–283, 2016.

[13] S. Chen *et al.*, "Lung sounds can be used as an indicator for assessing severity of chronic obstructive pulmonary disease at the initial diagnosis," *Nan Fang Yi Ke Da Xue Xue Bao*, vol. 40, no. 2, pp. 177–182, 2020.

[14] N. S. Haider, B. K. Singh, R. Periyasamy, and A. K. Behera, "Respiratory Sound Based Classification of Chronic Obstructive Pulmonary Disease: a Risk Stratification Approach in Machine Learning Paradigm," *J. Med. Syst.*, vol. 43, no. 8, p. 255, Jun. 2019.

[15] M. A. Fernandez-Granero, D. Sanchez-Morillo, and A. Leon-Jimenez, "Computerised analysis of telemonitored respiratory sounds for predicting acute exacerbations of COPD," *Sensors (Switzerland)*, vol. 15, no. 10, pp. 26978–26996, 2015.

[16] S. Vhaduri, "Nocturnal Cough and Snore Detection Using Smartphones in Presence of Multiple Background-Noises," in *COMPASS 2020 - Proceedings of the 2020 3rd ACM SIGCAS Conference on Computing and Sustainable Societies*, 2020, pp. 174–186.

[17] A. H. Morice *et al.*, "ERS guidelines on the assessment of cough.," *Eur. Respir. J.*, vol. 29, no. 6, pp. 1256–76, Jun. 2007.

[18] T. Drugman *et al.*, "Objective study of sensor relevance for automatic cough detection," *IEEE J. Biomed. Heal. Informatics*, vol. 17, no. 3, pp. 699–707, 2013.

[19] D. B. Rendón, J. L. Rojas Ojeda, L. F. Crespo Foix, D. S. Morillo, and M. A. Fernández, "Mapping the human body for vibrations using an accelerometer.," *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2007, pp. 1671–4, 2007.

[20] D. S. Morillo, A. L. Jiménez, and S. A. Moreno, "Computer-aided diagnosis of pneumonia in patients with chronic obstructive pulmonary disease," *J. Am. Med. Informatics Assoc.*, 2013.

[21] U. Jain, K. Nathani, N. Ruban, et al. "Cubic SVM classifier based feature extraction and emotion detection from speech signals," in *Proceedings - 2018 International Conference on Sensor Networks and Signal Processing, SNSP 2018*, 2019, pp. 386–391.

[22] X. Yan, H. Zhu, and J. Luo, "A novel kernel-free nonlinear SVM for semi-supervised classification in disease diagnosis," *J. Comb. Optim.*, pp. 1–18, Nov. 2019.

[23] Z. Abduh, E. A. Nehary, M. Abdel Wahed, and Y. M. Kadah, "Classification of heart sounds using fractional fourier transform based mel-frequency spectral coefficients and traditional classifiers," *Biomed. Signal Process. Control*, vol. 57, p. 101788, Mar. 2020.

[24] A. Leal *et al.*, "A Multi-feature Approach for Noise Detection in Lung Sounds," in *IFMBE Proceedings*, vol. 64, 2019, pp. 211–216.

[25] J. Knocikova, J. Korpas, M. Vrabec, and M. Javorka, "Wavelet analysis of voluntary cough sound in patients with respiratory diseases," *J. Physiol. Pharmacol.*, vol. 59, no. SUPPL. 6, pp. 331–340, 2008.

[26] G. Rudraraju *et al.*, "Cough sound analysis and objective correlation with spirometry and clinical diagnosis," *Informatics Med. Unlocked*, vol. 19, p. 100319, Jan. 2020.

Telemedicina y E-Salud

Miércoles, 25 de noviembre de 2020

Plataforma mHealth ALIBIRD para el apoyo en el tratamiento y seguimiento de pacientes con cáncer

J.M. Iniesta-Chamorro[1], M. Sereno Moyano[2], I. Espinosa-Salinas[3], B. López Plaza[4], M.E. Hernando[1,5], E.J. Gómez[1,5], Consorcio ALIBIRD2020-CM

[1] Grupo de Bioingeniería y Telemedicina, Centro de Tecnología Biomédica, ETSI de Telecomunicación, Universidad Politécnica de Madrid, España, {jm.iniesta, mariaelena.hernando, enriquejavier.gomez}@upm.es

[2] Oncología Médica, Hospital Universitario Infanta Sofía, Madrid, España,{maria.sereno}@salud.madrid.org

[3] IMDEA-Food Institute, CEI UAM + CSIC, Madrid, España, {mariaisabel.espinosa}@imdea.org

[4] Instituto de Investigación Sanitaria del Hospital Universitario La Paz (IdiPAZ), Madrid, España, {bricia.plaza}@idipaz.es

[5] Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, CIBER, Madrid, España

Resumen

Según la OMS, el cáncer es la segunda causa de muerte en el mundo. Alrededor de un tercio de las muertes por cáncer se deben a los cinco principales factores de riesgo conductuales y dietéticos: índice de masa corporal elevado, ingesta reducida de frutas y verduras, falta de actividad física, consumo de tabaco y consumo de alcohol. Pacientes que controlan y adoptan estilos de vida saludables, reportan una mejora en la calidad de vida, menor fatiga y una posible reducción del riesgo de recaída.

ALIBIRD2020-CM es un consorcio multidisciplinar cuyo objetivo principal es aplicar la nutrición de precisión para mejorar el pronóstico de enfermos de cáncer. Este trabajo presenta la plataforma mHealth con la que se pretende empoderar a pacientes y apoyar la toma de decisiones clínicas durante el tratamiento oncológico. Durante la fase de análisis y captura de requisitos se identificaron las necesidades de los usuarios, los componentes de intervención, así como las funcionalidades y características principales de la plataforma. Tras el proceso de desarrollo, se ha obtenido una primera versión de las aplicaciones que conforman el sistema y se procederá a la realización de una validación técnica. Paralelamente, se están definiendo los estudios piloto en los que la plataforma será evaluada con pacientes de dos hospitales.

1. Introducción y objetivo

1.1. Antecedentes y contexto

La vida hace unos siglos era muy diferente a la actual. Desde el siglo XVIII, los cambios experimentados a nivel industrial, científico, económico, político y social conllevaron un aumento en la esperanza de vida. Actualmente ningún país del mundo tiene una esperanza de vida inferior a la del país con mayor esperanza de vida del siglo XIX y la población de la mayoría de los países más desarrollados tiene expectativas de vida que se acercan a los 80 años e incluso la superan [1]. En 2050 se espera que la cantidad de personas mayores de 60 años duplique a la de 2017. Este envejecimiento de la población tendrá un impacto significativo en la sociedad, en los sistemas de salud y económico en las próximas décadas [2].

Un aspecto negativo de esta situación es el aumento de la prevalencia de enfermedades relacionadas con el envejecimiento, como es el caso del cáncer, que es la segunda causa de muerte a nivel mundial y la principal en los países desarrollados. Casi el 90% de las personas que murieron de cáncer en 2017, 9,6 millones en todo el mundo, tenían 50 años o más [1]. En España, el número absoluto de casos de cáncer diagnosticados sigue aumentando anualmente, probablemente por el aumento de la población y por diversos factores de riesgo tales como el consumo de tabaco, alcohol, aumento de los niveles de contaminación, obesidad y sedentarismo entre otros [3]. La OMS estima que un tercio de todos los cánceres serían prevenibles mediante cambios en el estilo de vida y aplicación de estrategias de base científica, abarcando tanto la detección precoz como el tratamiento de los pacientes. Pacientes que tras un diagnóstico de cáncer realizan cambios favorables en su estilo de vida (dieta saludable, actividad física y descanso adecuados) experimentan mejorías en su estado general, disminución notable de la fatiga e incluso una menor incidencia de recurrencia o metástasis [4-5].

El modelo genérico del proceso asistencial del cáncer consta de seis bloques principales: prevención, detección, diagnóstico, tratamiento, supervivencia y cuidados en la fase final de la vida. Pese a los avances tecnológicos, el registro y recopilación de las variables relevantes para el tratamiento se realiza de manera presencial, y entre las visitas presenciales el control es escaso. En 1999 el 'Institute of Medicine' (IOM) realizó un estudio para analizar la calidad de los cuidados de pacientes oncológicos, y se identificaron seis pilares fundamentales en la transformación hacia un modelo asistencial ideal: 1) pacientes empoderados en la autogestión de la enfermedad; 2) equipos interdisciplinares, entrenados y coordinados; 3) sistema sustentado en la evidencia, investigación y estudios clínicos que apoyen la toma de decisiones; 4) sistemas de información adecuados, aplicación de las TIC; 5) transición de las evidencias hacia las prácticas clínicas, verificación de procesos, mediciones de calidad y rendimiento; 6) sistemas de cuidados accesibles y centrados en el paciente.

Este enfoque del IOM está alineado con el concepto de Real World Evidence (RWE), o Evidencia del Mundo Real, definido como el uso de los datos generados por el propio paciente (Patient Reported Outcomes o PROs),

ISBN: 978-84-09-25491-0, pags. 223- 226

recogidos a través de tecnologías y soluciones de "Salud digital", aplicaciones móviles, redes sociales, dispositivos wearables, e incluso por medio de sensores o dispositivos domésticos, para tomar decisiones relativas a procedimientos médicos y apoyando el desarrollo de terapias altamente personalizadas [6]. En la actualidad, existe un gran interés por potenciar la utilización y generación de soluciones y/o servicios de salud digital aplicadas en el tratamiento, seguimiento y, en consecuencia, en dirección a mejorar la calidad de vida de los pacientes oncológicos, así como apoyar a los profesionales clínicos e investigadores relacionados.

ALIBIRD2020-CM es un consorcio multidisciplinar cuyo objetivo es aplicar la nutrición de precisión para mejorar el pronóstico de enfermos de cáncer. Se contempla el desarrollo de una plataforma mHealth que por un lado permita empoderar a los pacientes en la autogestión de la salud y en el cumplimiento de las pautas prescritas, apoyando la adherencia al tratamiento; y por otro lado, apoye al personal sanitario en la toma de decisiones clínicas durante el tratamiento y seguimiento oncológico, teniendo en cuenta diferentes dimensiones del paciente (genética, microbiota, metabolismo, hábitos de vida (dieta, actividad física, sueño), síntomas, medicación, etc.

La validación de la plataforma se realizará en 2021-2022 mediante dos estudios piloto que se desarrollarán en el Hospital Universitario Infanta Sofía (HUIS) y Hospital Universitario La Paz (HULP), de la Comunidad de Madrid, que participan en el consorcio del programa ALIBIRD.

- Estudio Piloto CNMP-HUIS: Análisis de la aplicabilidad de una plataforma eHealth en el manejo y seguimiento de pacientes con cáncer no microcítico de pulmón (CNMP) avanzado en tratamiento activo sobre pacientes con cáncer de pulmón. Seguimiento de 6 meses con 10 visitas presenciales. Entre los objetivos del estudio están: evaluar la experiencia reportada por el paciente, evaluar cambios de la satisfacción y usabilidad, cambios en la calidad de vida, análisis descriptivo y dinámico intra-paciente y cambios en diferentes dimensiones, correlación de variables, etc.
- Estudio Piloto CM-HULP: Evaluación y validación de una herramienta mHealth aplicada en el seguimiento de mujeres supervivientes de cáncer de mama (Estudio MAMANUT). Seguimiento de 3 meses con 2 visitas telemáticas. Entre los objetivos del estudio están: valorar los hábitos de la vida diaria (dieta, actividad física, etc.) tras dos años de diagnóstico, evaluar la usabilidad y satisfacción de los voluntarios.

Este trabajo presenta los avances realizados en el desarrollo de la plataforma mHealth que dará soporte a estos estudios.

1.2. Objetivo

Los objetivos y actividades de esta fase del proyecto son el estudio del estado del arte, la captura de los requisitos, el diseño de la arquitectura, el desarrollo iterativo e incremental de los servicios y/o aplicaciones de usuario, y la integración de los componentes en la plataforma mHealth ALIBIRD aplicable al tratamiento y seguimiento de pacientes con cáncer.

2. Materiales y métodos

2.1. Metodología empleada

El esquema planteado para la intervención y aplicación de la mHealth, mostrado en la Figura 1, está inspirado en una serie de pasos y procesos seguidos en intervenciones previas de la literatura [7]. Durante la ejecución de las tareas y actividades se ha seguido una metodología de desarrollo ágil, caracterizada por un enfoque de co-creación entre los clientes/usuarios de las herramientas y aplicaciones (pacientes-voluntarios, profesionales clínicos e investigadores), desarrolladores y resto de participantes del programa. El proceso iterativo e incremental ha incluido etapas de análisis, diseño, desarrollo/implementación, integración y pruebas técnicas.



Figura 1. *Procesos para la intervención mHealth*

2.2. Diseño, desarrollo e integración de componentes

Análisis y diseño

Inicialmente se realizó un estudio del estado del arte sobre intervenciones de eHealth/mHealth en cáncer, incluyendo experiencias documentadas en la literatura científica, así como los objetivos y resultados parciales obtenidos en otros proyectos internacionales como iManageCancer [8], que plantea una plataforma TIC para empoderar pacientes oncológicos con servicios y "serious games" que apoyen el seguimiento de la enfermedad y la toma de decisiones; o el proyecto eSMART [9] y la aplicación móvil LuCApp [10] enfocadas en la autogestión de síntomas.

En la etapa de diseño se capturaron y analizaron los requisitos para determinar la estructura y arquitectura del sistema, Figura 2, y las características y funcionalidades a desarrollar.

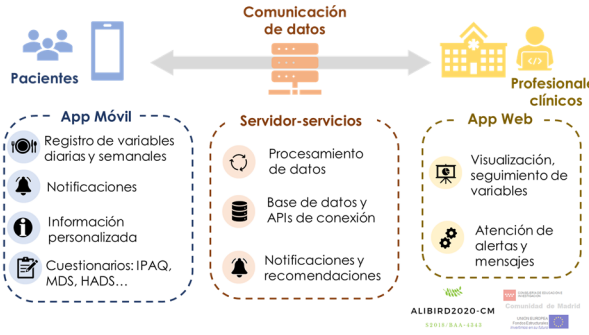


Figura 2*. Esquema de la plataforma mHealth*

Como metodología de diseño, se partió de un enfoque basado en el "Diseño Centrado en el Usuario o UCD (Norma ISO 9241-210:2019)", para incluir estrategias de "Co-design" o diseño participativo. La epidemia COVID-19 ha complicado algunas de las actividades planteadas,

como son las entrevistas individuales, las sesiones de trabajo en grupo o los workshops basados en estrategias como "storyboarding" o "card-sorting". En cualquier caso, ha sido posible realizar gran parte de las acciones con apoyo telemático y abordar la realización de bocetos, mockups y wireframes de las aplicaciones.

Desarrollo del sistema

En esta fase, partiendo de la arquitectura diseñada, se abordó el desarrollo de los componentes de la plataforma.

- Aplicación móvil para el paciente oncológico: incluye escenarios y módulos para el registro de variables clínicas y PROs relativos al estilo de vida de los pacientes (dieta, actividad física, sueño, consumo de tabaco y alcohol) y otros aspectos específicos del tratamiento como la aparición de síntomas, toxicidad y medicación. Algunas de las dimensiones se instrumentan a través de cuestionarios validados (MDS de adherencia a la dieta mediterránea, IPAQ de actividad física, EQ-5D-5L de calidad de vida, LC13 de síntomas, etc.) e integrados en la aplicación. Dentro de la aplicación se incluyó un apartado con recursos informativos sobre la enfermedad, y un asistente del tipo "Chatbot" para resolver dudas generales. Otra característica muy importante son las notificaciones que el paciente puede recibir en la aplicación, dado que no sólo mejoran la calidad del seguimiento y la monitorización, sino que apoyan el empoderamiento del paciente. Se optó por desarrollar una aplicación móvil de tipo híbrida con tecnologías Web (Ionic, HTML, CSS, JavaScript) multiplataforma.
- Aplicación Web para el profesional clínicos: Permite el acceso a la información registrada por cada uno de los pacientes con la aplicación móvil. Cuenta con un módulo de "alertas o alarmas" para enfocar la atención hacia aquellas variables, que ya sea por su condición o nivel de valor adquirido, requieren una acción por parte del profesional. Permite la generación de reportes con la información agrupada por semanas que los profesionales pueden descargar e incorporar en la Historia Clínica Electronica (HCE). Para el desarrollo de esta aplicación del tipo Web PWA o "Progressive Web Apps", también se utilizó IonicFramework a la que los profesionales accederán desde diferentes puestos del servicio de oncología.
- Infraestructura IT: Soporta tanto la comunicación y almacenamiento de datos como el procesamiento y análisis de estos para la generación de notificaciones (recordatorios, refuerzos positivos, apoyo e información, resumen semanal de cumplimiento) dirigidas a la aplicación móvil del paciente y alarmas e indicadores en la aplicación del profesional. La infraestructura IT está compuesta por una serie de servidores/servicios Web y otros componentes tales como APIs o interfaces de comunicación que interactúan con las aplicaciones u otros servicios de terceros. Para la implementación de los servidores Web y servicios se ha utilizado NestJS (NodeJS) y para las bases de datos MongoDB y MySQL.
- Otros servicios y subsistemas de terceros: a) Firebase para para el control de autenticación de usuarios en las aplicaciones, y para el envío de notificaciones tipo "Push" hacia la aplicación móvil; b) Plataforma de datos Fitbit, fabricante de los Wearables o pulseras de actividad física propuestos para que utilicen los pacientes; c) DialogFlow, plataforma de comprensión de lenguaje, utilizada para el desarrollo del asistente conversacional o Chatbot; d) Plataforma de eCRF REDCap [11] o sistema de captura de datos para estudios de investigación, utilizada para el desarrollo de los cuestionarios validados accesibles vía Web y como repositorio de datos.

Integración de los componentes

Para la integración de los distintos componentes se emplearon estos mecanismos: 1) se han desarrollado APIs o interfaces de conexión entre las aplicaciones y subsistemas que interactúan a través de comunicaciones seguras vía Internet (HTTPS); 2) en el caso de módulos internos de las aplicaciones, la comunicación se realiza mediante el acceso a métodos disponibles en librerías o plugins desarrollados específicamente para cada caso; 3) la interacción con servicios de terceros, tales como Firebase, para la autenticación, o la plataforma de Wearables Fitbit, se realizan mediante las APIs y métodos aportados por los fabricantes. Al finalizar la evolución de los aplicativos o sistemas, éstos eran integrados y se comprobaba su comportamiento para verificar que se cumplía con los requisitos especificados originariamente.

3. Resultados

A continuación, se describen los resultados obtenidos, desde la arquitectura del sistema a las aplicaciones desarrolladas.

3.1. Arquitectura del sistema

La Figura 3 muestra los componentes de intervención y las funcionalidades principales de la plataforma mHealth identificados tras la fase de análisis y captura de requisitos.

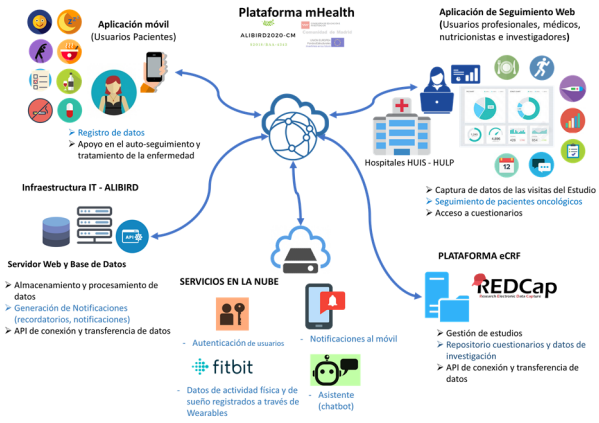


***Figura 3**. Arquitectura de la Plataforma mHealth*

3.2. Aplicación móvil del paciente oncológico

La aplicación móvil para los pacientes oncológicos incluye las funcionalidades para el registro de datos, la visualización del progreso de estos y gestiona la recepción de notificaciones relativas a recomendaciones y recordatorios. Cuenta con apartados de perfil del usuario, escenario con información relativa a la enfermedad y pautas recomendadas (Figura 4). Toda la información

registrada se transmite a través de internet y es procesada en tiempo real por otros componentes de la arquitectura.

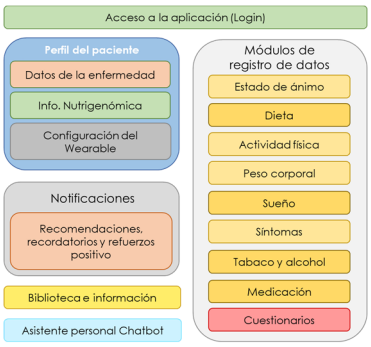


Figura 4. *Escenarios y módulos de la app móvil del paciente*

3.3. Aplicación Web del profesional clínico

La app del profesional incluye los módulos y funciones necesarias para la visualización y el seguimiento del tratamiento de los pacientes, apoyando la toma de decisiones (Figura 5).

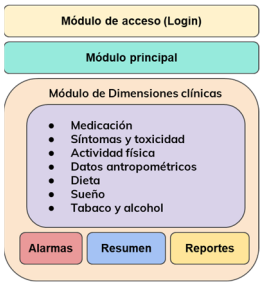


Figura 5. *Escenarios de la app del profesional clínico*

3.4. Infraestructura IT y otros servicios

La infraestructura IT planteada permite la captura de los datos, su procesamiento, y almacenamiento, e implementa la proactividad del sistema para reaccionar ante eventos de forma autónoma. Está compuesta por un servidor Web, servidor de base de datos y varios servicios que cumplen funciones específicas: 1) Servicio principal o "Core": incluye los mecanismos para la comunicación de datos a través de APIs con las aplicaciones e interactúa con otros sistemas/servicios internos y externos como la herramienta de cuestionarios REDCap; 2) Servicio de reglas: consta de una serie de pautas definidas por los profesionales y se encarga de analizar o evaluar las variables clínicas y datos registrados por los pacientes, cuyo resultado puede desencadenar el envío de notificaciones al paciente o alertas al profesional; 3) Servicio de Wearables: contempla los mecanismos necesarios para comunicarse y capturar los datos desde la plataforma del fabricante de los dispositivos de monitorización de actividad física y sueño; 4) Servicio de notificaciones: cuenta con las funciones para lanzar mensajes u recordatorios a la app de los pacientes.

4. Discusión y conclusiones

El cáncer es un gran problema de salud pública y empeora significativamente la calidad de vida de los pacientes, siendo de gran relevancia la adherencia al tratamiento para prevenir complicaciones de la enfermedad. En línea con las actividades del programa ALIBIRD, se han llevado a cabo las tareas de análisis, captura y especificación de requisitos, definición de escenarios e iniciado el desarrollo de los componentes del sistema. Se pretende que la aplicación móvil destinada al paciente oncológico, apoye su empoderamiento en la autogestión de la salud y calidad de vida. Por su parte, los profesionales clínicos vean facilitada su labor en el tratamiento y seguimiento de pacientes. La plataforma y sus componentes se encuentran en una fase avanzada de desarrollo, y en los próximos meses se evolucionarán la mayoría de los módulos para posteriormente realizar validación técnica, previa a la evaluación en los estudios piloto en los hospitales del consorcio. Teniendo en cuenta los alcances y las funcionalidades de la plataforma mHealth ALIBIRD, ésta se configura como una potente herramienta de empoderamiento del paciente oncológico y de apoyo durante el tratamiento y seguimiento de la enfermedad.

Agradecimientos

Al consorcio ALIBIRD2020-CM(S2018/BAA-4343): "Fórmulas terapéuticas de nutrición de precisión para el cáncer", financiado por la DG de Investigación e Innovación de la Comunidad de Madrid (DGII) y cofinanciado con Fondos Estructurales de la UE.

Referencias

[1] M. Roser et al., "Life Expectancy and Cancer," Our World in Data, 2019. https://ourworldindata.org (Consultada: Mayo 2020)

[2] United Nations, Department of Economic and Social Affairs, "World Population Prospects: The 2017 Revision," 2017. https://www.un.org/development/desa/publications/world-population-prospects-the-2017-revision.html

[3] Sociedad Española de Oncología Médica (SEOM). "Las cifras del cáncer en España 2020," 2020. https://seom.org/seomcms/images/stories/recursos/Cifras_del_cancer_2020.pdf

[4] J. Ligibel, "Lifestyle Factors in Cancer Survivorship," J. Clin. Oncol., vol. 30, no. 30, pp. 3697–3704, 2019, doi: 10.1200/JCO.2012.42.0638.

[5] A. I. Phipps et al., "Pre-diagnostic sleep duration and sleep quality in relation to subsequent cancer survival," J. Clin. Sleep Med., vol. 12, no. 4, pp. 495–503, 2016, doi: 10.5664/jcsm.5674.

[6] The Professional Society for Health Economics and Outcomes Research, "Real-World Evidence," https://www.ispor.org/strategic-initiatives/real-world-evidence (Consultada: Mayo 2020)

[7] L. Pfaeffli et al., "A mHealth cardiac rehabilitation exercise intervention: findings from content development studies," BMC Cardiovasc. Disord., vol. 12, no. 1, pp. 1–9, 2012.

[8] H. Kondylakis et al., "iManageCancer: Developing a Platform for Empowering Patients and Strengthening Self-Management in Cancer Diseases," 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS), Thessaloniki, 2017, pp. 755-760, doi: 10.1109/CBMS.2017.62.

[9] R. Maguire et al., "The eSMART study protocol: a randomised controlled trial to evaluate electronic symptom management using the advanced symptom management system (ASyMS) remote technology for patients with cancer", BMJ Open, 2017;7:e015016, doi: 10.1136/bmjopen-2016-015016

[10] O. Ciani et al., "Lung Cancer App (LuCApp) study protocol: A randomised controlled trial to evaluate a mobile supportive care app for patients with metastatic lung cancer," BMJ Open, vol. 9, no. 2, pp. 1–10, 2019.

[11] PA Harris et al., "The REDCap consortium: Building an international community of software partners," J Biomed Inform. 2019 May 9, doi: 10.1016/j.jbi.2019.103208

Implementación de mecanismos de seguridad en una plataforma IoT para la monitorización remota de la actividad física

J. Calvillo Arbizu[1,2], F. Ramos Rojas, I. Román Martínez[2], J. Reina Tosina[1,3]

[1] Grupo de Ingeniería Biomédica, Universidad de Sevilla, Sevilla, España, {jcalvillo, jreina}@us.es

[2] Departamento de Ingeniería Telemática, Universidad de Sevilla, Sevilla, España, isabel@trajano.us.es

[3] Departamento de Teoría de la Señal y Comunicaciones, Universidad de Sevilla, Sevilla, España

Resumen

El paradigma de Internet of things (IoT) supone una oportunidad para el despliegue de nuevos modelos asistenciales soportados por la tecnología como la monitorización remota domiciliaria. No obstante, entre los obstáculos que impiden su amplia aplicación está el riesgo de vulneración de la información de los pacientes. Diversas herramientas soportan mecanismos de autenticación y control de acceso para escenarios de IoT pero pocas iniciativas están específicamente orientadas al ámbito de la salud.

En este trabajo se diseña y desarrolla una serie de mecanismos de seguridad en una plataforma IoT para la monitorización remota de la actividad física. Las herramientas están basadas en FIWARE y el control de acceso se efectúa siguiendo el modelo basado en atributos y el lenguaje XACML. La autenticación se realiza mediante tokens y el control de acceso se garantiza tanto para publicadores de información como para consumidores. Los resultados muestran la viabilidad de las herramientas FIWARE para satisfacer los requisitos de seguridad de un escenario sanitario si bien ha sido necesario adaptarlas para cubrir aspectos específicos del dominio.

1. Introducción

Los escenarios de atención sanitaria domiciliaria se están beneficiando tanto de los avances en sensorización como del desarrollo de plataformas bajo el paradigma de *Internet of Things* (IoT) [1]. Uno de los beneficios de IoT es la posibilidad de disponer de un registro continuo de datos fisiológicos que puedan ser procesados de forma automática para detectar patrones de riesgo o favorecer las estrategias de medicina personalizada a partir de una supervisión avanzada de la evolución del paciente. En contraposición, los riesgos de seguridad (como la suplantación de identidad o *phishing*, la interceptación de datos por parte de terceros maliciosos, o el acceso no autorizado a información personal [2]) obstaculizan el uso de este tipo de plataformas.

La práctica regular de actividad física es una de las intervenciones que suponen mayores beneficios en el estado de salud de la población, así como en la reducción de patologías cardiovasculares. Así mismo, en diversas enfermedades crónicas como la enfermedad pulmonar obstructiva crónica (EPOC), la práctica de actividad física regular mejora la capacidad física y la reducción de síntomas [3]. Por ello, este tipo de intervenciones suelen formar parte del tratamiento de diversas patologías crónicas, aunque en el ámbito hospitalario suponen un alto consumo de recursos [4].

Los programas domiciliarios de actividad física supervisados remotamente se destacan como una alternativa con potencial para reducir los obstáculos derivados de su acceso y coste [5]. Varios estudios han sugerido que este tipo de intervenciones soportadas por la tecnología suponen un incremento en la cantidad y calidad de la actividad física realizada y una reducción del tiempo de sedentarismo cuando se comparan con poblaciones de control [5]. Sin embargo, también se ha concluido que las mejoras en la adherencia desaparecen a largo plazo si la intervención finaliza [6].

Las capacidades que brinda el paradigma de IoT pueden ser clave para superar las limitaciones de las actuales intervenciones remotas, aunque los riesgos de seguridad deben minimizarse con el objetivo de contar con soluciones seguras, fiables y eficientes. Aunque existen herramientas para minimizar los riesgos de seguridad en plataformas IoT, pocas tienen en cuenta los requisitos específicos de este dominio. Este trabajo presenta el diseño y desarrollo de una plataforma segura para la monitorización remota de actividad física basada en el paradigma IoT y utilizando herramientas de código abierto (FIWARE). En este trabajo se detallan las medidas de seguridad implementadas en la plataforma en cada uno de los procesos que se llevan a cabo en su ejecución atendiendo a los requisitos particulares del dominio.

En la sección II se describen los aspectos relevantes de seguridad y privacidad de los datos en el contexto que nos ocupa, se discute sobre el concepto IoT en el dominio sanitario y se introduce la iniciativa tecnológica FIWARE. La sección III muestra los resultados del diseño y desarrollo de la plataforma propuesta. El trabajo finaliza con las conclusiones en la sección IV.

2. Materiales y métodos

2.1. Conceptos básicos de seguridad y privacidad

Cualquier aplicación tecnológica en el dominio sanitario debe garantizar la privacidad de la información y evitar la participación de terceros maliciosos. Para conseguirlo, por un lado, debe verificar la identidad de cualquier usuario o sistema que pretenda acceder a recursos de salud. Esta identificación tiene que ser validada por alguna autoridad o servicio de identidad con potestad reconocida por todas las partes implicadas.

Una vez que el solicitante ha sido correctamente identificado debe evaluarse si tiene autorización para

ISBN: 978-84-09-25491-0, pags. 227- 230

acceder al recurso solicitado. Existen diversos modelos de control de acceso, destacando por su flexibilidad el modelo basado en atributos (ABAC, por sus siglas en inglés). En ABAC, un intermediario (el punto de cumplimiento de políticas o PEP) intercepta la petición del usuario y la deriva al PDP (punto de decisión de políticas) que es quien decide si esa petición debe ser permitida. La decisión se basa en un conjunto de atributos del usuario, del recurso protegido y de la acción, entre otros (Figura 1). Estos atributos pueden incluirse en la propia petición de decisión o recuperarse de repositorios específicos (o puntos de información de políticas, PIP). Las políticas que definen las reglas de acceso a los recursos protegidos se encuentran en uno o varios elementos (Punto de Administración de Políticas o PAP) que son consultados por el PDP cada vez que tiene que tomar una decisión.



Figura 1. *Esquema de modelo de control de acceso basado en atributos (ABAC)*

2.2. IoT y FIWARE

Diferentes tecnologías de sensorización, procesamiento de información y comunicación convergen para soportar el paradigma de IoT [7]. FIWARE es una iniciativa financiada por la Comisión Europea cuyo objetivo es contribuir al despliegue generalizado de plataformas IoT en la sociedad, siguiendo un modelo de código abierto y computación en la nube [8]. La arquitectura de referencia de FIWARE incluye un conjunto de funciones de propósito general, publicadas como APIs públicas (llamadas Generic Enablers, GE).

Uno de los componentes principales de la arquitectura FIWARE es el Context Broker (CB), elemento encargado de gestionar el ciclo de vida de la información intercambiada entre publicadores y suscriptores. Entre las implementaciones de CB existentes destaca Orion [9] que sirve de intermediario de la información mediante el uso de entidades, esto es, de elementos de información previamente definidos. Otros componentes de FIWARE proporcionan capacidades de procesamiento de eventos, seguridad o almacenamiento histórico de eventos. En este trabajo se hace uso de los siguientes elementos relacionados con la seguridad: el gestor de identidad Keyrock [10], el intermediario Wilma [11] y el punto de decisión Authzforce [12]. Para la gestión de la información que envían los sensores, FIWARE dispone de diversos componentes de procesamiento de eventos y persistencia que se pueden combinar con herramientas desarrolladas por terceros para obtener una solución avanzada y robusta.

3. Resultados

3.1. Diseño de la plataforma

En nuestra plataforma de monitorización los agentes IoT publican información sobre la actividad física y la frecuencia cardíaca de los pacientes a través de la entidad *PhysicalActivity* [13]. Esta entidad comprende los siguientes campos: identificador de paciente y de dispositivo, instante de registro de la medida, frecuencia cardíaca, gastos metabólicos instantáneo y acumulado. En la Figura 2 se muestra una solicitud de publicación al Orion CB a través de una acción POST.

```
POST http://orion.fiware.host:1026/v2/entities/
            PhysicalActivity1363/attrs?options=keyValues
{
            "Patient" : "User34562",
            "Device" : "ActivityMonitor1360",
            "date" : "2020-04-03T11:24:00",
            "instant_metabolic_expenditure": "0.52",
            "accumulated_metabolic_exp": "267.15",
            "heart_rate": "75"
}
```

Figura 2.*Ejemplo de mensaje POST con datos de actividad enviados desde un gestor IoT al Context Broker*

Mientras que los publicadores envían los datos a Orion siguiendo el patrón definido en la entidad, los suscriptores pueden suscribirse a cualquier información que se ajuste a dichas plantillas. Así, una vez que Orion CB recibe una nueva publicación, la envía a los suscriptores interesados.

En la Figura 3 se muestra un esquema de la plataforma así como los procesos de autenticación y control de acceso que se detallan a continuación. Una de las aportaciones de nuestra plataforma es que el control de acceso no se limita a la consulta de información por parte de los consumidores sino que también las publicaciones de los agentes IoT y las peticiones de suscripción de los consumidores deben ser aprobadas antes de ejecutarse. Estas acciones, por lo general, no suelen ser controladas por mecanismos de seguridad.

3.2. Autenticación

El proceso de comprobación de identidad está estrechamente relacionado con el control de acceso puesto que necesitamos verificar la identidad del solicitante para decidir si tiene permitido o no realizar la acción solicitada.

En nuestra plataforma, cualquier comunicación con Wilma (el PEP que recoge las peticiones de acciones de participantes) requiere un *token* de acceso con la identidad del solicitante. Para obtenerlo, el usuario, sistema o agente IoT realiza una petición POST a Keyrock indicándole el usuario y la contraseña del solicitante que habrán sido registrados previamente en su base de datos (puntos 1 de Figura 3). Si la información enviada corresponde con un usuario registrado, Keyrock devolverá un *token* de acceso que el solicitante utilizará para realizar acciones en la plataforma IoT.

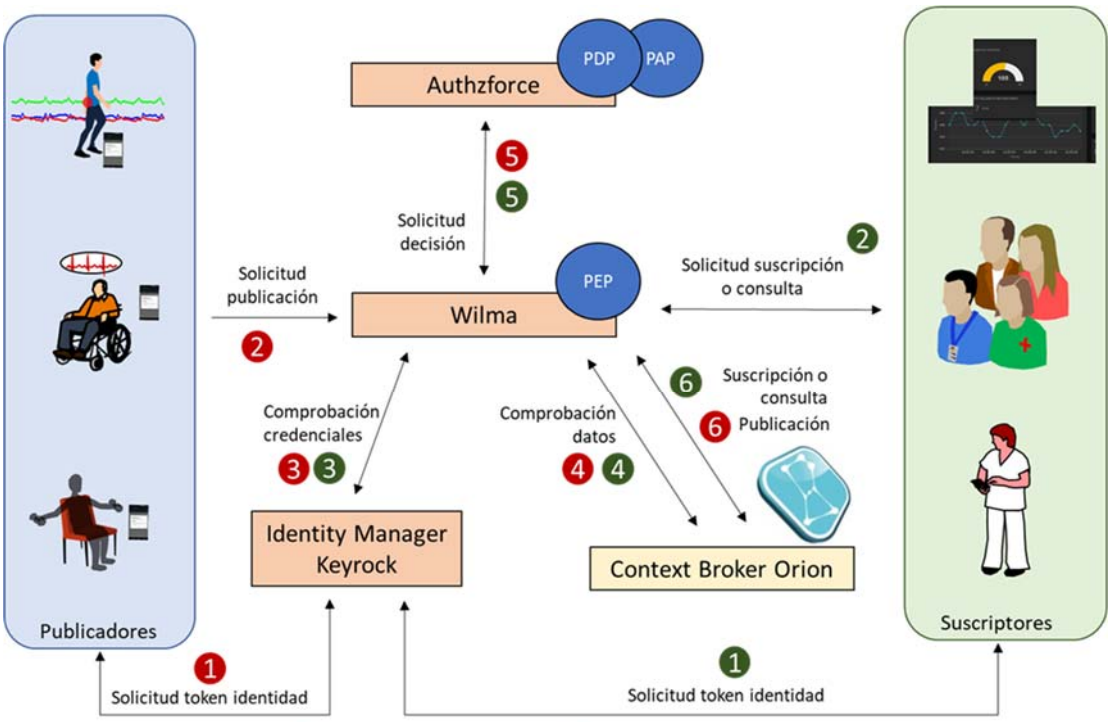


Figura 3. *Esquema de la plataforma segura para la monitorización remota de actividad física basada en IoT*

Cuando un participante solicite la realización de una acción a Wilma, adjuntará a la petición su *token* de acceso (puntos con el número 2). Wilma enviará dicho *token* a Keyrock el cual responderá certificando o no la validez del *token* y la identidad del solicitante (puntos con el número 3). Si la identidad es correcta Wilma lanza el proceso de autorización correspondiente.

3.3. Autorización para los publicadores

Los publicadores en FIWARE pueden realizar dos tipos de acciones: publicar una entidad nueva en Orion o actualizar una entidad previamente publicada. En el primer caso, el agente IoT envía a Wilma la petición POST con la entidad que desea publicar junto con su *token* de acceso (punto 2 rojo de la Figura 3). Wilma comprobará si las credenciales son correctas consultando a Keyrock y, en caso de que así sea, extrae de la petición aquellos datos necesarios para la toma de decisión de acceso. Actualmente, la decisión se basa en los atributos: identificador y organización del publicador, identificador del paciente objeto de la medida, dispositivo con el que se ha realizado esta y acción que se pretende realizar (publicación o actualización). Con estos datos Wilma desarrolla la petición de autorización utilizando el lenguaje XACMLv3 y la envía a Authforce.

Authforce evalúa la petición utilizando todas las políticas de control de acceso relevantes en cada caso y que han sido previamente definidas por el administrador de la plataforma. Las políticas están formalizadas también en XACMLv3 y se basan en atributos, pudiendo incluir (además de los atributos necesarios de usuario, recurso y acción) otros atributos adicionales como el periodo de tiempo en el que se permite la publicación (por ejemplo, solo en horario diurno). Las políticas definidas relacionan identificadores de agentes con pacientes y dispositivos de manera que se controle que la información publicada es coherente con las relaciones configuradas y no existe suplantación de identidad. Si la decisión de Authforce es positiva, Wilma redirecciona la petición original a Orion para que publique la entidad de datos de actividad física.

El segundo caso de publicación (actualización de una entidad previamente publicada) es análogo al anterior pero, como se trata de actualizar una entidad publicada anteriormente, en primer lugar Wilma recupera de Orion la entidad publicada que se quiere actualizar (punto 4 rojo de Figura 3) y esa información (junto a los datos de la petición del agente) es remitida a Authforce para que dictamine si ese agente puede actualizar esa entidad concreta y la petición que está realizando es coherente con la publicación previa.

3.4. Autorización para los consumidores

En FIWARE los consumidores pueden suscribirse a las entidades definidas en Orion para recibir un mensaje cuando se produzca una publicación en ellas, o consultar las entidades actualmente publicadas. En ambos casos el procedimiento es similar. En primer lugar, y siempre después de haber verificado la identidad del consumidor, Wilma consulta en Orion la entidad sobre la que quiere actuar (suscripción o consulta). Se recuperan así los datos de la entidad que serán necesarios para que Authforce dictamine si puede ser consultada o suscrita por el consumidor (por ejemplo, porque pertenecen a la misma organización o porque es el médico de referencia del paciente objeto de los datos). Wilma compondrá la petición de decisión con estos datos y la propia petición del consumidor y la enviará a Authforce.

Como en el caso anterior, se evaluarán las políticas almacenadas para tomar la decisión adecuada que se comunicará a Wilma. En caso de que se permita la acción, Wilma enviará los datos de la entidad al consumidor (en caso de petición de consulta) o comunicará a Orion la

petición de suscripción para que la almacene y la ejecute cuando se den nuevas publicaciones en esa entidad.

FIWARE permite que un consumidor consulte todas las entidades que en un momento dado tenga Orion en uso. Si consideramos que diferentes organizaciones sanitarias pueden hacer uso de la misma plataforma nos encontramos con un riesgo de seguridad ya que un consumidor perteneciente a una organización podría consultar información sobre pacientes que no tiene asignados e incluso que pertenecen a otra organización. Para resolver este potencial riesgo, si un consumidor realiza una consulta general (haciendo una petición GET a Wilma sin especificar ningún recurso, /v2/entities), se ha modificado Wilma para que derive toda la información del consumidor a Authforce y que este le indique a qué entidades puede acceder y a cuáles no. Una vez hecho esto, Wilma recogerá de Orion solo aquellas entidades autorizadas para el consumidor y se las devolverá como respuesta a su petición general. De esta manera, el consumidor nunca conocerá si existen más entidades aparte de aquellas a las que tiene derecho acceder.

4. Conclusiones

El paradigma IoT supone una oportunidad para el desarrollo de nuevos modelos asistenciales ya que el uso de dispositivos sensores permite recopilar información sobre el estado de salud del paciente, su entorno o la adherencia a planes de cuidados. No obstante, los riesgos a la vulneración de la información intercambiada requieren el despliegue de mecanismos que garanticen la privacidad y seguridad de los pacientes y los datos.

En este trabajo se presenta una prueba de concepto de una plataforma segura para la monitorización remota de actividad física basada en IoT. Los procesos de autenticación y autorización se basan en los componentes de FIWARE pero se han modificado para satisfacer los requisitos del dominio. Por ejemplo, aunque se suele realizar el control de acceso sobre los consumidores, en pocas ocasiones se garantiza también un control sobre los publicadores, evitando así que terceros maliciosos publiquen información falsa o suplanten la identidad de agentes habilitados. Además, se han implementado mecanismos de seguridad que facilitan la utilización simultánea de la plataforma por publicadores y consumidores pertenecientes a distintas organizaciones. La información de las entidades y los atributos utilizados permiten que la decisión de acceso tenga en cuenta los diferentes dominios administrativos e incluso se podrían configurar diferentes niveles de acceso dentro de la misma organización.

La aplicación del modelo ABAC y el lenguaje XACML permite un proceso de autorización flexible cuya complejidad está a voluntad de la necesidad de cada escenario ya que los atributos utilizados aquí (atributos de usuario, recurso, acción y entorno) podrían complementarse con cualquier información relevante para la toma de decisiones.

Por último cabe destacar que, atendiendo a los riesgos de seguridad, hemos satisfecho los procesos de autenticación y autorización. No obstante, quedan pendientes aquellas vulneraciones que se pueden producir en el intercambio de los datos (si son interceptados por terceros maliciosos) para lo cual se planea complementar este trabajo con mecanismos de encriptación y cifrado de la información intercambiada. Los trabajos futuros pasan por resolver esos niveles de seguridad complementarios y desplegar un escenario piloto que se pueda evaluar con usuarios reales.

Agradecimientos

Este trabajo ha sido soportado por el Grupo de Ingeniería Biomédica de la Universidad de Sevilla, la Fundación Progreso y Salud, Junta de Andalucía, bajo el proyecto EvaRehaResp PIN-0394-2017, y en parte por el Programa Eurostars de la Comisión Europea, bajo el proyecto FRAIL, y el proyecto VÍA-RENAL (financiado por la Fundación Mutua Madrileña).

Referencias

[1] Al-Fuqaha M, Guizani M, Mohammadi M, Aledhari M, Ayyash M. Internet of things: A survey on enabling technologies, protocols, and applications. *IEEE communications surveys & tutorials*, vol. 17(4), 2015, pp 2347-76.

[2] Khan MA, and Salah K. IoT security: Review, blockchain solutions, and open challenges. *Future Generation Computer Systems*, vol. 82, 2018, pp 395-411.

[3] ZuWallack R. The nonpharmacologic treatment of chronic obstructive pulmonary disease. *Proceedings of the American Thoracic Society*, vol. 4, 2007, pp 549–53.

[4] Hekler EB, King AC, Banerjee B, Robinson T, et al. A Case Study of BSUED: Behavioral Science-informed User Experience Design Online. http://v1.personalinformatics.org/docs/chi2011/hekler.pdf (Consultada:Septiembre 2020).

[5] Stockwell S, Schofield P, Fisher A, Firth J, et al. Digital behavior change interventions to promote physical activity and/or reduce sedentary behavior in older adults. *Experimental Gerontology*, vol. 120, 2019, pp 68-87.

[6] Rothman AJ. Toward a theory-based analysis of behavioral maintenance. *Health Psychology*, vol. 19(1), 2000, pp 64–9.

[7] Li S, Da Xu L, and Zhao s. The internet of things: a survey. *Inform Syst Front*, vol. 17, 2015, pp. 243–59.

[8] FIWARE. https://www.fiware.org/ (Consultada: Septiembre 2020).

[9] FIWARE Orion. https://fiware-orion.readthedocs.io/en/master/ (Consultada: Septiembre 2020).

[10] FIWARE Keyrock. https://fiware-idm.readthedocs.io/en/latest/#identity-manager-keyrock (Consultada: Septiembre 2020).

[11] FIWARE Wilma. https://fiware-academy.readthedocs.io/en/latest/security/wilma/index.html (Consultada: Septiembre 2020).

[12] FIWARE Authzforce. https://authzforce-ce-fiware.readthedocs.io/en/latest/ (Consultada: Septiembre 2020).

[13] Calvillo-Arbizu J, Román-Martínez I, Roa-Romero LM, Reina-Tosina J. Armonización ISO/IEEE 11073 - IoT en la monitorización de actividad física. XXXIV Simposium Nacional de la Unión Científica Internacional de Radio, Sevilla (España), 2019

Análisis emocional de tweets en castellano para personalizar tratamientos clínicos en telemedicina

M. Jiménez Hernando[1], R. Pérez Javato, P. Sánchez-González[1,2], D. Moreno Blanco[1], E. J. Gómez Aguilera [1,2]

[1] Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, España, {manuel.jimenez, p.sanchez, diego.morenob, enriquejavier.gomez } @upm.es

[2] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Zaragoza, España

Resumen

Este trabajo se fundamenta en la hipótesis de que el uso de emoticonos en las redes sociales puede ser un indicador de las emociones del usuario. Creamos una base de datos utilizando la API de Twitter para entrenar un modelo de inteligencia artificial capaz de clasificar un mensaje corto en 6 posibles emociones (alegría, tristeza, miedo, sorpresa, enfado y asco).
Se ha creado un entorno web para poder validar el modelo de clasificación multiclase entrenado, obteniendo una efectividad de 45,3%, triplicando la tasa esperada e igualando los resultados de los estudios de análisis sentimental o de polaridad de referencia en castellano. Los resultados obtenidos suponen un avance en el análisis de emociones en castellano, pues no existen trabajos similares en nuestro idioma. El objetivo último de esta investigación es contribuir a una mayor personalización de los tratamientos y a la mejora de los algoritmos de ayuda a la decisión en telemedicina.

1. Introducción

La personalización de los tratamientos es una pieza fundamental para conseguir aumentar la adherencia de los pacientes en telemedicina. La comprensión del estado emocional de los usuarios es un punto clave para poder llevar a cabo esta personalización. Una forma de abordar el objetivo de conocer el estado de ánimo de los pacientes es el análisis de textos escritos por los propios pacientes, a través del procesamiento del lenguaje natural.

Dentro del mundo del procesamiento de lenguaje natural (PNL) cabe destacar la diferencia entre los modelos de Análisis Sentimental, donde se predice el contenido sentimental de un texto en un eje de positividad-negatividad, mientras que el Análisis Emocional típicamente se refiere a la predicción de las diferentes emociones básicas de Ekmann [1] (alegría, tristeza, miedo, asco, sorpresa y enfado) que pueden aparecer en un texto. Un problema mucho más complejo, ya que pasamos de una clasificación binaria a una multiclase. Mientras que para el primero ya ha habido estudios muy prometedores en castellano [2][3], para el segundo, el campo en castellano está mucho menos explorado, principalmente por la falta de datasets.

Ante la ausencia de datasets en castellano de textos cortos ligados a emociones, en este trabajo se ha procedido a la creación de una base de datos customizada mediante la explotación de datos de la red social Twitter. Estableciendo la hipótesis de que los emojis son un indicador o aproximación de las emociones, se pueden etiquetar los tweets con la emoción que representa el emoji que los acompaña. Además, se ha creado una página web para realizar la validación del modelo desarrollado en el presente trabajo.

2. Materiales

Para la realización de este trabajo se ha utilizado el lenguaje Python, debido a su flexibilidad y la facilidad para conectar con la API de Twitter mediante la librería tweepy. La base de datos ha sido creada mediante la explotación de mensajes publicados en esta red social seleccionados por el uso de determinados emoticonos.

Se han utilizado las librerías keras y pandas para el procesamiento de datos y el entrenamiento del modelo de predicción. Gracias a dicha flexibilidad, también se ha podido crear el entorno web mediante este lenguaje con la librería django.

3. Trabajos relacionados

La principal inspiración de este trabajo es la herramienta Deepmoji, desarrollada por el MIT [4]. Encontramos una gran diferencia con este trabajo, por un lado, en la enorme cantidad de tweets que utilizan en su estudio (del orden de 10^9) y, por otro lado, el idioma. Esta diferencia idiomática es, de hecho, fundamental para entender el éxito de su herramienta respecto a la nuestra.

La investigación en temas de procesamiento de lenguaje natural en castellano es mucho menos abundante que para la lengua inglesa, no obstante, nuestro idioma es el tercero más hablado a nivel mundial. Esta enorme diferencia de producción científica respecto al inglés se debe a: la falta de bases de datos estructuradas en español, la hegemonía del inglés como lengua científica y a la falta de herramientas para el preprocesamiento de textos en castellano. Con este trabajo queremos aportar un granito de arena al campo del análisis emocional en castellano.

Los trabajos más relevantes en este campo en nuestro idioma son [2][3], ambos trabajos surgieron como respuesta a un mismo reto de la 4ª conferencia TASS de 2015, donde se proporcionó a los participantes un conjunto de tweets clasificados en distintos niveles de polaridad, desde muy negativo a muy positivo. Se utilizaron dos técnicas de aprendizaje profundo para clasificar los textos en 6 niveles de polaridad.

ISBN: 978-84-09-25491-0, pags. 231- 234

En nuestro caso la clasificación es más compleja, ya que las categorías que predecimos son menos antagónicas que positivo y negativo. La diferencia entre miedo y sorpresa, por ejemplo, puede ser muy compleja. Es por ello que en [1], se apoyan en un dataset enorme para que la red neuronal llegue a entender este tipo de relaciones sutiles del lenguaje.

4. Metodología

4.1. Creación de una base de datos.

Esta primera fase es un punto vital del proyecto. Tras la revisión del estado del arte, se ha llegado a la conclusión de que había que crear una base de datos propia, debido a la falta de colecciones en castellano. Inicialmente se ha hecho una selección de 22 emoticonos, relacionados con las 6 emociones ya mencionadas. Después, tras un proceso de análisis de las palabras más utilizadas en cada emoticono supervisado por el criterio humano, se seleccionan los 6 emoticonos más representativos para crear la colección de datos definitiva.

La base de datos final consta de 120.000 tweets, 20.000 para cada emoción. Para mejorar la calidad de los datos se han excluido retweets, menciones, tweets con más de un emoticono y tweets con enlaces.



Figura 1: Selección inicial y final de emoticonos

4.2 Preprocesamiento. El preprocesamiento de los datos es un punto crucial en las aplicaciones de PLN. Para que el algoritmo de clasificación funcione óptimamente se han de limpiar todos los datos guardados en nuestra colección. Una particularidad de adquirir los datos directamente de Twitter [1], es la gran cantidad de elementos informales, abreviaturas y faltas de ortografía en el lenguaje utilizado. Por lo que se ha dedicado más esfuerzo en la limpieza de cada texto corto de la base de datos, utilizando técnicas de limpieza de faltas de ortografía. Además del preprocesamiento estándar en PLN, que consiste en la eliminación de artículos, preposiciones, signos de puntuación y caracteres especiales.

4.3 Diseño e implementación del modelo. La clasificación de textos es un problema clásico del machine learning. Pese a que las redes neuronales convolucionales se han probado eficaces para resolver algunos de estos problemas [2], en este trabajo se ha optado por desarrollar una red neuronal recurrente. Concretamente del tipo Long Short Term Memory o LSTM, por su capacidad de almacenar información a largo plazo. Una vez seleccionado el modelo, se ha realizado el ajuste de los hiperparámetros, para asegurar el comportamiento óptimo de la red, es decir, que sea capaz de realizar una correcta clasificación de los datos evitando el overfitting. Al tener un conjunto de datos no masivo, la red neuronal no debe tener muchos parámetros. Es por ello que la red neuronal implementada tiene únicamente una capa oculta con 10 unidades LSTM.

4.4 Entrenamiento y evaluación del modelo. Para llevar a cabo esta fase, se ha dividido previamente la base de datos en dos colecciones, la colección de entrenamiento y la de evaluación. La colección de entrenamiento está compuesta por dos tercios de los tweets de la colección inicial, mientras que la de evaluación está formada por el tercio restante. Así, para esta fase, se introducen ambas colecciones en el algoritmo, de forma que se entrena con la colección de entrenamiento y posteriormente, se evalúa con la colección de evaluación, pudiendo comprobar si el algoritmo ha aprendido correctamente mediante la exactitud obtenida.

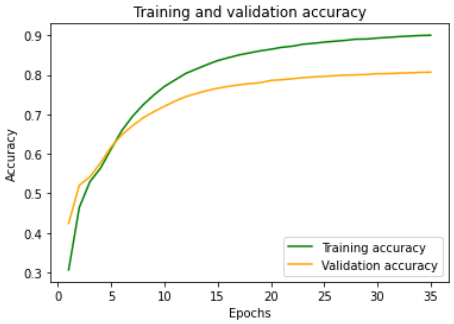


Figura 2: Evolución de la eficacia del modelo

En evaluación, para los datos de twitter se ha obtenido una eficacia predictora del 79,9%. Sin embargo, para juzgar el modelo en un escenario más realista, se ha creado un entorno web para realizar una validación con textos no extraídos de twitter.

4.5 Creación del entorno web y validación del modelo. El objetivo de este paso es que mediante la página web, las personas accedan, introduzcan una frase y el algoritmo les devuelva una respuesta en forma del emoticono que más se ajusta a la frase introducida (Figura 3). Tras esa respuesta, el usuario tiene a su disposición una encuesta de satisfacción que ha de rellenar. Por último, con las respuestas de la encuesta se verificará el funcionamiento del modelo.



Figura 3: Página principal del entorno web de validación

5. Validación y resultados

La validación del modelo se ha realizado con 150 frases que han sido escritas en parte por diferentes usuarios que han accedido a la página web y por otra parte por mensajes extraídos directamente de twitter con uno de los emoticonos de referencia. De esas 150 frases, 68 se han predicho correctamente, lo que supone un **45,3%** del total, mientras que 82 (el 54,7% restante) se han predicho de manera errónea. Estadísticamente, la probabilidad de cada una de las 6 emociones es del 16,66% (1/6), al tratarse de una clasificación multiclase, donde a la hora de predecir, el Emoji correcto "compite" contra otros 5 emojis posibles, lo que hace que la clasificación sea mucho más compleja que la clasificación binaria. Así, los resultados obtenidos triplican la tasa de acierto esperada, pasando de un 16,66% a un 45,33%.

Si se analiza más en detalle el número de aciertos y fallos (Figura 4), separando por los diferentes emojis, se puede observar que el modelo predice correctamente el 67,4% de las frases que expresan tristeza y las frases que expresan miedo (66,6%). Por otro lado, el modelo tiene una tasa de aciertos del 45,5% en las frases que expresan enfado, es decir, es capaz de predecir correctamente la mitad de las veces.

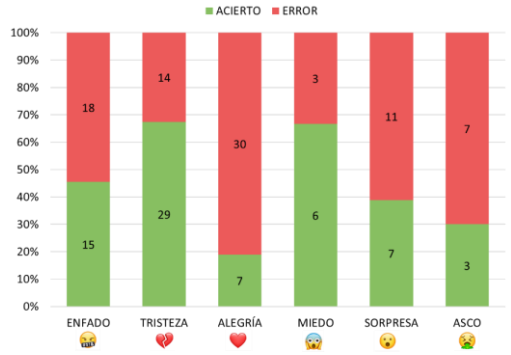


Figura 4: Desglose de resultados por emoticono

En las frases que expresan sorpresa, el modelo predice correctamente un 38,8% de las veces y en las frases en las que se expresa asco, la tasa de aciertos es del 30,0%. Por último, la emoción que peor predice el modelo es la alegría, con una tasa de aciertos del 18,9%.

Una hipótesis para explicar este resultado asimétrico puede ser el lenguaje asociado a cada emoji. Mientras que en el caso de la tristeza o el miedo las palabras que la gente expresa no dan lugar a duda, en el caso de la alegría o el asco, puede que el lenguaje empleado sea mucho más ambiguo, tenga palabras clave menos comunes o esté plagado de sarcasmo, o bien de casos especiales de palabras, como las erratas o abreviaturas.

Observando las frases que expresan alegría, hay una gran tendencia a clasificarlas con un sentimiento muy distinto, prácticamente opuesto (Figura 5). Así, el 45,9% de las veces, el modelo ha predicho que la frase expresaba tristeza o miedo, frente a un 18,9% de las veces que se ha clasificado con el sentimiento de alegría.

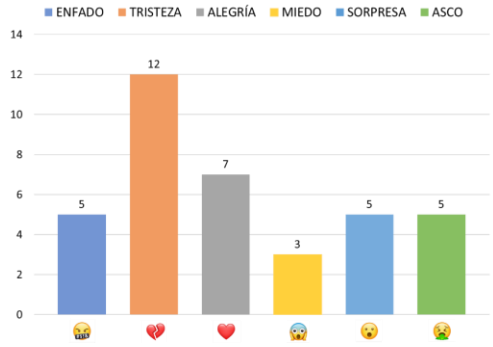


Figura 5: Predicciones para frases con emoji de alegría

Este resultado se puede explicar mediante el sarcasmo o la ironía, ya que utilizando el emoticono opuesto al que le correspondería al mensaje genera esa ironía que un ser humano puede entender, pero para una máquina es más difícil. Esta presencia del sarcasmo en los datos de entrenamiento provoca que se asigne en las predicciones el emoticono opuesto al que expresa. La manera de solucionar esto sería con muchos más datos de entrenamiento, ya que esto nos permitiría entrenar a una red neuronal con muchos más parámetros, con la certeza de que alcanzará un grado de aprendizaje del sarcasmo [1].

Como hemos mencionado antes, el lenguaje escrito es muy ambiguo y según la persona que lo lea puede ser interpretado con una u otra emoción. Si tomamos como ejemplo el siguiente mensaje: ***"Primera vez en la vida que cocino pollo y me queda crudooo"***, se puede observar que, aunque el usuario quería expresar asco, el algoritmo ha clasificado la frase como tristeza.

ALEGRÍA	TRISTEZA	MIEDO	ENFADO	ASCO	SORPRESA	ACIERTO	EMOCIÓN CORRECTA
0,37%	**90,78%**	0,67%	0,29%	7,45%	0,43%	**NO**	**ASCO**

No obstante, si esta frase es interpretada por otra persona, se podría interpretar las emociones de tristeza, enfado o asco. Con este ejemplo se muestra la complejidad del problema debida a la ambigüedad del lenguaje.

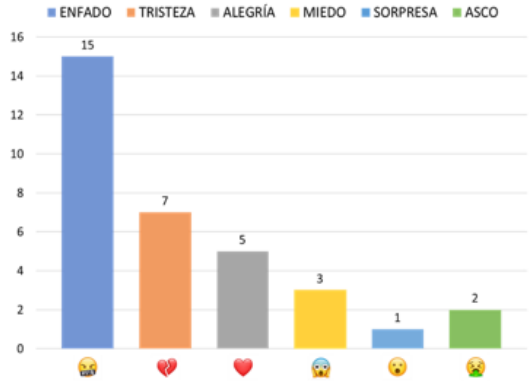


Figura 6: Predicciones para frases con emoji de enfado

También podemos ver que para las frases que expresaban enfado (Figura 6), el algoritmo las ha clasificado correctamente un 45,4% de las veces. Siendo la segunda

emoción más predicha la tristeza, en un 21,2%. Podemos interpretar que estos errores no son graves, debido a las similitudes emocionales entre enfado y tristeza.

A continuación, mostramos una comparación de los resultados obtenidos con los estudios de referencia en castellano [2][3]. Se puede observar que la exactitud del modelo desarrollado en este proyecto es similar a la de los otros dos trabajos en castellano de referencia de análisis de polaridad de sentimientos. Recordamos que para el análisis sentimental la clasificación ocurre en un eje positivo-negativo, mientras que en nuestro trabajo la clasificación se da entre 6 emociones distintas.

MÉTODO	DESCRIPCIÓN	EXACTITUD EN LA FASE EVALUACIÓN
LSTM (VILARES ET AL.)	6 niveles de polaridad	56,8%
CNN (SEGURA ET AL.)	6 niveles de polaridad	54,0%
LSTM	**6 emociones**	**45,3%**

Por otro lado, al comparar los resultados con la herramienta Deepmoji [1], se puede observar que los resultados que hemos obtenido en este trabajo podrían mejorar mucho más. En ese proyecto se obtuvo un 82,0% de exactitud en la validación con 64 emociones, mientras que en castellano se ha conseguido llegar a un 45,3% de exactitud con 6 emociones. La razón de la gran efectividad de esta herramienta es que utilizó una colección de datos enorme, formada por 1,2 billones de tweets. Mientras que, en este proyecto, se han utilizado alrededor de 120.000.

6. Conclusiones

Estableciendo la hipótesis de que los emoticonos son un indicador de las emociones en redes sociales, se ha podido crear una base de datos para el análisis emocional en castellano utilizando la API de Twitter y Python. Posteriormente se ha entrenado a una red neuronal del tipo LSTM para predecir qué emoción representa mejor a un texto corto, similar a un tweet. Se ha proporcionado una herramienta web para hacer la validación con 150 frases de diferentes usuarios. Los resultados han sido prometedores, con una tasa de acierto del 45,3% para 6 emociones, triplicando la tasa esperada y casi igualando a estudios anteriores en castellano sobre análisis sentimental o de polaridad.

El factor que más afecta a nuestro resultado es la cantidad de datos. Al tener una base de datos relativamente pequeño (del orden del megabyte y no del terabyte), no nos podemos permitir entrenar una red neuronal con un gran número de parámetros. El hecho de que la red no tenga suficientes parámetros tiene una repercusión directa en la efectividad sobre los casos especiales, como las abreviaturas, erratas, anglicismos y demás fenómenos lingüísticos extraños. También hemos planteado la hipótesis de que, con una red suficientemente grande, se podría llegar a alcanzar un entendimiento del sarcasmo y la ironía, tan presentes en el lenguaje de las redes sociales.

Cabe destacar que la mayoría de los avances en los modelos de procesamiento del lenguaje natural se desarrollan en lengua inglesa, tardando unos años en aparecer resultados equivalentes en castellano.

En conclusión, teniendo en cuenta la naturaleza del problema de clasificación multiclase, las limitaciones del tamaño de la base de datos y el acceso a librerías en castellano, los resultados obtenidos son prometedores.

Utilizando herramientas de procesamiento de lenguaje natural como la presentada en este trabajo, el horizonte del entendimiento emocional de los usuarios en telemedicina está un paso más cerca. Ya que se podrá incluir la respuesta emocional de los pacientes en la personalización de tratamientos y en los algoritmos de ayuda a la decisión en contextos clínicos [5], un tema que tendrá gran relevancia en el futuro.

7. Agradecimientos

Parcialmente financiado por FEDER/Ministerio de Ciencia, Innovacion y Universidades– Agencia Estatal de Investigación/_Proyecto DPI2017-86088-C3-1-R. Y la beca de Formación de Personal Investigador PRE2018-085736.

Referencias

[1] P. Ekmann, «An Argument for Basic Emotions,» de *Cognition & Emotion*, University of California , San Francisco, USA, pp. vol. 6, no. 3, pp. 169–200.

[2] I. Segura-Bedmar, A. Quiros y P. Martínez, «Exploring Convolutional Neural Networks for Sentiment Analysis of Spanish tweets,» de *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, 2015.

[3] D. Vilares, Y. Doval, M. A. Alonso y C. Gómez-Rodríguez, «LyS at TASS 2015: Deep Learning Experiments for Sentiment Analysis on Spanish Tweets,» *TASS 2015*, pp. 47-52, 2015

[4] B. Felbo, A. Mislove, A. Søgaard, I. Rahwan y S. Lehmann, «Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm». *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2017.

[5] J. Solana Sánchez, F.J. Gárate Barreiro, A. García-Rudolph, R. Sánchez-Carrión,E. Hernando Pérez, E. Opisso, E.J. Gómez. Plataforma PERSSILAA: Algoritmos y herramientas de ayuda a la decisión para la prevención de la fragilidad en personas mayores. *XXXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica, Valencia*, 2016

Validación de la electrónica y ubicación de un sistema para la obtención ambulatoria de constantes vitales en el oído externo

G.M. Besné [1,2], E. Irigoien[1,3], JD. Trigo[1,3], JJ. Astrain[1,3], A. Lopez-Martin[1,3], J. Artieda[1,2], M. Alegre[1,2], L. Serrano Arriezu[1,3], F. Falcone[1,3] y M. Valencia[1,2]

[1]Universidad de Navarra, CIMA, Pamplona, España; [2]IdiSNA, Instituto de Investigación Sanitaria de Navarra, Pamplona, España; [3] Institute for Smart Cities, Universidad Pública de Navarra, Pamplona, España

gbesne@alumni.unav.es, mvustarroz@unav.es

Resumen

La monitorización ambulatoria es de gran interés para servicios médicos e individuos particulares. Este tipo de tecnología resulta todavía más interesante en el contexto de la pandemia que estamos viviendo, dado que permiten generar redes de monitorización fácilmente escalables. En este trabajo se han identificado elementos para implementar un dispositivo que integre diferentes parámetros fisiológicos de interés en un amplio rango de patologías y se ha validado su utilización para la captura de constantes vitales en la región del oído externo. El prototipo actual integrar cuatro sensores alrededor de un sistema de comunicación inalámbrico que provee de información de temperatura, ritmo cardíaco, acelerometría y oxigenación de forma eficiente y mínimamente intrusiva. En el futuro los esfuerzos se concentrarán en la integración de los componentes en un dispositivo de bajo factor de forma y alta eficiencia energética que pueda colocarse en torno al área validada.

1. Introducción

La monitorización ambulatoria en pacientes crónicos tales cómo pacientes de epilepsia u otras enfermedades neurológicas y mentales, enfermedades cardiovasculares o personas dependientes, es de gran relevancia. Permite llevar a cabo un seguimiento del paciente más detallado y a su vez, ofrece la capacidad de equipar algunos de estos sistemas con alarmas para alertar a cuidadores, familiares o personal sanitario. En el contexto sanitario en el que nos encontramos, un sistema con estas características sería de gran utilidad para llevar a cabo la monitorización y el seguimiento ambulatorio de pacientes afectados por la COVID-19.

Los sistemas comerciales semejantes a las pulseras de actividad proporcionan cierta monitorización, extrayendo ritmo cardiaco, actigrafía y temperatura del paciente. Sin embargo, la posición de estos sistemas en el cuerpo no es la idónea, ya que la medición se ve afectada por la posición de la mano y el movimiento de la misma. Otros sistemas comerciales más fiables carecen de la ergonomía de los primeros, lo que los hace menos deseables, aún más para su uso en niños o personas dependientes. Aun así, el abanico de parámetros fisiológicos que pueden capturar es mucho mayor, proporcionando más información.

En este trabajo trataremos validar la medición de diferentes parámetros fisiológicos utilizando componentes de electrónica de consumo aplicados sobre la región periauricular. Ésta es un área corporal que permite el desarrollo de sistemas con alta estabilidad en las mediciones ante movimientos del paciente y cuya ergonomía y baja intrusividad permite un uso prolongado (incluso durmiendo) dando parámetros de fiabilidad suficientes para asistir en la evaluación clínica.

2. Materiales y métodos

2.1. Sensorización

Tras analizar diferentes patologías y evaluar las posibilidades técnicas y los niveles de obstrucción a la hora de obtener las señales, hemos decidido desarrollar un sistema capaz de adquirir información sobre la temperatura central (Tª), aceleración (Acc), pulsioximetría y electrocardiografía (ECG). Los sensores y electrodos requeridos para cada medición se colocarán en la oreja, más concretamente: el pabellón auricular, el canal auditivo o el área peri-auricular. Lo sensores utilizados para estas mediciones han de ser de bajo coste, fácil adquisición, con un factor de forma reducido, controlables por medio de microcontroladores semejantes a Arduino, de bajo consumo y comunicación digital para una mayor compactación.

Para la aceleración hemos identificado el ADXL345 por su gran versatilidad pudiendo ser programado para múltiples tareas y alarmas. A diferencia del resto de parámetros que se discuten en este apartado no hay ningún requerimiento con respecto al posicionamiento del acelerómetro, con lo que se posicionara allí donde no entorpezca el posicionamiento del resto de sensores.

La medición de la temperatura central se realizará por medio de dos sensores equipados con distintas tecnologías: el sensor de infrarrojos (IR) MLX90615 y el termistor TMP112. El MLX90165 se clocaría en el canal auditivo con el fin de registrar la temperatura timpánica, siendo esta una de las mediciones poco invasivas que mejor reflejan la temperatura central [1]. Este sensor está equipado con un termistor al que se puede acceder. Sin embargo, la implementación de este sensor obstruiría el canal auditivo. Por ello también se registrará la temperatura del canal auditivo por medio del TMP112 cuyo reducido tamaño permite colocarlo en contacto con el canal auditivo sin obstruirlo. A diferencia del resto de dispositivos este sensor no constaba de una placa de desarrollo para su control, por lo que se ha desarrollado una para su testeo.

ISBN: 978-84-09-25491-0, pags. 235- 238

Para la pulsioximetría hemos optado por el MAX30102. Este sensor se basa en la fotopletismografía (PPG) por reflectancia de LEDs rojo e IR. Con la adquisición de estas señales se puede calcular la saturación de oxígeno en sangre periférica (SpO_2) y el ritmo cardíaco (HR) [2,3]. Se han realizado múltiples registros con este sensor para determinar cuál de los posibles puntos de interés es más fiable a la hora de registrar el PPG.

Por último, para el registro del ECG, se utilizará el sensor de ECG de la plataforma Bitalino que permite registrar de forma simultanea múltiples canales analógicos. Estos sistemas se integrarán en un prototipo de desarrollo en el que se puedan ir incluyendo o retirando elementos según el diseño experimental lo requiera.

2.2. Diseño experimental

1. *Registros de ECG y PPG*. Estos registros se realizan colocando simultáneamente sensores de ECG en el pecho replicando la derivación 1-6, otro en el punto de interés y un MAX30102.
2. *Registros múltiples de PPG*. Estos registros se realizan colocando un MAX30102 en el dedo índice y un segundo sensor en el área de interés.
3. *Registros múltiples de temperatura*. El registro simultáneo de los distintos dispositivos y un tercer sensor para la temperatura ambiente determinará la medida en la que las condiciones ambientales afectan al registro. Como referencia, se toma la temperatura corporal por medio de un termómetro digital comercial en la axila al principio del registro.
4. *Registro multimodal*. Registros realizados con los distintos tipos de sensores para estudiar interacciones en distintas situaciones.

3. Resultados

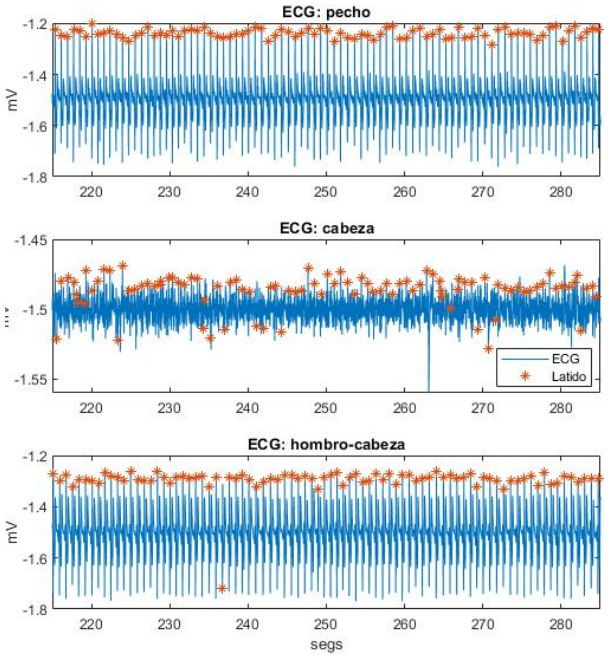


Figura 1 *Registros de ECG a)en el pecho (superior), b) entre ambas orejas (media) y c) entre y el hombro y la oreja contralateral (inferior)*

Los registros de ECG y PPG combinados han mostrado que el sistema propuesto no es capaz de registrar el ECG cuando se colocan los electrodos de oreja a oreja. Sin embargo, como se ve en la Figura 1 para poder registrar el ECG de forma aceptable se han de colocar los electrodos de oreja a hombro contralateral. Por ello, y dado que la detección de la onda R puede obtenerse también con el sensor de PPG, no se ha incluido el ECG en el sistema final.

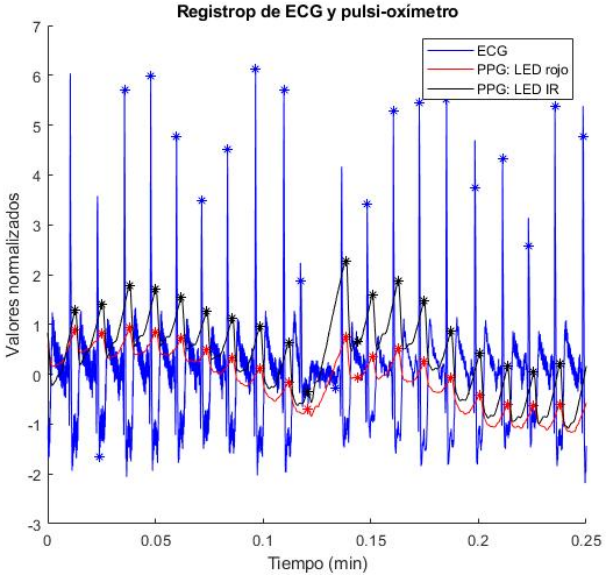


Figura 2 *Sección del registro simultáneo de ECG y PPG de LED rojo e IR. Los asteriscos indican la detección de un latido cardiaco en su registro correspondiente según el código de color*

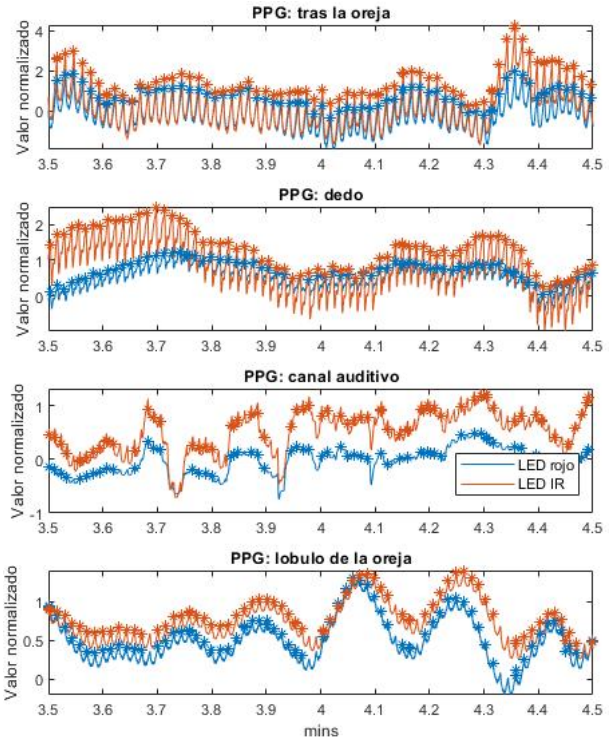


Figura 3 *Registros normalizados del MAX30102 (en orden descendente) sobre la apófisis mastoidea, en la yema del dedo, en el canal auditivo y lóbulo de la oreja. Latidos detectaos indicados con asterisco según el código de color.*

La Figura 2 contiene registros de PPG y ECG en los que se muestra cómo la localización de la onda R ya sea mediante PPG o ECG es coincidente (excepto por el retardo constante que muestra el PPG respecto al EEG) y, por lo tanto, parámetros como el ritmo cardiaco u otras variables relacionadas con la dinámica cardíaca no se verían afectadas por la modalidad de registro.

Los registros de múltiples MAX30102 (Figura 3), demuestran que la región peri-auricular constituye un área válida para la medición con MAX30102, mostrando registros de PPG de calidad adecuada en el pabellón auricular, el lóbulo de la oreja, en el área peri auricular (sobre la apófisis mastoidea), así como en el interior del canal auditivo. Este último registro resulta complicado e incómodo de realizar, debido al tamaño de los componentes que acompañan al sensor en la plataforma de desarrollo para microcontroladores. Sin embargo, estamos trabajando para implementar un nuevo sustrato para el sensor que permitirá introducirlo sin generar molestias ni obstrucción en el canal auditivo.

Por último, el sistema de registro multimodal se ha montado alrededor de una banda de plástico con el contorno ajustable proporcionando una sujeción estable a la cabeza por medio de un sistema de engranajes. En él se han anclado el microcontrolador, la fuente de alimentación y el módulo de comunicación inalámbrica por Bluetooth. Junto a ello, se han unido conectores para los sensores, el acelerómetro y el sensor de temperatura ambiente (colocado en una de las extensiones del armazón para no verse afectado por la temperatura ambiente). Los sensores de temperatura IR y por contacto (MLX90615 y TMP112 respectivamente) se han integrado en un mismo conector, y el MAX30102 en un segundo conector.

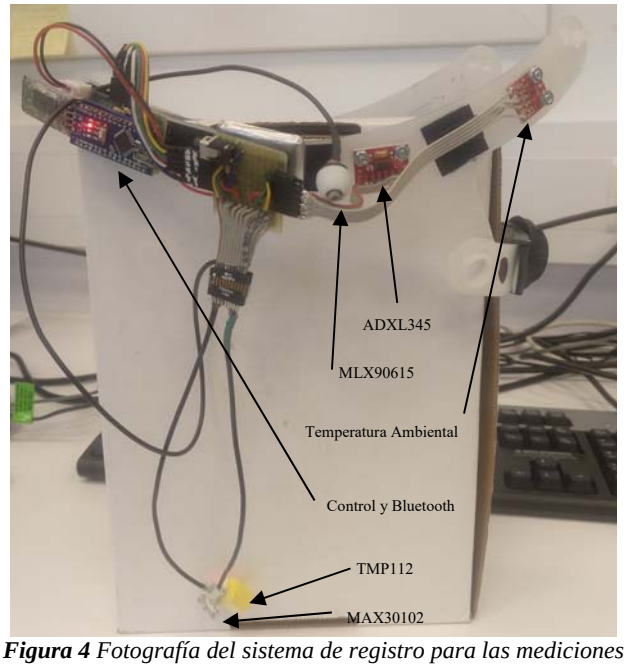


Figura 4 *Fotografía del sistema de registro para las mediciones multimodales con cada uno de los componentes resaltados.*

Con este sistema se han realizado registros para testar la respuesta de estos sensores ante el movimiento del sujeto y a los cambios de temperatura ambiente. Se ha observado que, al realizar actividades como cambio postural, andar o saltar, el patrón de actividad registrado por el acelerómetro difiere (Figura 5).

En estos registros no se ha colocado el MAX30102 en el canal auditivo debido a su factor de forma y se ha registrado en sobre la apófisis mastoidea (Figura 6). Los registros muestran que no se ve afectado significativamente cuando se producen movimientos de poca magnitud como caminar. Los saltos introducen ligeros artefactos en el registro mientras que los cambios posturales introducen grandes artefactos, debido al efecto de la gravedad sobre el sensor y las restricciones de la conexión.

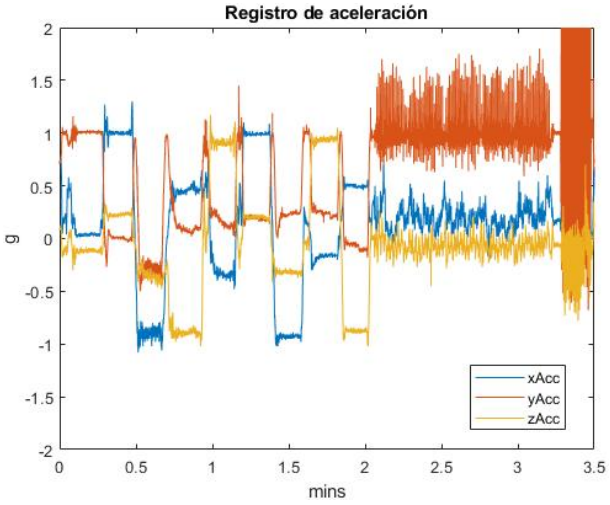


Figura 5 *Registro del acelerómetro. Muestra varios patrones de actividad: cambios posturales (hasta el minuto 2), caminando (del minuto 2 a 3) y saltos, tramo final.*

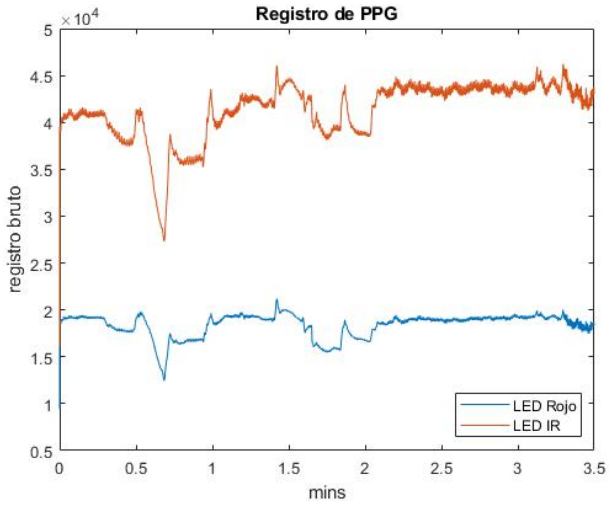


Figura 6 *Registro de PPG mostrando la señal sin procesar del LED rojo e IR. Se observan artefactos debido al movimiento del sujeto y a cambios posturales: cambios posturales (hasta el minuto 2), caminando (del minuto 2 a 3) y saltos, (final)*

Los registros de temperatura han sido llevados a cabo en un en torno a 30ºC y se ha estudiado la respuesta de los distintos sensores y sus correspondencias (Figura 7, superior).

Comparada con la temperatura axilar (tomada al a inicio del registros) de 36.5º se desvía únicamente en 0.2º de la tempera tomando por el sensor IR de 36.3º, confirmando

que es una medición aceptable de la temperatura corporal. La temperatura del termistor del mismo sensor se ve ligeramente afectada por el por la diferencia con la de temperatura ambiente de 30.2º, descendiendo hasta los 35.6º grados, mientras que el TMP112 se ve altamente afectada reduciéndose hasta los 33º.

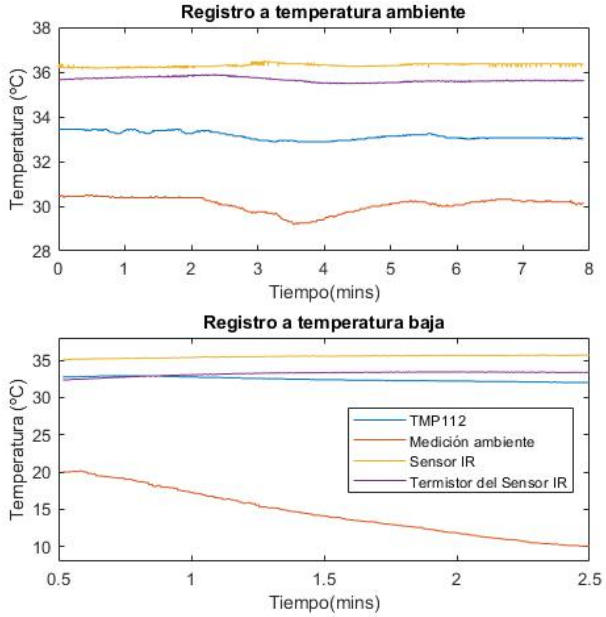


Figura 7 *Registro de temperatura realizado con el sistema integrado en tres áreas diferentes a temperatura ambiente (superior) y a temperatura baja (inferior).*

Estas diferencias se deben a que el sensor IR se ha posicionado a mayor profundidad que el TMP112 debido al factor de la forma del mismo. A diferencia de la medición por IR, ambas medidas se ven afectadas por variaciones en la temperatura ambiente.

Este comportamiento se ve exacerbado tras realizar el mismo registro ante menor temperatura. En la Figura 7, inferior se repite el registro, pero iniciando en una estancia a 20ºC y entrando en una estancia 2ºC, con una temperatura corporal de 36.2ºC. Dado que este registro es de menor duración debido a la exposición a bajas temperaturas, la temperatura IR no se estabiliza hasta el final, alcanzando valores de 36.1ºC. El resto de temperaturas se ven severamente afectadas por la temperatura externa, descendiendo en mayor medida que el registro a 30ºC.

4. Discusión

Esta comunicación presenta un sistema capaz de registrar los parámetros fisiológicos seleccionados en área auricular, a excepción del ECG. Esta circunstancia no es un problema dado que la localización del latido cardiaco también puede estimarse a partir del registro de PPG, incluso mejorándolo si se incluye un LED verde que es más eficaz detectando cambios volumétricos en el flujo sanguíneo

El acelerómetro es capaz de monitorizar los movimientos del sujeto al ser posicionado en la cabeza como se ve en los distintos patrones de actividad. La temperatura registrada por termistores se ve afectada en mayor medida por la temperatura externa a medida que su posición es más superficial, siendo el registro más fiable el realizado con tecnología IR.

El MAX30102 ha sido capaz de realizar los registros en el punto de mayor interés: en canal auditivo. Sin embargo, dado el tamaño de la placa de desarrollo no es posible fijar de forma estable ni ergonómica el sensor a las paredes del canal. Por ese motivo, se observan artefactos por movimiento en los registros comparativos y se decidió registrar sobre la apófisis mastoidea en el registro multimodal. Sin embargo, también se han visto afectado por movimiento, dada la posición del sensor, en un lugar expuesto y no bien fijado. Aun así, los artefactos que han afectado significativamente a la medición se deben a la construcción del sistema, no a la falta de idoneidad del punto de registro. En la actualidad estamos rediseñando dicha placa para que en el futuro el sistema posea una mejor ergonomía y factor de forma que permita introducir el sensor en el canal auditivo, con lo que se espera que estos artefactos desaparezcan.

Por todo ello, concluimos que la monitorización de los parámetros propuestos en el área auricular es factible y provee de información útil, confirmando nuestro interés en dedicar esfuerzos a llevar a cabo la integración de los sensores en un sistema de reducido tamaño se llevará a cabo en el seno de un futuro proyecto.

Agradecimientos

Es trabajo ha sido financiado por el Instituto de Salud Carlos III a través del proyecto “DTS19/00130”. (Co-financiado por *European Regional Development Fund; “A way to make Europe”*).

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 y la Asociación de Amigos de la Universidad de Navarra por financiar mis estudios doctorales.

Referencias

[1] Gasim G, Musa I, Abdien M, Adam I. Accuracy of tympanic temperature measurementusing an infrared tympanic membranethermometer. MCResearchNotes2013,6:194

[2] Sun Y, Thako, N. Photoplethysmography Revisited: From Contact to Noncontact, from Point to Imaging. IEEE Trans. Biomed. Eng. 63, 463–477 (2016).

[3] Allen, J. Photoplethysmography and its application in clinical physiological measurement. Physiol. Meas. 28, R1–R39 (2007).

Señales Biomédicas II

Jueves, 26 de noviembre de 2020

Influence of voluntary contractions on the basal sEMG activity of the pelvic floor muscles

M. Albaladejo-Belmonte[1], M. Tarazona[2], F. J. Nohales[2], J. Alberola-Rubio[3], J. Garcia-Casado[1]

[1] Centro de Investigación e Innovación en Bioingeniería, Universitat Politècnica de València, Valencia, Spain {moalbel, jgarciac}@ci2b.upv.es

[2] Servicio de Ginecología y Obstetricia, Hospital Politècnic i Universitari La Fe, Valencia, Spain, tarazona_marmot@gva.es, franoal@uv.es

[3] Unidad de Bioelectrónica, Procesamiento de señales y Algoritmia, Instituto de Investigación Sanitaria La Fe, Valencia, Spain, pepe_alberola@iislafe.es

Abstract

Chronic pelvic pain (CPP) is a complex clinical condition that affects many women, being sometimes misdiagnosed or mistreated, which can be treated with the infiltration of botulinum toxin (BoNTA). The pelvic floor musculature (PFM) condition from CPP patients can be assessed by means of surface electromyography (sEMG). The evaluation of the basal activity can help to detect a muscular dysfunction, therefore it is important to ensure that the PFM shows a minimum activation when its sEMG is being analysed. In this study, we recorded the sEMG of 25 women with CPP before and 8, 12 and 24 weeks after their treatment with BoNTA while they performed a protocol of 5 voluntary contractions. The root mean square (RMS) and sample entropy (SampEn) of basal segments pre- (B[PRE]), inter- (B[I]) and post- (B[POST]) contractions of the sEMG were computed and normalized according to the minimum (RMS_{norm}) and maximum ($SampE_{norm}$) of the recording. B(PRE) showed the lowest RMS_{norm} median both before and after the treatment with BoNTA, which proved that the activity of the PFM is minimum before the first contraction. As for $SampEn_{norm}$, although results were not so conclusive, they also indicated that B(PRE) should be taken as a reference to analyse the PFM function at its state of minimum activity. Future works aiming to characterize the effects of BoNTA in PFM by means of sEMG should consider basal segments before contractions to assess basal tone conditions.

1. Introduction

Chronic pelvic pain (CPP) is a complex clinical condition with a prevalence between 5.7% and 26.6%, higher for fertile women [1]. It is defined as a recurrent or intermittent pain that lasts more than 6 months and is usually associated to a sexual dysfunction [2]. Due to its complexity, CPP is often misdiagnosed or mistreated, the reason why it is important to manage it from a multidisciplinary scope.

CPP can be treated with surgical or non-surgical methods. In the last case, a possible therapy relies on the infiltration of botulinum neurotoxin type A (BoNTA) in the pelvic floor musculature (PFM). Once injected, the toxin binds to the nerve terminals of a presynaptic motoneuron, blocking the release of acetylcholine and thus synaptic processes. Hence, muscle fibres innervated by that neuron become unable to develop action potentials and, therefore, to contract themselves. Although this is its most well-known mechanism, BoNTA also inhibits the release of substances related to pain and inflammation [3].

CPP is usually associated to a muscular dysfunction, therefore surface electromyography (sEMG) can yield important information about PFM condition [4]. In particular, the dysfunction may be caused by a hypotonicity or a hypertonicity of the PFM at rest [5], so the analysis of the signal acquired while the muscle is relaxed (basal tone) is of great importance, although tests usually include voluntary contractions in the protocol [6]. Many women, even healthy ones, have a poor ability to voluntary contract and relax their PFM when they are asked to [7] and nerves due to the clinical setting and the test can lead to a muscle tension higher than in a normal physiological scenario. Therefore, it is important to ensure that the sEMG signal is analysed while their musculature is minimally activated. On the other hand, vaginal and anal probes are typically used to perform sEMG recordings, since they can detect the activity of the deepest PFM However, intracavitary probes are uncomfortable for patients and not robust to movement artifacts [4]. Our team is aiming to assess the feasibility of monitoring PFM condition and the effect of CPP treatment with BoNTA using adhesive electrodes. These electrodes can be more sensitive to crosstalk of surrounding muscle and hence muscular relaxation can be of greater importance. In this preliminary study, the aim was to determine in which segment/period, of a protocol with voluntary contractions, the PFM shows a state of maximum relaxation (before, between or after contractions).

2. Materials and methods

2.1. Signal acquisition

Twenty-five women with CPP who showed a sexual dysfunction were recruited at Hospital Politècnic i Universitari La Fe (Valencia). They were treated with 80 U.I. of BoNTA (Xeomin, Merz Pharma España S.L.) diluted in 2 mL of lidocaine at 1%, which was infiltrated at a single point in the puborectalis or pubococcygeus. Patients' sEMG was recorded before the infiltration and 8, 12 and 24 weeks later. The skin was gently exfoliated with an abrasive gel (Nuprep, Weaver and Company, USA) and

ISBN: 978-84-09-25491-0, pags. 240- 243

four monopolar signals were registered using six disposable Ag/AgCl electrodes (3M Red Dot 2660-5); four on labia majora (M1, M2, M3, M4) and two on both iliac spines (reference [REF] and ground [GND] electrodes), as shown in *Figure 1*. Signals were band-pass filtered between [3, 1000] Hz with a multipurpose amplifier (15 LT, Grass Technologies, USA) and digitalized at a sampling rate of 10 kHz. Patients were asked to perform 5 maximum voluntary contractions of 5s and resting periods of 10s between them during the recordings.

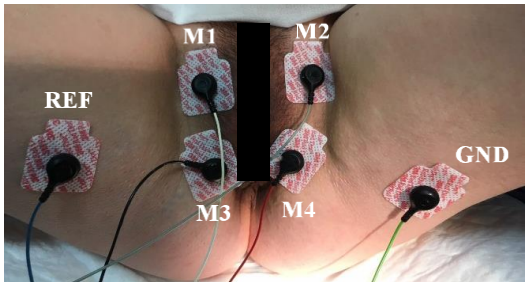


Figure 1. *Electrodes arrangement for the sEMG acquisition.*

2.2. Signal processing

A band-pass (bandwidth = [30, 450] Hz) and a comb filter (50 Hz) were applied to remove the noise and the power interference in the sEMG, respectively. Then, four bipolar signals were obtained: one for each vertical region (right: B1= M1-M3; left: B2 = M2-M4) and horizontal area (upper: B3 = M1-M2; lower: B4 =M3-M4) of the PFM. After that, we manually annotated different basal segments in signals: 1 pre-contractions (B[PRE]), 5 inter-contractions (B[I]) and 1 post-contractions (B[POST]).

When we computed the power spectrum density (PSD) of annotated signal segments, we observed anomalous tonal components in many of them, probably due to electromagnetic interferences at the recording room. We designed an algorithm that detected such tonal interferences and removed them by notch filtering (*Figure 2*).

Two parameters were computed to characterize the basal segments:

- ***Root mean square* (*RMS*):** it measures the power of the signal. Given a discrete time series of N samples, $x[n]$, it is computed as follows:

$$RMS\ (x[n]) = \sqrt{\frac{1}{N} \cdot \sum_{k=1}^{N} x[k]^2} \quad (1)$$

- ***Sample entropy* (*SampEn*):** it quantifies the complexity of the signal. It is calculated as the negative of the natural logarithm of the conditional probability that two series of N samples that are similar at m points keep being similar at the next point, given a tolerance r and without considering self-matchings [8]. To compute *SampEn*, signals were normalized to have zero mean and unit variance and $m = 2$ and $r = 0.15$ were chosen.

In the case of B(I), the value of the 5 segments of the recording was averaged to obtain only one per parameter.

Finally, a statistical analysis was performed to determine if the characteristics of a basal segment significantly vary depending on its relative position to contractions in the signal and in which one the PFM shows a minimal activity. To this end, for each patient, week and bipolar sEMG channel, parameters calculated for B(PRE), B(POST) and B(I) were first normalized. On the one hand, the energy of the sEMG is expected to be minimal with maximal muscle relaxation [9] . The *RMS* computed for B(PRE), B(POST) and B(I) in the same recording was thus divided by the minimum among them, creating the dataset RMS_{norm}. On the other hand, common synaptic input is delivered to multiple motor units of the muscle when it needs to generate force, therefore their action potentials are correlated in time [9]. In other words, the contraction of the PFM involves an increase of the motor units' synchronization, which is the reason why the value of *SampEn* of the three basal segments was divided by the maximum among them, obtaining the dataset $SampEn_{norm}$. After that, parameters from the 4 bipolar channels were grouped together and Kruskal-Wallis test ($\alpha = 0.05$) was performed with the alternative hypothesis that the median of RMS_{norm} (or $SampEn_{norm}$) of a basal segment at a given week depends on its relative position to contractions in the sEMG. An additional multiple comparison test based on the Tukey's honestly significant difference procedure was performed to find medians that are significantly different from each other.

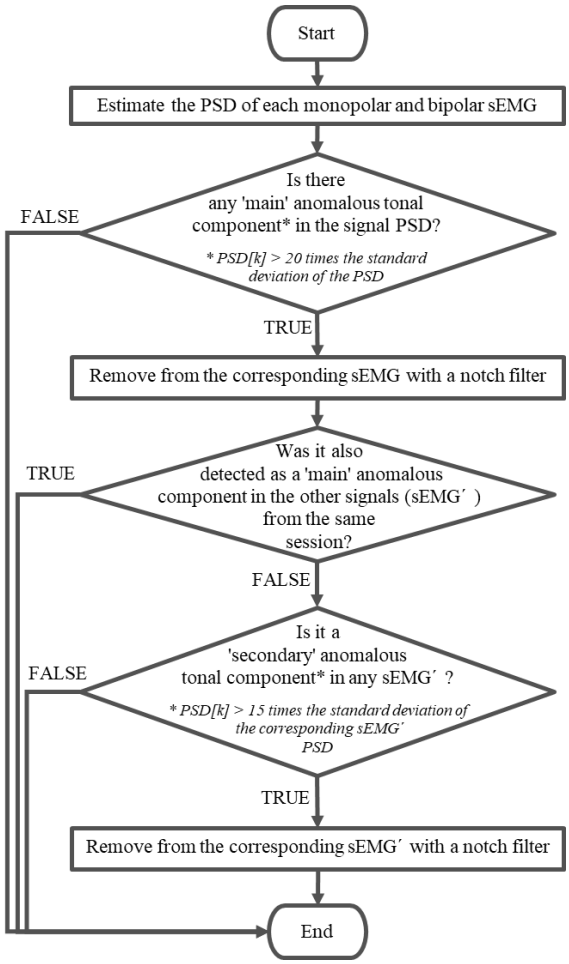


Figure 2. *Flow chart of the algorithm to remove anomalous tonal components from the sEMG.*

3. Results

Figure 3 depicts a patient's bipolar sEMGs, where different basal and contractile segments have been highlighted.

Figures 4-5 display the Box-Whisker plot of RMS_{norm} and $SampEn_{norm}$ datasets for each week of the monitoring, with B(PRE), B(POST) and B(I) as grouping variables. In *Table 1* their medians and 25 and 75 percentiles are shown.

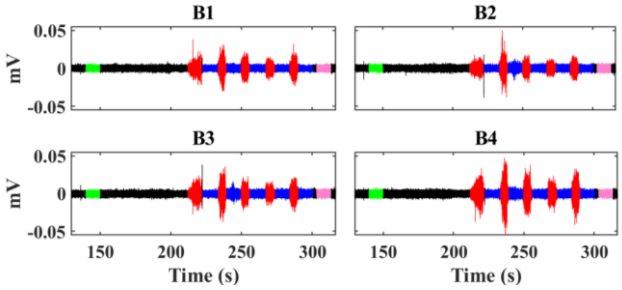


Figure 3. *sEMG of the right (B1), left (B2), upper (B3) and lower (B4) region of the PFM. Segments: contractions (red) and basal segments before (green), between (blue) and after them (pink).*

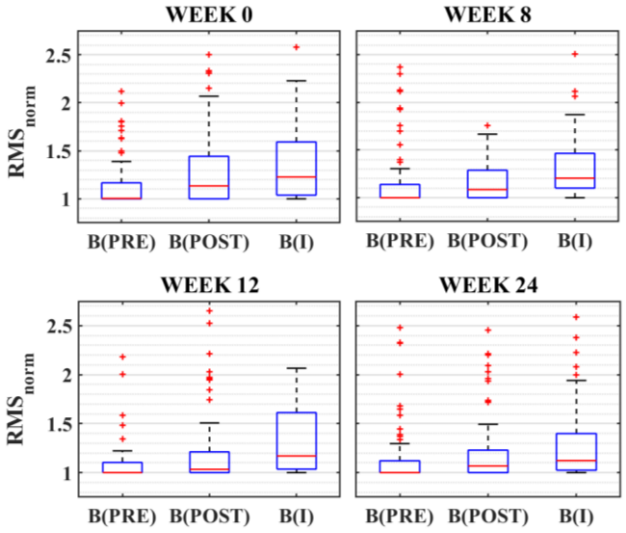


Figure 4. *Box-Whisker plot of RMS_{norm} in B(PRE), B(POST), B(I) for each week of the recording.*

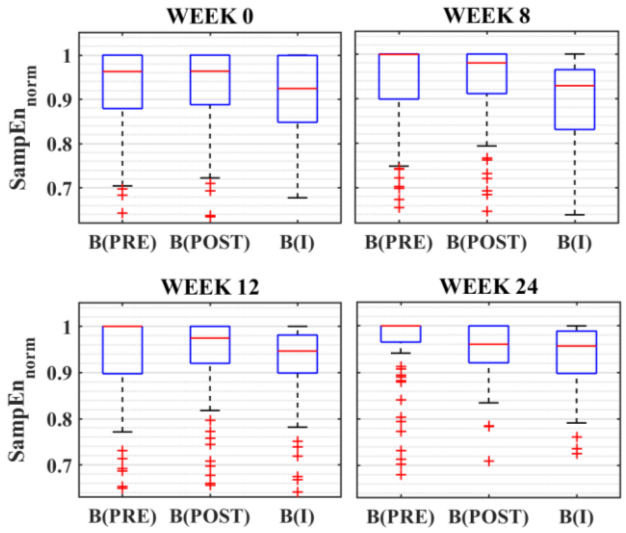


Figure 5. *Box-Whisker plot of $SampEn_{norm}$ in B(PRE), B(POST), B(I) for each week of the recording.*

Due to the presence of some pronounced outliers, it is difficult to visually determine from *Figures 4-5* the box of which basal segment is below or above the others at a given week. However, it can be appreciated in *Table 1* that the median and 75 percentile of RMS_{norm} are lower in B(PRE) than in B(POST) and B(I), both before (W0) and after (W8, W12, W24) the treatment with BoNTA. As for $SampEn_{norm}$, its median and percentiles are very similar between B(PRE) and B(POST) in the different recording weeks, while their values seem to be lower in the case of B(I).

Table 1. *Median (25-75 percentiles) of RMS_{norm} and $SampEn_{norm}$ for each basal segment and week (W).*

	W	B(PRE)	B(POST)	B(I)
RMS_{norm}	0	1.00 (1.00-1.17)	1.13 (1.00-1.44)	1.23 (1.04-1.59)
	8	1.00 (1.00-1.14)	1.09 (1.00-1.29)	1.21 (1.10-1.46)
	12	1.00 (1.00-1.10)	1.03 (1.00-1.21)	1.17 (1.04-1.61)
	24	1.00 (1.00-1.12)	1.07 (1.00-1.23)	1.12 (1.02-1.40)
$SampEn_{norm}$	0	0.96 (0.88-1.00)	0.96 (0.89-1.00)	0.92 (0.82-1.00)
	8	1.00 (0.90-1.00)	0.98 (0.91-1.00)	0.93 (0.83-0.97)
	12	1.00 (0.90-1.00)	0.97 (0.92-1.00)	0.95 (0.90-0.98)
	24	1.00 (0.97-1.00)	0.96 (0.92-1.00)	0.96 (0.90-0.99)

P-values from Kruskal-Wallis tests were lower than 0.05, except for $SampEn_{norm}$ at week 0, so it cannot be assumed that data from B(PRE), B(POST) and B(I) belong to the same distribution. Results provided by multiple pairwise comparison tests are displayed in *Table 2*. As for RMS_{norm}, there is a significant difference between the median of B(PRE) and the other groups before and after the treatment. In contrast, there are no statistical differences between any pair of basal segments at week 0 when $SampEn_{norm}$ is studied, while they appeared weeks later in B(PRE) vs. B(I) and to a lesser extent in B(PRE) vs. B(POST) and B(POST) vs. B(I).

Table 2. *P-value of pairwise comparison test at week W for RMS_{norm} and $SampEn_{norm}$. P-values<0.05 are shown in grey.*

	Segments compared	W0	W8	W12	W24
RMS_{norm}	B(PRE) vs. B(POST)	0.01	0.01	0.04	0.02
	B(PRE) vs. B(I)	<0.01	<0.01	<0.01	<0.01
	B(POST) vs. B(I)	0.14	<0.01	<0.01	0.13
$SampEn_{nor}$	B(PRE) vs. B(POST)	0.98	0.79	0.32	<0.01
	B(PRE) vs. B(I)	0.25	<0.01	<0.01	<0.01
	B(POST) vs. B(I)	0.18	<0.01	0.08	0.41

4. Discussion

It has been observed that the sEMG segment chosen to monitor the basal tone of the PFM may significantly influence the results obtained. We found that *RMS* value is significantly lower in basal segments prior to contractions than in between or after them, regardless the week of the recording (before and after the infiltration of BoNTA). This means that, if the energy of the sEMG is the criterion of evaluation, it can be concluded that the PFM is more relaxed at the beginning of the recording. Therefore, B(PRE) should be the interval selected to monitor the basal tone of the muscle in CPP before and after the treatment. In contrast, when the criterion used is the complexity of the sEMG, the selection is not so clear. The number of significant differences between groups is not so widespread on weeks, especially on week 0. However, the median of $SampEn_{norm}$ in B(PRE) is greater than in B(I) throughout the entire post-treatment period and than B(POST) at week 24. Therefore, we can consider that the basal segment of the sEMG before the contractions is also suitable to analyse the activity of the PFM at its maximum rest according to signal complexity criterion. The reason why the activity of the PFM seems to be higher in sEMG intervals between and after contractions could be that women cannot relax completely their musculature and they may need more time to return to a state of maximum rest after a voluntary activation.

On the other hand, we obtained a great number of outliers when we plotted Box-Whisker diagrams. Considering that the *RMS* and *SampEn* of each basal segment was normalized according to other segment of its same recording, the presence of so many anomalous data manifests the lack of control that some women have over their PFM. They were unable to maintain a maximum state of relaxation throughout the entire acquisition when they were not performing a contraction.

Results of this work do not agree with those by Naess & Bø [10]. They observed that women's intravaginal pressure at rest significantly diminished after the performance of three maximum voluntary contractions in both case and control groups. In the present study, the higher *RMS* values of sEMG obtained after the contractions would be expected to lead to a higher intravaginal pressure, which was not measured though. However, their study was performed on nulliparous women who were 24.3 ± 4.7 years old, had vestibulodynia and were monitored with anal probes. In contrast, women recruited in this work were 43.1 ± 9.3 years old, only one of them was nulliparous and signals were acquired with external adhesive electrodes, so results cannot be directly extrapolated.

Some sEMG segments were discarded due to the presence of intermittent artifacts, so further efforts should be made to improve the quality of recordings. In future work, changes in sEMG signals after BoNTA infiltration will be studied to assess their value to monitor CPP treatment.

5. Conclusions

The characteristics of the PFM sEMG at rest, both before and after the treatment of CPP with BoNTA, vary significantly depending on whether the patient has previously performed a contraction or not. In particular, the signal power is lower and its complexity is usually higher before the voluntary activation of the muscle, which proves that the PFM shows a minimum activity in that period of the sEMG recordings. Therefore, patients' basal tone should be evaluated before carrying out any contraction of the PFM at each clinical follow-up session.

Acknowledgments

This study was funded by ISCIII, MCIU, VLC Campus in Convocatoria Ayudas: UPV-La Fe (INBIO): 2016 SPEHG (ID:C18), 2019 sEMG_BONTAv (ID:C06) and with funds from private contracts with Merz Pharma España S.L.

References

[1] Ahangari A. Prevalence of chronic pelvic pain among women: An updated review. *Pain Physician*, vol. 17, no. 2, 2014, pp.E141–E147, .

[2] Doggweiler R *et al.* A standard for terminology in chronic pelvic pain syndromes: A report from the chronic pelvic pain working group of the international continence society. *Neurourology and urodynamics*, vol. 36, no. 4, 2017, pp.984–1008, doi: 10.1002/nau.23072.

[3] Oh HM, Chung ME. Botulinum toxin for neuropathic pain: A review of the literature. , *Toxins*, vol. 7, no. 8, 2015. pp.3127–3154, , 2015, doi: 10.3390/toxins7083127.

[4] Keshwani N, McLean L. State of the art review: Intravaginal probes for recording electromyography from the pelvic floor muscles. , *Neurourology and Urodynamics*, vol. 34, no. 2, 2015. pp.104–112, , 2015, doi: 10.1002/nau.22529.

[5] Cheong YC, Smotra G, Williams ACDC. Non-surgical interventions for the management of chronic pelvic pain. *Cochrane Database of Systematic Reviews*, vol. 2014, no. 3, 2014, doi: 10.1002/14651858.CD008797.pub2.

[6] Morrissey D, El-Khawand D, Ginzburg N, Wehbe S, O'Hare P, Whitmore K. Botulinum Toxin A Injections into Pelvic Floor Muscles under Electromyographic Guidance for Women with Refractory High-Tone Pelvic Floor Dysfunction: A 6-Month Prospective Pilot Study. *Female Pelvic Medicine and Reconstructive Surgery*, vol. 21, no. 5, 2015, pp.277–282, doi: 10.1097/SPV.0000000000000177.

[7] Henderson JW, Wang S, Egger MJ, Masters M, Nygaard I. Can women correctly contract their pelvic floor muscles without formal instruction?. *Female Pelvic Medicine and Reconstructive Surgery*, vol. 19, no. 1, 2013, pp.8–12, doi: 10.1097/SPV.0b013e31827ab9d0.

[8] Richman JS, Moorman JR. Physiological time-series analysis using approximate and sample entropy. *American Journal of Physiology - Heart and Circulatory Physiology*, vol. 278, no. 6 47-6, 2000, doi: 10.1152/ajpheart.2000.278.6.h2039.

[9] Farina D, Negro F, Muceli S, Enoka RM. Principles of motor unit physiology evolve with advances in technology. , *Physiology*, vol. 31, no. 2, 2016. pp.83–94, , 2016, doi: 10.1152/physiol.00040.2015.

[10] Naess I, Bø K. Can maximal voluntary pelvic floor muscle contraction reduce vaginal resting pressure and resting EMG activity?. *International Urogynecology Journal*, vol. 29, no. 11, 2018, pp.1623–1627, doi: 10.1007/s00192-018-3599-1.

Decodificación no lineal de los movimientos de la mano en tiempo real mediante un sistema Brain–Computer Interface

V. Martínez-Cagigal[1], R. J. Kobler[2], V. Mondini[2], R. Hornero[1], G. R. Müller-Putz[2]

[1] Grupo de Ingeniería Biomédica, E.T.S.I. Telecomunicación, Universidad de Valladolid, Valladolid, España

[2] Institute of Neural Engineering, Graz University of Technology, Graz, Austria

Resumen

La decodificación de los movimientos de las extremidades superiores a través de la actividad neuronal ya es una realidad en registros invasivos, mientras que la utilidad del electroencefalograma (EEG) para este mismo propósito aún no está clara. Algunos estudios han conseguido decodificar la posición y la velocidad usando métodos lineales. Sin embargo, las señales obtenidas exhibían amplitudes inferiores a los movimientos efectuados, afectando la realimentación del sistema. En un estudio preliminar demostramos la capacidad de un decodificador no lineal para combinar información direccional (e.g., velocidad) y no direccional (e.g., rapidez), resolviendo la variación de amplitud de forma offline. El objetivo de este estudio es evaluar si la decodificación no lineal de los movimientos de la mano del usuario puede realizarse en tiempo real. Para ello, se pidió a cinco sujetos de control que persiguiesen un objetivo en movimiento con un brazo robótico. Inicialmente, el control del brazo dependía de los movimientos de la mano, sustituyéndose por la decodificación EEG gradualmente. Las correlaciones entre los movimientos reales y decodificados superaron el umbral de aleatoriedad. Los resultados sugieren que la información no direccional también se codifica en el EEG, demostrando la capacidad del método para decodificar los movimientos de las extremidades superiores en tiempo real.

1. Introducción

Los sistemas *brain–computer interface* (BCI) tradicionalmente se han basado en la clasificación de diversas señales de control (e.g., potenciales evocados, ritmos sensoriomotores) para detectar las intenciones del usuario [1]. No obstante, durante los últimos años se ha investigado la posibilidad de decodificar directamente los movimientos de las extremidades utilizando registros invasivos [2]. Es natural, por tanto, preguntarse si los registros no invasivos como el electroencefalograma (EEG) codifican también esta información.

Estudios previos han demostrado que la información cinemática de las extremidades superiores se codifica en la banda de baja frecuencia del EEG para movimientos ejecutados [3], observados [4], [5] e imaginados [6]. Esto hace posible la decodificación de las posiciones y velocidades usando algoritmos lineales, como la regresión de mínimos cuadrados parciales (PLS, *Partial Least Squares*) [4], [6] o los filtros de Kalman [5], [7], [8]. Los primeros estudios se centraron en evaluar la posibilidad de decodificar esta información en el EEG [3], [6] y estudiar la localización de las fuentes corticales [4]. Un estudio reciente desarrolló un sistema BCI online para controlar un brazo robótico usando un filtro de Kalman lineal basado en modelos cinemáticos [8]. Aunque se demostró que la decodificación en tiempo real es posible, se observó una discordancia entre las amplitudes de los movimientos decodificados y los reales. Este problema también se vio reflejado en estudios anteriores [4], [9], poniendo de manifiesto la incapacidad de los modelos lineales para extraer dicha magnitud. Sin embargo, en un estudio preliminar demostramos que las cinemáticas direccionales (e.g., velocidad) y no direccionales (e.g., rapidez) aparecen codificadas en el magnetoencefalograma [10] y en el EEG [5], permitiendo su decodificación mediante métodos no lineales, como el *unscented Kalman filter* (UFK) [11].

El objetivo de este estudio es diseñar, desarrollar y evaluar un sistema BCI para controlar un brazo robótico mediante la decodificación de los movimientos de la mano derecha del usuario en tiempo real. El sistema hará uso del UKF para decodificar no solo las posiciones y velocidades, sino también la rapidez instantánea, con el objetivo de mantener la amplitud real de los movimientos.

2. Materiales y métodos

2.1. Sujetos

Cinco sujetos de control diestros participaron en el estudio (edad media: 28.2 ± 2.4 años; 2 hombres, 3 mujeres). Todos ellos dieron su consentimiento informado para participar en el mismo y fueron remunerados por el *Institute of Neural Engineering* (INE) de la Universidad Tecnológica de Graz (Austria).

2.2. Experimento y paradigma

La instalación del experimento se muestra en la Figura 1(a). El usuario se sienta cómodamente en frente de una pantalla donde se muestra un punto blanco que sigue una trayectoria. El sistema se compone de tres dispositivos: (1) el sistema de adquisición EEG, que registra la actividad cerebral; (2) el LeapMotion (LM, *LeapMotion Inc.*, USA), que monitoriza la posición de la mano del usuario; y (3) el brazo robótico (JACO, *Kinova Robotics Inc.*, *Canada*).

La tarea del usuario era seguir la trayectoria del punto con la vista y con el robot. Tal y como indica la Figura 1(b), el JACO se controlaba inicialmente con los movimientos de la mano capturados por el LM. Este control se sustituyó progresivamente por la decodificación EEG a lo largo de la sesión de evaluación hasta el punto de controlarse plenamente por la señal EEG. El decodificador se calibró con 5 bloques (100% LM, 0% EEG) y se testeó con los bloques 6–7 (66 % LM, 33% EEG), 8–9 (33% LM, 66% EEG) y 10–11 (0% LM, 100% EEG). Cada bloque estaba compuesto por 10 *trials* (i.e., trayectorias), de

ISBN: 978-84-09-25491-0, pags. 244- 247

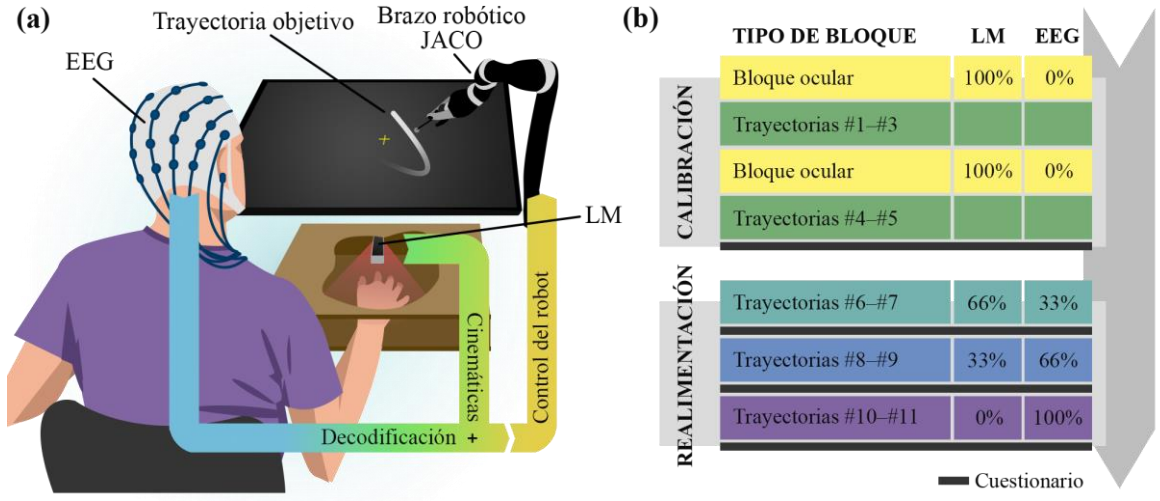


Figura 1.*(A) Instalación del experimento. El usuario se sienta frente a una pantalla que muestra una trayectoria que deberá seguir con el brazo robótico JACO. (B) Paradigma del estudio. Durante la calibración, el JACO se controla únicamente mediante LM. Durante la realimentación, el control mediante LM se reduce hasta que el sistema se controle enteramente por el EEG.*

aproximadamente 24 s cada uno. Entre los bloques de calibración se realizaron 2 bloques adicionales enfocados a registrar movimientos oculares [12]. Después de cada condición (100%, 66%, 33% y 0% LM), los usuarios rellenaron un cuestionario para recoger sus impresiones. Los *trials* se abortaban si el LM detectaba movimientos alejados a más de 12.5cm del punto central, evitando así artefactos musculares debido a la tensión en los hombros. Debido a esta restricción, se mapearon las posiciones reales a movimientos más amplios en el JACO, a escala 1:2. Las 96 posibles trayectorias se generaron *offline* y fueron las mismas para todos los usuarios. Doce de ellas se generaron mediante muestreo de ruido rosa filtrado entre 0.2–0.4 Hz. El resto eran versiones rotadas (90º, 180º y 270º) y volteadas del conjunto inicial. De esta manera se evitó cualquier correlación entre las posiciones y velocidades horizontales y verticales de todas las trayectorias [4].

El EEG se registró empleando 60 electrodos activos referenciados en el lóbulo de la oreja izquierda y usando Fpz como tierra. Se registraron 4 canales de electrooculografía (EOG) adicionales, tal y como muestra la Figura 2(a). Todos ellos se conectaron a un equipo BrainAmp (*Brain Products GmbH, Alemania*) con una frecuencia de muestreo de 500 Hz.

2.3. Procesado de señal

Como se puede observar en la Figura 2(b), las señales EEG se filtraron tanto frecuencialmente (0.18–1.5 Hz) como espacialmente (CAR, *common average filtering*) [8], y se diezmaron a 20 Hz. Se corrigieron las derivas y las posibles desconexiones de los electrodos mediante *high-variance electrode artifact removal* (HEAR) [13]. Los pestañeos y los artefactos oculares se corrigieron con el algoritmo *sparse generalized eye artifact subspace subtraction* (SGEYESUB) [12]. Posteriormente, PLS proyectó los datos EEG [5], [6] y se decodificó la señal con el método no lineal *square-root UKF* (SQ-UKF) [11]. Paralelamente, a la señal LM se le aplicó un filtro paso-bajo a 4 Hz (para reducir ruido ocasional) y se diezmó a 20 Hz. Finalmente, se ponderaron las señales y se escalaron (1:2) para adecuarse al movimiento del JACO sobre la pantalla.

2.4. Decodificación de los movimientos

La decodificación tenía como objetivo estimar el estado del movimiento de la mano del usuario mediante el análisis del EEG. En total, se decodificaron 5 cinemáticas: las posiciones p_x y p_y, las velocidades v_x y v_y, y la rapidez $\rho = \sqrt{v_x^2 + v_y^2}$ (incluida para obtener información sobre el rango de amplitudes de los movimientos de la mano).

La regresión PLS se empleó tanto para reducir la colinealidad de la señal EEG como para mantener la información cinemática. PLS permite obtener una estimación inicial encontrando la relación entre las señales EEG y LM. En este caso, manteniendo un total de $N = 40$ componentes (explica el 70% de la varianza) [4], [8]. El resultado se puede expresar como $\hat{\boldsymbol{L}} = \boldsymbol{EW}$, donde $\boldsymbol{E} \in \mathbb{R}^{\boldsymbol{n}\times\boldsymbol{m}}$ denota la señal EEG del periodo de calibración en el instante actual (n muestras) con muestras pasadas ($m = 64$ canales × 7 retrasos $\in \{-300{:}\,50{:}\,0\}$ ms) y $\boldsymbol{W} \in \mathbb{R}^{\boldsymbol{m}\times\boldsymbol{N}}$ es la matriz de pesos. Para ello, se empleó la implementación SIMPLS. Posteriormente, se aplicó SQ-UKF, un método no lineal basado en el filtro de Kalman, para mejorar la estimación $\hat{\boldsymbol{L}}$, que se supone ruidosa. Los parámetros iniciales se ajustaron a los datos de calibración usando una aproximación *encoding* [5]:

$$\boldsymbol{F} = C_{x_k x_{k-1}}(C_{x_k x_k})^{-1},$$

$$\boldsymbol{Q} = C_{\epsilon_q \epsilon_q}, \text{ donde } \boldsymbol{\epsilon_q} = \boldsymbol{F}\boldsymbol{x}_{k-1} - \boldsymbol{x}_k,$$

$$h(\boldsymbol{x}_k) = C_{\hat{l}_k x_k^+}(C_{x_k^+ x_k^+})^{-1}\boldsymbol{x}_k^+$$

$$\boldsymbol{R} = C_{\epsilon_r \epsilon_r}, \text{ donde } \boldsymbol{\epsilon_r} = h(\boldsymbol{x}_k) - \hat{\boldsymbol{l}}_k,$$

donde $\boldsymbol{F}$ es la matriz de transición de estados, Q la covarianza del error *a priori*, $h(\cdot)$ el modelo de observación, $\boldsymbol{R}$ la covarianza del ruido de las mediciones, y $C_{\boldsymbol{a},\boldsymbol{b}}$ denotaría la matriz de covarianza entre las señales $\boldsymbol{a}$ y $\boldsymbol{b}$. El estado en el momento k se indica como $\boldsymbol{x}_k = [p_x, v_x, p_y, v_y]_k^T$, convirtiéndose en $\boldsymbol{x}_k^+ = [\boldsymbol{x}, \rho\,]_k^T$ con la rapidez. La no-linearidad se incluye solamente en la ecuación de la medición, mientras que la expresión de la estimación *a priori* se mantiene lineal. SQ-UFK retorna

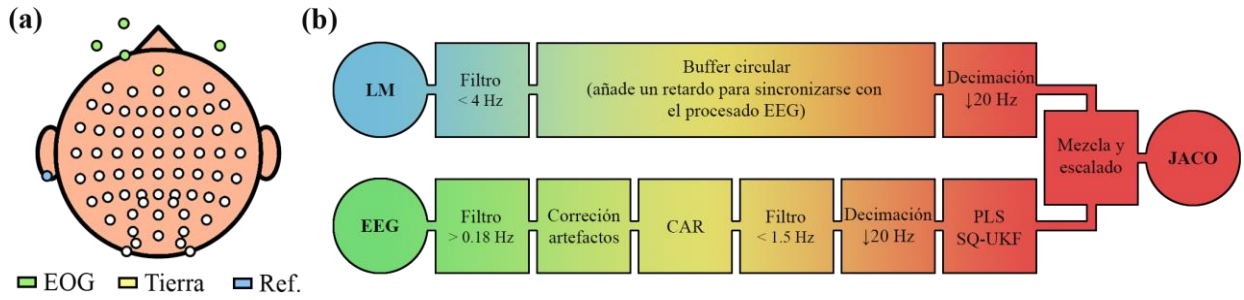


Figura 2.*(A) Distribución de los 64 canales empleados, incluyendo los 4 EOG, la tierra y la referencia. (B) Procesado de señal empleado sobre las señales LM y EEG para decodificar los movimientos y controlar el robot JACO.*

una estimación de los estados $\hat{\boldsymbol{x}}_k$ en tiempo real, que se usa posteriormente para controlar el robot JACO. En la condición de 100% LM empleamos una validación cruzada para simular la decodificación EEG en la etapa de calibración (bloque de trayectorias #1–#5).

3. Resultados

En la Figura 3(a) se muestran las correlaciones entre los estados decodificados $\hat{\boldsymbol{x}}_k$ y los movimientos LM reales, incluyendo el umbral de aleatoriedad. Además, se muestra un *trial* representativo para comparar visualmente la amplitud y forma de los movimientos LM y las decodificaciones PLS y SQ-UKF (Figura 3(b)). La media del ratio entre los movimientos LM y los decodificados fue 1.07 ± 0.09, indicando un desajuste de amplitud nimio. Los resultados de los cuestionarios de satisfacción se resumen en la Figura 3(c). Los usuarios indicaron que la tarea era interesante e intuitiva, y que no les resultó exigente ni física ni mentalmente. Como se puede observar, los usuarios prestaron atención principalmente a la trayectoria de la pantalla, seguido por sus propias manos y, finalmente, al robot JACO. Tal y como se esperaba, el nivel de control percibido decayó a medida que se dio más peso a la señal EEG. Algunos usuarios indicaron que sintieron el control incluso en la condición de 100% EEG.

4. Discusión y conclusiones

Las correlaciones entre los movimientos decodificados y los reales sugieren que la decodificación en tiempo real es posible. Las correlaciones más altas se encontraron en la condición 100% LM (p_x: 0.63, v_x: 0.68, p_y: 0.51, v_y: 0.55), que decrecieron a medida que el control mediante EEG ganaba importancia. Sin embargo, las correlaciones para el control 100% EEG (p_x: 0.39, v_x: 0.48, p_y: 0.35, v_y: 0.44) indican que las cinemáticas aún pueden decodificarse en un sistema BCI controlado enteramente por el EEG. Nótese que el modelo se entrenó con la condición 100% LM, alcanzando una correlación media de 0.45. Es natural que los usuarios empezasen a compensar los errores de la realimentación ofrecida por el JACO en los bloques sucesivos, causando una ligera disminución del rendimiento. Además, la inclusión del JACO implicó un retraso adicional entre los movimientos de la mano y el robot que, a pesar de ser el mismo en todas las condiciones, fue percibido por los usuarios como mayor a medida que aumentó el control basado en EEG. En general, las correlaciones son comparables a las obtenidas en el estudio *offline* previo [5], y el ratio de amplitudes mejora en comparación con los resultados reportados en el estudio *online* con un filtro de Kalman lineal [8].

Debido a la naturaleza de la tarea; i.e., una banda de frecuencias muy bajas y estrecha, la posibilidad de observar correlaciones espúreas es alta. Por esa razón, el límite superior de las correlaciones obtenidas aleatoriamente ($\alpha = 0.05$) ha sido calculado para asegurar que la decodificación es real. También cabe destacar que, mientras que SQ-UKF es ideal para este tipo de tareas, es de esperar una disminución del rendimiento en una tarea de "movimiento libre", donde las posiciones y las velocidades no tengan media nula o posean una varianza mayor.

A pesar de los resultados positivos, observamos que las correlaciones del eje vertical (Y) son inferiores a las obtenidas para el eje horizontal (X). Este decaimiento influye especialmente en las condiciones de 33% y 0% LM, causando correlaciones cercanas al umbral de aleatoriedad. La razón podría deberse a la instalación del experimento. Tal y como muestra la Figura 1(a), la pantalla se inclinó hacia el eje Y para favorecer los movimientos del robot. Esto provoca una percepción de los movimientos verticales ambigua en comparación con los horizontales. Es decir, para que el usuario perciba un desplazamiento vertical de Δd, JACO necesita desplazarse $d = \Delta d/\sin(\theta)$, donde $\theta \sim 45º$ es el ángulo de inclinación de la pantalla (i.e., para el eje horizontal, $d = \Delta d$). Este desajuste podría influenciar la percepción de la trayectoria y dificultar la codificación de los movimientos verticales en el EEG. No obstante, es necesario realizar experimentos adicionales para determinar la causa real de este fenómeno, que no fue observado anteriormente [8].

Una de las desventajas principales de los estudios previos con modelos lineales era el desajuste en la amplitud, causando amplitudes decodificadas más bajas que las reales [8]. Tal y como se muestra en la Figura 3(b) y en [5], este problema se ha aliviado mediante la inclusión de la rapidez en los estados del filtro de Kalman, obteniendo un ratio de amplitud medio de 1.07. Aunque se asume que la rapidez se codifica linealmente en la señal EEG, la relación con respecto a las velocidades no es lineal. Por tanto, un enfoque no lineal como SQ-UKF ha demostrado ser adecuado para combinar dicha información y mejorar la estimación general [8].

Con respecto a los resultados de los cuestionarios, los usuarios se mostraron satisfechos con el experimento. También reportaron que pudieron sentir el control del JACO incluso en las condiciones 33% y 0% LM.

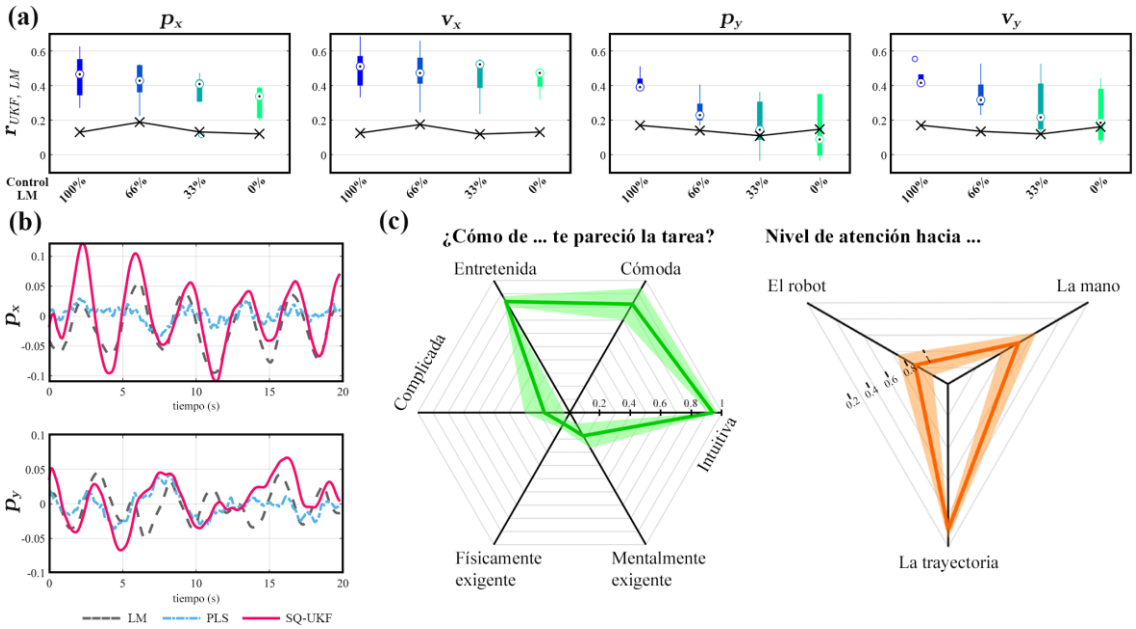


Figura 3.*(A) Correlaciones entre la decodificación SQ-UKF y los movimientos LM reales para cada condición y variable (el umbral de aleatoriedad se representa como una línea negra). (B) Trial que muestra la señal LM y las decodificaciones PLS y SQ-UFK. (C) Resultados normalizados de los cuestionarios en relación con el experimento (izquierda) y el nivel de atención (derecha).*

Asimismo, indicaron que la tarea era entretenida, intuitiva y fácil, mostrándose dispuestos a participar en futuros estudios. La estrategia general de los usuarios fue prestar la mayor atención a la trayectoria, seguido de sus manos y, finalmente, en el robot. Mediante esta estrategia pudieron ignorar posibles errores en la realimentación del JACO para las condiciones de 66%, 33% y 0% LM, evitando así los movimientos de compensación.

Los resultados del estudio no solo demuestran que la información sobre el movimiento de las extremidades superiores se encuentra codificado en el EEG, sino que la información sobre la rapidez también está presente y puede extraerse en tiempo real para mejorar la decodificación. En conclusión, el método no lineal SQ-UKF es viable para decodificar dichos movimientos en el EEG.

Agradecimientos

Este estudio se ha financiado mediante el CIBER-BBN del Instituto de Salud Carlos III y FEDER, así como los proyectos 'Feel Your Reach' del European Research Council (Consolidator Grant 681231), 'DPI2017-84280-R' del MCIU y FEDER, y '0378_AD_EEGWA_2_P' (POCTEP 2014-2020) de la Comisión Europea y FEDER. V. Martínez-Cagigal cuenta con una ayuda FPI-UVa.

Referencias

[1] G. R. Müller-Putz, A. Schwarz, J. Pereira, and P. Ofner, "From classic motor imagery to complex movement intention decoding: The noninvasive Graz-BCI approach," *Prog. Brain Res.*, vol. 228, pp. 39–70, 2016.

[2] L. R. Hochberg *et al.*, "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm," *Nature*, vol. 485, no. 7398, pp. 372–375, 2012.

[3] T. J. Bradberry, R. J. Gentili, and J. L. Contreras-Vidal, "Reconstructing three-dimensional hand movements from noninvasive electroencephalographic signals," *J. Neurosci.*, vol. 30, no. 9, pp. 3432–3437, 2010.

[4] R. J. Kobler, A. I. Sburlea, and G. R. Müller-Putz, "Tuning characteristics of low-frequency EEG to positions and velocities in visuomotor and oculomotor tracking tasks," *Sci. Rep.*, vol. 8, no. 1, pp. 1–14, 2018.

[5] R. J. Kobler, A. I. Sburlea, V. Mondini, H. Masayuki, and G. R. Müller-Putz, "Distance- and speed-informed kinematics decoding improves M/EEG based upper-limb movement decoder accuracy," *J. Neural Eng.*, vol. In press, 2020.

[6] P. Ofner and G. R. Müller-Putz, "Using a noninvasive decoding method to classify rhythmic movement imaginations of the arm in two planes," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 3, pp. 972–981, 2015.

[7] J. Lv, Y. Li, and Z. Gu, "Decoding hand movement velocity from electroencephalogram signals during a drawing task," *Biomed. Eng. Online*, vol. 9, pp. 1–21, 2010.

[8] V. Mondini, R. J. Kobler, A. I. Sburlea, and G. R. Müller-Putz, "Continuous low-frequency EEG decoding of arm movement for closed-loop, natural control of a robotic arm," *J. Neural Eng.*, vol. 17, no. 4, p. 046031, 2020.

[9] J. M. Antelis, L. Montesano, A. Ramos-Murguialday, N. Birbaumer, and J. Minguez, "On the Usage of Linear Regression Models to Reconstruct Limb Kinematics from Low Frequency EEG Signals," *PLoS One*, vol. 8, no. 4, 2013.

[10] R. J. Kobler, H. Masayuki, and H. Hashimoto, "Simultaneous decoding of velocity and speed during executed and observed tracking movements: a MEG study," in *Proceedings of the 8th Graz BCI Conference 2019*, 2019.

[11] R. Van Der Merwe and E. A. Wan, "The square-root unscented Kalman filter for state and parameter-estimation," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 6, pp. 3461–3464, 2001.

[12] R. J. Kobler, A. I. Sburlea, C. Lopes-Dias, A. Schwarz, M. Hirata, and G. R. Müller-Putz, "Corneo-retinal-dipole and eyelid-related eye artifacts can be corrected offline and online in electroencephalographic and magnetoencephalographic signals," *Neuroimage*, vol. 218, no. May, p. 117000, 2020,.

[13] L. Paninski, M. R. Fellows, N. G. Hatsopoulos, and J. P. Donoghue, "Spatiotemporal Tuning of Motor Cortical Neurons for Hand Position and Velocity," *J. Neurophysiol.*, vol. 91, no. 1, pp. 515–532, 2004.

Choosing strategies to deal with artifactual EEG data in children with cognitive impairment

Ana Tost[1], Carolina Migliorelli[2,1,3], Alejandro Bachiller[1,2,3*], Sergio Romero[1,2,3], Miguel A. Mañana[1,2,3*], M. Ángeles García-Cazorla[4,3]

[1] Department of Automatic Control (ESAII), Biomedical Engineering Research Centre (CREB), Universitat Politècnica de Catalunya (UPC), Barcelona, Spain.

[2] CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Spain.

[3] Institut de Recerca Sant Joan de Déu, Barcelona, Spain.

[4] Neurology Department, Neurometabolic Unit and Synaptic Metabolism Lab, Institut Pediàtric de Recerca, Hospital Sant Joan de Déu, metabERN and CIBERER-ISCIII, Barcelona, Spain.

Abstract

Rett syndrome is a disease that involves acute cognitive impairment and, consequently, a complex and varied symptomatology. This study evaluates the EEG signals of twenty-nine patients in order to implement an effective classification method to find the optimal artifact reduction strategy in each case. The classification has been made based on the mean and standard deviation (SD), allowing to differentiate patients with stereotyped and constant movements from those with a greater number of spasm or sudden movements. Since the various signal patterns may require diverse treatments, two artifact reduction methods have been analyzed. The first one is based on the distribution, using again the mean and SD, and the second one is based on an energy function which, theoretically, should be more robust to outliers and more stable to signal to noise ratio. The results confirm the existence of three groups of signals differentiated by having: low mean and low SD, high mean and low SD and high mean and high SD. However, despite finding three different patterns, the energy-based method is the one that works best for all them, offering adequate adaptation to each type of signal without losing robustness and stability. In conclusion, its implementation improves the detection of outliers without compromising artifact-free data segments, which allows to maintain the quality and quantity of the records.

1. Introduction

Electroencephalogram (EEG) has been consolidated over the years as one of the main techniques to identify brain activity and behavior. Measurement of neurophysiological changes related to postsynaptic activity in the neocortex has proven to be a powerful tool for the study of complex neuropsychiatric disorders, since variations in EEG signals depict a definite type of brain abnormality [1]. Information must be extracted by designing and developing signal processing algorithms in order to use it for diagnosis, monitoring and treatment of the identified brain pathologies. It is fundamental to make sure that the signal that will be analyzed corresponds only to the brain activity, and therefore, results will be reliable. This is incompatible with the presence of artifacts, which lead to an obstacle to interpret EEG signals due to their high amplitudes [2].

The challenge comes when signals to be dealt with are from awake pediatrics patients with a severe neurodegenerative disease [2]. Literature covers all known artifact reduction methods but, mostly, under controlled circumstances. For example, during children's sleeping periods, where the sensitivity to artifacts is highly reduced because movement, eye blinks and muscle artifacts mostly disappear [3]. Constrained signals without uncontrolled big artifacts are well studied and numerous solutions are already presented [4]. But, what happens when patients present movement disorders, seizures or autistic behavior? Or even all at the same time? This is what occurs in the case of patients with Rett Syndrome.

Rett Syndrome (RS) is a neurodevelopmental disorder caused by a mutation in the X-linked dominant MECP2 gene. The occurrence varies between 1/10.000 and 1/15.000 new-born females that follow a normal development during the first 6 to 18 months of age. After this period, patients will lose the abilities acquired until that moment and will initiate a degenerative process that is divided in 4 stages. Principal symptoms that will appear are: microcephaly, intellectual disability, seizures, autistic behavior, hyperventilation, spasticity, hyperreflexia and stereotyped movements [5]. Despite the severity and variability of the symptoms, it is possible to evaluate the level of cognitive performance objectively [6]. Analyzing the power of the EEG signals at the different frequency bands before, during and after visual stimulation tasks, allows to know the awareness, learning and comprehension capacities of patients. The combination of cognitive training and EEG is a powerful tool to better understand the disease and to incorporate communication improvements in the patient's day-to-day life.

For the purpose of the study and given its relevance, ensure the quality of the data constitutes the first step of the analysis. Previous studies with children with similar pathologies have shown difficulties in obtaining an EEG without artifacts. Many patients must be discarded due to excess artifacts, studies are based on signals no longer than 1 minute due to the lack of artifact-free epochs or visual inspections to manually remove artifacts must be done, which is highly time consuming [7], [8]. The most common solution to remove artifacts is to reject the data segments that contain those to get an artifact-free signal to work with. Thus, the aim of the study is to optimize this procedure to find the appropriate method to detect outliers automatically and accurately.

ISBN: 978-84-09-25491-0, pags. 248- 251

2. Methods

2.1. Participants, data acquisition and pre-processing

The database is formed by 29 patients with Rett Syndrome recruited from the pediatric neuroscience unit at the Sant Joan de Déu Hospital, Barcelona. A twenty-channel EEG (according to 10/20 system) was recorded with a Starstim 20 wireless device from Neuroelectrics, Barcelona, at a sampling frequency of 500 Hz. An accelerometer was placed at the back of the head to record the movement in x, y and z axes. During the session, the subject was using an eye-tracking device from Tobii Technology to record its visual response to a given stimulus on the screen. A total of 85 records containing basal and activity periods were registered. All the obtained signals are filtered using an elliptic bandpass filter between 1 and 50 Hz.

The study was approved by the local ethics committee following the Declaration of Helsinki.

2.2. Influence of patient's symptomatology in data variability

The clinical characteristics of the subjects are important to understand the high variability of the data: there are subjects who remain fairly stable during the session with slight random movements, patients who have a constant and continuous repetitive movement and patients who produce very strong and sudden movements at times but who remain mostly stable throughout the session. This is due to the symptoms of each patient as well as the stage of the disease in which they are. The degree of attention, understanding, motor control, autism or epilepsy varies in each child and influences their behavior during the activities.

In order to classify the data and group subjects according to their degree of movement, the mean and the standard deviation (*SD*) of the modulus of the derivative of the accelerometer signal in the 3 axes are analyzed. The derivative is used to maximize the slope changes of the accelerometer signal to better detect the head movements. Finding the relationship between the mean and the *SD* is the key point for ranking. It is expected to find a group with low mean and *SD*, due to the lack of sudden movements and high stability; another group with high mean but low *SD*, due to constant repetitive movements and, finally, a third group with high mean and *SD* due to sporadic strong movements. Figure 1 shows the mean-*SD* relationship found in the eighty-five recordings from the twenty-nine patients.

Two classification methods are used to analyze the mean and the SD: The Kernel Density Estimation (KDE), which represents the data using a continuous two-dimensional probability density curve that is analogous to a histogram, and the Gaussian Mixture Model (GMM), which depicts the density representation as the weighted sum of Gaussian distributions. The GMM algorithm is applied to the dataset for fitting three mixture-of-Gaussian models and to assign each record to the Gaussian it mostly belongs to [9]. As it is a probabilistic model, we can filter by probabilities to keep only those records with a probability (p) greater than 95% of belonging to its group and, therefore, obtain only those groups with clearly differentiated data.

2.3. Artifact rejection methods

In this study, two artifact rejection methods will be analyzed. The first one is based on the mean and the standard deviation and needs a *k-factor* to fit the data correctly. It is calculated as the mean + *k-factor* · *SD* of the filtered signal. In this study, 3 *k-factors* are tested: 3, 4 and 5. The second one is based on an energy function that has entropy as its main component. The purpose of the algorithm is to determine the baseline of the signal without the influence that movement peaks may add. It has the advantage of not having a threshold that depends on some events such as artifacts, as the first artifact detection method does by using the data distribution. Therefore, and in theory, this method will be more robust to outliers and more stable to signal-to noise ratio [9]. The methodology consists of: the filtered signal is divided into 100-ms segments and the wavelet entropy of the autocorrelation of each segment is calculated. Then, the baseline is computed as the *95%* percentile of the entropy value distribution and the threshold is determined as the *k- value* % of the obtained baseline. For this method, five *k-values* are tried. The five *k-values* are 0.995, 0.996, 0.997, 0.998 and 0.999. As the threshold is determined by the *k- value*-quantile of the absolute value of the baseline, the higher the *k-value* is, the higher the threshold. In order to decide the *k*-value, the percentage of artifact epochs versus the *k-value* will be analyzed.

3. Results

3.1. Data classification

Figure 1 shows the mean and *SD* distribution of the modulus of the accelerometer derivative in the three axes (x, y, z).

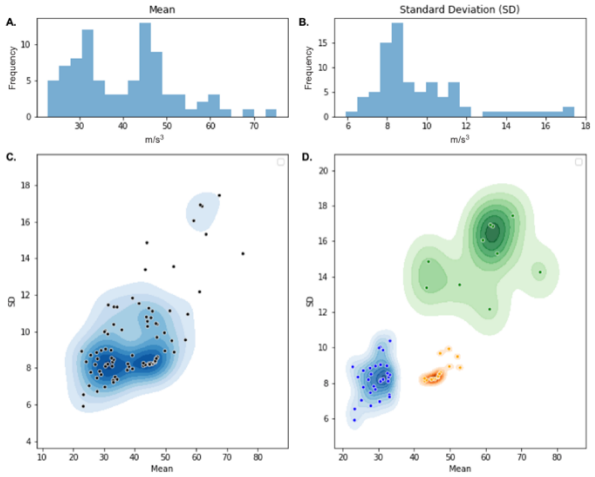


Figure 1. A. *Histogram showing the distribution of the means.* ***B.*** *Histogram showing the distribution of the standard deviations.* ***C.*** *Mean-SD ratio distribution using kernel density estimation (KDE).* ***D.*** *Density representation as the weighted sum of Gaussian distributions using the Gaussian Mixture Model (GMM). Three groups are differentiated: 1. LmLsd (blue): low mean and low SD. 2. HmLsd (orange): high mean and low SD. 3. HmHsd (green): high mean and high SD.*

Figure 1a shows the histogram of the mean values of the accelerometer where three main peaks can be observed: the first, which is the second highest, has the lowest mean around 30 m/s^3, the second, that is the highest, represents the group with an intermediate mean around 45 m/s^3, and finally, the smallest peak shows a more limited group of records that have the highest mean around 60 m/s^3. Figure 1b shows the histogram of the standard deviation (SD). There is one main peak around the value of 8 m/s^3 indicating that most records have a low *SD*, then, there is a lower concentration of records between 10 and 12 m/s^3, and finally, the smallest part of the distribution between 14 and 18 m/s^3 that represents the registers with the highest *SD*.

So far, and based on histograms, the presence of the three expected groups is observed. Figure 1c shows the mean-*SD* ratio distribution using KDE. Three groups are differentiated according to the density distribution: the area of highest density with low mean and low *SD*, the second highest density area with a higher mean and the same low *SD* and the smallest group with high mean and high *SD*.

After recognizing those three distinct groups, Figure 1d shows the results from the application of the GMM algorithm: LmLsd (Low mean Low *SD*), low mean due to stable behavior during the session with a low *SD* due to the lack of strong sudden movements. HmLsd (High mean Low *SD*): high mean due to constant repetitive movement during the session but low *SD* due to the lack of strong sudden movements. HmHsd (High mean High *SD*): both the mean and the *SD* are high because subjects present very strong and sporadic movements, which means that they remain mainly calm during the session but the outliers caused by these movements have extremely high amplitude, which increases the global mean and the *SD*.

By clustering with GMM, three groups could be obtained according to the behavior of the children during the sessions. The variability of the signals can cause each group to require a different solution when calculating the threshold to detect artifacts, for this reason, in this study two different methods of artifact rejection will be evaluated.

3.2. Selection of the *k*-value for the energy threshold

The energy threshold was obtained as the *k-value*% of the baseline. Five different values have been tested to find the optimum. Figure 2 shows the relationship between the *k-values* and the percentage of artifact epochs for the three groups classified in section 3.1.

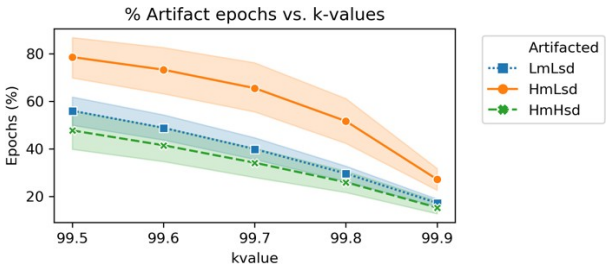


Figure 2. *K-values (x axis) versus the percentage of epochs with artifacts (y axis) for the three groups LmLsd, HmLsd and HmHsd.*

Both LmLsd and HmHsd, show a linear decrease in the percentage of artifactual epochs as *k-value* increases, while in the case of HmLsd, the variation in the percentage of artifactual epochs is more abrupt and is quantized with notable changes in the slope between each *k-value*. The HmLsd is the one that presents a differentiated behavior and the response to each *k-value* implies a greater change in the final selection of data. Before taking a decision, the evaluation will continue in the next section 3.3, where the thresholds with *k-values from* 99.7 to 99.9% will be studied in depth, as they are the ones that present the most abrupt changes in the HmLsd slope.

3.3. Threshold examination: mean-*SD* vs. energy

Once all procedures are defined, it is time to see how the different thresholds fit the data and what percentage of artifacts are detected in each case.

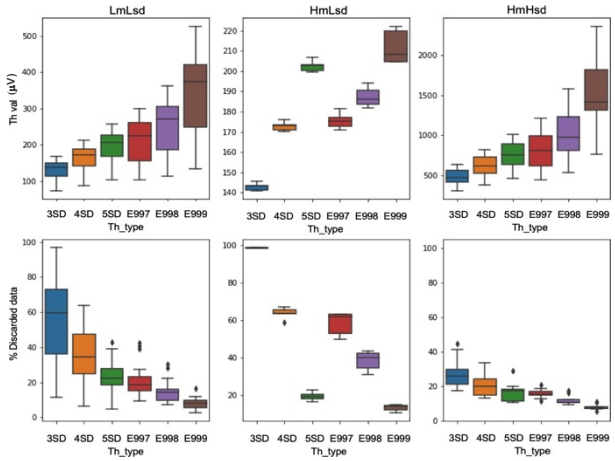


Figure 3. *The figure shows the threshold values and the percentage of discarded data, both versus each type of threshold. The mean plus SD method is tested with three "k-factors" of 3, 4 and 5 and the energy-based threshold is tested for three of the five k-values shown in Figure 2, due to preliminary results. The k-values are: 0.997, 0.998 and 0.999.*

The observation and analysis of Figure 3 in detail is crucial to understand how both methodologies are working and how they affect data. The first column contains the information regarding the LmLsd group. It can be seen how all three energy-based thresholds have higher median values (from 220 to 380 μV) than the mean-*SD* thresholds, which increase more smoothly and linearly and go from 120 to 200 μV. Therefore, the mean plus *SD* method seems to have a lower variation of the median threshold values for the different *k-factors*, and it also shows less variability between records (the range of values between quartiles of 25 and 75% is smaller), however, its influence on the percentage of discarded data is the opposite. It ranges from 60% of the data discarded with 3*SD* to 20% in 5*SD*, showing in this case a greater variation of the median threshold values and a greater variability between records than the energy-based threshold. This second one has a greater disparity between *k-values* in the median threshold calculation, but it is more precise selecting outliers and shows less variability between records, going from 20% of the discarded data with k=0.997 to 10% with k=0.999.

The second column, which describes the behavior of the HmLsd group, shows that the energy-based thresholds vary from 175 to 205 μV while for the mean-*SD* method go from 142 to 203 μV. In this case, the energy threshold presents a smaller variation with respect to the calculation of the median thresholds, unlike the LmLsd group, but greater variability between registers, as in the previous group. Following the same line, the percentage of discarded data is again more consistent for the energy thresholds, which vary from 60% for the lowest *k-value* to 15% for the highest, while for the 3*SD* threshold the percentage of discarded data is almost the 100% and, for the 5*SD*, it is around 20%. Therefore, the mean plus *SD* method is less consistent when setting the thresholds and discarding the data.

Finally, the third column shows the boxplots for the HmHsd group, whose behavior is very similar to LmLsd. All three energy-based thresholds are greater than the mean-*SD* ones with respect to the threshold's median values (750 to 1450 μV versus 500 to 750 μV, respectively), and again, this second method is more consistent calculating the median threshold values for the different *k-factors*.

Likewise, the variability between registers when calculating the thresholds is greater for the energy-based threshold. On the contrary, and also coinciding with the first group, the variability between records in the percentage of discarded data is greater for the mean-*SD* method. Regarding the median values of the percentage of discarded data, both methods show a similar behavior, going from 26 to 18 μV for the mean-*SD* method and from 17 to 8 μV for the energy-based method.

In summary, the variability in the threshold calculation is greater for the energy-based method while it is also more consistent discarding data. For this reason, it would be the most suitable method for all three groups. The last decision to make is which *k-value* to use. Observing Figure 3 and analyzing Figure 2 again, the decision will be made based on the calculation of the slopes between the *k-values* of the HmLsd group (due to its differentiated behavior). It is observed that the steepest change in the slope occurs between 99.8% and 99.9%, so the *k-value* that will be used to define the energy-based threshold is 0.998.

Figure 4 translates these results to observe the effect of the threshold chosen as optimal (energy threshold with *k-value* of 0.998) against the three thresholds of mean plus *SD,* which are used as the reference classification because were already used in the previous study to obtain the preliminary results that lead to the actual one.

For the purpose of the study, it is considered that at least 20% of the data must be free of artifacts. In Figure 4 all those records that conserve less than this 20% are shown in blue, while the ones that conserve enough data are shown in black. The difference is visible between the two main methods, energy and mean-*SD*. For the energy threshold, only 3 records would be eliminated, while for the 3*SD* method more than 50% of the registers would be removed, specifically 56 out of 85. In the case of the 4*SD*, a total of 47 registers would be eliminated (also over 50%) and finally, for the 5*SD*, 21 records would be lost. So, for the less restrictive mean-*SD* method, the percentage of records deleted is 24.7% versus 3.5% for the energy method.

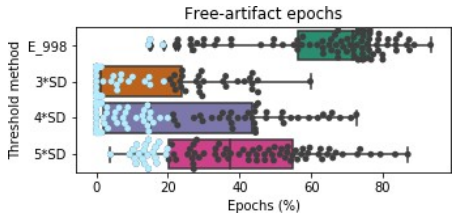


Figure 4. *The figure presents a boxplot for each threshold method (x-axis) with the information regarding the percentage of free-artifact epochs (y-axis) for all records.*

4. Conclusions and limitations

The complexity and variability of symptoms in patients with Rett Syndrome require an equally complex and diverse data analysis. The initial idea of the study about the existence of different groups of patients, due to their clinical stage and behavior, has been exhibited by using the mean and standard deviation. Three groups of registries have been identified, showing the importance of knowing the origin of movement, since it influences the performance of the different artifact detection methods. However, after evaluating both methods, if a single solution has to be offered for all types of data, the energy-based method is the optimal for all them.

The greater consistency discarding data together with a greater variability in the calculation of the thresholds, implies a more precise adaptation to each signal, obtaining more robustness in the selection of artifacts. This helps to reduce the loss of artifact- free data by removing only outliers, leading to a reduction in the percentage of records deleted.

The main limitations of the study lie in the signal recordings, since the artifacts must be annotated in real time to contrast automatically with the classification of the data and the detection of outliers. So far it has been done with video recordings, which is a manual method that can include more errors and involves a greater investment of time.

5. Acknowledgements

We would like to acknowledge specific funding support from the Spanish Patient Associations Mi Princesa Rett and Rettando al Síndrome de Rett. This project has received also funding from Torrons Vicens and the Ministry of Economy and Competitiveness (MINECO), Spain, under contract DPI2017-83989-R. CIBER-BBN is an initiative of the Instituto de Salud Carlos III, Spain. Alejandro Bachiller is a Serra Húnter Fellow.

6. References

[1] J. Wang, J. Barstein, L. E. Ethridge, M. W. Mosconi, Y. Takarae, and J. A. Sweeney, "Resting state EEG abnormalities in autism spectrum disorders," *Journal of Neurodevelopmental Disorders*, vol. 5, no. 1, pp. 1–14, Dec. 2013, doi: 10.1186/1866-1955-5-24.

[2] M. Dovgialo *et al.*, "Assessment of Statistically Significant Command-Following in Pediatric Patients with Disorders of Consciousness, Based on Visual, Auditory and Tactile Event-Related Potentials," *International Journal of Neural Systems*, vol. 8, no. 3, p. 1850048, 2019, doi: 10.1142/S012906571850048X.

[3] A. Kaminska, F. Cheliout-Heraut, M. Eisermann, A. Touzery de Villepin, and M. D. Lamblin, "EEG in children, in the laboratory or at the patient's bedside," *Neurophysiologie Clinique*, vol. 45, no. 1, pp. 65–74, Mar. 2015, doi: 10.1016/j.neucli.2014.11.008.

[4] D. W. Klass, "The continuing challenge of artifacts in the EEG," *American Journal of EEG Technology*, vol. 35, no. 4. pp. 239–269, Dec. 10, 1995, doi: 10.1080/00029238.1995.11080524.

[5] N. P. Verma, R. L. Chheda *'*, M. A. Nigro, and Z. H. Hart, "Electroencephalographic findings in Rett Syndrome" 1986.

[6] P. M. Baptista, P. M. Baptista, M. T. Mercadante, E. C. Macedo, and J. S. Schwartzman, "Correspondence: Cognitive performance in Rett syndrome girls: a pilot study using eyetracking technology," *Journal of Intellectual Disability Research*, doi: 10.1111/j.1365-2788.2006.00818.x.

[7] R. A. Fabio, L. Billeci, G. Crifaci, E. Troise, G. Tortorella, and G. Pioggia, "Cognitive training modifies frequency EEG bands and neuropsychological measures in Rett syndrome," *Research in Developmental Disabilities*, vol. 53–54, pp. 73–85, 2016, doi: 10.1016/j.ridd.2016.01.009.

[8] K. J. Roche, J. J. Leblanc, A. R. Levin, H. M. O'Leary, L. M. Baczewski, and C. A. Nelson, "Electroencephalographic spectral power as a marker of cortical function and disease severity in girls with Rett syndrome," *Journal of Neurodevelopmental Disorders*, vol. 11, no. 1, p. 15, Jul. 2019, doi: 10.1186/s11689-019-9275-z.

[9] C. Migliorelli *et al.*, "SGM: a novel time-frequency algorithm based on unsupervised learning improves high-frequency oscillation detection in epilepsy," *J. Neural Eng*, vol. 17, p. 26032, 2020, doi: 10.1088/1741- 2552/ab8345.

Análisis del flujo aéreo nocturno mediante wavelets para la ayuda en el diagnóstico de la apnea del sueño infantil

V. Barroso García[1,2], G.C. Gutiérrez Tobal[1,2], L. Kheirandish Gozal[3], D. Álvarez González[1,2,4], F. Vaquerizo Villar [1,2], F. del Campo Matías[1,2,4], D. Gozal[3], 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] *Department of Child Health, The University of Missouri School of Medicine, Columbia, Missouri, USA,* gozald@health.missouri.edu

[4] Servicio de Neumología, Hospital Universitario Río Hortega, Valladolid, España, fsas@telefonica.net

Resumen

Este estudio se centra en el análisis automático de la señal de flujo aéreo monocanal (FA) para la ayuda en el diagnóstico de la Apnea Obstructiva del Sueño (AOS) infantil. En concreto, se ha evaluado la utilidad diagnóstica de las medidas extraídas de los diferentes niveles de descomposición obtenidos al aplicar la transformada wavelet discreta a la señal de FA. Además, se ha evaluado la complementariedad entre la información wavelet de FA y el índice de desaturación de oxígeno en sangre ≥3% (ODI_3), que fue directamente obtenido de la señal de saturación de oxígeno en sangre (SpO_2). Se han analizado 946 registros de FA y SpO_2 procedentes de niños de 0 a 13 años. El análisis se realizó en dos etapas metodológicas: (i) descomposición wavelet y extracción de características y (ii) estimación del índice de apnea hipopnea mediante LSBoost con árboles de decisión como clasificador base. Por tanto, dos modelos fueron entrenados: $LSBoost^{FA}$ alimentado con las características wavelet de FA y $LSBoost^{FA,ODI3}$ alimentado con las características wavelet de FA y el ODI_3. $LSBoost^{FA,ODI3}$ alcanzó mayor rendimiento diagnóstico que $LSBoost^{FA}$ en el grupo de test, con una precisión del 78.5% para 1 evento/h, 77.4% para 5 eventos/h, y 88.8% para 10 eventos/h, mostrando así la complementariedad de ambos enfoques. Estos resultados sugieren que el análisis de la señal FA mediante la transformada wavelet podría ser usado junto con el ODI_3 como una potente herramienta para ayudar a diagnosticar la AOS pediátrica, especialmente en los sujetos con un grado de severidad severo.

1. Introducción

La Apnea Obstructiva del Sueño (AOS) es un trastorno respiratorio de alta prevalencia, que afecta hasta al 4% de la población infantil [1]. Los niños afectados por esta enfermedad presentan repetidos episodios de apnea (ausencia de flujo aéreo) e hipopnea (reducción del flujo aéreo acompañado de una desaturación de oxígeno en sangre ≥3% o arousal) mientras duermen [2]. La AOS tiene graves consecuencias en los niños que la padecen, ya que afecta a su sistema cardiovascular y nervioso, así como a su desarrollo cognitivo [1]. Por tanto, es esencial que ésta enfermedad sea diagnosticada y tratada a tiempo, antes de que las consecuencias empeoren y lleguen a ser potencialmente irreversibles.

La polisomnografía nocturna (PSG) es el método de referencia para diagnosticar la AOS pediátrica [3]. Mediante esta prueba se registran múltiples señales biomédicas, que permiten calcular el índice de apnea-hipopnea (IAH: número de eventos de apnea e hipopnea por hora de sueño) y determinar la presencia y la severidad de la enfermedad [2]. A pesar de la eficacia de la prueba, la PSG es compleja, requiere de costosos equipos para adquirir las señales y es especialmente incómoda para los niños [4]. Además, presenta disponibilidad limitada, lo que ocasiona largas listas de espera y demoras en el diagnóstico [4].

Para lidiar con estas limitaciones inherentes de la PSG, se están explorando técnicas de diagnóstico más simples. Un enfoque común es el análisis de un conjunto reducido de señales biomédicas, como la fotopletismografía (PPG), la saturación de oxígeno en sangre (SpO_2) o el flujo aéreo (FA) [5–9]. En este estudio, se propone el análisis wavelet de la señal de FA para ayudar en el diagnóstico de la AOS en niños. El FA está directamente involucrado en las definiciones de apnea e hipopnea [2], por lo que la simplificación de la PSG conduce de forma natural a su análisis. Además, esta señal ha mostrado una gran utilidad en el diagnóstico de la AOS pediátrica [8–10]. Dado que la repetida ocurrencia de eventos apneicos conduce a alteraciones del FA en el dominio de la frecuencia, su análisis espectral ha sido ampliamente estudiado en este contexto [9,10]. Típicamente estos estudios aplican métodos basados en la transformada de Fourier. Sin embargo, dichos métodos presentan diversas limitaciones, como la asunción de estacionariedad de la señal y el establecimiento de una resolución tiempo-frecuencia fija [11,12]. Por el contrario, la transformada wavelet discreta (TWD) permite realizar un análisis multi-resolución de señales no estacionarias, superando así estas limitaciones [12]. Estas propiedades de la TWD hacen de él un método muy adecuado para realizar la caracterización de señales biomédicas no estacionarias como el FA. Además, TWD ya ha sido aplicado con éxito a otras señales polisomnográficas [11,13]. Sin embargo, este es el primer estudio en el que TWD es aplicado a la señal de FA en contexto de la AOS infantil.

De acuerdo con las consideraciones anteriores, el presente estudio parte de la hipótesis de que la TWD de la señal de FA podría proporcionar información útil para ayudar a diagnosticar la AOS en niños. Por tanto, nuestro objetivo fue evaluar la utilidad y la capacidad diagnóstica de las características derivadas del análisis wavelet de FA para ayudar a determinar la presencia y la severidad de AOS.

ISBN: 978-84-09-25491-0, pags. 252- 255

Para ello, se realizó el proceso de descomposición wavelet y se calcularon las características de la señal de detalle de FA obtenida en cada nivel de descomposición. Además, el índice de desaturación de oxígeno en sangre ≥3% (ODI_3), un parámetro ampliamente utilizado en la práctica clínica, fue incorporado al estudio para evaluar su complementariedad con la información aportada por las características wavelet. Finalmente, estas características se combinaron mediante un modelo de regresión LSBoost con árboles de decisión como modelo base, con el propósito de mejorar la capacidad diagnóstica de las características individuales.

2. Sujetos y señales

En este estudio se ha llevado a cabo el análisis de la señal de FA de un total de 946 sujetos pediátricos con sospecha de AOS. Todos ellos fueron remitidos a la Unidad Pediátrica del Sueño del *Comer Children's Hospital* de la Universidad de Chicago (EE.UU.), donde se les realizó la PSG. El Comité Ético del Hospital aprobó los protocolos de estudio (#11-0268-AM017, #09-115-B-AM031 y #IRB14-1241) y los representantes legales de los niños dieron su consentimiento informado para participar en la investigación.

Todos los niños fueron diagnosticados de acuerdo con las reglas de la *American Academy of Sleep Medicine* (AASM) [2]. Para determinar la severidad de cada sujeto, se usaron los umbrales de IAH 1, 5 y 10 eventos/h [1,14]. La Tabla 1 muestra los datos demográficos y clínicos de los sujetos bajo estudio. La población fue distribuida aleatoriamente en dos grupos: entrenamiento (60%) y test (40%), para diseñar y validar la metodología propuesta, respectivamente. No se encontraron diferencias estadísticamente significativas en edad, sexo, índice de masa corporal (IMC) e IAH entre ambos grupos (*p*-valor > 0.01 al aplicar el test de Mann-Whitney).

Los niños fueron monitorizados usando el polisomnógrafo digital Polysmith, Nihon Kohden America Inc., Irvine, CA, USA. Los registros de FA fueron adquiridos con un termistor y remuestreados a 100 Hz [2]. Después, estas señales fueron filtradas, mediante un filtro paso bajo con una frecuencia de corte de 1.5 Hz, y normalizadas [9]. Los registros de SpO_2, utilizados para obtener el ODI_3, fueron adquiridos con un pulsioxímetro y remuestreados a 25 Hz [2]. Además, las señales de FA y SpO_2 fueron sometidas a un proceso de rechazo automático de artefactos como en estudios previos [8,9]. Todos los registros incluidos en nuestro estudio tuvieron más de 3 horas de duración después de la eliminación de artefactos. El ODI_3 se obtuvo directamente de los registros de SpO_2, siguiendo la metodología propuesta por Taha *et al*. [15].

3. Metodología

3.1. Descomposición wavelet y extracción de características de FA

TWD se puede ver como una extensión del análisis clásico de Fourier donde, en vez de analizar la señal con una resolución tiempo-frecuencia fija, se realiza un análisis multi-resolución [12]. Esta propiedad hace que TWD sea capaz de realizar una descomposición de la señal en diferentes rangos de frecuencia, proporcionando alta resolución temporal a altas frecuencias y alta resolución frecuencial a bajas frecuencias [11,12]. Este análisis puede aplicarse utilizando la transformada wavelet continua o discreta [12]. Dado que la TWD sólo emplea escalas diádicas, reduciendo así el coste computacional, en nuestro estudio utilizamos la forma discreta.

Característica	Todos	Entrenamiento	Test
Sujetos (*n*)	946	570	376
Edad (años)	6 [6]	6 [5]	6 [6]
Varones (*n*)	584 (61.7%)	339 (59.5%)	245 (65.2%)
IMC (kg/m^2)	17.9 [6.2]	17.7 [6.7]	18.1 [6.0]
IAH (e/h)	3.8 [7.8]	4.2 [8.3]	3.3 [6.4]

Los datos son presentados como número de sujetos (%) o mediana [rango intercuartil]

Tabla 1. *Datos demográficos y clínicos de los sujetos.*

De acuerdo con la Figura 1, el proceso de descomposición de una señal $x[n]$ mediante la TWD consiste en un banco de filtros recursivos seguidos de un submuestreo de la señal filtrada [12]. Así, en cada nivel de descomposición j se aplica un filtro paso alto $g[n]$ (correspondiente a la función wavelet) y un filtro paso bajo $h[n]$ (correspondiente a la función de escalado), seguidos de un diezmado diádico para obtener la señal de detalle D_j y de aproximación A_j, respectivamente [12]. Este proceso de descomposición continúa hasta alcanzar el nivel de detalle máximo $N=\log_2(M)$, donde M es el tamaño de la señal.

En este estudio, la TWD se aplicó a segmentos de FA de $M = 2^{16}$ muestras, utilizando para ello la wavelet madre Daubechies-5 [16]. Por tanto, se realizó una descomposición en N=16 niveles de detalle. Dado que la señal de FA se filtró previamente a 1.5 Hz, en esta investigación se analizaron las señales de detalle de D_6 a D_{16} (de 0 a 1.5625 Hz). Para cuantificar la información contenida en cada una de ellas, las siguientes características fueron extraídas [11]:

- Los momentos estadísticos del primer al cuarto orden (M_{1Dj} – M_{4Dj}) fueron calculados para evaluar la tendencia central, dispersión, asimetría y concentración de la información wavelet contenida en cada nivel de detalle.
- El máximo (Max_{Dj}) y mínimo (Min_{Dj}) valor de los coeficientes de D_j fueron obtenidos para determinar la máxima y mínima amplitud de cada una de las señales de detalle generadas en el proceso de descomposición.
- La energía (E_{Dj}) fue extraída para medir la amplitud cuadrática de la señal D_j y estimar así la actividad producida en el nivel de resolución asociado a cada banda de frecuencias.
- La entropía wavelet (*WE*) fue calculada para evaluar los cambios de distribución de energía producidos en el proceso de descomposición y cuantificar su irregularidad [11].

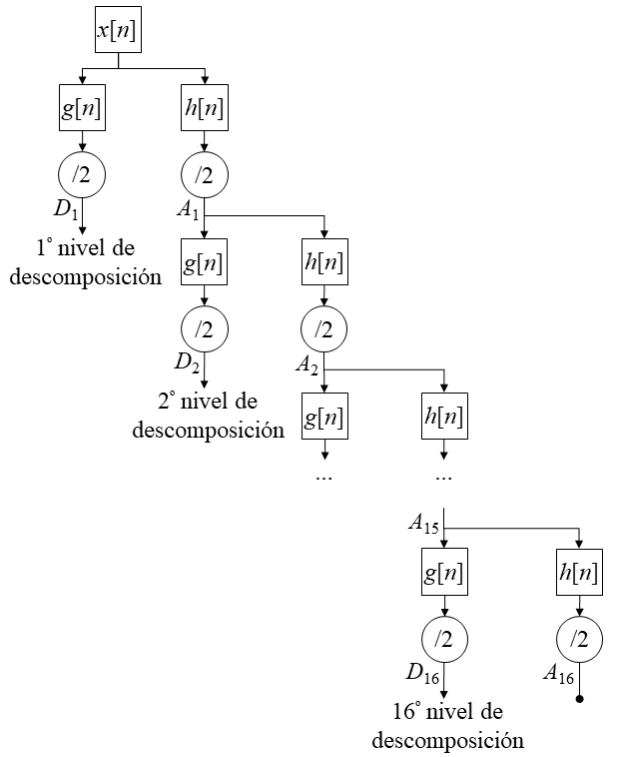


Figura 1. *Proceso de descomposición de la transformada wavelet discreta.*

De este modo, se obtuvo un total de 78 características wavelet (7 de cada nivel de detalle y la *WE* del perfil TWD completo) de la señal de FA.

3.2. Estimación del IAH

Para estimar el IAH se utilizó el método LSBoost (*Least Squares Boosting*), un algoritmo de boosting especialmente diseñado para problemas de regresión de variables continuas [17,18]. LSBoost utiliza modelos predictivos base que son ajustados y entrenados en L iteraciones para obtener una predicción más robusta de la variable. Para ello, aplica el método del gradiente descendente a la función de pérdida cuadrática, lo que equivale a un ajuste repetido de los residuos [18]. De este modo, en cada iteración se realiza un ajuste por mínimos cuadrados que permite ajustar la mejor variable predictora reduciendo la suma residual de cuadrados [17].

En nuestro estudio, se emplearon árboles de decisión como modelo base para incrementar la generalización de LSBoost [18]. Este método también requiere el uso de la tasa de aprendizaje α para reducir el sobreajuste de los modelos [18]. Tanto L como α requieren ser optimizados. Así, se eligieron los valores de L y α que maximizaron el kappa de Cohen obtenido al aplicar el algoritmo a 1000 réplicas *bootstrap* del grupo de entrenamiento.

3.3. Análisis estadístico

Las características wavelet extraídas de FA no pasaron el test de normalidad de Lilliefors. Por este motivo, se utilizó el test no paramétrico de Kruskal-Wallis para evaluar la existencia de diferencias estadísticamente significativas (p-valor < 0.01 después de la corrección por Bonferroni) entre los grupos de severidad de AOS. El rendimiento diagnóstico de los modelos LSBoost se evaluó en términos de sensibilidad (S), especificidad (E), precisión (P), valor predictivo positivo y negativo (VPP y VPN) y razón de verosimilitud positiva y negativa (LR+ y LR-).

4. Resultados

4.1. Grupo de entrenamiento

Dos modelos LSBoost fueron diseñados y entrenados con el grupo de entrenamiento: LSBoost^{FA}, alimentado con las características wavelet extraídas de FA, y LSBoost^{FA,ODI3}, cuyas entradas fueron las características de FA y el ODI_3. En máximo kappa se obtuvo para valores de L = 250 y α = 0.125 tanto para LSBoost^{FA} como para LSBoost^{FA,ODI3}.

4.2. Grupo de Test

Se evalúo el rendimiento diagnóstico de los modelos LSBoost^{FA} y LSBoost^{FA,ODI3}, así como el rendimiento individual del ODI_3. Los resultados obtenidos para los umbrales 1, 5 y 10 eventos/h se muestran en la Tabla 2.

5. Discusión y conclusiones

En este estudio se ha evaluado la utilidad de las características wavelet extraídas de los diferentes niveles de descomposición de FA para detectar la AOS pediátrica y su severidad. También se ha evaluado la complementariedad entre la información proporcionada por el análisis wavelet de FA y el índice oximétrico ODI_3. Así, la combinación de ambos enfoques permitió obtener un modelo LSBoost capaz de estimar el IAH de los sujetos pediátricos con una moderada-alta precisión diagnóstica (78.5%, 77.4% y 88.8% para 1, 5 y 10 eventos/h, respectivamente). Además, destaca el valor LR+ = 10.73 alcanzado por este modelo en 10 eventos/h, ya que un LR+ ≥ 10 es un robusto indicador para establecer la presencia de una enfermedad [19].

Diversos estudios han evaluado el uso de un conjunto reducido de señales biomédicas para detectar la AOS en niños. Gil *et al*. analizaron 21 PPGs obteniendo una precisión del 80.0% para 5 eventos/h [5]. Garde *et al*. basó su estudio en el análisis de 207 registros SpO_2, alcanzando una precisión de 75.0%, 82.0% y 89.0% para 1, 5 y 10 eventos/h, respectivamente [6]. Gutiérrez-Tobal *et al*. analizaron características espectrales de FA e índices convencionales de SpO_2 procedentes de 50 sujetos

Modelo	IAH	S (%)	E (%)	P (%)	VPP (%)	VPN (%)	LR+	LR-
LSBoost^{FA}	1	94.1	12.5	78.5	81.9	33.3	1.08	0.47
	5	67.4	58.7	62.0	49.5	75.0	1.63	0.56
	10	58.8	83.8	78.5	49.5	88.3	3.62	0.49
LSBoost^{FA,ODI3}	1	96.1	4.2	78.5	80.9	20.0	1.00	0.95
	5	71.6	80.9	77.4	69.2	82.6	3.74	0.35
	10	72.5	93.2	88.8	74.4	92.6	10.73	0.29
ODI_3	1	59.9	86.1	64.9	94.8	33.7	4.31	0.47
	5	69.5	89.4	81.9	79.7	83.0	6.53	0.34
	10	81.3	88.5	87.0	65.7	94.6	7.07	0.21

Tabla 2. *Rendimiento diagnóstico de los modelos LSBoost^{FA} y LSBoost^{FA,ODI3} y del ODI_3.*

[10]. Para el umbral 3 eventos/h su propuesta alcanzó una precisión del 86.3%. No obstante, el bajo número de sujetos empleados en estas investigaciones limita la generalización de sus resultados. Por el contrario, nosotros hemos validado nuestra propuesta empleando una amplia base de datos.

Otros estudios incluyeron un mayor número de registros, como Hornero *et al*. [7], Barroso-García *et al*. [8], y Jiménez-García *et al*. [9], que analizaron 4191 señales SpO_2, 946 señales FA y 974 registros FA y SpO_2, respectivamente. Estos trabajos obtuvieron precisiones de 75.1 – 83.2% en 1 evento/h, 78.5 – 82.0% en 5 eventos/h, y 90.2 – 91.0 % en 10 eventos/h. En el presente estudio alcanzamos precisiones dentro de este rango para 1 evento/h (78.5%), así como ligeramente inferiores para 5 y 10 eventos/h (77.4% y 88.8%, respectivamente). Sin embargo, nuestra propuesta superó a estos estudios en diversas métricas, como en S a Hornero *et al*. y Jiménez-García *et al*. en los 3 umbrales, o en E, VPP y LR+ a Barroso-García *et al*. en 5 eventos/h.

Este estudio presenta algunas limitaciones. Aunque el número de sujetos es elevado, con una base de datos aún mayor nuestros resultados serían más generalizables. Además, aunque Daubechies-5 ha mostrado gran utilidad, sería interesante aplicar otras wavelets madre y analizar cuál se ajusta mejor al comportamiento del FA nocturno. Cómo línea de investigación futura, la caracterización de FA podría centrarse en el análisis de único nivel de detalle, o una combinación de estos, que fuera más representativo del comportamiento del FA en presencia de eventos apneicos. También podría llevarse a cabo una etapa de selección de características que permitiera evaluar la relevancia y redundancia de los parámetros wavelet extraídos, así como reducir la dimensionalidad del modelo predictivo. Finalmente, podría ser útil el uso de otros métodos de estimación del IAH o clasificación multiclase que mejoren el rendimiento diagnóstico de FA, lo que constituye otra interesante línea de investigación futura.

En resumen, la señal FA se descompuso en 16 niveles de detalle y se caracterizó mediante características wavelet. El modelo LSBoost alimentado con estas características y el ODI_3 alcanzó un rendimiento diagnóstico moderado-alto, superando el rendimiento individual del ODI_3 en 1 y 10 eventos/h. Por ello, estos resultados sugieren que el análisis wavelet de la señal FA podría ser usado junto con el ODI_3 para ayudar a diagnosticar la AOS en niños, sobre todo en los casos más severos.

Agradecimientos

Este estudio ha sido financiado por el Ministerio de Ciencia, Innovación y Universidades - Agencia Estatal de Investigación y FEDER (DPI2017-84280-R y RTC-2017-6516-1), la '*European Commission*' y FEDER (POCTEP 0702_MIGRAINEE_2_E y 0378_AD_EEGWA_2_P), y por el CIBER-BBN mediante el Instituto de Salud Carlos III cofinanciado con fondos FEDER. V. Barroso García es beneficiaria de una ayuda concedida por la Consejería de Educación de la Junta de Castilla y León y el FSE. F. Vaquerizo Villar cuenta con un contrato predoctoral FPU (FPU16/02938) del Ministerio de Educación, Cultura y Deporte. L. Kheirandish Gozal y D. Gozal cuentan con ayudas del 'National Institutes of Health' (HL130984, HL140548 y AG0618).

Referencias

[1] A.G. Kaditis *et al*., Obstructive sleep disordered breathing in 2- to 18-year-old children: Diagnosis and management, Eur. Respir. J. 47 (2016) 69–94.

[2] 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. 08 (2012) 597–619.

[3] C. Jon, Polysomnography in Children, in: R.B. Mitchell, K.D. Pereira (Eds.), Pediatr. Otolaryngol. Clin., Humana Press, Totowa, NJ, 2009: pp. 35–47.

[4] 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. 50 (1995) 972–5.

[5] 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. 57 (2010) 1079–1088.

[6] A. Garde *et al*., Pediatric pulse oximetry-based OSA screening at different thresholds of the apnea-hypopnea index with an expression of uncertainty for inconclusive classifications, Sleep Med. 60 (2019) 45–52.

[7] R. Hornero *et al*., Nocturnal Oximetry-based Evaluation of Habitually Snoring Children., Am. J. Respir. Crit. Care Med. 196 (2017) 1591–1598.

[8] V. Barroso-García *et al*., Usefulness of recurrence plots from airflow recordings to aid in paediatric sleep apnoea diagnosis, Comput. Methods Programs Biomed. 183 (2020) 105083.

[9] J. Jiménez-García *et al*., Assessment of Airflow and Oximetry Signals to Detect Pediatric Sleep Apnea-Hypopnea Syndrome Using AdaBoost, Entropy. 22 (2020) 670.

[10] 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. 18 (2015) 401–7.

[11] F. Vaquerizo-Villar *et al*., Wavelet analysis of oximetry recordings to assist in the automated detection of moderate-to-severe pediatric sleep apnea-hypopnea syndrome, PLoS One. 13 (2018) e0208502.

[12] O. Rioul *et al*., Wavelets and Signal Processing, IEEE Signal Process. Mag. 8 (1991) 14–38.

[13] R. Lin *et al*., A new approach for identifying sleep apnea syndrome using wavelet transform and neural networks, Biomed. Eng. - Appl. Basis Commun. 18 (2006) 138–143.

[14] M. Luz Alonso-Álvarez *et al*., Consensus document on sleep apnea-hypopnea syndrome in children, Arch. Bronconeumol. 47 (2011) 2–18.

[15] B.H. Taha *et al*., Automated Detection and Classification of Sleep-Disordered Breathing From Conventional Polysomnography Data, Sleep. 20 (1997) 991–1001.

[16] J. Lee *et al*., Classification of healthy and abnormal swallows based on accelerometry and nasal airflow signals, Artif. Intell. Med. 52 (2011) 17–25.

[17] P. Bühlmann *et al*., Boosting Algorithms: Regularization, Prediction and Model Fitting, Stat. Sci. 22 (2008) 477–505.

[18] J.H. Friedman, Greedy function approximation: A gradient boosting machine, Ann. Stat. 29 (2001) 1189–1232.

[19] J.J. Deeks *et al*., Diagnostic tests 4: likelihood ratios, BMJ. 329 (2004) 168–169.

Análisis de la respuesta electrodermal para la identificación de pacientes con depresión

Martinez M. [1], Lazaro J. [1, 2], Kontaxis S. [1], Laguna P.[1, 2], Gil E[1, 2]., Bernal Ruíz ML. [1], Siddi S.[2], de la Cámara C. [1], Aguiló J. [2], García E. [2], Haro JM [2]. and Bailón R [1, 2].

[1]Universidad de Zaragoza, Zaragoza, España

[2] Centro de Investigación Biomédica en Red, Madrid, España

Resumen

La Actividad Electrodermal (AED) se relaciona directamente con la activación del sistema sudomotor controlado por el sistema nervioso simpático. De esta forma, las variaciones originadas en la respuesta al estrés pueden ser analizadas. Esto permite su utilización en el estudio de los trastornos conductuales, entre ellos, los trastornos depresivos. En el presente estudio se pretende comprobar la utilidad del análisis de la AED en el manejo de pacientes con trastorno depresivo unipolar.

Se realizó un estudio de cohortes donde incluyeron un total de 56 sujetos que fueron sometidos a cuestionarios de salud mental y test estresores. Tras un preprocesado se descompuso la señal en sus componentes tónica y fásica, y se extrajo información de la actividad nerviosa sudomotora. Se realizó estadística descriptiva e inferencial sobre las distintas variables.

El análisis de la señal de AED mostró diferencias estadísticamente significativas entre los pacientes afectos de depresión y los controles eutímicos. Además, se obtuvieron resultados con una relevancia clínica más destacable para los datos de la componente tónica y las características extraídas de la actividad nerviosa sudomotora. Estos resultados indican que el análisis de la señal AED puede contribuir a caracterizar las diferencias en la respuesta a un estímulo de pacientes con depresión respecto a sujetos sanos.

1. Introducción

Durante los últimos 30 años los trastornos distímicos se están convirtiendo en una de las principales fuentes de incapacidad médica. Desde 1990 el número de casos diagnosticados ha aumentado a un ritmo imparable [1]. Según la base de datos mundial de la Organización Mundial de la Salud, los trastornos depresivos afectan a 254 millones de personas y son la causa de aproximadamente 800.000 muertes por suicidio al año [2].

A pesar de lo alarmante de estos datos, llama la atención, la ausencia de criterios diagnósticos objetivos, reproducibles y fiables de la depresión.

El Manual Diagnóstico y Estadístico de los Trastornos Mentales [3] sienta las bases para el diagnóstico de la depresión mediante un cuestionario con lo que, el sujeto será diagnosticado o no, en consonancia con su percepción subjetiva de sus sentimientos y sus actitudes ante la vida.

Numerosos estudios han intentado establecer un nivel objetivo de múltiples marcadores bioquímicos con respecto al diagnóstico de depresión [4]. Aunque muchos de ellos alcanzaron altos niveles de significación, todavía no hay evidencia suficiente para que los resultados sean consistentes. Además, la toma de las muestras biológicas para el análisis de la presencia de estos biomarcadores dificulta establecer rutinas de seguimiento de los sujetos.

En respuesta a determinados estímulos, el eje Hipotalámico-Pituitario-Adrenal (HPA), mediado por el Sistema Nervioso Simpático (SNS), aumenta la producción de cortisol. Este pico de cortisol es detectado por receptores de glucocorticoides, de modo que la producción de cortisol se inhibe a través de un neurotransmisor específico, el ácido gamma-aminobutírico (GABA) [5]. En pacientes con depresión, este sistema regulador se ve alterado: niveles altos de cortisol no generan una respuesta inhibitoria vía GABA, lo que genera una situación de estrés crónico y, secundariamente, una falta de respuesta al mismo [6].

La actividad electrodermal (AED) describe los cambios en la resistencia que ofrece la piel al paso de una corriente eléctrica. Estos cambios se deben a las variaciones de salinidad producidas por el flujo de sudoración y su contenido de iones de sodio. Se distinguen dos componentes fundamentales en la señal de AED:

- Las componentes de evolución lenta conocidas como componente tónica o nivel de conductancia dérmica.
- Las componentes de mayor frecuencia conocidas como componente fásica o respuestas de conductancia dérmica (RCD).

La activación del SNS en respuesta a estímulos externos, provoca un aumento en el trabajo eficaz de las glándulas ecrinas. El control del sistema sudomotor depende únicamente del sistema nervioso simpático [7]. Por esta razón, la AED se considera posiblemente el marcador más fiable de activación simpática [8].

Un nivel elevado y constante de hormonas segregadas por el SNS es un importante indicador de trastornos conductuales. Los bajos niveles del neurotransmisor inhibitorio GABA, conducen a la hiperactivación del eje HPA, teniendo como consecuencia la saturación de los receptores que desencadenan las respuestas del SNS [5-6] ante un estímulo de estrés.

Esto se traduce en la idea de que pacientes afectados por desórdenes depresivos reaccionarán de manera menos intensa a los estímulos que los sujetos sanos [9]. La activación de las glándulas sudoríparas ecrinas sigue el mismo patrón, es decir: la señal de AED se comportará de manera distinta en pacientes y sujetos sanos [10].

Este estudio comprobará si existen diferencias en la AED de pacientes deprimidos y una muestra de población sana, al someter a los sujetos a diferentes situaciones de estrés.

ISBN: 978-84-09-25491-0, pags. 256- 259

2. Material y Métodos

2.1. Protocolo experimental y recogida de muestras

Se configuraron dos cohortes de población: un grupo de 24 sujetos diagnosticados de trastorno unipolar depresivo y otro grupo de 32 sujetos eutímicos, todos ellos firmaron el correspondiente consentimiento informado. El conjunto de la muestra es homogéneo en cuanto a distribución por edades y sexo.

Cada uno de los sujetos fue citado para la realización de una sesión experimental [9]. La sesión se configuró en las siguientes fases:

1) Cuestionarios de salud mental:
   a. Inventario de depresión de Beck (BDI-II)
   b. Test psicométricos
2) Test estresores:
   a. Test del trazo.
   b. Test de Stroop.
3) Recuperación

Durante el procedimiento se recogieron registros de diferentes señales fisiológicas (electrocardiograma, electromiograma, AED, …).

2.2. Preprocesado de la señal.

Las señales obtenidas, con frecuencia de muestreo $f_m = 200\ Hz$, fueron interpoladas linealmente para subsanar las pérdidas puntuales en la medición.

Se analizaron tanto la representación gráfica como los valores de cada una de las señales de AED durante la realización de cada una de las pruebas y de manera individualizada, comprobando si las características correspondían a las esperadas para este tipo de registros [7]. En caso contrario, se ha descartado la muestra. En particular, se han eliminado los registros que incumplían:

- Rango de oscilación lenta entre 1 y 40 μS.
- Rango de oscilación rápida entre 0.1 y 0.5 μS.

Finalmente, la muestra y sus cohortes quedaron compuestos de la siguiente manera:

	PACIENTES	CONTROLES	TOTAL
BDI	24	22	48
PSICOMÉTRICOS	23	27	50
TMT	22	28	50
STROOP	23	31	54
RECOVERY	22	32	54

Tabla 1. *Población a estudio para cada una de las fases.*

En este estudio se utilizó un filtrado paso bajo de fase cero con un filtro de respuesta máximamente plana (Butterworth) con N = 3 coeficientes y frecuencia de corte de $f_C = 0{,}5\ Hz$ para la eliminación de ruido.

2.3. Procesado. El modelo cvxEDA

El modelo cvxEDA [11] describe la actividad electrodermal como la salida de un sistema Lineal Invariante en el Tiempo (LTI). Esta descripción permite la separación de la AED en sus componentes tónica y fásica. Esta descomposición permite la observación de la actividad neurológica que subyace en los datos de las señales de AED recogidas.

Inicialmente la AED es definida como la superposición de la componente tónica (x_t) , la componente fásica (x_f) y un término de ruido Gaussiano (ε):

$$y[n] = x_t[n] + x_f[n] + \varepsilon$$

La actividad fásica, relacionada con la respuesta inmediata a un estímulo inducido, se define como la convolución entre la actividad del SNS, $p[n]$, y la función de Bateman, que modela el proceso de difusión del sudor.

La componente tónica se modela a través de una función *B-Spline*, una curva parametrizada acorde con la dinámica de baja frecuencia de esta componente.

Una vez se ha definido el modelo de la señal de AED se plantea un problema de interferencia bayesiana para estimar el máximo a posteriori de x_t y la actividad nerviosa sudomotora (ANSM). Al haber definido tres variables marcadas ($x_t[n]$, $p[n]$ y ε) el conocimiento o estimación de sus distribuciones de probabilidad convierte el problema bayesiano en un problema de optimización convexa.

Los parámetros que controlan la penalización aplicada sobre las componentes x_t y x_f se han seleccionado conforme a lo propuesto en [11], tomando valores 0,04 y 0,01, respectivamente.

2.4. Caracterización de la señal

Para las diferentes fases del estudio, cada señal y sus componentes x_t y x_f son caracterizadas a través de sus parámetros estadísticos.

Para determinar la posición y el valor de los picos de ANSM se buscan los máximos locales de la señal y se comprueba que no se haya producido otra activación en un tiempo inferior la duración mínima de una RCD [7].

Medidas de tendencia central	Medidas de dispersión
Mediana, M_e	Rango intercuartílico, IQR
Potencia media picos ANSM, P_{ASMN}	Frecuencia picos ANSM, f_{ANSM}

Tabla 2. *Características extraídas de cada componente.*

2.5. Análisis estadístico

Sobre los conjuntos de datos obtenidos se aplicó el test de Kolmogorov-Smirnov para asegurar o descartar el ajuste de los mismos a la distribución normal.

Se realizaron comparaciones estadísticas de las distintas características obtenidas para las diferentes cohortes. Para la comparación entre los dos grupos, sujetos diagnosticados de depresión contra sujetos eutímicos, se realizó la prueba de "U" de Mann-Whitney con corrección de Bonferroni para adaptar el nivel de significancia al número de comparaciones efectuadas.

Para determinar la relevancia clínica de las comparaciones realizadas para cada una de las variables, se calcularon las curvas de rendimiento diagnóstico (ROC). La curva ROC proporciona una representación gráfica de la capacidad del análisis de una variable para ejercer una discriminación entre dos grupos. El área bajo la curva ROC (AUC)

proporciona un parámetro numérico de gran relevancia clínica para evaluar la bondad de las pruebas.

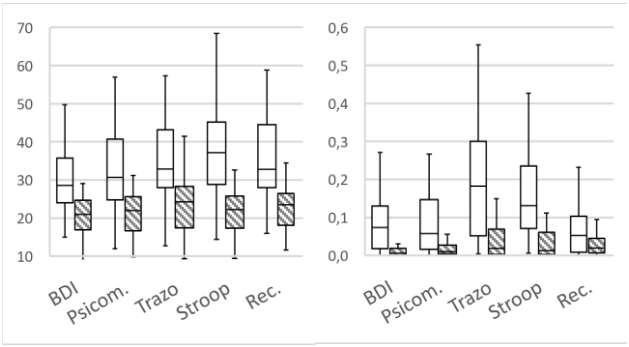
3. Resultados

Se encontraron diferencias significativas con una relevancia clínica elevada (AUC > 0,75) en los valores medianos M_e de las dos componentes de la AED, en todas las fases del estudio, a excepción de la fase de recuperación. En la fase de recuperación no se encontró significación estadística suficiente para distinguir entre las medianas de x_f de ambos grupos.

	Tónica	Fásica
BDI	0.79	0.83
Test psicométricos	0.79	0.76
Trazo	0.80	0.85
Stroop	0.84	0.85
Recuperación	0.81	< 0,65

Tabla 3. *Resultados de AUC para los valores de M_e de las componentes tónica y fásica*

Además, la distribución de M_e de la componente tónica muestra un aumento gradual a lo largo de las diferentes fases del estudio.



Tónica *Fásica*

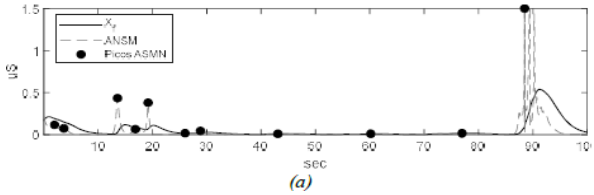
Gráfico 1. *Distribuciones de M_e para controles (blanco) y pacientes (trama).*

En todas las fases del estudio, exceptuando recuperación, se encontraron diferencias con significación estadística y relevancia clínica elevada entre ambas cohortes para las distribuciones IQR de ambas componentes.

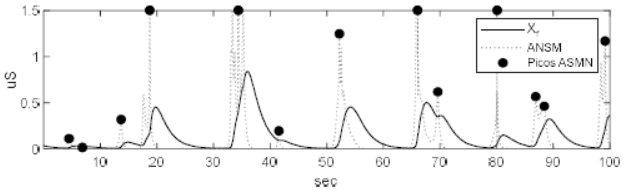
Se detectaron diferencias significativas con alto valor de relevancia clínica entre los grupos de pacientes y controles para las distribuciones de f_{SMNA} y de P_{SMNA} en todas las fases del estudio, exceptuando la fase de recuperación. Se comprobó que los controles presentaban una mayor frecuencia y potencia de picos de SMNA.

	P_{SMNA}	f_{SMNA}
BDI	0.84	0.80
Psicométricos	0.82	0.78
Trazo	0.83	0.82
Stroop	0.86	0.78

Tabla 4. *Resultados de AUC para los valores de las variables de ANS*



(a)



(b)

Ilustración 1*. Ejemplo comparativo de x_f y ANSM entre pacientes (a) y controles (b)*

4. Discusión

Estos resultados son acordes a las expectativas del planteamiento de nuestro estudio, ya que mientras los controles proporcionan respuestas a la estimulación "estresante" de las pruebas [7, 12-13], los pacientes al encontrarse en estado de sobreestimulación de su SNS son incapaces de proporcionar una respuesta electrodermal consecuente al estímulo recibido.

Un aspecto llamativo de la distribución de medianas de la componente tónica es que conforme las distintas pruebas van progresando en el tiempo los registros controles y pacientes van diferenciándose de manera más marcada. Este fenómeno puede deberse a la acumulación de Na^+ procedente de las glándulas ecrinas como respuesta del SNS y a la termorregulación fisiológica del organismo según las condiciones de humedad y temperatura [14-16]. Sin embargo, esta condición debería afectar por igual a pacientes y controles, cuando en los datos de nuestro estudio se observa esta tendencia claramente en los controles y es difícilmente perceptible en los pacientes. Este fenómeno explicaría porqué continúa habiendo diferencias estadísticamente significativas durante el periodo de recuperación cuando ni pacientes ni controles están recibiendo estímulo alguno.

Como era de esperar, hay mayor variabilidad de IQR en los controles que en los pacientes. La atenuación de la respuesta electrodermal en los pacientes deprimidos genera una menor variabilidad que en los controles eutímicos.

A diferencia de lo que ocurre en la componente tónica la componente fásica no se ve afectada por los efectos acumulativos o de adaptación por lo que proporciona una medida más consistente de la activación del SNS.

Se observa en los controles que la activación inducida en los test estresores, en los que se requiere una reacción ante un determinado estímulo, es mucho mayor que la respuesta obtenida durante la realización de los cuestionarios de salud mental, que, al ser autoadministrados, no requieren de reacciones rápidas o precisas.

En los pacientes, cuyas reacciones, como ya se ha señalado, son mucho menores que las de los controles, se aprecia el mismo patrón de activación (valores más elevados para los test estresores que para los cuestionarios de salud mental) pero de manera mucho más atenuada.

A diferencia de lo que ocurría en las comparaciones realizadas para la componente tónica, en la fase de recuperación no se encontraron diferencias con significación estadística entre pacientes y controles. Este hecho pudo ser debido a la misma razón: la progresión en

la activación de los controles tiene escaso efecto sobre la componente fásica.

Si bien los valores obtenidos de relevancia clínica para la componente fásica fueron ligeramente superiores, se encontraban en el mismo rango de calidad.

Como era de esperar, hay mayor variabilidad en los controles que en los pacientes en las medidas de desviación. La atenuación de la respuesta electrodermal en los pacientes deprimidos genera una menor variabilidad que en los controles eutímicos.

Si bien, las diferencias existentes no fueron siempre significativas, llama poderosamente la atención que sí lo son, consistentemente, cuando estudiamos la componente fásica de la señal (si exceptuamos la fase de recuperación).

La señal de ANSM se vincula con la activación del SNS como respuesta a un estímulo puntual. Esta señal es de gran utilidad en aquellos estudios que cuentan con las referencias temporales exactas de cuando se ha producido un estímulo puesto que permite relacionarlo con su respuesta correspondiente [17]. En este caso, no se cuenta con estas referencias temporales, los estímulos son pruebas de "larga" duración, pero las medidas obtenidas de ANSM pueden cuantificar el nivel de respuesta de los sujetos.

Las diferencias existentes, fueron similares a las obtenidas para la componente fásica: con significación estadística constante para frecuencia y potencia de los picos entre pacientes y controles en todas las fases del estudio salvo en la recuperación.

5. Conclusiones

El análisis realizado de la AED muestra diferencias estadísticamente significativas entre pacientes afectos de depresión y controles eutímicos

El modelo cvxEDA ha permitido la descomposición de la señal en sus componentes tónica y fásica y la obtención de la Señal de Actividad Nerviosa Sudomotora, proporcionando un número elevado de marcadores simpáticos que han resultado clínicamente relevantes a la hora de distinguir pacientes de controles.

El análisis independiente de la componente fásica y la señal de actividad nerviosa sudomotora permiten obviar el posible sesgo de la habituación y el patrón de termorregulación del organismo.

Agradecimientos

Este trabajo ha sido financiado por AEI, FEDER, EU y Gobierno de Aragón, mediante los proyectos RTI2018-097723-B-I00, LMP44-18, T39-20R, y por el ISCIII a través del Centro de Investigación Biomédica en Red (CIBER) y por el Programa Marco de Investigación e Innovación de la Unión Europea Horizonte 2020 (2014-2020) bajo el Acuerdo de subvención Marie Skłodowska-Curie nº 745755.

Referencias

[1] James SL, Abate D, Abate KH, et al. A systematic analysis for the Global Burden of Disease Study. *The Lancet*, vol. 392, 2018, pp 1789-1858 (ISSN: 0140-6736).

[2] World Health Organization. WHO International. https://www.who.int/news-room/fact-sheets/detail/depression (Consultada: Enero 2020).

[3] American Psychiatric Association. DSM-5. 2016 (ISBN: 9780890425541).

[4] Strawbridge R, Young AH, Cleare AJ. Biomarkers for depression: recent insights, current. *Neuropsychiatric Disease and Treatment*, vol 13, 2017, pp 1245-1262 (ISSN:1178-2021).

[5] Nuss P. Anxiety disorders and GABA neurotransmission: a disturbance of modulation. *Neuropsych. Disease and Treatment*, vol 11, 2015, pp 165-175 (ISSN: 1178-2021).

[6] Luscher B, Shen Q, Sahir N. The GABAergic deficit hypothesis of major depressive disorder. *Molecular Psychiatry*, vol 16, 2011, pp 383-406 (ISSN: 1359-4184)

[7] Hasset J. A Primer of Psychophysiology. W.H.Freeman & Co, 1978 (ISBN: 0716700379)

[8] Braithwaite, J. J., Watson, D. G., Jones, R., Rowe, M. A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs). *Psychophysiol.* vol. 49, 2013, pp 1017-1034 (ISSN 1469-8986).

[9] Kontaxis S, Gil E, Marozas V, et al. Photoplethysmographic Waveform Analysis for Autonomic Reactivity Assessment in Depression. *IEEE Transactions on Biomedical Engineering*, 2020 (ISSN: 0018-9294).

[10] Hu, Y., Converse, C, Lyons M.C., Hsu, W.H. Neural control of sweat secretion: a review. *British Journal of Dermatology*, vol 178, 2017, pp. 1246-1256 (ISSN: 1365-2133)

[11] Greco A., Valenza G., Lanata A., Scilingo E. P., Citi L. cvxEDA: A Convex Optimization Approach to Electrodermal Activity Processing. *IEEE Transactions on Biomedical Engineering*, vol. 63m, 2016, pp 797-804 (ISSN: 1558-2531).

[12] McCorry LK. Physiology of the Autonomic Nervous System. *American Journal of Pharmaceutical Education*, vol 71, 2007 (ISSN: 1553-6467)

[13] Cacioppo T, Tassinary G, Berntson G, editors. Handbook of Psychophysiology.Cambridge University Press, 2017, pp 217-243 (ISBN: 110705852X).

[14] Bari D, Aldosky H, Tronstad C, Kalvøy H, Martinsen Ø. Influence of Relative Humidity on Electrodermal Levels and Responses. *Skin Pharmacology and Physiology*, vol 31, 2018, pp 298-307 (ISSN: 1660-5527).

[15] Scholander T. Some measures of electrodermal activity and their relationships as affected by varied temperatures. *Journal of Psychosomatic Research*, vol. 7, 1963, pp 151-158 (ISSN: 0022-3999).

[16] Barrya RJ, Sokolovb EN. Habituation of phasic and tonic components of the orienting reflex. *International Journal of Psychophysiol.*, vol. 15, 1993, pp 39-42 (ISSN: 0167-8760).

[17] Greco A, Valenza G, Citi L, Scilingo EP. Arousal and Valence Recognition of Affective Sounds Based on Electrodermal Activity. *IEEE Sensors Journal*, vol. 17, 2016, pp 716-725 (ISSN: 1558-1748).

Evaluación de la robustez de la red neuronal funcional en la demencia por enfermedad de Alzheimer

M. Revilla-Vallejo [1], C. Gómez [1,2], V. Gutiérrez-de Pablo [1], M. A. Tola-Arribas [2,3], M. Cano [4], R. Hornero [1,2,5], J. Poza [1,2,5]

[1] Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, marcos.revilla@gib.tel.uva.es.
[2] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, (CIBER-BBN), Valladolid, España.
[3] Departamento de Neurología, Hospital Universitario Río Hortega, Valladolid, España.
[4] Departamento de Neurofisiología Clínica, Hospital Universitario Río Hortega, Valladolid, España.
[5] IMUVA, Instituto de Investigación en Matemáticas, Universidad de Valladolid, Valladolid, España.

Resumen

La demencia debida a enfermedad de Alzheimer (EA) es el tipo de demencia más habitual en los países occidentales. Caracterizar la red neuronal desde puntos de vista novedosos puede ser de gran interés para conocer las dinámicas subyacentes a la progresión de esta enfermedad. En este estudio, se pretende evaluar la robustez de la red neuronal construida a partir de registros de electroencefalograma. Para ello, se empleó una base de datos de 174 sujetos divididos en tres grupos: 43 sujetos de control cognitivamente sanos, 67 pacientes con deterioro cognitivo leve (DCL) y 64 pacientes con demencia por EA. Se construyó la red neuronal funcional a partir de la medida de conectividad funcional Amplitude Envelope Correlation. Para caracterizar la robustez de la red, se evaluó la variación de la eficiencia global de la misma a medida que era atacada empleando diferentes estrategias. Se observó que las redes de los pacientes con DCL y demencia por EA son más vulnerables a los ataques que las de los sujetos de control con todas las estrategias aplicadas. Se puede concluir por tanto que la progresión de la EA provoca cambios en la robustez de la red neuronal funcional, cuya cuantificación puede ser muy útil para entender los mecanismos neurodegenerativos subyacentes.

1. Introducción

La demencia debida a enfermedad de Alzheimer (EA) es el tipo de demencia más habitual en los países occidentales [1]. Es posible diferenciar tres etapas en lo que a la progresión de la enfermedad respecta: estadio preclínico, deterioro cognitivo leve (DCL) por EA y demencia debida a EA. El deterioro en la memoria episódica, característico del estadio preclínico, constituye uno de los primeros síntomas de la patología. Tras este, aparece una disfunción ejecutiva que afecta a las tareas cotidianas. A medida que la enfermedad avanza, los cambios estructurales a nivel cerebral son más notables y, en consecuencia, aparecen dificultades de comunicación y trastornos en la comprensión, consecuencia de la atrofia global del cerebro [1].

Las señales biomédicas obtenidas a partir de los registros de la actividad neuronal tienen un gran potencial para caracterizar las alteraciones subyacentes a estas patologías [2]. Tal es el caso de la electroencefalografía (EEG), que consiste en el registro de las oscilaciones del potencial eléctrico cerebral a través de electrodos colocados en el cuero cabelludo. Su bajo coste y elevada resolución temporal lo hace una herramienta muy adecuada para su uso en la práctica clínica diaria frente a otras, como la resonancia magnética funcional (fMRI), la tomografía por emisión de positrones (PET) o la magnetoencefalografía (MEG) [3].

La EA se ha identificado como un síndrome de desconexión en el que los patrones de conectividad cerebral se encuentran alterados [2]. Los registros de EEG permiten obtener patrones de conectividad funcional entre las diferentes áreas cerebrales. Las matrices de conectividad resultantes se han caracterizado habitualmente con parámetros de Teoría de Grafos, que dan cuenta de las alteraciones asociadas a las diferentes fases de progresión de la demencia por EA [4]. Sin embargo, para comprender de forma más profunda las complejas alteraciones que provoca esta patología en la red neuronal, es necesaria una caracterización que permita simular y cuantificar cómo esas redes neuronales funcionales se comportan ante ataques y perturbaciones. Surge en consecuencia el concepto de robustez, como la habilidad de una red para resistir a perturbaciones o ataques [5].

Nuestra hipótesis es que las alteraciones estructurales que provoca la progresión de la EA conllevan complejos cambios de la red cerebral que los parámetros de red tradicionales no son capaces de cuantificar completamente [2, 4]. Por tanto, el objetivo de este estudio es plantear una nueva metodología que permita simular las perturbaciones en la red neuronal funcional asociadas a la progresión de la EA y cuantificar su robustez.

2. Materiales

2.1. Sujetos

La base de datos empleada ha sido registrada en el Hospital Universitario Río Hortega de Valladolid y consta de un total de 174 sujetos: 43 sujetos de control cognitivamente sanos, 67 pacientes con DCL por EA y 64 pacientes con demencia debida a EA. La clasificación de los sujetos se realizó en base a los criterios NIA-AA (*National Institute on Aging and Alzheimer's Association*) [6]. Todos los pacientes, o sus familiares en caso de incapacidad, dieron su consentimiento expreso para la participación en el estudio, que se diseñó siguiendo el Código Ético de la Asociación Médica Mundial (Declaración de Helsinki) y de acuerdo con el protocolo aprobado por el

ISBN: 978-84-09-25491-0, pags. 260- 263

Comité Ético del Hospital Universitario Río Hortega.

La Tabla 1 resume los datos sociodemográficos de los tres grupos anteriormente citados. No existen diferencias estadísticamente significativas en la edad (p-valor > 0.05, test U de Mann-Whitney), ni en la distribución de sexos (p-valor > 0.05, test Chi-Cuadrado). En cambio, en los valores de *Mini-Mental State Examination* (MMSE) existen diferencias estadísticamente significativas entre los tres grupos (p-valor < 0.05, test U de Mann-Whitney), revelando el deterioro en las capacidades cognitivas consecuencia de la progresión de la EA.

2.2. Registro de EEG

El registro de la actividad eléctrica cerebral se realizó con un electroencefalógrafo de 19 canales (modelo XLTEK®, *Natus Medical*). Se solicitó a los sujetos que permanecieran sentados, en reposo y con los ojos cerrados, obteniendo un total de 5 min. de registro. Los electrodos se colocaron según el Sistema Internacional 10-20 con un referenciado promedio común [7].

Todas las señales se sometieron a un preprocesado consistente en los siguientes pasos [7]: (i) filtrado digital con un filtro FIR (*Finite Impulse Response*) en la banda 1-70 Hz; (ii) filtrado de ranura para eliminar la interferencia de 50 Hz; (iii) análisis de componentes indepedientes (ICA) para la detección y eliminación de artefactos [8]; y (iv) selección de épocas de 5 s libres de artefactos mediante inspección visual.

A partir de los registros, se calcularon las fuentes cerebrales correspondientes a 68 regiones corticales, haciendo uso del algoritmo de localización sLORETA (*Standarized Low - Resolution Brain Electromagnetic Tomography*) y del atlas Desikan-Killiany [9].

3. Métodos

3.1. Estimación de la conectividad

Para construir la red neuronal, se calculó la conectividad funcional entre las 68 regiones corticales consideradas. Para ello, se empleó la *amplitude envelope correlation* (AEC) ortogonalizada que caracteriza la correlación existente entre las amplitudes de las envolventes de dos señales [10]. En consecuencia, para cada segmento de 5 s se dispone de una matriz de tamaño 68×68 que da cuenta de las interacciones existentes entre las diferente regiones corticales. Las redes neuronales funcionales se construyeron únicamente para las bandas zeta (θ), 4-8 Hz; y beta (β), 13-30 Hz, dada la relevancia de las mismas a la hora de caracterizar la progresión de la EA [11].

3.2. Estimación de la robustez de la red

Un campo actual de estudio en el ámbito de la neurociencia está relacionado con el análisis de la respuesta de las redes cerebrales frente a ataques o perturbaciones externas. Esto es de gran interés, por ejemplo, para modelar las alteraciones inducidas por enfermedades neurodegenerativas como la EA. Sin embargo, para poder cuantificar estos aspectos, son necesarias nuevas medidas que permitan estimar la habilidad de la red para resistir ante ataques o fallos [5]. Surge así el concepto de robustez de red. La caracterización de la robustez es un campo abierto, si bien ya existen diversas medidas dirigidas a cuantificar esta propiedad [5, 12]. En este estudio, se considera la robustez basada en la eficiencia global (R_{Eff}).

Datos	Controles	DCL	EA
N	43	67	64
Edad (años)	76[74 79]	77[72 81]	82[77 83]
Sexo (H:M)	13:30	19:38	28:36
MMSE	29[28 30]	27[26 28]	22[19 24]
Bayer-ADL	1[1 1]	3[2 3]	6[5 7]
T@M	45[42 48]	32[27 35]	21[17 26]

Tabla 1. *Datos sociodemográficos y clínicos de los grupos analizados; N: número de sujetos; H: hombre; M: mujer; MMSE: Mini-Mental State Examination (rango: [0 30]); Bayer-ADL: Bayer Activities of Daily Living (rango [0 10]); T@M: Test de Alteración de Memoria (rango [0 50]). Edad, MMSE, Bayer-ADL, T@M: mediana [Q1 Q3]*

La eficiencia global es el inverso de la longitud de camino característico, entendiendo este último como el camino más corto promedio entre todos los pares de nodos en la red [4]. La eficiencia evita la limitación del valor infinito de longitud de camino característico cuando dos nodos se encuentran desconectados. Para la evaluación de la robustez, la red es atacada empleando una serie de estrategias que se detallan en el Apartado 3.3. Tras cada ataque, la eficiencia global de la misma se calcula. De esta forma, se obtiene la evolución de esta medida a lo largo de esa secuencia de ataques [5].

3.3. Estrategias de ataque a la red

En el marco de este estudio se han empleado cinco estrategias de ataque: eliminación de nodos de forma aleatoria (eliminación de nodos aleatoria, ENA) o con probabilidad proporcional al grado de nodo (eliminación de nodos primarios, ENP), disminución a la mitad del valor de todas las conexiones del nodo seleccionado de forma aleatoria (ataque a nodos aleatorio, ANA) o con probabilidad proporcional al grado (ataque a nodos primarios, ANP), y disminución a la mitad del valor de las conexiones seleccionadas con probabilidad proporcional a su valor (ataque a enlaces primarios, AEP). Para las estrategias de eliminación de nodos, la red es atacada hasta que todos los nodos son eliminados, un total de 68 veces. En el caso de las estrategias de ataque tanto a nodos como enlaces, se fija 150 como el número de ataques. Este valor se escoge para poder observar una alteración de la red cuantificable mediante la medida de robustez.

3.4. Análisis estadístico

Los datos no cumplían las hipótesis de normalidad y homocedasticidad por lo que se aplicaron tests no paramétricos para evaluar las diferencias estadísticas. Concretamente, se empleó el test U de Mann-Whitney a la variación de la R_{Eff} con respecto al valor inicial, a medida que la red era atacada. De esta forma fue posible

cuantificar la alteración de la red. Se muestran únicamente los resultados para la iteración 65 (estrategias de eliminación) y 150 (estrategias de ataque) dado que no existe apenas variación en los p-valores obtenidos a lo largo de todo el proceso de ataque, manteniéndose el patrón de diferencias que se muestra en el Apartado 4.

4. Resultados y discusión

En la Figura 1 se muestra la evolución de la R_{Eff} a medida que la red es modificada bajo las cinco estrategias de ataque en las bandas θ y β. De forma global, las redes neuronales responden de forma similar a los ataques en los dos rangos de frecuencia considerados, tal y como se observa en las tendencias de evolución de la robustez de red. Sin embargo, cabe destacar que las redes obtenidas en la banda β son menos eficientes inicialmente y muestran, comparativamente, una menor variación tras las 150 iteraciones de ataque. Esto sugiere que la red neuronal funcional es más robusta (*i.e.* más resistente a ataques) en la banda β que en la banda θ, hecho que da cuenta de la distinta afectación de la demencia por EA en diferentes bandas de frecuencia [10, 13]. La demencia por EA, a lo largo de sus diferentes fases, afecta a diversas funciones cognitivas. La banda β se encuentra estrechamente relacionada con el pensamiento activo y la atención, aspectos que se deterioran en fases más tardías de la enfermedad. En cambio, la banda θ tiene un papel importante en la resolución de problemas, aspecto que se deteriora de forma más temprana en la progresión de la demencia por EA [1]. Considerando de forma comparativa las diferentes estrategias de ataque, es posible apreciar como la red neuronal funcional responde de forma diferente a cada una de ellas. La mayor similitud se observa entre la ENP y la ENA. Las eliminaciones de nodos con grado elevado provocan variaciones similares a las que se producen como consecuencia de la eliminación aleatoria. Esta evolución se puede justificar considerando la uniformidad de la distribución de grado de nodo que caracteriza a las redes neuronales [5]. Es decir, no existen nodos con un grado excesivamente elevado o bajo. En definitiva, la funcionalidad cerebral no se concentra en una región concreta, sino que se reparte por toda la red, si bien también es posible encontrar determinadas regiones estrechamente relacionadas con ciertas funciones cognitivas [1]. A pesar de los cambios en la respuesta de la red antes los ataques, la evaluación de la robustez obtenida mediante estos es muy similar. En definitiva, y como era de esperar, las diferentes estrategias llevan a resultados similares, al ser la robustez una característica inherente a la red que no depende de cómo esta sea perturbada.

En la Tabla 2 se muestran las diferencias estadísticamente significativas en la variación de la eficiencia de la red de los tres grupos para la iteración final. Los resultados sugieren que la robustez se modifica significativamente a medida que la demencia por EA progresa. Las diferencias estadísticamente significativas se concentran en las comparaciones entre controles vs. sujetos con DCL y, fundamentalmente, entre controles vs. pacientes con EA. Los cambios estructurales característicos de la enfermedad [3] se manifiestan en la red neuronal funcional [2] y, por tanto, en la robustez de la misma. La pérdida de complejidad, la lentificación del espectro o la desconexión progresiva características de la EA [1] pueden provocar alteraciones en la robustez de la red cerebral siendo, en consecuencia, más vulnerable a los ataques que la red de los sujetos cognitivamente sanos.

	C vs. DCL	DCL vs. EA	C vs. EA
θ - ENP	**0.0396**	0.3607	**0.0027**
θ - ENA	**0.0414**	0.3656	**0.0028**
θ - ANP	0.0592	0.3930	**0.0053**
θ - ANA	**0.0408**	0.4059	**0.0026**
θ - AEP	**0.0402**	0.3607	**0.0026**
β - ENP	0.0507	0.1489	**0.0061**
β - ENA	0.0507	0.1489	**0.0027**
β - ANP	**0.0045**	0.1707	**0.0074**
β - ANA	0.0584	0.1281	**0.0032**
β - AEP	**0.0486**	0.1582	**0.0062**

Tabla 2. *p-valores resultado de aplicar el test U de Mann-Whitney a los valores de R_{Eff} de los tres grupos de sujetos en la iteración 65 (estrategias de eliminación) y 150 (estrategias de ataque). Los p-valores significativos ($p<0.05$) se resaltan en negrita.*

Además, la estrategia de ataque empleada condiciona la existencia de diferencias estadísticamente significativas entre los grupos. Atacar, por ejemplo, una red de un sujeto de control con cierta estrategia de ataque, puede hacer que esa red sea más similar a la red de un paciente con DCL o bien, a la de un paciente con demencia por EA. Esto sucederá si la estrategia de ataque se alinea con las alteraciones a nivel de red originadas por los cambios neurofisiológicos característicos de la demencia por EA [2, 3]. Por tanto, aquellas estrategias de ataque en las que apenas existen diferencias significativas en la robustez de la red pueden ser las que deterioren la misma de forma similar a como lo hace la progresión de la demencia por EA. Esto sería de utilidad para determinar los mecanismos de neurodegeneración subyacentes a esta enfermedad.

En el presente estudio se han encontrado ciertas limitaciones que es necesario comentar. En primer lugar, se ha escogido una única medida para la evaluación de la robustez: la eficiencia global. Existen variedad de aproximaciones a la robustez que, empleadas de forma complementaria, pueden aportar aspectos también interesantes. En segundo lugar, restringir la localización de fuentes a 68 ROIs y promediar las mismas, reduciendo la resolución espacial. En tercer y último lugar, analizar de forma desagregada las redes neuronales funcionales correspondientes a los diferentes ritmos cerebrales. El desempeño global del cerebro es consecuencia de la integración de todos ellos, por lo que una evaluación de la robustez de red considerando las contribuciones de forma conjunta puede ser más precisa.

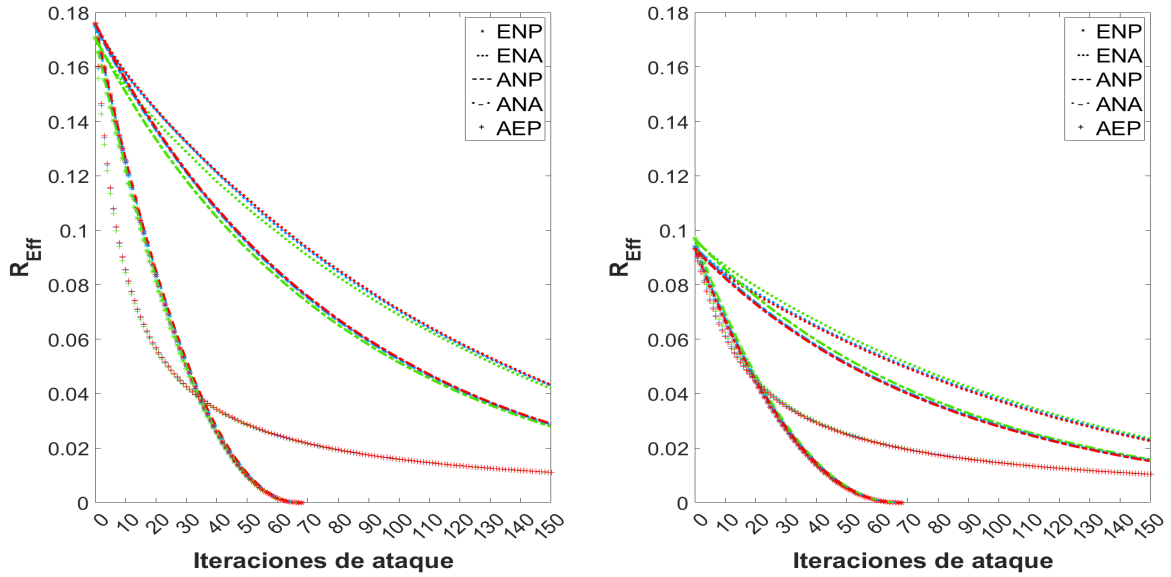


Figura 1. *Evolución de la eficiencia global (R_{Eff}) en la bandas θ (izquierda) y β (derecha) para las cinco estrategias de ataque. En verde, sujetos cognitivamente sanos; en azul, pacientes con DCL; en rojo, pacientes con EA.*

5. Conclusiones

Las diferencias observadas mediante la eficiencia global en los tres grupos analizados sugieren que la progresión de la demencia por EA provoca una reducción progresiva de la robustez de la red neuronal funcional, reflejando los cambios estructurales subyacentes y adquiriendo un papel clave en el conocimiento de los mecanismos neurodegenerativos asociados a la patología.

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 PGC2018-098214-A-I00, por la Comisión Europea y el FEDER bajo los proyectos 'Análisis y correlación entre el genoma completo y la actividad cerebral para la ayuda al diagnóstico de la enfermedad de Alzheimer' y 'Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres' ('Programa de Cooperación Interreg V-A España-Portugal POCTEP 2014-2020'), y por el CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto de Salud Carlos III co-financiado con fondos FEDER.

Referencias

[1] Alzheimer's Association. Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 15:321–387, 2019.

[2] C.J. Stam *et ál.* Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease. *Brain*, 132(Pt1):213–224, 2009.

[3] L. Rossini *et ál.* Early diagnosis of Alzheimer's disease: the role of biomarkers including advanced EEG signal analysis. Report from the IFCN-sponsored panel of experts. *Clin. Neurophysiol.*, 131(6):1287–1310, 2020.

[4] M. Rubinov and O. Sporns. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52(3):1059–1069, 2010.

[5] M. Bellingeri *et ál.* Robustness of weigthed networks. *Physica A*, 489:47–55, 2018.

[6] G.M. McKhann. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*, 7(3):263–269, 2012.

[7] A. Maturana-Candelas *et ál.* EEG Characterization of the Alzheimer 's Disease Continuum by Means of Multiscale Entropies. *Entropy*, 21(544):1–15, 2019.

[8] P. Common. Independent component analysis, A new concept? *Sign. Process.*, 36:287–314, 1994.

[9] V. Rodríguez-González *et ál.* Consistency of local activation parameters at sensor- and source-level in neural signals. *J. Neural Eng.*, Aceptado, 2020.

[10] P. Núnez *et ál.* Characterizing the fluctuations of dynamic resting-state electrophysiological functional connectivity: Reduced neuronal coupling variability in mild cognitive impairment and dementia due to Alzheimer's disease. *J. Neural Eng.*, 16(5):1–21, 2019.

[11] S. Ruiz-Gómez *et ál.* Measuring Alterations of Spontaneous EEG Neural Coupling in Alzheimer 's Disease and Mild Cognitive Impairment by Means of Cross-Entropy Metrics. *Front. Neuroinform.*, 12:1–11, 2018.

[12] Y. Jiang. The evaluation of complex networks'robustness based on entropy measure. *IET Conf. Publications*, 2014(CP653), 2014.

[13] L. Koelewijn *et ál.* Alzheimer's disease disrupts alpha and beta-band resting-state oscillatory network connectivity. *Clin. Neurophysiol.*, 128(11):2347–2357, 2017.

Imágenes Biomédicas I

Jueves, 26 de noviembre de 2020

Citometría de Flujo con Hoja Plana para el Conteo de Bacterias de Tuberculosis

S. Laguna[1], C. Gómez-Cruz[1], I. Albert-Smet[1], J.J. Vaquero[1,2], J.Ripoll[1,2], A. Muñoz-Barrutia[1,2]

[1] Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid, España.

[2] Instituto de Investigación Sanitaria del Hospital Gregorio Marañón, Madrid, España.

Resumen

La citometría de flujo es una tecnología láser para clasificar y analizar partículas y sus propiedades [1]. Dentro de sus aplicaciones está el conteo de bacterias en disolución. En este trabajo presentamos una nueva plataforma de citometría de flujo basada en dispositivos de microfluídica y microscopía de hoja plana para cuantificar bacterias de tuberculosis, Mycobacterium tuberculosis (Mtb), causantes de la tuberculosis en humanos, la enfermedad infecciosa más letal a nivel global. Este proyecto se desarrolla como parte de un esfuerzo por reducir el tiempo transcurrido en la maduración de nuevos compuestos contra la tuberculosis. Se plantea como una alternativa a los métodos tradicionales de conteo de bacterias, para obtener el número de aquellas vivas en disolución después de ser expuestas a nuevos compuestos. A su vez, este trabajo propone una arquitectura nueva de microscopía de hoja-plana para adquirir imágenes y un método de procesado dedicado al conteo de las bacterias a partir de los fotogramas obtenidos. Los resultados preliminares de los experimentos de prueba de concepto han sido evaluados dando lugar a una sensibilidad del 93.8 %.

1. Motivación y estado del arte

Cuantificación de bacterias

Tradicionalmente, el conteo de bacterias se realiza mediante unidades de formación de colonias (UFC), lo que conlleva un periodo de cultivo de semanas [2]. El objetivo de este proyecto es reducir ese tiempo a minutos para agilizar el proceso de desarrollo de nuevos compuestos contra la tuberculosis. Adicionalmente, proponemos una plataforma de microscopía que permite una rápida adquisición de imágenes en 3D con baja fototoxicidad.

La plataforma de citometría propuesta alberga dos componentes principales: Un microscopio de hoja plana y los dispositivos microfluídicos. En el estado del arte se han planteado plataformas similares para el conteo de partículas a diferentes niveles, como: Regmi *et al.* [3] con un sistema con microscopio de hoja plana que analiza partículas a nivel sub-celular sin enfocarlas espacialmente y Nawaz *et al.* [4], que analiza y enfoca en 3D partículas de un tamaño de 2.5 µm usando un microscopio confocal. En este proyecto se han adaptado estas ideas a una nueva arquitectura de microscopía de hoja plana y con el fin de medir partículas del tamaño y densidad de bacterias *Mtb*.

Microscopio SPIM

La microscopía plana de la iluminación (SPIM, por sus siglas en inglés), es un método de microscopía de fluorescencia de hoja plana. Las imágenes se adquieren al ser iluminadas por un plano de luz que incide lateralmente la muestra. El aspecto diferencial de SPIM es el uso de una lente cilíndrica para crear el plano de luz al ser atravesada por un rayo de laser [5]. Este sistema se caracteriza por una rápida adquisición de imágenes con poca fototoxicidad.

Dispositivos de microfluídica

La microfluídica es la tecnología que procesa o manipula fluidos a nivel microscópico mediante dispositivos con canales de diferentes geometrías y propiedades. En este proyecto se han utilizado dispositivos de PDMS (polidimetilsiloxano), un polímero impermeable al agua, transparente, no toxico para las células y permeable a gases, lo que lo hace ideal para ensayos biológicos [6]. El dispositivo microfluídico utilizado se encarga de enfocar espacialmente la disolución de bacterias para facilitar su detección y discretización y permitir su cuantificación.

2. Materiales y métodos

2.1. Modelos de microfluídica y simulaciones

Los dispositivos de microfluídica propuestos para la plataforma se han fabricado con PDMS y se basan en un diseño de Mao *et al.* [7] que consigue un enfocado hidrodinámico en 3D mediante un vórtice de Dean [8]. Esto asegura que todas las partículas fluyan en una única línea dentro del canal para una correcta discretización. El diseño propuesto ha sido estudiado teóricamente mediante simulaciones de flujo con el programa Ansys Fluent, para encontrar la configuración de flujos que optimizan el enfoque de las bacterias, resultando en: 0,65 µl/min para la entrada de bacterias, 3,3 µl/min para los laterales y 3 µl/min para la última entrada. La Figura 1 muestra la geometría del canal del dispositivo de microfluídica, así como la simulación llevada a cabo en la que se aprecia el enfoque de bacterias en el plano x-y. Estos dispositivos se probarán en el siguiente paso del proyecto.

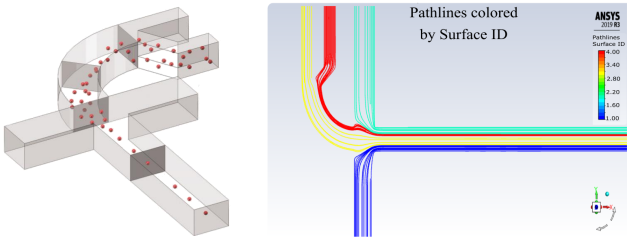


Figura 1. *(Izqda.) Canal del dispositivo de microfluídica propuesto por Mao et al. [7] con partículas enfocadas en rojo. (Dcha.) Resultado de la simulación de flujo. Las líneas rojas representan el canal de entrada de las bacterias que es enfocado en el plano por los canales adyacentes. Ansys Fluent es el software empleado.*

Como prueba de concepto de la plataforma se ha empleado un dispositivo con canales en forma de "Y" que no permite

ISBN: 978-84-09-25491-0, pags. 265- 268

el enfoque de las bacterias, por lo que no todas quedan discretizadas. Pese a ello, los primeros experimentos muestran un correcto flujo de partículas y permiten desarrollar un método de cuantificación. La Figura 2 enseña la estructura del microdispositivo en "Y".

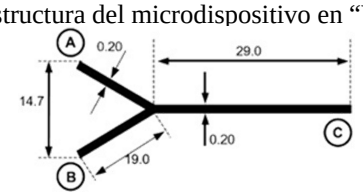


Figura 2. *Canales del dispositivo de microfluídica en forma de Y usado en los experimentos iniciales. Medidas en mm.*

2.2. SPIM personalizado

La plataforma SPIM propuesta está esquematizada en la Figura 2 con sus componentes, enumerados en la Tabla 1.

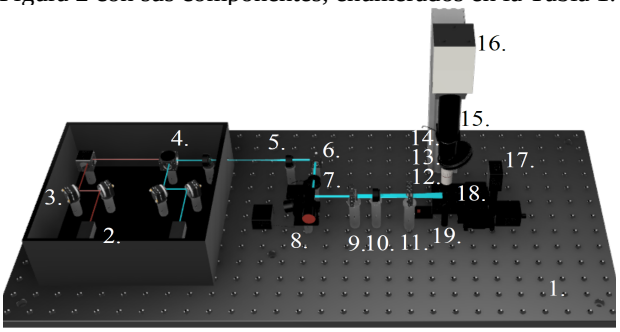


Figura 3. *Estructura del SPIM y componentes diseñada con Autodesk Fusion 360. Los ficheros CAD de componentes han sido obtenidos de Thorlabs y reproducidos con permiso. Copyright © 2020 Thorlabs. Los componentes están detallados en la Tabla 1. Un ordenador y un DAQ (no dentro del esquema) controlan la adquisición de imágenes.*

Nº	Componente	Modelo
1	Tabla óptica	Thorlabs PBG51506
2	Láseres	iBeam smart Ultra Compact
3	Espejos cinemáticos	Thorlabs KCB1/M
4	Filtro dicroico	Edmund optics 25mm Diam.
5	L. esférica,	Thorlabs AC254-030-A-ML
6	Espejo cinemático	Thorlabs KMSS/M
7	1ª Lente Optotune	Ed. EL-16-40-TC-VIS-5D-C
8	Galvo	Thorlabs GVS012/M
9	L. cilíndrica, f=50mm	Thorlabs LJ1695RM-A
10	L. esférica,f=100 mm	Thorlabs AC254-100-A-ML
11	L. cilíndrica, f=20mm	Thorlabs LJ1328L2-A
12	Objetivo 5x,10x	Mitutoyo 378-802-6,
13	2ª Lente Optotune	Ed. EL-16-40-TC-VIS-5D-C
14	Filtro de emisión	Edmund optics Dual-Band
15	Lente de tubo	Thorlabs TL200-CLS2
16	Cámara (sCMOS)	Hamamatsu C13440-20CU
17	Escenarios lineares	Zaber T-LSM050A
18	Plataforma de soporte	ZABER AP102B
19	LED	

Tabla 1. *Componentes y modelo específico del sistema SPIM.*

Estructura global

Los elementos de SPIM han sido montados sobre 2 tablas de aluminio de tamaños 600 x 900 mm y 450 x 600 mm. Estas tablas minimizan las inestabilidades térmicas y amortiguan las vibraciones externas para optimizar la calidad de imagen. La muestra se sujeta por una plataforma que está directamente unida a tres escenarios lineales que permiten el movimiento en 3 grados de libertad y la navegación debajo del objetivo. Adicionalmente, un dispositivo de adquisición de datos (DAQ) y un ordenador son centros de control, pese a no aparecer en la Figura 3.

Camino de excitación

La fuente de iluminación del sistema son dos láseres con longitudes de onda de 488 nm y 633 nm, un diámetro de 1 mm y una potencia máxima de 100 mW. Además, un LED colocado debajo de la muestra sirve como fuente suplementaria de luz blanca durante la adquisición. Los láseres están alineados y colimados sobre el mismo eje óptico usando dos espejos cinemáticos y un filtro dicroico de paso largo que refleja el láser de 488 nm y transmite el de 633 nm.

A continuación, hay una expansión de haz Kepleriana de 1 a 5 mediante dos lentes esféricas planoconvexas y acromáticas con distancias focales de 30 y 150 mm. En este punto, el eje óptico es rotado 90º con un espejo cinemático. Después, una lente de enfoque electrónico variable (Optotune) favorece el enfoque del haz debajo del objetivo.

Por último, el láser cruza el Galvo, un sistema de dos espejos motorizados que permiten el escaneado en 3D de la muestra, así como reducir la sombra en la imagen sin necesidad de mover la muestra. De esta forma se evitan fuerzas actuando sobre la plataforma. El láser sale de este sistema rotado 90º y vuelve a su orientación original. Al salir del Galvo, el láser cruza tres lentes: Una cilíndrica plano-convexa con distancia focal de 50 mm, encargada de crear el plano de luz; una esférica con distancia focal de 100 mm, para enfocar debajo del objetivo y finalmente, una segunda lente cilíndrica para redirigir la hoja de luz. Aquí, la hoja de luz ilumina lateralmente una sección de la muestra, permitiendo una fototoxicidad baja cuando se compara con otros sistemas de microscopía confocal.

Camino de emisión

La luz es emitida por la muestra y recogida en la torre de la cámara colocada perpendicularmente al plano de excitación. El primer elemento que incluye la torre es el objetivo. Concretamente, dos magnificaciones diferentes han sido caracterizadas: Un objetivo 5x de aire plano-apocromático y un 10x de agua semiapocromático. Ambos se pueden cambiar usando la rueda de objetivos. A continuación, hay una segunda lente de enfoque electrónico variable que permite que haya un cambio coordinado en el enfoque en la dirección z y el movimiento del láser, mejorando el sistema de control del escáner en 3D. El siguiente componente es un filtro óptico de emisión de banda de paso dual, en un rango de 577 ± 12nm y 690 ± 25 nm, que corresponden a las longitudes de onda de emisión de los fluoróforos excitados por los láseres de 488 nm y 633 nm, respectivamente.

Posteriormente, hay una lente de tubo, un tipo de lente esférica con la imagen en el infinito, cuyo fin es recibir los rayos paralelos del objetivo y reenfocarlos en una cámara. Finalmente, la luz llega a la cámara del sistema. Concretamente, es un sensor digital CMOS científico (sCMOS) que combina las arquitecturas de CCD y CMOS. Una vez el sensor lee la señal, un voltaje proporcional a la intensidad recibida se envía al ordenador mediante dos conectores. La cámara utilizada tiene una frecuencia de adquisición de 100 fps, un sensor de 2048 x 2048 píxeles y

un rango dinámico de 37000:1. Por último, en los experimentos se ha usado un binning de 2 x 2.

Software

El software que controla el microscopio es Opt3D, programado en LabView y que realiza el primer post-procesamiento para la reconstrucción de imágenes en 3D.

2.3. Experimentos de citometría

El sistema de microscopía ha sido adaptado a un citómetro de flujo mediante los dispositivos de microfluídica y un sistema de tubos y bombas (Jeringa motorizada Graseby MS 16A) para la circulación del flujo de bacterias. Durante los experimentos de prueba de concepto, se han usado microesferas fluorescentes de 1 µm de diámetro en lugar de bacterias, por razones de seguridad. Todas las imágenes han sido tomadas usado el objetivo de magnificación 5x, el láser de 488 nm y el microdispositivo en forma de Y que aparece en la Figura 2. Este dispositivo no enfoca las partículas en 3D, por lo que se esperan alteraciones en la cuantificación que se arreglarán con el microdispositivo final, visto en la Figura 1. En la adquisición, la hoja plana queda fija en una sección por la que cruzan las bacterias enfocadas en 3D. De esta forma se aumenta la frecuencia de adquisición, clave para cuantificar todas las bacterias que cruzan los canales.

Procesamiento de imagen

Las imágenes obtenidas de los experimentos iniciales permiten desarrollar un método de conteo de bacterias a partir de los fotogramas. El método propuesto incluye 6 pasos, esquematizados y con ejemplos en la Figura 4.

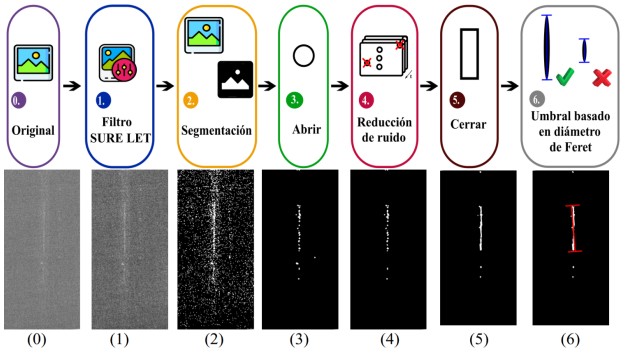


Figura 4. *(Sup.) Pasos del método de procesado para la cuantificación, explicados del 1-6 en el texto a continuación. (Inf.) Ejemplo de una imagen del SPIM en cada paso. El rectángulo de señal en (6) corresponde a una microesfera fluyendo en el canal.*

1. Filtro SURE LET: La imagen es filtrada para reducir el ruido con el filtro SURE-LET propuesto por Blu et al. [9]. Se basa en una aproximación a la minimización por error cuadrático medio y la transformada Wavelet.

2. Segmentación por umbral: Un umbral de 424 au ha sido determinado experimentalmente con una potencia de laser de 10 mW para segmentar globalmente la imagen.

3. Morfología matemática, abrir: Abrimos la imagen con un elemento estructural circular con radio de 2 píxeles para eliminar las partículas que no corresponden a la señal.

4. Reducción de ruido con fotograma anterior: Elimina las partículas cuyas coordenadas coinciden con aquellas en fotogramas anteriores en el tiempo, que corresponden a ruido constante de la adquisición o a esferas pegadas en las paredes del canal.

5. Morfología matemática, cerrar: Las microesferas aparecen como rectángulos de luz, ya que la imagen es obtenida mientras fluyen por el canal. Por ello, cerrar con un elemento estructural rectangular de tamaño [20,1] actúa como un filtro de paso bajo con las frecuencias verticales. Esto hace que las partículas de la imagen que aparecen muy cercanas verticalmente se unan para dar lugar a una única partícula final rectangular. Esta partícula corresponderá a la microesfera desplazándose a lo largo del canal.

6. Umbralización basada en el diámetro de Feret: El diámetro de Feret de las partículas es comparado con un umbral de 307 píxeles. Si el diámetro de Feret es mayor, una microesfera se añade a la cuenta total. Este valor se ha determinado considerando la velocidad de la disolución, 0.1 m/s, la frecuencia de la cámara, 100 fps y el campo de visión de 1,6 mm. Por consiguiente, la esfera recorre alrededor de 1 mm por fotograma, cruzando un 60% el campo de visión, 614 píxeles (1024*0.6). Estableciendo de umbral la mitad, 307, se cuantifica la esfera si ha estado en la imagen durante más de la mitad de la adquisición.

7. Cuantificación y análisis de datos: Cuatro vídeos de 50 fotogramas han sido procesados y los resultados analizados frente a referencias manualmente etiquetadas.

3. Resultados

Caracterización del microscopio

La configuración final del SPIM aparece en la Figura 5.

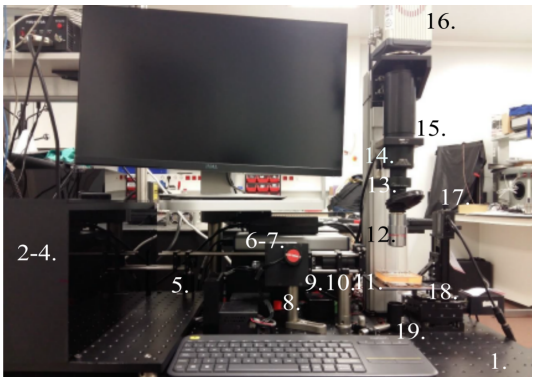


Figura 5. *Configuración del SPIM en el laboratorio, componentes numerados detallados en la Tabla 1 e incluye el ordenador, centro de contro . Del 2-4 aparecen cubiertos para proteger el sistema de láseres y 6-7 quedan tapados por el Galvo (8). El esquema en la Figura 3 puede ser usado para una mejor visualización del sistema.*

Los dos objetivos caracterizados dan lugar a un campo de visión de 0,84 x 0,84 mm para el 10x y de 1,61 x 1,61 mm para el 5x. Estas áreas permiten el desarrollo de la plataforma de citometría. La resolución espacial de este sistema óptico se obtiene mediante el ancho completo a la mitad del máximo (FWHM) de la función de dispersión de punto (point spread function, PSF) de una microesfera fluorescente. En este caso, se han usado esferas de 1 µm de diámetro disueltas en agua destilada. La resolución espacial se ha obtenido con la PSF de los perfiles de intensidad de las diagonales que cruzan el píxel de máxima intensidad, con un ejemplo del proceso en la Figura 6.

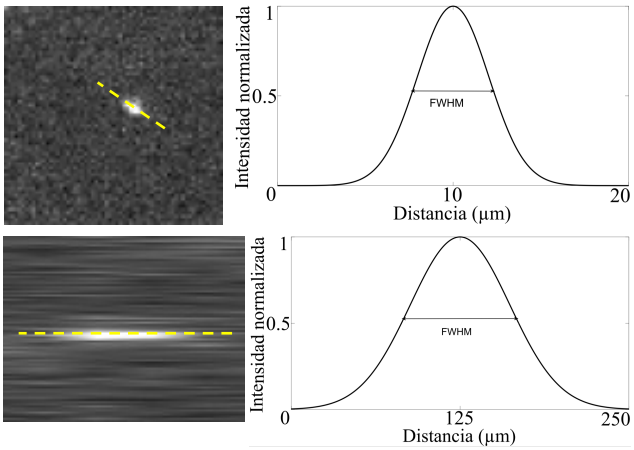


Figura 6. *Microesferas de 1 µm usadas en la computación de la resolución espacial con el objetivo de 5x. Las gráficas representan la función gaussiana que se ajusta a la media de los perfiles de intensidad de las microesferas, marcados en amarillo. (Sup.) Sección x-y de la microesfera. (Inf.) Sección x-z de la microesfera.*

La resolución espacial obtenida con el objetivo de magnificación 5x y microesferas de 1 µm es de 3,71 µm en el plano x-y y de 63,52 µm en la dirección z. Por otro lado, usando el objetivo de magnificación 10x con microesferas de 0.5 µm, se ha obtenido una magnificación de 2,64 µm en el plano x-y y de 34,19 µm en la dirección z.

Citometría

La plataforma ha sido adaptada correctamente a un citómetro de flujo y un método de cuantificación es capaz de detectar las partículas. Cuatro videos se han analizado para evaluar el método de procesado. La precisión (A_{cc}), sensibilidad (S_e), especificidad (S_p), valor predictivo positivo (P^+) y valor predictivo negativo (P^-) se han calculado globalmente para evaluar su rendimiento con los verdaderos positivos (T_P), verdaderos negativos (T_N), falsos positivos (F_P) y falsos negativos (F_N). En nuestro caso, T_P=30, T_N=152, F_P=14 and F_N=2, resultando en:

$$A_{cc} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} = \mathbf{91,9\%} \quad (1)$$

$$S_e = \frac{T_P}{T_P + F_N} = \mathbf{93,8\%} \quad S_p = \frac{T_N}{T_N + F_P} = \mathbf{91,6\%} \quad (2)$$

$$P^+ = \frac{T_P}{T_P + F_P} = \mathbf{68,2\%} \quad P^- = \frac{T_N}{T_N + F_N} = \mathbf{98,7\%} \quad (3)$$

4. Discusión y conclusiones

El sistema de microscopía SPIM desarrollado está listo para ser empleado en diversas aplicaciones como estudios de organoides, lab-on-chip o la caracterización anatómica de organismos [10]. Además, el desarrollo de la plataforma de citometría permite tener dos herramientas muy valiosas en un único sistema, reduciendo el coste y espacio ocupado en el laboratorio, un factor clave en muchos grupos.

Los resultados del conteo de partículas muestran valores generales altos, con excepción de un P^+ de 68.2%. Esto significa que hay una sobredetección de partículas debida a la falta de enfoque de las partículas en el dispositivo de microfluídica de los experimentos iniciales (Figura 2) que hace que algunas se cuenten de manera repetida. Esta sobredetección se muestra ligeramente en la especificidad del 91.6% debido a los falsos positivos. Pese a ello, una sensibilidad de 93.8% muestra que la mayoría de las partículas son correctamente detectadas y que el método propuesto es adecuado. El dispositivo de microfluídica propuesto (Figura 1) tiene resultados teóricos satisfactorios que consiguen enfocar partículas del tamaño de bacterias en 3D, y será utilizado para obtener más vídeos con los que optimizar el protocolo de análisis de imagen descrito para hacerlo más robusto. Tras probar experimentalmente este dispositivo, se esperan resultados favorables del método de conteo y que el sistema pueda ser usado con bacterias *Mtb*.

En conclusión, este proyecto muestra que la microscopía de hoja plana puede ser adaptada correctamente para funcionar como una plataforma de citometría de flujo. Además, esta plataforma cuantifica partículas en disolución satisfactoriamente y después de ser probada con los dispositivos finales, podría reemplazar los métodos tradicionales de conteo de *Mtb* y agilizar el proceso de desarrollo de nuevos compuestos contra la tuberculosis.

Agradecimientos

Ministerio de Ciencia, Innovación, PID2019-109820RB-I00; Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación, TEC2016-78052-R / MINECO/FEDER, UE; This project has received funding from 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.

Referencias

[1] A. Adan et al., "Flow cytometry: Basic principles and applications", Critical reviews in biotechnology, vol. 37, no. 2, pp. 163–176, 2017.

[2] G. J. De Knegt et al., "Assessment of treatment response by colony forming units, time to culture positivity and the molecular bacterial load assay compared in a mouse tuberculosis model", Tuberculosis, vol. 105, pp. 113–118, 2017.

[3] R. Regmi, et al., "High resolution light-sheet based highthroughput imaging cytometry system enables visualization of intra-cellular organelles", AIP Advances, vol. 4, no. 9, p. 097 125, 2014.

[4] A. A. Nawaz et al., "Sub-micrometer-precision, three-dimensional (3d) hydrodynamic focusing via "microfluidic drifting"", Lab on a Chip, vol. 14, no. 2, pp. 415– 423, 2014.

[5] J. Huisken et al., "Selective plane illumination microscopy techniques in developmental biology", Development, vol. 136, no. 12, pp. 1963–1975, 2009.

[6] S. K. Sia et al., "Microfluidic devices fabricated in poly (dimethylsiloxane) for biological studies", Electrophoresis, vol. 24, no. 21, pp. 3563–3576, 2003.

[7] X. Mao et al., "Single-layer planar on-chip flow cytometer using microfluidic drifting based threedimensional (3D) hydrodynamic focusing," Lab Chip, vol. 9, no. 11, pp. 1583–1589, Jun. 2009.

[8] N. Nivedita, P. Ligrani, and I. Papautsky, "Dean flow dynamics in low-aspect ratio spiral microchannels", Scientific reports, vol. 7, no. 1, pp. 1–10, 2017.

[9] T. Blu and F. Luisier, "The sure-let approach to image denoising", IEEE Transactions on Image Processing, vol. 16, no. 11, pp. 2778–2786, 2007.

[10] I. Albert-Smet et al., "Applications of light-sheet microscopy in microdevices",Frontiers in Neuroanatomy, vol. 13, p. 1, 20

Reconstrucción y validación de imágenes electrocardiográfica: Problema inverso

M. Borràs Argemí[1], J. Chamorro Servent[1]

[1] BCN-MedTech, Departamento de Tecnologías de la Información y la Comunicación, Universitat Pompeu Fabra, Barcelona, España, mborrasargemi@gmail.com, judit.chamorro.servent@gmail.com

Resumen

Cada año un número considerable de personas fallecen en los países desarrollados a causa de problemas cardíacos, precisando así de técnicas de diagnóstico y tratamiento. La técnica de imagen electrocardiográfica (ECGI), se utiliza para reconstruir de manera no-invasiva señales electrofisiológicas del corazón y mapear la estimulación cardíaca en relación con su anatomía. La solución del problema inverso del ECGI depende de la relación entre los potenciales en la superficie cardíaca y los potenciales en el torso (problema directo). En este trabajo, hemos aplicado el método de soluciones fundamentales (MFS) para resolver el problema directo. Seguidamente, hemos transferido y adaptado cuatro métodos de problema inverso a la reconstrucción del ECGI: la técnica de reconstrucción algebraica (ART), la ART con inicialización aleatoria (randART), la ART Split Bregman (ART-SB) y el método de restricción de rango del mínimo residuo generalizado (RRGMRES). Todos estos métodos los hemos testeado con datos del repositorio de datos experimentales y análisis geométrico (EDGAR). Hemos comparado las respectivas reconstrucciones con los potenciales registrados en el epicardio, proporcionados por EDGAR, y con la correspondiente reconstrucción obtenida vía el método del residuo mínimo generalizado (GMRES), utilizado con éxito previamente en ECGI. Los resultados muestran que la ART proporciona una solución más estable y con diferencias despreciables respecto la randART. La mejor reconstrucción (morfológicamente) se obtiene con el método RRGMRES. En conclusión, se ha probado que ART, randART y RRGMRES mejoran GMRES.

1. Introducción

En Europa y Norte América, por cada 100000 personas, se producen, cada año, entre 50 y 100 muertes repentinas causadas por problemas de corazón [1].

Actualmente se están estudiando técnicas no invasivas para identificar pacientes en riesgo de problema cardíaco, proporcionar un diagnóstico preciso y realizar terapias de guía de ablación [2]. Una de las técnicas no invasivas más utilizadas, y en la cual hemos centrado este trabajo es el estudio de imágenes electrocardiográficas (ECGI) [3]. El ECGI es una modalidad de imágenes que se utiliza para reconstruir las señales electrofisiológicas del corazón y delimitar una zona del epicardio, considerada de interés diagnóstico o terapéutico, a través del mapeo de dichas señales [2-4].

A pesar del éxito de la técnica ECGI en los últimos años, la comprensión y el tratamiento de muchas enfermedades cardíacas aún no es factible sin una mejora de la solución del problema inverso [5]. Dicho problema inverso es muy mal condicionado y requiere de métodos de regularización o algoritmos iterativos para que su solución sea única y estable.

El objetivo principal de este artículo es transferir métodos de problema inverso, usados en diferentes aplicaciones de bioingeniería, para aplicarlos al ECGI. Queremos adaptar y probar métodos existentes en otros campos (centrándonos en algoritmos iterativos), para reconstruir potenciales cardíacos en el epicardio y comparar los resultados obtenidos en términos de amplitud y morfología de los potenciales, así como los correspondientes mapas isocrónicos.

Para ello, primero hemos resuelto el problema directo aplicando el método de soluciones fundamentales (MFS) [6], un método sin mallado. Seguidamente, hemos transferido y adaptado cuatro métodos de problema inverso al ECGI, nunca usados en dicho campo. Y finalmente, hemos aplicado el método del residuo mínimo generalizado (GMRES), el cual ya había sido probado en ECGI, mostrando un mejor rendimiento que los métodos de regularización vía Tikhonov [7]. Estos métodos han sido testeados con diferentes datos proporcionados por el repositorio de datos experimentales y análisis geométrico (EDGAR): un archivo de código abierto con datos revisados, distribuido libremente a la comunidad internacional de investigación, proporcionando así una herramienta para la aplicación y validación de técnicas de ECGI [8].

Las reconstrucciones de los potenciales hechas por cada uno de los métodos mencionados anteriormente, para cada conjunto de datos experimentales, han sido comparadas con sus respectivos potenciales de referencia (proporcionados por EDGAR) y los resultados reconstruidos con GMRES como se explica en la sección 2.4.

Los resultados se han expuesto en la sección 3 y las conclusiones han sido extraídas en la sección 4.

2. Metodología

2.1. Problema directo

Los potenciales eléctricos, $\mathbf{\Phi}$, se pueden expresar mediante el MFS [6] como una combinación lineal de las

ISBN: 978-84-09-25491-0, pags. 269- 272

soluciones fundamentales de Laplace, $\boldsymbol{f}$, mediante un conjunto $\boldsymbol{Ns}$ de fuentes (o puntos) virtuales ($\left(\boldsymbol{y_j}\right)_{\boldsymbol{j=1..N_S}} \notin \Omega$):

$$\boldsymbol{\Phi(x) = a_0 + \sum_{j=1}^{N_S} f\left(|x - y_j|\right) a_j,} \qquad (1)$$

para todo punto $\boldsymbol{x} \in \Omega$, y sus respectivos coeficientes $\left(\boldsymbol{a_j}\right)_{\boldsymbol{j=1..N_S}}$. Siendo $\boldsymbol{f(x, y_j) = \frac{1}{4\pi r}}$, donde $\boldsymbol{r = |x - y_j|}$ es la distancia euclídea 3D.

Las fuentes virtuales $\left(\boldsymbol{y_j}\right)_{\boldsymbol{j=1..N_S}}$, son previamente situadas fuera del dominio de interés, Ω, y de sus respectivos contornos (siendo $\boldsymbol{\Gamma_E}$ el contorno del epicardio y $\boldsymbol{\Gamma_T}$ el contorno del torso, respectivamente). Su localización se obtiene mediante una deflación de los puntos $\left(\boldsymbol{x_i^E}\right)_{\boldsymbol{i=1,2,\cdots,N_E}}$ de $\boldsymbol{\Gamma_E}$ respecto el centro geométrico del corazón (utilizando un factor numérico de 0.8). Y, similarmente, inflando las posiciones de los electrodos $\left(\boldsymbol{x_i^T}\right)_{\boldsymbol{i=1,2,\cdots,N_T}}$ de $\boldsymbol{\Gamma_T}$, des del mismo centro, usando un factor numérico de 1.2 [6].

Mediante (1) y aplicando las condiciones de contorno de Dirichlet y Neumann homogéneas para todo $\boldsymbol{x} \in \boldsymbol{\Gamma_T}$, se puede escribir un sistema matricial [6].

Finalmente, un algoritmo de problema inverso se debe aplicar para encontrar los coeficientes de las fuentes virtuales, $\left(\boldsymbol{a_j}\right)_{\boldsymbol{j=1..N_S}}$, y reconstruir los potenciales eléctricos en el epicardio, $\boldsymbol{\Phi}_{\mathrm{E}}$, aplicando (1) para todo $\boldsymbol{x} \in \boldsymbol{\Gamma_E}$.

2.2. Problema inverso

Los métodos iterativos usados para resolver el problema inverso del ECGI (nunca antes aplicados en este campo) son: i) la técnica de reconstrucción algebraica (ART) y la ART con inicialización aleatoria (randART) [9], ambas usadas en imagen médica para sistemas indeterminados y mal condicionados [10]; ii) la ART Split Bregman (ART-SB) [11], que fue aplicada para observar los resultados de combinar la ART con una eliminación de ruido en cada iteración; iii) el método de restricción de rango del mínimo residuo generalizado (RRGMRES), que ha sido aplicado en otros campos para eliminar el ruido proveniente de los datos e introducido por la técnica del GMRES [12].

El método GMRES fue también aplicado para usarlo como referencia, y poder comparar así los métodos propuestos respecto estudios realizados sobre el problema inverso del ECGI.

2.3. Datos experimentales

Las bases de datos usadas en este trabajo para probar el método de problema directo, así como los diferentes algoritmos propuestos para resolver el problema inverso del ECGI, han sido obtenidos del repositorio de datos experimentales y análisis geométrico (EDGAR) [8].

Se han usado cuatro datos diferentes de dos bases de datos creadas por diferentes grupos de investigación: la base de datos de la universidad de Maastricht, cuyos potenciales fueron tomadas a un perro normal anestesiado [13] y la creada por el instituto de bioingeniería de Auckland, en el que los potenciales fueron tomados a un cerdo [14]. Ambas disponen de señales tomadas simultáneamente en el torso y en el epicardio, para el ritmo sinusoidal normal "sinus", y para el ritmo estimulado "paced" des del ápice epicárdico del ventrículo izquierdo. Estas bases de datos tienen características diferentes en cuanto al número de nodos del mallado tan del torso como del corazón (equivalente al número de electrodos usados en la toma de señales): la base de datos de Maastricht dispone de un mallado del torso con 135 nodos y del corazón con 83 nodos para "sinus" y 65 para "paced"; mientras que la de Auckland está formada por un mallado con 142 nodos en el torso y 239 en el corazón. Por simplicidad, dichos nodos fueron utilizados en nuestro método sin mallado (2.1) como x_i^T y x_i^E, respectivamente.

2.4. Análisis estadístico

Primeramente, hemos realizado un análisis de convergencia de cada una de las técnicas iterativas usadas para resolver el problema inverso. Una vez hemos verificado que todas ellas convergían (anotando la iteración de convergencia), hemos procedido con el análisis estadístico.

Para comparar los resultados de los diferentes métodos de problema inverso propuestos hemos calculado los coeficientes de correlación (CC) y el error cuadrado medio relativo (RE) entre los potenciales reconstruidos ($\boldsymbol{\Phi_{CE_i}}$) y sus correspondientes potenciales de referencia medidos en el epicardio ($\boldsymbol{\Phi_{TE_i}}$). Se calcularon ambos para cada paso del tiempo de escala proporcionado, siguiendo las ecuaciones siguientes:

$$\boldsymbol{CC = \frac{\sum_{i=1}^{N_{L^*}} \left(\Phi_{TE_i} - \overline{\Phi_{TE_t}}\right)\left(\Phi_{CE_i} - \overline{\Phi_{CE_t}}\right)}{\sqrt{\sum_{i=1}^{N_{L^*}} \left(\Phi_{TE_i} - \overline{\Phi_{TE_t}}\right)^2} \sqrt{\sum_{i=1}^{N_{L^*}} \left(\Phi_{CE_i} - \overline{\Phi_{CE_t}}\right)^2}}} \qquad (2)$$

$$\boldsymbol{RE = \sqrt{\frac{\sum_{i=1}^{N_{L^*}} \left(\Phi_{CE_i} - \Phi_{TE_i}\right)^2}{\sum_{i=1}^{N_{L^*}} \left(\Phi_{TE_i}\right)^2}}} \qquad (3)$$

El CC más alto representa una mejor morfología y el RE más bajo una mejor amplitud de los potenciales reconstruidos. Hemos calculado también los cuartiles (Q1, Q2 = mediana y Q3) del CC y del RE para cada técnica aplicada en cada base de datos.

De esta manera, hemos comparado los resultados de las diferentes técnicas de problema inverso entre ellos. Hemos tomado como "método estándar" el GMRES, para poder comparar así las otras técnicas con un método ya aplicado antes en ECGI.

3. Resultados

En las Tablas 1 y 2 se muestran los cuartiles obtenidos al calcular los CC y los RE de cada técnica de problema inverso aplicada a todas las bases de datos mencionadas en el apartado 2.3.

Base de datos de Maastricht		Q1		Mediana		Q3	
		CC	RE	CC	RE	CC	RE
GMRES	Sinus	0,28	0,80	0,51	0,94	0,72	1,13
	Paced	0,21	0,91	0,33	0,98	0,69	1,06
ART	Sinus	0,48	0,70	0,67	0,87	0,76	1,09
	Paced	0,34	0,77	0,61	0,88	0,75	1,05
Rand-ART	Sinus	0,48	0,70	0,67	0,87	0,76	1,09
	Paced	0,34	0,77	0,61	0,88	0,75	1,05
ART-SB	Sinus	0,39	0,75	0,65	0,88	0,76	1,14
	Paced	0,39	0,82	0,67	0,90	0,80	0,98
RRGMRES	Sinus	0,30	0,76	0,61	0,95	0,75	1,13
	Paced	0,33	0,77	0,60	0,92	0,78	1,08

Tabla 1. *Cuartiles de los errores relativos (RE) y de los coeficientes de correlación (CC) para cada técnica de reconstrucción aplicada a la base de datos de Maastricht*

Base de datos de Auckland		Q1		Mediana		Q3	
		CC	RE	CC	RE	CC	RE
GMRES	Sinus	0,53	0,74	0,73	0,83	0,83	0,92
	Paced	0,65	0,49	0,87	0,73	0,94	0,88
ART	Sinus	0,52	0,64	0,79	0,79	0,89	0,95
	Paced	0,69	0,51	0,87	0,67	0,94	0,86
Rand-ART	Sinus	0,52	0,68	0,79	0,79	0,89	0,96
	Paced	0,69	0,50	0,88	0,68	0,94	0,86
ART-SB	Sinus	0,34	0,71	0,70	0,88	0,85	1,00
	Paced	0,65	0,58	0,84	0,75	0,92	0,87
RRGMRES	Sinus	0,58	0,62	0,80	0,76	0,91	0,90
	Paced	0,74	0,39	0,88	0,62	0,95	0,84

Tabla 2. *Cuartiles de los RE y de los CC para cada técnica de reconstrucción aplicada a la base de datos de Auckland*

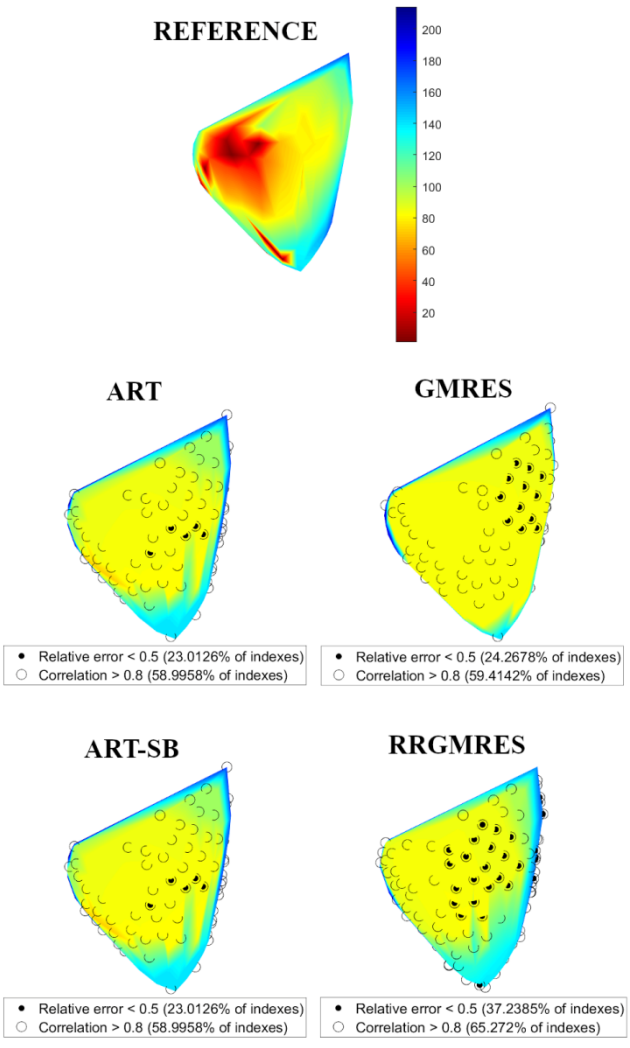


Figura 1. *Mapas isocrónicos de la cara diestra del epicardio, obtenidos para la base de datos de Auckland con el ritmo estimulado "paced", y localizando las mejores reconstrucciones*

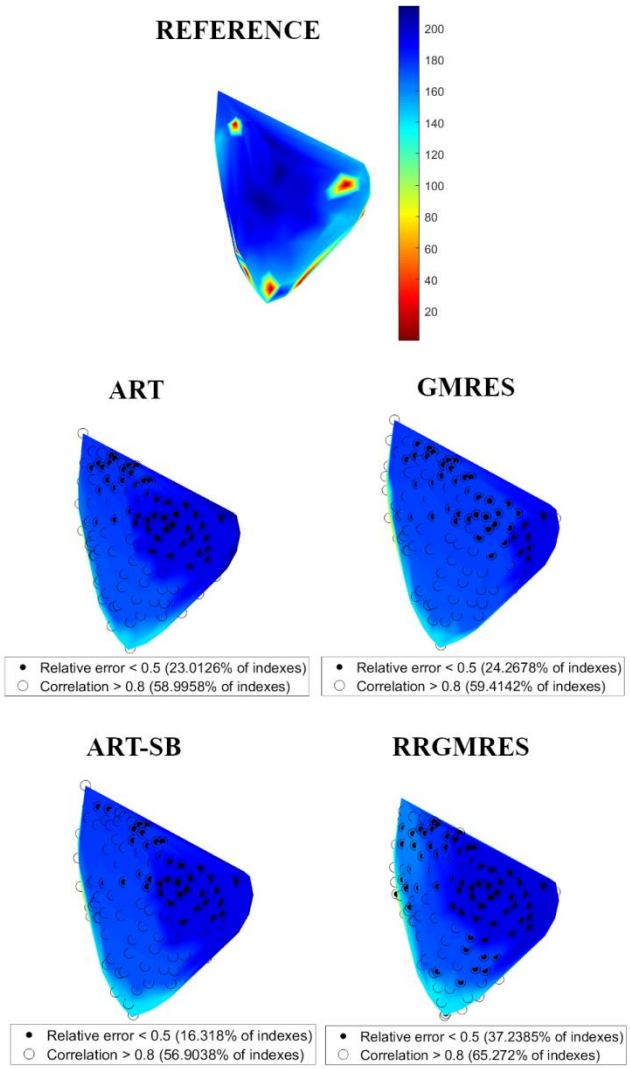


Figura 2. *Mapas isocrónicos de la cara izquierda del epicardio, obtenidos para la base de datos de Auckland con el ritmo estimulado "paced", y localizando las mejores reconstrucciones*

En las Figuras 1 y 2 se muestran los mapas isocrónicos obtenidos para la base de datos de Auckland con el ritmo estimulado "paced". Además, también hemos localizado las mejores reconstrucciones (CC más altos y RE más bajos), dentro de los mapas isocrónicos adquiridos a partir de la reconstrucción de los potenciales, mediante las diferentes técnicas de problema inverso.

4. Discusión y conclusiones

El objetivo de este trabajo era adaptar métodos iterativos de problema inverso por primera vez al campo del ECGI. Una vez aplicados, hemos podido observar en las Tablas 1 y 2 que todas las técnicas usadas en la solución del problema inverso mejoran las reconstrucciones obtenidas por la técnica GMRES, que era el único método ya probado en este ámbito [7]. Si ordenamos las técnicas según sus resultados, vemos que ART, randART y RRGMRES lograron las mejores reconstrucciones de los potenciales. Cabe destacar que la ART ha resultado ser la técnica más estable entre bases de datos, ya que, no obteniendo siempre los mejores resultados, sus reconstrucciones siempre han sido relativamente precisas (Tabla 1 versus Tabla 2).

Los mapas isocrónicos de la parte derecha del epicardio (Figura 1) muestran que tan en la zona de activación temprana (color rojo), como en la zona del ápex las reconstrucciones son menos precisas. Además, se puede apreciar que la transición entre la activación temprana y la tardía no se representa adecuadamente en las reconstrucciones, a diferencia de la referencia. En la Figura 2, que representa la cara opuesta a la Figura 1, se puede observar una mejor reconstrucción de las zonas problemáticas mencionadas arriba. Pero, aun así, estas zonas siguen mostrando una peor reconstrucción que las otras, mayormente en términos de amplitud (RE). Además, ambas figuras muestran que RRGMRES presenta mejores reconstrucciones en los ventrículos, especialmente en términos de morfología (CC). Siendo los ventrículos la cavidad cardíaca más difícil de reconstruir en problemas de ECGI hasta el momento [5].

En conclusión, los resultados demuestran que las técnicas ART, randART y RRGMRES mejoran la reconstrucción obtenida por GMRES.

Finalmente, cabe destacar la necesidad de trabajo futuro para analizar estas técnicas aplicadas en más bases de datos, y probar así su eficacia en señales con una relación ruido-señal (SNR) más baja.

Referencias

[1] John, R. M., et al. Ventricular arrhythmias and sudden cardiac death. *The* Lancet, vol 380, 2012, pp 1520-1529.

[2] Shah, A. Frontiers in Noninvasive Cardiac Mapping, An Issue of Cardiac Electrophysiology Clinics. *Elsevier Health Sciences*, vol 7, 2015 (ISBN: 9780323356510).

[3] Wang, Y., et al. Noninvasive Electroanatomic Mapping of Human Ventricular Arrhythmias with Electrocardiographic Imaging. *Science translational medicine*, vol 3, 2011, pp 98ra84.

[4] Dubois, R., Labarthe, S., Coudière, Y., Hocini, M. & Haïssaguerre, M. Global and Directional Activation Maps for Cardiac Mapping in Electrophysiology. *Computing in Cardiology (CinC)*, vol 39, 2012, pp 349-352.

[5] Duchateau, J., et al. Performance and limitations of noninvasive cardiac activation mapping. *Heart Rhythm*, vol 16, 2018, pp 435-442.

[6] Wang, Y. & Rudy, Y. Application of the Method of Fundamental Solutions to Potential-based Inverse Electrocardiography. *Annals of Biomedical Engineering*, vol 34, 2006, pp 1272-1288.

[7] Ramanathan, C., et al. Noninvasive Electrocardiographic Imaging (ECGI): Application of the Generalized Minimal Residual (GMRes) Method. *Annals of Biomedical Engineering*, vol 31, 2003, pp 981-994.

[8] Aras, K., et al. Experimental Data and Geometric Analysis Repository (EDGAR). *Journal of electrocardiology*, vol 48, 2015, pp 975-981.

[9] Hansen, P. C., & Jørgensen, J. S. AIR Tools II: algebraic iterative reconstruction methods, improved implementation. *Numerical Algorithms*, vol 79, 2018, pp 107-137.

[10] Ziegler, A., Grass, M. & Koehler, T. Method and device for the iterative reconstruction of cardiac images. U.S. Patent No. 7,596,204. Eindhoven, NL, 2009.

[11] Chamorro-Servent, J., et al. Use of Split Bregman denoising for iterative reconstruction in fluorescence diffuse optical tomography. *Journal of Biomedical Optics*, vol 18, 2013.

[12] Hansen P.C. Discrete inverse problems: insight and algorithms. *SIAM*, vol 7, 2010.

[13] Cluitmans, M. J., de Jong, M. M., Volders, P. G., Peeters, R. L. & Westra, R. L. Physiology-based Regularization Improves Noninvasive Reconstruction and Localization of Cardiac Electrical Activity. *Computing in Cardiology*, vol 41, 2014, pp 1-4 (ISBN: 978-94-6159-594-2).

[14] Bear, L. R., et al. Forward Problem of Electrocardiography. *Circulation: Arrhythmia and Electrophysiology*, vol 8, 2015, pp 677-684.

Efecto de los parámetros geométricos de adquisición en tomosíntesis digital lineal

A. del Valle Hidalgo[1], A. Galán González,[1,2], F. J. García Blas[3], M. Desco Menéndez[1,2,4,5], M. Abella García[1,2,4]

[1] Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España, {adelvalle, mabella}@hggm.es

[2] Dept. Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid, España,

[3] Dept. Informática, Universidad Carlos III de Madrid, 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

La tomosíntesis digital ofrece una alternativa a la tomografía computarizada (TAC) y a la radiografía convencional para el diagnóstico clínico, sobre todo frente al reciente brote de COVID-19. Su principal ventaja es que permite obtener información tomográfica manteniendo una dosis de radiación inferior al TAC. Su resolución espacial es alta en los cortes coronales, pero se ve limitada a lo largo del eje antero-posterior (grosor de corte), con gran dependencia de los parámetros geométricos de adquisición. Este trabajo investiga la influencia del ángulo de barrido, del paso angular y del plano focal en la resolución espacial del volumen reconstruido. Para ello, se simulan proyecciones de tomosíntesis de un maniquí de tórax con tres nódulos y se reconstruyen mediante el algoritmo SART. Los resultados muestran que la resolución espacial en el corte coronal es independiente de estos parámetros mientras que la resolución en profundidad aumenta proporcionalmente con el ángulo de barrido y cuando el plano focal es posterior al plano de interés.

1. Introducción

El reciente brote de COVID-19 ha puesto de manifiesto la importancia de la radiografía de tórax para el diagnóstico rápido y temprano de la enfermedad, así como de su seguimiento. La infección por el virus SARS-CoV-2 provoca infecciones del tracto respiratorio inferior con anomalías tales como opacidades en vidrio esmerilado ('*ground-glass opacity*'), patrones de empedrado ('*crazy-paving pattern*') y consolidaciones que se detectan mediante radiografía convencional o tomografía axial computerizada (TAC), siendo esta última la más usada en investigación por su mayor resolución espacial [1].

La radiografía plana es poco sensible e infravalora el alcance de las lesiones pulmonares, respecto a un TAC, pero es inviable realizar un TAC a todos los pacientes con sospecha COVID-19 tanto por la alta dosis de radiación que supone como por el hecho de que el paciente queda colocado dentro del gantry, lo que imposibilita su uso en casos con dificultades de movilidad del paciente.

La tomosíntesis digital (TD) ofrece una alternativa de diagnóstico al TAC con un menor coste de equipo, y, lo que es más importante, una menor dosis de radiación [2-4]. La TD es comúnmente utilizada en estudios de mama [4, 5] y, más recientemente, de tórax, estando en fases de investigación para la detección de cáncer de pulmón [3].

Los equipos de TD están formados por una fuente de rayos X que se mueve siguiendo una trayectoria lineal, para estudios de tórax, o circular, en estudios de mama [3], mientras el detector permanece fijo. La TD permite reconstruir un número arbitrario de cortes paralelos al plano del detector a partir de un conjunto de proyecciones que cubren un ángulo limitado. Es decir, la reconstrucción con TD genera un volumen 3D del paciente en el que se reduce la influencia de estructuras anatómicas superpuestas [2-4]. La principal limitación de la TD es su menor resolución espacial que la TAC en el eje antero-posterior (grosor de corte) [3].

La calidad de las imágenes reconstruidas depende en gran medida del algoritmo de reconstrucción [6] y de los parámetros de adquisición relativos a la geometría del sistema: dirección de barrido, ángulo de barrido, número de proyecciones y plano focal [7]. Ambos determinan la dosis mínima necesaria para mantener una resolución espacial clínicamente válida. Para optimizar el protocolo de adquisición, es necesario comprender la influencia de los parámetros de adquisición y la resolución espacial.

Estudios previos [8-11] han investigado la influencia de los parámetros de adquisición pero ninguno cubre todos los parámetros simultáneamente ni el mismo algoritmo de reconstrucción que el resto. Por otro lado, todos definen el plano focal sobre el plano del detector, al igual que los equipos comerciales [4], lo que implica que el detector es estacionario. Sin embargo, no se han explorado lo efectos de que éste sea móvil, como proponemos en este estudio. El objetivo de este estudio es determinar la influencia de los principales parámetros de adquisición (ángulo de barrido, paso angular y plano focal) en la resolución espacial de las imágenes reconstruidas en un mismo estudio y bajo unas mismas condiciones (algoritmo de reconstrucción, sistema y maniquí comunes). Para ello, se usa un algoritmo de reconstrucción (SART), que se ve menos restringido ante la geometría del sistema que los empleados anteriormente [5].

ISBN: 978-84-09-25491-0, pags. 273- 276

2. Métodos

El estudio se basa en simulaciones con FUX-SIM, un simulador de rayos X que permite explorar distintas configuraciones de fuente y detector [12].

2.1. Geometría del sistema

La geometría de simulación se basa en el equipo de TD lineal que se muestra en la Figura 1, en el que la fuente de rayos X y el detector se pueden mover siguiendo trayectorias lineales de forma sincronizada. Las distancias fuente-detector en la posición central de la fuente (1800 mm) y objeto-detector (109 mm) se mantienen constantes en todo el estudio. El detector tiene 2150×2150 pixeles con un tamaño de pixel de 0.2×0.2mm. El eje antero-posterior (AP) une el centro del detector con la fuente en la posición central, siendo el eje sobre el que se mide el grosor de corte. El ángulo de barrido, definido como el ángulo respecto del punto CoR entre la primera y la última proyección, y el paso angular entre proyecciones se mantienen a 30 y 0.5 grados, respectivamente. El plano focal define el plano sobre el que se orienta la fuente de forma que las estructuras contenidas en él son proyectadas en la misma posición del detector, y se sitúa sobre el plano del detector (0 mm) para el caso mostrado en la Figura 1.

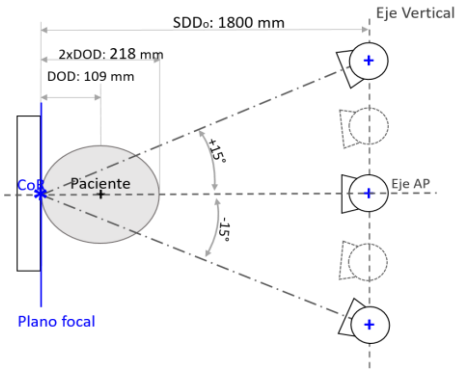


Figura 1. *Geometría del sistema de tomosíntesis lineal manteniendo el detector fijo. SDDo: distancia fuente-detector en la posición central de la fuente. DOD: distancia objeto-detector. AP: eje antero-posterior. CoR: punto de intersección entre el eje AP y el centro del panel en la posición central.*

En total se realizaron tres experimentos, variando sólo un parámetro (ángulo de barrido, paso angular y plano focal) cada vez y manteniendo el resto constante, en los valores mencionados. El ángulo de barrido, varió entre 20 grados (i.e. ±10 grados) y 50 grados (i.e. ±25 grados), con un paso de 1 grado (i.e. ±0.5 grados) y una distribución simétrica en torno al eje AP. Estos valores son los típicamente usados en TD de tórax [3]. El paso angular varió entre 0.5 y 2 grados. El plano focal varió entre el plano del detector, como está implementado en los equipos comerciales, y el plano anterior del maniquí con un paso de DOD/2.

2.2. Maniquíes

Las simulaciones se basan en dos maniquíes. Los dos primeros experimentos emplean un TAC del maniquí de tórax *PBU-60* (KYOTO KAGAKU, Kyoto, JP), colocado en posición antero-posterior de forma que la dirección de barrido es paralela a la columna vertebral (Figura 2A). El tamaño de matriz es de 512×259×430 voxeles, con 0.843×0.843×0.8 mm/voxel. Para evaluar el efecto de los parámetros de adquisición a distinta profundidad, se incluyeron digitalmente en el TAC tres esferas de 6 mm de diámetro y 1.4 g/cm^3 de densidad, valores típicos de un nódulo pulmonar [13], en tres planos coronales (anterior, medial y posterior) del lóbulo superior del pulmón.

El experimento relativo al plano focal emplea adicionalmente un maniquí no antropomórfico para explorar planos coronales fuera del pulmón y evitar posibles interferencias entre nódulos. Este maniquí se compone de una esfera de 6 mm y 1.4 g/cm^3 colocada sobre el eje AP. La posición de la esfera varía a lo largo del eje AP, situándose en distintos planos coronales (Figura 2B-C). El tamaño de matriz es 259x259x200 con 0.843×0.843×0.8 mm/voxel.

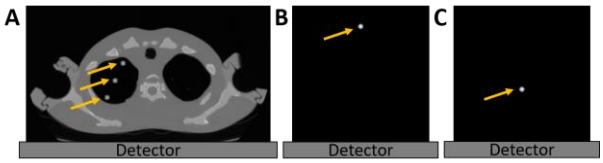


Figura 2. *Corte axial del maniquí de tórax (A) y del maniquí no antropomórfico (B-C) con nódulos (flechas) en distintos planos.*

2.3. Reconstrucción y análisis

Los datos simulados se reconstruyen mediante el algoritmo iterativo '*Simultaneous algebraic reconstruction technique*' (SART) [5], resultando en un volumen de iguales dimensiones que la referencia:, 259 planos coronales con un grosor de corte de 0.843 mm (Figura 3).

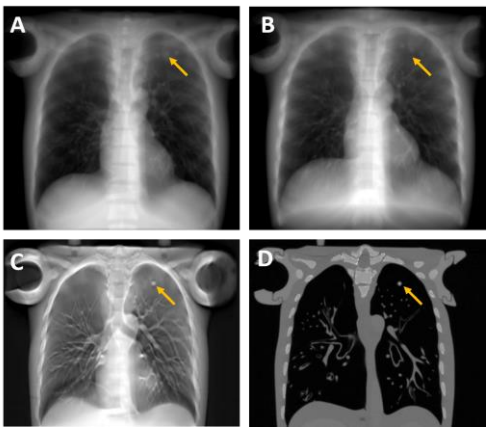


Figura 3. Ejemplos de proyección (A y B), reconstrucción (C) y maniquí de referencia (D).Las flechas señalan los nódulos.

Para evaluar el efecto en la resolución espacial, utilizamos el cambio en el diámetro del nódulo, dado por el ancho a media altura (FWHM) del perfil obtenido a lo largo de los tres ejes anatómicos del mismo (Figura 4), al igual que se hace en Kuramoto et al. [9]. El error se obtiene como la diferencia de diámetro entre los nódulos reconstruidos y los de referencia (TAC).

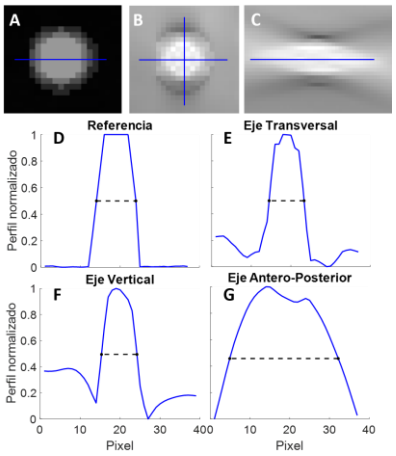


Figura 4. *Perfiles medidos sobre el nódulo de referencia (A) y el nódulo reconstruido en vista coronal (B) y sagital (C). Perfiles correspondientes al nódulo de referencia en la imagen de TAC (D) y al reconstruido (E-G). Las vistas en cada eje del nódulo de referencia son iguales al tratarse de una esfera. Medida del diámetro (FWHM) en línea discontinua.*

3. Resultados

3.1. Influencia del ángulo de barrido

La Figura 5A muestra que el error entre los nódulos reconstruidos y la referencia es constante para distintos ángulos de barrido en los ejes transversal y vertical. En el eje AP, el aumento del ángulo de barrido reduce el error en un 50%, lo que indica un aumento significativo de la resolución espacial. Este comportamiento es similar en los tres planos coronales evaluados (Figura 5D).

3.2. Influencia del paso angular

La Figura 5B muestra que variaciones del paso angular no influyen en el diámetro de los nódulos a lo largo de los ejes transversal y vertical. Se aprecia el mismo efecto en el eje AP en los planos anterior y medial. Sin embargo, en el plano posterior se observa una ligera disminución del error (15%) para pasos angulares inferiores a 1 grado.

3.3. Influencia del plano focal

La Figura 6 muestra que el error en el diámetro del nódulo reconstruido en distintos planos coronales de los maniquíes no antropomórficos se mantiene constante cuando el plano focal se sitúa en el plano anterior del maniquí (2×DOD mm). Cuando el plano focal se sitúa a una distancia intermedia (DOD mm), el error aumenta cuando el nódulo se sitúa en planos coronales anteriores al plano focal y disminuye en un 14% para planos posteriores al mismo.

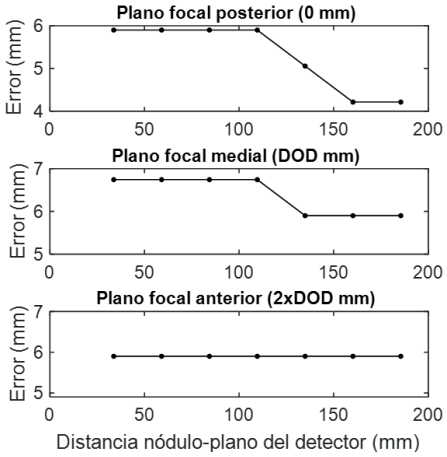


Figura 6. *Error del diámetro en el eje AP según el plano focal para nódulos en distintos planos coronales.*

En el maniquí antropomórfico (Figura 5C y F) se observa un efecto similar. Para el plano medial y anterior, el error aumenta (hasta en un 28%) según el plano focal se sitúa por delante del nódulo. El error en los ejes transversal y vertical se mantiene cercano a 0 mm en todos los casos.

4. Discusión y conclusiones

Todos los experimentos muestran un error en el eje AP mayor que en los otros dos ejes, confirmando la literatura previa [2-4, 11]. Ésto se debe a que la geometría del sistema restringe el ángulo de barrido, de forma que existe un muestreo incompleto que limita la reconstrucción [2]. La resolución espacial en los ejes vertical y transversal, es decir en el plano coronal, es independiente de los parámetros de adquisición estudiados. Además, esta resolución es alta considerando que el diámetro de los nódulos reconstruidos es cercano al de referencia (errores inferiores al 1%). En cambio, la resolución espacial a lo largo del eje AP mejora con mayores ángulos de barrido, confirmando las

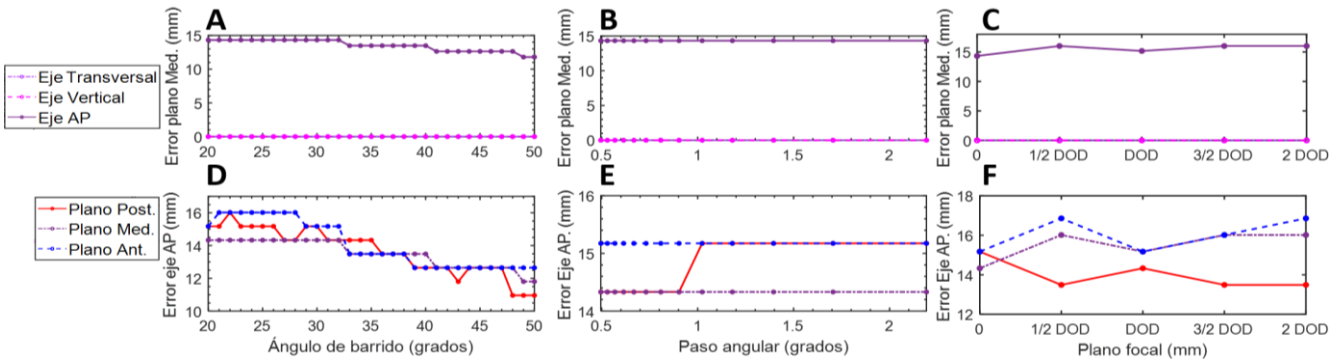


Figura 5. *Error del diámetro para el plano medial (arriba) y el eje AP (abajo) del maniquí de tórax según el ángulo de barrido (A y D), el paso angular (B y E) y la posición del plano focal (C y F).*

observaciones de Deller et al. [8] y Li et al. [14], que indican que la resolución espacial en el eje AP es inversamente proporcional a la tangente del semiángulo de barrido. Además, nuestros resultados muestran que el efecto del ángulo de barrido en la resolución espacial es constante a lo largo del eje AP.

Respecto al plano focal, los resultados muestran que la resolución en el eje AP aumenta cuando el plano focal se encuentra en un plano posterior al plano coronal de interés. Este fenómeno se debe a que la posición del plano focal, geométricamente equivalente a variar el ángulo de barrido junto con la distancia objeto-fuente, determina el ángulo de incidencia efectivo del haz de rayos X. Así, el ángulo de incidencia será mayor, además de la región muestreada y la resolución espacial, para planos focales más próximos al detector que el plano de interés, independientemente de la distancia entre estos dos planos. Cabe destacar que en el nódulo posterior del tórax se observa un efecto opuesto. Una posible causa es que el nódulo se encuentra más alejado del eje AP que los otros nódulos, de forma que quede en el límite de la región muestreada para planos focales alejados del detector. Trabajos futuros han de aclarar este punto.

Los sistemas comerciales actuales [4, 11] definen el plano focal en el del detector, lo que implica que éste sea estacionario. El uso de un detector móvil requeriría sistemas robóticos más avanzados, impondría mayores requisitos a la calibración geométrica e implicaría errores de reconstrucción debidos a la mecánica. La evaluación variando el plano focal (detector móvil) muestra que el uso de un detector móvil para variar el plano focal no mejoraría significativamente la resolución espacial.

En cuanto al paso angular, los resultados parecen indicar que la resolución espacial en el eje AP no se ve afectada. En el plano posterior la resolución se mantiene constante, pero se da un ligero salto para pasos angulares menores a 1 grado. Este salto se podría explicar con el estudio de Sechopulos y Ghetti que sugiere que un aumento del número de proyecciones, equivalente a reducir el paso angular, por encima de cierto límite no conlleva mejoras en la resolución [15]. En todo caso, esta independencia del paso angular es beneficiosa para la aplicación clínica puesto que permite reducir la dosis de radiación absorbida por el paciente (aumentando el paso angular) y reducir el ruido. Trabajos futuros han de validar estos resultados en experimentos reales y definir el conjunto de parámetros que optimizan la resolución espacial, principalmente en la dirección de profundidad, para una dosis mínima.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia, Innovación y Universidades (Instituto de Salud Carlos III, proyecto DTS17/00122; Agencia Estatal de Investigación, proyecto DPI2016-79075-R-AEI/FEDER, UE), cofinanciado por Fondos de la Unión Europea (FEDER), "A way of making Europe". Además, ha sido financiado por el Programa de apoyo a la realización de proyectos interdisciplinares de I+D para jóvenes investigadores de la Universidad Carlos III de Madrid 2019-2020 en el marco del Convenio Plurianual Comunidad de Madrid- Universidad Carlos III de Madrid (proyecto DEEPCT-CM-UC3M) y por CRUE Universidades, CSIC y el Banco Santader (Fondo Supera, proyecto RADCOV19). El CNIC está financiado por el Ministerio de Ciencia, Innovación y Universidades y la fundación PRO-CNIC y es un centro de excelencia Severo Ochoa (SEV-2015-0505).

Referencias

[1] Ye, Z., et al., *Chest CT manifestations of new coronavirus disease 2019 (COVID-19): a pictorial review.* Eur Radiol, 2020. 30(8): pp 4381-4389.

[2] Dobbins, J.T., 3rd, and D.J. Godfrey, *Digital x-ray tomosynthesis: current state of the art and clinical potential.* Physics in medicine and biology, 2003. 48(19): pp 65-106.

[3] Ferrari, A., et al., *Digital chest tomosynthesis: the 2017 updated review of an emerging application.* Ann Transl Med, 2018. 6(5): pp 91.

[4] Machida, H., et al., *Whole-body clinical applications of digital tomosynthesis.* Radiographics, 2016. 36(3): pp 735-750.

[5] Zhang, Y., et al., *Tomosynthesis reconstruction using the simultaneous algebraic reconstruction technique (SART) on breast phantom data.* Medical Imaging 2006. 6142.

[6] Gomi, T., et al., *Comparison of Reconstruction Algorithms for Decreasing the Exposure Dose During Digital Tomosynthesis for Arthroplasty: a Phantom Study.* J Digit Imaging, 2016. 29(4): pp 488-95.

[7] Machida, H., et al., *Optimizing parameters for flat-panel detector digital tomosynthesis.* Radiographics, 2010. 30(2): pp 549-562.

[8] Deller, T., et al. *Effect of acquisition parameters on image quality in digital tomosynthesis*. *Medical Imaging 2007: Physics of Medical Imaging*. 2007. pp 65101L.

[9] Kuramoto, T., et al., *Variations in slice sensitivity profile for various height settings in tomosynthesis imaging: Phantom study.* Phys Med, 2018. 53: pp 108-117.

[10] Li, B., et al., *Optimization of slice sensitivity profile for radiographic tomosynthesis.* Med Phys, 2007. 34(7): pp 2907-16.

[11] Zhao, C., Weber, T., Herbst, M., Ritschl, L., Vogt, S., Kappler, S., Siewerdsen, J. H., & Zbijewski, W., *Advanced Scan Trajectories for Chest Tomosynthesis on Robotic X-Ray Systems*, in *6th International Conference on Image Formation in X-Ray Computed Tomography*. 2020: Regensburg, Germany. pp. 210-213.

[12] Abella, M., et al., *FUX-Sim: Implementation of a fast universal simulation/reconstruction framework for X-ray systems.* PloS one, 2017. 12(7): pp e0180363.

[13] Silverman, J.D., N.S. Paul, and J.H. Siewerdsen, *Investigation of lung nodule detectability in low-dose 320-slice computed tomography.* Med Phys, 2009. 36(5): pp 1700-10.

[14] Li, B., et al., *The impact of acquisition angular range on the z-resolution of radiographic tomosynthesis.* International Congress Series, 2004. 1268: pp 13-18.

[15] Sechopoulos, I. and C. Ghetti, *Optimization of the acquisition geometry in digital tomosynthesis of the breast.* Med Phys, 2009. 36(4): pp 1199-207.

Detection of PD-L1 expression in non-small cell lung cancer with ^{18}F-FDG PET imaging

R. Pérez-Domínguez[1], A.P. Seiffert[1], E. Martínez-Albero[2], S. Ruiz-Solís[2], A. Gómez-Grande[2], E.J. Gómez[1,3], P. Sánchez-González[1,3]

[1] Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España;
raquel.perez.dominguez@alumnos.upm.es, {ap.seiffert, enriquejavier.gomez, p.sanchez}@upm.es

[2] Servicio de Medicina Nuclear, Hospital Universitario 12 de Octubre, Madrid, España;
{emalbero, sebastian.ruiz, adolfo.gomez}@salud.madrid.org

[3] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Madrid, España

Abstract

Lung cancer is the leading cause of cancer-related death, being non-small cell lung cancer (NSCLC) the most frequent type. Programmed cell death ligand-1 (PD-L1) expression in cancer cells trick the immune system to continue its proliferation, thus leading to a faster and more aggressive disease growth. Currently, a tumour sample must be extracted through biopsy and posterior IHC testing is employed for PD-L1 detection. ^{18}F-fluorodeoxyglucose (FDG) PET could provide a non-invasive method to detect PD-L1 expression in the tumour. The use of quantitative image features extracted from ^{18}F-FDG PET is studied in a group of 60 patients with NSCLC. A total of 52 image features characterising the metabolic activity of the primary tumour in the ^{18}F-FDG PET images are extracted. Statistical differences between patients with and without PD-L1 expression are studied and classification models are generated based on these image features. 37 of 52 features show statistically significant differences between PD-L1 expression and non-expression. A hyperparameter-optimised K-nearest neighbour classification model result in accuracy, sensitivity and specificity values of 73.69%, 33.3% and 92.31%, respectively. The results indicate that ^{18}F-FDG PET could detect PD-L1 expression in non-small cell lung cancer.

1. Introduction

Lung cancer is the leading cause of cancer deaths worldwide with 2 million new cases diagnosed and nearly 1.8 million deaths in 2018, which represents 18.4% of total cancer diseases [1]. Non-Small Cell Lung Cancer (NSCLC) is the most common type and comprises over 80% of all known lung cancer diagnoses [2].

Inmunotherapy is a novel method of lung cancer treatment option for non-small cell lung cancer. Many clinical trials have proven high effectiveness and improvement in the median overall survival of lung cancer patients with metastatic stages of up to 13.8 months [3, 4].

Inmunotherapy uses monoclonal antibodies directed against the immune checkpoint molecules such as programmed cell death 1 (PD-1) or its ligand (PD-L1). PD-L1 is expressed on the surface of various cancer and immune cells and PD-1 on T-lymphocytes. PD-L1 interacts with PD-1 and downregulates T cell effector functions, thus leading to a faster and more aggressive growth of tumoral cells [3].

Up until now, immunohistochemistry (IHC) is the only validated testing method to detect PD-L1 in tumour samples. Due to several factors, it is desirable to evaluate the protein in the entire primary tumour. However, not all tumours are resectable, nor is it worthwhile for all affected patients to undergo surgery. Its evaluation may be also performed through biopsy extracted samples [5]. Therefore, non-invasive techniques are of great interest for the detection of PD-L1 expression.

^{18}F-fluorodeoxyglucose (FDG) PET images are used in staging and treatment assessment of lung cancer. Metabolic parameters from ^{18}F-FDG PET in lung cancer and its relationship to PD-L1 expression have been studied previously [6–9]. Mainly, the semi-quantitative standardized uptake value (SUV) is measured, and higher metabolic uptake is related to PD-L1 expression. However, the studies only evaluated a limited number of image features.

The aim of this study is to evaluate the use of image features from ^{18}F-FDG PET for the non-invasive detection of PD-L1 expression. Apart from metabolic parameters based on the SUV and metabolically active volume, texture features are extracted. Statistical differences of the feature values between PD-L1 expression and non-expression in tumour cells are studied and classification models are generated to further evaluate their potential.

2. Material and Methods

2.1. Materials

A total of 60 patients diagnosed with non-small cell lung cancer were included in the study group. The inclusion criteria were as follows:

- ^{18}F-FDG PET scan at diagnosis.
- Confirmed diagnosis of NSCLC.
- PD-L1 expression evaluated by IHC testing.

The exclusion criteria were:

- Maximum SUV (SUVmax) of the tumour below 1 due to software limitations.
- Metabolically active tumour volume below 64 voxels.

ISBN: 978-84-09-25491-0, pags. 277- 280

The demographic and tumour characteristics are shown in Table 1. The study group is divided in patients with PD-L1 expression and non-expression defined by IHC testing of a tumour sample extracted by biopsy. Moreover, each case is defined by 12 clinicopathological features like tumour size or stage.

Characteristic	PD-L1 < 1%	PD-L1 ≥ 1%
Gender		
Male	27	18
Female	12	3
Age		
< 60	4	4
≥ 60	35	17
Smoking history		
Smoker	31	19
Non-smoker	8	2
TNM stage		
I	20	8
II	2	1
III	3	8
IV	14	4
Histology		
Squamous cell carcinoma	5	8
Adenocarcinoma	33	12
Large cell carcinoma	1	1
Total	39	21

Table 1. *Demographic and tumour characteristics.*

For image acquisition, blood glucose level was measured before the dose injection. Levels <180 mg/dl were required to avoid unintelligible images and false negative interpretation. ^{18}F-FDG doses were injected intravenously and were adjusted by patient's weight (4-5 MBq/kg). Whole-body ^{18}F-FDG PET/CT images were obtained with an integrated PET/CT scanner (Siemens Biograph 6 scanner, Siemens Healthineers AG, Erlangen Germany). PET images were reconstructed using an iterative algorithm with image matrix of 168 x 168. Reconstructed images were attenuation corrected using the CT scan.

Tumour segmentation and feature extraction is performed in LIFEx software[1] (Local Image Feature Extraction) [10]. The necessary code for statistical analysis and the creation of classification models is implemented in Python. This process is carried out through Anaconda Navigator (v. 1.9.12, by Anaconda Inc.) and Jupyter (v. 6.0.3). Several Python libraries are used for the database management and the development of the different analyses. In particular, databases were handled with *Pandas* library[2], the statistical analysis was carried out with different functions of *SciPy* library[3], and the classification model was implemented using *scikit-learn* library[4], which is developed especially for predictive data analysis.

2.2. Methods

The methodology of this work is divided into three steps: (1) tumour segmentation, (2) feature extraction, and (3) statistical analysis.

Tumour segmentation. The tumour segmentation is carried out to limit the region of interest, in this case the primary tumour, from which the desired features are extracted. First, a 3D volume of interest (VOI) including the complete tumour is manually drawn. To ensure a correct segmentation, some surrounding tissue is included in this VOI. Then, the metabolically active tumour tissue is segmented by thresholding using as threshold 41% of the SUVmax of the selected VOI [11].

Feature extraction. Previously to the feature extraction, a preprocessing, consisting in resampling voxel intensities using a fixed number of discrete values, is performed. Absolute resampling is used following equation (1), where the upper bound is the highest SUVmax of the study group and D is the number of discrete levels [12].

$$R(x) = round\left(D \times \frac{I(x)}{UpperBound}\right) \quad (1)$$

Then, a total of 52 image features are extracted from the segmented tumours. These are grouped into first, second, and higher order features based on the number of pixels whose interrelationship is evaluated [10, 12]:

- First order features include: (1) SUV-based features, (2) metabolic tumour volume (MTV), (3) histogram-based, and (4) shape-based features.
- Second order features: Calculated from the gray-level co-occurrence matrix (GLCM).
- Third and higher order features: Calculated form the gray-level run length matrix (GLRLM), neighbourhood gray-level difference matrix (NGLDM) and gray-level size zone matrix (GLSZM).

Statistical analysis. The normality of the data is evaluated by a Shapiro-Wilk test. Differences between the two groups (PD-L1 expression and non-expression) are evaluated by t-tests or Mann-Whitney U-tests. A p-value < 0.05 is considered statistically significant.

Statistically significant features are used to create classification models based on support vector machines (SVM) and K-nearest neighbour (KNN). Test/training split (30%/70%) and cross-validation with hyperparameter selection are performed to find the best model. Hyperparameters are optimised for accuracy and then for sensitivity using a grid search. The performance is evaluated by calculating the accuracy, sensitivity and specificity of the classification model after applying the test set. The different steps of the classification model training and testing are shown in Figure 1.

[1] https://www.lifexsoft.org/ Last access: Aug 2020

[2] https://pandas.pydata.org/ Last access: Aug 2020

[3] https://www.scipy.org/ Last access: Aug 2020

[4] https://scikit-learn.org/stable/ Last access: Aug 2020

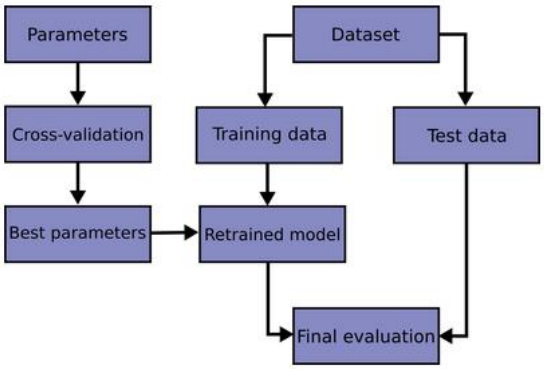


Figure 1. *Classification model methodology* [13].

3. Results and Discussion

The tumour segmentation in LIFEx is exemplified in the following images. Figure 2 shows the manually drawn VOI to select the primary tumour and surrounding tissue. Figure 3 shows the segmented tumour after thresholding from which the image features are extracted.

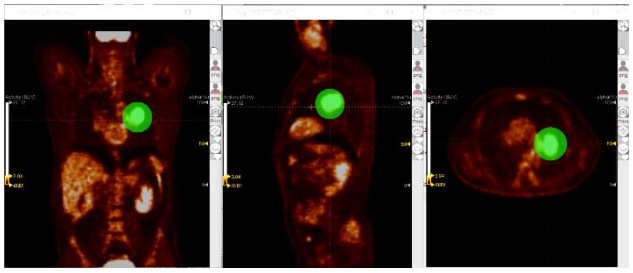


Figure 2. *VOI selection including the tumour.*

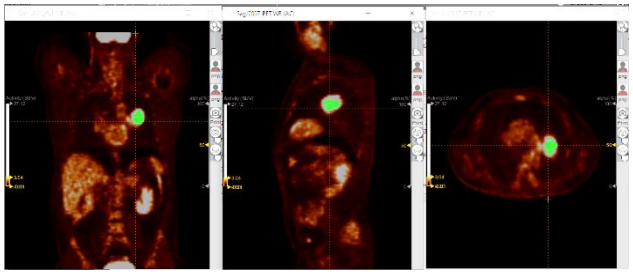


Figure 3. *Segmented tumour after thresholding.*

After the feature extraction, the statistical differences between PD-L1 expression and non-expression are studied. Out of 64 evaluated features (12 clinicopathological and 52 image features), 37 present statistically significant differences between both groups. The distribution is as follows:

- 2 of 12 clinicopathological features.
- 11 of 20 first order features.
- 6 of 7 GLCM features.
- 18 of 25 third or higher order features.

As can be observed, most textural features might be able to detect PD-L1 expression or differentiate between expression and non-expression. First order features describe a greater variety of tumour characteristics and none of the 4 shape-based features are statistically significant.

The SVM and KNN models using 70% of the data in the training with default hyperparameters resulted in accuracies of 63.15% in both cases. The results for both algorithms are shown in Table 2.

	Accuracy	Sensitivity	Specificity
SVM	63.15%	0.00%	92.31%
KNN	63.15%	50.00%	69.23%

Table 2. *Testing performance of model trained with default parameters.*

After applying the hyperparameter optimisation using the training set, the performance of the model improves. A summary of the performance evaluation is shown in Table 3 and Table 4 for accuracy-optimised and sensitivity-optimised hyperparameter selection, respectively. The KNN model with optimised accuracy during hyperparameter selection is considered the best model as it presents the highest accuracy and specificity.

	Accuracy	Sensitivity	Specificity
SVM	63,15%	33.33%	76.92%
KNN	73.68%	33.33%	92.31%

Table 3. *Testing performance of model trained with accuracy-optimised hyperparameters.*

	Accuracy	Sensitivity	Specificity
SVM	68.42%	33.33%	84.61%
KNN	68.42%	50.00%	76.92%

Table 4. *Testing performance of model trained with sensitivity-optimised hyperparameters.*

Following the testing of the created classification model, the misclassified cases of the best model are analysed. These cases and some of their characteristics are shown in Table 5.

Patient	SUV-max	T stage (mm)	TNM stage	PD-L1	Classification result
1	2.8	T1 (13)	I	1	0
2	4.6	T1 (15)	I	1	0
3	11.02	T2 (47)	II	1	0
4	14	T2 (45)	III	1	0
5	27.9	T2 (46)	IV	0	1

Table 5. *Characteristics of incorrectly classified cases.*

First, the SUVmax levels of every misclassified case are analysed. Thus, its mean value is calculated for both groups (PD-L1 expression and non-expression), resulting in mean SUVmax values of 16.41 and 9.66, respectively. As can be seen, patients 1, 2 and 3 are closer to the mean SUVmax of the opposite group and patient 5 presents a high SUVmax which is closer to those expressing PD-L1.

On the other hand, the patients incorrectly classified as PD-L1 non-expression (first four patients in Table 5), have a T1 or T2 stage. In turn, all correctly classified cases as PD-L1 positive are T3 or T4 stage. The overall TNM stage of these cases is also on the lower end of the scale.

In the case of patient 5, the evaluation was performed over a sample extracted through biopsy. Due to the advanced

metastatic stage of the disease, the patient could not undergo surgery. Hence, there is a possibility that the patient may have been misdiagnosed by the specialist as non-expression of the PD-L1 protein, due to the characteristic heterogeneity of tumour masses. Thus, the protein may be expressed in the tumour, but not exactly in the extracted sample. However, the patient died 16 months after its diagnosis, hence there is no additional PD-L1 protein evaluation.

Regarding previous studies, SUVmax has been shown to be a significantly higher in NSCLC patients with PD-L1 expression [6–9]. In the present study, SUVmax present statistically significant differences between PD-L1 expression and non-expression (p = 0.0056). Moreover, other SUV-based features as well as texture features show significant differences between both groups.

The main limitation of the present study is the low number of patients whose PD-L1 protein expression in tumoral cells is known, and who hence meet the inclusion criteria. The fact that the tumour tissue is characteristically heterogeneous causes that the only way to obtain a totally reliable result is by performing an IHC test on the entire tumour.

Finally, a future research line may be carried out with a larger database, in which the patients diagnosed as PD-L1 negative through biopsy are excluded, and the number of patients whose tumour cells express the PD-L1 protein is increased. Moreover, other classification models may be generated to increase its performance even though the results in the present study already indicate that ^{18}F-FDG PET might be of use in the detection of PD-L1 expression.

4. Conclusion

^{18}F-FDG PET imaging is evaluated as a non-invasive method for the detection of PD-L1 expression in NSCLC. In this work, image features present statistically significant differences between PD-L1 expression and non-expression. Moreover, a classification model is generated that present high accuracy and specificity. It can therefore be concluded that ^{18}F-FDG PET image features can be useful to facilitate the detection of PD-L1 protein expression. By improving the overall measurements of the classification model, it could be used as a complementary tool for PD-L1 evaluation in patients in a non-invasive way.

Acknowledgments

This work is partly supported by an FPU fellowship (*Beca de Formación de Profesorado Universitario*) from the Spanish Ministry of Science, Innovation and Universities (FPU16/06487).

References

[1] Bray F *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, vol 68, no 6, 2018, pp 394–424, (ISSN: 1542-4863).

[2] Kanitkar AA *et al.* Causes of death in long-term survivors of non-small cell lung cancer: A regional surveillance, epidemiology, and end results study. *Annals of Thoracic Medicine*, vol 13, no 2, 2018, pp 76–81, (ISSN: 19983557).

[3] Pawelczyk K *et al.* Role of PD-L1 expression in non-small cell lung cancer and their prognostic significance according to clinicopathological factors and diagnostic markers. *International Journal of Molecular Sciences*, vol 20, no 4, 2019, p 824, (ISSN: 14220067).

[4] Mascaux C *et al.* Personalised medicine for nonsmall cell lung cancer. *European Respiratory Review*, vol 26, no 146, 2017, (ISSN: 16000617).

[5] Kim I *et al.* Reliability of PD-L1 assays using small tissue samples compared with surgical specimens. *Medicine*, vol 98, no 14, 2019, p e14972, (ISSN: 15365964).

[6] Martínez Albero E *et al.* Association between PD-L1 expression and histology with SUVmax on 18F-FDG PET/CT in non small-cell lung cancer patients. *European Journal of Nuclear Medicine and Molecular Imaging*, vol 46, no Supplement 1, 2019, pp S579–S580.

[7] Hu B *et al.* 18F-FDG maximum standard uptake value predicts PD-L1 expression on tumor cells or tumor-infiltrating immune cells in non-small cell lung cancer. *Annals of Nuclear Medicine*, vol 34, no 5, 2020, pp 322–328, (ISSN: 18646433).

[8] Wu X *et al.* PD-L1 expression correlation with metabolic parameters of FDG PET/CT and clinicopathological characteristics in non-small cell lung cancer. *EJNMMI Research*, vol 10, no 1, 2020, (ISSN: 2191219X).

[9] Zhao L *et al.* Relationship between SP142 PD-L1 Expression and 18F-FDG Uptake in Non-Small-Cell Lung Cancer. *Contrast Media & Molecular Imaging*, vol 2020, 2020, pp 1–11, (ISSN: 1555-4309).

[10] Nioche C *et al.* Lifex: A freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Research*, vol 78, no 16, 2018, pp 4786–4789, (ISSN: 15387445).

[11] Boellaard R *et al.* FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *European journal of nuclear medicine and molecular imaging*, vol 42, no 2, 2015, pp 328–54, (ISSN: 1619-7089).

[12] Ha S *et al.* Radiomics in Oncological PET/CT: a Methodological Overview. *Nuclear Medicine and Molecular Imaging*, vol 53, no 1, 2019, pp 14–29, (ISSN: 1869-3474).

[13] scikit-learn developers. 3.1. Cross-validation: evaluating estimator performance. 2020. https://scikit-learn.org/stable/modules/cross_validation.html#cross-validation (Last access: Aug 2020).

Geometry performance study on scintillator crystal for hexagonal SiPM application

David Pérez-Benito[1], Rigoberto Chil[1], Leandro A. Hidalgo-Torres[1], Juan José Vaquero[1,2]

1: Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid, España, {david.perez.benito@uc3m.es}
2: Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España

Abstract

Scintillation light collection influences essential characteristics in PET such as energy resolution and timing. The type of reflector, finish treatment and geometry of the crystal, both shape and ratio width-height, greatly affects its efficiency. This work presents a comparative study of the light output, scintillation pulse timing and energy resolution for four geometries (cylinder, hexagon, square and triangle) of the same cross section with two border roughness and two reflector types: specular and diffuse. This aims to determine the optimal choice for the segmented crystals of a new PET scanner design of hexagonal scintillation blocks. Rough crystal provides higher light output, but it is more sensitive to scintillation height with respect to the photodetector. The wider angular distribution of reflected rays degrades the time and energy resolution. According to the evaluated performance characteristics and the favorable tessellation of the hexagonal block, polished hexagon with specular reflector is found to be the best suiting choice.

1. Introduction

Position Emission Tomography (PET) is a functional nuclear imaging technique capable of estimating the spatial distribution of molecules of biological interest within the subject. These molecules are bound to radioisotopes, which upon nuclear decay, emit two gamma photons isotopically. Their interaction with the detectors serves to finally conform the distribution of the molecules.

From the geometrical perspective, the optimal spatial arrangement of the detectors is a sphere. However, in every commercial PET scanner, detectors are disposed in rings with the subject laying along the longitudinal axis. State of the art of new PET scanner designs aims to improve sensitivity, spatial and temporal resolution by enlarging the axial extent of the scanner [1]. A feasible alternative to the ideal spherical geometry is a truncated icosahedron, as it was proposed in our previous work [2]. The subsequent geometrical consequence is that its triangular faces should not be packed with traditional square scintillator crystals so empty gaps are minimized. Instead, both scintillators and photodetectors are hexagons of 15.25 mm of edge as shown on the left-hand side of Figure 1.

The scintillator crystal is a dense solid material that emits light isotopically when it is excited by an ionizing event. Light is finally sensed by the photodetector placed in the opposite face of the crystal. It is common practice from the 80‘s to segment the scintillator block into small regular crystals to set an optical barrier around them and therefore confine the light into the crystal where the burst occurs [3]. This way, the scintillation position can be estimated with higher precision. Conventional square crystals on the hexagonal scintillation do not provide the best coverage, Figure 1 (center). Hence, it is necessary to explore other geometries; hexagons or triangles can be laid out to shape a larger hexagon, as shown in Figure 1 (right).

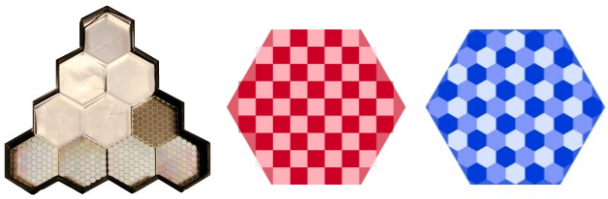


Figure 1. *Left: one of the twenty faces of the icosahedron. For demonstration purposes, in the upper half are placed the scintillations and in the lower half one can find the photodetectors. Middle: segmentation of the hexagonal block with square crystals. Right: tessellation with hexagonal crystals.*

The desirable characteristics for the scintillator-photodetector system are uniformity, fast pulse rise and decay time, good energy resolution and high light output. In this context, uniformity stands for similar pulse amplitude and width for ionizing particles of the same energy, in other words, number of photons and timing, independently of the scintillation point. Fast decay time enhances the scanner counting capabilities. Large photon collection also improves energy and timing resolution [4].

The most significant enhancement in light output and timing has come with the development of new scintillation materials such as Lutetium-yttrium oxyorthosilicate (LYSO). However, the geometry of the crystal has not been matter of profound analysis and square crystal was embraced as the standard solution. The main role of geometry in light output is the light trapped within the crystal due to total internal reflection. For regular bodies like parallelepipeds or cylinders, analytical studies were subject of research in the middle of the past century [5]. It is well known since then, that high symmetric crystals have lower light output and in the case of cylindrical bodies, it also suffers from radial variance [6]. Upon consideration of complex physical phenomena such as Compton scatter, absorption or the smoothness of the borders of the crystal, analytical calculations are not feasible and Monte Carlo simulations are widely used. Studies for PET applications have proposed the use of triangular crystals to improve the edge identification of the scintillation block [7] or crystals with slant planes to break the symmetry [8].

In this work, a Monte Carlo based study of the role of the geometry together with two reflector types (specular and

ISBN: 978-84-09-25491-0, pags. 281- 284

Lambertian) and finish treatments (polished and rough) is presented. This analysis aims to find the best suiting combination for the hexagonal block application in terms of light output, timing and energy resolution.

2. Methods

Monte Carlo simulations were performed in GATE [9] with LUT Davis model for optical border interaction [10]. This model accounts for the uneven elevation and irregular orientation of the surface of the crystal which has a strong dependency of the surface treatment. Coarse ground finish can be achieved for instance with saw cut and grit paper polish. The range of the coarseness this type of treatment ranges in the order of micrometers. Further mechanical polish can reduce the differences to nanometers. Hereafter, the first is referred as rough whereas the latter is referred as polished. The diffuse reflector is Teflon (PTFE), wrapped 5 times around the crystal and the specular reflector is ESR (Enhanced Specular Reflector 3M^{TM}). The first has a reflectivity index of 0.95 and the latter 0.98 in the spectrum range of the LYSO scintillator. The light yield is 35000 photos/MeV and the attenuation length is 800 mm for all wavelengths.

Four geometries are studied: hexagon (1.25 mm of edge), square (2.01 mm of edge), cylinder (diameter of 2.27 mm) and equilateral triangle (3.06 mm of edge), Figure 2 (left). Every crystal has a height of 10 mm. All figures have the same cross section to equate the subtended solid angle for equal scintillation point. Other options such as geometries inscribed in a cylinder of common radius would bias the study of the light output toward wide geometries. In principle, hexagon and triangle are good candidates because they can fit into the hexagonal block without gaps as shown in Figure 1. Additionally, in the approach towards the ideal spherical geometry, the current truncated icosahedron design could be further refined into Catalan solid such as hexakis icosahedron, whose triangular faces can be tessellated with triangles. Square is the common geometry and it is valuable to assess its characteristics with respect to the other figures. Cylinders have good light transfer characteristics as it is known in the field of optical fibers. Even though circular segmentation with cylinder would lead to empty space between crystals, it is included as matter of comparison.

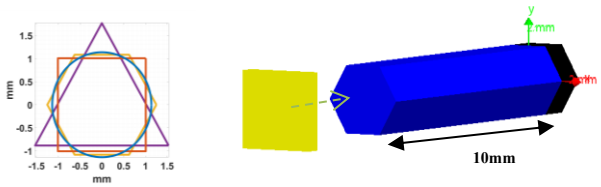


Figure 2. *Left: Cross section view of the four geometries. Right: representation of the hexagonal crystal (blue), radioactive source (yellow) and photodetector (black). The dashed green arrow represents the emission direction of gamma photons.*

To reduce the study to the mere influence of the geometry, the simulation has been simplified as follows. The source is defined as flat and radiates uniformly the crystal with monoenergetic 511keV gamma photons in perpendicular direction with respect to the upper surface of the crystal, Figure 2 (right). Only gamma photons that have made full energy deposition by photoelectric effect are considered. With this, dispersion in the number of detected photons due Compton interaction is excluded. For the four geometries, the scintillator material is LYSO with refraction index of 1.81. In every simulation, the upper surface of the crystal is considered polished and covered with diffuse reflector. The coupling media of the crystal and the photodetector is optical grease with refraction index of 1.5. Reflectors are air-coupled to the crystal.

Light output is estimated as the mean and standard deviation of the number of optical photons arriving at the photodetector interface. Results are sliced by the height of the scintillation position with respect to the photodetector. Timing is evaluated by means of the decay time of the scintillation pulse. The pulse itself is a time histogram of the number of photons. The decay time is defined as the time it takes for the pulse to reach *1/e* of the total decay time. An estimation of the energy resolution is obtained by counting the number of photons carried by the scintillation pulse over a firing threshold. This resembles to pulse integration on real experiment scenario. The number of photons are histogrammed and the energy resolution is presented as the FWHM. No distinction on the scintillation position within the crystal is considered. In every experiment, the same number of photoelectric events per height slice are considered for each geometry and combination of reflector and surface finish.

3. Results

3.1. Light Output

Light output, estimated as the average number of photons, is presented for the four geometries and surface finish in Figure 3 for Teflon, Lambertian or diffuse reflector and in Figure 4 for specular reflector. Results are shown for ten different scintillation heights.

In both cases, crystals with rough finish output larger number of photons. This is contrary to purely geometrical considerations of polished bare crystals. It is well-known that for symmetrical bodies, only rays emitted at certain angles can escape the crystal due of internal reflection [6]. The rough finish partially randomizes the orientation of the reflection borders and larger number of rays escape the crystal independently of their direction. However, in the presence of reflector, photons are redirected towards the crystal and it causes an enormous improvement in light collection, especially for rough crystal. Furthermore, this finish breaks the crystal symmetry which enhances light collection [5][11].

Scintillation height dependency is due to the number of photons that undergo reflection. If no reflector was considered, the number of detected photons would be proportional to the subtended solid angle and independent on the reflector type or surface treatment. No substantial variation in height is observed for polished crystal. It also has shorter variation among pulses (see the error bars which stands for the standard deviation).

Diffuse reflector improves light collection for rough crystals, especially for the lower slices of the crystal. The

favorable direction of diffuse reflectors towards its normal, promotes the re-entrance of the ray towards the detector.

In every of the four combinations and every single slice depth, the highest light output is given by the cylinder, followed by the hexagon, square and triangle. The first behaves like a light guide and the fraction of internal reflection is larger. This also explains the order of geometries.

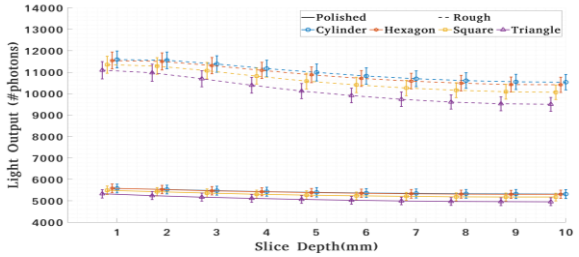


Figure 3. *Light output with 5 times Teflon wrapping*

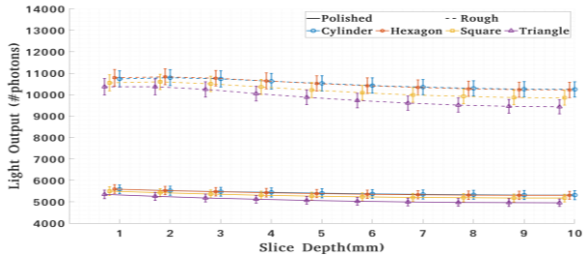


Figure 4. *Light output with ESR reflector*

3.2. Timing

Decay time is presented for the four geometries with diffuse reflector in Figure 5 and specular reflector in Figure 6 for both rough and polished finish.

Scintillation pulses of rough crystal have longer tail as a result of the variable reflection direction that reflected photons suffer. This is directly translated in longer decay time. Its dispersion in the middle section of the crystal is significantly high for rough crystal. The effect of wide angular distribution is added to the distance travelled by photons emitted away from the photodetector. In the lower half, due to the larger solid angle subtended by the detector from the scintillation point, the pulse is composed by direct photons and photons that reflected only few times. Thus, the role of the surface is not remarkable.

Time decay is unaffected by the reflector type for polished crystals. Although the angular distribution of both reflectors is certainly different, photons not reflected in a favorable direction are finally absorbed or escape the crystal. In the case of rough finish, decay time is approximately 5 picoseconds larger for diffuse reflector for every geometry in the upper half. This is expected for scintillation far from the photodetector since photons that escape the crystal are directed back into the crystal.

Time decay with respect to the geometry only revels dependency for rough crystals with increasing effect from the middle of the crystal towards the furthest side. The apparent independency on the geometry for specular reflector, suggests that polished crystals have larger number of reflections and part of the photons are either absorbed or escape. The comparative choice of equal cross section does introduce certain bias to study pulse timing. For a triangle, a photon crossing the complete figure will travel longer distance than in a cylinder. On the other hand, the cylinder tend to concentrate higher photon density on the outer part of the geometry which in turn implies long distances as well. These effects seem to be balanced.

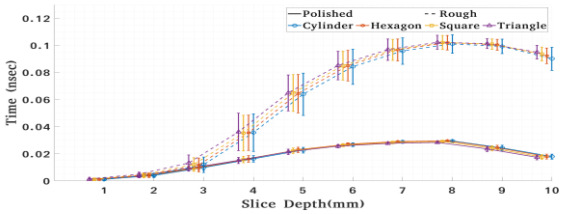


Figure 5*. Time decay with diffuse reflector.*

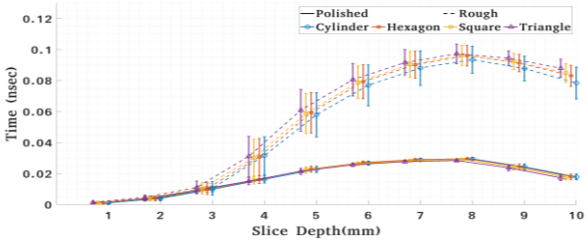


Figure 6. *Time decay with specular reflector.*

3.3. Energy Resolution

The energy resolution, estimated as the FWHM of simple Gaussian fitting for each geometry, finish and reflector, is presented in
Table ***1***. The favorable properties of the polished cylinder in terms of light output are not consistent with the energy resolution. The reason is the radial variation of the light output: the closer the scintillation to the surface the higher trapping fraction. This is not the case for the hexagonal scintillator since photons can eventually escape the crystal after few internal reflections, providing better energy resolution. The lower performance of the triangle can be argued by the selection of the cross section as equalizing parameter since theoretically the geometry does not prone light trapping or non-uniformity.

Finish	***Rough***	***Polished***	***Rough***	***Polished***
Reflector	***ESR***		***Teflon***	
Cylinder	*9.72*	*10.49*	*11.98*	*10.48*
Hexagon	*10.06*	*9.82*	*12.30*	*9.78*
Square	*10.80*	*10.03*	*13.32*	*10.03*
Triangle	*12.33*	*10.56*	*12.81*	*10.50*

Table 1*. Energy resolution estimated as the FWHM.*

As expected, rough finish has poorer energy resolution. This is again explained by the randomized angular

distribution of the reflected rays. However, it might be striking that polished finish with specular reflector has worse energy resolution for cylinders. As stated above, light output is radial dependent in cylinders. Roughing the surface, breaks total internal reflection and alleviates the non-uniformity. This improvement is not observed with diffuse reflector. The variation on light collection with height is thought to be the reason.

4. Discussion and Conclusions

Scintillation crystal characteristics are affected by various parameters such as ratio width-height of the crystal, surface treatment, reflector type, coupling media, geometry of the crystal and, to a lesser extent, by attenuation length. Thus, the combination of them should be studied for each application. In this work four geometries, with polished and rough finish combined with diffuse and specular reflectors, are compared in terms of its light output, scintillation decay time and energy resolution.

The results of the light collection and timing shows the effect of the wide reflector angular distribution of rough crystals. This breaks the symmetry of the crystal and the roles of the internal reflection and the trapped rays are less significant. In contrast, photons travel longer distances to finally reach the photodetector. This might be counterproductive in the computation of the energy carried by the pulse since in case of longer and low tails, part of the pulse might be under the firing threshold. Instead pulse integration, the energy estimation by the maximum of the amplitude of the pulse should be considered.

Despite of its lower light output, polished finish reveals favorable timing. This is of special important for the application in high-sensitivity scanner. The arrangement of hexagons within the larger hexagonal scintillation block is a significant advantage to finally decide on hexagons. Furthermore, the good results in energy resolution will also enhance the imaging scanner capabilities. Therefore, results suggest selecting hexagons with polish and specular reflector.

The immediate future work is to validate the simulation results with experimental data. The number of photons cannot be directly counted but it can be estimated upon calibration with the analog-to-digital converter (ADC) signal. For a complete evaluation, the spatial resolution in a tomography experiment will be also considered. The extent of optical crosstalk will significantly affect the image quality. In this sense, splitting the hexagon in six triangles will be further investigated to promote spatial resolution.

Acknowledgment

Ministerio de Ciencia, Innovación, Proyecto "PID2019-109820RB-I00"; Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación, proyecto "TEC2016-78052-R / MINECO/FEDER, UE"; This project has received funding from 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 organization, Bill & Melinda Gates Foundation and University of Dundee.

Authors would like to thank the high-performance computer Vlaams Supercomputer Centrum and the Medical Image and Signal Processing (MEDISIP) at Ghent University for granting access to run the simulations.

References

[1] S. R. Cherry, T. Jones, J. S. Karp, J. Qi, W. W. Moses, and R. D. Badawi, "Total-body PET: Maximizing sensitivity to create new opportunities for clinical research and patient care," *J. Nucl. Med.*, vol. 59, no. 1, pp. 3–12, 2018.

[2] D. Perez-Benito, R. Chil, J. M. Udias, M. Desco, and J. J. Vaquero, "SiPM-based PET detector module for a 4π span scanner," *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 936, no. August 2018, pp. 18–21, 2019.

[3] M. E. Casey and R. Nutt, "A Multicrystal Two Dimensional BGO Detector System for Positron Emission Tomography," *IEEE Trans. Nucl. Sci.*, vol. 33, no. 1, pp. 460–463, 1986.

[4] S. E. Derenzo, W. S. Choong, and W. W. Moses, "Fundamental limits of scintillation detector timing precision," *Phys. Med. Biol.*, vol. 59, no. 13, pp. 3261–3286, 2014.

[5] W. Shurcliff and C. Jones, "The trapping of fluorescent light produced within objects of high geometrical symmetry," *J. Opt. Soc. Am.*, vol. 39, no. 11, pp. 912–916, 1949.

[6] Y. Xiaoguang, "A study of light collection in scintillation," vol. 228, pp. 101–104, 1984.

[7] P. Peng, C. Liu, J. Du, X. Bai, and S. R. Cherry, "Improving edge crystal identification in flood histograms using triangular shape crystals Improving edge crystal identi fi cation in fl ood histograms using triangular shape crystals," 2018.

[8] C. Carrier and R. Lecomte, "Theoretical modelling of light transport in rectangular parallelepipedic scintillators," *Nucl. Inst. Methods Phys. Res. A*, vol. 292, no. 3, pp. 685–692, 1990.

[9] S. Jan *et al.*, "GATE : a simulation toolkit for PET and SPECT," *Phys. Med. Biol.*, vol. 49, pp. 4543–4561, 2004.

[10] E. Roncali, M. Stockhoff, and S. R. Cherry, "An integrated model of scintillator-reflector properties for advanced simulations of optical transport," *Phys. Med. Biol.*, vol. 62, no. 12, pp. 4811–4830, 2017.

[11] G. Keil, "Design principles of fluorescence radiation converters," *Nucl. Instruments Methods*, vol. 89, no. C, pp. 111–123, 1970.

Corrección del Artefacto de Truncamiento en TAC mediante Aprendizaje profundo

P. M. Berdón[1,2], C. F. Del Cerro[1,2], R. C. Gimenez[2], M. Desco[1,2,3,4], M. Abella[1,2,3]

[1] Dept. Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid, España

[2] Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España

[3] Centro Nacional Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, España

[4] Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid, España

Resumen

La adquisición de proyecciones incompletas debido a que parte de la muestra se extiende fuera del campo de visión, resulta en inconsistencias en los datos que dan lugar a lo que se conoce como artefacto de truncamiento. Se han propuesto varios métodos para la compensación de la falta de datos, basados en la extrapolación de las proyecciones, pero ninguno consigue recuperar completamente los datos truncados.

Este trabajo presenta un nuevo método de compensación del artefacto de truncamiento en imágenes de TAC basado en técnicas de aprendizaje profundo. La evaluación en datos simulados a partir de estudios de roedores muestra la viabilidad de la propuesta.

1. Introducción

La obesidad es una enfermedad mundial con una incidencia del 13% de la población adulta y un 18% en niños y adolescentes en 2016 según la OMS [1]. Existen países en los que esta enfermedad tiene una mayor prevalencia como en Estados Unidos. Esto puede ser un problema en estudios de Tomografía Axial Computarizada (TAC) ya que las posibilidades de que el paciente se extienda fuera del campo de visión son mayores, produciendo un truncamiento en los datos.

Esta falta de datos hace que no sea cuantitativa la imagen reconstruida y aparezcan artefactos como el halo brillante que se muestra en la Figura 1.

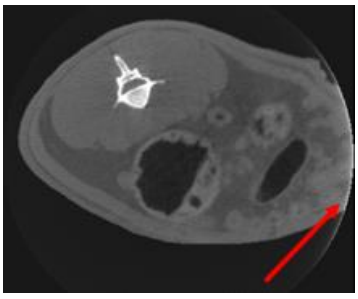


Figura 1.*Corte axial de un TAC de roedor con proyecciones truncadas. La flecha roja señala el artefacto brillante debido al truncamiento*

En la literatura encontramos algoritmos para la corrección del truncamiento basados en la extrapolación de los datos tanto en el dominio la proyección como en el dominio de imagen reconstruida. En [2], los autores proponen una extrapolación de espejo simétrico de las proyecciones truncadas basada en un modelo de sinusoide, que asegura una caída de los valores de atenuación suave, garantizando la continuidad en el punto de truncamiento. En [3, 4] se propone un método que se basa en modelar la muestra como si estuviera inmersa en un cilindro lleno de agua. De esta forma, se ajusta el cilindro basándose en la intensidad de las proyecciones no truncadas colindantes. Otro método explorado en [5] se basa en completar las sinusoides del sinograma truncado. A diferencia de los métodos anteriores, que buscan corregir el truncamiento extendiendo las proyecciones, en [6] los autores tratan de compensar la falta de datos en la imagen reconstruida mediante extrapolación de las proyecciones truncadas usando la técnica *Deformable Image Registration*, de tal forma que sus proyecciones coincidan con las proyecciones del TAC de referencia.

En este trabajo se propone un método automático de corrección de truncamiento en imágenes de TAC basado en la extrapolación de las proyecciones haciendo uso de técnicas de aprendizaje profundo. Para el entrenamiento de la red se usan datos simulados de truncamiento en a partir de estudios TAC de roedor.

2. Algoritmo propuesto

El método propuesto hace uso de la arquitectura U-Net [7] siguiendo el diseño de la implementación de [8], basada en la arquitectura codificador-decodificador, diseñada para la segmentación de imágenes biomédicas (Figura 2). La razón de elegir esta red es que permite mantener la resolución espacial entre la imagen de entrada y salida. El número de bloques sigue lo descrito en [8], pero se modifica el número de canales de características a 64, tras probar diferentes valores, para extraer más detalles de las imágenes y propagar la información útil a las capas de mayor resolución. Se usa el optimizador de Adam, que converge más rápido y es más estable que otros optimizadores [9], con una tasa de aprendizaje de 0.001 después de probar valores en el rango de 0.0001 a 0.03. La función de coste es el error cuadrático medio (MSE), comúnmente usada [10] y fácil de implementar.

ISBN: 978-84-09-25491-0, pags. 285- 288

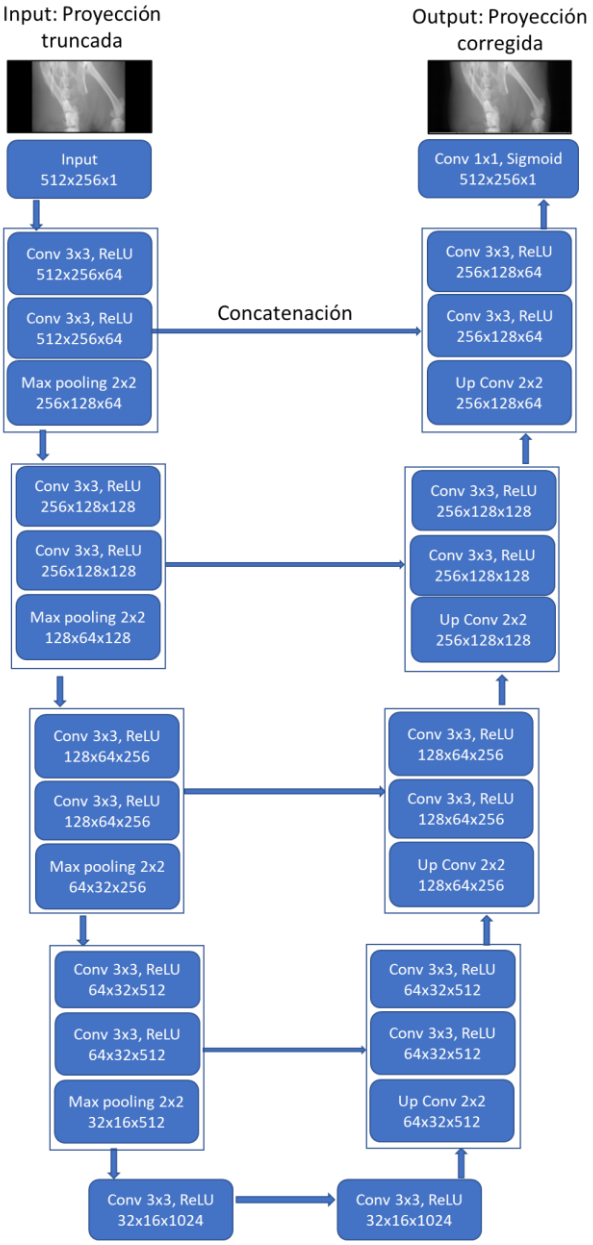


Figura 2.*Arquitectura de la red de U-Net utilizada para la corrección del truncamiento. Debajo de la especificación de cada capa aparece la dimensionalidad de un solo ejemplo que emite esta capa*

3. Generación de la base de datos

Para generar la base de datos se ha utilizado un estudio TAC de roedor adquirido con el sistema micro-TAC Add-On de SEDECAL [11], obteniendo 360 proyecciones de tamaño 512×512 píxeles, cubriendo todo el rango angular de 360 grados.

Para ampliar la base de datos, el estudio se divide a lo largo del eje axial en tres subconjuntos de proyecciones de 512×256 píxeles, como se muestra en la Figura 3.

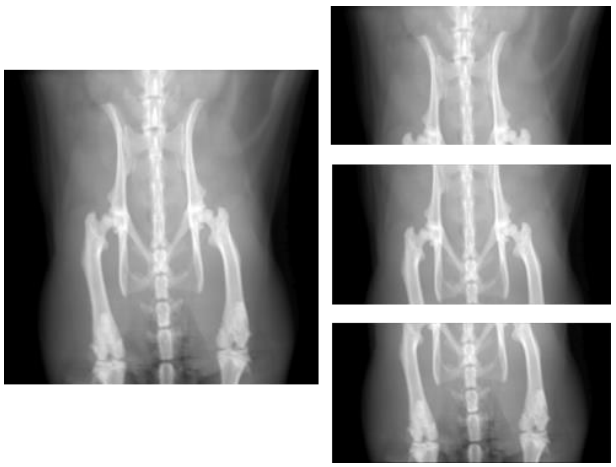


Figura 3. *Proyección original del estudio de TAC (izquierda) y subconjuntos de datos asociados (derecha)*

Se han simulado once casos realistas de truncamiento para cada subvolumen, poniendo a cero las bandas externas de la proyección (Figura 4). Como resultado, se dispone de un total de 2880 imágenes de entrenamiento y 1080 imágenes para validación.

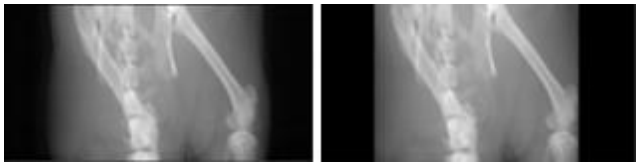


Figura 4.*Proyección sin truncar (izquierda), proyección truncada (derecha)*

4. Evaluación

La evaluación del método se hace con un TAC de roedor que no se ha usado para el entrenamiento de la red, con un tamaño de la matriz de las proyecciones de 512×256 píxeles, adquirida con el mismo sistema. La reconstrucción de la imagen se hace con Raptor, un algoritmo basado en FDK [12, 13], resultando en un volumen de 512×512×256. La Figura 5 muestra la zona truncada sobre una proyección de este estudio y la imagen reconstruida.



Figura 5.*Ejemplo de proyección y corte axial del estudio de TAC usado para evaluación*

El método propuesto se compara con el método de Ohnesorge et al. [2], el algoritmo más citado en la literatura, en base al *Root Mean Square Error (*RMSE) entre la reconstrucción ideal y la que se obtiene con las proyecciones truncadas corregidas.

5. Resultados

La Figura 6 muestra la recuperación de los datos truncados superior con el método propuesto que con el algoritmo de

Ohnesorge, que extiende demasiado el tejido blando de la proyección.

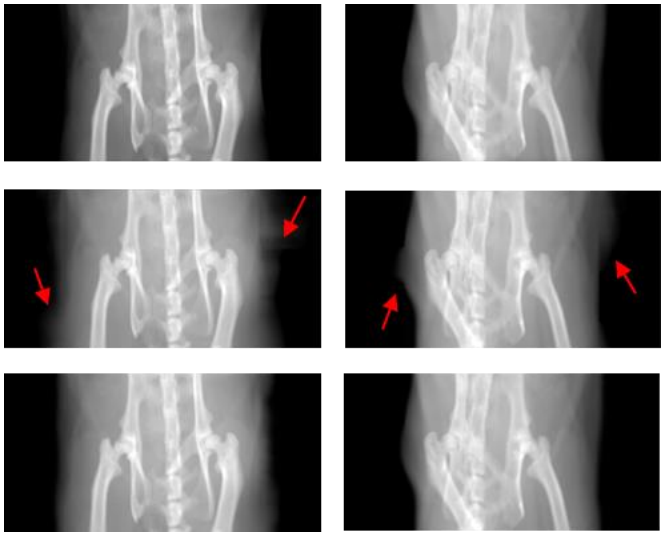


Figura 6. *Proyecciones originales (arriba), proyecciones truncadas corregida por el algoritmo de Ohnesorge (centro) y proyecciones truncadas corregida por el método propuesto (abajo)*

La Figura 7 muestra una mejor recuperación de los datos truncados con el método propuesto, consiguiendo eliminar la banda de más intensidad que introduce el algoritmo de Ohnesorge.

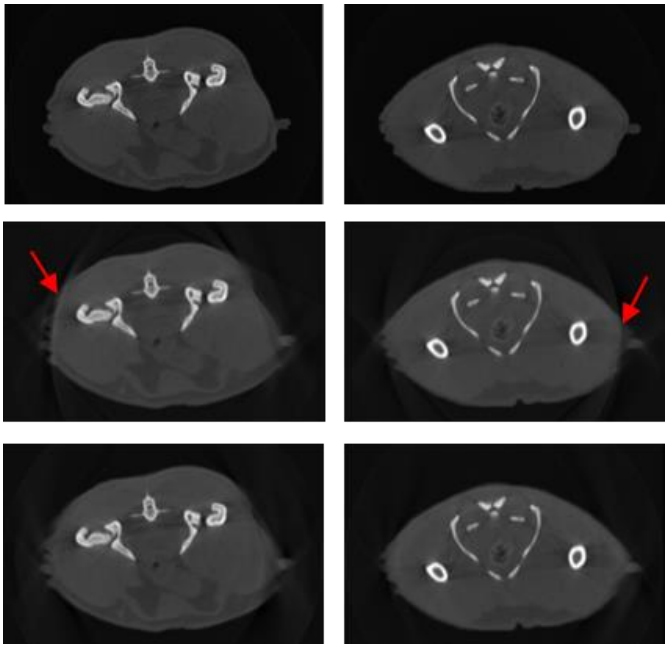


Figura 7. *Reconstrucción proyecciones originales (arriba), reconstrucción de las proyecciones truncadas corregidas mediante el algoritmo de Ohnseorge (centro) y reconstrucción de las proyecciones truncadas corregidas por el método propuesto (abajo)*

La Tabla 1 muestra un error 15% inferior con el método propuesto respecto al algoritmo de Ohnseorge, resultando en valores más parecidos a la imagen original.

	RMSE
Algoritmo de Ohnesorge	0.0203
Método propuesto	0.0176

Tabla 1. *RMSE de la reconstrucción para los dos métodos de evaluación*

6. Discusión y trabajo futuro

Este trabajo presenta un algoritmo de corrección del artefacto de truncamiento basado en la extrapolación de las proyecciones usando técnicas de aprendizaje profundo.

La evaluación en datos de TAC reales muestra una recuperación significativa del artefacto de truncamiento, un 15% superior que la obtenida con el método clásico propuesto por Ohnesorge, ampliamente citado. Sin embargo, se observa un salto en el borde entre datos medidos y datos extrapolados. Esto se puede deber a la una poca variabilidad en los datos de entrenamiento o a un limitado número de canales. Por otro lado, sería interesante evaluar casos con truncamiento de zonas el hueso, ya que sus valores son muy altos y no hay una transición suave entre los bordes.

Se observa una pérdida de resolución de la imagen, debida seguramente a la función de coste utilizada. La elección de esta función de coste se hizo en base a su simplicidad, que permitía una primera aproximación para obtener unos primeros resultados y evaluar la viabilidad del método. En la siguiente versión del método se explorará su sustitución por una función de coste basada en *perceptual loss* [14], que compara la salida de los filtros de las capas convolucionales para conseguir más detalles.

En conclusión, se ha propuesto un nuevo método de corrección del artefacto de truncamiento que permite compensar las inconsistencias en los datos, eliminando los artefactos y recuperando la cuantitividad de la imagen reconstruida, en menos de 16 ms.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia, Innovación y Universidades (Instituto de Salud Carlos III, proyecto DTS17/00122; Agencia Estatal de Investigación, proyecto DPI2016-79075-R-AEI/FEDER, UE), cofinanciado por Fondos de la Unión Europea (FEDER), “A way of making Europe”. Además, ha sido financiado por el Programa de apoyo a la realización de proyectos interdisciplinares de I+D para jóvenes investigadores de la Universidad Carlos III de Madrid 2019-2020 en el marco del Convenio Plurianual Comunidad de Madrid- Universidad Carlos III de Madrid (proyecto DEEPCT-CM-UC3M). El CNIC está financiado por el Ministerio de Ciencia, Innovación y Universidades y la fundación PRO-CNIC y es un centro de excelencia Severo Ochoa (SEV-2015-0505).

Referencias

[1] Salud OMdl. Obesidad y sobrepeso 2020 [cited 2020]. Available from: https://www.who.int/es/news-room/fact-sheets/detail/obesity-and-overweight.

[2] Ohnesorge B, Flohr T, Schwarz K. Efficient correction for CT image artifacts caused by objects extending outside the scan field of view. *Med Phys*. Vol. 27, Sup. 1, 2000, pp. 39-46.

[3] Hsieh J, Chao E, Thibault J. A novel reconstruction algorithm to extend the CT scan field-of-view. *Med Phys*. Vol. 31, Sup. 9, 2004, pp. 2385-91.

[4] Baojun Li, Junjun Deng, Albert H. Lonn, Hsieh J. An enhanced reconstruction algorithm to extend CT scan field-of-view with z-axis consistency constraint. *Med Phys*. Vol. 39, Sup. 10, 2012.

[5] Chityala R, Hoffmann K.R., S R. Artifact reduction in truncated CT using Sinogram completion. *Proc SPIE 5747, Medical Imaging*. 2005.

[6] Zhen X, Yan H, Zhou L, Jia X, Jiang S. Deformable image registration of CT and truncated cone-beam CT for adaptive radiation therapy. *Phys Med Biol*. Vol. 55, Sup. 22, 2013, pp. 7979-93.

[7] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, LNCS*. Vol. 9351, 2015, pp. 234-41.

[8] Buda M, Saha A, Mazurowski MA. Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm. *Computers in Biology and Medicine*. Vol. 109, 2019.

[9] Diederik P. Kingma, Ba J. Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations*. 2014.

[10] Mohamed M. Zahra, Mohamed H. Essai, Ellah ARA. Performance Functions Alternatives of Mse for Neural Networks Learning. *International Journal of Engineering Research & Technology (IJERT)*. Vol. 3, Sup. 1, 2014.

[11] Vaquero JJ, Redondo S, Lage E, Abella M, Sisniega A, Tapias G, et al. Assessment of a New High-Performance Small-Animal X-ray Tomograph. *IEEE Trans Nucl Sci*. Vol. 55, Sup. 3, 2008, pp. 898-905.

[12] Abella García M, Vaquero López JJ, Sisniega A, Pascau González Garzón J, Udías Á, García V, et al. Software architecture for multi-bed FDK-based reconstruction in X-ray CT scanners. *Computer Methods and Programs in Biomedicine*. Vol. 107, Sup. 2, 2012, pp. 218-32.

[13] Blas JG, Abella M, Isaila F, Carretero J, Desco M. Surfing the optimization space of a multiple-GPU parallel implementation of a X-ray tomography reconstruction algorithm. *Journal of Systems and Software*. Vol. 95, 2014, pp. 166-75.

[14] Justin Johnson, Alexandre Alahi, Fei-Fei L. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. *European Conference on Computer Vision*. 2016.

Modelado de Sistemas Médicos

Jueves, 26 de noviembre de 2020

Estimulación medular: optimización de la distancia entre polos de estimulación a partir de modelos de simulación

Jose-Luis Durá[1], Carmen Solanes[1], José De Andres[2,3], Javier Saiz[1]

1. Centro de Investigación e Innovación en Bioingeniería, Universitat Politécnica de Valencia, Valencia
2. Servicio de Anestesia Reanimación y Tratamiento del Dolor. Consorcio Hospital General Universitario. Valencia.
3. Unidad de Anestesia-Departamento de Cirugía Facultad de Medicina. Universidad de Valencia.

Resumen

La estimulación medular eléctrica es una terapia de elección en el tratamiento del dolor crónico severo de origen neuropático o isquémico. En la estimulación medular tónica, es condición necesaria pero no suficiente que las fibras Aβ correspondientes a los dermatomas del área de dolor sean estimuladas. La localización de dichas fibras estimuladas en los cordones posteriores depende de diversos factores como son el tipo de electrodo, su posición, y los parámetros de programación del neuroestimulador. El objetivo del desarrollo de un modelo de simulación de la neuroestimulación medular es el de suministrar a los profesionales un medio objetivo para optimizar la elección de estos parámetros: tipo y número de electrodos a utilizar, localización en el espacio epidural y programación de la señal eléctrica (definición de los polos activos y duración del impulso). En este estudio se evalúa cómo la distancia entre contactos en un mismo electrodo influye sobre los parámetros calculados por un modelo matemático, realizando el estudio sobre distancias entre 1 y 15 mm. Hemos observado unos resultados similares a otros estudios en cuanto a los umbrales de percepción, pero en lo relativo al área de fibras estimuladas en los Cordones Posteriores el máximo ha sido hallado para una distancia entre contactos de 9 mm de borde a borde, lo que contradice la creencia que hay respecto a que reduciendo la distancia entre polos se aumenta el área de estimulación. Hemos observado igualmente que una mayor distancia entre contactos aumenta las posibilidades de estimulación de raíces dorsales, generalmente no deseada.

1. Introducción

La estimulación medular eléctrica (EME) es una técnica utilizada para el tratamiento del dolor crónico severo desde hace más de 50 años. Actualmente unos 34.000 pacientes son tratados al año mediante esta terapia [1]. Existen dos tipos de EME en la actualidad: (i) la basada en la teoría de la puerta de control de Melzack y Wall [2] y aplicada por primera vez por Shealy [3], cuyo efecto es la creación de potenciales de acción en los Cordones Posteriores (CP) de la médula y/o las Raíces Dorsales (RD). Dicha estimulación provoca una sensación en el paciente denominada parestesia y cuya zona debe cubrir los dermatomas correspondientes al dolor en el paciente, y se conoce como Estimulación Medular Tónica. (ii) La estimulación denominada sub-umbral, que no provoca la creación de un potencial de acción y por tanto ninguna sensación en el paciente. Existen diversos tipos de estimulación sub-umbral (alta frecuencia, alta densidad, ráfaga, Differential Target Multiplexed -DTM, etc..) con un número importante de estudios clínicos demostrando su eficacia, pero no existe en la actualidad ningún mecanismo de acción descrito, haciendo imposible su estudio a partir de modelos matemáticos, puesto que no se conoce ni los elementos nerviosos sobre los que actúa ni su mecanismo de acción [4].

Nuestro modelo se centra en el estudio de diversos parámetros relativos a la estimulación medular tónica, que sigue siendo la más utilizada, y la única de la que disponemos de un modelo de fibra.

En esta investigación hemos estudiado cual es la distancia óptima entre polos en un mismo electrodo situado en la línea media anatómica de la médula, para una polaridad de Doble Cátodo Guardado, considerada la que genera mayor área de fibras estimuladas, en términos de Área de Fibras estimuladas en los Cordones Posteriores [5].

2. Material y métodos

Nuestro modelo, como el resto de los publicados por distintos grupos, se basa en dos modelos conectados. El primero es un modelo de Volumen Conductor consistente en la representación tridimensional del segmento de médula que deseamos estudiar [6]. Las dimensiones se han tomado a partir de mediciones reales en pacientes sanos mediante la técnica de Resonancia Magnética Nuclear [7]. En el modelo se incluye además el electrodo epidural responsable de crear un campo eléctrico en la medula, de forma que se pueden establecer las características eléctricas habitualmente utilizadas en la práctica clínica en forma de condiciones de contorno, tal y como se muestra en la Figura 1. Además de los tejidos y del electrodo, se ha incluido un tejido alrededor de los contactos activos que simulan la impedancia de la interfase electrodo-tejido, al igual que se ha hecho en algunos modelos de estimulación cerebral profunda [8], [9].

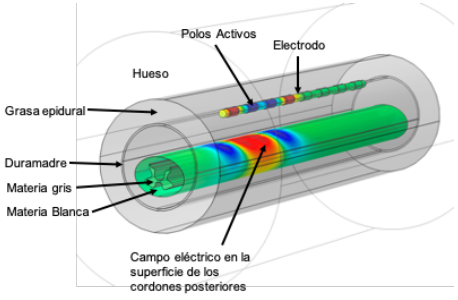


Figura 1. Modelo anatómico 3D en Comsol Multiphysics.

ISBN: 978-84-09-25491-0, pags. 290- 293

Utilizamos un mallado adaptativo con elementos tetraédricos cuadráticos y 3.2 millones de grados de libertad. En la Figura 2 puede apreciarse el modelo medular.

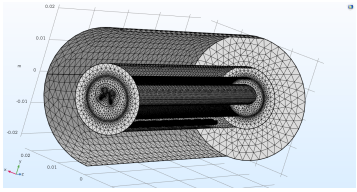


Figura 2. Malla generada en el modelo de conductor volumétrico.

Se calcula mediante el método de los elementos finitos, utilizando el software COMSOL MULTIPHYSICS v.5.3a, el potencial eléctrico como una solución aproximada a la ecuación de Laplace: $\nabla^2 V = 0$.

La densidad de corriente J(x,y,z) se calcula mediante la versión generalizada de la ley de Ohm:

$$J = \sigma * E$$
$$E = -\nabla V_e$$

Siendo V_e el potencial eléctrico en un punto del volumen conductor, E el campo eléctrico, J la densidad de corriente y σ el tensor de conductividad eléctrica.

Las conductividades de los distintos tejidos han sido extraídas de distintas publicaciones [10]-[13]. Se establecen como condiciones de contorno el potencial en los polos del electrodo de estimulación, en este caso como hemos dicho un doble cátodo guardado (+--+).

El modelo considera que, en este tipo de estimulación, dada la baja frecuencia a la que trabajamos, puede considerarse como cuasi-estática y despreciar los efectos capacitivos [14].

A partir de estos resultados se extrae el valor del campo Potencial Eléctrico a lo largo de un axón de los CP, así como en las raíces dorsales. Se exportan los valores para 40 filas de 100 axones a partir de la superficie de los CP, y en cada axón se calcula dicho campo para los nodos de Ranvier correspondientes a una fibra de 12.8 µm, considerándose 65 nodos de Ranvier por fibra. Del mismo modo se exportan los valores del campo potencial para las RD de ambos lados considerando los axones de un diámetro de 15 µm.

Estos datos se procesan en un programa desarrollado por nuestro grupo y funcionando bajo MATLAB v.R2018a basado en el modelo de fibra de MRG (McIntyre, Richardson y Grill) [15] puesto que es uno de los más utilizados en los modelos actuales, y de esta forma determinamos si un axón, con los datos de potencial en los nodos de Ranvier extraídos del conductor volumétrico, genera un potencial de acción, determinando tanto los umbrales de activación como el mapa de fibras estimuladas. En la Figura 3 se aprecia el potencial en los nodos de Ranvier a partir del modelo MRG, tanto en el caso de no llegar al umbral de generación del potencial de acción (izquierda) como en el caso de generarlo (derecha). En la Figura 4 se puede apreciar el mapa de fibras estimuladas en los CP.

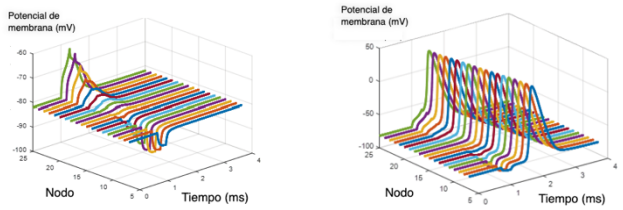


Figura 3. Potencial en nodos de Ranvier para un estímulo sub-umbral y supra-umbral

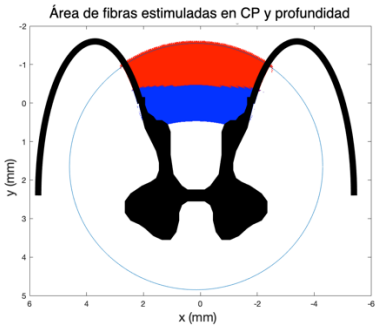


Figura 4. Área de fibras estimuladas en los CP.

De cada simulación extraemos los siguientes parámetros: Umbral de Percepción (Amplitud para la que se activa la primera fibra) bien en CP (UP_{CP}) o en RD (UP_{RD}), Ratio entre ambas, y el Área de fibras estimuladas en los CP (A_{CP}) así como la profundidad de éste área (P_{CP}). En realidad, cuando hablamos de área de CP estimulados, dada la baja población de fibras mayores de 12.8 µm en los CP [16] el significado de esta área es que, si se encuentra en ella una fibra de esas características, será "estimulada".

El parámetro que variamos en esta investigación es la distancia entre los polos del electrodo medido entre los bordes del contacto (Distancia entre Contactos o DEC) tal y como muestra la Figura 5, evaluando los parámetros descritos para una DEC entre 1 y 15 mm, con pasos de 1 mm entre 1 y 8 mm (además del valor de 1.5 mm DEC debido a que existe un electrodo en el mercado con esta medida), y a partir de 10 mm solamente estudiamos 12 y 15 mm DEC porque actualmente no existen electrodos en el mercado con estas distancias. Para todas estas medidas se estudia la posición en línea media anatómica del electrodo, programación de doble cátodo guardado y una duración de impulso de 200 µs.

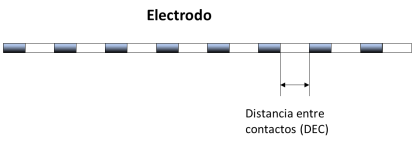


Figura 5: Distancia entre contactos (DEC) en un electrodo de estimulación medular.

El modelo matemático puede calcular el umbral de estimulación (en CP y en RD) para cualquier geometría y programación de la señal eléctrica, pero esta amplitud no es a la que el neuroestimulador va a trabajar habitualmente, puesto que en ese caso no se cubriría la zona deseada. Por ello se introduce el Umbral de Molestia (UM), que es el máximo valor al que puede programarse la amplitud sin que la parestesia se convierta en una sensación molesta para el paciente. Dicha amplitud es necesaria para el cálculo del Área de fibras estimuladas en los CP, uno de los parámetros más importantes a la hora de comparar distintas geometrías y programaciones. Prácticamente todas las investigaciones que utilizan simulaciones utilizan la amplitud del UM como amplitud del estímulo para el cálculo de A_{CP}, y se establece como 1.4 veces el Umbral de Percepción (ya que es el valor deducido a partir de estudios clínicos en pacientes). Lo que difiere es si se toma el Umbral de Percepción en los CP o en las RD. Algunos autores utilizan siempre UP_{RD} para calcular el umbral de molestia, independientemente del valor de UP_{CP}. En este estudio se ha decidido tomar el menor de ellos porque en general hemos observado que los pacientes refieren la parestesia dolorosa en la misma zona de inicio de la estimulación en el umbral de percepción, y en muchos casos la altura metamérica del cátodo resulta incompatible con una estimulación en las raíces dorsales. Esta hipótesis respecto al cálculo de la amplitud de la simulación puede cambiar por completo los resultados, como se verá más adelante.

3. Resultados

Se ha calculado para las distancias entre bordes de los contactos de 1 mm hasta 15 mm los parámetros que nos proporciona el modelo: UP_{CP}, UP_{RD}, Ratio entre ambos, A_{CP} para una amplitud igual a 1.4 veces el menor umbral de percepción (en CP o en RD) y P_{CP}, con los diámetros de fibra descritos anteriormente.

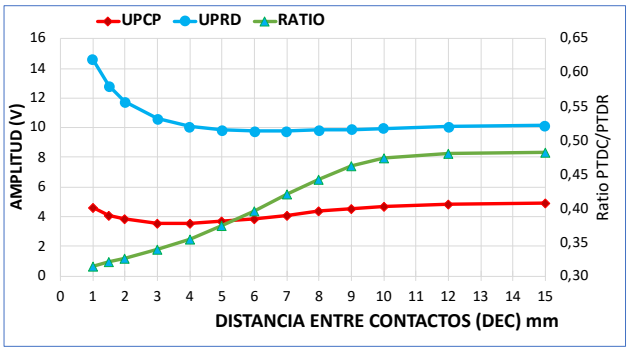


Figura 6. Variación de los Umbrales de Percepción y de la ratio entre ambos en función de la distancia entre contactos

En cuanto al U_{CP} observamos, tal y como muestra la Figura 6, un mínimo para 4 mm de DEC de 3.57 V. En cuanto al UP_{RD}, el mínimo está para una DEC de 6 mm (9.78 V) y se mantiene hasta 15 mm de DEC. Ambas posiciones presentan un patrón similar salvo en el caso de valores de DEC menores de 7 mm, donde el UP_{RD} es muy superior, haciendo imposible la estimulación de Raíces Dorsales en la posición medial para una DEC de 1 mm.

La ratio UP_{CP}/UP_{RD} para la posición central muestra un crecimiento constante entre 1 y 15 mm de DEC (desde 0.31 hasta 0.48) aunque la curva se aplana a partir de 10 mm DEC.

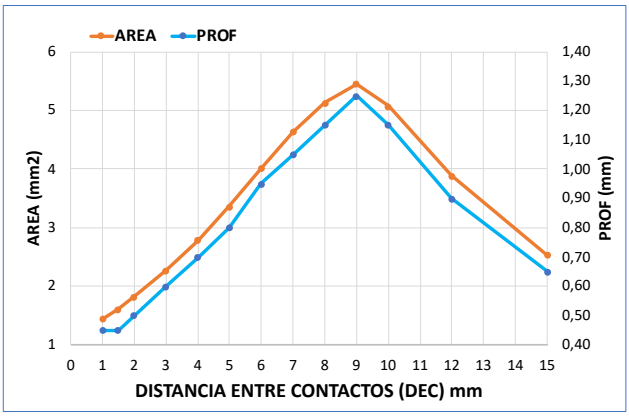


Figura 7. Variación del Área de Fibras estimuladas en los CP y de la Profundidad en función de la distancia entre contactos

En cuanto a A_{CP}, parámetro que consideramos fundamental para la optimización de la estimulación, el comportamiento que observamos es un aumento de dicha área desde el valor para 1 mm de DEC (1.4 mm^2) hasta un máximo para 9 mm de DEC (5.1 mm^2) y un descenso posterior hasta los 15 mm de DEC (2.5 mm^2), tal y como se muestra en la Figura 7. Esto significa que la diferencia de área de fibras estimuladas con un electrodo de 9 mm de DEC es un 350% mayor que para 1 mm de DEC. El comportamiento de la profundidad de dicha área es similar, yendo desde 0.45 mm para 1 mm DEC a 1.15 mm para 9 mm de DEC.

4. Discusión

Desde la publicación de las investigaciones de Manola [17] se ha dado por supuesto que el área de fibras estimuladas en los Cordones Posteriores era inversamente proporcional a la distancia entre polos activos. Esto se debe a que Manola siempre realiza los cálculos de área con una amplitud 1.4 veces el UP_{DR}, independientemente del valor del UP_{CP}. Dado que este valor crece exponencialmente para valores cercanos a 1 mm de distancia entre contactos, las amplitudes utilizadas para valores de DEC cercanas a 1 mm son muy superiores a las que nosotros utilizamos, resultando áreas (que son proporcionales a la amplitud de la simulación) muy superiores a las halladas por nosotros.

Sin embargo, si consideramos que el umbral de molestia se debe a una sobreexcitación de fibras tanto en las RD como en los CP, resulta un área máxima de activación de fibras para una DEC de 9 mm.

Otra consecuencia del aumento de la DEC es la mayor facilidad para activar RD (efecto en general no deseado) tal y como refleja la variación de la ratio entre umbrales.

Viendo los resultados resulta paradójico que en los últimos años se haya ido reduciendo la distancia entre contactos en los electrodos disponibles en el mercado. En aquellos

estimuladores en lo que puede definirse la corriente entre polos de forma independiente esto se explica porque el sistema de barrido que utilizan necesita una distancia mínima entre contactos para ser efectivo [18], pero resulta difícil de explicar en sistemas en que la corriente se reparte en función de la resistencia entre polos (independientemente de que se basen en un generador de corriente o de potencial).

5. Conclusiones

Creemos firmemente que el uso de modelos matemáticos de simulación es una herramienta muy potente para el análisis y comparación de distintas configuraciones de estimulación medular. Sin embargo, es necesario tener muy claro que, aunque pueda parecer un método perfectamente objetivo, presenta limitaciones a tener en cuenta. Las primeras se refieren al hecho de modelizar un fenómeno físico, lo que entraña necesariamente una simplificación del fenómeno. Pero en este estudio hemos visto que además las hipótesis de trabajo, como en este caso la elección de la amplitud de simulación, puede llevar a conclusiones opuestas. Es necesario por tanto ser muy meticuloso a la hora de analizar bajo qué hipótesis se efectúan los cálculos.

Con respecto al estudio de la distancia óptima entre polos, podemos concluir que la creencia de que una menor distancia implica una mayor área de fibras estimuladas en los CP, está basada en un artículo cuya hipótesis de trabajo no son las únicas posibles. Esto ha provocado que durante años se haya tendido a comercializar electrodos con los polos más cercanos, mientras que partiendo de otras hipótesis de trabajo las conclusiones serían que una distancia de 9 mm entre polos crearía la mayor superficie de estimulación. Hay que tener en cuenta que esta distancia crearía grandes diferencias en la zona estimulada al estar tan distanciados los cátodos, dificultando la localización idónea dede la zona de parestesia, de forma que la solución ideal será con toda probabilidad un compromiso entre ambos efectos contrapuestos.

Referencias

[1] Thomson, S. Spinal Cord Stimulation. Its Role in Managing Chronic Disease Symptoms. 2019. https://www.neuromodulation.com/assets/documents/Fact_Sheets/fact_sheet_spinal_cord_stimulation.pdf

[2] Melzack R, Wall PD. Pain mechanisms: a new theory. Science. 1965; 150:971-9.

[3] Shealy CN, Mortimer JT, Reswick JB. Electrical inhibition of pain by stimulation of the dorsal columns: preliminary clinical report.Anesth Analg. 1967 Jul-Aug; 46(4):489-91

[4] Capogrosso M, Lempka SF. A computational outlook on neurostimulation. Bioelectron Med. 2020;6:10

[5] Durá J.L., Solanes C., De Andrés J., Saiz J. Computational Study of the Effect of Electrode Polarity on Neural Activation Related to Paresthesia Coverage in Spinal Cord Stimulation Therapy. Neuromodulation 2019; 22: 269–279

[6] Plonsey, R. "Volume Conductor Theory." The Biomedical Engineering Handbook: Second Edition. Ed. Joseph D. Bronzino. Boca Raton: CRC Press LLC, 2000

[7] Fradet L, Arnoux PJ, Ranjeva JP, Petit Y, Callot V. Morphometrics of the entire human spinal cord and spinal canal measured from in vivo high-resolution anatomical magnetic resonance imaging. Spine (Phila Pa 1976). 2014 Feb 15;39(4): E262-9

[8] Butson CR, Cooper SE, Henderson JM, McIntyre CC. Patient-specific analysis of the volume of tissue activated during deep brain stimulation. Neuroimage. 2007 Jan 15;34(2):661-70

[9] Butson CR, Maks CB, McIntyre CC. Sources and effects of electrode impedance during deep brain stimulation.Clin Neurophysiol. 2006;117(2):447-54

[10] Struijk JJ, Holsheimer J, van Veen BK. Analysis of Dorsal Column Stimulation. Engineering in Medicine and Biology Society, 1988. Proceedings of the Annual International Conference of the IEEE

[11] Holsheimer J. Computer modelling of spinal cord stimulation and its contribution to therapeutic efficacy . Spinal Cord 1998;36:531-540

[12] Arle JE, Carlson KW, Mei L, Shils JL. Modeling Effects of Scar on Patterns of Dorsal Column Stimulation. Neuromodulation 2014; 17: 320–333

[13] Ladenbauer J, Minassian K, Hofstoetter US, Dimitrijevic MR, Rattay F. Stimulation of the human lumbar spinal cord with implanted and surface electrodes: a computer simulation study. IEEE Trans Neural Syst Rehabil Eng. 2010; 18(6):637-45

[14] Bossetti CA, Birdno MJ, Grill WM. Analysis of the quasi-static approximation for calculating potentials generated by neural stimulation J. Neural Eng. 5 (2008) 44–53

[15] McIntyre CC, Richardson AG, Grill WM. Modeling the excitability of mammalian nerve fibers: influence of afterpotentials on the recovery cycle.J Neurophysiol. 2002;87(2):995-1006

[16] Feirabend HK, Choufoer H, Ploeger S, Holsheimer J, van Gool JD. Morphometry of human superficial dorsal and dorsolateral column fibres: significance to spinal cord stimulation. Brain. 2002; 125(Pt5): 1137-49

[17] Manola L, Holsheimer J, Veltink PH, Bradley K, Peterson D. Theoretical investigation into longitudinal cathodal field steering in spinal cord stimulation. Neuromodulation. 2007;10(2):120-32

[18] Oakley J, Varga C, Krames E, Bradley K. Real-Time Paresthesia Steering Using Continuous Electric Field Adjustment. Part I: Intraoperative Performance. Neuromodulation. 2004;7(3):157-67

Modelado de la propagación óptica a través de tejidos biológicos anisotrópicos utilizando las funciones de Green

J. L. Ganoza-Quintana, F. Fanjul-Vélez, J. L. Arce-Diego

Grupo de Ingeniería Biomédica, Departamento TEISA, Universidad de Cantabria, Av. de los Castros s/n, 39005 Santander (Cantabria), España

fanjulf@unican.es, arcedj@unican.es

Resumen

Los tejidos biológicos son medios turbulentos debido a que las células que los componen tienen una gran cantidad de orgánulos, tales como las mitocondrias, los ribosomas, el citoplasma, etc. Todas estas estructuras tienen diferentes índices de refracción, por tanto, estos orgánulos actúan como esparcidores en la propagación de un haz óptico incidente en un tejido biológico. La influencia de una distribución particular de esparcidores puede ser analizada por las funciones de Green. En este trabajo se introduce un parámetro de anisotropía geométrica en el espectro de potencia del tejido biológico y se estudian sus efectos sobre las estadísticas de los haces ópticos que se propagan por el mismo. Además de los efectos de otros parámetros, como la variación del índice de refracción, la pendiente del espectro de potencia y las escalas internas / externas del tejido. También se presentan algunas técnicas de obtención de estos parámetros biológicos que influyen en el modelo matemático de la propagación de la luz.

1. Introducción

Las técnicas ópticas biomédicas de tratamiento, caracterización y cirugía dependen de la propagación de la luz en los tejidos biológicos [1]. Estos tejidos biológicos están formados por células que contienen diferentes orgánulos que actúan como esparcidores en la propagación de la luz. La precisión de la propagación de la luz y las estimaciones a priori son particularmente significativas en las técnicas de diagnóstico. Como el tejido biológico es un medio turbio, con gran influencia del esparcimiento, es necesario analizar adecuadamente su influencia sobre parámetros de propagación óptica, como la coherencia. La influencia de una distribución de esparcidores puede analizarse mediante las funciones de Green [2]. En este trabajo, en primer lugar, se define el marco teórico; a continuación, se estudia la propagación de haces Schell-Gaussianos por tejidos biológicos anisotrópicos. Luego se presentan algunas técnicas en la obtención de algunos parámetros biológicos que influyen en el modelo de propagación de la luz y finalmente se presentan los resultados y las conclusiones.

2. Modelo teórico

Sea una onda escalar $U^{(i)}(\vec{\boldsymbol{r}};\omega)$ viajando en el espacio libre que incide sobre un objeto, confinado a un volumen V, de índice de refracción $n(\vec{\boldsymbol{r}};\omega)$. Se determina la forma del campo esparcido $U^{(s)}(\vec{\boldsymbol{r}};\omega)$ y se define formalmente el campo total en la región del objeto de esparcimiento como $U^{(t)}(\vec{\boldsymbol{r}};\omega) = U^{(i)}(\vec{\boldsymbol{r}};\omega) + U^{(s)}(\vec{\boldsymbol{r}};\omega)$. El campo total satisface la ecuación de Helmholtz con un índice de refracción no uniforme y, por lo tanto, un número de onda no uniforme k [3],

$$[\nabla^2 + k^2 n^2(\vec{\boldsymbol{r}};\omega)]U^{(t)}(\vec{\boldsymbol{r}};\omega) = 0. \quad (1)$$

Entonces,

$$U^{(t)}(\vec{\boldsymbol{r}};\omega) = U^{(i)}(\vec{\boldsymbol{r}};\omega) + \int_D F\left(\overrightarrow{\boldsymbol{r'}};\omega\right) U^{(i)}\left(\overrightarrow{\boldsymbol{r'}};\omega\right) G\left(\vec{\boldsymbol{r}} - \overrightarrow{\boldsymbol{r'}};\omega\right) d^3r' \quad (2)$$

Donde la función $F(\vec{\boldsymbol{r}};\omega)$ se denomina potencial de esparcimiento del medio y $G\left(\vec{\boldsymbol{r}} - \overrightarrow{\boldsymbol{r'}};\omega\right)$ es la función de Green de la ecuación tridimensional de Helmholtz.

El modelo Gaussiano Schell (MGS) está basado en el siguiente modelo para la función de densidad espectral cruzada en 2 vectores espaciales $\overrightarrow{\boldsymbol{\rho_1'}}$ y $\overrightarrow{\boldsymbol{\rho_2'}}$, dada por,

$$W^{(G)}\left(\overrightarrow{\boldsymbol{\rho_1'}}, \overrightarrow{\boldsymbol{\rho_2'}};\omega\right) = A_0^2(\omega) e^{-\frac{\rho_1'^2 + \rho_2'^2}{4\sigma^2(\omega)}} e^{-\frac{\left(\overrightarrow{\boldsymbol{\rho_1'}} - \overrightarrow{\boldsymbol{\rho_2'}}\right)^2}{2\delta^2(\omega)}}, \quad (3)$$

Donde el superíndice (G) representa el modelo Gaussiano Schell, A_0^2 es el máximo valor de la densidad espectral y las anchuras de media cuadrática (r.m.s) σ^2 y δ^2 son independientes de la posición, pero generalmente dependen de la frecuencia [4].

La turbulencia óptica se explica bien por la presencia de irregularidades en el índice de refracción o los denominados "remolinos turbulentos", que aparecen debido a fluctuaciones en diversas propiedades físicas de la materia, como la temperatura, la presión y la concentración de contenido químico no homogéneo. Tales remolinos se crean en diferentes tipos de materia a través de ciertos mecanismos físicos / químicos / biológicos. El tamaño más grande posible de un remolino en el proceso turbulento se toma como la definición de la escala externa L_0 de turbulencia. Los remolinos más grandes se descomponen en otros más pequeños con energía, hasta que el tamaño del remolino alcanza el límite inferior cuando la energía se disipa. El tamaño del remolino más pequeño antes de la disipación define la escala interna l_0.

Las características estadísticas más importantes del índice de refracción en el espacio tridimensional son los dos primeros momentos: el valor medio de un campo,

$$n_0(\vec{\boldsymbol{r}}) = \langle n(\vec{\boldsymbol{r}})\rangle_M, \quad (4)$$

Y su función de covarianza,

ISBN: 978-84-09-25491-0, pags. 294- 297

$$B_n(\overrightarrow{r_1},\overrightarrow{r_2}) = \langle [n(\overrightarrow{r_1}) - n_0(\overrightarrow{r_1})][n(\overrightarrow{r_2}) - n_0(\overrightarrow{r_2})] \rangle_M, \quad (5)$$

Donde los corchetes angulares con subíndice M denotan el promedio del conjunto sobre las realizaciones del medio.

La interacción de los campos electromagnéticos con los medios turbulentos es un proceso muy complejo en el caso general cuando estos últimos son anisotrópicos y no homogéneos. En el rango de inercia de las escalas, a menudo se supone que los medios aleatorios son homogéneos (los momentos estadísticos del campo son invariantes a la translación).

En tales circunstancias, la relación entre la función de covarianza espacial $B_n(\vec{r})$ y el espectro de potencia o densidad espectral de potencia espacial $\Phi_n(\vec{\kappa})$, que determina la distribución de energía entre los remolinos de diferentes tamaños, tiene la forma del par tridimensional de la transformada de Fourier [3],

$$B_n(\vec{r}) = \iiint_{-\infty}^{\infty} e^{i\vec{\kappa}\cdot\vec{r}} \Phi_n(\vec{\kappa}) d^3\kappa, \quad (6)$$

$$\Phi_n(\vec{\kappa}) = \left(\frac{1}{2\pi}\right)^3 \iiint_{-\infty}^{\infty} e^{i\vec{\kappa}\cdot\vec{r}} B_n(\vec{r}) d^3 r, \quad (7)$$

Aquí $\vec{\kappa} = (\kappa_x, \kappa_y, \kappa_z)$ es el vector tridimensional, cuyos componentes tienen las unidades m^{-1}, representando frecuencias espaciales.

3. Propagación de un haz MGS por tejidos biológicos anisotrópicos

Se supone un haz de luz escalar modelo MGS incidente sobre un tejido biológico anisotrópico con factores de anisotropía $\mu_x = \mu_z \neq \mu_y$, en el plano $z = 0$, denominado plano fuente o plano de incidencia y se propaga a través de él en el semi-espacio positivo $z > 0$. Para los tejidos biológicos con anisotropía geométrica, el espectro de potencia tridimensional se puede escribir de la siguiente forma [5],

$$\Phi_n(\kappa_x, \kappa_y, \kappa_z) = \frac{(2\pi)^3 \sigma_n^2 \mu_x \mu_y \mu_z e^{-\frac{\kappa_x^2}{\kappa_{mx}^2} - \frac{\kappa_y^2}{\kappa_{my}^2} - \frac{\kappa_z^2}{\kappa_{mz}^2}}}{\kappa_0^{3-\alpha}\left[\kappa_0^2 + 4\pi^2\left(\mu_x^2\kappa_x^2 + \mu_y^2\kappa_y^2 + \mu_z^2\kappa_z^2\right)\right]^{\frac{\alpha}{2}}}, \quad \mathbf{3 < \alpha < 4}. \quad (8)$$

Donde α es la pendiente del espectro de potencia, σ_n^2 es la varianza del índice de refracción del tejido biológico, μ_x, μ_y y μ_z son los coeficientes de resistencia anisotrópica en cada dirección. Además, $\overrightarrow{\kappa_0}$ es el vector de frecuencia de corte a gran escala con magnitud $\kappa_0 = \sqrt{\kappa_{x0}^2 + \kappa_{y0}^2 + \kappa_{z0}^2}$ y componentes $\kappa_{x0} = \frac{2\pi}{L_x}, \kappa_{y0} = \frac{2\pi}{L_y}, \kappa_{z0} = \frac{2\pi}{L_z}$ en donde L_x, L_y y L_z son las escalas exteriores a lo largo de las direcciones x, y y z, respectivamente: $L_x = \mu_x L_0, L_y = \mu_y L_0$ y $L_z = \mu_z L_0$. El vector de frecuencia de corte a escala pequeña $\overrightarrow{\kappa_m}$ tiene magnitud $\kappa_m = \sqrt{\kappa_{xm}^2 + \kappa_{ym}^2 + \kappa_{zm}^2}$ y componentes $\kappa_{xm} = \frac{2\pi}{l_x}, \kappa_{ym} = \frac{2\pi}{l_y}, \kappa_{zm} = \frac{2\pi}{l_z}$ con $l_x = \mu_x l_0, l_y = \mu_y l_0$ y $l_z = \mu_z l_0$. Se empleará la aproximación de Markov, lo que implica que las fluctuaciones en el índice de refracción en cualquier par de puntos a lo largo de la dirección de propagación están correlacionadas. Entonces, el espectro de potencia es,

$$\Phi_n(\kappa_x, \kappa_y, 0) = \frac{(2\pi)^3 \sigma_n^2 \mu_x \mu_y \mu_z e^{-\frac{\kappa_x^2}{\kappa_{mx}^2} - \frac{\kappa_y^2}{\kappa_{my}^2}}}{\kappa_0^{3-\alpha}\left[\kappa_0^2 + 4\pi^2\left(\mu_x^2\kappa_x^2 + \mu_y^2\kappa_y^2\right)\right]^{\frac{\alpha}{2}}}, \quad \mathbf{3 < \alpha < 4}. \quad (9)$$

En la Figura 1 el espectro de potencia de un típico tejido biológico anisotrópico es representado a lo largo de las direcciones x e y. Los parámetros del tejido biológico se han seleccionado de la siguiente manera: $\sigma_n = 2 \times 10^{-2}, L_0 = 5\mu m, l_0 = 0.2\mu m, \mu_x = 1, \mu_y = 3, \mu_z = 1, \alpha = 3.5$. Se observa que el espectro de potencia es mayor a lo largo de la dirección x que en la dirección y.

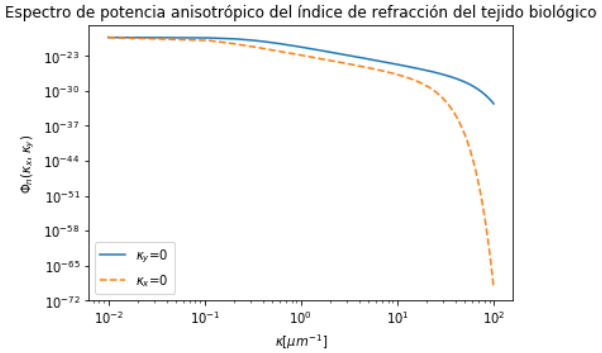


Figura 1.*Espectro de potencia anisotrópico del índice de refracción a lo largo de las direcciones x e y.*

Para un medio turbulento anisotrópico y el vector con coordenadas cartesianas $\vec{\rho} = (\xi, \eta)$, entonces, la densidad espectral $S(\vec{\rho}, z)$ está dada por,

$$S(\vec{\rho}, z) = \frac{1}{\sqrt{\Delta_x(z)\Delta_y(z)}} e^{-\frac{\xi^2}{2\sigma_0^2\Delta_x(z)}} e^{-\frac{\eta^2}{2\sigma_0^2\Delta_y(z)}} \quad (10)$$

Donde σ_0 denota la media cuadrática de la anchura del haz de la fuente transversal y,

$$\Delta_x(z) = 1 + \left[\frac{1}{4k^2\sigma_0^4} + \frac{1}{k^2\sigma_0^2}\left(\frac{1}{\delta_0^2} + \frac{2\pi^2 k^2 Tz}{3\mu_x^2}\right)\right] z^2 \quad (11)$$

$$\Delta_y(z) = 1 + \left[\frac{1}{4k^2\sigma_0^4} + \frac{1}{k^2\sigma_0^2}\left(\frac{1}{\delta_0^2} + \frac{2\pi^2 k^2 Tz}{3\mu_y^2}\right)\right] z^2 \quad (12)$$

Con δ_0 como la media cuadrática de la anchura de coherencia de la fuente transversal, $k = 2\pi/\lambda$ como el número de onda angular y,

$$T = \frac{\mu_z \sigma_n^2 \kappa_0}{4\pi(\alpha-2)}\left[\left(+4\pi^2(-2)\left(\frac{\kappa_m}{\kappa_0}\right)^2 (2\pi)^{2-\alpha}\left(\frac{\kappa_m}{\kappa_0}\right)^{2-\alpha} e^{\frac{\kappa_0^2}{4\pi^2\kappa_m^2}} \Gamma\left(2 - \frac{\alpha}{2}, \left(\frac{2\pi\kappa_m}{\kappa_0}\right)^{-2}\right)\right) - 2\right] \quad (13)$$

En donde $\Gamma(.,.)$ es la función Gamma incompleta. Además, el valor $1/e$ de S de los haces MGS se usan convencionalmente para los cálculos de las medias

cuadráticas de las anchuras de estas distribuciones. En particular, a lo largo de las direcciones x e y,

$$\sigma_i(z) = \sigma_0\sqrt{\Delta_i(z)} = \sqrt{\sigma_0^2 + \frac{z^2}{4k^2\sigma_0^2} + \frac{z^2}{k^2\delta_0^2} + \frac{2\pi^2k^2Tz^3}{3\mu_i^2}}, \quad (\boldsymbol{i} = \boldsymbol{x}, \boldsymbol{y}) \tag{14}$$

$$\delta_i(z) = \left[\frac{1}{\delta_0^2\Delta_i(z)} + \frac{2\pi^2k^2Tz}{3\mu_i^2} + \frac{2\pi^2k^2Tz}{3\mu_i^2\Delta_i(z)}\left(2 - \frac{\pi^2Tz^3}{6\mu_i^2\sigma_0^2}\right)\right]^{-\frac{1}{2}}, \quad (i = x, y) \tag{15}$$

4. Obtención de parámetros biológicos del modelo de propagación de la luz

Los parámetros biológicos del modelo que anteriormente se ha presentado son: la escala exterior L_0, la varianza del índice de refracción σ_n^2 y los coeficientes de resistencia anisotrópica en cada dirección μ_x, μ_y y μ_z.

Para estudiar sus características cuantitativamente de la escala externa L_0, se ajusta el espectro de inhomogeneidades de índice a la ecuación, [6],

$$\Phi(\kappa) = \frac{4\pi\sigma_n^2L_0^2(m-1)}{\left(1+\kappa^2L_0^2\right)^m} \tag{16}$$

que tiene la misma forma que el espectro de Von Karman, excepto que el exponente m puede asumir valores distintos de 4/3 (el exponente que describe las fluctuaciones bidimensionales en la turbulencia clásica en el régimen de inercia). Aquí m es aproximadamente igual a la mitad de la pendiente medida en el rango de escala de la ley de potencial. Su magnitud está relacionada con la dimensión fractal d_f de una superficie bidimensional,

$$d_f = 4 - m \tag{17}$$

La dimensión fractal de una imagen puede estimarse mediante diversas técnicas: (a) contaje de cajas; (b) correlación; (c) sandbox; (d) espectro de Fourier; y (e) otros.

El método del contaje de cajas consiste en una cuadrícula con cajas de tamaño ε superpuestas en una imagen, el número de cajas que contienen cualquier parte de la figura se registra como $N(\varepsilon)$. El procedimiento se repite para diferentes tamaños de ε, y la dimensión D_{box} es,

$$D_{box} = -\frac{\log[N(\varepsilon)]}{\log(\varepsilon)} \tag{18}$$

Entonces,

$$\log[N(\varepsilon)] = -D_{box}\log(\varepsilon) \tag{19}$$

D_{box} se calcula a partir del valor negativo de la pendiente de la regresión lineal de $\log[N(\varepsilon)]$ en $\log(\varepsilon)$.

El contaje de cajas se puede extender a otras dimensiones lo que implica cubrir un fractal con una cuadrícula de cajas n-dimensionales o hipercubos con una longitud lateral δ y contar el número de cajas no vacías $N(\delta)$. Para señales, la cuadrícula es de cuadrados y, para imágenes, una cuadrícula de cubos.

Se utilizan cajas de tamaños recursivamente diferentes para cubrir el fractal. Una señal de entrada con N elementos o una imagen de tamaño $N \times N$ se usa como entrada, donde N es una potencia de 2. La pendiente p obtenida en una gráfica bi-logarítmica del número de cajas usadas contra su tamaño entonces produce la dimensión fractal (también conocida como la dimensión Minkowski) donde $D = -p$.

Las divisiones sucesivas por un factor de dos se utilizan para que el tamaño de la caja proporcione un espaciado regular en la gráfica bi-logarítmica y el ajuste de mínimos cuadrados. En la práctica, generalmente se aplica una cuadrícula regular a los datos y se cuentan los cuadros no vacíos. En la Figura 2 se observa el cálculo de la dimensión fractal de una imagen de un tejido mediante el método del contaje cajas.

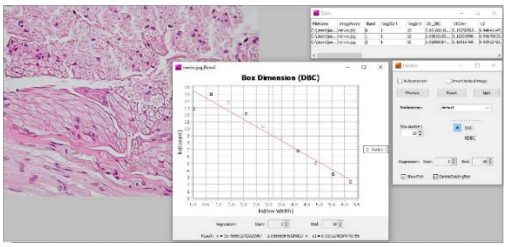


Figura 2.*Cálculo de la dimensión fractal de una imagen de un tejido mediante el método del contaje de cajas. En esta figura se muestra una gráfica logarítmica de una recta cuya pendiente es la dimensión fractal de la imagen del tejido nervioso introducido al sistema.*

Una vez se obtiene la dimensión fractal d_f y por tanto $m = 4 - d_f$, se define la densidad espectral de potencia normalizada,

$$\Phi_{nor}(\kappa) = \frac{\Phi(\kappa)}{\Phi(0)} = \frac{1}{\left(1+\kappa^2L_0^2\right)^m} \tag{20}$$

Se aproxima por la serie binomial por Taylor hasta solamente el cuarto grado:

$$\Phi_{nor}(\kappa) = 1 - mL_0^2\kappa^2 + \frac{(m^2+m)}{2}L_0^4\kappa^4 \tag{21}$$

Se realiza un ajuste de mínimos cuadrados de la densidad espectral de potencia normalizada con un polinomio de la siguiente forma:

$$f(\kappa) = 1 + b\kappa^2 + c\kappa^4 \tag{22}$$

Por tanto, a partir de los valores de los coeficientes b y c, se obtiene el valor de la escala exterior L_0. En la Figura 3 se representa el ajuste de la curva de la densidad espectral de potencia normalizada para calcular el valor de la escala exterior L_0.

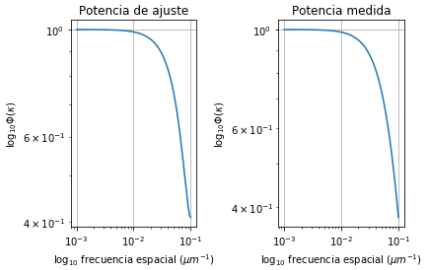


Figura 3.*Ajuste de la curva de la densidad espectral de potencia normalizada medida.*

La distribución del índice de refracción de las células y los tejidos gobierna su interacción con la luz y puede informar sobre las modificaciones morfológicas asociadas con la enfermedad. A través de mediciones basadas en la intensidad, la información del índice de refracción solo se puede extraer a través de modelos de esparcimiento que se aproximan a la propagación de la luz. Las propiedades ópticas de los diferentes bio-tejidos están determinadas por las variaciones del índice de refracción local, típicamente en el rango de 0.04 - 0.01 con el índice de fondo de ≈1.34 para fluido extracelular y ≈1.36 para citoplasma.

Teniendo en cuenta la ecuación de Helmholtz,

$$\nabla^2 U(\vec{r},\omega) + n^2\beta_0^2 U(\vec{r},\omega) = 0 \qquad (23)$$

Se obtiene la relación de dispersión estadística para un campo en medio de esparcimiento débil, [7],

$$\langle k^2\rangle = n_0^2\beta_0^2\left(1+\frac{\sigma_n^2}{n_0^2}\right) \qquad (24)$$

Donde $\langle k^2\rangle$ es el momento de segundo orden del vector de onda, $\langle k^2\rangle = \langle k_x^2\rangle + \langle k_y^2\rangle + \langle k_z^2\rangle$, los corchetes angulares indican el promedio del conjunto, n_0 es el índice de refracción promedio, β_0 es el número de onda en el vacío, $\beta_0 = \omega/c$, y σ_n^2 es la varianza del índice de refracción en el volumen 3D. Entonces, la varianza del índice de refracción asociado con una muestra transparente se puede recuperar de la imagen de fase medida como,

$$\sigma_n^2 = \frac{1}{\beta_0^2}\langle|\nabla\phi|^2\rangle \qquad (25)$$

Donde ϕ representa el cambio de fase medible de la imagen. La anterior ecuación muestra que se puede extraer la variación del índice de refracción de la intensidad media del gradiente de las imágenes de fase medidas. En la Figura 4 se presenta el gradiente de la imagen de un tejido para el cálculo de la varianza del índice de refracción.

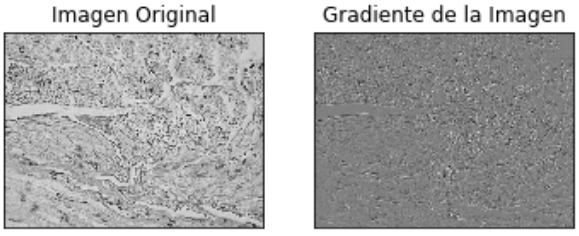


Figura 4.*Gradiente de la imagen de un tejido.*

5. Resultados y discusión

En la Figura 5 se exploran los efectos de los parámetros del bio-tejido en la evolución de las medias cuadráticas de las anchuras del haz σ_i $(i = x, y)$. Se observa que cuando σ_n^2 aumenta, entonces σ_i también aumenta. Pero cuando cualquiera de los siguientes parámetros α, L_0 o l_0 aumenta, entonces σ_i disminuye.

Como conclusiones, se ha observado la influencia de una distribución particular de esparcidores puede ser analizada por las funciones de Green. Se ha proporcionado una variedad de ejemplos numéricos que ilustran los efectos de la fuente y los parámetros del tejido en la evolución del haz. Se ha expuesto que para que el modelo matemático de la propagación de la luz tenga una mayor precisión y así poder modelar los diferentes fenómenos ópticos, como el esparcimiento o la coherencia, se debe obtener una mayor precisión cuantitativa de los parámetros de los tejidos biológicos. Los resultados de este estudio pueden ser útiles en el diagnóstico médico y el tratamiento de bio-tejidos anisotrópicos mediante radiación óptica.

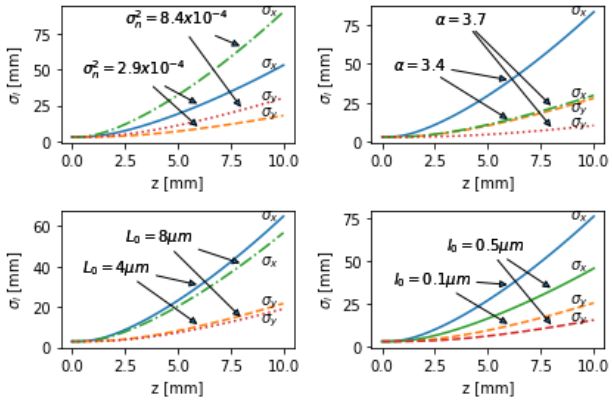


Figura 5.*Medias cuadráticas de las anchuras σ_i $(i = x, y)$ para diferentes parámetros de los tejidos.*

Agradecimientos

Este Trabajo ha sido parcialmente financiado por el proyecto del Plan Nacional de I+D+i "Modificación mediante alta presión de las propiedades plasmónicas y luminiscentes de nanopartículas de metales y óxidos desnudas y recubiertas" (PGC2018-101464-B-I00), del Ministerio de Ciencia, Innovación y Universidades, cofinanciado con fondos FEDER y por la Fundación San Cándido.

Referencias

[1] Brezinski, M. E., Tearney, G. J., Bouma, B., Boppart, S. A., Pitris, C., Southern, J. F. and Fujimoto, J. G., "Optical biopsy with optical coherence tomography," Ann. Ny. Acad. Sci. 838(1), 68-74 (1998).

[2] Arce-Diego, JL, Fanjul-Velez, F., Ganoza-Quintana, JL. Optical coherence propagation in biological tissues with significant scattering by Green's functions. Frontiers in Optics / Laser Science. OSA Technical Digest (Optical Society of America, 2018), paper JTu3A.106.

[3] Korotkova O. Random Light Beams: Theory and Applications. CRC Press, 2013 (ISBN: 1439819505).

[4] Andrews LC, Phillips RL. Laser beam propagation through random media. SPIE Press, 2005 (ISBN: 0819459488).

[5] Chen X, Korotkova O. Optical beam propagation in soft anisotropic biological tissues. OSA Continuum Vol. 1, Issue 3, pp. 1055-1067 (2018)

[6] Schmitt J. M., Kumar, G., Turbulent nature of refractive-index variations in biological tissue, OSA Publishing, Optics Letters, Vol. 21, Issue 16, pp. 1310-1312 (1996)

[7] Shan M., Kandel M. E., Popescu G., Refractive index variance of cells and tissues measured by quantitative phase imaging

Computational Simulations of Atrial Electrophysiology Remodelling under Metabolic Syndrome Driven by Up-regulation of Potassium Repolarizing Ionic Currents

A.J. Rodríguez Almeida[1], A.Bizy[2], M. Zarzoso-Muñoz[1], C.J. Calvo[1,3]

[1]Joint Unit Electrophysiology and Bioengineering, Dts. Physiology and Physiotherapy, Facultat de Medicina i Odontología, Universitat de València – Universitat Politècnica de València, València, España; {anrodal7@etsii.upv.es , conrado.calvo@uv.es}

[2]Department of Biomedical Sciences (F.Health Sciences), U. CEU Cardenal Herrera, Moncada, Valencia, España;

[3] CIBERcv Enfermedades Cardiovasculares, Valencia, España

Abstract

Prevalence of metabolic disorders such as diabetes mellitus, obesity, or metabolic syndrome has not stopped increasing during the last years in developed countries. These alterations have been experimentally and clinically associated with the development and enhancement of cardiovascular diseases, like heart failure, atrial and ventricular arrhythmias, or sudden cardiac death. Specifically, MS can be defined as the presence of central obesity together with two or more of the following factors: triglycerides increase, decrease of high-density lipoprotein cholesterol (CHDL), blood pressure increase and/or blood glucose increase. This disorder induces a structural, functional, and electrophysiological cardiac remodelling that is related to cardiac dysfunction. Despite lacking extensive experimental data prior studies have proven that, at a cellular level, MS and obesity may mainly result in up-regulation and modification of the biophysics of specific repolarizing potassium channels. Therefore, electrophysiological modifications by MS remodelling may involve altered activation-repolarization and restitution dynamics in the atrium, enhancing the propensity of appearance of automatic activity at high frequency and the sudden onset of atrial fibrillation (AF). We surmise that electrical remodelling observed in MS driven by combined potassium channels up-regulation will modify restitution dynamics in a frequency dependent manner and may suffice for the appearance of automatic activity, increasing the vulnerability for triggering AF. In this work, a computational simulation study is carried out based on Koivumäki model of atrial electrophysiology and prior knowledge, to study the effects that such electrical remodelling may have on atrial restitution dynamics. Implications under MS+AF remodelling are also explored. We conclude that modifications in atrial electrophysiology due to up-regulation of repolarizing currents under MS may partially explain an increased susceptibility for AF development.

1. Introduction

Cardiovascular diseases are the leading cause of death worldwide according to the World Health Organization, accounting for about 31% (17.9 million/year). Prevalence of metabolic disorders such as diabetes mellitus, obesity, or metabolic syndrome (MS) has not stopped increasing during the last years in developed countries [1]. These alterations have been experimentally and clinically associated to the enhancement of cardiovascular diseases, like heart failure or atrial and ventricular arrythmias that, under aggravated conditions could result in sudden cardiac death [2]. In fact, it has been reported that these alterations could increase by two-fold the probabilities of suffering from cardiovascular diseases [3].

MS can be defined as the presence of central obesity together with two or more of the following factors: triglycerides increase, decrease of high density lipoprotein cholesterol (CHDL), blood pressure increase and/or blood glucose increase [4]. MS has been clinically and experimentally associated to structural, functional and electrophysiological cardiac remodelling. Left ventricle hypertrophy, appearance of fibrosis on the left atrium, prolonged QT interval dispersion and ventricular repolarization alterations are amongst the main modifications induced in the heart under MS [5-7].

However, underlying electrophysiological echanisms that associate this remodelling with the propensity of suffering from cardiovascular disease and cardiac arrhythmias are poorly understood. Up-regulation of repolarizing potassium currents in cardiac cells of MS-induced heart animal models has been experimentally observed. Altogether, prior knowledge suggests that MS remodelling may have an effect on activation-repolarization and restitution dynamics in the atrium, enhancing the propensity of appearance of high frequency automatic activity and the onset of atrial arrhythmias, for instance, Atrial Fibrillation (AF) [8].

In this work, we seek to predict and analyse in detail the frequency-dependent implications of the remodelling associated to MS for atrial rhythm disorders, at cellular level, by means of predictive computational modelling and simulations.

2. Methodology

2.1. Computational simulations of atrial physiology

Human atrial electrophysiology was studied following Koivumäki-2014 model under control and AF conditions, widely validated in the literature [9]. All carried simulations at cellular level were solved using an implicit method for stiff differential equation system arrays using as initial conditions steady-state values, under each specific condition, at 1Hz. Automatic features calculation, restitution analyses and extraction of sensibility and

ISBN: 978-84-09-25491-0, pags. 298- 301

parametric biomarkers were performed using the same platform environment.

2.2. Model validation and sensitivity analysis

Basic AP features, ionic currents, long-term stability, and ability to mimic the Atrial Fibrillation (cAF) were evaluated. The AP can be defined as the depolarization of the cell that allows the propagation of the electrical stimulus between cells. Specifically, the studied AP properties were Action Potential Duration (APD) at 90% of repolarization (APD_{90}), APD_{50} (APD at 50% of repolarization), Resting Membrane Potential (RMP), amplitude and depolarization velocity. Simulations at cellular level were carried by pacing the model during 20-50 seconds with a fixed Pacing Cycle Length (PCL) protocol at 1Hz to achieve steady-state. The PCL is the time between two consecutive stimulus. Action potential amplitude, RMP and APD long-stability was assessed. In these simulations, the model was paced during 20 minutes with a BCL of 1000 ms [10].

Then, a sensitivity analysis was assessed. This technique has been proven useful in different cardiac electrophysiology simulations experiments to get in-depth knowledge about the dependence of computational model outputs with variations of its inputs parameters (i.e., ionic currents conductivities) [11],[12]. The studied parameters were $APD_{50,}$ $APD_{90,}$ RMP, AP triangularization, amplitude, depolarization velocity, intracellular calcium amplitude and excitability. The latter is a surrogate measurement; it is the maximum value of the product of two variables that represents the probability of the two sodium current inactivation gates to transit to open-state. Here, close to zero-values represent a completely close gating, whereas 1 represents a completely open-state, thus is intrinsically related to true excitability and refractoriness.

The modified variables in this analysis are the main current conductivites: $I_{Ca,L}$, I_{Kr}, I_{Ks}, I_{Kur}, I_{to}, I_{Na}, I_{K1}. All their conductivities have been multiplied by a varying factor from 0.25 to 2, with fixed steps equals to 25%. After this, linear regression was applied for quantification for every peer variable-parameter.

2.3. Analysis of influence of repolarizing currents on atrial restitution

MS remodelling has been shown to be mainly driven by up-regulation of different repolarizing potassium currents. Thus, the analysis of the individual frequency-dependent modulation of these currents in atrial electrophysiology is assessed.

With this purpose, I_{Kr}, I_{Ks}, I_{to} and I_{K1} conductivities have been modified from its 25% of initial value until the 300%, with steps of 25% of increment on them. Under these conditions, Koivumäki model has been paced during 50 seconds with the following PCLs: 500, 400, 300, 250, 200, 175, 150 and 125 and 100 ms. Then, APD rate adaptation has been calculated, resulting in a set of restitution curves. Thus, individual influence, considering the main potassium repolarizing currents, woth varying changes on their conductivities, in the restitution dynamics are compared at high and low frequencies, with interest in analysing such transition and the maximal restitution slope changes.

2.4. Model adaptation under metabolic syndrome

Atrial computational model has been adapted based on the analysis of the experimental data extracted from realistic MS-induced by high-fat high-sucrose diet in a rabbit animal model [13]. Overexpression of specific potassium currents, an abbreviation of the Effective Refractory Periods (ERP) and Functional Refractory Periods (FRP) was observed. Table 1 illustrates the two different adaptations that have been simulated to evaluate the implications for atrial electrophysiology considering two different degrees of severity of MS remodelling in the atrium.

Maximum conductance	MS Severity 1	MS Severity 2
I_{Kr}	30%	50%
I_{Ks}	50%	50%
I_{to}	50%	75%
I_{Kur}	-	30%
I_{K1}	-	30%

Table 1. Modelled up-regulation of MS adaptations upon specific severity depending on its components.

Once adapted, the model was paced for 50 seconds with BCLs equals to 1000, 900, 800, 700, 600, 500, 400, 300, 250, 200, 175, 150, 125 and 100. Then, AP modifications on relevant biomarkers, restitution dynamics changes and excitability oscillations between control and MS induced conditions were analysed.

2.5. Model adaptation under metabolic syndrome condition in a situation of atrial fibrillation

A preliminary study on the relationship between MS, the increased propensity for atrial arrhythmias, and its maintenance, Koivumäki model was adapted to simulate the MS remodelling under AF conditions [9]. Thus, ionic channels conductivities were modified by combining prior knowledge about AF ionic currents remodelling on the atrium and the observed changes on MS. Table 2 shows the introduced changes for this set of simulations, only considering modifications on the maximal current conductances.

Maximum conductance	Regulation
I_{to}	-50%
I_{Ks}	350%
I_{Kur}	+55%
$I_{Ca,L}$	+26%
I_{K1}	+275%

Table 2. Introduced changes in Koivumaki model to simulate MS + AF condition.

Simulations have been carried out by pacing the model for 50 seconds with BCLs equals to 1000, 850, 750, 675, 600, 500, 450, 350, 250, 200 and 150 ms. Then, AP changes, restitution dynamics changes and excitability

variations between control and MS condition under AF situation were analysed.

3. Results

3.1. Model's validation and sensitivity analysis

Modified Koivumäki model was validated replicating atrial electrophysiology under specific conditions, which rapidly reached physiological amplitudes, RMP and APD in steady-state. Then, sensitivity analysis provided valuable insight regarding model dependence on changes on its main conductivities (not shown).

3.2. Effects of potassium currents' up regulation in repolarization restitution dynamics

Restitution curves obtained after parametrical analyses by varying different potassium currents' conductivities are illustrated in Figure 1. This could represent different degrees of remodelling, and overall, the individual contribution to frequency dependent modifications on atrial electrophysiological restitution.

At low frequencies, I_{Kr}, I_{Ks} and I_{to} show a constant variation of APD for all conductivities' variations, but I_{to} variations produce slightly greater changes on the atrial restitution curve. Changes on conductivities of this currents do not affect restitution aggressively at any frequency. In contrast, I_{K1} variations induce frequency dependent modifications in the restitution curve.

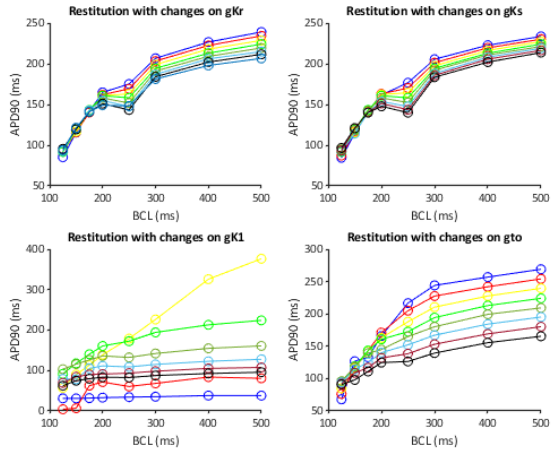


Figure 1. Restitution dynamics modulation by main potassium repolarizing currents.

Hence, this simulation study shows that I_{K1} plays a major role in the modulation of frequency-dependent restitution dynamics of the atrium, whereas I_{Kr}, I_{Ks} and I_{to} modifications contribute rapidly abbreviating the APD, and thus, shortening refractory periods.

3.3. Effects of metabolic alterations involving remodelling in potassium currents

Figure 2 shows the electrical modifications under MS on atrial electrophysiology. It can be observed that, in both MS severity models implemented, reduction of APD is constant at low frequencies. This APD shortening directly involves ERP abbreviation, which meets the reference experimental data. On the contrary, APD is similar in both MS severity types and in the control case at high frequencies. This phenomenon is also observed experimentally. Furthermore, MS excitability is greater than control at high and low frequencies.

MS severity type-2 model, which includes I_{Kur} and I_{K1} remodelling and a greater up-regulation of I_{Kr} and I_{to}, shows a greater APD shortening compared to severity type-1 implementation, as it was expected from the analysis described in the previous section. Also, its excitability is greater than in MS severity type-1.

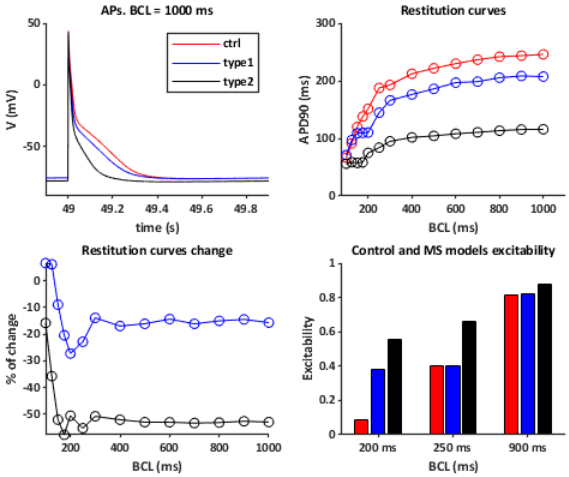


Figure 2. Effect of remodelling of MS in atrial electrophysiology. Top, left: APD; Top, right: Restitution curves; Bottom, left: Restitution curve percentage of change; Bottom, right: Excitability.

3.4. Effects of fast atrial pacing and atrial fibrillation remodelling

Figure 3 illustrates the electrophysiological modifications under MS and AF simultaneous remodelling.

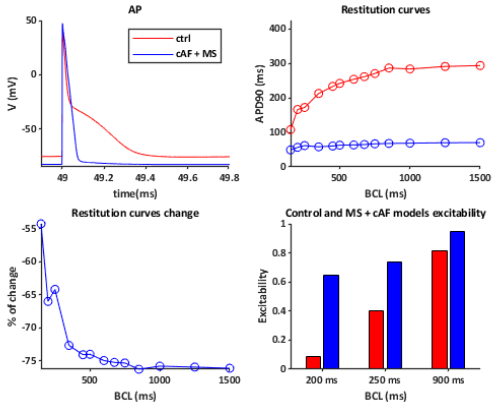


Figure 3. Effects of remodelling of MS under cAF condition. A: APD; B: Restitution curves; C: Restitution curve percentage of change; D: Excitability.

Firstly, a drastic abbreviation of APD is observed under this condition. Restitution curve is flattened at all frequencies. Despite this event occurs more drastically at low frequencies, at high frequencies the APD shortening exceeds 50% of its control value, which is a significant decrease. These events involve a serious decrease of refractory period at all frequencies, which could potentially enhance propensity to produce rhythm disorders in the atrium.

APD restitution curve is aggressively flattened at all frequencies. MS remodelling under AF conditions induces a practically flat restitution curve.

Excitability of the remodelled cell is greater than the control at all frequencies. As it is expected, this excitability decreases with frequency in both cases. However, reduction of this parameter is noticeably lower at high frequencies under MS combined to AF remodelling.

Therefore, under AF remodelling MS produces a remarkable and drastic APD abbreviation, and more aggressive flattening of the restitution curve together with greater excitability being able to harbour high-frequency automaticity and fibrillatory activity. Therefore, this suggests that MS remodelling in a AF substrate could potentially increase the risk for AF acceleration and maintenance.

4. Conclusions

We demonstrated the individual implications of potassium currents up-regulation on atrial electrophysiology adaptation. This simulation study demonstrates that MS remodelling by means of up-regulation of potassium repolarizing currents, based on realistic experimental data, induce significant changes in atrial electrophysiology and its dynamic adaptation. The most remarkable changes are APD abbreviation at low and medium frequencies and the increase of excitability at all frequencies. The former involves the shortening of the ERP, which would make atrial tissue more vulnerable to fast activation. Furthermore, I_{K1} has demonstrated to play a key role in the onset and maintenance of AF. Up-regulation of this current produces more drastic frequency-dependent modifications on atrial repolarization and restitution. This rate dependent adaptation under MS remodelling could potentially increase the propensity to create re-entrant circuits that could evolve into sustained arrhythmias.

MS remodelling under AF conditions increased the risk for shortened refractory periods able to harbour long-term AF. APD and refractoriness are considerably lowered. Hence, simulations results suggest that MS remodelling under AF conditions could potentially provide a functional substrate to increase the propensity of the onset of rhythm disorders and to maintain the occurrence of AF through the creation of more complex re-entrant circuits.

In conclusion, results obtained in this preliminary simulation work show association between MS remodelling and atrial repolarization and restitution modifications related to AF risk. However, further studies should be accomplished to gain a more detailed insight regarding this matter. It is critical to use enriched realistic experimental date and current stimulation protocols to reliably reproduce the dynamic modifications and adaptation observed under MS predicting in tissue and organ levels its implications for atrial remodelling and risk.

References

[1] M. G. Saklayen, "The Global Epidemic of the Metabolic Syndrome," *Curr. Hypertens. Rep.*, vol. 20, no. 2, pp. 1–8, 2018, doi: 10.1007/s11906-018-0812-z.

[2] J. Duflou, R. Virmani, I. Rabin, A. Burke, A. Farb, and J. Smialek, "Sudden death as a result of heart disease in morbid obesity," *Am. Heart J.*, vol. 130, no. 2, pp. 306–313, 1995, doi: 10.1016/0002-8703(95)90445-X.

[3] S. M. Grundy, "Pre-diabetes, metabolic syndrome, and cardiovascular risk," *J. Am. Coll. Cardiol.*, vol. 59, no. 7, pp. 635–643, 2012, doi: 10.1016/j.jacc.2011.08.080.

[4] G. Alberti, P. Z. Zimmet, and J. Shaw, "Metabolic syndrome-a new world-wide definition. A Consensus Statement from the International Diabetes Federation," *Diabet. Med.*, vol. 23, pp. 469–480, Jun. 2006, doi: 10.1111/j.1464-5491.2006.01858.x.

[5] E. D. Abel, S. E. Litwin, and G. Sweeney, "Cardiac remodeling in obesity," *Physiol. Rev.*, vol. 88, no. 2, pp. 389–419, Apr. 2008, doi: 10.1152/physrev.00017.2007.

[6] K. K. Kadappu *et al.*, "Changes in left atrial volume in diabetes mellitus: more than diastolic dysfunction?," *Eur. Heart J. Cardiovasc. Imaging*, vol. 13, no. 12, pp. 1016–1023, Dec. 2012, doi: 10.1093/ehjci/jes084.

[7] M. A. Fraley, J. A. Birchem, N. Senkottaiyan, and M. A. Alpert, "Obesity and the electrocardiogram.," *Obes. Rev. an Off. J. Int. Assoc. Study Obes.*, vol. 6, no. 4, pp. 275–281, Nov. 2005, doi: 10.1111/j.1467-789X.2005.00199.x.

[8] U. Ravens and E. Cerbai, "Role of potassium currents in cardiac arrhythmias," *Europace*, vol. 10, no. 10, pp. 1133–1137, 2008, doi: 10.1093/europace/eun193.

[9] J. T. Koivumäki, G. Seemann, M. M. Maleckar, and P. Tavi, "In Silico Screening of the Key Cellular Remodeling Targets in Chronic Atrial Fibrillation," *PLoS Comput. Biol.*, vol. 10, no. 5, 2014, doi: 10.1371/journal.pcbi.1003620.

[10] M. Wilhelms, H. Hettmann, M. M. Maleckar, J. T. Koivumäki, O. Dössel, and G. Seemann, "Benchmarking electrophysiological models of human atrial myocytes," *Front. Physiol.*, vol. 3 JAN, no. August 2014, 2013, doi: 10.3389/fphys.2012.00487.

[11] Y. S. Lee *et al.*, "The contribution of ionic currents to rate-dependent action potential duration and pattern of reentry in a mathematical model of human atrial fibrillation," *PLoS One*, vol. 11, no. 3, pp. 1–17, 2016, doi: 10.1371/journal.pone.0150779.

[12] E. T. Y. Chang, M. Strong, and R. H. Clayton, "Bayesian sensitivity analysis of a cardiac cell model using a Gaussian process emulator," *PLoS One*, vol. 10, no. 6, pp. 1–20, 2015, doi: 10.1371/journal.pone.0130252.

[13] Calvo C. J. *et al.*, "Modifications of short-term intrinsic pacemaker variability in diet-induced metabolic syndrome: a study on isolated rabbit heart," *J. Physiol. Biochem.*, vol. 75, no. 2, pp. 173–183, 2019, doi: 10.1007/s13105-019-00667-7.

Prototipo de modelo anatómico impreso en 3d para la exploración del oído humano.

L. Ramírez Palomino[1], J. Vásquez López [2]

[1]Laboratorio de Impresión 3D, FabLab Málaga, España, linarp3d@gmail.com

[2] Grupo GREMO, Departamento de Morfología, Universidad del Cauca, Popayán, Colombia, jalfonvasquez@unicauca.edu.co

Resumen

En este artículo, se da a conocer el desarrollo de un prototipo de herramienta de exploración anatómica del oído humano, el cual fue realizado en el laboratorio de impresión 3D FabLab Málaga-España. En el desarrollo del proyecto fue indispensable el uso de la impresión 3d para la fabricación de los modelos de los huesos del oído medio (Martillo, yunque y estribo) y del oído interno, emulando las características anatómicas del oído. Se obtuvieron los modelos de estas regiones, conservándose con precisión los detalles anatómicos y morfométricos de estas estructuras.

Es importante resaltar el potencial de las tecnologías de impresión 3d para la fabricación de prototipos en el área médica, por ser un medio adecuado, para la obtención de modelos anatómicos de regiones corporales de difícil abordaje quirúrgico o exploratorio, como el caso del oído medio y del oído interno.

1. Introducción

Las impresoras 3d son dispositivos tecnológicos, capaces de generar un objeto solido a partir de un modelo CAD, y por su fácil uso, se han convertido en parte fundamental para desarrollar trabajos en distintas áreas. Impresión 3d, es un término general que abarca un amplio grupo de procesos y materiales disponibles para las tecnologías 3d, tales como: a) Proceso de impresión 3d de extrusión de materiales, cuya tecnología más común es el modelado de deposición fundida (FDM), el cual consiste en la adaptación de un carrete de material termoplástico a la impresora 3d, que expulsa el material por una boquilla que lo funde, para finalmente, depositarlo en la plancha de construcción de la impresora, siguiendo una ruta determinada para obtener la forma sólida del objeto deseado; b) Polimerización de VAT, donde se aplica la tecnología 3d de estereolitografía (SLA), basado en un proceso llamado fotopolimerizacion, que a partir de resinas sensibles, se seleccionan capas ultra finas de ellas y un láser ultravioleta va solidificando el material hasta la formación del modelo 3d. Además de los anteriores, en la industria existen otros procesos comunes, como: Fusión en polvo (Polímeros), Material Jetting, binder Jetting y Fusión de cama en polvo (Metales). [1]

Es importante resaltar, que el campo de la medicina es uno de los más avanzados en cuanto al uso de las impresoras 3d, debido al desarrollo de múltiples investigaciones y proyectos, como la fabricación de: prótesis personalizadas; prototipos; creación de modelos anatómicos para la enseñanza y la planeación quirúrgica a partir de modelos obtenidos mediante imágenes médicas [2].

En el área de la educación médica, la impresión 3d ha facilitado la manera de acceder a conocimientos anatómicos, permitiendo la reproducción y/o copia exacta de órganos y regiones del cuerpo humano, llamados modelos anatómicos. Los modelos anatómicos, son estructuras tridimensionales con gran semejanza a la anatomía humana. Estos, se presentan en tamaño real o ampliado, convirtiéndose en herramientas que permiten alcanzar mejores resultados en el aprendizaje de la anatomía humana, que tiene un alto grado de dificultad, tanto para estudiantes universitarios como para los médicos en adiestramiento y cirujanos.[4,3] El grado de dificultad es mayor, si los puntos o zonas de interés son complejos, como por ejemplo el abordaje exploratorio del oído medio o del interno, puesto que su estructura anatómica es muy pequeña, no siendo fácilmente observable [5]. Actualmente, se dispone de modelos anatómicos del oído, pero con grandes limitaciones anatómicas.

El oído es el órgano encargado de conducir la información sonora, procedente del exterior hasta el cerebro, se localiza en el hueso temporal y está dividido en tres regiones anatómicas: a) Oído externo, b) Oído medio y c) Oído interno. [5]

El oído externo formado por el pabellón auricular y el conducto auditivo externo (CAE), permite focalizar las ondas sonoras para conducirlas a la membrana timpánica o tímpano, que vibra ante ellas, esta membrana es el límite con el oído medio. En el oído medio se encuentra el sistema de huesecillos o también llamada cadena osicular, conformado por el martillo, yunque y el estribo, cuya función es trasmitir las vibraciones a la ventana oval. Además, el oído medio por intermedio de la trompa de Eustaquio regula las presiones sobre la membrana timpánica. Por su parte, en la ventana oval el oído interno, recibe las vibraciones procedentes del estribo y las trasduce en perturbaciones de la perilinfa y la endolinfa que estimularán a las células pilosas del órgano de corti, que convierte las señales mecánicas en impulsos eléctricos, que son conducidos a la corteza cerebral para ser interpretados y percibidos como un sonido. [5,6,7].

Teniendo en cuenta que el estudio de la anatomía macroscópica y quirúrgica del oído, es compleja, debido su tamaño y a la falta de modelos anatómicos cercanos a la realidad, se propone la creación de una herramienta mediante impresión 3d, que permita el entrenamiento y la

ISBN: 978-84-09-25491-0, pags. 302- 305

adquisición de habilidades, para realizar de manera segura, idónea y realista, la exploración del oído humano, contribuyendo de esta manera a la seguridad del paciente.

2. Metodología

Este trabajo fue realizado en el laboratorio de impresión 3d FabLab Málaga, previo entrenamiento en tecnologías de punta en impresión 3D. Este proyecto fue realizado siguiendo las siguientes etapas:

2.1. Etapa Uno: Implementación del Software

Se realizó una búsqueda de los ficheros en 3D del oído medio e interno, basados en reconstrucción tridimensional de imágenes médicas, encontrados en una plataforma web libre, implementada para enseñanza anatómica del oído, desarrollado por la Universidad de McGill (Canadá) [8]. Esta reconstrucción permitió navegar virtualmente sobre las zonas anatómicas de interés, con diferentes perspectivas 3d, acción que no es posible realizarla al utilizar solo imágenes 2D. Se encontraron los tres huesecillos que conforman la cadena osicular (Figura 7), y una microsegmentación del laberinto óseo, provenientes de un lenguaje de modelado de realidad virtual, los cuales se importaron al software de Rhinoceros para exportarlos en un formato Stl, compatible con la impresora (Figura 1).

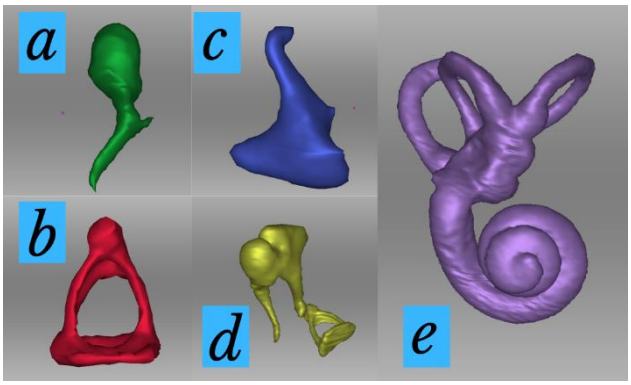


Figura 1.*Modelos stl: a) Martillo, b) Estribo, c) Yunque, d) Cadena Osicular. e) oído interno.*

Considerándose las medidas promedio del Martillo, Yunque y Estribo [9], se exportaron al software de *Rhinoceros 6* (Herramienta software de modelado 3D), los modelos en formato *stl*, con las dimensiones adecuadas. Posteriormente, se ajustaron las posiciones de los modelos en el software 3Dsprint (Software exclusivo de *3DSystem*®) con el fin de encontrar una posición adecuada de los modelos para proceder a la impresión de ellos.

2.2. Etapa Dos: Impresión 3d de los Modelos

En esta etapa se utilizó la impresora ProJet MJP 2500 de 3DSystem® (Figura 2 a), que permite la impresión de piezas con alta calidad, por su precisión con el modelo CAD y por su velocidad de impresión; tres veces mayor, al ser comparadas con máquinas similares.
El material utilizado para la impresión de estos modelos fue VisiJet M2R-WT® (MJP) (Figura 2 b), material de color blanco, con aspecto y tacto plástico, ideal para objetos que requieran ser utilizados para función mecánica rígida y/o flexión moderada. Este material es certificado por la United States Pharmacopeia (USP) como clase VI, el cual tiene biocompatibilidad ISO 10993, por lo que su uso se recomienda para aplicaciones médicas tales, como guías quirúrgicas y/o prototipos de modelos anatómicos para simulaciones.

2.3. Etapa Tres: Limpieza de piezas.

Al obtener las piezas mediante la impresión 3d, se les realizó un proceso de limpieza, debido a que la impresión genera un soporte de VisiJet M2 Sup (Figura 2 d), que es un material similar a la cera. Para el procedimiento de limpieza, se elimina la cera mediante un sistema llamado Easy Clean de 3DSystem®(Figura 2 c), que disuelve los residuos del soporte en aceite mineral, obteniéndose una pieza totalmente limpia.



Figura 2. *3DSystem®, a) ProJet MJP 2500, b) VisiJet M2R-WT® (MJP) c) Easy Clean d) VisiJet M2 Sup.*

2.4. Etapa Cuatro: Estructura del prototipo

Finalizada las impresiones y su procedimiento de limpieza, se continúa a realizar la estructura donde reposaran los modelos de los huesos del oído medio y el oído interno, con el fin de explorar su anatomía. Se utilizó un molde de silicona del pabellón auricular derecho, como simulación del oído externo y se adaptó un mecanismo en el interior, al final del CAE para simular el tímpano y la cavidad timpánica. (Ver figura 3) Anatómicamente, la membrana timpánica limita el acceso al oído medio y en razón a ello, se optó por desarrollar un mecanismo de "tímpano móvil", para así, dar paso a la cámara de exploración del oído medio. En la exploración se utilizó Visual Ear Cleaner®, dispositivo endoscópico de exploración anatómica, que consta de: cámara con diámetro de 5,5mm; 3 MP de resolución, e iluminación mediante 6 led que permiten la visibilidad en cavidades oscuras. La imagen generada por este dispositivo endoscópico se puede observar en tiempo real en la pantalla del ordenador.

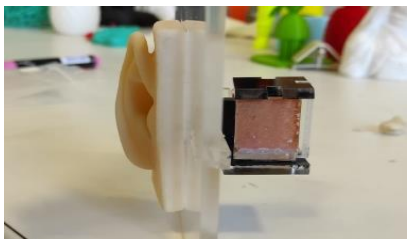


Figura 3.*Estructura de explotación.*

3. Resultados

Los archivos en formato *stl* fueron fundamentales para lograr la impresión del modelo de la cadena osicular y del oído interno, en la (Figura 4) se observa la placa de impresión al ser extraída de la máquina, cuyo tiempo de impresión fue de 4 horas. El residuo de soporte de cera, fue eliminado casi en su totalidad de las piezas que se utilizaron para el prototipo de exploración.

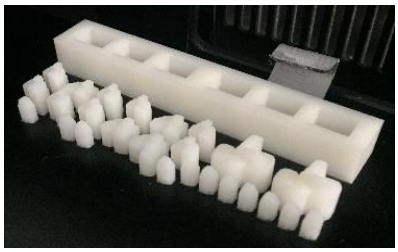


Figura 4.*Placa de impresión.*

En la tabla 1,2,3,4., se observa que las dimensiones de los huesecillos impresos, al ser comparados, con las dimensiones morfométricas del sistema de huesecillos del oído medio del humano, no tienen una diferencia estadísticamente significativa, por lo cual es posible la implementación de este prototipo, para la visualización acercada a la realidad de esta región anatómica compleja.

Dimensiones	Estribo.	Estribo impreso.	Error relativo .
Largo	3,18 mm	3mm	5,6%

Tabla 1. *Comparación dimensiones del Estribo*

Dimensiones	Yunque.	Yunque impreso.	Error relativo .
Largo	6.06 mm	6 mm	0.9%

Tabla 2. *Comparación dimensiones del yunque.*

Dimensiones	Martillo.	Martillo impreso.	Error relativo .
Largo	8 mm	7.6 mm	5%

Tabla 3. *Comparación dimensiones del Martillo.*

Dimensiones	Cadena osicular.	Cadena osicular impreso.	Error relativo .
Largo	18 mm	17 mm	5.5%

Tabla 4. *Comparación dimensiones de la cadena osicular.*

Debido al pequeño tamaño de las piezas impresas, el procedimiento de limpieza afectó algunos modelos, causando fractura o deformidades, sumado a la exposición a altas temperaturas durante este proceso. A pesar de los inconvenientes, al mejorar la técnica de este procedimiento, se logró extraer los modelos de la cadena de huesecillos en perfecto estado, conservándose los detalles anatómicos (Figura 5). En la imagen de la izquierda, se aprecia el estribo cubierto con el soporte de cera y luego del procedimiento de limpieza, se obtiene como resultado la imagen de la derecha, el mismo procedimiento fue realizado en cada uno de los modelos impresos (Figura 6,7).



Figura 5.*Estribo impreso en 3d antes y después del procedimiento de limpieza.*



Figura 6.*Superficie del oído interno impresa en 3d.*



Figura 7.*Huesos impresos en 3D.*

El ensamble de la cadena osicular y la superficie del oído interno fue desarrollado en una placa de metacrilato con dimensiones de 1,5cm x 2cm. Al oído interno le fue adicionado un efecto cromático para simularlo con la realidad anatómica, mientras que los huesecillos

continuaron con el color blanco original de la impresión, para facilitar su identificación (Figura 8)



Figura 8.*Cadena osicular ensamblada al oído interno.*

Finalizado el montaje de la estructura, como se observa en la (figura 9), se procedió a realizar la exploración anatómica, esta se dividió en 4 puntos de interés: pabellón auricular; membrana timpánica; la cadena osicular y en nuestro caso, el punto final de interés, fue posible apreciar los detalles anatómicos externos del oído interno: con la cóclea y los canales semicirculares.



Figura 9.*Herramienta de exploración.*

Al hacer la intervención exploratoria en el modelo anatómico, se realizó la comparación con los cuatro puntos de interés del oído de un ser humano (Figura 10). Las imágenes de la izquierda a, b, c, corresponden a imagen reales *(in vivo)*, adquiridas mediante otoscopia, mientras que las imágenes de la derecha d, f, g., corresponden a imágenes de los modelos impresos, adquiridas con la cámara del explorador anatómico (Visual Ear Cleaner®).

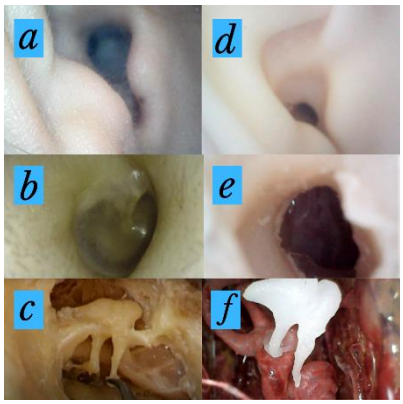


Figura 10. *Comparación de los puntos de interés de la exploración anatómico Reales y del prototipo*

4. Conclusión.

Mediante la utilización de técnicas de impresión 3D, se logró la obtención de modelos del sistema de huesecillos del oído y la superficie del oído interno, conservándose con precisión los detalles anatómicos y morfométricos de estas estructuras. El desarrollo de estos modelos, evidencia que es posible aseverar que la impresión 3D puede seguir utilizándose en la fabricación de modelos anatómicos que ofrezcan dificultad, como es el caso de la compleja anatomía del oído. De acuerdo con los resultados obtenidos, es posible que como trabajo futuro, este prototipo pueda ser utilizado en el entrenamiento de médicos residentes de otorrinolaringología y para estudiantes de pregrado. Así como también, se pueda convertir en un simulador quirúrgico de entrenamiento, para facilitar el proceso de adquisición de destrezas requeridas en intervenciones quirúrgicas del oído, como: Estapedectomia o Timpanoplastia entre otras.

Agradecimientos

Los autores expresan su agradecimiento a Angelo Di Placido y Fernando García del FabLab Málaga, por su apoyo incondicional para la realización del presente trabajo. De igual manera agradecen al Departamento de Morfología de la Facultad de Ciencias de la Salud de la Universidad del Cauca-Colombia y especialmente al docente Oscar H. Ríos Ramírez, por facilitar el acceso al material óseo del oído.

Referencias

[1] Gomez S., Impression 3D. Vol 2. 2020 (ISBN-13:978-8426727121).

[2] Bucco M., La impresión 3d y sus aplicaciones en servicios médicos (prótesis, fármacos, órganos) trabajo *de fin de master, Universidad de San Andrés, Ciudad* Autónoma de Buenos Aires, junio 2016.

[3] Collipal Larre, E, & Silva Mella, H., (2011). Estudio de la Anatomía en Cadáver y Modelos Anatómicos: Impresión de los Estudiantes. *International Journal of Morphology*, *29*(4),11811185. https://dx.doi.org/10.4067/S0717-95022011000400018.

[4] Inzunza, O, Caro I, Mondragón G, Baeza F, Burdiles A. & Salgado, G., Impresiones 3d, nueva tecnología que apoya la docencia anatómica. Int. J. Morphol., 33(3):1176-1182, 2015.

[5] Basterra J., Otorrinolaringología y patología cervicofacial: Manual Ilustrado. Vol 2. 2015 (ISBN: 9788445825969).

[6] Moore KL, Dalley II AF, Agur AM., Fundamentos de anatomia con orientación clinica., Editorial Wolters Kluwer, Vol 6. 2019 (ISBN: 9788417602512).

[7] Pawlina, W., ROSS. HISTOLOGIA: TEXTO Y ATLAS Editorial Wolters Kluwer, Vol 8. 2020 (ISBN:9788417602659).

[8] Audilab.bmed.mcgill.ca. 2020. 3D Ear - An Ear Anatomy Tutorial. [online] Available at: <http://audilab.bmed.mcgill.ca/~daren/3Dear/index.html> [Accessed 21 October 2020].

[9] Maestro M. Propuesta de un nuevo modelo de prótesis para timpanoplastia. Trabajo de fin de grado. Escuela Técnica Superior de Ingenieros Industriales UPM, Madrid, febrero 2017.

Predicción de la activación neuronal en estimulación tónica en la terapia de EME: uso de modelos 3D de médula espinal personalizados al paciente frente a modelos generalizados

C. Solanes[1], J. L. Durá[1], M. A. Canós[2], J. De Andrés[3], L. Martí-Bonmatí[4], J. Saiz[1]

[1] Ci2B, Universitat Politècnica de València, Valencia, España, carsogal@etsii.upv.es

[2] Unidad del Dolor, Hospital Universitari i Politècnic La Fe, Valencia, España, canos_marver@gva.es

[3] SART, Hospital General de Valencia, Valencia, España, deandres_jos@gva.es

[4] Departamento de Radiología, Hospital Universitari i Politècnic La Fe, Valencia, España, marti_lui@gva.es

Resumen

La terapia de Estimulación de Médula Espinal (EME) es una técnica utilizada para el tratamiento del dolor crónico. En estimulación tónica, el objetivo es estimular las fibras sensitivas Aβ que se encuentran en los cordones posteriores de la médula espinal. Esta estimulación produce una sensación de parestesia que bloquea la señal del dolor. Sin embargo, el 35% de los pacientes perciben una estimulación molesta, pues los cambios posturales, la migración de los electrodos o una incorrecta posición de los mismos produce cambios en la parestesia. En los últimos 40 años se han desarrollado modelos computacionales de médula espinal generalizados que permiten entender mejor el efecto de los parámetros de estimulación y de la geometría de los electrodos en la activación neuronal. Pero, el avance hacia una medicina personalizada obliga a desarrollar modelos computacionales personalizados al paciente para optimizar la estimulación de forma individual. Para conocer las ventajas de usar un modelo personalizado frente a uno generalizado, en este estudio se ha comparado la predicción de la activación neuronal y de los parámetros de estimulación que se obtiene de ambos modelos con un caso real. Los resultados muestran que el modelo personalizado predice de forma más aproximada la activación neuronal y, por tanto, el efecto de la estimulación en un paciente. Por tanto, el uso de modelos personalizados podría suponer en un futuro la mejora de la eficacia de la terapia, pues se podría optimizar la programación y la posición de los electrodos antes del implante del sistema de EME.

1. Introducción

La terapia de Estimulación de Médula Espinal (EME) es un tratamiento mínimamente invasivo que se utiliza para reducir la intensidad, la frecuencia y la duración del dolor que sienten los pacientes que sufren dolor crónico neuropático [1]. Esta terapia consiste en aplicar corriente eléctrica en las estructuras nerviosas de la médula espinal (las fibras de los cordones posteriores (DC, por sus siglas en inglés) y de las raíces dorsales (DR, por sus siglas en inglés)) a través de unos electrodos que se implantan en el espacio epidural y que se conectan a una batería [2].

La estimulación más usada y que mayor evidencia ha dado a la eficacia clínica de la terapia de EME es la estimulación tónica [3]. Los mecanismos de acción en los cuales se basa su efecto clínico se conocen desde 1965, cuando Melzack y Wall plantearon la "teoría de la puerta control del dolor" [4]. Esta teoría explica que, cuando se estimulan las fibras aferentes sensitivas Aβ grandes, hay una "puerta" que se cierra impidiendo que las señales de dolor lleguen al cerebro produciendo, además, una sensación de hormigueo (parestesia) en los dermatomas estimulados [4]. Estas fibras se encuentran tanto en los DC como en las DR. Sin embargo, como en los DC se puede activar todos los dermatomas de un nivel vertebral, el objetivo de la terapia de EME se centra en maximizar la activación neuronal en los DC [2], [5].

Aunque la tasa de éxito de la terapia de EME se encuentra entre el 62% y el 84% (dependiendo de la condición de dolor crónico) [6], el 35% de los pacientes sufren los efectos secundarios de la estimulación tónica, como la pérdida de parestesia en los dermatomas dolorosos, la aparición de shocks eléctricos o la neuralgia intercostal [7]. Por tanto, existe una necesidad de aumentar la precisión a la hora de programar los electrodos y de posicionarlos adecuadamente para minimizar los efectos secuandarios.

Durante los últimos 40 años, el desarrollo de modelos computacionales generalizados ha permitido estudiar el efecto de los parámetros de estimulación (frecuencia, anchura de pulso, amplitud y polaridad) sobre la activación neuronal [8]. Sin embargo, en cada nivel vertebral, el tamaño y la forma de la médula espinal y del canal espinal, la posición relativa de la médula espinal en el canal espinal y la cantidad de líquido cefalorraquídeo (CSF, por sus siglas en inglés) entre la localización epidural de los electrodos y las fibras nerviosas de los DC varía en cada paciente [9]. Dado que estos parámetros geométricos influyen en la activación de las fibras, actualmente se están desarrollando modelos computacionales personalizados, que parecen predecir umbrales de estimulación que se aproximan mejor a los valores clínicos [10].

Con el objetivo de estudiar la precisión en la predicción de la activación neuronal y las ventajas de modelizar de manera más detallada o general la médula espinal, en este trabajo se compara los resultados obtenidos de un modelo personalizado y un modelo generalizado con los de un caso clínico.

2. Metodología

2.1. Modelo de médula espinal generalizado

El modelo generalizado, desarrollado por nuestro grupo [11], se compone de los siguientes tejidos: materia gris, materia blanca, CSF, grasa y hueso (ver Figura 1). El

ISBN: 978-84-09-25491-0, pags. 306- 309

modelo incluye dos electrodos percutáneos en el espacio epidural (grasa), con la interfase electrodo-tejido en los polos activos. La posición de los electrodos se aproxima a su posición real en el paciente de este estudio. Cada estructura tiene asignada la conductividad eléctrica correspondiente [11].

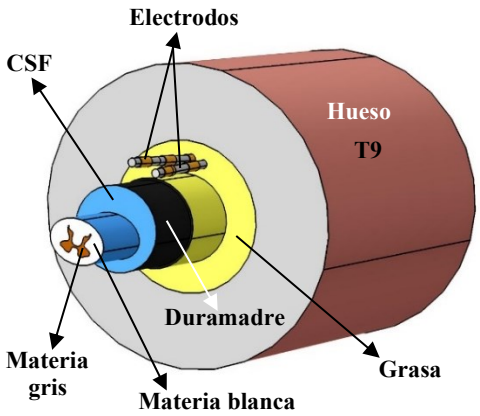


Figura 1. *Modelo conductor volumétrico generalizado de la médula espinal desarrollado en COMSOL Multiphysics* [11]. *CSF: líquido cefalorraquídeo.*

La geometría se basa en valores medios de los parámetros geométricos, obtenidos a partir de resonancias magnéticas de alta resolución *in vivo* de médulas espinales humanas [12], correspondientes al nivel vertebral T9, nivel donde el paciente de este estudio recibía la estimulación eléctrica.

2.2. Modelo de médula espinal personalizado al paciente

Para obtener las estructuras anatómicas de la médula espinal del paciente hemos utilizado resonancias magnéticas de alta resolución (T2* y T1*). Con el software *Spinal Cord Toolbox (SCT)* [13], se ha segmentado automáticamente la materia blanca y el CSF. El hueso se ha segmentado manualmente usando el software Seg3D. Estas tres segmentaciones se han importado a COMSOL Multiphysics. La duramadre se ha incluido en COMSOL Multiphysics como una capa de 0.3 mm que envuelve el CSF; y la grasa como un cilindro que cubre todo el espacio epidural. La materia gris no se incluye en el modelo para reducir el coste computacional, ya que se ha demostrado que la materia gris no afecta a la distribución del campo eléctrico [14]. El modelo considera tres niveles vertebrales distintos, desde T8 hasta T10 (ver Figura 2), que es donde el paciente real tiene los electrodos implantados. Los electrodos percutáneos se sitúan en el espacio epidural, en la misma posición donde los tiene el paciente de este estudio.

2.3. Modelo de fibra nerviosa sensitiva mielínica

El modelo de fibra que se ha utilizado y programado en MATLAB es una combinación de dos modelos: el modelo de Wesselink [15], que contiene la dinámica de membrana correspondiente al de una fibra sensitiva mielínica humana en los nodos de Ranvier; y el modelo B de McIntyre, Richardson and Grill (MRG) [16], que modeliza de forma más realista los internodos, pues se consideran como aislantes imperfectos, es decir, se tiene en cuenta las pérdidas de corriente. El esquema eléctrico equivalente se puede observar en la Figura 3.

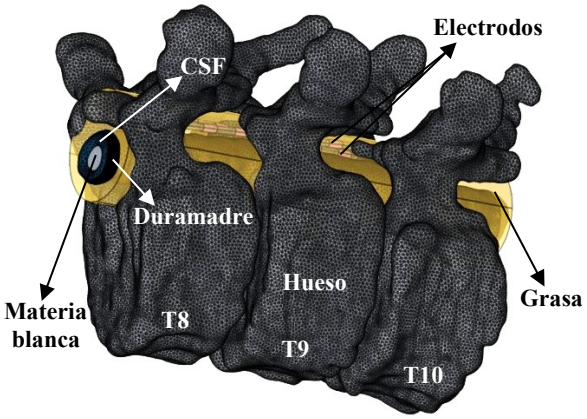


Figura 2. *Modelo conductor volumétrico personalizado al paciente, desarrollado en COMSOL Multiphysics. CSF: líquido cefalorraquídeo.*

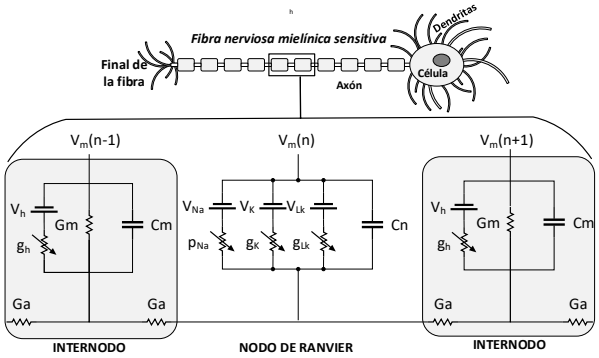


Figura 3. *Esquema eléctrico del modelo de fibra nerviosa sensitiva.*

Aplicando la ley de Kirchhoff, se obtiene una ecuación diferencial para los nodos de Ranvier (ecuación (1)), y una ecuación diferencial para los internodos (ecuación (2)):

$$\frac{dV_{m,n}}{dt} = \frac{1}{C_n}\left[G_a\left(V_{m,n-1} - 2V_{m,n} + V_{m,n+1} + V_{e,n-1} - 2V_{e,n} + V_{e,n+1}\right) - \pi dl I_{ion,n}\right] \quad (1)$$

$$\frac{dV_{m,n}}{dt} = \frac{1}{C_m}\left[G_a\left(V_{m,n-1} - 2V_{m,n} + V_{m,n+1} + V_{e,n-1} - 2V_{e,n} + V_{e,n+1}\right) - G_m(V_n - V_{rest})\right] \quad (2)$$

Siendo V_m el potencial de membrana, V_e el potencial extracelular, n el compartimento a analizar, C_n y C_m las capacidades de la membrana en el nodo y en el internodo respectivamente (μF), I_{ion} la suma total de las corrientes iónicas (mA), G_a la conductancia entre los centros de dos compartimentos adyacentes (mS) y G_m la conductancia en los internodos (mS). La resolución de las ecuaciones diferenciales permite calcular el potencial de acción y determinar si una fibra queda activada o no ante un estímulo eléctrico.

2.4. Parámetros de salida del modelo

Se calculan los siguientes parámetros para comparar cuantitativamente los resultados de los modelos [11]:

- *Umbral de percepción (PT).* Voltaje necesario para activar las fibras Aβ de 12.8 µm de diámetro en los DC (PT_{DC}) o de 15 µm en las DR (PT_{DR}). Expresado en V; corresponde al potencial programado en un polo del electrodo con respecto a una referencia remota.
- *Umbral de molestia (DT).* Voltaje a partir del cual la estimulación es molesta. Se define como 1.4*PT.
- *Área de activación (AA).* Máxima área de fibras nerviosas de 12.8 µm de diámetro que se activan en los DC en DT, en un corte transversal de la médula.
- *Profundidad de activación (PA).* Máxima distancia longitudinal en un corte transversal de la médula en la que fibras nerviosas de 12.8 µm de diámetro han sido activadas en los DC en DT.
- *Ratio de reclutamiento ($R_{DC/DR}$).* Ratio entre PT_{DC} y PT_{DR}.

3. Resultados

La paciente de este estudio es una mujer de 47 años diagnosticada de síndrome de cirugía de espalda fallida; presentaba dolor crónico en el miembro inferior y las lumbares del lado izquierdo (ver Figura 4(C)). Los parámetros de estimulación aplicados para cubrir con parestesia los dermatomas dolorosos son: anchura de pulso de 300 µs, frecuencia de 50 Hz, amplitud de 2.2 mV y, como polaridad, cátodo guardado transverso (CGT) en los polos 5(+) 6(-) 7(+) y 14(+) (ver Figura 4(A)).

Con los modelos de médula espinal generalizado y personalizado al paciente, se ha calculado la distribución del campo eléctrico en la superficie de los DC (ver Figura 4(A)), se ha simulado la activación neuronal en los DC (ver Figura 4(B)) y se ha calculado los parámetros de estimulación tras aplicar un estímulo rectangular de 300 µs de duración. Los resultados obtenidos se pueden observar en la Tabla 1. En cuanto al PT_{DC} y el DT, el modelo generalizado da valores que son aproximadamente 170% superiores a los valores del caso real (4.28 y 6 mV frente a 1.6 y 2.2 mV, respectivamente), mientras que el modelo personalizado solo muestra un 5% de diferencia (1.52 y 2.1 mV frente a 1.6 y 2.2 mV).

Fuente	PT_{DC} (V)	PT_{DR} (V)	DT (V)	$R_{DC/DR}$	AA (mm^2)	PA (µm)
Paciente real	1.6	-	2.2	-	-	-
Modelo generalizado	4.28	9.1	6	0.47	1.46	450
Modelo personalizado	1.52	3.2	2.1	0.475	1.32	398

Tabla 1. *Parámetros de evaluación obtenidos del paciente real, del modelo generalizado y del modelo personalizado.*

PT_{DR} también es superior en el modelo personalizado (184% mayor que el modelo personalizado). Sin embargo, solo existe una diferencia del 1% entre ambos modelos al calcular $R_{DC/DR}$. AA y AD son valores que no se pueden obtener clínicamente. En este caso, ambos modelos dan resultados similares, aunque los valores del modelo generalizado también son superiores al modelo personalizado: un 10.6% en AA (1.46 frente a 1.32 mm^2) y un 13% en AD (450 frente a 398 µm).

Si observamos el número de fibras activadas en cada lado de los DC en ambos modelos (ver Figura 4(B)), el modelo personalizado activa 75 fibras del lado derecho, llegando hasta las fibras que se encuentran a 1 mm de la línea media, mientras que el modelo generalizado activa 140 fibras, llegando a fibras que se encuentran a 1.5 mm de la línea media. Por otra parte, la relación entre las fibras activadas en el lado izquierdo y el lado derecho en el modelo personalizado es de 4.3 (320 frente a 75), mientras que en el modelo generalizado es de 3 (422 frente a 140). Dado que la paciente no reportó sensación de parestesia en el lado derecho del cuerpo, es probable que la mayor parte de la activación se produjera en el lado izquierdo. Por lo tanto, en cuando a la predicción de la activación neuronal, el modelo personalizado parece aproximarse mejor a la realidad que el modelo generalizado.

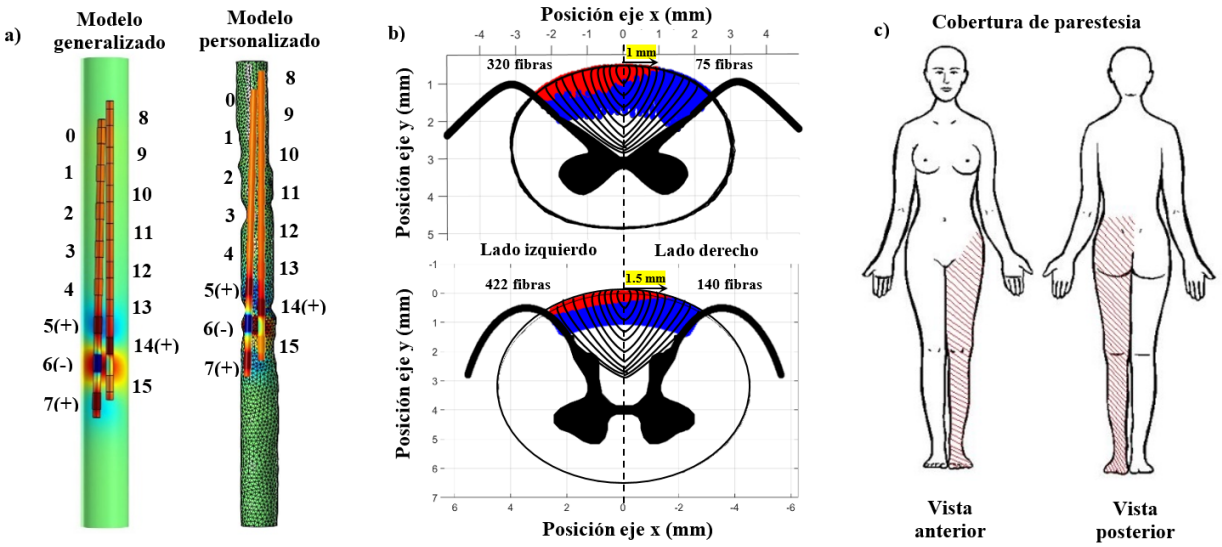


Figura 4. *a) Segunda derivada del potencial eléctrico en la superficie de los DC. b) Activación neuronal en los DC obtenidos del modelo personalizado al paciente y del modelo generalizado. c) Cobertura de parestesia en la paciente de este estudio.*

4. Discusión y conclusiones

En este trabajo se ha comparado la precisión en la predicción de la activación neuronal mediante un modelo de médula espinal personalizado al paciente frente a un modelo generalizado.

Los resultados muestran que el modelo personalizado predice umbrales de estimulación (PT_{DC} y DT) que se aproximan más a los valores clínicos, al igual que Lempka et al. [10]. Por otra parte, la cobertura de parestesia de la paciente de este estudio se encuentra en sus dermatomas dolorosos, es decir, en el lado izquierdo del cuerpo. En este caso, el modelo personalizado predice un menor número de fibras activadas en el lado derecho que el modelo generalizado (casi un 50% menos (75 fibras frente a 140)). Además, el número de fibras activadas en el lado izquierdo es 4 veces mayor que en el lado derecho, mientras que en el modelo generalizado es 3 veces mayor. Por tanto, el modelo personalizado también parece ser más preciso en el sentido de poder predecir dónde va a notar el paciente la estimulación, es decir, la sensación de hormigueo.

En cuanto a los parámetros AA, AD y $R_{DC/DR}$, la diferencia máxima observada entre los valores obtenidos con el modelo el personalizado y el modelo generalizado es del 13%. Este resultado también es importante pues, cuando el objetivo es estudiar el efecto de los parámetros de estimulación o de los parámetros geométricos sobre la activación neuronal de forma genérica o cualitativa, un modelo generalizado ofrece la ventaja de tener un coste computacional bajo, por lo que su uso permitiría trabajar más rápidamente que con un modelo personalizado.

Según los resultados, un modelo personalizado demuestra ser más preciso a la hora de predecir la activación neuronal y su posible efecto en términos de parestesia que los modelos generalizados. Sin embargo, este estudio solo se ha realizado con un paciente. Para validar esta conclusión, un estudio con un número más elevado de pacientes debería realizarse. Si esta hipótesis resulta ser cierta, entonces los modelos personalizados podrían ayudar a mejorar la cobertura de parestesia en la práctica clínica a través de aplicaciones clínicas futuras como: la optimización de los parámetros de estimulación para seleccionar los dermatomas de interés o la planificación de la posición de los electrodos antes del implante del sistema de EME en cada uno de los pacientes.

Referencias

[1] A. W. Lee and J. G. Pilitsis, “Spinal cord stimulation: indications and outcomes.,” *Neurosurg. Focus*, vol. 21, no. 6, pp. 1–6, 2006.

[2] G. Molnar and G. Barolat, “Principles of cord activation during spinal cord stimulation,” *Neuromodulation*, vol. 17, no. SUPPL. 1, pp. 12–21, 2014.

[3] J. Caylor *et al.*, “Spinal cord stimulation in chronic pain: evidence and theory for mechanisms of action,” *Bioelectron. Med.*, vol. 5, no. 1, pp. 1–41, 2019.

[4] R. Melzack and P. D. Wall, “Pain Mechanisms: A New Theory,” *Science (80-.).*, vol. 150, no. 3699, pp. 971–979, 1965.

[5] J. Holsheimer, “Which neuronal elements are activated directly by spinal cord stimulation,” *Neuromodulation*, vol. 5, no. 1, pp. 25–31, 2002.

[6] T. Cameron, “Safety and efficacy of spinal cord stimulation for the treatment of chronic pain: A 20-year literature review,” *J. Neurosurg.*, vol. 100, no. 3 SUPPL., pp. 254–267, 2004.

[7] J. C. Kleiber, B. Marlier, M. Bannwarth, E. Theret, P. Peruzzi, and F. Litre, “Is spinal cord stimulation safe? A review of 13 years of implantations and complications,” *Rev. Neurol. (Paris).*, vol. 172, no. 11, pp. 689–695, 2016.

[8] R. Vallejo, K. Bradley, and L. Kapural, “Spinal Cord Stimulation in Chronic Pain,” *Spine (Phila. Pa. 1976).*, vol. 42, no. 14, pp. S53–S60, 2017.

[9] R. M. Levy, “Anatomic considerations for spinal cord stimulation,” *Neuromodulation Technol. Neural Interface*, vol. 17, pp. 2–11, 2014.

[10] S. F. Lempka, H. J. Zander, C. J. Anaya, A. Wyant, J. G. Ozinga, and A. G. Machado, “Patient-Specific Analysis of Neural Activation During Spinal Cord Stimulation for Pain,” *Neuromodulation Technol. Neural Interface*, vol. 2019, 2019.

[11] J. L. Durá, C. Solanes, J. De Andrés, and J. Saiz, “Computational Study of the Effect of Electrode Polarity on Neural Activation Related to Paresthesia Coverage in Spinal Cord Stimulation Therapy,” *Neuromodulation*, vol. 22, no. 3, pp. 269–279, 2019.

[12] L. Fradet, P. J. Arnoux, J. P. Ranjeva, Y. Petit, and V. Callot, “Morphometrics of the entire human spinal cord and spinal canal measured from in vivo high-resolution anatomical magnetic resonance imaging,” *Spine (Phila. Pa. 1976).*, vol. 39, no. 4, pp. E262–E269, 2014.

[13] B. De Leener, J. Cohen-Adad, and S. Kadoury, “Automatic Segmentation of the Spinal Cord and Spinal Canal Coupled with Vertebral Labeling,” *IEEE Trans. Med. Imaging*, vol. 34, no. 8, pp. 1705–1718, 2015.

[14] J. J. Struijk, J. Holsheimer, B. K. van Veen, and H. B. K. Boom, “Epidural Spinal Cord Stimulation: Calculation of Field Potentials with Special Reference to Dorsal Column Nerve Fibers,” *IEEE Trans. Biomed. Eng.*, vol. 38, no. 1, pp. 104–110, 1991.

[15] W. A. Wesselink, J. Holsheimer, and H. B. K. Boom, “A model of the electrical behaviour of myelinated sensory nerve fibres based on human data,” *Med. Biol. Eng. Comput.*, vol. 37, no. 2, pp. 228–235, 1999.

[16] A. G. Richardson, C. C. McIntyre, and W. M. Grill, “Modelling the effects of electric fields on nerve fibres: Influence of the myelin sheath,” *Med. Biol. Eng. Comput.*, vol. 38, no. 4, pp. 438–446, 2000.

Resultados del reto “Biomarker and Endpoint Assessment to Track Parkinson's Disease (BEAT-PD) DREAM Challenge”

J. Aragón Arqués [1,2], M. Valencia Ustárroz [1,2,3]

[1] Universidad de Navarra, CIMA, Pamplona, España, [2] Universidad de Navarra, DATAI, Pamplona, España

[3] IdiSNA, Instituto de Investigación Sanitaria de Navarra, Pamplona, España

jaragonarqu@alumni.unav.es, mvustarroz@unav.es

Resumen

El desafío “The Biomarker and Endpoint Assessment to Track Parkinson's Disease (BEAT-PD) Challenge” [1] es el primer reto diseñado para evaluar diferentes técnicas de procesamiento de datos no estructurados provenientes de sensores móviles (relojes o móviles inteligentes) con el objetivo de estudiar la progresión de la Enfermedad de Parkinson (EP). En esta comunicación presentamos los resultados obtenidos mediante la aplicación de un método implementado para identificar del estado motor de pacientes de EP a partir de las señales proporcionadas en el BEAT-PD. La aproximación implementa algoritmos de ML (Machine Learning) para detectar los niveles de gravedad de discinesia y temblor y a continuación determina el estado motor (ON/OFF) del paciente. Se han implementado modelos de GBM (Gradient Boosting Machine) SVM (Support Vector Machine) y RF (Random Forest). Las métricas obtenidas demuestran un rendimiento muy competitivo, teniendo en cuenta que para la implementación del método se han utilizado única y exclusivamente medidas de acelerometría, excluyéndose informaciones clínicas o demográficas.

1. Introducción

La enfermedad de Parkinson (EP) es una enfermedad neurodegenerativa, crónica e invalidante, que afecta en la actualidad a 160.000 personas en España y a más de 7 millones de personas en todo el mundo [2]. Se caracteriza por una pérdida de las neuronas dopaminérgicas de la sustancia negra dando lugar a un amplio espectro de alteraciones motoras, autonómicas y cognitivas. Posiblemente el síntoma más identificable de la EP son los temblores en reposo, que se manifiestan principalmente en las extremidades superiores. También pueden verse afectados el mentón, los labios, la cara y las extremidades inferiores, desapareciendo durante la ejecución de movimientos voluntarios. Paradójicamente, el temblor es uno de los síntomas motores que menos imposibilita al enfermo de EP. Por el contrario, la bradicinesia, del griego antiguo bradys (“lento”) y kinesis (“movimiento”), es el síntoma caracterizado por el enlentecimiento del movimiento y causa dificultades en la planificación, el inicio y la ejecución de movimientos, así como la realización de tareas secuenciales y simultáneas. Con el progreso de la enfermedad los síntomas se empeoran y aparecen otros trastornos que la agravan.

Algunos de los síntomas de la EP pueden tratarse mediante tratamientos farmacológicos orientados a incrementar los niveles de dopamina en el cerebro. En general, y en los primeros años de la enfermedad estos tratamientos resultan eficaces y seguros y en torno a un 80 % de los pacientes tratados con precursores o agonistas de la dopamina manifiestan una mejoría inicial y un 20 % llega a recuperar por completo la función motora. Sin embargo, este tipo de terapias no frenan la progresión de la EP y con el tiempo la respuesta se reduce o da lugar a complicaciones y efectos secundarios [3] [4]. Una de las principales complicaciones que se deriva del tratamiento prolongado es la aparición de discinesias; movimientos anormales involuntarios de pies y manos (vagamente comparables a bailar o tocar el piano) que en la mayoría de los casos resultan tanto o más invalidantes de la propia enfermedad. Con los años aparecen también fluctuaciones motoras como consecuencia de cambios en la respuesta a la medicación y se manifiestan mediante episodios en los que reaparecen los síntomas de la enfermedad. Los períodos OFF son aquéllos en los pacientes manifiestan los síntomas propios de la EP. En contrapartida, durante los períodos ON los pacientes recuperan el control de la función motora, aunque en algunos casos, pueden acompañarse de la aparición de discinesias. Así, la medicación se ajusta tratando de mantener los niveles adecuados de dopamina con el fin de maximizar el tiempo en ON sin discinesia y minimizar la severidad de los períodos OFF y su duración [3], [5], [6]. Por lo tanto, la obtención de información precisa sobre la evolución de las fluctuaciones motoras o la aparición de efectos no deseados resulta imprescindible para proporcionar un tratamiento óptimo del paciente, minimizando el tiempo en OFF y la aparición de discinesias [3], [5].

Por otro lado, los avances recientes en la mHealth o salud móvil han demostrado un gran potencial para aprovechar las tecnologías basadas en sensores que permitirían llevar a cabo una monitorización continua y remota de los pacientes incluso en el entorno domiciliario. Aunque muchas de las aproximaciones se llevan a cabo utilizando sensores portátiles “de ámbito científico/medico”, cada vez más se presentan iniciativas que emplean los sensores que incorporan teléfonos inteligentes u otros dispositivos portátiles de consumo, como relojes inteligentes y pulseras de actividad física. Estos dispositivos no sólo permitirían medir los fenotipos de la enfermedad de forma mucho más detalladas, sino que además abren la posibilidad de llevar a cabo el seguimiento de los pacientes con una frecuencia mayor que la proporcionada por los exámenes clínicos. Sin embargo, el proceso de conversión de los flujos de datos basados en sensores en biomarcadores digitales es complejo y aún no se han desarrollado estándares metodológicos para guiar este proceso.

ISBN: 978-84-09-25491-0, pags. 310- 313

En esta comunicación presentamos un método para la identificación del estado motor de pacientes de EP a partir de los datos del desafío "The Biomarker and Endpoint Assessment to Track Parkinson's Disease (BEAT-PD) Challenge" [1]. El método implementa una primera fase de procesamiento de señal, seguida del empleo de algoritmos de ML (Machine Learning) que aplicados exclusivamente sobre señales de acelerometría permite predecir la gravedad de las discinesias y el nivel de temblor, además de identificar el estado motor de los pacientes

2. Materiales y Métodos

2.1. BEAT-PD Challenge

Financiado y gestionado por la Fundación Michael J. Fox para la Investigación del Parkinson, el reto BEAT-PD fue publicado a principios de 2020 e incluye datos recopilados por investigadores de la Universidad Northwestern, la Universidad de Rochester, la Universidad de Alabama, la Universidad de Cincinnati y el Centro Médico de la Universidad Radboud. Los datos fueron obtenidos en el contexto de dos ensayos clínicos, REAL-PD [7] y CIS-PD (Clinician Input Study Parkinson's Disease) [8] e incluyen parámetros clínicos e información de sensores inerciales obtenidos mediante relojes o teléfonos inteligentes en el entorno domiciliario de pacientes de EP (Figuras 1 y 2). Para más detalles e información actualizada, puede consultarse la página web: https://www.synapse.org/#!Synapse:syn20825169.

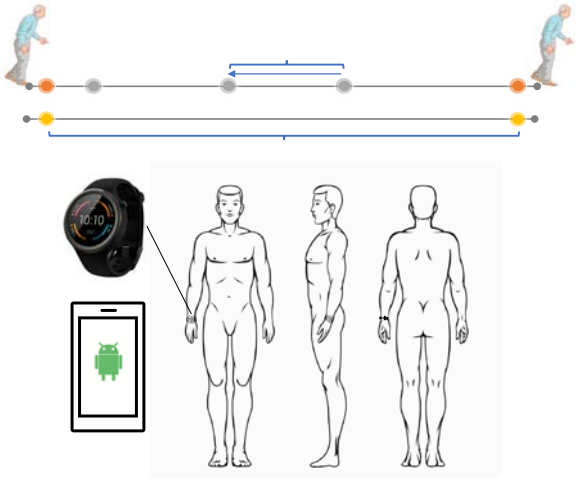


Figura 1.*El estudio REAL-PD tuvo una duración de 2 semanas en las que se monitorizó de forma remota a 12 pacientes domiciliarios utilizando un teléfono Android y un reloj Motorola Moto 360. Los pacientes informaran sus síntomas a lo largo de dos días durante bloques de tiempo de 30'. Los pacientes auto informaron y clasificaron el temblor en una escala de gravedad de 0 a 4 (menor a mayor gravedad), el estado de la medicación ON/OFF (0/1) y discinesia (0 a 4).*

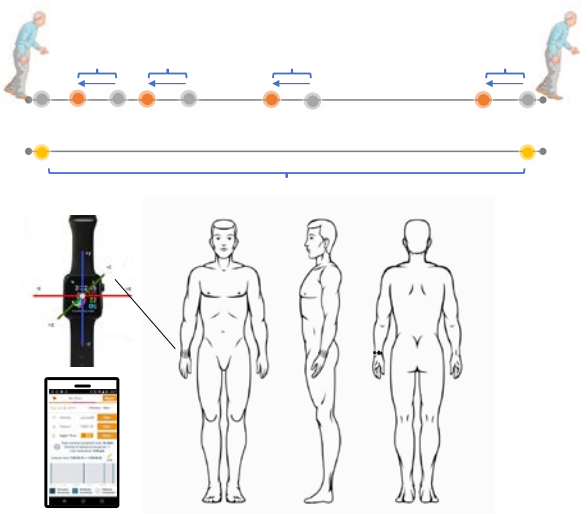


Figura 2. *En el estudio CIS-PD 16 pacientes con EP usaron un Apple watch series 2 junto con la aplicación Fox Wearable Companion en un teléfono móvil para y transmitir los datos recogidos del acelerómetro, informar de la gravedad de los síntomas y de los niveles de medicación. El estudio tuvo una duración de 6 meses en los que durante dos semanas los participantes clasificaron los síntomas (ON/OFF discinesia y temblor) en periodos de 30' en escalas de 0 a 4*

En este trabajo sólo se han utilizado los datos provenientes del acelerómetro triaxial de los relojes inteligentes, y más concretamente las componentes de aceleración lineal muestreadas a 50Hz. El BEAT-PD Challenge también proporciona datos demográficos y clínicos que no se han utilizado en el desarrollo de los modelos.

2.2. Preprocesamiento y extracción de parámetros

De forma cualitativa y a través de la simple inspección visual de las señales, puede apreciarse cómo en función del estado motor del paciente, las señales de acelerometría difieren completamente en lo que respecta a los cambios de la amplitud de la señal en el tiempo (Figura 3). Por ello, resulta intuitivo incluir algunos parámetros estadísticos obtenidos en el dominio del tiempo (RMS, rango, sesgo, kurtosis, sobre aceleración, media, mediana valor absoluto de la mediana y desviación típica) de la resultante y de los tres ejes de aceleración Adicionalmente se han incluido parámetros obtenidos mediante la estimación de la correlación cruzada entre las señales de los tres ejes del acelerómetro.

Además de los parámetros asociados a cada uno de los tres ejes de aceleración, se obtuvieron descriptores de la aceleración resultante, los valores de pico y frecuencia correspondientes a los máximos en la FFT, PSD y función de autocorrelación. Finalmente, y dado que un espectro de alta potencia en la banda de (1, 4) Hz es un indicador de que el sujeto tiene discinesia o que camina o sube o baja escaleras, es necesario analizar la siguiente banda (8, 20) Hz para distinguir entre ambos casos [10], [11] y las frecuencias superiores a 4 Hz corresponden al temblor (Figura 4) [3] [9], se han extraído como parámetros los valores de PSD en estas bandas.

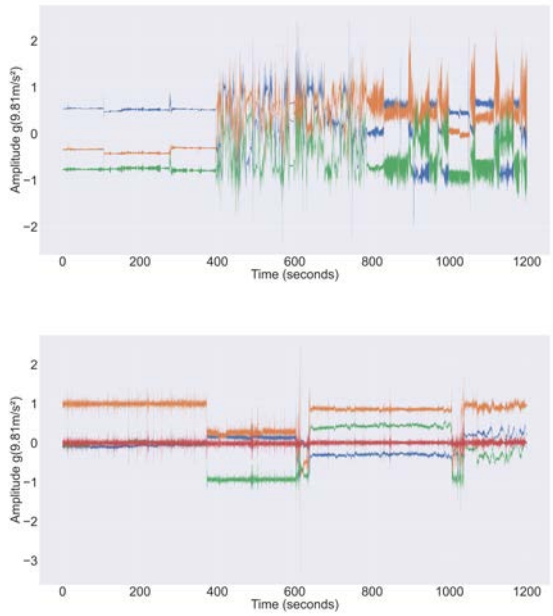


Figura 3. *Señales de acelerometría con/sin sintomatología*

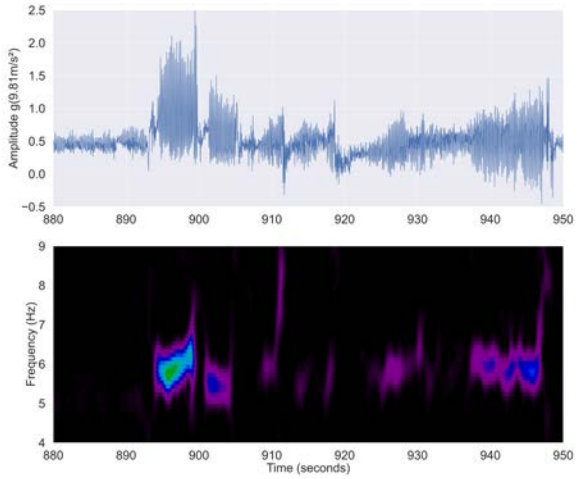


Figura 4. *Componente Y durante un episodio de temblor y su representación tiempo-frecuencia. Puede verse cómo la frecuencia del temblor se centra en el rango de 5 a 8 Hz*

2.3. Preparación de los datasets entrenamiento/test

Las variables para el conjunto de datos entrenamiento y test se corresponden con las características extraídas del análisis de las ventanas completas de cada una de las grabaciones de las señales de acelerometría. En total se extrajeron 211 parámetros asociados al temblor, 216 a discinesias y 385 relativos al estado ON/OFF. Sobre el total de variables para los modelos se realizó una selección calculando la correlación entre los regresores y la variable objetivo, se les asignó una puntuación F y un valor P, en función de la cual se han seleccionado 40 características de temblor, 40 de discinesias y 15 características para el ON/OFF en el estudio CIS-PD, y 20 en el REAL-PD. En ningún momento se incluyeron variables clínicas o demográficas de los informes UPDRS de los sujetos.

Los conjuntos de datos de entrenamiento/test de cada sujeto, se repartieron con el criterio 70% entrenamiento y 30% test. Para mitigar el efecto del gran desbalance entre clases, se realizó un sobre muestreo aleatorio con reemplazo. De esta forma se logra entrenar los modelos con una distribución de datos equilibrada en clases.

2.4. Diseño y parametrización de los modelos de ML

Para cada una de las variables a predecir se entrenaron 3 modelos de regresión basados en los algoritmos GBM, RF, SVM. El entrenamiento y test de cada modelo se realizó de forma personalizada utilizando únicamente los datos de cada sujeto por separado. El ajuste de los hiper parámetros y optimización los algoritmos también se realizó a nivel de sujeto.

2.5. Evaluación de los modelos

Para cada sujeto del estudio, *k*, la precisión del modelo se evalúa como el error cuadrático medio *(MSE)*:

$$MSE_k = \frac{1}{n_k}\sum_{i=1}^{n_k}(y_{i,k} - \hat{y}_{i,k})^2$$

Donde n_k es el número de observaciones de cada *k(sujeto)*, $y_{i,k}$, es el valor de la i^{th} variable objetivo para el individuo *k* e $\hat{y}_{i,k}$ es su predicción correspondiente.

La puntuación estadística final se calcula al ponderar los MSE's, según la fórmula:

$$FinalScore_{sqrt} = \frac{\sum_{k=1}^{N}\sqrt{n_k}MSE_k}{\sum_{k=1}^{N}\sqrt{n_k}}$$

3. Resultados

RF ha sido el algoritmo que mejor comportamiento ha tenido en las rondas de entrenamiento y test (MSE y WMSE), SVM fue el que arrojó los segundos mejores resultados y por último GBM (Tabla 1).

	CIS-PD (Weighted MSE)			REAL-PD (Weighted MSE)		
	on_off	discinesia	temblor	on_off	discinesia	temblor
RF	0.9271	0.5893	0.4683	0.2023	0.2508	0.3214
SVM	1.0403	0.5750	0.4892	0.2227	0.1890	0.3724
GBM	0.9128	0.6387	0.5596	0.2324	0.2230	0.3425

Tabla 1. *Comparativa de métricas WMSE durante las fases de entrenamiento y test*

Ya fuera de plazo solicitamos a la organización evaluar los resultados de esta implementación. Para ello, los miembros de la organización del desafío BEAT-PD evaluaron nuestros modelos a partir de un subconjunto de datos de entrenamiento sin etiquetar que puso a disposición de los equipos participantes. Los resultados obtenidos para el modelo y la posición ocupada respecto al resto de participantes se muestra en la Tabla 2.

Subchallenge	Ranking	Weighted MSE
S1 - on/off	14	1.0679
S2 - discinesia	8	0.4372
S3 - temblor	11	0.4407

Tabla 2. *Resultados obtenidos utilizando RF. En total participaron 37 equipos, formados por científicos de datos e ingenieros biomédicos de centros de investigación y Universidades de todo el mundo*

4. Discusión y conclusiones

En esta comunicación presentamos la implementación de modelos de ML aplicados a la identificación de síntomas motores en enfermos de Parkinson a partir de señales de acelerometría. Constituye un ejemplo de cómo el análisis de la información recogida mediante dispositivos digitales de consumo podría ofrecer soporte para mejorar el tratamiento de los pacientes de EP. Es más, brinda la oportunidad de investigadores clínicos y básicos recaben información acerca de la enfermedad en entornos domiciliarios, posibilitando la monitorización remota; lo que redunda en una mejora de la asistencia sanitaria, los resultados en salud de los pacientes de EP y el progreso de la investigación médica. Los resultados presentados sugieren la posibilidad de implementar sistemas fácilmente escalables que permitan realizar un seguimiento de la progresión de la EP más detallado.

Comparando el resultado con los de los equipos que han participado en el BEAT-PD Challenge, los resultados de este trabajo presentan un grado de precisión comparable, teniendo en cuenta que no se han utilizado en los modelos variables clínicas o demográficas, salvo su código de identificación. Es muy probable que la inclusión de los datos UPDRS habría incrementado la precisión de los modelos, pero dicha utilización introduce también un sesgo en los modelos (por ejemplo, un sujeto con niveles de temblor alto en los informes clínicos puede producir que el modelo siempre tienda a predecir niveles altos de temblor en ese sujeto). El objetivo en esta comunicación ha sido trabajar única y exclusivamente a partir de las señales de acelerometría, a fin de dotar al sistema de variables dinámicas que permitan recoger los cambios en la progresión de la EP con mayor precisión y flexibilidad. Dividir las grabaciones de las medidas de los sensores en fragmentos o ventanas de tiempo más pequeñas es otra estrategia que también podría mejorar los resultados. Y sería interesante también explorar en un futuro el uso de redes neuronales convolucionales (CNN). Los datos del acelerómetro son similares en estructura a los datos de imágenes una vez que se les aplica una transformada tiempo frecuencia como la de la Figura 3. Por otro lado, faltaría por realizar un análisis de la correspondencia entre los niveles auto informados con ciertas características de las señales del sensor; esta información permitiría establecer umbrales y corregir valores auto informados incoherentes (por ejemplo, un nivel de temblor igual 0 frente a una energía muy alta en su banda asociada) ayudando a mejorar los modelos.

Agradecimientos

La Fundación Michael J. Fox para la Investigación del Parkinson ha financiado el estudio Clinician Input Study (CIS-PD) Wireless Adhesive Sensor Sub-Study (Grant No. 14702), y ha hecho posible que los datos analizados en este trabajo estuvieran públicamente disponibles gracias a la organización del “The Biomarker and Endpoint Assessment to Track Parkinson's Disease (BEAT-PD) Challenge”.

Referencias

[1] BEAT-PD DREAM Challenge - syn20825169». https://www.synapse.org/#!Synapse:syn20825169/wiki/596118.

[2] Ray Dorsey, E., ET AL. (2018). Global, regional, and national burden of Parkinson’s disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet Neurology, 17(11), 939–953. https://doi.org/10.1016/S1474-4422(18)30295-3

[3] C. Pérez-López et al., «Assessing Motor Fluctuations in Parkinson’s Disease Patients Based on a Single Inertial Sensor», Sensors, vol. 16, n.o 12, p. 2132, dic. 2016

[4] Stowe R, Ives N, Clarke CE, Deane K, van Hilten , Wheatley K, Gray R, Handley K, Furmston A. Evaluation of the efficacy and safety of adjuvant treatment to levodopa therapy in Parkinson´s disease patients with motor complications. Cochrane Database of Systematic Reviews 2010, Issue 7. Art. No.: CD007166. DOI: 10.1002/14651858.CD007166.pub2.

[5] A. Antonini y E. Tolosa, «Apomorphine and levodopa infusion therapies for advanced Parkinson’s disease: selection criteria and patient management», Expert Rev. Neurother., vol. 9, n.o 6, pp. 859-867, jun. 2009.

[6] A. Rodríguez-Molinero et al., «Remote control of apomorphine infusion rate in Parkinson’s disease: Real-time dose variations according to the patients’ motor state. A proof of concept», Parkinsonism Relat. Disord., vol. 21, n.o 8, pp. 996-998, ago. 2015.

[7] Y. P. Raykov et al., «Probabilistic modelling of gait for remote passive monitoring applications», ArXiv181202585 Eess Math, ene. 2019. Disponible en: http://arxiv.org/abs/1812.02585.

[8] J. J. Elm et al., «Feasibility and utility of a clinician dashboard from wearable and mobile application Parkinson’s disease data», Npj Digit. Med., vol. 2, n.o 1, Art. n.o 1, sep. 2019, doi: 10.1038/s41746-019-0169-y.

[9] G. Deuschl, P. Bain, y M. Brin, «Consensus statement of the Movement Disorder Society on Tremor. Ad Hoc Scientific Committee», Mov. Disord. Off. J. Mov. Disord. Soc., vol. 13 Suppl 3, pp. 2-23, 1998, doi: 10.1002/mds.870131303.

[10] A. J. Manson, P. Brown, J. D. O’Sullivan, P. Asselman, D. Buckwell, y A. J. Lees, «An ambulatory discinesia monitor», J. Neurol. Neurosurg. Psychiatry, vol. 68, n.o 2, pp. 196-201, feb. 2000, doi: 10.1136/jnnp.68.2.196.

[11] A. Samà Monsonís, «Human movement analysis by means of accelerometers : aplication to human gait and motor symptoms of Parkinson’s Disease», TDX Tesis Dr. En Xarxa, oct. 2013. https://upcommons.upc.edu/handle/2117/95248

Señales Biomédicas III

Jueves, 26 de noviembre de 2020

Análisis de flujo aéreo y saturación de oxígeno en sangre mediante transformada *wavelet* para la detección de la apnea obstructiva del sueño infantil

J. Jiménez García [1], G.C. Gutiérrez Tobal [1,2], M. García Gadañón [1,2], D. Álvarez González [1,2,3], A. Martín Montero [1], F. del Campo [1,2,3], L. Kheirandish-Gozal [4], D. Gozal [4], R. Hornero Sánchez [1,2]

[1] Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, jorge.jimenez@gib.tel.uva.es

[2] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Instituto de Salud Carlos III, Madrid, España

[3] Servicio de Neumología, Hospital Universitario Río Hortega, Valladolid, España, fsas@telefonica.net

[4] *Dept. of Child Health*, The University of Missouri School of Medicine, Columbia, Missouri, EEUU, gozald@health.missouri.edu

Resumen

El objetivo de este estudio fue evaluar la utilidad de la transformada wavelet de las señales de flujo aéreo (FA) y saturación de oxígeno en sangre (SpO_2) para detectar la apnea obstructiva del sueño (AOS) infantil. Para ello, se emplearon las señales de 974 sujetos pediátricos con sospecha de padecer AOS, que se dividieron aleatoriamente en conjuntos de entrenamiento y test. Se extrajeron características procedentes de la transformada wavelet discreta de ambas señales en diferentes niveles de detalle. Se formaron tres conjuntos de características para evaluar la complementariedad de la información de ambas señales: FA, SpO_2 y FA+SpO_2. Finalmente, se evaluó la capacidad diagnóstica de ambas señales por separado y de manera conjunta mediante modelos de regresión logística entrenados para detectar la AOS en los umbrales del índice de apnea-hipopnea de 1, 5 y 10 eventos/h. Los resultados logrados en el conjunto de test fueron superiores en el caso del conjunto FA+SpO_2, alcanzando precisiones del 68.72%, 82.31% y 88.72% en los umbrales de 1, 5 y 10 eventos/h, respectivamente. Los resultados obtenidos al combinar las características de ambas señales sugieren que el análisis wavelet de las señales de FA y SpO_2 de manera conjunta puede contribuir a simplificar el diagnóstico de la AOS infantil.

1. Introducción

La apnea obstructiva del sueño (AOS) se caracteriza por la recurrencia de eventos de apnea e hipopnea durante el sueño. Cuando este trastorno respiratorio se manifiesta en los niños, puede provocar déficits neurocognitivos y diferentes problemas en el desarrollo, el crecimiento y la conducta [1], [2]. La AOS infantil es una enfermedad que afecta hasta a un 5% de los niños en todo el mundo [1]. La polisomnografía nocturna (PSG) es la prueba de referencia para diagnosticarla [2]. Mediante la PSG, el sueño de los niños se analiza usando el registro de múltiples señales biomédicas como electrocardiograma, electroencefalograma, flujo aéreo (FA), saturación de oxígeno en sangre (SpO_2), etc. El índice de apnea-hipopnea (IAH) determina la presencia y severidad de la AOS como la cantidad de eventos de apnea y/o hipopnea registrados por cada hora de sueño (e/h) [2]. No obstante, la PSG es una prueba compleja, con un coste elevado y la disponibilidad de unidades del sueño en los hospitales es baja. Además, es incómoda para los niños debido al elevado número de sensores instalados [3]. Las alternativas más aceptadas se han centrado en limitar el número de señales, reduciendo así la complejidad e intrusividad de la prueba [3], [4]. Los eventos de apnea e hipopnea se definen principalmente a partir de las señales de FA y SpO_2 [5], por lo que el análisis de estas señales puede contribuir a la simplificación del diagnóstico y a la comodidad del diagnóstico de la AOS en los niños. Finalmente, la información extraída de estas señales puede ser útil para el desarrollo de métodos automáticos de ayuda al diagnóstico de la AOS.

En estudios previos se han propuesto diversos métodos de análisis automático de las señales de FA y SpO_2 [6]–[10]. Muchos de estos trabajos incluyeron el análisis de variables clínicas obtenidas de la señal de SpO_2, como el índice de desaturaciones de oxígeno del 3% (ODI 3%) [8]–[10]. Se ha mostrado en estudios previos que el ODI 3% proporciona información relevante sobre la presencia de la AOS en los niños, pero que puede no ser suficiente para diagnosticar la AOS infantil con precisión [11], [12]. Varios métodos de análisis automático de las señales de FA y SpO_2 se han centrado en parámetros estadísticos en el dominio temporal, no lineales y en el análisis espectral [8], [12], [13]. Sin embargo, el análisis espectral está limitado por una resolución temporal y espectral fija que dificulta el análisis de estas señales en bandas de baja frecuencia. En estas bandas se encuentra información relacionada con la recurrencia y duración de los eventos de apnea e hipopnea [10], [12]. La transformada *wavelet* (WT, *Wavelet Transform*) proporciona buena resolución espectral en frecuencias bajas, a la vez que ofrece buena resolución temporal en frecuencias altas, por lo que resulta adecuada para el análisis de las señales de FA y SpO_2 [13]. Hasta donde tenemos conocimiento, la WT se ha propuesto en estudios previos sobre la AOS infantil empleando solo la señal de SpO_2 [13]. En este estudio se propone el análisis conjunto de las señales de FA y SpO_2 mediante la WT para detectar la AOS en niños.

ISBN: 978-84-09-25491-0, pags. 315- 318

Por lo tanto, la hipótesis de partida de este estudio es que el análisis de las señales de FA y SpO_2 mediante la WT proporciona información valiosa y complementaria al ODI 3% para diagnosticar la AOS. Este estudio se centra en evaluar la utilidad de la información obtenida de la WT de las señales de FA y SpO_2 para detectar la AOS en niños de manera conjunta con el ODI 3%. Para lograr este objetivo se ha extraído y seleccionado la información más relevante y no redundante de ambas señales y se ha evaluado el rendimiento diagnóstico de las mismas mediante regresión logística (RL).

2. Base de datos

La base de datos utilizada en este estudio está formada por las señales de FA y SpO_2 procedentes de la PSG de 974 sujetos pediátricos. Todos los sujetos fueron evaluados mediante PSG en la unidad del sueño del *Comer's Children's Hospital, University of Chicago Medicine* (Chicago, IL, EEUU) ante la sospecha de sufrir AOS. El Comité Ético del hospital aprobó el protocolo de investigación y los tutores legales de todos los sujetos proporcionaron el consentimiento informado. Se emplearon las reglas de la *American Academy of Sleep Medicine* (AASM) para la evaluación de cada uno de los sujetos bajo estudio [5]. Se han utilizado los umbrales de IAH de 1, 5 y 10 e/h para establecer tres niveles de severidad de la AOS: leve, moderado y severo [4], [8]. En la Tabla 1 se muestran los datos sociodemográficos y clínicos de los sujetos. Los sujetos de la base de datos se dividieron aleatoriamente en los conjuntos de entrenamiento (60%) y test (40%). No se encontraron diferencias estadísticamente significativas (p>0.01, test U de Mann-Whitney) en edad, sexo, índice de masa corporal (IMC) e IAH entre ambos conjuntos.

3. Métodos

La frecuencia de muestreo (f_s) de las señales de FA fue 100 Hz y en las señales de SpO_2, 25 Hz, según la recomendación de la AASM [5]. Para evaluar la complementariedad de la información de las señales de FA y SpO_2, se evaluaron tres configuraciones: empleando cada señal por separado y empleando ambas señales de manera conjunta.

	Todos	Entrenamiento	Test
Sujetos (n)	974	584	390
Edad[(1)] (años)	6.07±3.43	6.08±3.43	6.06±3.45
Sexo (M/F)	599/375	346/238	253/137
IMC[(1)] (kg/m^2)	20.21±6.63	20.28±6.88	20.11±6.25
IAH[(1)] (e/h)	9.34±17.24	10.32±19.61	7.88±12.79
No AOS (n)	171	96	75
AOS Leve (n)	398	229	169
AOS Moderado (n)	176	113	63
AOS Severo (n)	229	146	83

[(1)]: Edad, IMC e IAH expresados como media ± desviación estándar.

Tabla 1. *Datos sociodemográficos y clínicos de los sujetos.*

3.1. Extracción de características

Se empleó la WT discreta (DWT, *Discrete Wavelet Transform*) debido a su mayor eficiencia computacional frente a la WT continua [14]. El algoritmo de la DWT calcula la WT de la señal de entrada A_0 únicamente en escalas potencia de 2 de manera recursiva. En cada nivel de descomposición i, la señal de entrada A_{i-1} se descompone en dos señales: aproximación (A_i) y detalle (D_i), según una función base llamada *wavelet* madre [14]. Los coeficientes D_i se obtienen tras un filtrado paso alto cuya respuesta al impulso, $g[n]$, es la *wavelet* madre. Los coeficientes A_i se obtienen mediante un filtrado paso bajo $h[n]$, relacionado con $g[n]$ mediante [14]:

$$g[L-1-n] = (-1)^n \cdot h[n] \quad (1)$$

Las señales A_i y D_i finales se obtienen tras diezmar las señales filtradas con un factor 2. Los coeficientes A_i y D_i se obtienen tras aplicar el mismo proceso a la señal A_{i-1} obtenida en la iteración anterior [14]:

$$A_i[k] = \sum_n A_{i-1}[n] \cdot h[2k-n] \quad (2)$$

$$D_i[k] = \sum_n A_{i-1}[n] \cdot g[2k-n] \quad (3)$$

Se calculó la DWT de segmentos de 2^{16} y 2^{14} muestras para las señales de FA y SpO_2 respectivamente. El nivel de descomposición más alto fue 16 en las señales de FA y 14 en las señales de SpO_2 [13]. Las *wavelets* madre empleadas fueron *'Daubechies5'* (FA) y *'Haar'* (SpO_2). Estas *wavelets* madre se han empleado con éxito en estudios previos con las señales de FA y SpO_2 [13], [15]. Tras calcular la DWT de las señales, se extrajeron características procedentes de los coeficientes D_8 (0.1953 Hz - 0.3906 Hz) y D_9 (0.0977 - 0.1953 Hz) de la señal de FA y de D_9 (0.0244 - 0.0488 Hz) en la señal de SpO_2. Estas bandas de frecuencia se corresponden con las duraciones habituales de los ciclos respiratorios ($D8_{FA}$), los eventos apneicos ($D9_{FA}$) y las desaturaciones ($D9_{SpO2}$) [10], [12]. Concretamente, se calcularon los momentos estadísticos de primer al cuarto orden (*M1-M4*), máximo (*Max*) y energía (*Ene*) de los coeficientes. Se calculó además la entropía wavelet (*WE*), una medida de irregularidad de la DWT a través de la entropía de la distribución de *Ene* a lo largo de los niveles de detalle D_i [13]. Por último, se obtuvo el ODI 3% a partir de la señal de SpO_2 como la cantidad de desaturaciones del nivel de oxígeno en sangre inferiores o iguales al 3% con respecto al nivel de base por cada hora de registro [10]. Se estudió la relación de las características extraídas con el IAH empleando el coeficiente de correlación de Spearman (ρ_S).

3.2. Selección de características

Tras la fase de extracción de características se obtuvieron 13 características de la señal de FA y 8 características de la señal de SpO_2. Se aplicó el método *Fast Correlation Based Filter* (FCBF) para seleccionar las características más relevantes, descartando las que sean redundantes entre ellas [16]. FCBF funciona en dos fases: (1) se eligen las características más relevantes y (2) se descartan las que sean redundantes con respecto a una de mayor relevancia [16]. Se aplicó el método FCBF sobre 1000 réplicas

bootstrap obtenidas mediante muestreo con reemplazo de los datos del conjunto de entrenamiento [8], [13]. Las características seleccionadas al menos 500 veces fueron finalmente seleccionadas. El proceso de selección se aplicó en tres conjuntos de características para evaluar su complementariedad: (1) procedentes de la señal de FA, (2) de la señal de SpO_2 y (3) ambas señales (FA+SpO_2).

3.3. Reconocimiento de patrones

Las características seleccionadas de los conjuntos de características descritos anteriormente fueron clasificados empleando regresión logística (RL) [17]. Este enfoque de clasificación binaria se ha utilizado ampliamente en el contexto de la ayuda al diagnóstico de la AOS pediátrica [7], [9], [10]. El algoritmo LR estima la probabilidad a posteriori de clasificar correctamente un sujeto como positivo a partir de una función logística dependiente de las variables de entrada. Para estimar los coeficientes óptimos del modelo, RL emplea el método de máxima verosimilitud [17]. Tras obtener los modelos de RL en el conjunto de entrenamiento, se evaluaron sobre el conjunto de test. La capacidad diagnóstica de los modelos para detectar la presencia de AOS empleando los umbrales de 1, 5 y 10 e/h se evaluó en términos de sensibilidad (*Se*), especificidad (*Sp*), precisión (*Acc*), valor predictivo positivo (*VPP*) y valor predictivo negativo (*VPN*).

4. Resultados

En la Tabla 2 se muestran los valores de correlaciones (ρ_S) de las características extraídas en este estudio. Las correlaciones con el IAH más elevadas se obtuvieron con las características obtenidas de la señal de SpO_2, siendo el ODI 3% la que mayor ρ_S alcanzó. Las correlaciones con el IAH de las características de la señal de FA más destacables fueron las de los coeficientes D_8. Tras la fase de selección de características, las características escogidas de la señal de FA fueron *M3-D8* y *M2-D9*, mientras que las que se seleccionaron de la señal de SpO_2 fueron ODI 3% y *M3-D9*. Al combinar todas las características en el conjunto FA+SpO_2, las características seleccionadas fueron 3: $M3\text{-}D8_{FA}$, ODI 3% y $M3\text{-}D9_{SpO2}$.

Los resultados de clasificación binaria empleando RL se muestran en la Tabla 3. Para cada uno de los umbrales, la mayor *Acc* se consiguió empleando el conjunto de características FA+SpO_2. Además, los valores de *Acc* fueron más elevados para los umbrales de IAH superiores.

5. Discusión y conclusiones

En este estudio se ha evaluado la capacidad diagnóstica de las señales de FA y SpO_2 de manera individual y conjunta empleando la DWT. Para lograrlo, se han desarrollado y aplicado modelos de RL sobre características de las señales de FA y SpO_2 procedentes de la DWT de ambas señales. Estudios previos han combinado la información procedente de la señal de FA con el ODI 3% y con características espectrales y no lineales extraídas de la señal de SpO_2 [11], [12]. La aplicación de la WT a las señales de FA y SpO_2 para la ayuda al diagnóstico de la AOS infantil es una contribución novedosa. Únicamente un estudio previo ha aplicado la DWT sobre la señal de SpO_2 en este contexto y, hasta donde tenemos conocimiento, no se ha analizado la DWT de las señales de FA y SpO_2 de manera conjunta.

	D_8 FA	D_9 FA.	D_9 SpO_2
M1	-0.3426	0.0826*	**0.6632**
M2	-0.2335	0.1428	**0.6377**
M3	**0.4057**	-0.0602*	-0.3483
M4	0.3369	-0.0968*	**-0.4476**
Max	-0.1039*	0.1312	**0.6555**
Ene	0.3154	0.1049*	**0.6645**
	WE FA	*WE* SpO_2	*ODI 3%*
	0.2741	**0.6213**	**0.6918**

*: Correlación no significativa (p>0.01).

Tabla 2. *Coeficientes de correlación de Spearman de las características extraídas con el IAH*

	Señal	*Se*	*Sp*	*Acc*	*VPP*	*VPN*
1 e/h	FA	66.35	54.67	64.10	86.01	27.89
	SpO_2	56.83	86.67	62.56	94.71	32.34
	FA+SpO_2	65.71	81.33	68.72	93.67	36.09
5 e/h	FA	68.49	61.48	64.10	51.55	76.53
	SpO_2	78.77	80.33	79.74	70.55	86.34
	FA+SpO_2	75.34	86.48	82.31	76.92	85.43
10 e/h	FA	59.04	70.03	67.69	34.75	86.35
	SpO_2	83.13	88.93	87.69	66.99	95.12
	FA+SpO_2	79.52	91.21	88.72	70.97	94.28

Tabla 3. *Resultados de clasificación empleando regresión logística. Valores expresados en porcentaje.*

Las correlaciones con el IAH de las características obtenidas a partir de la DWT de las señales de FA y SpO_2 revelan que estas pueden proporcionar información valiosa para diagnosticar la AOS infantil. No obstante, la correlación más elevada se logró con el ODI 3%. Tras la fase de selección de características, es destacable que se seleccionaron características de ambas señales. Este resultado indicaría que la información de la DWT es complementaria entre señales y con el ODI 3%, una característica útil en el contexto de la AOS infantil. En cuanto a la capacidad diagnóstica de los modelos de RL, destaca que el conjunto FA+SpO_2 obtuvo la mayor *Acc* en los tres umbrales del IAH. En el umbral de 1 e/h, los pares *Se-Sp* y *VPP-VPN* fueron más equilibrados al combinar ambas señales. Además, se consiguieron valores más elevados de *Sp* y *VPP* en los umbrales de 5 y 10 e/h con respecto al conjunto SpO_2. Estos resultados sugieren que la información complementaria de las señales de FA y SpO_2 contribuye a incrementar su capacidad diagnóstica de manera conjunta.

Estos resultados son similares a los logrados en trabajos previos. Garde *et al.* obtuvieron *Acc* 71.0% (1 e/h), 78.0% (5 e/h) y 88.0% (10 e/h) empleando RL sobre señales de pulsioximetría [7]. Los resultados del presente estudio son superiores en 5 y 10 e/h, aunque inferiores en 1e/h. En Hornero *et al.* se aplicó una red neuronal alimentada por

características de SpO_2, logrando *Acc* 75.2% (1 e/h), 81.7% (5 e/h) y 90.2% (10 e/h) [8]. En el estudio de Vaquerizo-Villar *et al.* se combinaron el análisis espectral y *wavelet* de la señal de SpO_2 con el ODI 3%, logrando *Acc* 84.0% (5 e/h) [13], superior a la lograda en el presente estudio. Trabajos recientes han combinado la información de FA con el ODI 3%, obteniendo *Acc* 78.5% (5 e/h), 91.0% (10 e/h) y *Acc* 82.1% (5 e/h), 90.3% (10 e/h) empleando redes neuronales y *ensemble learning*, respectivamente [11], [12]. En este caso, los resultados del presente estudio son ligeramente superiores en 5 e/h.

Es necesario indicar las limitaciones de este estudio. En primer lugar, se ha estudiado la DWT de las señales de FA y SpO_2. En estudios futuros se podría combinar la información de la DWT con otros parámetros espectrales y/o no lineales. Por otra parte, se ha empleado RL en la etapa de reconocimiento de patrones. Aunque el objetivo era evaluar la utilidad y la complementariedad de la DWT, los resultados son moderados en comparación con el estado del arte. Podría ser interesante evaluar el rendimiento diagnóstico de clasificadores más avanzados como los basados en *ensemble learning*.

Como conclusión, el análisis conjunto de las señales de FA y SpO_2 mediante la WT reveló la complementariedad de ambas señales para diagnosticar la AOS en niños. El rendimiento diagnóstico de los modelos de RL entrenados a partir de la información extraída de ambas señales fue superior al logrado utilizando una señal en los tres umbrales de IAH. De esta forma, el análisis *wavelet* de estas señales sería útil para desarrollar métodos de ayuda al diagnóstico de la AOS simplificados.

Agradecimientos

Este estudio ha sido financiado por el Ministerio de Ciencia, Innovación y Universidades - Agencia Estatal de Investigación y FEDER (DPI2017-84280-R, RTC-2017-6516-1), por la Comisión Europea y FEDER con el proyecto 'Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres' ('Interreg V-A Spain-Portugal POCTEP 2014–2020'), y por CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto de Salud Carlos III, cofinanciado con fondos FEDER. J. Jiménez García cuenta con un contrato predoctoral concedido por la Universidad de Valladolid. A. Martín Montero cuenta con un contrato predoctoral para la Formación de Personal Investigador (PRE2018-085219) del Ministerio de Ciencia, Innovación y Universidades.

Referencias

[1] C. L. Marcus *et al.*, "Diagnosis and Management of Childhood Obstructive Sleep Apnea Syndrome," *Pediatrics*, vol. 130, no. 3, pp. e714–e755, Sep. 2012.

[2] A. G. Kaditis *et al.*, "Obstructive sleep disordered breathing in 2- to 18-year-old children: diagnosis and management," *Eur. Respir. J.*, vol. 47, no. 1, pp. 69–94, Jan. 2016.

[3] H.-L. Tan *et al.*, "Pediatric Home Sleep Apnea Testing," *Chest*, vol. 148, no. 6, pp. 1382–1395, Dec. 2015.

[4] H.-L. Tan *et al.*, "Overnight Polysomnography versus Respiratory Polygraphy in the Diagnosis of Pediatric Obstructive Sleep Apnea," *Sleep*, vol. 37, no. 2, pp. 255–260, 2014.

[5] 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. 8, no. 5, pp. 597–619, Oct. 2012.

[6] A. Garde *et al.*, "Development of a Screening Tool for Sleep Disordered Breathing in Children Using the Phone Oximeter™," *PLoS One*, vol. 9, no. 11, p. e112959, Nov. 2014.

[7] A. Garde *et al.*, "Pediatric pulse oximetry-based OSA screening at different thresholds of the apnea-hypopnea index with an expression of uncertainty for inconclusive classifications," *Sleep Med.*, vol. 60, pp. 45–52, Aug. 2019.

[8] R. Hornero *et al.*, "Nocturnal Oximetry–based Evaluation of Habitually Snoring Children," *Am. J. Respir. Crit. Care Med.*, vol. 196, no. 12, pp. 1591–1598, Dec. 2017.

[9] A. Crespo *et al.*, "Assessment of oximetry-based statistical classifiers as simplified screening tools in the management of childhood obstructive sleep apnea," *Sleep Breath.*, vol. 22, no. 4, pp. 1063–1073, Dec. 2018.

[10] D. Álvarez *et al.*, "Automated Screening of Children With Obstructive Sleep Apnea Using Nocturnal Oximetry: An Alternative to Respiratory Polygraphy in Unattended Settings," *J. Clin. Sleep Med.*, vol. 13, no. 05, pp. 693–702, May 2017.

[11] V. Barroso-García *et al.*, "Usefulness of recurrence plots from airflow recordings to aid in paediatric sleep apnoea diagnosis," *Comput. Methods Programs Biomed.*, vol. 183, p. 105083, 2020.

[12] 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, p. 670, Jun. 2020.

[13] F. Vaquerizo-Villar *et al.*, "Wavelet analysis of oximetry recordings to assist in the automated detection of moderate-to-severe pediatric sleep apnea-hypopnea syndrome," *PLoS One*, vol. 13, no. 12, p. e0208502, Dec. 2018.

[14] O. Rioul *et al.*, "Wavelets and signal processing," *IEEE Signal Process. Mag.*, vol. 8, no. 4, pp. 14–38, Oct. 1991.

[15] J. Lee *et al.*, "Classification of healthy and abnormal swallows based on accelerometry and nasal airflow signals," *Artif. Intell. Med.*, vol. 52, no. 1, pp. 17–25, May 2011.

[16] L. Yu *et al.*, "Efficient Feature Selection via Analysis of Relevance and Redundancy," *J. Mach. Learn. Res.*, vol. 5, pp. 1205–1224, 2004.

[17] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.

Análisis de la variabilidad temporal de la onda P para predecir el resultado de la ablación por catéter en fibrilación auricular paroxística

A. Ruiz Moreno[1], M. A. Arias Palomares[2], A. Puchol Calderón[2], M. I. Pachón Iglesias[2], J. J. Rieta Ibáñez[3], R. Alcaraz Martínez[1]

[1] Grupo de Investigación en Ingeniería Electrónica, Biomédica y de Telecomunicación, Universidad de Castilla-La Mancha, Cuenca, España, {antonio.ruiz, raul.alcaraz}@uclm.es
[2] Servicio de Cardiología, Hospital Virgen de la Salud, Toledo, España
[3] BioMIT.org, Departamento Ingeniería Electrónica, Universitat Politecnica de Valencia, Valencia, España, jjrieta@upv.es

Resumen

El aislamiento de las venas pulmonares es el principal tratamiento de la fibrilación auricular (FA) paroxística. Sin embargo, presenta una alta tasa de recurrencia y el desarrollo de predictores no invasivos de esta cuestión, es esencial para permitir un tratamiento más personalizado de cada paciente. Para tal propósito, se han analizado varias características de la onda P a partir de registros de electrocardiografía (ECG), pero, aún no se ha estudiado su variabilidad a lo largo del tiempo. Por tanto, este trabajo explora la capacidad de la variabilidad temporal de la duración y la amplitud de la onda P, para anticipar el fallo de la crioablación. Paea este estudio antes de la ablación, se registró un ECG de 5 minutos en 45 pacientes con FA paroxística. Las ondas P en la derivación II se delinearon con un algoritmo automático y se calcularon la duración y amplitud de todas ellas. A continuación, se obtuvieron la media, la desviación estándar y el coeficiente de variación (CV) para ambos parámetros y se correlacionaron con el resultado de la ablación después del seguimiento de 9 meses. El CV para ambas características de las ondas P presentó la capacidad predictiva más alta de la recurrencia de la FA. Además, obtuvo mejoras de alrededor del 10 % en la clasificación con respecto a la media, alcanzando valores del 70 %. Estos resultados sugieren que la variabilidad temporal de la onda puede proporcionar nueva información sobre la heterogeneidad de la conducción auricular, lo cual podría ser útil para mejorar la estrategia de selección de candidatos para ser sometidos a crioablación.

1. Introducción

La fibrilación auricular (FA) es una de las arritmias más comunes en la población anciana y se asocia a complicaciones graves, como la insuficiencia cardíaca o el accidente cerebrovascular [1]. Aunque los mecanismos que desencadenan y apoyan estas arritmia todavía no se comprenden bien [2], se ha demostrado que los músculos auriculares que rodean las venas pulmonares (VPs) son la fuente más frecuente de actividad auricular focal que inicia la arritmia [3]. Por tanto, en la actualidad, la primera línea en el tratamiento de la FA paroxística sintomática consiste en el aislamiento de las VPs mediante ablación por catéter [3]. El procedimiento más común para ese propósito es la ablación por radiofrecuencia, que aplica corriente punto a punto en el tejido auricular para lograr la necrosis celular mediante el calentamiento del tejido [4]. Sin embargo, en los últimos años la crioablación está recibiendo una atención creciente porque es significativamente más rápida y menos compleja, y además ha reportado tasas de eficacia a largo plazo muy similares [5]. En este caso, la energía criogénica se aplica con un balón de un solo paso, lo que conduce a la necrosis por congelación [4].

Inicialmente, ambos tipos de ablaciones con catéter son muy efectivos, restaurando el ritmo sinusal (RS) en más del 80 % de los pacientes [6]. Sin embargo, en ambos casos la recurrencia de la FA es común durante el período de seguimiento, y alrededor del 30-40 % de los pacientes requieren uno o varios procedimientos adicionales durante el primer año [6]. Este resultado ocurre incluso con el uso de fármacos antiarrítmicos potentes, así como con el aislamiento total de las VPs [5]. Por tanto, además de la reconexión de las VPs, otras causas de recurrencia de la FA incluyen la existencia de focos extra-VP y un remodelado avanzado del sustrato auricular [7]. En este contexto, el desarrollo de predictores del fracaso del procedimiento es esencial para un tratamiento más personalizado de los pacientes y, por tanto, para hacer más eficiente la gestión de los mismos [5].

Con tal propósito, en las últimas dos décadas se han estudiado y presentado varios marcadores ecocardiográficos, tales como el diámetro auricular izquierdo, la velocidad de conducción auricular, etc. [5]. Sin embargo, desde un punto de vista clínico, no han sido suficientemente predictivos, y muchos autores se han centrado en analizar la onda P reflejada en el electrocardiograma (ECG) [5]. De hecho, el remodelado del sustrato auricular a menudo ofrece como resultado retrasos en la conducción, los cuales se reflejan

ISBN: 978-84-09-25491-0, pags. 319- 322

en una onda P prolongada [5]. Además, la duración de esta y otras de sus características típicas, incluyendo la dispersión, la morfología y el tiempo de pico (entre otras) han proporcionado una capacidad de predicción del resultado de la ablación con catéter de entre el 60 y el 70 % [5]. Asimismo, varios autores han corroborado que algunas características de la onda P se alteran significativamente tanto después de ablaciones por radiofrecuencia como por ablación con criobalón [8].

Sin embargo, es posible que la presencia de un remodelado estructural auricular leve o moderado no cambie de forma global y permanente la morfología de la onda P, y sus características más comunes, por lo que recientemente se ha sugerido el estudio de cómo esta onda cambia con el tiempo [9]. Así pues, el objetivo principal del presente trabajo es explorar la capacidad de la variabilidad temporal de dos parámetros comunes de la onda P, tales como la duración y la amplitud, para estimar el remodelado auricular, así como para anticipar el fracaso de la crioablación en pacientes que padecen FA paroxística.

2. Método

2.1. Población bajo estudio

La base de datos utilizada en este trabajo consistió en 45 pacientes (14 mujeres y 31 hombres) con FA paroxística sometidos a crioablación con catéter en el Hospital Universitario de Toledo. Brevemente, los pacientes fueron primero sedados con anestesia general o sedación consciente. A continuación, se utilizó un acceso venoso femoral para introducir los catéteres y el acceso a la aurícula izquierda se logró mediante punción transeptal. La anticoagulación se mantuvo a través de un bolo de heparina inicial, y la administración de otros adicionales de acuerdo a la monitorización de la coagulación durante el procedimiento. Finalmente, se realizaron una o varias crioaplicaciones para aislar cada VP. El procedimiento finalizó cuando todas las VP se aislaron con éxito. Después de la intervención, los sujetos se monitorizaron durante algunas horas y, si no hubo complicaciones, fueron dados de alta.

El procedimiento fue inicialmente exitoso en todos los pacientes, quienes no presentaron ninguna complicación durante un seguimiento de nueve meses. Transcurrido este tiempo, 31 sujetos mantuvieron el ritmo sinusal y los 14 restantes recayeron a FA. Todos los pacientes recibieron anticoagulantes y fármacos antiarrítmicos según el criterio del personal clínico.

2.2. Adquisición y preprocesado de la señal de ECG

Se registró una señal de ECG estándar de 12 derivaciones de forma continua durante 5 minutos justo antes de la ablación de cada paciente. Los registros se adquirieron con una frecuencia de muestreo de 1 kHz y 16 bits de resolución. Solo se analizó la derivación II, porque la onda P presenta una mayor amplitud que en las señales restantes y la mayoría de los trabajos anteriores la han recomendado para evaluar la morfología de las ondas P [10]. A continuación, esta señal se preprocesó para eliminar la linea base, la interferencia de la red eléctrica y el ruido de alta frecuencia. Más concretamente, la línea base se estimó utilizando un filtrado paso bajo con una frecuencia de corte de 0,8 Hz y se restó de la señal original [11]. Posteriormente, la interferencia de la red eléctrica se eliminó mediante un algoritmo basado en la transformada wavelet [12]. Este método también fue capaz de eliminar la mayor parte del ruido de alta frecuencia, pero, aún así se utilizó un filtrado paso bajo con una frecuencia de corte de 70 Hz para obtener una señal lo más limpia posible [11].

2.3. Obtención de los puntos fiduciales de la onda P

Para identificar los puntos fiduciales (es decir, el inicio, el pico y el fin) de las ondas P se utilizó un algoritmo publicado anteriormente [13]. Este algoritmo se ha validado haciendo uso de varias bases de datos anotadas manualmente por cardiólogos expertos, proporcionando una sensibilidad superior al 94 % . Además, el algoritmo ha podido delinear la onda P con errores de ubicación inferiores a 12 ms con respecto a las anotaciones manuales realizadas por expertos [14]. Brevemente, el método se basa en la generación de un modelo gaussiano de cada onda P para ayudar a su delineación y un umbral de pendiente adaptativo que tiene en cuenta la morfología de las ondas anteriores. No obstante, cabe señalar que la delimitación de todas las ondas P se supervisó visualmente y los errores claros se corrigieron manualmente.

A partir de los puntos fiduciales obtenidos para cada onda P, se calcularon su duración y su amplitud. Luego, se obtuvieron la media, la desviación estándar (std) y el coeficiente de variación (CV), definido como la relación entre la std y la media, de las series temporales resultantes para ambos parámetros y se correlacionaron con el resultado de la ablación.

2.4. Evaluación del rendimiento

Los resultados se expresan como la media $\pm$ la std para todos los índices propuestos para anticipar el resultado de la ablación por catéter. Además, las diferencias estadísticas entre ellos se analizaron mediante una prueba U de Mann-Whitney. Se consideró un valor $p < 0,05$ como estadísticamente significativo.

Por otro lado, se evaluó la capacidad de cada característica para discriminar entre pacientes que recayeron a FA y mantuvieron RS al final del seguimiento mediante una curva de característica operativa del receptor (ROC). La sensibilidad (Se) se consideró como el porcentaje de pacientes que recayeron en FA correctamente clasificados. Asimismo, se consideró como especificidad (Sp) la tasa de pacientes que mantenían el RS. Se seleccionó el umbral óptimo como aquel que proporcionó el mayor porcentaje de pacientes correctamente clasificados, es decir, el que maximizó la exactitud (Acc).

3. Resultados

Los valores medios y de std para los tres parámetros calculados a partir de la series temporales de duración y amplitud de la onda P se muestran en la Tabla 1. Como puede

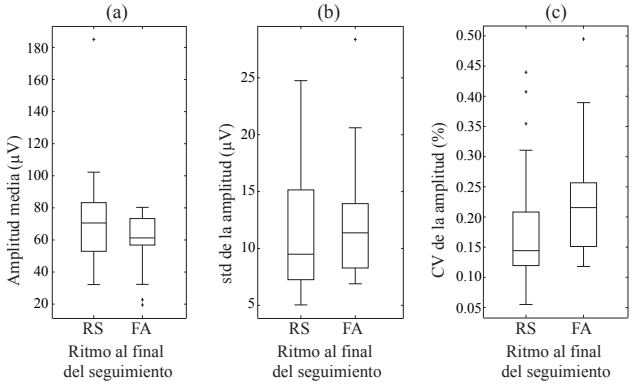


Figura 2. *Diagrama de cajas y bigotes de las distribuciones de los valores de (a) media, (b) std y (c) CV calculados para la amplitud de la onda P.*

Onda P	Parámetro	Grupo de pacientes que mantienen un RS	Grupo de pacientes que recaen en FA	p-valor
Duración	Media (ms)	154.7±21.4	156.3±30.7	0.607
	Std (ms)	15.1±5.3	21.4± 9.6	0.020
	CV (%)	9.7± 3.0	13.4±4.7	0.003
Amplitud	Media (μV)	71.7±28.3	57.9±19.6	0.211
	Std (μV)	11.3± 0.5	12.4±0.6	0.447
	CV (%)	17.3±9.5	23.4±11.2	0.035

Tabla 1. *Media y std para los valores de media, std y CV calculados a partir de las series temporales de duración y amplitud de la onda P.*

verse, el valor medio de duración de la onda P fue mayor en los pacientes que recayeron a FA que en los que mantuvieron el RS durante el seguimiento, pero no se observaron diferencias estadísticamente significativas. De hecho, la Figura 1 (a) muestra datos completamente superpuestos para ambos grupos de pacientes. Por el contrario, se observaron diferencias estadísticamente significativas entre ambos grupos de pacientes tanto para los valores de std como de CV obtenidos a partir de la serie temporal de duración de la onda P, mostrando también valores mayores para los pacientes recayeron a FA que para los que mantuvieron un RS, como las Figuras 1 (b) y (c) muestran.

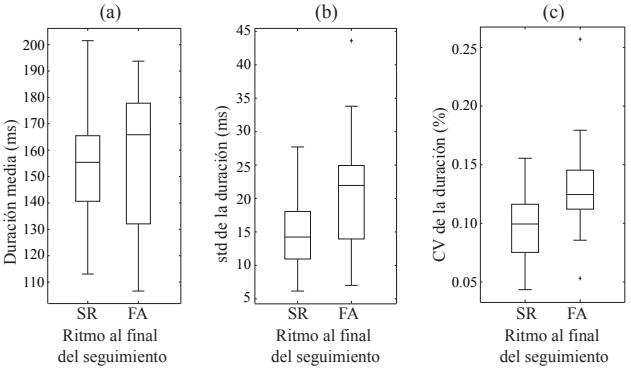


Figura 1. *Diagrama de cajas y bigotes de las distribuciones de los valores de (a) media, (b) std y (c) CV calculados para la duración de la onda P.*

En cuanto a la serie temporal de la amplitud de la onda P, el valor medio fue mayor para los pacientes que mantuvieron el RS durante el seguimiento que para los que recayeron a FA. Sin embargo, como muestra la Figura 2 (a), los datos se superpusieron para ambos grupos y no se notaron diferencias estadísticamente significativas. De manera similar, la Figura 2 (b) también muestra cómo los valores estándar se superpusieron principalmente para ambos grupos y tampoco se observó una separación estadística. Finalmente, en la Figura 2 (c) se observó una mayor separación entre los dos grupos de pacientes para los valores del CV, siendo los calculados para los pacientes que recayeron a FA mayores que los obtenidos para los pacientes que mantenían el RS. De hecho, en este caso se obtuvo un valor de p ligeramente inferior a 0,05.

De acuerdo con estos hallazgos, los mejores resultados de clasificación fueron proporcionados por los parámetros CV y std calculados a partir de las series de tiempo de duración de la onda P, como se puede observar en la Tabla 2. Ambos índices han reportado valores de Acc alrededor del 73 y 71 %, respectivamente, con valores bien equilibrados de Se y Sp. El siguiente mejor resultado se logró mediante el CV obtenido a partir de la serie temporal de amplitud de la onda P. Sin embargo, en este caso los valores de Acc, Se y Sp fueron notablemente más bajos, en torno al 65 %. Los demás parámetros presentaron valores de Acc iguales o inferiores al 60 %.

Onda P	Parámetro	Se (%)	Sp (%)	Acc (%)
Duración	Media	57.1	61.4	60.0
	Std	71.4	71.0	71.1
	CV	71.4	74.2	73.3
Amplitud	Media	64.3	52.6	56.6
	Std	57.1	54.8	55.6
	CV	64.3	64.5	64.4

Tabla 2. *Resultados de clasificación obtenidos por los parámetros calculados a partir de las series temporales de duración y amplitud de la onda P.*

4. Discusión

En algunos trabajos previos se ha identificado una duración prolongada de la onda P antes de la ablación como un predictor de recurrencia de la FA [5]. Más concretamente, puntos de corte entre 120 y 150 ms para la duración de la onda P han sido propuestos para clasificar pacientes con FA recurrente y los que mantenían RS durante su seguimiento, con valores de sensibilidad y especificidad alrededor del 70 % [5]. De manera similar, una amplitud de onda P baja antes de la ablación también se ha asociado con una alta probabilidad de recurrencia de la arritmia después de la ablación por catéter y un valor de aproximadamente de 0,1 mV ha alcanzado una sensibilidad del 70 % [15] . En comparación con estos resultados, en el presente trabajo los valores medios obtenidos de las series temporales de duración y amplitud de la onda P han mostrado tasas de clasificación alrededor de un 10-15 % más bajas. Esta discrepancia entre los resultados podrían explicarse por los diferentes enfoques utilizados para estimar las características de la onda P. Mientras que en trabajos previos la duración y la amplitud de la onda P se estimaron a partir de señales de ECG promediadas o de registros de ECG muy cortos de 12 derivaciones [5, 15], en el presente ambos parámetros se han calculado onda a onda en señales de 5 minutos de duración y finalmente se han promediado.

A pesar de estos resultados y de que los valores medios de

duración y amplitud de la onda P no proporcionaron diferencias estadísticamente significativas entre los grupos, se observó una tendencia hacia ondas P con mayor duración y menor amplitud para los sujetos que recayeron a FA durante los primeros meses después de la intervención. Este hallazgo concuerda con la mayoría de trabajos previos [5, 15] y sugiere que la existencia de un remodelado estructural auricular más grande antes de la ablación aumenta la probabilidad de desarrollar FA después del procedimiento, incluso cuando todas las VP se encuentran aisladas correctamente.

No obstante, es interesante destacar que los parámetros std y CV obtenidos de las series temporales tanto de duración como de amplitud de la onda P han presentado una mejor separación entre grupos de pacientes que la media, revelando así diferencias estadísticamente significativas y una mayor capacidad discriminante. Este resultado sugiere que la variabilidad temporal de las características de la onda P puede estimar nueva información sobre el remodelado auricular. En este sentido, la reducción de la variabilidad en la duración y amplitud de la onda P podría significar una propagación más rígida e invariable de la activación auricular en pacientes sin FA recurrente, lo que implica un moderado remodelado auricular y, en consecuencia, una mayor probabilidad de responder exitosamente a la ablación con catéter. Aunque este hallazgo es muy interesante, para su confirmación todavía se requieren más estudios con bases de datos más amplias.

5. Conclusiones

En el presente trabajo se ha estudiado por primera vez la variabilidad a lo largo del tiempo de duración y la amplitud de la onda P para anticipar el resultado de la ablación por catéter con crioenergía. Los resultados han proporcionado que la desviación estándar y el coeficiente de variación de la serie temporal derivada para ambos parámetros de la onda P logran una mejor identificación de los pacientes que recaerá a FA después de varios meses de seguimiento que la media. Por lo tanto, el análisis de la variabilidad temporal de la onda P podría proporcionar nueva información sobre la heterogeneidad de la conducción auricular, lo que podría ser útil para mejorar la estrategia de selección de candidatos óptimos para la crioablación.

Agradecimientos

Esta investigación ha sido apoyada por los fondos DPI2017-83952-C3 de MINECO/AEI /FEDEREU,SBPLY/17/180501/000411 de la Junta de Comunidades de Castilla-la Mancha y AICO/2019/036 de la Generalitat Valenciana.

Referencias

[1] Lippi G, Sanchis-Gomar F, Cervellin G. Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge. Int J Stroke Jan 2020; 1747493019897870.

[2] Schotten U, Dobrev D, Platonov PG, Kottkamp H, Hindricks G. Current controversies in determining the main mechanisms of atrial fibrillation. J Intern Med May 2016; 279(5):428–38.

[3] Cheniti G, Vlachos K, Pambrun T, Hooks D, Frontera A, Takigawa. Atrial fibrillation mechanisms and implications for catheter ablation. Front Physiol 2018;9:1458.

[4] Yano M, Egami Y, Yanagawa K, Nakamura H, Matsuhiro. Comparison of myocardial injury and inflammation after pulmonary vein isolation for paroxysmal atrial fibrillation between radiofrequency catheter ablation and cryoballoon ablation. J Cardiovasc Electrophysiol Jun 2020; 31(6):1315–1322.

[5] Pranata R, Yonas E, Vania R. Prolonged P-wave duration in sinus rhythm pre-ablation is associated with atrial fibrillation recurrence after pulmonary vein isolation-a systematic review and meta-analysis. Ann Noninvasive Electrocardiol 09 2019;24(5):e12653.

[6] Inaba O, Metzner A, Rottner L, Mathew S, Lemes C, Maurer. Radiofrequency or cryoballoon ablation for index pulmonary vein isolation: What is the impact on long-term clinical outcomes after repeat ablation? J Cardiovasc Electrophysiol May 2020;31(5):1068–1074.

[7] Redfearn DP, Skanes AC, Gula LJ, Griffith MJ, Marshall HJ, Stafford PJ, Krahn AD, Yee R, Klein GJ. Noninvasive assessment of atrial substrate change after wide area circumferential ablation: a comparison with segmental pulmonary vein isolation. Ann Noninvasive Electrocardiol Oct 2007;12(4):329–37.

[8] Furniss GO, Panagopoulos D, Kanoun S, Davies EJ, Tomlinson DR, Haywood GA. The effect of atrial fibrillation ablation techniques on P-wave duration and P-wave dispersion. Heart Lung Circ Mar 2019;28(3):389–396.

[9] Huo Y, Holmqvist F, Carlson J, Gaspar T, Hindricks G, Piorkowski C, Bollmann A, Platonov PG. Variability of P-wave morphology predicts the outcome of circumferential pulmonary vein isolation in patients with recurrent atrial fibrillation. J Electrocardiol 2015;48(2):218–25.

[10] De Bacquer D, Willekens J, De Backer G. Long-term prognostic value of P-wave characteristics for the development of atrial fibrillation in subjects aged 55 to 74 years at baseline. Am J Cardiol Sep 2007;100(5):850–4.

[11] Sörnmo L, Laguna P. ECG Signal Processing in Bioelectrical Signal Processing in Cardiac and Neurological Applications. 1st edition. Elsevier Academic Press, 2005. ISBN 978-0-12-437552-9.

[12] García M, Martínez-Iniesta M, Ródenas J, Rieta JJ, Alcaraz R. A novel wavelet-based filtering strategy to remove powerline interference from electrocardiograms with atrial fibrillation. Physiol Meas 11 2018;39(11):115006.

[13] González F, Alcaraz R, Rieta JJ. Electrocardiographic P-wave delineation based on adaptive slope gaussian detection. In Computing in Cardiology. IEEE, 2017; 1–4.

[14] Ruiz A, Arias MA, Pachón MI, Langley P, Rieta JJ, Alcaraz R. Thorough assessment of a P-wave delineation algorithm through the use of diverse electrocardiographic databases. In 2019 E-Health and Bioengineering Conference (EHB). IEEE, 2019; 1–4.

[15] Park JK, Park J, Uhm JS, Joung B, Lee MH, Pak HN. Low P-wave amplitude ($<$0.1 mV) in lead I is associated with displaced inter-atrial conduction and clinical recurrence of paroxysmal atrial fibrillation after radiofrequency catheter ablation. Europace Mar 2016;18(3):384–91.

Diseño de un prototipo de manga sensorizada para el registro de electromiografía de alta densidad en el antebrazo

D. Invernon Garrido[1], J. F. Alonso López[1,2]

[1] Departamento de Ingeniería de Sistemas, Automática e Informática Industrial (ESAII), Centre de Recerca en Enginyeria Biomèdica (CREB), Universitat Politènica de Catalunya (UPC), Barcelona, España, {david.invernon, joan.francesc.alonso}@upc.edu

[2] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), España

Resumen

Este trabajo presenta el diseño, desarrollo y validación de una manga sensorizada reutilizable para el registro de electromiografía de alta densidad en el antebrazo. Permite el registro en tiempo real de la actividad de los músculos flexores y extensores, y puede ser de gran utilidad para el seguimiento y valoración objetiva de la rehabilitación en patologías neuromusculares como el ictus. En los resultados obtenidos se observa la activación diferenciada de los músculos superficiales, así como la evaluación de la fatiga muscular mediante este sensor económico y que no requiere una excesiva preparación previa de la piel para el registro de electromiografía.

1. Introducción

Las enfermedades neurológicas y neuromusculares constituyen una de las principales causas de discapacidad y pérdida de funcionalidad en las extremidades. El ictus es la primera causa de mortalidad entre las mujeres españolas y la segunda en los varones, unas 40.000 muertes de un total de 120.000 casos anualmente [1]. Únicamente el 40% de las personas que sobreviven a un ictus puede valerse por sí mismas, y un 30% de ellos presenta problemas de discapacidad asociados a parálisis, problemas de equilibrio, trastornos del habla y otros déficits cognitivos.

El plan de actuación sobre ictus en Europa 2018-2030 tiene entre sus diferentes objetivos la prevención y atención del ictus, así como la rehabilitación y la evaluación de resultados y vida después del ictus [2].

A día de hoy, los procesos de rehabilitación neuromuscular para pacientes con ictus se centran en el movimiento de las extremidades afectadas [3]. Durante los ejercicios, el terapeuta puede dirigir los movimientos que el paciente debe realizar, pero no tiene ninguna herramienta que le proporcione la información relativa a la activación muscular ni puede conocer en qué medida el esfuerzo realizado se corresponde con el movimiento real. Entre las herramientas tecnológicas más avanzadas existen dispositivos que permiten realizar estos ejercicios de manera autónoma sin la supervisión constante como los robots y los exoesqueletos, pero estos tienen un coste elevado y no suelen estar disponibles para todos los pacientes.

La activación muscular puede monitorizarse de manera no invasiva mediante la electromiografía de superficie. Las señales electromiográficas (EMG) captadas mediante electrodos sobre la piel registran la superposición de potenciales de acción causados por la despolarización-repolarización de las fibras musculares cuando son estimuladas por neuronas motoras para ejercer la contracción. La amplitud de la señal registrada en el electromiograma dependerá de la cantidad de unidades motoras reclutadas (que son las unidades mínimas de control funcional y estructural del músculo esquelético). La fuerza ejercida también condiciona la frecuencia con la que son reclutadas las unidades motoras; a más fuerza, mayor frecuencia de activación de las unidades motoras. Existen otras variables que afectan a la señal registrada como la fisionomía del paciente (grasa subdérmica, vello corporal, etc.) el material del electrodo y el contacto piel-electrodo [4].

La detección de la activación muscular puede realizarse con un simple registro bipolar siempre que la colocación de los electrodos sea adecuada. Sin embargo, las señales electromiográficas de alta densidad (HD-EMG) son más adecuadas para analizar los grupos musculares involucrados en movimientos complejos, como por ejemplo los del antebrazo, ya que se distribuyen uniformemente los electrodos sobre el área ocupada por los músculos a monitorizar, de manera que se pueden registrar sus patrones de activación temporal y espacial de manera conjunta [5].

El objetivo de este trabajo es el diseño, implementación y validación de una manga sensorizada con electrodos y reutilizable para el registro de señales HD-EMG en el antebrazo. De esta manera se pretende obtener información cuantitativa sobre el patrón de activación muscular durante los ejercicios de rehabilitación de pacientes que han sufrido un ictus incluso en los casos en los que no es posible observar el movimiento de la extremidad superior.

2. Materiales y métodos

2.1. Músculos del antebrazo

Para el diseño de la matriz de electrodos se ha tenido en cuenta tanto la disposición de los músculos como los movimientos que permiten una activación selectiva de los mismos. Los músculos del antebrazo (Figura 1) dirigen el movimiento de la mano.

Se pueden diferenciar principalmente en dos grupos:

ISBN: 978-84-09-25491-0, pags. 323- 326

- Músculos extensores, situados en la parte posterior, se encargan de los movimientos de extensión de la mano y los dedos. Aunque se agrupan principalmente en dos planos, superficial y profundo, para el diseño de la matriz de electrodos se prioriza el plano superficial, cuyas señales será más fácil registrar de manera adecuada sobre la piel mediante electromiografía de superficie. Los principales extensores que se pretende registrar son el *Extensor Carpi Ulnaris*, el *Extensor Digitorum Communis*, el *Extensor Carpi Radialis Longus* y el *Extensor Carpi Radialis Brevis*.
- Músculos flexores, situados en la parte anterior, se encargan de los movimientos de flexión de la mano y los dedos. Se agrupan en los planos superficial, intermedio y profundo. Siguiendo el mismo razonamiento que para los músculos extensores, para el diseño de la matriz de electrodos se tienen en cuenta principalmente los músculos del plano superficial. Los principales flexores que se pretende registrar son el *Flexor Carpi Ulnaris*, el *Flexor Carpi Radialis*, el *Flexor Digitorum Superficialis* y el *Flexor Pollicis Longus*.

2.2. Selección del material de los electrodos

Dado que el número de electrodos necesario para cubrir todos los músculos es elevado, se realizó un estudio previo con materiales económicos:

- Remaches con un baño de plata, de 6 mm de diámetro.
- Discos de acero inoxidable AISI 316, de 6 mm de diámetro.

Se confeccionaron matrices de 4 filas y 5 columnas con distancia entre electrodos de 1 cm, siguiendo las recomendaciones para el registro de electromiografía de superficie [6]. Las matrices se colocaron sobre el músculo *Biceps Brachii* de una decena de voluntarios reclutados en el propio laboratorio, todos ellos sin desórdenes neuromusculares (rango de edad entre 21 y 35 años) que realizaron varias contracciones a diferentes niveles de esfuerzo hasta manifestar fatiga muscular. No se realizó ningún tipo de preparación especial de la piel más allá su limpieza (sin usar pasta abrasiva ni gel conductor).

Se registraron señales diferenciales a una frecuencia de muestreo de 2048 Hz con un amplificador EMG-USB de 128 canales (OT Bioelettronica, Turín, Italia), colocando las columnas en la dirección de las fibras musculares. Las señales resultantes se procesaron usando MATLAB r2019b (The Mathworks, Natick, Massachussets, Estados Unidos). Se utilizo un filtro de Butterworth de tercer orden, paso banda entre 15 y 350 Hz, con doble pasada para cancelar la distorsión de fase.

Se realizó una inspección visual de las señales obtenidas, así como el cálculo de la relación señal a ruido (teniendo en cuenta los intervalos de contracción muscular y silencio, descartando transitorios), la velocidad de conducción, y la frecuencia media y mediana del espectro [7].

Con ambos tipos de electrodos se pudo observar la propagación de los potenciales de acción de unidad motora. La velocidad de conducción obtenida se mantuvo dentro del rango fisiológico del bíceps (entre 3 y 6 m/s) [4], permitiendo la estimación de la fatiga muscular. Los espectros obtenidos a partir de las señales registradas también permitieron observar la fatiga muscular a partir de sus correspondientes frecuencias media y mediana.

Finalmente, se escogió el acero inoxidable para la construcción del prototipo de manga, tanto por su coste reducido como por la facilidad de obtención en comparación con los remaches bañados en plata.

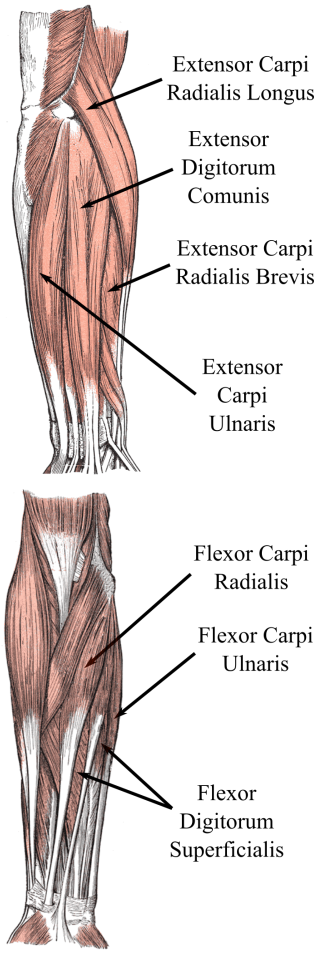


Figura 1. *Disposición de los músculos superficiales de interés en el antebrazo derecho: extensores en la parte posterior (arriba) y flexores en la anterior (abajo). Adaptado de [6].*

			61	46	31	16	1
99	90	76	62	47	32	17	2
100	91	77	63	48	33	18	3
101	92	78	64	49	34	19	4
102	93	79	65	50	35	20	5
103	94	80	66	51	36	21	6
	95	81	67	52	37	22	7
	96	82	68	53	38	23	8
	97	83	69	54	39	24	9
	98	84	70	55	40	25	10
		85	71	56	41	26	11
		86	72	57	42	27	12
		87	73	58	43	28	13
		88	74	59	44	29	14
		89	75	60	45	30	15

1	11	21	31	41	51	61
2	12	22	32	42	52	62
3	13	23	33	43	53	63
4	14	24	34	44	54	64
5	15	25	35	45	55	65
6	16	26	36	46	56	66
7	17	27	37	47	57	67
8	18	28	38	48	58	68
9	19	29	39	49	59	69
10	20	30	40	50	60	70

Figura 2. *Disposición de electrodos en las matrices de los extensores (izquierda) y flexores (derecha) del antebrazo.*

Descripción	Músculo activo
Aducción de mano y muñeca	*Extensor Carpi Ulnaris*
Extensión de dedos y muñeca	*Extensor Digitorum Communis*
Abducción de mano y muñeca	*Extersor Carpi Radiales (Longus* y *Brevis)*
Flexión y abducción de la muñeca	*Flexor Carpi Ulnaris* y *Flexor Carpi Radialis*
Flexión de las articulaciones interfalángicas y metacarpofalángicas	*Flexor Digitorum Superficialis* y *Flexor Digitorum Profundus*
Flexión de la articulación interfalángica y metacarpofalángica del pulgar	*Flexor Pollicis Longus*

Tabla 1. *Protocolo de ejercicios.*

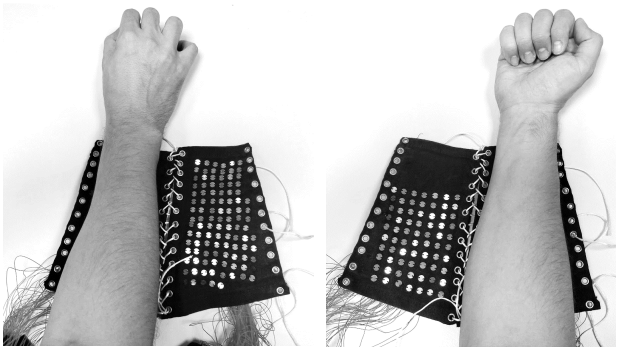


Figura 2. *Disposición de electrodos en las matrices de los extensores (izquierda) y flexores (derecha) del antebrazo.*

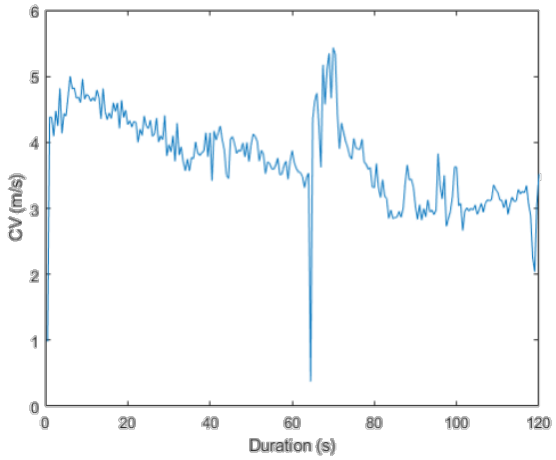


Figura 4. *Velocidad de conducción medida sobre el músculo Extensor Carpi Ulnaris (contracción de 120 s con 5 s de descanso). Se observa el descenso debido a la fatiga muscular.*

2.3. Diseño de la matriz de electrodos

La disposición de los músculos flexores y extensores superficiales es el criterio principal para dimensionar las matrices de electrodos que tiene que cubrir las partes anterior y posterior del antebrazo, respectivamente.

Las columnas de la parte posterior se dispusieron paralelamente al cúbito con el brazo está en reposo, de manera que la primera columna se encuentre justo por encima del hueso. Considerando la disposición de los músculos extensores superficiales del antebrazo, se diseñó una matriz no rectangular con una distancia entre electrodos de 1 cm, y un máximo de 15 filas y 8 columnas (Figura 2).

Para la parte anterior también se tomó el cúbito como referencia, con la primera columna justo por debajo del hueso. Al no ser tan importante la zona distal del antebrazo se diseñó una matriz rectangular de 10 filas y 7 columnas. La manga resultante puede verse en la Figura 3.

2.4. Validación del prototipo

Con 166 canales a registrar, se utilizaron dos amplificadores EMG-USB de 128 canales, uno para cada una de las matrices, configurados en modo diferencial *daisy chain*. Descartando las señales diferenciales no válidas (por ejemplo, entre los canales 15 y 16 de la matriz de extensores) se obtuvieron 95 señales diferenciales para los músculos extensores y 63 para los flexores.

Dado que no es posible adaptar el software del fabricante para los nuevos sensores, se utilizó el software OTBiolabLight v1.3 (OT Bioelettronica, Turín, Italia) para adquirir las señales a través TCP/IP directamente desde MATLAB. Se programaron diversas rutinas de adquisición, procesado y visualización para poder observar las señales y los mapas de activación (valor cuadrático medio) en tiempo real durante la adquisición.

Se diseñó un protocolo de ejercicios que permitiese la activación selectiva de los músculos del antebrazo (Tabla 1), de manera que fuese posible validar la actividad registrada con la manga sensorizada tanto en el mapa como en las señales registradas. Se realizaron pruebas con 5 voluntarios reclutados en el propio laboratorio y sin desórdenes neuromusculares (rango de edad entre 23 y 36 años), cuyos antebrazos tuvieran unas dimensiones similares o superiores a las del prototipo desarrollado para evitar problemas de solapamiento al colocar la manga.

3. Resultados

Las matrices construidas permiten registrar señales EMG de alta densidad con las que evaluar la actividad y fatiga muscular (Figura 4). Los mapas de valor cuadrático medio obtenidos en tiempo real permiten valorar visualmente la activación de cada músculo superficial del antebrazo de manera rápida para saber si se están realizando los ejercicios correctamente.

Las Figuras 5 y 6 muestran, a modo de ejemplo, algunos mapas de los músculos extensores y flexores, así como las señales asociadas en las cuales se observa la propagación de los potenciales de acción.

4. Conclusiones

Los resultados obtenidos son esperanzadores puesto que se ha logrado adquirir y mostrar en tiempo real señales y mapas de EMG de alta densidad del antebrazo con sensores de bajo coste. Se trata de una región complicada con una gran cantidad de músculos y un gran problema de *crosstalk*, teniendo en cuenta de que no se ha realizado

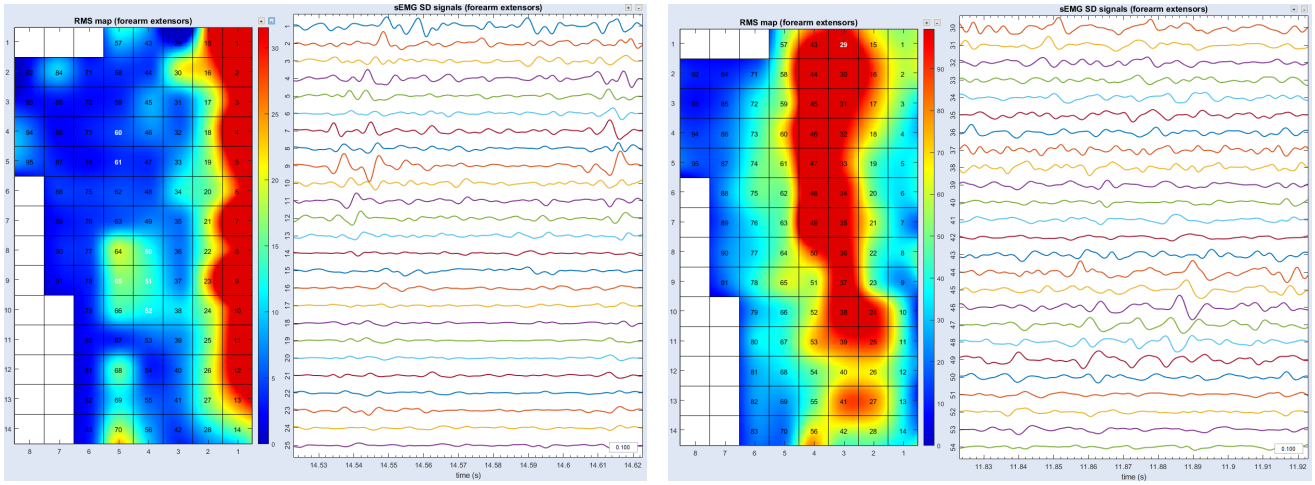


Figura 5. *Mapas correspondientes a la aducción de mano y muñeca (izquierda) y la extensión de dedos y muñeca (derecha).*

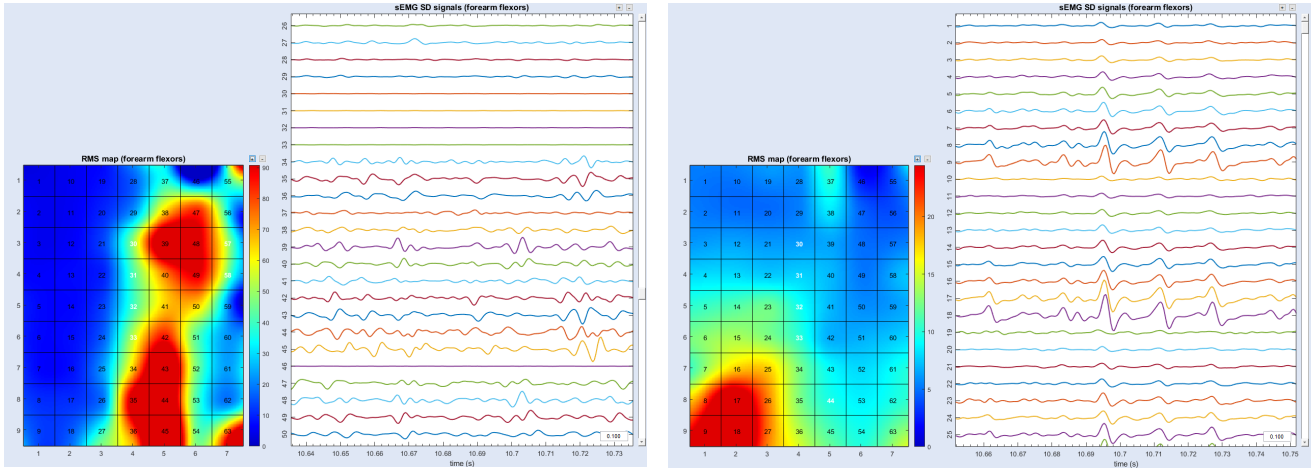


Figura 6. *Mapas correspondientes a la flexión de articulaciones interfalángicas y metacarpofalángicas (izquierda) y la flexión de la articulación interfalángica y metacarpofalángica del pulgar (derecha).*

ningún tipo de preparación de la piel previa al registro, pero es posible evaluar las señales EMG utilizando algoritmos validados para el cálculo de la velocidad de conducción y variables espectrales con las que valorar la fatiga muscular.

En individuos con mucho vello o mucha grasa subdérmica la calidad de la señal se degrada ligeramente, pero en general el principal problema sigue siendo el movimiento de la articulación que provoca falta de contacto entre la piel y algún electrodo de manera intermitente.

Las mejoras a implementar en el futuro más próximo van dirigidas a mejorar el ajuste de la manga y la segmentación automática de los músculos activos para ayudar a los terapeutas a valorar la rehabilitación de pacientes con ictus, así como detectar la correcta colocación de la manga.

Agradecimientos

Este trabajo ha sido financiado parcialmente por los proyectos *MyoSleeve* (CaixaImpulse 2018) y DPI2017-83989-R (Ministerio de Economía y Competitividad).

J. F. Alonso es un profesor Serra Húnter.

Los autores agradecen la colaboración de E. Pérez, M. Jordanic y M. A. Mañanas durante las pruebas preliminares con diferentes electrodos y en la definición del protocolo de ejercicios.

Referencias

[1] Página web de la Federación española del ictus. https://ictusfederacion.es/infoictus/codigo-ictus (Consultada: Agosto 2020).

[2] Página web de la Stroke Alliance For Europe (SAFE). https://safestroke.eu (Consultada: Agosto 2020).

[3] Domínguez D, Grau M. Entrenamiento aeróbico y de fuerza en la rehabilitación del ictus. *Fisioterapia,* vol 33, 2011, pp 210-6 (ISSN: 0211-5638).

[4] Merletti R, Farina D. Surface Electromyography: Physiology, Engineering, and Applications. Wiley, 2016 (ISBN: 9781118987025).

[5] Rojas-Martínez M, Mañanas MA, Alonso JF. High-density surface EMG maps from upper-arm and forearm muscles. *Journal of NeuroEngineering and Rehabilitation,* vol 9, 2012 (ISSN: 1743-0003).

[6] Gray, H. Anatomy of the Human Body.Lea & Febiger, 1918, Bartleby.com, 2000. www.bartleby.com/107/. (Consultada: Octubre de 2020).

[7] Merletti R, Muceli S. Tutorial. Surface EMG detection in space and time: Best practices. *Journal of Electromyography and Kinesiology*, vol 49, 2019 (ISS: 1050-6411).

Predicción a Largo Plazo del Resultado de la Ablación Quirúrgica en Fibrilación Auricular Mediante el Uso de una Única Derivación del Electrocardiograma Preoperatorio

M. Martínez Iniesta[1], P. Escribano Cano[1], M. García Teruel[1], J. Ródenas García[1], F. Hornero Sos[2], J.J. Rieta Ibáñez[3], R. Alcaraz Martínez[1]

[1] Grupo de Investigación en Electrónica, Telecomunicación y Bioingeniería, DIEEAC, Universidad de Castilla-La Mancha, Albacete, España, {Miguel.Martínez, Pilar.Escribano, Manuel.García, Juan.Rodenas, Raul.Alcaraz}@uclm.es

[2] Dpto. Cirugía Cardiovascular, Hospital Clínico Universitario, Valencia, España, hornero_fer@gva.es

[3] BioMIT.org, Dpto. Ingeniería Electrónica, Universidad Politécnica de Valencia, Valencia, España, jjrieta@upv.es

Resumen

La ablación quirúrgica (AQ), realizada concomitantemente en pacientes que requieren cirugía a corazón abierto, es el procedimiento más eficaz para controlar la fibrilación auricular (FA). Considerando los importantes efectos secundarios de esta técnica, junto a los beneficios de anticipar una decisión sobre el tratamiento antiarrítmico, la predicción preoperatoria de su resultado a largo plazo supone un interesante reto clínico. Este trabajo introduce un novedoso algoritmo de predicción del resultado de la AQ, transcurrido un año de seguimiento, basado en el análisis de la derivación V1 del electrocardiograma (ECG) preoperatorio. El método extrae de éste las ondas fibrilatorias, segmenta la señal resultante en tramos de 1 s y calcula la desviación estándar de la energía wavelet relativa para la séptima escala (EWR7_DE) contenida en cada uno de ellos. Este índice consiguió una capacidad discriminatoria del 75% entre los pacientes que mantuvieron un ritmo sinusal y aquellos que recayeron a FA durante el seguimiento, superando así en casi un 20% la habilidad de otros parámetros clásicos y convirtiéndose en un prometedor predictor de recurrencia de la FA tras la AQ.

1. Introducción

La fibrilación auricular (FA) es la arritmia cardíaca más común en la práctica clínica. Afecta globalmente a más de 33 millones de personas, lo que supone más del 0,5% de la población mundial [1]. Su prevalencia está directamente relacionada con la edad, es decir, mientras que solo afecta al 0,16% de la población menor de 50 años, este porcentaje se eleva hasta el 17% entre personas mayores de 80 años [2]. Las actuales proyecciones sobre el crecimiento de la población mayor de 60 años indican que ésta pasará de 962 millones en 2017 a más de 2.000 millones en 2050. Ello, asociado al aumento de la prevalencia de enfermedades crónicas que predisponen la FA en personas mayores y las mejoras que continuamente se están produciendo en su detección, apuntan a un importante aumento de la incidencia de esta arritmia en los próximos años [3]. Teniendo en cuenta que se estima que la gestión de la FA representa más del 15% del gasto destinado a patologías cardiovasculares y aproximadamente el 1% del presupuesto sanitario mundial, esta situación supondrá un enorme impacto económico en todos los sistemas sanitarios [4].

Los mecanismos que desencadenan y mantienen la FA no son todavía completamente conocidos y las terapias disponibles no son tan eficaces como se desearía desde un punto de vista clínico [5,6]. De hecho, aunque los tratamientos actuales, tales como la cardioversión eléctrica y farmacológica o la ablación por catéter, alcanzan tasas de éxito iniciales superiores al 90%, la recurrencia de la FA suele ser común transcurridas varias semanas o meses, presentando tasas de fallo del 50-70% [3]. Por el contrario, la ablación quirúrgica (AQ) ha alcanzado tasas de éxito a largo plazo notablemente mayores, próximas al 80%, incluso tras un seguimiento de varios años [7]. Este procedimiento consiste en una intervención quirúrgica a corazón abierto a través de la cual se crean líneas de ablación en la aurícula, mediante crioterapia o radiofrecuencia, para redireccionar los impulsos del nodo sinusal al nodo auriculoventricular a través de una ruta específica [8]. Dicha intervención puede aplicarse de forma independiente o concomitante con otros procedimientos cardíacos, incluyendo el injerto de bypass de la arteria coronaria o la reparación de las válvulas mitral y aórtica [9].

Tras la AQ, los pacientes son habitualmente tratados con antiarrítmicos y anticoagulantes para evitar la recurrencia de la FA, siendo evaluados rutinariamente a los 3, 6, y 12 meses y, posteriormente, cada año. El tratamiento se retira si se mantiene el ritmo sinusal (RS) estable durante varios meses, pero en el caso de recurrencia de FA durante el seguimiento, se aplica cardioversión eléctrica para restaurar el RS [7, 10]. En este contexto, una predicción preoperatoria del ritmo del paciente transcurrido un año de la intervención podría ser de utilidad clínica para planificar un seguimiento personalizado de cada paciente, programando con anticipación la cardioversión eléctrica en aquellos pacientes con un alto riesgo de recurrencia de la FA y evitando tratamientos farmacológicos agresivos en aquellos otros con alta probabilidad de mantenimiento del RS [11].

Determinados parámetros clínicos, tales como la edad, el tamaño de la aurícula izquierda o el tiempo preoperatorio en FA, propuestos para tal propósito en los últimos años, han proporcionado solamente una moderada capacidad para predecir el resultado de la AQ a largo plazo [12, 13].

ISBN: 978-84-09-25491-0, pags. 327- 330

Por el contrario, otros índices calculados a partir del electrocardiograma de superficie (ECG) han demostrado una gran efectividad para anticipar el resultado inmediato del procedimiento en el momento del alta [14]. Así pues, el presente trabajo introduce un nuevo método, basado en la estimación de la variabilidad temporal presentada por las ondas fibrilatorias (*f*) reflejadas en la señal de ECG, para predecir de una manera preoperativa el resultado a largo plazo de la AQ.

2. Metodología

2.1. Base de datos

En este estudio participaron 53 pacientes que se encontraban en FA persistente o permanente, al menos, cuatro meses antes de someterse a la AQ concomitantemente con una cirugía a corazón abierto. Para todos ellos se obtuvo, antes de las 48 horas previas a la cirugía, una grabación estándar del ECG de 12 derivaciones con una frecuencia de muestreo de 1 kHz y una resolución en la amplitud de 0.4 μV. Para cada uno de los pacientes, se extrajo un extracto de ECG de la derivación V1 de 20 segundos de duración. Únicamente fue evaluada esta derivación porque refleja las mayores ondas *f* comparadas con la actividad ventricular [15]. Tras un año de seguimiento, 23 de estos pacientes recayeron a FA, mientras que los 30 restantes permanecieron en RS.

2.2. Procesamiento de la señal de ECG

Cada segmento de ECG extraído fue inicialmente preprocesado para reducir diferentes tipos de ruido relacionados con la adquisición de la señal y, de esta manera, mejorar su posterior análisis. Con este propósito, para eliminar la interferencia de la red eléctrica y preservar la morfología original de las ondas *f* [16], se utilizó inicialmente una metodología de filtrado basada en la umbralización de los coeficientes resultantes de aplicar la transformada wavelet estacionaria (TWE). A continuación, la línea de base fue eliminada por medio de un filtro paso alto con 0.5 Hz de frecuencia de corte [17]. Finalmente, el ruido de alta frecuencia fue reducido mediante un filtro paso bajo de tipo IIR con frecuencia de corte en 70 Hz [18]. Ambos filtros fueron diseñados mediante una ventana Chebyshev con una atenuación relativa del lóbulo lateral de 40 dB.

A continuación, las ondas *f* fueron extraídas de cada extracto de ECG a través de un algoritmo previamente publicado de cancelación del complejo QRST [19]. Brevemente, se detectaron en primer lugar los picos R haciendo uso de un eficiente método basado en la transformada fasorial [20]; a continuación, todos los complejos QRST fueron alineados y, para cada complejo bajo cancelación, fueron seleccionados los 10 más similares mediante correlación cruzada entre ellos; por último, se obtuvo una plantilla representativa de cada complejo mediante un análisis de componentes principales, que fue sustraída del ECG preprocesado para obtener las ondas *f*. A modo de ejemplo, la Figura 1 muestra un extracto de un ECG, junto a las ondas *f* extraídas mediante el algoritmo descrito.

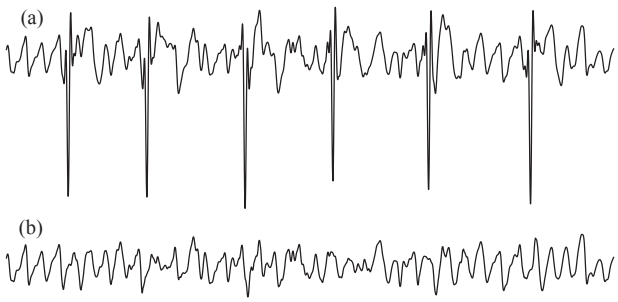


Figura 1. *Ejemplo de un intervalo de un ECG (a), junto con las ondas f extraídas usando un algoritmo previamente publicado de cancelación del complejo QRST (b).*

2.3. Descomposición wavelet de las ondas *f*

Las ondas f se caracterizan por una morfología que varía en el tiempo, presentando así formas de onda de diferente tamaño, amplitud y duración [21]. No obstante, en los últimos años, el grado de organización de estas ondas ha estado estrechamente asociado con un resultado exitoso en varios tratamientos de la FA, tales como como la ablación por catéter y la cardioversión eléctrica [22]. En ese sentido, en el presente trabajo se propone un novedoso algoritmo para estimar la trayectoria temporal de la organización de las ondas *f* con el fin de anticipar el resultado a largo plazo de la AQ. Para ello, inicialmente, la señal que contenía las ondas *f* fue dividida en tramos de 1 segundo de duración. Este período de tiempo fue obtenido experimentalmente tras varias pruebas. Posteriormente, cada segmento se descompuso en varias escalas de tiempo y frecuencia usando la TWE. Esta transformación es una herramienta adecuada para el análisis de transitorios, aperiodicidades y otras características de señales no estacionarias, ya que los cambios sutiles en la morfología de la señal pueden ser fácilmente resaltadas sobre las escalas de interés [23].

Teniendo en cuenta que las señales de ECG fueron adquiridas con una frecuencia de muestreo de 1 kHz y que las ondas *f* presentan la información más relevante en las bajas frecuencias, se eligió una descomposición wavelet de ocho niveles. Al igual que en trabajos previos [24], se utilizó una función Daubechies de sexto orden como wavelet madre. Posteriormente, se calculó la energía wavelet relativa (EWR) para la séptima escala, dado que cubre el rango de frecuencia típica de la FA comprendido entre 4 y 8 Hz. Dicho cálculo se realizó como [25]:

$$EWR7 = \frac{\sum_{k=1}^{N} C_j(k)^2}{\sum_{j=1}^{8}\sum_{k=1}^{N} C_j(k)^2}\,, \qquad (1)$$

donde C_j es un vector que contiene los coeficientes wavelet para la escala j y N es la longitud de la señal original (esto es, el número de muestras correspondientes a 1 segundo). Por último, se calculó la desviación estándar de la EWR7 (EWR7_DE) para todos los tramos como medida de la variabilidad temporal de las ondas *f*.

2.4. Análisis estadístico

Para poder establecer una referencia, se ha analizado dos características clínicamente conocidas de las ondas *f*: la frecuencia auricular dominante (FAD), estimada como la frecuencia con mayor amplitud de densidad espectral de

potencia dentro del rango de 3-12 Hz [25], y la amplitud de las ondas *f* (AOF), obtenida como el valor cuadrático medio de las ondas *f* [26].

Para las tres métricas indicadas, los resultados se expresaron en media ± desviación estándar. Además, las diferencias estadísticas entre ambos grupos de pacientes, esto es, aquellos que recayeron a FA y los que mantuvieron RS, fueron evaluadas mediante un test *U* de Mann-Whitney. Se consideró una significación estadística $p<0.05$ como estadísticamente significativa. Además, la habilidad de cada métrica para discriminar entre ambos grupos fue evaluada mediante una curva ROC (Receiver Operating Characteristic), que es el resultado de representar gráficamente la sensibilidad frente a 1 – especificidad, para varios umbrales de discriminación. La sensibilidad (Se) se consideró como el porcentaje de pacientes que mantuvieron el RS que fueron correctamente clasificados, en tanto que la especificidad (Es) se obtuvo como la tasa de pacientes que recayeron en FA correctamente clasificados. El umbral óptimo de discriminación se seleccionó como aquel que proporcionó el mayor porcentaje de pacientes correctamente clasificados, esto es, la mayor exactitud (Ex). También se obtuvieron los valores de predictividad positiva (P^+) y de predictividad negativa (P^-), calculados como la proporción de verdaderos pacientes que mantuvieron el RS y recayeron a FA, respectivamente, de entre todos los identificados dentro de cada uno de los dos grupos. Finalmente, se calculó el área bajo la curva ROC (AROC), como una medida agregada del rendimiento de una métrica a través de todos los umbrales de clasificación posibles.

3. Resultados

En la Tabla 1 se muestran, para ambos grupos de pacientes, los valores medios y de desviación estándar obtenidos para los tres índices analizados. Se muestra también el valor de la significancia estadística. Como puede observarse, aunque la FAD proporcionó un valor de *p* muy próximo a ser significante, la EWR7_DE fue el único índice que proporcionó diferencias estadísticamente significativas entre pacientes que recayeron a FA y aquellos otros que se mantuvieron en RS durante el seguimiento. De hecho, para estos últimos, los valores obtenidos con este índice fueron notablemente inferiores a los conseguidos para aquellos pacientes que recayeron a FA.

De acuerdo con estas observaciones, la Tabla 2 muestra que los mejores resultados de clasificación fueron alcanzados mediante la EWR7_DE, que consiguió valores de Ex y AROC entre un 16 y 22% mayores que los obtenidos con la FAD y la AOF, alcanzando una capacidad discriminante del 75%. Este índice también consiguió mejoras comprendidas entre el 12 y el 30% en Se, Es, P^+ y P^-, en comparación con las otras dos métricas.

4. Discusión y conclusiones

Recientemente, tanto la FAD como la AOF han demostrado ser prometedores predictores preoperatorios del fracaso temprano de la AQ, mostrando menores y mayores valores, respectivamente, para los pacientes que mantienen un RS que para aquellos que recaen a FA [14].

Índice	Grupo de pacientes que		*p*
	mantuvieron RS	recayeron a FA	
FAD	6.10±1.40 Hz	6.70±0.90 Hz	0.060
AOF	48.70±17.80 µV	44.70±22.70 µV	0.201
EWR7_DE	0.102±0.026	0.131±0.012	<0.001

Tabla 1. *Valor medio y desviación estándar para los índices analizados para los dos grupos de pacientes. Se muestra también la significación estadística*

Índice	*P*+	*P*-	Se(%)	Es(%)	Ex(%)	AROC(%)
FAD	67	48	63	52	58	58
AOF	53	52	59	46	53	57
EWR7_DE	83	65	76	75	75	74

Tabla 2. *Resultados de la clasificación para los tres índices analizados*

Un comportamiento similar se ha observado para estas dos métricas en el presente trabajo, si bien es cierto que no se han encontrado diferencias estadísticas entre ambos grupos de pacientes. Además, aunque ambos índices alcanzaron tasas de clasificación superiores al 70% en el momento del alta [14], los valores de Ex y AROC obtenidos en este estudio, tras un año de seguimiento, no lograron superar el 60%. Estas diferencias entre ambos resultados se podrían deber a varias causas, pero quizás la más relevante podría ser la alteración del sustrato auricular provocada por los fármacos antiarrítmicos e incluso la cardioversión eléctrica durante el seguimiento. En cualquier caso, esta suposición requiere de una investigación más profunda en trabajos posteriores.

A diferencia de estos dos predictores, la EWR7_DE sí obtuvo diferencias estadísticamente significativas entre ambos grupos de pacientes, observándose una notable menor variabilidad en la energía de las ondas *f* para aquellos pacientes que mantuvieron el RS que en aquellos otros que recayeron a FA. Este resultado señala que la presencia de ondas *f* más desorganizadas e inestables aumenta la probabilidad de recaída en FA a largo plazo tras la AQ. A este respecto, algunos trabajos previos también han asociado una conducción eléctrica más homogénea dentro de las aurículas con una mayor organización de las ondas *f* [27,28]. En consecuencia, podría considerarse que el análisis de cómo las ondas *f* evolucionan en el tiempo podría proporcionar información novedosa sobre el sustrato auricular tras la AQ, lo que podría resultar de gran utilidad en la selección de aquellos candidatos más adecuados para este tratamiento. Por tanto, este trabajo abre una nueva puerta a la obtención de nuevos conocimientos acerca de la remodelación del sustrato auricular tras la AQ, que merece ser cuidadosamente explorada en el futuro utilizando una base de datos con un mayor número de pacientes.

Agradecimientos

Trabajo financiado por los proyectos DPI2017-83952-C3 MINECO/AEI/FEDER, UE, SBPLY/17/180501/000411 de la Junta de Comunidades de Castilla-La Mancha y AICO/2019/036 de la Generalitat Valenciana.

Referencias

[1] Chugh SS, Roth GA, Gillum RF, Mensah GA. Global burden of atrial fibrillation in developed and developing nations. *Glob Heart*, vol. 9, no. 1, pp. 113–9, Mar 2014.

[2] Zoni-Berisso M, Lercari F, Carazza T, Domenicucci S. Epidemiology of atrial fibrillation: European perspective. *Clin Epidemiol*, vol. 6, pp. 213–20, 2014.

[3] Chung MK, Refaat M, Shen WK, Kutyifa V, Cha YM, Di Biase L, Baranchuk A, Lampert R, Natale A, Fisher J, Lakkireddy DR, ACC Electrophysiology Section Leadership Council. Atrial fibrillation: Jacc council perspectives. *J Am Coll Cardiol*, vol. 75, no. 14, pp. 1689–1713, Apr 2020.

[4] Pathak RK, Evans M, Middeldorp ME, Mahajan R, Mehta AB, Meredith M, Twomey D, Wong CX, Hendriks JML, Abhayaratna WP, Kalman JM, Lau DH, Sanders P. Cost-effectiveness and clinical effectiveness of the risk factor management clinic in atrial fibrillation: The CENT study. *JACC Clin Electrophysiol*, vol. 3, no. 5, pp. 436–447, 05 2017.

[5] Nattel S, Dobrev D. Controversies about atrial fibrillation mechanisms: Aiming for order in chaos and whether it matters. *Circ Res*, vol. 120, no. 9, pp. 1396-1398, 04 2017.

[6] Schotten U, Dobrev D, Platonov PG, Kottkamp H, Hindricks G. Current controversies in determining the main mechanisms of atrial fibrillation. *J Intern Med*, vol. 279, no. 5, pp. 428–38, May 2016.

[7] Khiabani AJ, MacGregor RM, Bakir NH, Manghelli JL, Sinn LA, Maniar HS, Moon MR, Schuessler RB, Melby SJ, Damiano Jr RJ. The long-term outcomes and durability of the cox-maze iv procedure for atrial fibrillation. *J Thorac Cardiovasc Surg*, May 2020.

[8] Ad N, Holmes SD, Rongione AJ, Massimiano PS, Fornaresio LM. Does surgical ablation energy source affect long-term success of the concomitant cox maze procedure? *Ann Thorac Surg*, vol. 104, no. 1, pp. 29–35, Jul 2017.

[9] Churyla A, Desai A, Kruse J, Cox J, McCarthy P. Concomitant atrial fibrillation ablation in patients undergoing coronary artery bypass and cardiac valve surgery. *J Cardiovasc Electrophysiol*, Feb 2020.

[10] Harky A, Bithas C, Chan JSK, Snosi M, Pousios D, Muir AD. Surgical treatment of atrial fibrillation: Cutting through the edges. *Braz J Cardiovasc Surg*, vol. 35, no. 3, pp. 375–386, Jun 2020.

[11] Ad N, Holmes SD. Prediction of sinus rhythm in patients undergoing concomitant cox maze procedure through a median sternotomy. *J Thorac Cardiovasc Surg*, vol. 148, no. 3, pp. 881–6; discussion 886–7, Sep 2014.

[12] Kim YH, Lee SC, Her AY, Kim HJ, Choi JO, Shin DH, Cho SW, Lee WS, Park SW, Park PW. Preoperative left atrial volume index is a predictor of successful sinus rhythm restoration and maintenance after the maze operation. *J Thorac Cardiovasc Surg*, vol. 134, no. 2, pp. 448–53, Aug 2007.

[13] Chen MC, Chang JP, Chang HW. Preoperative atrial size predicts the success of radiofrequency maze procedure for permanent atrial fibrillation in patients undergoing concomitant valvular surgery. *Chest*, vol. 125, no. 6, pp. 2129–34, Jun 2004.

[14] Hernández A, Alcaraz R, Hornero F, Rieta JJ. Preoperative study of the surface ECG for the prognosis of atrial fibrillation maze surgery outcome at discharge. *Physiol Meas*, vol. 35, no. 7, pp. 1409–23, Jul 2014.

[15] Petrenas A, Marozas V, Sörnmo L. Lead systems and recording devices, in *Atrial Fibrillation from an Engineering Perspective.* Springer, 2018, pp. 25–48.

[16] García M, Martínez-Iniesta M, Ródenas J, Rieta JJ, Alcaraz R. A novel wavelet-based filtering strategy to remove powerline interference from electrocardiograms with atrial fibrillation. *Physiol Meas*, vol. 39, no. 11, p. 115006, 11 2018.

[17] Dotsinsky I, Stoyanov T. Optimization of bi-directional digital filtering for drift suppression in electrocardiogram signals. *J Med Eng Technol*, vol. 28, no. 4, pp. 178–80, 2004.

[18] Sörnmo L, Laguna P. Bioelectrical signal processing, in *cardiac and neurological applications.* Academic Press, 2005, vol. 8.

[19] Alcaraz R, Rieta JJ. Adaptive singular value cancelation of ventricular activity in single-lead atrial fibrillation electrocardiograms. *Physiol Meas*, vol. 29, no. 12, pp. 1351–69, Dec 2008.

[20] Martínez A, Alcaraz R, Rieta JJ. Application of the phasor transform for automatic delineation of single-lead ecg fiducial points. *Physiol Meas*, vol. 31, no. 11, pp. 1467–85, Nov 2010.

[21] Sörnmo L, Alcaraz R, Laguna P, Rieta JJ. Characterization of f waves in *Atrial Fibrillation from an Engineering Perspective*. Springer, 2018, pp. 221–279.

[22] Alcaraz R, Rieta JJ. Signal analysis in atrial fibrillation, in *Cardiovascular Computing—Methodologies and Clinical Applications*. Springer, 2019, pp. 331–350.

[23] Addison PS. The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance. CRC press, 2017.

[24] García M, Ródenas J, Alcaraz R, Rieta JJ. Application of the relative wavelet energy to heart rate independent detection of atrial fibrillation. *Comput Methods Programs Biomed*, vol. 131, pp. 157–68, Jul 2016.

[25] Szilágyi J, Walters TE, Marcus GM, Vedantham V, Moss JD, Badhwar N, Lee B, Lee R, Tseng ZH, Gerstenfeld EP. Surface ecg and intracardiac spectral measures predict atrial fibrillation recurrence after catheter ablation. *J Cardiovasc Electrophysiol*, vol. 29, no. 10, pp. 1371–1378, 10 2018.

[26] Alcaraz R, Hornero F, Rieta JJ. Noninvasive time and frequency predictors of long-standing atrial fibrillation early recurrence after electrical cardioversion. *Pacing Clin Electrophysiol*, vol. 34, no. 10, pp. 1241–50, Oct 2011.

[27] Bollmann A, Husser D, Mainardi L, Lombardi F, Langley P, Murray A, Rieta JJ, Millet J, Olsson SB, Stridh M, Sörnmo L. Analysis of surface electrocardiograms in atrial fibrillation: techniques, research, and clinical applications. *Europace*, vol. 8, no. 11, pp. 911–26, Nov 2006.

[28] Lankveld T, Zeemering S, Scherr D, Kuklik P, Hoffmann BA, Willems S, Pieske B, Haïssaguerre M, Jaïs P, Crijns HJ, Schotten U. Atrial fibrillation complexity parameters derived from surface ecgs predict procedural outcome and long-term follow-up of stepwise catheter ablation for atrial fibrillation. *Circ Arrhythm Electrophysiol*, vol. 9, no. 2, p. e003354, Feb 2016.

Estudio de la relación entre el gen MAPT y la conectividad cerebral en la demencia por enfermedad de Alzheimer

A. Maturana-Candelas[1], J. Poza [1,2,3], R. Hornero[1,2,3], P. Núñez[1], S.J. Ruiz-Gómez[1], V. Gutiérrez-de Pablo[1], M. Revilla-Vallejo[1], C. Gómez[1,2]

[1] Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, aaron.maturana@gib.tel.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

Resumen

El objetivo de este estudio es analizar la relación entre la actividad electroencefalográfica (EEG) y el alelo MAPT en pacientes con demencia por enfermedad de Alzheimer (EA), sujetos con deterioro cognitivo leve (DCL) y controles. Para ello, se registró la actividad EEG en 109 sujetos con alelo normal y 90 con alelo protector. Posteriormente, se calculó la conectividad entre 68 regiones de interés mediante el método phase lag index (PLI). Nuestros resultados muestran valores de conectividad significativamente más altos en los grupos control y DCL con alelos normales en las bandas bajas de frecuencia (delta y zeta), mientras que en bandas altas (beta para controles y gamma para DCL) se obtiene el efecto contrario. Este comportamiento se invierte y se reduce en el grupo EA, con valores de conectividad significativamente más altos en el grupo con alelo normal en bandas altas de frecuencias y viceversa. Estos resultados sugieren que el alelo protector de MAPT mantiene patrones de conectividad propios de estados cognitivos de menor deterioro.

1. Introducción

La demencia debida a la enfermedad de Alzheimer (EA) es una patología caracterizada por la alteración de los patrones de conducta y pérdida de memoria a corto plazo [1]. Estas perturbaciones cognitivas provocadas por la destrucción de tejido nervioso y defectos sinápticos se reflejan en patrones y dinámicas aberrantes en la actividad neuronal [2,3]. Normalmente, los pacientes con EA atraviesan por un periodo previo a la demencia, denominado deterioro cognitivo leve (DCL) [4]. El DCL debido a EA se considera una forma prodrómica de la enfermedad, asociada a ligeros déficits cognitivos, pero insuficientes para precisar un diagnóstico de demencia [4].

El DCL y la demencia por EA han sido ampliamente estudiados por medio de diversas técnicas. Entre ellas, destaca la electroencefalografía (EEG), que permite registrar el potencial eléctrico generado a partir de la actividad cerebral [5]. Este método cuenta con múltiples ventajas, como alta resolución temporal, portabilidad y coste reducido. Estas propiedades hacen del EEG una técnica empleada en una extensa variedad de estudios centrados en la caracterización de las alteraciones cerebrales provocadas por la EA [5]. El EEG registra la información en la superficie cortical, siendo inaccesible la medición en zonas más profundas del cerebro. Con el fin de estimar la actividad eléctrica en regiones profundas se han desarrollado procedimientos de localización de fuentes. Uno de ellos es la tomografía electromagnética de baja resolución estandarizada (sLORETA) que se basa en el supuesto de que la actividad de neuronas vecinas está fuertemente correlacionada, lo que implica un error de localización nulo [6]. sLORETA ya ha sido empleada en el estudio de alteraciones neurofisiológicas provocadas por la EA, tanto en el córtex como en zonas más profundas del cerebro [7,8].

Uno de los biomarcadores más típicos de la EA es la agregación de ovillos neurofibrilares de proteína tau hiperfosforilada intracelular [1]. La presencia de estos depósitos es el resultado de alteraciones estructurales, asociadas a degeneración axonal. Ello provoca disrupciones en los mecanismos sinápticos que pueden provocar alteraciones en el comportamiento eléctrico neuronal [9]. El gen MAPT (*microtubule-associated protein tau*) se encarga de codificar la proteína tau [10], la cual está relacionada con los procesos fisiológicos asociados a la neurodegeneración en la EA [1]. En investigaciones previas, se determinó el polimorfismo de nucleótido único (SNP) rs8070723-G, perteneciente al gen MAPT, como un posible factor protector en el desarrollo de la EA [10]. Puesto que MAPT está íntimamente ligado a la producción de tau, nuestra hipótesis es que los sujetos con el alelo normal (*wild type*, WT) mostrarán diferentes patrones de actividad cerebral que aquellos que tienen el alelo protector. Así pues, el objetivo de este estudio es evaluar si existen diferencias en los patrones de conectividad funcional, mediante la medida *phase lag index* (PLI), en pacientes con EA, sujetos con DCL y controles de ambos grupos genéticos.

2. Métodos

2.1. Sujetos

En este estudio participaron 199 sujetos. Todos los pacientes con DCL y EA fueron diagnosticados según los criterios del *National Institute of Aging and Alzheimer's Association* (NIA-AA) [11]. Los participantes fueron clasificados en seis grupos, atendiendo a su estado cognitivo (EA, DCL y control) y al tipo de alelo presente en el gen MAPT (normal WT o protector). En la tabla 1 se resumen los datos sociodemográficos de cada grupo.

ISBN: 978-84-09-25491-0, pags. 331- 334

Se obtuvo el consentimiento informado de cada participante o, de no ser posible, de su familiar o representante legal, acorde con las recomendaciones del Código Ético de la Asociación Médica Mundial (Declaración de Helsinki). El protocolo fue aprobado por el Comité de Ética de la Universidad de Oporto (Oporto, Portugal. Informe nº 38 / CEUP / 2018).

2.2. Análisis genético

Se recogió una muestra de saliva de cada sujeto mediante un kit *DNA Genotek Oragene* (OG-500). Las muestras biológicas se genotiparon utilizando *Thermo Fisher Scientific Axiom™ Spain Biobank Arrays* en el Centro Nacional de Genotipado (CeGEN, Santiago de Compostela, España). Se implementó un control de calidad en el genotipado. Este protocolo fue seguido por análisis individuales y de marcadores, de acuerdo con la guía de mejores prácticas de Affymetrix. Los sujetos con tasas de heterocigosidad superiores al umbral de aceptación definido fueron descartados. El SNP de MAPT (rs8070723) calificó como válido en el análisis de control de calidad.

2.3. Registro y preprocesado de EEG

Para cada sujeto, se registraron 5 minutos de actividad EEG en estado de reposo con un equipo Nihon Kohden Neurofax JE-921A de 19 canales, en los electrodos F3, F4, F7, F8, Fp1, Fp2, T3, T4, T5, T6, C3, C4, P3, P4, O1, O2, Fz, Cz y Pz del sistema internacional 10-20. La frecuencia de muestreo se estableció en 500 Hz. Durante la adquisición se aplicó referenciado promedio. Se solicitó a los participantes que se mantuvieran relajados durante el procedimiento en un ambiente libre de ruido. Los investigadores controlaron el estado de vigilia del sujeto para prevenir la somnolencia.

Los datos de EEG se almacenaron como archivos ASCII y se preprocesaron siguiendo estos pasos [12]: (i) sustracción de la media; (ii) filtro de ranura de 50 Hz; (iii) filtro FIR (*finite impulse response*) paso banda de ventana Hamming entre 1 y 70 Hz; (iv) análisis de componentes independientes (ICA) para eliminar componentes ruidosas; y (v) segmentación en épocas de 5 s de duración y rechazo visual de aquellas con artefactos.

Tras la adquisición de EEG, se estimó la actividad a nivel de fuente utilizando el algoritmo de localización de fuentes sLORETA. Este algoritmo asume máxima correlación entre generadores neuronales vecinos para obtener la solución al problema inverso [6]. La localización de fuente se llevó a cabo trabajando con 15000 posibles fuentes que fueron agrupadas en 68 regiones corticales de interés (ROIs) de acuerdo con el atlas de Desikan-Killiany, una clasificación de regiones basada en giros correspondientes a su función dominante [13]. El preprocesado y el análisis de las señales se llevaron a cabo con MATLAB® (versión R2018a, Mathworks, Natick, MA) y *Brainstorm toolbox*.

2.4. Procesado de EEG

La conectividad entre las 68 ROIs propuestas se estimó mediante el PLI en las bandas clásicas de frecuencia: delta (δ, 1-4 Hz), zeta (θ, 4-8 Hz), alfa (α, 8-13 Hz), beta (β, 13-30 Hz) y gamma (γ, 30-70 Hz). El PLI cuantifica la asimetría de las distribuciones de diferencia de fase de dos señales [14]. Esta medida tiene la ventaja de estar poco afectada por la conducción de volumen [14]. A mayor acoplamiento en fase, valores más próximos a 1. El PLI se define matemáticamente de la siguiente forma:

$$PLI_{X,Y} = |\langle sign\,(\Delta\Phi_{X,Y})\rangle|, \quad (1)$$

donde $\langle\cdot\rangle$ es la esperanza matemática y $\Delta\Phi_{X,Y}$ es la diferencia de fase o fase relativa entre las señales *X* e *Y*. Para cada sujeto, se promediaron los resultados obtenidos en todas las épocas libres de artefactos, obteniéndose una matriz de conectividad para cada sujeto y banda de frecuencia. El tamaño de cada una de estas matrices es 68x68, donde cada elemento refleja el valor de conectividad entre dos ROIs.

3. Resultados

En la figura 1 se muestran las conexiones entre ROIs en las que se obtuvieron diferencias estadísticamente significativas (p-valor < 0.01, test U de Mann-Whitney) entre sujetos con alelo WT y sujetos con alelo protector. Las conexiones rojas muestran interacciones con valores medios de conectividad más elevados en el grupo con alelo normal WT, mientras que las conexiones azules indican interacciones con valores de conectividad mayores en el grupo con alelo protector.

Los resultados obtenidos muestran un aumento de conectividad en sujetos con alelo WT con respecto al grupo con alelo protector en las bandas delta y zeta para los grupos control y DCL, y en la banda beta para el grupo de pacientes con EA. Por otro lado, una mayor conectividad es observable en los sujetos con alelo protector en controles en la banda beta, en el grupo de sujetos con DCL en la banda gamma y, en menor medida, en el grupo de pacientes con EA en bandas zeta y alfa.

	Controles		DCL		EA	
	WT	PR	WT	PR	WT	PR
N	23	22	24	21	62	47
Edad ($\bar{x}$±SD) (años)	77.73 ± 6.22	81.77 ± 7.83	85.12 ± 8.26	84.52 ± 5.74	81.13 ± 6.98	79.83 ± 7.58
Sexo (H:M)	10:13	12:10	9:15	5:16	16:46	11:36
MMSE ($\bar{x}$±SD)	28.73 ± 1.18	29.00 ± 1.02	22.37 ± 2.90	24.14 ± 2.70	15.72 ± 7.58	14.35 ± 8.36

Tabla 1. Base de datos sociodemográficos y clínicos: x: media, SD: desviación típica. WT: alelo wild type, PR: alelo protector. H: hombre, M: mujer. MMSE: puntuación del test Minimental State Examination.

4. Discusión

En este estudio se ha analizado la relación entre la conectividad entre regiones cerebrales y los alelos WT y protector del gen MAPT. Los hallazgos obtenidos más destacables son: (i) el aumento significativo de conectividad de los sujetos con alelo WT con respecto al protector de los grupos control y DCL en bandas bajas (delta y zeta); (ii) el aumento significativo de conectividad en los sujetos con alelo protector en el grupo control en la banda beta, y en el grupo DCL en gamma; y (iii) el comportamiento opuesto en el grupo de pacientes con EA en relación a los grupos anteriores, además de un impacto del gen notablemente más reducido.

En el grupo de pacientes con DCL se obtuvo un aumento significativo de conectividad en los sujetos con alelo WT con respecto al protector en banda zeta, para un total de 108 interacciones. La condición clínica de DCL es conocida por un aumento de actividad eléctrica cerebral anormal en esta banda de frecuencia [15], comúnmente atribuido a mecanismos compensatorios para equilibrar las primeras perturbaciones neurofisiológicas desencadenantes de la EA [16]. Nuestros resultados podrían interpretarse no ya como un aumento significativo de conectividad en el grupo con el alelo WT, sino como un efecto modulador de dichas compensaciones inducido por el alelo protector. Puesto que esta forma de tau se asocia a desestabilización del axón y con alteraciones funcionales en las dendritas [17], las siguientes hipótesis podrían explicar la reducción de conectividad en el grupo de sujetos con DCL. En primer lugar, la proteína tau como componente esencial del citoesqueleto axonal podría perdurar más y ser más estable, produciendo menores concentraciones de tau hiperfosforilada en el proceso de neurodegeneración temprano. En segundo lugar, la proteína en estado hiperfosforilado podría presentar una disminución de la capacidad citotóxica en la neurona. Los resultados obtenidos en el grupo de control para la misma banda sugieren que este proceso podría haber comenzado en una fase preclínica de la enfermedad en algunos sujetos.

Adicionalmente, se obtuvieron valores de conectividad significativamente más altos para los controles con alelo protector en la banda beta y para los pacientes con DCL con este mismo alelo en la banda gamma, en comparación con los grupos con alelo WT. Puesto que las bandas beta y gamma están asociadas con procesos de alto nivel cognitivo, como tareas que requieren uso de memoria [18], nuestros resultados sugieren que podría haber relación entre la presencia del alelo protector de MAPT y la conservación de capacidades intelectuales en la senescencia no patológica, así como en la fase prodrómica de la enfermedad. El efecto a nivel molecular del alelo protector en altas frecuencias puede deberse a la asociación entre la presencia de tau hiperfosforilada y alteraciones de sistemas dopaminérgicos [19], ya que la dopamina es un elemento relacionado con la modulación de la actividad beta y gamma [20].

Por último, es interesante señalar el efecto de ambas versiones del gen MAPT en los pacientes con EA. Por un lado, se puede observar una disminución del impacto de la presencia de ambos alelos (WT y protector) asociada a la conectividad en este estadio, como refleja el menor número de conexiones estadísticamente significativas, en comparación con los grupos control y DCL. Por otro lado, el grupo de enfermos con EA con alelo protector muestra aumentos significativos de conectividad en zeta y alfa, mientras que esto no ocurre en los otros grupos. Además, las bandas beta y gamma, con interacciones de valores más elevados asociados al alelo WT para controles y DCL respectivamente, manifestaron el efecto opuesto en el grupo EA. Esta observación podría indicar un efecto moderado de protección por parte del alelo protector de MAPT en la conectividad cerebral en estas bandas de frecuencia. Asimismo, el menor número de interacciones estadísticamente significativas podría indicar que el efecto

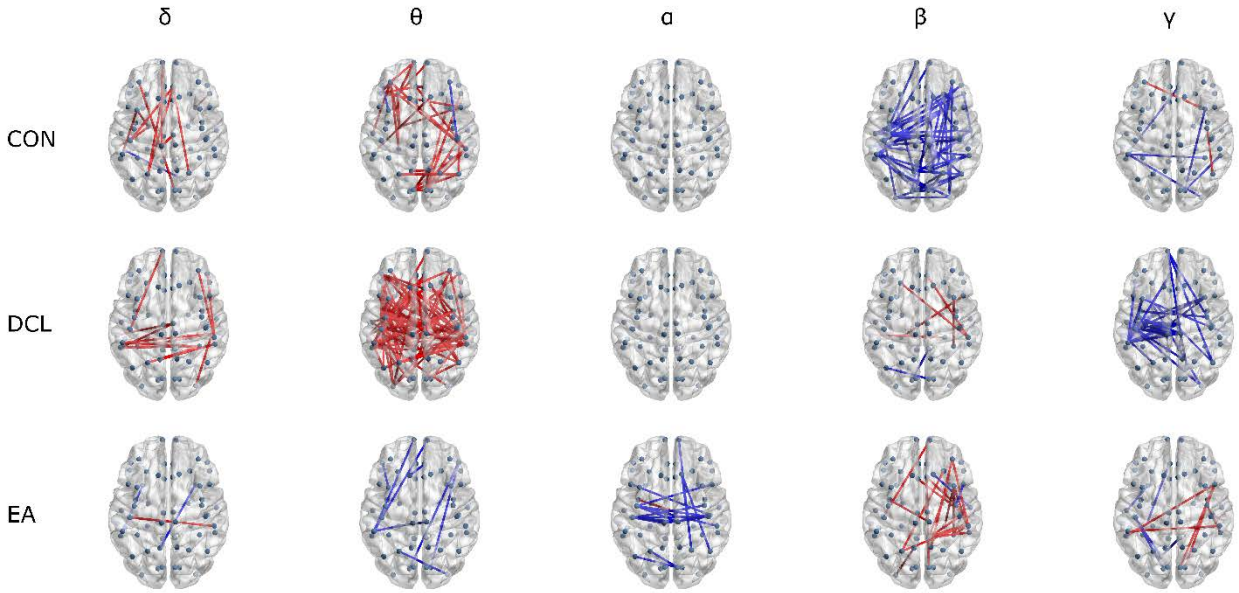


Figura 1. Diferencias estadísticamente significativas en los valores de conectividad entre regiones de interés (p-valor < 0.01, test U de Mann-Whitney). Líneas rojas indican un valor de PLI mayor en sujetos con el alelo normal WT, mientras que las líneas azules indican un valor de PLI mayor en los sujetos con el alelo protector.

de este alelo tiene un impacto más notorio en fases tempranas y preclínicas de la enfermedad.

Finalmente, es necesario mencionar ciertas limitaciones de este estudio. En primer lugar, la base de datos empleada cuenta con un número reducido de sujetos. En segundo lugar, es preciso tener en cuenta que los resultados obtenidos pueden deberse a un efecto secundario del alelo protector de MAPT, y no a directamente con la proteína tau. Finalmente, no se ha considerado la correlación entre MAPT y otros genotipos asociados a procesos citotóxicos que puedan desembocar en efectos de conectividad similares. Para subsanar estas limitaciones, incrementar la cantidad de participantes y estudiar otros genes vinculados a alteraciones sinápticas sería prioritario.

5. Conclusión

En este estudio, se han analizado las diferencias de conectividad entre grupos de estudio en diversos estados cognitivos con dos versiones del gen MAPT. Los resultados sugieren que el efecto del alelo protector puede estar asociado a una modulación de los mecanismos de compensación en fases tempranas de la EA y a la conservación del comportamiento fisiológico propio de estados cognitivos saludables. Puesto que MAPT está íntimamente relacionado con la síntesis de proteína tau, la versión del alelo protector podría reflejar un incremento de la estabilidad axonal y presentar una menor citotoxicidad en su forma hiperfosforilada.

Agradecimientos

Esta investigación ha sido financiada por la ‘Comisión Europea’ y el ‘Fondo Europeo de Desarrollo Regional’ (FEDER) a través de los proyectos ‘Análisis y correlación entre el genoma completo y la actividad cerebral para la ayuda en el diagnóstico de la enfermedad de Alzheimer’ y ‘Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres’ (‘Programa de Cooperación Transfronteriza Interreg V-A España-Portugal 2014-2020’), por el ‘Ministerio de Ciencia e Innovación - Agencia Estatal de Investigación’ y FEDER a través del proyecto PGC2018-098214-A-I00, y por el ‘CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN)’ a través del ‘Instituto de Salud Carlos III’ cofinanciado con fondos FEDER. A. Maturana-Candelas y S.J. Ruiz-Gómez son beneficiarios de una ayuda concedida por la Consejería de Educación de la Junta de Castilla y León y el Fondo Social Europeo. P. Núñez disfruta de un contrato predoctoral ‘Ayuda para contratos predoctorales para la Formación de Profesorado Universitario (FPU)’ del ‘Ministerio de Educación, Cultura y Deporte’ (FPU17/00850).

Referencias

[1] Alzheimer's Association, “2019 Alzheimer’s disease facts and figures”, *Alzheimer’s Dement.*, vol. 15, no. 3, pp. 321–387, 2019.

[2] C. Babiloni *et al.*, “Brain neural synchronization and functional coupling in Alzheimer’s disease as revealed by resting state EEG rhythms”, *Int. J. Psychophysiol.*, vol. 103, pp. 88–102, 2016.

[3] J. Dauwels *et al.*, “Diagnosis of Alzheimers Disease from EEG Signals: Where Are We Standing?”, *Curr. Alzheimer Res.*, vol. 7, no. 6, pp. 487–505, 2010.

[4] R. C. Petersen, “Mild cognitive impairment as a diagnostic entity”, *J. Intern. Med.*, vol. 256, no. 3, pp. 183–194, 2004.

[5] J. Poza *et al.*, “Spatio-Temporal Fluctuations of Neural Dynamics in Mild Cognitive Impairment and Alzheimer’s Disease”, *Curr. Alzheimer Res.*, vol. 14, no. 9, 2017.

[6] R. D. Pascual-Marqui, “Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details”, *Methods Find. Exp. Clin. Pharmacol.*, 24D, 1–16, 2002.

[7] C. Babiloni *et al.*, “Cortical sources of resting state EEG rhythms are sensitive to the progression of early stage Alzheimer’s disease”, *J. Alzheimer’s Dis.*, vol. 34, no. 4, pp. 1015–1035, 2013.

[8] A. Kabbara *et al.*, “Reduced integration and improved segregation of functional brain networks in Alzheimer’s disease”, *J. Neural Eng.*, vol. 15, no. 2, p. 026023, 2018.

[9] B. R. Hoover *et al.*, “Tau Mislocalization to Dendritic Spines Mediates Synaptic Dysfunction Independently of Neurodegeneration”, *Neuron*, vol. 68, no. 6, pp. 1067–1081, 2010.

[10] M. Allen *et al.*, “Association of MAPT haplotypes with Alzheimer’s disease risk and MAPT brain gene expression levels”, *Alzheimer’s Res. Ther.*, vol. 6, no. 4, pp. 1–14, 2014.

[11] G. M. McKhann *et al.*, “The diagnosis of dementia due to Alzheimer’s disease: Recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease”, *Alzheimer’s Dement.*, vol. 7, no. 3, pp. 263–269, 2011.

[12] A. Maturana-Candelas *et al.*, “EEG Characterization of the Alzheimer’s Disease Continuum by Means of Multiscale Entropies”, *Entropy*, vol. 21, no. 6, p. 544, 2019.

[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] C. J. Stam *et al.*, “Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources”, *Hum. Brain Mapp.*, vol. 28, no. 11, pp. 1178–1193, 2007.

[15] J. C. McBride *et al.*, “Spectral and complexity analysis of scalp EEG characteristics for mild cognitive impairment and early Alzheimer’s disease”, *Comput. Methods Programs Biomed.*, vol. 114, no. 2, pp. 153–163, 2014.

[16] A. Ashraf *et al.*, “Cortical hypermetabolism in MCI subjects: a compensatory mechanism?”, *Eur. J. Nucl. Med. Mol. Imaging*, vol. 42, no. 3, pp. 447–458, 2015.

[17] B. R. Hoover *et al.*, “Tau Mislocalization to Dendritic Spines Mediates Synaptic Dysfunction Independently of Neurodegeneration”, *Neuron*, vol. 68, no. 6, pp. 1067–1081, 2010.

[18] S. M. Staufenbiel *et al.*, “Effect of beta and gamma neurofeedback on memory and intelligence in the elderly”, *Biol. Psychol.*, vol. 95, no. 1, pp. 74–85, 2014.

[19] H. L. Melrose *et al.*, “Impaired dopaminergic neurotransmission and microtubule-associated protein tau alterations in human LRRK2 transgenic mice”, *Neurobiol. Dis.*, vol. 40, no. 3, pp. 503–517, 2010.

[20] K. Kondabolu *et al.*, “Striatal cholinergic interneurons generate beta and gamma oscillations in the corticostriatal circuit and produce motor deficits”, *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 22, pp. 3159–3168, 2016.

Evaluación de la Calidad en Registros Electrocardiográficos Procedentes de Pacientes con Fibrilación Auricular Mediante una Red Neuronal Convolucional

A. Huerta Herraiz[1], A. Martínez Rodrigo[1], M.A. Arias Palomares[2], P. Langley[3], J.J. Rieta Ibáñez[4], R. Alcaraz Martínez[1]

[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] Departamento de Arritmias Cardíacas, Hospital Virgen de la Salud, Toledo, España, maapalomares@secardiologia.es
[3] Facultad de Ciencias e Ingeniería, Universidad de Hull, Reino Unido, p.langley@hull.ac.uk
[4] BioMIT.org, Dep. de Ingeniería Electrónica, Universitat Politecnica de Valencia, Valencia, España, jjrieta@upv.es

Resumen

La fibrilación auricular (FA) es la arritmia cardíaca más común en la práctica clínica. Su aparición se asocia con un mayor riesgo de eventos cardiovasculares, siendo su detección temprana un reto todavía sin resolver. Esta arritmia suele comenzar con episodios asintomáticos y muy breves, necesitándose por tanto para su detección sistemas de registro electrocardiográfico (ECG) con capacidad para la monitorización de largos periodos de tiempo. El principal problema que presentan estos sistemas es su fuerte exposición al ruido, por ser adquiridos en entornos no controlados. Por lo tanto, es necesario disponer de sistemas automáticos que evalúen la calidad del registro de ECG adquirido, como primer paso para una detección precisa y temprana de la FA. Así pues, en este trabajo se introduce un algoritmo basado en aprendizaje profundo, con el objetivo de discernir entre segmentos ECG de alta y baja calidad en registros de ECG de una sola derivación, adquiridos en pacientes con FA. El método se basa en la alta capacidad de aprendizaje de una red neuronal convolucional, la cual se ha entrenado con señales de ECG sin preprocesar convertidas a imagen. Los resultados muestran una gran capacidad del método para discernir entre los segmentos de ECG de alta y baja calidad, alcanzando una exactitud del 95 %. Además, el método solo confunde el 6 % de los segmentos con FA como intervalos de baja calidad. Estos resultados mejoran en más de un 20 % el rendimiento de la mayoría de los algoritmos previamente publicados para la evaluación de la calidad del ECG que también tratan con señales de FA.

1. Introducción

Actualmente, la fibrilación auricular (FA) es la arritmia cardíaca más común en la práctica clínica diaria, afectando aproximadamente a 37,5 millones de personas en todo el mundo [1]. Su prevalencia está relacionada directamente con la edad, y se prevé que alcance proporciones epidémicas a mediados de este siglo [2]. No es una enfermedad mortal por sí misma, pero aumenta el riesgo de accidentes cerebrovasculares, así como otros trastornos cardiovasculares [3]. A la hora de realizar un diagnóstico temprano y acertado de esta arritmia, el principal inconveniente es la corta duración que presentan sus episodios iniciales, de los cuales se calcula que un 90 % son asintomáticos [4]. Para intentar solucionar este problema, los pacientes se monitorizan con dispositivos portátiles o vestibles, capaces de capturar señales electrocardiográficas (ECG) de muy larga duración, mientras el paciente sigue realizando sus vida diaria normal.

El principal inconveniente de los dispositivos portátiles es su sensibilidad al entorno en el que se registra la señal, la cual puede estar afectada por distintas perturbaciones, tales como artefactos de movimiento, interferencias de la red eléctrica, o ruido de alta frecuencia, entre otros [5]. A pesar de que la eliminación de estos ruidos en el ECG ha sido motivo de diferentes trabajos de investigación [6], el rendimiento de los métodos propuestos es todavía limitado. Esto se debe a que algunas componentes de frecuencia del ruido, así como los artefactos de movimiento se solapan con la forma de onda del ECG en los dominios del tiempo y frecuencia [7]. Además, en el caso de la FA, las regulares ondas P se reemplazan por otras mucho más irregulares y caóticas, llamadas ondas fibrilatorias, las cuales presentan un aspecto muy parecido al ruido [8]. Por lo tanto, la evaluación automática de la calidad del ECG es de vital importancia para una detección temprana y precisa de la FA, previniendo la mala interpretación de intervalos de ECG ruidosos y evitando una mala interpretación del registro, un mal diagnóstico y un mal tratamiento, con los consiguientes riesgos asociados.

Para discernir adecuadamente entre intervalos de ECG de alta y baja calidad en registros de muy larga duración, se requieren algoritmos automáticos que no necesiten supervisión durante su funcionamiento, de hecho, los análisis realizados automáticamente pueden ser más eficientes y son mucho más rápidos que los realizados manualmente, presentando una mayor coherencia en sus decisiones [9].

ISBN: 978-84-09-25491-0, pags. 335- 338

Así pues, hasta la fecha, se han publicado diversos métodos automáticos basados en características dependientes del tiempo y de eventos morfológicos [7,10]. Sin embargo, el rendimiento de estos trabajos está estrechamente ligado con su habilidad para detectar los puntos más relevantes en el ECG, los cuales se ven muy deteriorados en presencia de ruido [11]. No obstante, recientemente se han presentado algunos métodos automáticos basados en redes neuronales, las cuales han obtenido un alto rendimiento a la hora de detectar tramos de alta calidad en registros de ECG. Con esta técnica, no es necesaria la extracción de características del ECG ni otro tipo de intervención externa [6, 12].

El principal problema de estos últimos trabajos es que solamente se han validado con registros de ECG obtenidos de personas sin patologías conocidas, evitando así un contexto más complejo, tal como el correspondiente a la presencia de FA, donde episodios de ritmo sinusal y de la arritmia se van alternando a lo largo del tiempo [10]. Otro inconveniente en estos trabajos es el relativo al aprendizaje de la red neuronal propuesta, la cual se ha entrenado desde cero con un número relativamente bajo de muestras, lo cual hace que pueda estar sobreentrenada. Por lo tanto, en el presente trabajo se propone el uso de una red neuronal preentrenada para discernir entre tramos de ECG de alta y baja calidad, extraídos de señales de muy larga duración y una sola derivación, pertenecientes a pacientes con FA. Concretamente, se ha empleado la bien conocida red AlexNet [13] bajo el concepto de aprendizaje de transferencia y con un pequeño ajuste para desempeñar la tarea descrita.

2. Base de datos

Para el desarrollo del presente trabajo y la validación del método propuesto, se ha empleado una base de datos compuesta por 36 registros de una sola derivación con una duración de entre una y dos horas. Las señales se adquirierion en pacientes con FA intermitente empleando un dispositivo Holter vestible, de la marca Nuubo^{TM}, trabajando con una frecuencia de muestreo de 250 Hz y una resolución de 12 bits en un rango dinámico de $\pm$5 mV. Los episodios de FA se detectaron mediante un método automático aplicado a las señales en bruto, sin ningún tipo de procesamiento [8]. Además, dos expertos etiquetaron de manera independiente las señales de ECG en tramos de alta y baja calidad. El criterio establecido para considerar un intervalo de baja calidad fue que ambos no pudiesen identificar con seguridad todos los picos R en el mismo. Finalmente, los registros se segmentaron en tramos de 5 segundos, y fueron agrupados en dos categorías, es decir, alta y baja calidad. El número total de tramos disponibles para el análisis fue de 46.709, de los cuales 9.663 pertenecían a la clase de baja calidad y 37.046 a la categoría de alta calidad. Cabe destacar que, del total de segmentos de alta calidad, 10.338 pertenecían a FA y 26.708 a ritmo sinusal normal.

3. Métodos

El método propuesto para evaluar la calidad del ECG se compone de dos etapas. En primer lugar, los segmentos de ECG de 5 segundos de duración se transforman en matrices de color de dos dimensiones y, en segundo lugar, se entrena la red neuronal utilizando como entrada dichas imágenes. A continuación, se detallan en profundidad cada uno de los dos pasos.

3.1. Transformada Wavelet Continua

A través de la transformada Wavelet continua (TWC) se consiguió una representación en imagen de cada segmento de ECG de 5 segundos que contenía su información tiempo-frecuencia. Esta transformación ha obtenido muy buenos resultados a la hora de tratar con señales aperiódicas, no estacionarias y con transitorios como el ECG [14]. A grandes rasgos, se basa en la descomposición de la señal en una serie de escalas temporales, las cuales representan un determinado rango de frecuencia concreto en el plano tiempo-frecuencia. Esta transformación se consigue correlando una señal inicial con versiones escaladas y desplazadas de una función wavelet, denominada *wavelet madre*. De forma mas precisa, para calcular la TWC de una señal $x(t)$ se emplea la siguiente expresión:

$$TWC(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\psi^*\Big(\frac{t-b}{a}\Big)dt, \quad (1)$$

donde $a, b \in \Re, a \neq 0$ son, respectivamente, el factor de escalado y el parámetro de desplazamiento de la wavelet madre, y $*$ denota el operador complejo conjugado. Como resultado, se obtiene una matriz de dos dimensiones llamada escalograma en la que los coeficientes calculados se muestran en las escalas y posiciones correspondientes. Este tipo de representación es muy común en trabajos que utilizan redes neuronales artificiales cuya entrada es una imagen. Como ejemplo, en la Figura 1 se muestran dos segmentos de ECG de 5 segundos de duración, en el que se observa la señal original y su correspondiente escalograma.

3.2. La red AlexNet

Las redes neuronales convolucionales se forman mediante la combinación de diversas capas que emulan las conexiones neuronales del cerebro. Los tres tipos más comunes son las capas de convolución, las capas de agrupación o *pooling* y las capas totalmente conectadas. En primer lugar, las capas de convolución se encargan de extraer las características más relevantes de la señal o imagen de entrada. En segundo lugar, se utilizan las capas de agrupación para reducir las dimensiones espaciales del mapa de características, sin perder información relevante. Por último, las capas totalmente conectadas se utilizan para conectar todas las neuronas de la capa anterior con todas las neuronas de la actual. Cabe destacar que el número total de neuronas de la última capa totalmente conectada define el número de clases de salida.

Para el desarrollo del presente trabajo se empleó la conocida arquitectura AlexNet [13]. Se adaptó para discernir entre segmentos de ECG de alta y baja calidad y se reentrenó para que adquiriese el nuevo conocimiento requerido. La arquitectura original de AlexNet se compone de ocho capas con capacidad de aprender, cinco convolucio-

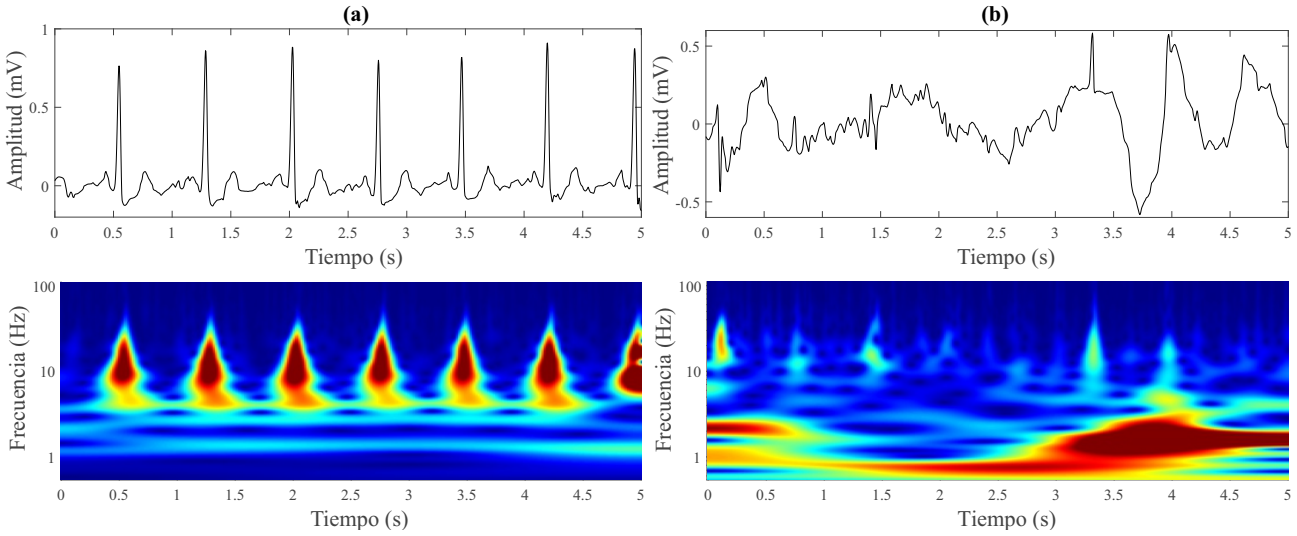


Figura 1. *Ejemplo de ECG típico de alta (izquierda) y baja (derecha) calidad. Se representa la señal original junto con el escalograma obtenido mediante la TWC.*

nales y tres totalmente conectadas, además de dos capas de agrupación. También, como función de activación emplea una función no lineal aplicada después de todas las capas convolucionales y totalmente conectadas. El modelo contiene dos funciones de *dropout* para reducir el sobreentrenamiento de la red.

Finalmente, la capa de salida de AlexNet se modificó para trabajar con dos clases y la red se re-entrenó con la base de datos descrita en la Sección 2, siendo necesarias 10 épocas con un tamaño de lote de 10 escalogramas. El proceso se desarrolló utilizando un algoritmo de descenso de gradiente estocástico con un momento de 0,9 y una tasa de aprendizaje de 0,0001 invariante durante todo el tiempo de entrenamiento.

3.3. Desarollo experimental

A la vista de la base de datos descrita en la Sección 2, el número de segmentos de ECG de ambas clases se encuentra notablemente desbalanceado, pudiendo ocasionar resultados poco fiables o sesgados. Para solucionar este problema se ha planteado una validación basada en repeticiones. Así pues, la red neuronal se sometió a 20 ciclos de validación, en los que se utilizaron como datos de entrada 1.200 muestras seleccionadas al azar de cada uno de los grupos. Hay que destacar que el subconjunto de alta calidad estuvo siempre compuesto por 600 segmentos de ritmo sinusal normal y 600 intervalos de FA. Con cada uno de estos 20 conjuntos de 2400 muestras, el algoritmo propuesto se entrenó y validó 5 veces empleando un 80 % y un 20 % de las señales, respectivamente. Así pues, finalmente se condujeron 100 procesos de entrenamiento y test

En cada uno de estos procesos, se obtuvo el resultado de la clasificación en términos de sensibilidad (Se), especificidad (Es) y exactitud (Ex). Además también se obtuvo el máximo, el mínimo, la media y la desviación estándar (std) de estos tres índices de funcionamiento para todos los ciclos de validación. Cabe mencionar que la sensibilidad se definió como la tasa de segmentos de ECG de alta calidad clasificados correctamente, la especificidad como el porcentaje de intervalos de baja calidad identificados correctamente y la exactitud como el número total de segmentos de ECG detectados correctamente. Además, se obtuvieron y promediaron las tasas de acierto para los tramos pertenecientes a ritmo sinusal normal ($\mathcal{R}_{SR}$) y FA ($\mathcal{R}_{AF}$), correctamente clasificados dentro del grupo de alta calidad.

4. Resultados

Los resultados obtenidos en cuanto a la media, std, máximo y mínimo de las métricas de funcionamiento evaluadas se encuentra en la Tabla 1. Se observa una exactitud en torno al 95 %, alcanzando un máximo del 96,88 %. También se aprecia una pequeña diferencia del 5 % entre sensibilidad y especificidad. En cuanto a las tasas de acierto dentro del grupo de ECG de alta calidad de intervalos de ritmo sinusal y de segmentos de FA, se observan valores superiores al 93 %. No obstante, se puede destacar que mientras que la tasa media de acierto de los segmentos de FA es de un 94 %, este valor sube hasta el 99,5 % para los segmentos de ritmo sinusal. Por último, en la tabla también se observan unos valores de std muy bajos para todos los índices, entre el 1 y el 3 %, por lo que no se aprecian diferencias significativas entre los diferentes ciclos de validación.

Índice	$Se.\,(\%)$	$Es.\,(\%)$	$Ex.\,(\%)$	$\mathcal{R}_{SR}\,(\%)$	$\mathcal{R}_{AF}$
Media	97.17	92.42	94.79	99.58	93.86
Std	1.57	3.14	1.41	0.94	2.75
Máximo	98.75	96.25	96.88	100	97.09
Mínimo	94.58	88.75	93.33	97.90	89.69

Tabla 1. *Valores de media, std, máximo y mínimo obtenidos para los índices de funcionamiento de los 100 ciclos de validación conducidos.*

5. Discusión

En este trabajo se ha presentado un novedoso algoritmo para discernir automáticamente entre tramos de alta y baja calidad en registros de ECG de una sola derivación obtenidos mediante un sistema portátil en pacientes con FA

intermitente. A pesar de que se han analizado las señales de ECG en bruto, se han alcanzado tasas de clasificación media comparables o incluso superiores a la mayoría de las técnicas anteriores que se han propuesto para evaluar la calidad en registros de ECG de una sola derivación [15].

Sin embargo, a la hora de compararse con estos métodos previos, hay que tener en cuenta ciertos aspectos, como por ejemplo que la base de datos empleada es diferente en cada trabajo. Además, mientras que en el presente estudio se desarrolló un robusto sistema de validación empleando más de 46.000 segmentos de ECG en 100 iteraciones, en la mayoría de los trabajos anteriores se ha realizado una validación con un solo conjunto de datos bastante limitado [15]. Por ejemplo, Zhao y Zhang [16] tan solo utilizaron 300 intervalos de ECG de 30 segundos de duración de una base de datos pública para el desarrollo de su trabajo. Claramente, con un análisis de la base de datos completa (más de 8.000 segmentos) se habrían obtenido resultados más robustos y una comparación más justa.

También hay que destacar que la mayoría de los métodos anteriores solamente se han validado con registros de ECG obtenidos de pacientes sin patologías [10]. Por lo tanto, considerar de manera conjunta segmentos de ritmo sinusal y FA dentro del grupo de alta calidad hace más compleja la tarea de discernir el grado de calidad que presenta el ECG que se procesa. De hecho, los pocos métodos que han considerado registros de FA para su validación han mostrado un descenso considerable en su capacidad de identificar la calidad del ECG, en comparación a cuando solo analizaban registros de pacientes sanos, obteniendo exactitudes en torno al 50 y 70 % [10, 11, 15]. Sin embargo, el presente trabajo ha demostrado una gran capacidad para discernir entre segmentos de FA de alta calidad y segmentos de ruido alrededor del 94 %, mientras que la exactitud global del método para discernir entre intervalos de ECG de alta y baja calidad también se ha mantenido en torno al 95 %.

Finalmente, cabe mencionar algunas limitaciones del estudio. Por un lado, la principal desventaja del método propuesto es la falta de una interpretación clara de los resultados en términos fisiopatológicos. En este sentido, los algoritmos basados en redes neuronales convolucionales se caracterizan por mantener ocultas las dependencias funcionales entre las entradas y las salidas [17]. Por otro lado, a pesar de que en el presente trabajo se ha analizado una base de datos con un elevado número de registros la cual se ha adquirido en un entorno totalmente realista, un estudio más exhaustivo incluyendo diferentes tipos de ruidos y artefactos sería muy útil para completar los resultados obtenidos.

6. Conclusiones

Se ha propuesto un método novedoso basado en una red neuronal convolucional pre-entrenada para discernir de manera fiable entre intervalos de alta y baja calidad en registros de ECG de una sola derivación, incluyendo episodios de ritmo sinusal normal y FA. El método se ha verificado de manera exhaustiva mediante múltiples ciclos de validación, obteniendo un rendimiento superior a la mayoría de métodos propuestos en trabajos previos.

Agradecimientos

Este trabajo ha sido cofinanciado por los proyectos de investigación competitiva DPI2017-83952-C3 de MINECO-AEI-FEDER-UE, SBPLY/17/180501/000411 de la JCCLM, AICO/2019/036 de la GVA y FEDER 2018/11744.

Referencias

[1] Lippi G, Sanchis-Gomar F, Cervellin G. Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge. Int J Stroke Jan 2020;1–5.

[2] Krijthe BP, Kunst A, Benjamin EJ, Lip GYH, Franco OH, Hofman A, Witteman JCM, Stricker BH, Heeringa J. Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060. Eur Heart J Sep 2013;34(35):2746–51.

[3] Khoo CW, Krishnamoorthy S, Lim HS, Lip GYH. Atrial fibrillation, arrhythmia burden and thrombogenesis. Int J Cardiol Jun 2012;157(3):318–23.

[4] Stachon P, Ahrens I, Faber T, Bode C, Zirlik A. Asymptomatic atrial fibrillation and risk of stroke. Panminerva Med Dec 2015;57(4):211–5.

[5] Nagai S, Anzai D, Wang J. Motion artefact removals for wearable ECG using stationary wavelet transform. Healthcare Technology Letters 2017;4(4):138–141.

[6] Yoon D, Lim HS, Jung K, Kim TY, Lee S. Deep learning-based electrocardiogram signal noise detection and screening model. Healthcare Informatics Research 2019;25(3):201–211.

[7] Oster J, Behar J, Sayadi O, Nemati S, Johnson AE, Clifford GD. Semisupervised ECG ventricular beat classification with novelty detection based on switching kalman filters. IEEE Trans Biomed Eng 2015;62(9):2125–2134.

[8] Ródenas J, García M, Alcaraz R, Rieta JJ. Combined nonlinear analysis of atrial and ventricular series for automated screening of atrial fibrillation. Complexity 2017;2017.

[9] Faust O, Hagiwara Y, Hong TJ, Lih OS, Acharya UR. Deep learning for healthcare applications based on physiological signals: A review. Computer Methods and Programs in Biomedicine 2018;161:1–13.

[10] Behar J, Oster J, Li Q, Clifford GD. ECG signal quality during arrhythmia and its application to false alarm reduction. IEEE Transactions on Biomedical Engineering 2013;(6).

[11] Satija U, Ramkumar B, Manikandan MS. Automated ECG noise detection and classification system for unsupervised healthcare monitoring. IEEE J Biomed Health Inform 05 2018;22(3):722–732.

[12] Huerta A, Martínez-Rodrigo A, González VB, Quesada A, Rieta J, Alcaraz R. Quality assessment of very long-term ECG recordings using a convolutional neural network. In IEEE E-Health and Bioengineering Conference. 2019; 1–4.

[13] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems. 2012; 1097–1105.

[14] Shoeb A, Cliford G. Chapter 16—wavelets; multiscale activity in physiological signals. Biomedical Signal and Image Processing Springer New York NY USA 2005;.

[15] Satija U, Ramkumar B, Manikandan MS. A review of signal processing techniques for electrocardiogram signal quality assessment. IEEE Rev Biomed Eng 2018;11:36–52.

[16] Zhao Z, Zhang Y. SQI quality evaluation mechanism of single-lead ECG signal based on simple heuristic fusion and fuzzy comprehensive evaluation. Front Physiol 2018;9:727.

[17] Mincholé A, Rodriguez B. Artificial intelligence for the electrocardiogram. Nat Med 01 2019;25(1):22–23.

Informática Biomédica

Jueves, 26 de noviembre de 2020

Modelo de deep learning basado en la arquitectura Inception para el diagnóstico de la apnea del sueño infantil mediante la señal de oximetría

F. Vaquerizo Villar[1,4], D. Álvarez González[2,4], L. Kheirandish Gozal[3], G.C. Gutiérrez Tobal[1,4], V. Barroso García[1,4], E. Santamaría Vázquez[1,4], F. del Campo[2,4], D. Gozal[3], R. Hornero[1,4]

[1] Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, fernando.vaquerizo@gib.tel.uva.es

[2] Servicio de Neumología, Hospital Universitario Río Hortega, Valladolid, España, fsas@telefonica.net

[3] *Department of Child Health, The University of Missouri School of Medicine*, Columbia, Missouri, EEUU, gozald@health.missouri.edu

[4] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, (CIBER-BBN), España

Resumen

En este estudio se evalúa la utilidad de una arquitectura deep learning basada en módulos Inception para mejorar la capacidad diagnóstica de la señal de saturación de oxígeno en sangre (SpO_2) en la ayuda al diagnóstico de la apnea obstructiva del sueño (AOS) infantil. Estudios recientes demandan la aplicación de nuevas arquitecturas de deep learning que permitan solventar las limitaciones de los métodos convencionales. En este sentido, los módulos Inception permiten realizar un análisis multiescala, que es capaz de obtener una caracterización más precisa de las señales bajo estudio que otros métodos de deep learning. Nuestra propuesta consiste en la aplicación de una arquitectura Inception a la señal de SpO_2 para estimar el índice de apnea-hipopnea (IAH). El modelo Inception obtenido fue validado en una base de datos compuesta por 1628 registros de SpO_2 de sujetos pediátricos. El IAH estimado por el modelo Inception alcanzó un elevado rendimiento diagnóstico, mostrando un Cohen's kappa de cuatro clases de 0.520, así como precisiones del 79.2%, 95.7% y 97.8% para los umbrales del IAH de 1, 5 y 10 eventos por hora. Además, el nuevo modelo superó la capacidad diagnóstica de dos métodos convencionales: el índice de desaturación de oxígeno del 3% y una red neuronal perceptrón multicapa entrenada en un estudio previo con características de la señal de SpO_2. Estos resultados sugieren que la arquitectura Inception es capaz de obtener nueva información diagnóstica de la señal de oximetría y mejorar la estimación del IAH en el contexto de la AOS infantil.

1. Introducción

La apnea obstructiva del sueño (AOS) pediátrico es un trastorno de alta prevalencia en la población infantil (1%-5%) [1]. La AOS está caracterizada por la recurrencia de episodios de ausencia total (apneas) o parcial (hipopneas) de respiración durante el sueño de los niños, que derivan en un sueño no reparador [1,2]. En caso de no ser tratada, los niños afectados pueden sufrir múltiples consecuencias negativas para su salud y desarrollo, como alteraciones neuroconductuales, cardiovasculares y metabólicas [1,3]. La prueba de diagnóstico estándar es la polisomnografía (PSG) nocturna [1]. Esta prueba requiere que los niños pasen al menos una noche en una unidad especializada del sueño, donde se registran múltiples señales biomédicas mientras duermen [1,3]. Estos registros son analizados por especialistas médicos para anotar las apneas e hipopneas con el objetivo de estimar el índice de apnea-hipopnea (IAH), que es el parámetro utilizado para diagnosticar la AOS [2]. A pesar de su efectividad, la PSG es una prueba compleja, con un coste elevado, altamente intrusiva para los niños y que presenta disponibilidad limitada, produciéndose así un retardo en el diagnóstico y tratamiento de la AOS en los niños afectados [4].

Debido a estas limitaciones, la comunidad científica se ha centrado en los últimos años en la búsqueda de pruebas diagnósticas más sencillas como alternativa a la PSG [4]. En este sentido, la señal de saturación de oxígeno en sangre (SpO_2) procedente de la oximetría nocturna ha sido frecuentemente utilizada, ya que se trata de una técnica sencilla, fiable y no intrusiva [4]. Estudios previos han obtenido resultados prometedores utilizando enfoques de análisis automático de la señal de SpO_2 basados en la extracción y selección de características (*feature engineering*) [5-7]. Sin embargo, estos enfoques tradicionales tienen dos grandes limitaciones [8]: (*i*) requieren un gran conocimiento para identificar, *a priori*, qué características deben ser extraídas; y (*ii*) proporcionan un nivel de abstracción bajo, limitando así su capacidad de detectar patrones complejos en los datos. Estos inconvenientes pueden llevar a omitir información relevante de la señal de SpO_2 relacionada con los eventos de apnea e hipopnea (eventos apneicos).

Las limitaciones anteriores pueden ser minimizadas mediante la aplicación de algoritmos de aprendizaje profundo (*deep learning*), que permiten aprender de manera automática patrones complejos a partir de la señal en crudo mediante el empleo de arquitecturas con múltiples niveles de abstracción [8]. Estos algoritmos han mejorado el rendimiento de los métodos convencionales en áreas como el reconocimiento de imágenes, el procesado del lenguaje o el análisis de series temporales. [8]. Además, estudios previos han demostrado la utilidad de los métodos de *deep learning* aplicados a señales fisiológicas procedentes de la PSG para detectar fases del sueño, eventos apneicos y/o estimar el IAH en pacientes adultos de AOS [9]. En el contexto de la AOS infantil, solamente dos estudios preliminares de nuestro grupo de investigación han utilizado técnicas de *deep learning*, concretamente redes neuronales convolucionales (CNNs)

ISBN: 978-84-09-25491-0, pags. 340- 343

[10,11]. Estos estudios demandaron la necesidad de nuevas arquitecturas de *deep learning* que permitieran corroborar la superioridad de estos algoritmos con respecto a los enfoques tradicionales de *feature engineering*.

En este estudio, se propone el uso de una arquitectura *Inception* de *deep learning*, que, a diferencia de las CNNs, permite aprender características a múltiples escalas temporales de manera simultánea [12]. Trabajos anteriores han mostrado que la señal de SpO_2 contiene información relevante a múltiples escalas [4,7]. Por consiguiente, este estudio parte de la hipótesis de que la arquitectura *Inception* podría ayudar a extraer de manera automática características relevantes de la señal de SpO_2 a diferentes escalas en el contexto de la AOS infantil. El objetivo del estudio es, por tanto, evaluar la utilidad de una arquitectura *Inception* para estimar el IAH a partir de la señal de SpO_2 en niños con sospecha de AOS.

2. Sujetos y señales

En este estudio se ha analizado la base de datos pública *Childhood Adenotonsillectomy Trial* (CHAT), cuyo protocolo se encuentra en el material suplementario de [13]. CHAT está compuesta por 1628 estudios del sueño de niños de 5 a 10 años. Las apneas e hipopneas de cada sujeto fueron marcadas según las reglas de la Academia Americana de Medicina del Sueño [2]. El IAH, calculado como el número de apneas e hipopneas por hora de sueño, fue utilizado para diagnosticar la severidad de la AOS [2]. En este estudio, se han empleado los puntos de corte habituales de 1, 5 y 10 eventos por hora (e/h) del IAH para determinar los grupos de severidad no-AOS (IAH<1 e/h), AOS leve (1≤ IAH<5 e/h), AOS moderado (5≤IAH<10 e/h) y AOS severo (IAH≥10 e/h) [5-7].

Los registros de SpO_2 fueron adquiridos durante la PSG utilizando frecuencias de muestreo de 1 a 512 Hz, que posteriormente se remuestrearon a una frecuencia común de 1 Hz. Para cada sujeto, se dividieron las señales de SpO_2 en segmentos de 20 minutos (1200 muestras), que permiten detectar clusters de desaturaciones, cuya duración mínima es de diez minutos. Para cada segmento, se estimó el IAH parcial con las anotaciones de eventos apneicos proporcionadas por los especialistas [11]. La población bajo estudio se dividió en tres grupos: grupo de entrenamiento, utilizado para entrenar la arquitectura *Inception*; validación, empleado para la optimización de hiperparámetros; y test, utilizado para evaluar el rendimiento diagnóstico de la metodología. El conjunto de entrenamiento se compuso de 859 registros (25587 segmentos) de SpO_2 de los grupos *baseline* (453) y *follow-up* (406) de CHAT, mientras que el tercer grupo, *nonrandomized*, se dividió en grupo de validación (primeros 500 sujetos: 14372 segmentos) y grupo de test (últimos 279 sujetos: 8132 segmentos). La tabla 1 muestra las características clínicas y sociodemográficas de la población bajo estudio.

3. Metodología

3.1. Arquitectura *Inception*

La arquitectura propuesta está inspirada en el trabajo de Szegedy *et al*. [12] para la clasificación de imágenes. Esta arquitectura incorpora módulos *Inception* que procesan los datos de entrada mediante varios caminos en paralelo, cada uno de ellos compuesto de capas convolucionales con distinto tamaño de filtro, concatenando sus salidas [12]. Esto permite aprender características a distintas escalas en el mismo nivel de abstracción, lo que ha llevado a mejorar arquitecturas tradicionales de *deep learning* en clasificación de imágenes [12]. La figura 1 muestra la arquitectura *Inception* empleada en este estudio.

La sección de entrada de la red son los segmentos de 20 minutos de la señal de SpO_2. A continuación, la arquitectura propuesta procesa la entrada mediante λ_I módulos *Inception*. Cada módulo *Inception* consta de cuatro bloques convolucionales (*conv blocks*) en paralelo, cada uno de ellos compuesto por: (*i*) una convolución 1D (conv), que extrae los mapas de características (*feature maps*) [8]; (*ii*) *batch normalization* (BN), que normaliza los *feature maps* [8]; (*iii*) *rectified linear unit* (ReLU), que aplica una función de activación para introducir no-

	Todos	Entrenamiento	Validación	Test
Sujetos (n)	1628	859	500	279
Edad (años)	7[6-8]	7[6-8]	7[6-8]	7[6-8]
Varones (n)	777 (47.7%)	417 (48.5%)	225 (46.2%)	135 (49.8%)
IMC (kg/m^2)	17.2 [15.4-21.5]	17.3 [15.5-22.3]	17.1 [15.8-20.7]	17.5 [15.6-20.5]
IAH (e/h)	1.6 [0.6-4.7]	3.1 [1.4-6.9]	0.8 [0.4-1.5]	0.8 [0.4-1.7]

Los datos se presentan como mediana [rango intercuartil] o n (%).
IMC: Índice de masa corporal; IAH: índice de apnea-hipopnea

Tabla 1. *Características sociodemográficas y clínicas de la población bajo estudio*

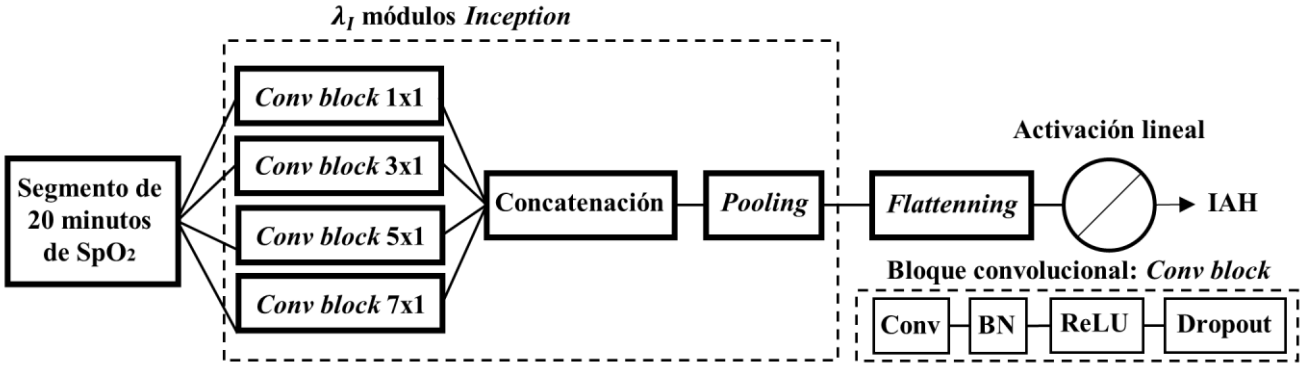


Figura 1. *Descripción de la arquitectura Inception propuesta.*

linealidad [8]; y (*iv*) *dropout*, que elimina conexiones entre elementos de la red durante el entrenamiento con probabilidad p_{drop} para minimizar el *overfitting* [8]. La diferencia entre los distintos *conv blocks* de cada módulo *Inception* es el tamaño de los filtros de la convolución 1D de cada bloque: 1x1, 3x1, 5x1 y 7x1. La salida de los distintos *conv blocks* de cada módulo *Inception* es concatenada [12], seguida de una capa de *max-pooling* con un factor *K*=3 que reduce la dimensionalidad [8]. Después de los λ_I módulos *Inception*, una capa de *flattenning* transforma los 2-D *feature maps* en datos 1-D [8]. Finalmente, una unidad de activación lineal es utilizada para obtener el IAH estimado para cada segmento de SpO_2 a la entrada. Para cada sujeto, se obtuvo el IAH como el valor promedio del IAH estimado para cada segmento de 20 minutos de la señal de SpO_2 correspondiente.

La implementación de la arquitectura *Inception* se llevó a cabo en la librería *Keras*. Los pesos de la red se inicializaron con el método He-normal, empleando el algoritmo de Adam con una tasa de aprendizaje inicial de 0.001 para actualizar los pesos en cada iteración [9]. Como función de error se utilizó la *Huber loss*, ya que tiene un hiperparámetro, delta (δ), que se puede ajustar para controlar la importancia de los *outliers*.

3.2. Comparación con métodos convencionales

Los siguientes métodos convencionales también han sido aplicados para comparar el rendimiento diagnóstico de la arquitectura *Inception*:

- Enfoque clínico: índice de desaturación de oxígeno del 3% (ODI3). ODI3 fue calculado como el número de desaturaciones de al menos el 3% por hora para cada señal de SpO_2. Este parámetro se suele utilizar en el contexto clínico de la AOS, así como con fines comparativos [4].
- Enfoque de *feature engineering* clásico: red neuronal perceptrón multicapa (MLP), entrenada por Hornero *et al* [5] para estimar el IAH con características extraídas de la señal de SpO_2. El modelo MLP diseñado por Hornero *et al* fue validado en una base de 4191 sujetos pediátricos [5], la mayor en la literatura. En el presente estudio, el modelo MLP ha sido aplicado a nuestro conjunto de test.

3.3. Análisis estadístico

El rendimiento diagnóstico del IAH estimado por la arquitectura *Inception* fue evaluado por medio de la matriz de confusión (de 4 clases), la precisión global (suma de la diagonal principal), *Cohen's kappa* (kappa), sensibilidad (S), especificidad (E) y precisión (P).

4. Resultados

4.1. Optimización de la arquitectura *Inception*

Los hiperparámetros de la arquitectura a optimizar fueron el número de módulos *Inception* (λ_I), la probabilidad de dropout (p_{drop}) y el parámetro delta de la *Huber loss* (δ). El conjunto de entrenamiento se utilizó para entrenar distintos modelos *Inception* variando estos parámetros, mientras que en el conjunto de validación se obtuvo el valor de kappa para cada uno de estos modelos. La combinación óptima de hiperparámetros que alcanzó un mayor kappa en validación fue: λ_I=5, p_{drop}=0.1 y δ=8.

4.2. Rendimiento diagnóstico del modelo *Inception*

La figura 2 muestra la matriz de confusión del modelo *Inception*, el ODI3 y la red MLP en el conjunto de test. El modelo *Inception* clasificó correctamente al 73.7% (205/279) de los sujetos según su severidad de la AOS, superando al ODI3 (64.2%, 179/279) y a la red MLP (63.8%, 178/279). En la tabla 2 se pueden ver las métricas de rendimiento diagnóstico de los tres métodos para los

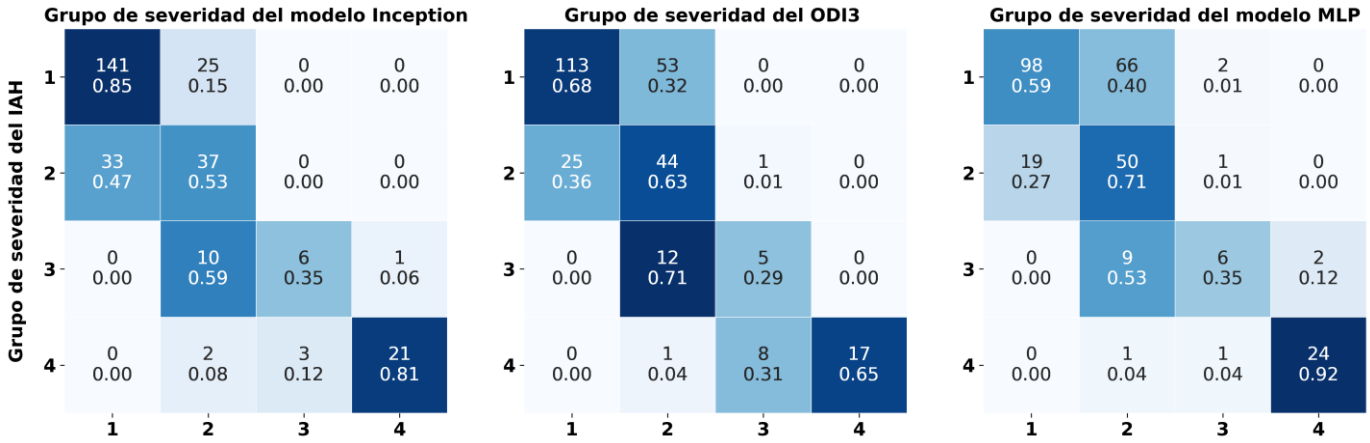


Figura 2. *Matriz de confusión de la red Inception, el ODI3, y la red MLP en el conjunto de test. 1: no-AOS (IAH<1 e/h); 2: AOS leve (1≤ IAH<5 e/h); 3: AOS moderado (5≤ IAH<10 e/h); 4: AOS severo (IAH≥10 e/h)*

IAH estimado	Inception			ODI3			MLP		
	IAH = 1 e/h	IAH =5 e/h	IAH = 10 e/h	IAH = 1 e/h	IAH =5 e/h	IAH = 10 e/h	IAH = 1 e/h	IAH =5 e/h	IAH = 10 e/h
S (%)	70.8	72.1	80.8	77.9	69.8	65.4	83.2	76.7	92.3
E (%)	84.9	100	99.6	68.1	99.6	100	59.0	98.7	99.2
P (%)	79.2	95.7	97.8	72.0	95.0	96.8	68.8	95.3	98.6
kappa	0.520			0.401			0.422		

Tabla 2. *Rendimiento diagnóstico de la red Inception, el ODI3, y la red MLP en el conjunto de test para los puntos de corte de 1, 5 y 10 e/h.*

puntos de corte del IAH de 1, 5 y 10 e/h. Nuestra propuesta alcanzó un kappa de 0.520, superior al del ODI3 (0.401) y la red MLP (0.422). Además, el modelo *Inception* superó al ODI3 y a la red MLP en la mayoría de las métricas de rendimiento para los tres puntos de corte del IAH, presentando una carga computacional similar en el conjunto de test.

5. Discusión y conclusiones

En este trabajo se ha evaluado la utilidad de una arquitectura novedosa de *deep learning* compuesta por módulos *inception* para mejorar la capacidad diagnóstica de la señal de SpO_2 en el contexto de la AOS infantil. El modelo propuesto alcanzó un elevado rendimiento diagnóstico para predecir la severidad de la AOS en niños, mejorando al ODI3 y a la red MLP diseñada en un relevante estudio previo [5]. En las matrices de confusión de la figura 2, se observa cómo nuestro modelo *Inception* supera de forma más notable al ODI3 y a la red MLP en la detección del grupo 1 (no-AOS). Por tanto, nuestra propuesta ayudaría a la reducción de las listas de esperas y los costes médicos asociados al diagnóstico, ya que en niños con un IAH<1 e/h sólo se recomienda realizar la PSG si persisten los síntomas de la AOS [3].

En los últimos años, varios estudios se han centrado en el uso de la señal de SpO_2 como ayuda al diagnóstico de la AOS infantil [4]. Estudios recientes han evaluado el rendimiento diagnóstico de metodologías de *feature engineering* para estimar el IAH en sujetos pediátricos [5-7]. Estos estudios alcanzaron precisiones en un rango de 75.2%-82.7%, 79.4%-81.9% y 88.2%-91.1% para los puntos de corte de 1, 5 y 10 e/h, respectivamente [5-7]. Sin embargo, solamente un estudio previo desarrollado por nuestro grupo aplicó una red CNN, la técnica más común de *deep learning*, a la señal de SpO_2 para estimar el IAH en la misma base de datos que el presente trabajo [11]. Este estudio alcanzó un kappa de 0.31 y precisiones del 74.8%, 90.7% y 95.1% para los puntos de corte de 1, 5 y 10 e/h [11], métricas que han sido superadas por la nueva arquitectura *Inception* propuesta.

Este estudio presenta algunas limitaciones. En primer lugar, el número de sujetos del grupo no-AOS es elevado en comparación con los demás grupos de severidad, especialmente los grupos AOS moderado y AOS severo. En este sentido, la validación de nuestra propuesta en otras bases de datos aumentaría la generalización de los resultados. Otra limitación es que el entrenamiento de la arquitectura *Inception* se ha realizado desde cero, al no existir arquitecturas de *deep learning* preentrenadas para señales 1D. Finalmente, la composición de los módulos *Inception* fue escogida siguiendo las recomendaciones de Szegedy *et al* [12]. Por consiguiente, se requieren pruebas adicionales para encontrar la configuración óptima de estos módulos para nuestro problema particular (señal de oximetría y detección de AOS).

En resumen, una arquitectura de *deep-learning* compuesta por módulos *Inception* ha demostrado ser útil para aprender de manera automática características de la señal de oximetría asociadas a los eventos apneicos en múltiples escalas temporales. Además, la arquitectura propuesta alcanzó un elevado rendimiento diagnóstico, superando al ODI3 y la red MLP, así como a los resultados obtenidos en estudios previos. Estos resultados sugieren que los módulos *Inception* pueden ser utilizados para mejorar la capacidad diagnóstica de la señal de oximetría en el contexto de la AOS infantil.

Agradecimientos

Este estudio ha sido financiado por el Ministerio de Ciencia Innovación y Universidades-Agencia Estatal de Investigación y FEDER (DPI2017-84280-R y RTC-2017-6516-1), por el CIBER-BBN (ISCII), por la Sociedad Española de Neumología y Cirugía Torácica (SEPAR 1649/218) y por la Sociedad Española de Sueño (Beca de Investigación SES 2019). F. Vaquerizo Villar cuenta una ayuda para la Formación de Profesorado Universitario del Ministerio de Educación, Cultura y Deporte (FPU16/02938). V. Barroso y E. Santamaría cuentan con una ayuda concedida por la Consejería de Educación de la Junta de Castilla y León y el Fondo Social Europeo.

Referencias

[1] C. L. Marcus *et al.*, "Diagnosis and Management of Childhood Obstructive Sleep Apnea Syndrome," *Pediatrics*, vol. 130, no. 3, pp. e714–e755, 2012.

[2] C. Iber *et al*, "The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specification," *J. Clin. Sleep Med.*, vol. 3, no. 7, p. 752, 2007.

[3] M. L. Alonso-Álvarez *et al.*, "Documento de consenso del síndrome de apneas-hipopneas durante el sueño en niños," *Arch Bronconeumol*, vol. 47, no. Supl 5, pp. 2–18, 2011.

[4] F. del Campo, *et al*, "Oximetry use in obstructive sleep apnea," *Expert Rev. Respir. Med.*, vol. 12, no. 8, pp. 665–681, 2018.

[5] R. Hornero *et al.*, "Nocturnal Oximetry–based Evaluation of Habitually Snoring Children," *Am. J. Respir. Crit. Care Med.*, vol. 196, no. 12, pp. 1591–1598, 2017.

[6] Z. Xu *et al.*, "Cloud algorithm-driven oximetry-based diagnosis of obstructive sleep apnoea in symptomatic habitually snoring children," *Eur. Respir. J.*, vol. 53, no. 2, 2019.

[7] F. Vaquerizo-Villar *et al.*, "Detrended fluctuation analysis of the oximetry signal to assist in paediatric sleep apnoea-hypopnoea syndrome diagnosis," *Physiol. Meas.*, vol. 39, no. 11, p. 114006, 2018.

[8] I. Goodfellow *et al*, *Deep Learning*. MIT Press, 2016.

[9] S. S. Mostafa *et al*, "A systematic review of detecting sleep apnea using deep learning," *Sensors (Switzerland)*, vol. 19, no. 22, pp. 1–26, 2019.

[10] F. Vaquerizo-Villar *et al.*, "Convolutional Neural Networks to Detect Pediatric Apnea- Hypopnea Events from Oximetry," in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2019, pp. 3555–3558.

[11] F. Vaquerizo-Villar *et al*, "Automatic Assessment of Pediatric Sleep Apnea Severity Using Overnight Oximetry and Convolutional Neural Networks," in *42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2020, pp. 633–636.

[12] C. Szegedy *et al.*, "Going Deeper with Convolutions," *arxiv*, 2015. [Online]. Available: https://arxiv.org/abs/1409.4842v1.

[13] 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, 2013.

Caracterización *in silico* de la duración de la repolarización y su variabilidad en el síndrome de QT Largo Tipo1 bajo estimulación β adrenérgica

D.A. Sampedro-Puente[1], F. Raphel[2], J. Fernandez-Bes[1], P. Laguna[1], D. Lombardi[2], E. Pueyo[1]

[1] Universidad de Zaragoza, IIS Aragón y CIBER-BBN, Zaragoza, España, {sampedro, jfbes, laguna, epueyo}@unizar.es

[2] INRIA y LJLL Sorbonne Université, París, Francia, {damiano.lombardi, fabien.raphel}@inria.fr

Resumen

El aumento de la variabilidad temporal de la repolarización se ha relacionado con el desarrollo de arritmias ventriculares en el síndrome de QT largo tipo 1 (LQT1), en particular bajo estimulación β-adrenérgica (β-AS). Sin embargo, los mecanismos iónicos subyacentes no se conocen completamente. La investigación computacional de tales mecanismos requiere primero del desarrollo de métodos capaces de reproducir las observaciones experimentales. En este trabajo se propone un método para la identificación de modelos computacionales de potencial de acción (AP) a partir de trazas de voltaje y se aplica a la investigación de la variabilidad de la repolarización en LQT1. El enfoque presentado combina el método Double Greedy Dimension Reduction (DGDR) junto con Unscented Kalman Filter (UKF) para la estimación de conductancias iónicas y niveles de fosforilación de modelos acoplados de AP y β-AS. La aplicación de este método DGDR-UKF sobre trazas de AP sintéticamente generadas para una población de células LQT1 permite, estimar con precisión los parámetros del modelo, con valores reducidos de incertidumbre y tiempo de convergencia en la estimación. Además, se comprueba que DGDR-UKF es capaz de replicar de forma fiable la distribución estadística de la duración del AP y de su variabilidad a corto plazo, tanto en condiciones basales como bajo β-AS, siendo los errores relativos inferiores al 4%. Finalmente, la estimación DGDR-UKF puede reproducir fenómenos arritmogénicos como las alternancias de AP. En conclusión, el método propuesto permite la caracterización de la duración de la repolarización y su variabilidad en LQT1, lo que se espera pueda ayudar a dilucidar los mecanismos subyacentes al desarrollo de arritmias inducidas por β-AS en LQT1.

1. Introducción

El síndrome de QT largo (LQT) es una enfermedad congénita que genera una prolongación anormal del intervalo QT y predispone a un alto riesgo de desarrollar arritmias potencialmente mortales. La variante más común de LQT es LQT tipo 1 (LQT1), en la cual más del 90% de los eventos letales ocurren durante episodios de estrés físico o emocional, generalmente asociados a una elevada estimulación simpática [1]. Diferentes estudios *in vivo* e *in vitro* de LQT1 han descrito un vínculo entre el aumento de la variabilidad espacio-temporal de la repolarización y el mayor riesgo de sufrir arritmias ventriculares y muerte cardíaca súbita, en particular en respuesta a la estimulación adrenérgica (β-AS) [2]. Se requiere una mayor investigación acerca de los mecanismos iónicos implicados en estas observaciones para poder establecer marcadores capaces de anticipar el riesgo de arritmias, y desarrollar terapias específicas para el síndrome LQT1.

El modelado y la simulación computacional se han convertido en una herramienta de gran valor para complementar la investigación experimental y clínica, contribuyendo a avanzar en la comprensión de la electrofisiología cardiaca, tanto en condiciones fisiológicas como patofisiológicas. En particular, se ha demostrado que las metodologías para la estimación de parámetros de modelos computacionales de potencial de acción (AP) a partir de datos experimentales tienen gran potencial en el desarrollo de modelos personalizados *in silico*. Estos modelos permiten investigar mecanismos iónicos que subyacen a ciertas propiedades electrofisiológicas y su relación con el riesgo arrítmico, reduciendo así el número de experimentos, que se guiarán de acuerdo con las predicciones obtenidas mediante simulación computacional. En estudios recientes, se ha demostrado que los métodos de estimación de parámetros basados en Double Greedy Dimension Reduction (DGDR) [3] y Kalman Filter (UKF) [4] pueden identificar, con altos niveles de precisión, los parámetros y variables de estado de modelos ventriculares humanos de AP con estocasticidad [5, 6].

En este trabajo evaluamos los métodos UKF y DGDR, tanto de manera combinada como independiente, para la identificación de parámetros de modelos acoplados de electrofisiología y señalización β-AS a partir de trazas de voltaje AP de entrada. Sobre la base de esta identificación, establecemos la capacidad de los métodos para reproducir la variabilidad de la repolarización espacio-temporal en el síndrome LQT1 y caracterizar fenómenos arritmogénicos relacionados con el aumento de la variabilidad.

2. Materiales y Métodos

2.1. Población de modelos de AP estocásticos

Tomando como base el modelo ventricular humano de O'Hara *et al.* [7], se construyó una población de modelos AP estocásticos, con representación de variabilidad espacio-temporal, variando las conductancias de las ocho principales corrientes iónicas (I_{Ks}, I_{Kr}, I_{to}, I_{CaL}, I_{K1}, I_{Na}, I_{NaCa} e I_{NaK}) en el rango ±100% de sus valores nominales. Inicialmente se generaron 8000 modelos por el método de Monte-Carlo, los cuales se redujeron a 2373 modelos tras un proceso de calibración basado en rangos fisiológicos de propiedades electrofisiológicas [6]. Cada uno de estos modelos calibrados representa una célula virtual con distintas propiedades iónicas.

ISBN: 978-84-09-25491-0, pags. 344- 347

La respuesta a β-AS se introdujo en los modelos computacionales utilizando la formulación descrita por Xie *et al.* [8, 9], junto con las actualizaciones posteriores descritas en [10]. Consideramos que los efectos de β-AS son iguales para todas las células, excepto por la fosforilación de los tres sustratos celulares con un impacto más notable en el AP, es decir, I_{Ks}, I_{CaL} e I_{NaK} [10]. Los factores de multiplicación aplicados a los niveles de fosforilación por PKA de estos tres sustratos, denominados f_{CaL}, f_{Ks} y f_{NaK}, se variaron en rangos que iban desde los valores de referencia (sin β-AS) hasta los valores alcanzados tras la aplicación de una dosis de isoproterenol de 1μM, la cual produce efectos máximos en el AP. Para la generación de esta población de factores multiplicadores también se usó el método de Monte-Carlo.

Por medio de la combinación de las variaciones descritas en las 8 conductancias iónicas y en 3 niveles de fosforilación, se construyó una población de 2373 modelos celulares estocásticos que representaban tanto las condiciones basales como las de respuesta a β-AS. Además, se modeló el síndrome LQT1 a través de la reducción en un 90% de la conductancia de la corriente de I_{Ks}.

2.2. Estimación de parámetros

Para cada traza de voltaje de entrada, que podría haber sido eventualmente obtenida de forma experimental o generada sintéticamente, se aplicó el método de estimación que combina DGDR y UKF para inferir los parámetros del modelo computacional subyacente. Tal y como se ha descrito en la sección 2.1, estos parámetros representan las diferentes conductancias iónicas y niveles de fosforilación de un modelo electrofisiológico de AP con representación de β-AS.

DGDR es un método diseñado para el análisis de datos de alta dimensión [3]. En este trabajo se aplicó en una primera etapa para la estimación de los parámetros del modelo AP por medio de la proyección de los datos de entrada en un sub-espacio a través de una combinación *sparse* lineal de las entradas de un diccionario previamente definido. Este diccionario estaba formado por diversos marcadores derivados del AP, entre los que se incluían la duración del AP para diferentes porcentajes de repolarización, la variabilidad a corto plazo (STV) de la duración del AP, el voltaje máximo, el potencial de reposo o las diferentes componentes de la descomposición *wavelet*, así como las combinaciones lineales de estos marcadores [6].

Las estimaciones de parámetros obtenidas mediante la aplicación de DGDR se utilizaron en una segunda etapa para inicializar las estimaciones del método basado en el UKF. Este método UKF permite estimar los parámetros del modelo AP por medio de la formulación de representaciones estado-espacio no lineales y de la aproximación de las distribuciones a posteriori de dichos estados mediante la propagación de los denominados *Sigma-Points* [11]. Además, al final de cada latido, las estimaciones obtenidas por UKF se actualizaron, en términos de media y desviación estándar, combinándolas con la estimación por DGDR, evitando así que la estimación pudiera caer en mínimos locales lejanos de los valores reales de los parámetros.

Se evaluó el comportamiento del método combinado DGDR-UKF en la población calibrada experimentalmente de modelos AP, previamente descrita en la sección 2.1, comparándose con los obtenidos por DGDR y UKF individualmente, tanto en condiciones basales como en respuesta a β-AS. Para este fin, la población de modelos AP se dividió en una subpoblación de entrenamiento de 2.000 células virtuales y una subpoblación de validación de 373 células virtuales. En un paso posterior, se aplicaron las condiciones LQT1 a las poblaciones de entrada y a las generadas a partir de las estimaciones obtenidas, comparando los resultados en términos de reproducción de los marcadores derivados del AP descritos en la siguiente sección.

2.3. Marcadores del AP

Los marcadores derivados del AP se calcularon a partir de N latidos de las trazas de AP de entrada y las estimadas e incluyeron: APD, calculada como la media de las duraciones de AP al 90% de repolarización (APD_{90}) de los latidos individuales, y STV, calculada como la distancia media perpendicular a la línea de identidad en el gráfico de Poincaré del APD_{90} [5]:

$$\mathrm{APD} = \frac{1}{N}\sum_{n=1}^{N} \mathrm{APD}_{90}(n)$$

$$\mathrm{STV} = \sum_{n=1}^{N-1} \frac{|\mathrm{APD}_{90}(n+1) - \mathrm{APD}_{90}(n)|}{(N-1)\sqrt{2}}$$

3. Resultados y Discusión

3.1. Variabilidad de repolarización en LQT1

La Fig. 1 muestra las gráficas de Poincaré del APD_{90} para dos células virtuales de la población de modelos descrita en la Sección 2 (modelos #5 y #330) en los cuatro escenarios considerados: basal, β-AS, LQT1 y LQT1 con β-AS. A partir de las trazas de voltaje de cada una de estas células se obtuvieron las diferentes estimaciones de las conductancias iónicas y de los niveles de fosforilación, tanto en situación basal como en respuesta a β-AS, para cada uno de los métodos de estimación DGDR, UKF y DGDR-UKF descritos en la Sección 2.2. A partir de dichas estimaciones se generaron nuevas trazas de AP estimadas, para las cuales se presentan los gráficos de Poincaré de APD_{90} en la misma figura. Como puede observarse, el método combinado DGDR-UKF permitió estimar con mayor precisión los datos de entrada que los métodos DGDR o UKF individualmente. Esta mejor estimación puede apreciarse no solo en términos del APD_{90} medio, sino también de su variabilidad temporal. Tal como puede observarse en la figura, el método combinado DGDR-UKF proporcionó unas nubes de puntos de los datos estimados mucho más semejantes a las de los datos de entrada. Tras la simulación de condiciones de LQT1 por medio de la inhibición de la corriente I_{Ks}, las diferencias entre los distintos métodos de estimación se acentuaron todavía más, pudiendo observarse que el método DGDR-UKF

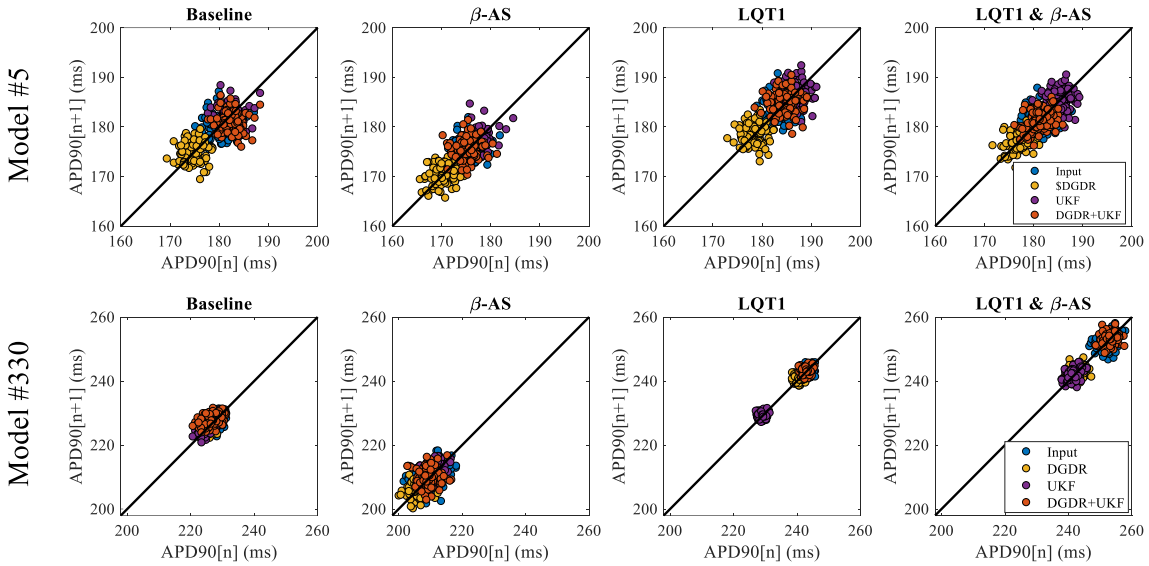


Figura 1. *Gráficos de Poincaré del APD90 para dos células virtuales diferentes (modelo #5 y #330 de la población definida en la sección 2.1), calculados a partir de la traza de voltaje de entrada (azul) y a partir de las trazas generadas a partir de las estimaciones por DGDR (amarillo), UKF (morado) y DGDR-UKF (naranja) bajo cuatro condiciones: basal, β-AS, LQT1 y LQT1 con β-AS*

reproduce con más alta fiabilidad tanto la duración, APD, como la variabilidad, STV, de la repolarización ventricular.

El rendimiento superior del método combinado DGDR-UKF ilustrado en la Fig. 1 para dos células particulares, se confirmó de forma generalizada en toda la población de células del estudio. La Fig. 2 muestra la función de densidad de probabilidad del error calculado como la diferencia del APD medio (análogamente para STV) de las trazas de AP de entrada y las trazas de AP estimadas. En el caso de LQT1 con β-AS se produjo una reducción del error absoluto medio del APD de 45% y 25% con respecto a los métodos individuales DGDR y UKF, respectivamente, y de

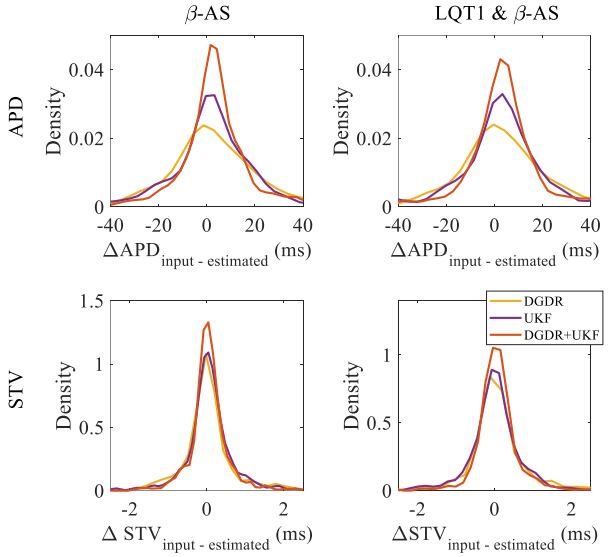


Figura 2. *Funciones de densidad de probabilidad del error entre el APD (paneles superiores) y STV (paneles inferiores) de entrada y el estimado para dos diferentes escenarios, β-AS (izquierda) y LQT1 con β-AS (derecha)*

47% y 27% para la desviación estándar de dicho error absoluto. También en la evaluación de STV se produjo una disminución notable del error. Estas funciones de densidad se presentan para dos de los escenarios analizados: condición β-AS y condición LQT1 con β-AS. A partir de estos resultados, puede concluirse que la metodología DGDR-UKF es capaz de representar con mayor precisión la duración y la variabilidad de la repolarización del AP en comparación con los métodos individuales DGDR y UKF. Es importante resaltar que otros métodos propuestos en la literatura, como los basados en Moment-Matching, Genetic Algorithms o Gaussian Process Emulators, no permiten la estimación de los valores óptimos de los parámetros del modelo para cada individuo (célula) de la población. Por otra parte, métodos basados en Markov Chain Monte Carlo, que sí permiten la estimación individual, tienen como gran desventaja el alto coste computacional asociado a la estimación.

Es interesante destacar que la combinación DGDR-UKF pudo reproducir fielmente los valores de APD y STV en todos los modelos de la población, independientemente de si β-AS y LQT1 se asociaron con un aumento, disminución o preservación del APD. Esto es relevante, particularmente considerando el alto nivel de variabilidad inter-individual en tales respuestas, con cambios en el APD que van desde los -86 ms a 12 ms en respuesta a β-AS en comparación con condiciones basales, de 1 ms a 196 ms en LQT1 en comparación con la condición basal y de -24 ms a 178 ms bajo LQT1 con β-AS en comparación con basal sin β-AS.

3.2. Alternancias proarrítimicas de la repolarización bajo condiciones de LQT1

Para evaluar en mayor profundidad el rendimiento de los métodos de estimación analizados, se analizó su capacidad para reproducir alternancias en el APD inducidas por la estimulación adrenérgica en condiciones de LQT1. La Fig.

3 muestra las trazas de entrada y las trazas de AP estimadas para una célula virtual de la población cuando se estimula a una frecuencia de 5.3Hz. Como puede observarse en dicha figura, los APD presentan valores alternados alto-bajo-alto-bajo a esta frecuencia. Para esa misma célula, se observaron alternancias de APD para todas las frecuencias en el rango de 4.5 a 5.5 Hz. Este fenómeno pudo reproducirse apropiadamente por el AP generado a partir de la estimación por DGDR-UKF, mostrando el patrón alterno de la serie temporal de APD y a su vez reproduciendo la morfología de los APs de la traza de entrada. Por otra parte, para esta misma célula virtual de la Fig. 3, los métodos individuales DGDR y UKF no produjeron alternancias para ninguna frecuencia de

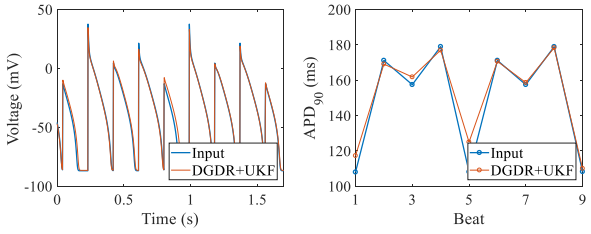


Figura 3. *Izquierda: AP de Entrada (azul) y AP generado a partir de la estimación por DGDR (naranja) para una célula virtual de la población de modelos descrita en la Sección 2, mostrando las alternancias del APD inducidas por la β-AS bajo LQT1 al estimularse a 5.3Hz. Derecha: serie temporal de APD correspondiente.*

estimulación entre 2 y 10 Hz. Este resultado pone de relieve la necesidad de contar con métodos de identificación robustos y precisos, ya que pequeños errores en la estimación de las conductancias de las corrientes iónicas y/o de los niveles de fosforilación en respuesta a β-AS pueden dar lugar a que los modelos AP no sean capaces de reproducir los fenómenos proarrítmicos relevantes observados en los datos de entrada originales, como son las alternancias de repolarización, o de reproducirlos pero en un rango diferente de frecuencias de estimulación, como se observa en otras células de nuestra población calibrada experimentalmente. Finalmente es interesante resaltar que aunque el patrón de alternancias que se muestra en la Fig. 3 no es el típico patrón 1:1 de valores corto-largo, con todos los APD cortos (respectivamente, largos) teniendo valores similares, la serie temporal de APD muestra diferencias entre los APD de latidos consecutivos superiores a los 14 ms en todos los casos.

4. Conclusiones

En este trabajo se ha evaluado una metodología que combina el método de reducción de dimensiones DGDR y el filtro UKF no lineal para la identificación robusta y fiable de modelos computacionales de electrofisiología y señalización β-adrenérgica a partir de trazas de voltaje de entrada. Se ha probado que el método propuesto permite la caracterización *in silico* de la duración de la repolarización y de su variabilidad espacio-temporal en condiciones que representan el síndrome LQT1. De manera importante, el método permite replicar fenómenos proarrítmicos observados en LQT1 bajo β-AS, tales como las alternancias de APD. Este trabajo sienta las bases para futuras investigaciones encaminadas a dilucidar los mecanismos iónicos que subyacen a la generación de arritmias en el síndrome LQT1, particularmente en respuesta a β-AS.

Agradecimientos

Este trabajo ha sido financiado por los proyectos ERC-StG 638284 (ERC), PID2019-105674RB-I00 y PID2019-104881RB-I00 (MICINN), LMP124-18, y grupo T39-20R, por el contrato predoctoral personal (Gobierno de Aragón y FEDER 2014-2020 "Construyendo Europa desde Aragón"), las ayudas a la movilidad del Campus Iberus Erasmus+ y de la Fundación CAI-Ibercaja a D.A., y por la financiación del INRIA. Los cálculos se han realizado en el ICTS NANBIOSIS (Unidad HPC, U. Zaragoza).

Referencias

[1] Toivonen L. *et al.* Genotype-Phenotype Correlation in the Long-QT Syndrome. *Circulation* 2001;103(1):89–95. ISSN 0009- 7322.

[2] Gallacher DJ. *et al.* In vivo mechanisms precipitating Torsades de Pointes in a canine model of drug-induced long-QT1 syndrome. *Cardiovas. Res.* 2007; 76(2):247–256.

[3] Lombardi D, Raphel F. A greedy dimension reduction method for classification problems. *HAL* 2019;hal–02280502.

[4] Julier SJ, Uhlmann JK. A new extension of the Kalman filter to Nonlinear systems. *In AeroSense '97. Proc. of AeroSense: The 11th Int. Symp. on Aerospace/Defence Sensing. Simulation and Controls*., 1997; 182–193.

[5] Sampedro-Puente DA. *et al.* Mechanisms Underlying Interactions Between Low-Frequency Oscillations and Beat-to-Beat Variability of Celullar Ventricular Repolarization in Response to Sympathetic Stimulation: Implications for Arrhythmogenesis. *Front. Physiol.* 2019;10:916.

[6] Sampedro-Puente DA. *et al.* Characterization of spatiotemporal cardiac action potential variability at baseline and under β-adrenergic stimulation by combined unscented kalman filter and double greedy dimension reduction. *IEEE J. Biomed.Health Inform*. 2020;ISSN 2168-2194.

[7] O'Hara T. *et al*. Simulation of the undiseased human cardiac ventricular action potential: model formulation and experimental validation. PLoS Comput Biol 2011;7(5):e1002061. ISSN 1553-7358.

[8] Xie Y. *et al.* β-adrenergic stimulation activates early afterdepolarizations transiently via kinetic mismatch of PKA targets. *J. Mol. Cell. Cardiol.* 2013;58:153–161.

[9] Soltis AR, Saucerman JJ. Synergy between CaMKII substrates and β-adrenergic signaling in regulation of cardiac myocyte Ca2+ handling. *Biophys. J.* 2010;99(7):2038–2047.

[10] Pueyo E. *et al.* Interactive effect of beta-adrenergic stimulation and mechanical stretch on low-frequency oscillations of ventricular action potential duration in humans. *J. Mol. Cell. Cardiol.* 2016;97:93–105.

[11] Sampedro-Puente DA. *et al.* Data-driven identification of stochastic model parameters and state variables: Application to the study of cardiac beat-to-beat variability. *IEEE J. Biomed. Health Inform.* 2020;24(3):693–704.

Segmentación automática de instancias aplicada a retinografías para la segmentación del disco óptico y la excavación

O. Kovalyk[1], J. Morales Sánchez[1], R. Verdú Monedero[1], I. Sellés Navarro[2]

[1] Universidad Politécnica de Cartagena, Cartagena, España,

oleksandr.kovalyk8@gmail.com, {juan.morales, rafael.verdu@upct.es

[2] Hospital Universitario Reina Sofía, Murcia, España, inmasell@um.es

Resumen

El glaucoma es una enfermedad ocular silenciosa que sin el tratamiento adecuado produce ceguera. Uno de los signos clínicos que sugieren esta enfermedad en sus fases iniciales son los niveles elevados de presión intraocular que producen alteraciones en la morfología del nervio óptico. La relación entre el diámetro de la excavación papilar y del disco óptico se considera, para valores superiores a 0.5, un indicador de posible glaucoma. En este artículo se propone el algoritmo de aprendizaje máquina Mask R-CNN para la segmentación del disco óptico y la excavación en retinografías de la base de datos RIMONEr3. Los resultados muestran el rendimiento del algoritmo así como la capacidad de delimitar los contornos con gran precisión, comparado con segmentaciones de expertos y otros métodos del estado del arte.

1. Introducción

El desarrollo de aplicaciones para el procesado de imágenes representa una ayuda importante en la detección de enfermedades oculares como el glaucoma [1]. Sin el debido tratamiento, esta enfermedad ocular, puede conducir a un deterioro irreversible de la vista y finalmente a una ceguera total, por lo que el desarrollo de herramientas para la detección del glaucoma en sus fases tempranas es esencial para preservar la salud ocular y minimizar los costes sanitarios [2].

Se sabe que los indicadores para el diagnóstico de la enfermedad aparecen años antes que las lesiones irreversibles en el campo visual. Los estudios principales con sus indicadores son la tonometría con la presión intraocular, la campimetría con el campo visual, y las retinografías con la relación entre el diámetro de la excavación y el diámetro del disco óptico (*cup to disc ratio, CDR*). Muchos métodos se basan inicialmente en la morfología del nervio óptico mediante el CDR en retinografías, ya que el aumento de la presión intraocular hace que la excavación se agrande, resultando en un incremento del valor de CDR a lo largo del tiempo. Para la medida del CDR es necesario disponer de la segmentación del disco y de la excavación, que se puede hacer de forma manual por expertos o mediante métodos automáticos de procesado de imagen [1]. Entre los métodos automáticos se pueden distinguir aquellos basados en técnicas de procesado de imagen como level-sets [3], umbrales [4], etc.

Por otro lado, los últimos avances en el campo del aprendizaje máquina, junto a la gran capacidad computacional a un precio accesible, han traído gran cantidad de algoritmos capaces de realizar segmentaciones automáticas. Como ejemplos, podemos encontrar métodos basados en aprendizaje máquina como redes M-net [5], parches [6], etc. En este trabajo se propone la red Mask R-CNN [7] para el análisis de las prestaciones de la segmentación automática de la excavación y del disco óptico en retinografías. A diferencia de otras redes utilizadas con más frecuencia en imágenes médicas, Mask R-CNN se emplea normalmente para detección y seguimiento de objetos en aplicaciones donde los requerimientos de tiempo de cómputo son prioritarios en relación a la propia segmentación del objeto, por lo que no es de uso común en aplicaciones médicas.

2. Método

2.1. Base de datos de retinografías

Las retinografías empleadas en este trabajo proceden de la base de datos de RIMONEr3 [8], la cual contiene 159 retinografías, 74 de las cuales corresponden a pacientes sospechosos de glaucoma, y las 85 restantes a pacientes sanos. Cada retinografía tiene asociadas cuatro segmentaciones manuales, dos segmentaciones del disco óptico (una por cada experto) y otras dos de las excavaciones o copas ópticas.

2.2. Arquitectura de la red neuronal

Mask R-CNN [7] es un algoritmo de segmentación semántica, que se encarga de localizar la ubicación de diferentes objetos en una imagen, clasificar dichos objetos, y por último calcular las máscaras binarias que permitirían delimitar o extraer cada uno de estos objetos de la imagen. Partiendo de una imagen RGB de dimensiones $N \times M \times 3$, si esta imagen contiene K objetos, entonces el tensor que representa el conjunto de máscaras tendría dimensiones $N \times M \times K$. Cada uno de los K canales es una matriz o máscara binaria $N \times M$ que determina la forma y posición de un objeto en la imagen.

El algoritmo Maks R-CNN [7] se compone de 4 subredes, que se muestran en el diagrama de la Figura 1: la red CNN (*Convolutional Neural Network*) [9] + FPN (*Feature Pyramid Network*) [10], la red RPN (*Region Proposal*

ISBN: 978-84-09-25491-0, pags. 348- 351

Network) [11], la red RoI (*Region of Interest*) Align [7] y la red del cálculo de la máscara.

En primer lugar, la imagen se procesa por la red CNN + FPN [9, 10]. FPN [10] es una red que extrae de una imagen un determinado mapa de características. La red FPN tiene dos salidas, una dirigida hacia la red RPN y otra hacia la red RoI Align.

La segunda red en actuar es la RPN [11], que procesa el mapa de características y propone posibles regiones de ubicación tanto para el disco óptico como para la excavación. Los cuadriláteros mostrados en la Figura 2 representan todo este conjunto de posibles localizaciones generadas.

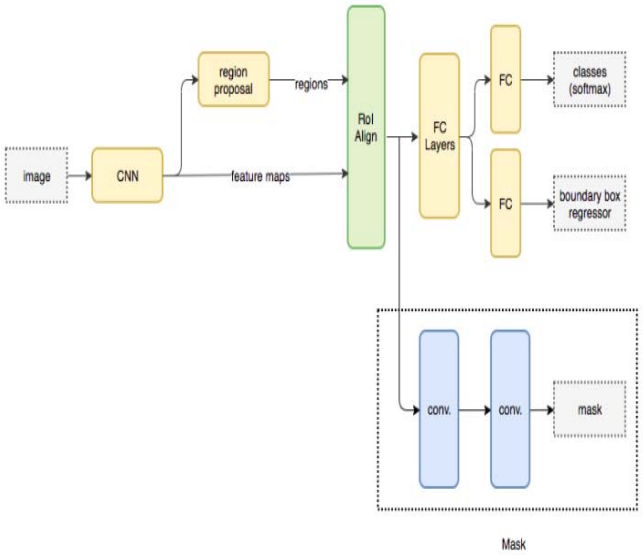


Figura 1. *Diagrama de bloques de Mask R-CNN.*

Estas potenciales ubicaciones, que pueden ser positivas o negativas, dependiendo respectivamente de si se proponen como candidatas a contener el objeto o a no contenerlo, pasan a la entrada de la posterior red RoI Align.

RoI Align es una red que se encarga de dos tareas básicas: clasificar la región propuesta por RPN (disco óptico o excavación), y ajustar el tamaño y la posición de las regiones positivas (regiones de interés) para maximizar el valor de la Intersección sobre la Unión (IoU) con la región verdadera (GT) delimitada por el experto (ver Figura 3), y finalmente se selecciona la RoI con mayor IoU, de entre aquellas regiones de interés que superen un cierto umbral mínimo de IoU, como se puede ver en la Figura 4.

Como se observa en la Figura 1, en paralelo a la red RoI Align se obtiene la máscara binaria correspondiente a la RoI con mayor índice de IoU, lo cual supone la segmentación en sí misma. Para ello se clasifican los píxeles que son parte del objeto y los que no lo son mediante una red convolucional, llegando a un resultado similar al ilustrado en la Figura 5.

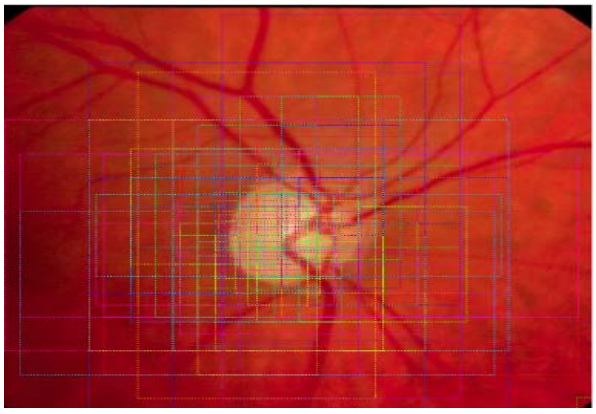


Figura 2. *Cuadriláteros generados por la red RPN, que en aras a la claridad se representan sobre la la retinografía, aunque en realidad se generan a partir del mapa de características.*



Figura 3. *Ajuste de las regiones de interés por la red RoI Align.*



Figura 4. *RoI seleccionada tras umbralizar por el valor de IoU. En línea discontinua la región de interés antes del ajuste de tamaño y en línea continua la versión final.*

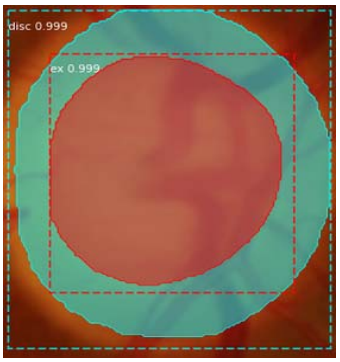


Figura 5. *Segmentaciones del disco óptico y la excavación. Máscara binaria de forma circular y región de interés en línea discontinua.*

2.3. Implementación

La implementación del algoritmo Mask R-CNN utilizada en este trabajo es una variante [12] de la propuesta inicialmente [7], y está programada en lenguaje Python, con soporte de la biblioteca de código abierto para aprendizaje automático TensorFlow de Google.

Los experimentos se han realizado en una misma máquina Manjaro Linux, con kernel 5.8.6, CPU Intel Core i7-8700K, GPU Nvidia RTX 2080Ti, y versión 1.15 de TensorFlow.

2.4. Entrenamiento

Para validar las prestaciones del algoritmo sobre la base de datos completa se ha utilizado una estrategia de validación cruzada de 5 iteraciones, de forma que se ha dividido la base de datos en 5 subconjuntos. Cada subconjunto está formado por 32 imágenes (a excepción de uno que tendrá 31), lo que supone aproximadamente un 20% de la base de datos. Cada uno de estos conjuntos se emplea en una iteración para test, mientras que el resto de la base de datos (4 subconjuntos) se utiliza para entrenamiento y validación.

3. Resultados

Esta sección muestra los resultados obtenidos con el método propuesto, centrándonos en medidas comparativas que reflejen el aspecto que se pretende evaluar: las prestaciones del algoritmo en cuanto a segmentación del disco óptico y la excavación. Para este fin, utilizaremos los habituales índices de Jaccard y Dice. Adicionalmente se corroborará la eficiencia en cuanto a tiempo de cómputo en la segmentación, principal factor por el que en otras aplicaciones se prefiere este algoritmo.

3.1. Segmentación

Como se ha indicado anteriormente, para medir la calidad de la segmentación obtenida se utilizarán los índices de Jaccard, definido mediante la expresión (1) como el valor de la Intersección sobre la Unión (IoU), y el índice de Dice, definido mediante la expresión (2).

$$J(Y,\hat{Y}) = \frac{|Y \cap \hat{Y}|}{|Y \cup \hat{Y}|} \quad (1)$$

$$D(Y,\hat{Y}) = \frac{2|Y \cap \hat{Y}|}{|Y| + |\hat{Y}|} \quad (2)$$

En ambos casos $\hat{Y}$ representa el valor predicho e Y es el valor verdadero.

	Disco		Copa	
	Dice	Jacc	Dice	Jacc
Exp. #1 vs #2	**97.44**	**95.17**	**89.96**	**84.13**
Predicción	**96.83 ± 1.64**	**94.09 ±2.79**	**77.31 ± 29.40**	**71.54 ±27.63**

Tabla 1. *Prestaciones (%) en media±desviación típica de Mask R-CNN sin técnicas de aumento de datos.*

La Tabla 1 muestra los resultados proporcionados por Mask R-CNN sobre RIMONEr3, sin hacer uso de técnicas adicionales de aumento de datos. Como se puede observar la predicción media en 5 iteraciones se queda muy cerca de los valores medidos entre el Experto #1 y Experto #2.

La Tabla 2 muestra de nuevo los resultados añadiendo nuevas imágenes al conjunto de entrenamiento utilizando aumento de datos mediante efecto espejo (simulando el ojo contrario). Con ello se consigue una mejora del 10% respecto a los valores de la Tabla 1.

	Disco		Copa	
	Dice	Jacc	Dice	Jacc
Exp. #1 vs #2	97.44	95.17	89.96	**84.13**
Predicción	**97.16 ±1.34**	**94.66 ±2.32**	**87.11 ± 14.51**	**81.05 ± 14.50**

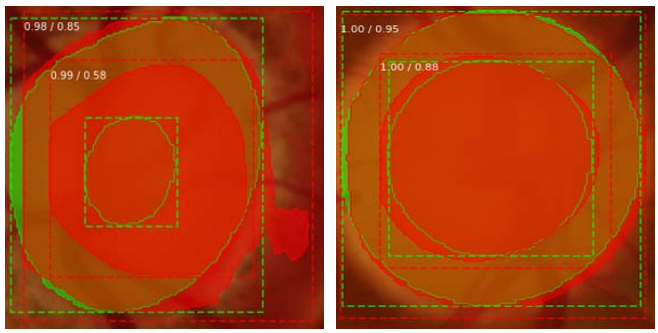
Tabla 2. *Prestaciones (%) en media±desviación típica de Mask R-CNN con técnicas de aumento de datos.*

La Tabla 3 muestra una comparativa con algunas de las técnicas de segmentación más recientes y que mejores resultados ofrecen para las mismas tareas de segmentación y base de datos de retinografías utilizada en este trabajo. Como podemos ver, Mask R-CNN con técnicas de aumento no solo supera en media a todos los algoritmos, sino que además en 3 de las 4 medidas mejora la diferencia o discrepancia entre las anotaciones de ambos expertos.

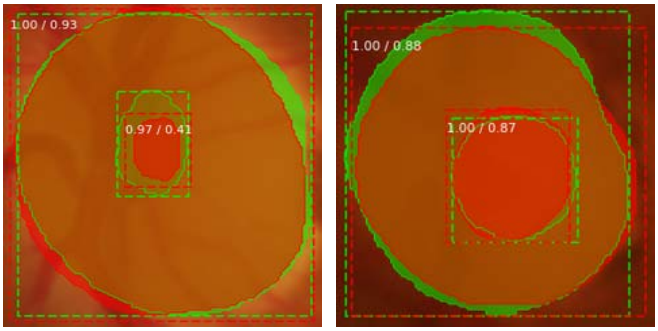
Dado que a la vista de la Tabla 1 y Tabla 2 las segmentaciones del disco óptico son en general de gran exactitud, en la Fig. 6 se ilustran ejemplos de segmentación de la excavación óptica, sobre la cual existen diferencias más sustanciales entre retinografías. En concreto podemos observar los dos peores casos de segmentación de la excavación óptica, así como los dos mejores. En dichas imágenes se pueden distinguir dos regiones: la más exterior corresponde al disco óptico y la más interior a la excavación. A su vez cada región incluye dos segmentaciones: la del experto en color verde y la automática en color rojo. Cada región aparece etiquetada con dos valores asociados: el valor de la izquierda es la certeza con la cual el algoritmo cree que el objeto encontrado pertenece a dicha clase (disco óptico o excavación) y el valor de la derecha es la concordancia en términos de IoU entre la segmentación proporcionada por la red y la segmentación objetivo (del experto).

	Disco		Copa	
	Dice	Jacc	Dice	Jacc
Exp. #1 vs #2	97.44	95.17	89.96	84.13
Sevastopolsky (2017)	95	89	82	69
Zilly et al.(2017)	94.2	89.0	82.4	80.2
Al-Bander and Zheng (2018)	90.36	82.89	69.03	55.67
U-Net media (2019)	94.91	90.65	79.32	68.28
Propuesto Mask R-CNN	**97.16 ±1.34**	**94.66 ± 2.32**	**87.11 ± 14.51**	**81.05 ± 14.50**

Tabla 3. *Comparativa con otros algoritmos de segmentación (%). Datos extraídos de otros trabajos comparativos [13].*



a) *b)*



c) *d)*

Figura 6. *Ejemplos de segmentaciones realizadas por el algoritmo Mask R-CNN. a) y c) peor caso y b) y d) mejor caso.*

3.2. Rendimiento computacional

La Tabla 4 y Tabla 5 recogen respectivamente los tiempos de cómputo en entrenamiento y en modo inferencia del algoritmo sobre la base de datos RIMONEr3.

	Sin aumento	Con aumento
Tiempo por época	16s	30s
Tiempo total	7m 57s	15m 2s

Tabla 4. *Tiempo medio de aprendizaje de Mask R-CNN, en base a 116 imágenes de entrenamiento y 11 de validación.*

Tiempo para una imagen	0.6375s
Tiempo para 32 imágenes	20.84s

Tabla 5. *Tiempo medio de predicción de Mask R-CNN.*

Como se podía esperar, es destacable la eficiencia de cómputo en modo inferencia, que permitiría emplear esta red para predicción en tiempo real o para cribados masivos de retinografías.

4. Conclusiones

En este artículo se ha evaluado el algoritmo Mask R-CNN en las tareas de segmentación automática del disco óptico y de la excavación en retinografías de la base de datos RIMONEr3. Además de las conocidas prestaciones de esta red en cuanto a eficiencia computacional, este método ha confirmado unas medidas de calidad de segmentación comparables o superiores a los algoritmos más recientes en el estado actual de la técnica, y que son de uso común en aplicaciones médicas. Los resultados muestran que, en media, Mask R-CNN alcanza mejores resultados que otros algoritmos, incluso con similar precisión a las segmentaciones manuales realizadas entre expertos del campo de la oftalmología, aunque esta cuestión debe contrastarse en bases de datos que incluyan segmentaciones de un mayor número de expertos. Los resultados de segmentación obtenidos mediante esta técnica se podrían aplicar clínicamente para calcular de forma automática el CDR, y a partir de él emitir un prediagnóstico de glaucoma.

Agradecimientos

Este trabajo está financiado parcialmente por los proyectos nacionales AES2017-PI17/00771 y AES2017-PI17-00821 del Instituto de Salud Carlos III, y por el proyecto regional 20901/PI/18 de la Fundación Séneca.

Referencias

[1] A. Almazroa, et al., "Optic disc and optic cup segmentation methodologies for glaucoma image detection: a survey," *Journal of Ophthalmology*, pp. 1–28, 2015.

[2] A. Sommer, et al., "Clinically Detectable Nerve Fiber Atrophy Precedes the Onset of Glaucomatous Field Loss," *Archives of Ophthalmology*, vol. 109, pp. 77–83, 1991.

[3] N. Thakur and M. Juneja, "Optic disc and optic cup segmentation from retinal images using hybrid approach," *Expert Systems with Applications*, vol. 127, 2019.

[4] M. K. Dutta, et al., "Glaucoma detection by segmenting the super pixels from fundus colour retinal images," *Int. Conf. on Medical Imaging, m-Health and Emerging Communication Systems (MedCom)*, 2014.

[5] H. Fu, et al., "Joint optic disc and cup segmentation based on multilabel deep network and polar transformation," *IEEE Trans on Medical Imaging*, vol. 37, no. 7, 2018.

[6] Z. Fan, et al., "Optic disk detection in fundus image based on structured learning," *IEEE journal of biomedical and health informatics*, vol. 22, no. 1, pp. 224–234, 2017.

[7] Piotr Dollár Ross Girshick Kaiming He, Georgia Gkioxari. Mask r-cnn. 2018. https: //arxiv.org/abs/1703.06870.

[8] F. Fumero, et al., "Interactive tool and database for optic disc and cup segmentation of stereo and monocular retinal fundus images," *23rd Int Conf in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG)*, 2015.

[9] D. Hunter Hubel, T. Nils Wiesel https://en.wikipedia.org/wiki/Convolutional_neural_network #History

[10] Ross Girshick Kaiming He Bharath Hariharan Serge Belongie Tsung-Yi Lin, Piotr Dollár. Feature pyramid networks for object detection. 4 2017. https://arxiv.org/pdf/1612.03144.pdf.

[11] Ross Girshick Jian Sun Shaoqing Ren, Kaiming He. Faster R-CNN. 2016. https://arxiv.org/abs/1506.01497.

[12] Waleed Abdulla. Mask R-CNN for object detection and instance segmentation on keras and tensorflow. https://github.com/matterport/Mask_RCNN, 2017.

[13] Shuang Yu, Di Xiao, Shaun Frost, Yogesan Kanagasingam. Robust optic disc and cup segmentation with deep learning for glaucoma detection. Computerized Medical Imaging and Graphics, Volumen 74, 2019, páginas 61-71.

Algoritmo multietapa para la detección de ventilaciones en la impedancia torácica durante la resucitación cardiopulmonar

X. Jaureguibeitia Lara[1], U. Irusta Zarandona[1], E. Aramendi Ecenarro[1], H.E. Wang[2], A.H. Idris[3]

[1] BioRes, Dpto. de Ingeniería de Comunicaciones, Universidad del País Vasco (UPV/EHU), Bilbao, España, {xabier.jaureguibeitia, unai.irusta, elisabete.aramendi}@ehu.eus

[2] University of Texas Health Science Center, Houston (TX, EEUU), Henry.E.Wang@uth.tmc.edu

[3] University of Texas Southwestern Medical Center, Dallas (TX, EEUU), ahamed.idris@utsouthwestern.edu

Resumen

La resucitación cardiopulmonar (RCP) es clave en el tratamiento de la parada cardiorrespiratoria extra-hospitalaria (PCREH). La impedancia torácica (IT) adquirida a través de los parches de un desfibrilador permite detectar las ventilaciones para proveer al rescatador de realimentación sobre el manejo de la vía aérea, pero presenta artefactos debidos a las compresiones torácicas. El objetivo de este trabajo fue el desarrollo de un algoritmo para la detección de ventilaciones en la IT durante compresiones concurrentes. Se analizaron un total de 152 episodios de PCREH, y se anotaron 9665 ventilaciones de referencia en el capnograma. El método constó de tres etapas: procesado de señal para la extracción de la componente de ventilación de la IT, incluyendo un bloque de filtrado adaptativo para eliminar el artefacto de compresiones, detección y caracterización de formas de onda de ventilación, y clasificación mediante una máquina de vectores de soporte para discriminar falsos positivos. Los pacientes, uno por episodio, fueron divididos en grupos de entrenamiento (70%) y evaluación (30%). Se utilizaron 100 particiones diferentes a fin de reducir el sesgo. Las métricas de desempeño finales mostraron valores medianos por paciente de 87.6% de sensibilidad y 85.0% de valor predictivo positivo. El algoritmo podría utilizarse para proporcionar realimentación al rescatador en términos de tasa de ventilación y volúmenes de aire insuflado.

1. Motivación

La parada cardiorrespiratoria extra-hospitalaria (PCREH), con una tasa de incidencia en torno a 55 casos por 100.000 habitantes y año, y una ratio de supervivencia cercana al 10%, representa un problema de salud pública de primer orden [1]. Un evento de parada se origina, generalmente, debido a un mal funcionamiento del sistema eléctrico del corazón, y resulta en la pérdida repentina de la función circulatoria. Una resucitación cardiopulmonar (RCP) de calidad, consistente en compresiones torácicas (CTs) y ventilaciones a fin de inducir un flujo mínimo de sangre oxigenada al paciente, constituye un elemento clave en el tratamiento de la PCREH, retrasando la aparición de daño celular, y ganando así tiempo para la aplicación de otras terapias [2]. Las guías de resucitación aconsejan una tasa de CTs de 100-120 min^{-1} [2]. Si las CTs se administran de forma ininterrumpida, es decir, sin pausas para ventilaciones, éstas últimas deben suministrarse con una tasa en torno a 10 min^{-1}, procurando no hiperventilar al paciente [3]. El cumplimiento de estas recomendaciones mejora cuando el rescatador recibe realimentación en tiempo real sobre la RCP practicada [4]. En este sentido, se han implementado diversas soluciones para mejorar la calidad de la terapia, incluyendo acelerómetros integrados para monitorizar la tasa y la profundidad de compresión, e incluso dispositivos mecánicos capaces de administrar CTs de forma autónoma y muy regular. Las soluciones para la monitorización de las ventilaciones son más limitadas, recomendándose el uso de la capnografía [3], que mide la presión parcial de CO_2 en los gases expirados. La capnografía, sin embargo, requiere de un manejo avanzado de la vía aérea, así como de equipo de monitorización especializado, y no aporta información sobre el volumen de aire insuflado.

La impedancia torácica (IT), medida como la resistencia del cuerpo al paso de corriente, representa una alternativa para la monitorización de ventilaciones, pues es sensible a cambios de volumen de aire en los pulmones. La mayoría de equipos de resucitación básicos registran la señal de IT a través de los parches de desfibrilación, por lo que ésta tiende a estar disponible una media de 13 min antes que la capnografía durante un evento de PCREH [5]. La amplitud de las fluctuaciones de IT debidas a las ventilaciones puede emplearse además como medida subrogada del volumen de aire insuflado, que a su vez puede ser relacionado con la efectividad de la terapia y las opciones de supervivencia del paciente [6]. En el lado negativo, la señal es sensible a diversas fuentes de ruido, tales como el movimiento de los electrodos o del paciente. Las CTs aplicadas como parte de la RCP producen, además, un artefacto que enmascara las ventilaciones y debe eliminarse. Las soluciones presentes en el estado del arte para este escenario hacen uso frecuente de filtros adaptativos originalmente concebidos para la eliminación del artefacto de CTs en el ECG [7].

Este trabajo parte de una solución anterior, diseñada para la detección de ventilaciones durante CTs aplicadas por un dispositivo mecánico [8], y trata de adaptarla al caso más general de compresiones manuales administradas por un rescatador. En este escenario, el artefacto debido a las CTs es más variable, y la prevalencia de ruido de movimiento es también mayor.

2. Materiales

Se analizaron los ficheros electrónicos de 152 episodios de PCREH, tratados por los servicios médicos de emergencia de Dallas - Fort Worth (TX, EEUU) entre febrero de 2016 y marzo de 2017, y adscritos al estudio clínico aleatorizado

ISBN: 978-84-09-25491-0, pags. 352- 355

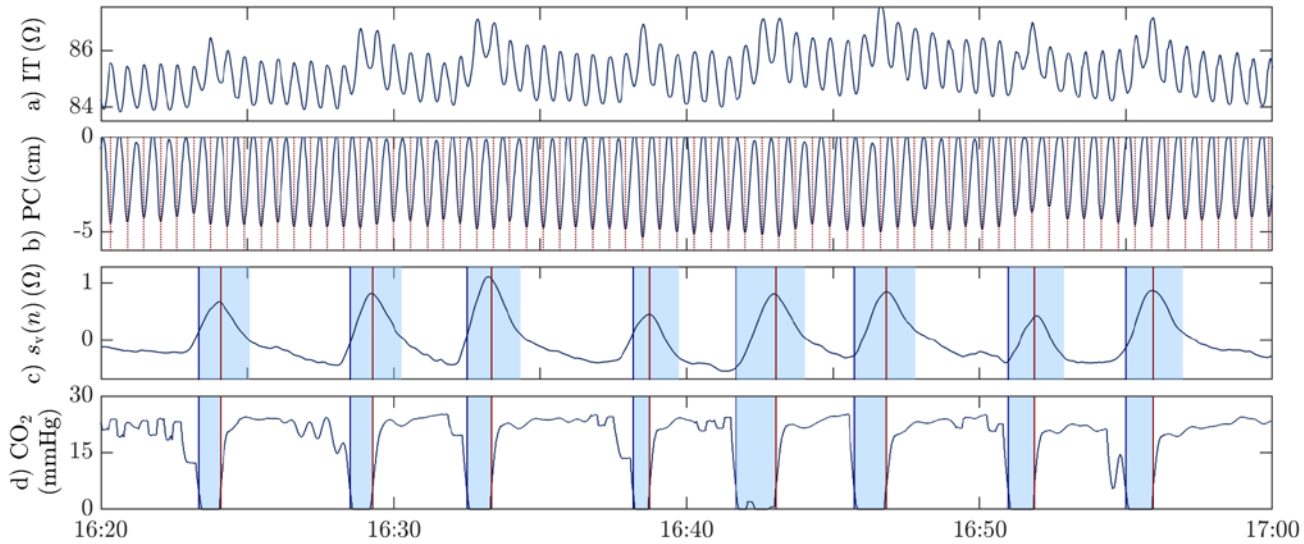


Figura 1. *Ejemplo de las señales implicadas en el proceso de anotación: (a) señal de IT original. (b) señal de PC, incluyendo los instantes de CT anotados. (c) componente de ventilaciones de la IT, obtenida tras filtrado adaptativo en la primera etapa del algoritmo. (d) señal de capnografía sincronizada con la IT, incluyendo las anotaciones de ventilaciones de referencia.*

Pragmatic Airway Resuscitation Trial [9]. Todos los casos fueron tratados con un monitor-desfibrilador Philips MRx (Philips Medical Systems, Andover, MA), que registra la señal de IT con una frecuencia de muestreo de 200Hz y una resolución de 0.74mΩ. Los ficheros fueron exportados y convertidos a un formato adecuado para su análisis usando MATLAB (MathWorks Inc., Natick, MA), y se exploraron en busca de segmentos de IT apropiados para el desarrollo del algoritmo. Se seleccionaron segmentos de al menos 60s durante CTs, con información concurrente de capnografía y profundidad de compresión (PC). Se excluyeron pausas de CTs mayores a 20s, así como intervalos de IT mostrando excursiones abruptas u otros artefactos significativos. Los instantes de CT se anotaron de forma automática en la PC para compresiones superiores a 1.5cm. Las ventilaciones se anotaron de forma manual en el capnograma a modo de referencia. Se corrigió el retardo entre las señales de IT y capnografía debido al transporte de gas en la vía aérea. La Figura 1 muestra un ejemplo de las señales implicadas en el proceso de anotación. La base de datos final constó de 423 segmentos, para un total de 1210 min de grabaciones de IT y 9665 ventilaciones de referencia, el 97.7% de ellas durante CTs concurrentes.

3. Métodos

El algoritmo de detección propuesto constó de tres etapas. En primer lugar, la señal de IT fue filtrada a fin de eliminar el artefacto de compresiones y de extraer la componente de ventilación. A continuación, se detectaron y delimitaron las fluctuaciones individuales asociadas a ventilaciones. Esta etapa fue diseñada tratando de maximizar la detección de ventilaciones reales, por lo que produjo también un elevado número de falsos positivos. Cada una de las detecciones se representó en términos de un vector $\boldsymbol{x_i}$ de 16 características representativas. La etapa final consistió en un clasificador basado en aprendizaje máquina para la discriminación de los falsos positivos.

3.1. Extracción de la componente de ventilación

Como primer paso, la señal de IT se remuestreó a 50Hz, lo que redujo la carga computacional del algoritmo y facilitó el diseño de los filtros: un primer filtro paso-alto con una frecuencia de corte de 0.06Hz para eliminar la componente continua, seguido de un filtro de media móvil (L=50) para suprimir el ruido de alta frecuencia y atenuar el artefacto de compresiones. La supresión final del artefacto se llevó a cabo mediante un esquema de filtrado adaptativo LMS (*Least Mean Squares*), con los instantes de CT previamente anotados empleados a modo de referencia [10].

Sea $s[n]$ la entrada a la etapa de filtrado adaptativo, ésta se consideró formada por dos términos incorrelados, $s_v[n]$ y $s_c[n]$, debidos respectivamente a ventilaciones y CTs, de modo que $s[n] \approx s_v[n] + s_c[n]$. La componente de CTs se modeló a su vez como una serie de Fourier truncada de la forma

$$\hat{s}_c[n] = A[n] \sum_{k=1}^{N} [a_k[n] \cos(k\phi[n]) + b_k[n] \sin(k\phi[n])]$$

donde N representa el número de armónicos considerado. La fase instantánea $\phi[n]$ se derivó de los instantes de CT, considerando una variación lineal de fase de 2π entre cada par de compresiones contiguas. $A[n]$, a su vez, se definió tal que diferenciase intervalos con (A=1) y sin (A=0) CTs, de acuerdo a una duración mínima de pausa de 1.5s. Los coeficientes de amplitud de los términos en fase y cuadratura, $a_k[n]$ y $b_k[n]$, se estimaron haciendo uso de la configuración LMS para la supresión de interferencia armónica [11]. La señal de error $e[n] = s[n] - \hat{s}_c[n]$ a la salida del filtro se aproxima en este caso a la componente de ventilación $s_v[n]$. Tras unos experimentos preliminares, se estableció un tamaño de paso μ=0.05 para el algoritmo LMS, así como un número de armónicos N=3.

3.2. Detección y caracterización de formas de onda

Las formas de onda potencialmente debidas a ventilaciones se identificaron empleando un algoritmo de detección de picos con una separación mínima de 1.5s. Para cada pico $t_{p,i}$ detectado, se aplicaron a continuación un conjunto de reglas heurísticas a fin de localizar los instantes de inicio $t_{s,i} \in [t_{p,i} - 5.5\ ,\ t_{p,i} - 0.45]$ y final $t_{e,i} \in [t_{p,i} + 0.45\ ,\ t_{p,i} + 5.5]$ de la forma de onda. Se descartó cualquier detección para la que $t_{s,i}$ o $t_{e,i}$ no pudieran ser determinados.

Cada fluctuación detectada se representó en términos de un total de 16 características extraídas de su morfología y de las inmediaciones de señal. Las 4 primeras consistieron en las amplitudes, Z_u y Z_d, y duraciones, T_u y T_d, de las fases de insuflación y exhalación, que se corresponden a su vez con las secciones ascendente y descendente de la forma de onda. Para cada fase se realizó también una aproximación de mínimos cuadrados en base a polinomios de Legendre de hasta orden 4, resultando en 10 nuevas características. Finalmente, las dos últimas características vinieron dadas por las excursiones de señal $\Delta_{90} = P_{95} - P_5$ y $\Delta_{95} = P_{97.5} - P_{2.5}$, donde P_x representa el percentil x de la señal $s_v[n]$ en una ventana de hasta un minuto centrada en la posición de pico.

3.3. Clasificación

Cada una de las ventilaciones potenciales identificadas en la etapa anterior se etiquetó como ventilación real ($y_i = 1$) o ausencia de ventilación ($y_i = 0$) en base a las anotaciones de referencia tomadas en el capnograma. Las detecciones cuyo instante de pico $t_{p,i}$ se ubicó dentro de un intervalo de referencia (áreas sombreadas en Figura 1.c) se etiquetaron como $y_i = 1$. En caso de que varias detecciones cumplieran la condición para un mismo intervalo, sólo aquella con $t_{p,i}$ más próximo al comienzo de expiración (líneas verticales en Figura 1.c) fue etiqueta como ventilación real. De este modo, la entrada a la etapa constituyó un conjunto de pares $\{\boldsymbol{x_i}, y_i\} \in \mathbb{R}^{16x(0,1)}$.

Se consideró que algunas de las formas de onda detectadas podrían estar afectadas por transitorios debidos a las pausas de compresiones, artefactos de ruido o similares, y de este modo deteriorar el funcionamiento del clasificador. En este sentido, se optó por eliminar de los datos de entrenamiento un *5* % de las muestras para ambas clases. La detección de valores extremos se llevó a cabo empleando la distancia de Mahalanobis.

Se escogió como algoritmo de clasificación una máquina de vectores de soporte (s*upport vector machine*, SVM) [12] con *kernel* gaussiano, para la que se optimizaron los hiper-parámetros C y γ, que controlan la frontera de decisión del algoritmo.

3.4. Evaluación del modelo

El conjunto final de ventilaciones potenciales se dividió en grupos de entrenamiento, para el aprendizaje de la etapa de clasificación, y evaluación. Esta división se realizó en base a paciente, de modo que todas las muestras asociadas a un mismo caso, con probables similitudes morfológicas entre sí, se asignasen a un mismo grupo. El 70% de los pacientes fue asignado al grupo de entrenamiento y el 30% restante a evaluación. Se exigió una proporcionalidad similar en el número de ventilaciones de referencia asignadas a cada uno de los grupos. Se generaron 100 particiones diferentes de entrenamiento/evaluación para minimizar cualquier sesgo debido a un reparto concreto de los datos.

Se determinaron sobre el grupo de evaluación el número de verdaderos (VP) y falsos positivos (FP), así como de falsos negativos (FN). Se contabilizó también como FN cualquier ventilación perdida en la etapa de detección de formas de onda. Se computaron, para cada paciente, los valores de sensibilidad (Se), valor predictivo positivo (VPP) y F_1:

$$\mathrm{Se} = \frac{\mathrm{VP}}{\mathrm{VP} + \mathrm{FN}} \qquad \mathrm{VPP} = \frac{\mathrm{VP}}{\mathrm{VP} + \mathrm{FP}} \qquad \mathrm{F1} = 2\,\frac{\mathrm{Se}\ \cdot \mathrm{PPV}}{\mathrm{Se} + \mathrm{PPV}}$$

Estas métricas de rendimiento se promediaron entre todas las particiones en las que el paciente objetivo formaba parte del grupo de evaluación. El análisis final se realizó entre pacientes, en términos de mediana (cuartil 1 – cuartil 3).

4. Resultados

La detección preliminar de formas de onda sobreestimó las ventilaciones, con Se de 94.1 (89.0 - 98.4) % y VPP de 56.0 (44.9 - 66.7) %. La etapa de clasificación logró corregir la mayoría de falsos positivos, resultando en unas métricas de desempeño finales de 87.6 (75.2 - 93.0) % para Se, 85.0 (71.5 - 92.5) % para VPP, y 83.7 (73.6 - 92.2) % para F_1. La Figura 2 muestra una distribución más detallada de las métricas de desempeño para los distintos pacientes.

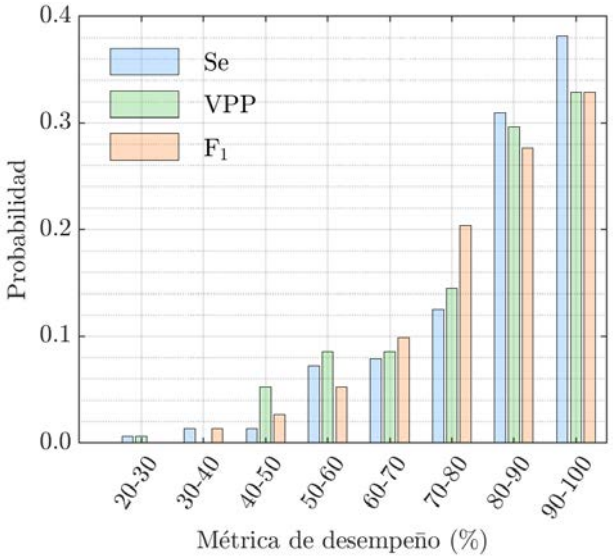


Figura 2. *Distribuciones de métricas de desempeño por paciente, normalizadas como probabilidad.*

5. Discusión

Este trabajo describe y evalúa un método automático para la detección de ventilaciones en IT durante compresiones concurrentes de carácter manual.

El algoritmo propuesto, adaptado de una solución previa para uso durante RCP mecánica [8], mostró un desempeño manifiestamente inferior al del trabajo original, en parte debido a una mayor irregularidad en frecuencia y amplitud de las CTs, pero, fundamentalmente, a la aparición de una

componente de ruido de baja frecuencia, que distorsiona y enmascara las ventilaciones, y que se estima originada por el movimiento del paciente y/o los electrodos durante la RCP. La Figura 3 muestra un ejemplo de dicho fenómeno. Aún así, el método se mostró efectivo en una mayoría de casos, con métricas de desempeño superiores al 85%. Estas prestaciones podrían mejorarse en un futuro mediante el diseño de nuevos de predictores o el uso de técnicas más avanzadas de filtrado y/o clasificación.

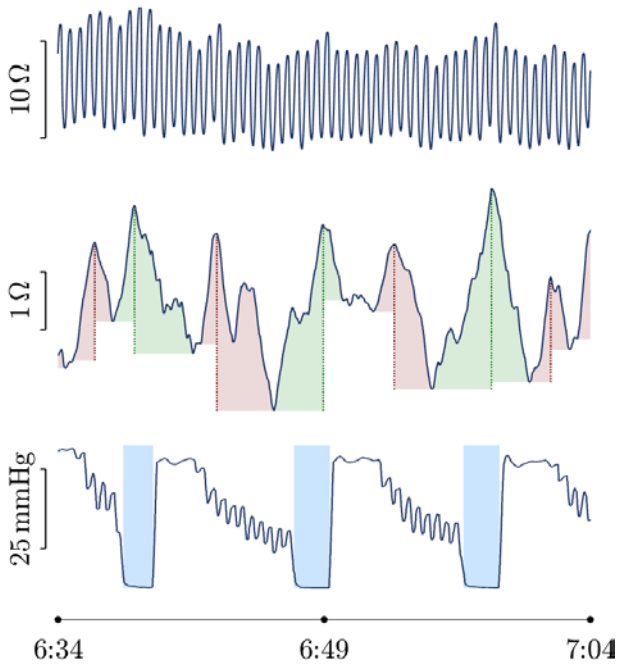


Figura 3. *Ejemplo de segmento de IT afectado por ruido de baja frecuencia.*

El algoritmo expuesto mejoró el rendimiento de soluciones más simples, como la propuesta por Alonso et al [13], para la que se obtuvo una Se mediana de 88.8%, pero un VPP muy inferior, de 63.4%. Esta reducción de rendimiento con respecto de los resultados publicados se atribuye a la fuerte presencia de ruido de movimiento, una característica muy dependiente de la base de datos de estudio. Cabe resaltar también que no se llevó a cabo ningún tipo de optimización de parámetros. Otras soluciones más complejas, como la propuesta por Risdal et al [7], no pudieron ser evaluadas por requerir para su entrenamiento de anotaciones tomadas directamente sobre la IT. La existencia en nuestra base de estudio de episodios fuertemente artefactados, para los que dicha anotación no resulta viable, sugiere, no obstante, un rendimiento inferior al reportado (con métricas superiores al 90%), más en línea con el mostrado por nuestra solución.

6. Conclusiones

Con unas métricas de rendimiento medianas alrededor del 85%, el algoritmo de detección expuesto podría emplearse para proveer de realimentación útil acerca de la terapia de ventilación en una mayoría de casos. Se estima necesario de cara al futuro el desarrollo de predictores que permitan identificar escenarios fuertemente afectados por ruido de movimiento, evitando así una realimentación inadecuada con posibles efectos adversos en la terapia RCP.

Agradecimientos

Este trabajo ha recibido ayuda financiera del Ministerio de Ciencia, Innovación y Universidades, proyecto RTI2018-101475-BI00, junto con el Fondo Europeo de Desarrollo Regional (FEDER), así como del Gobierno Vasco a través de la subvención a grupos de investigación IT-1229-19 y la beca pre-doctoral PRE-2019-1-0209.

Referencias

[1] Berdowski J, Berg RA, Tijssen JGP, Koster RW. Global incidences of out-of-hospital cardiac arrest and survival rates: Systematic review of 67 prospective studies. *Resuscitation*, vol 81, 2010, pp 1479-1487.

[2] Perkins GD, Handley AJ, Koster RW, et al. European Resuscitation Council Guidelines for Resuscitation 2015 Section 2. Adult basic life support and automated external defibrillation. *Resuscitation*, vol 95, 2015, pp 81-99.

[3] Soar J, Nolan JP, Böttiger BW, et al. European Resuscitation Council Guidelines for Resuscitation 2015 Section 3. Adult advanced life support. *Resuscitation*, vol 95, 2015, pp 100-147.

[4] Hostler D, Everton-Stewart S, Rea TD, et al. Effect of real-time feedback during cardiopulmonary resuscitation outside hospital: prospective cluster-randomized trial. BMJ (Clinical research ed.), vol 342, 2011.

[5] Aramendi E, Lu Y, Chang MP, et al. A novel technique to assess the quality of ventilation during pre-hospital cardio-pulmonary resuscitation. *Resuscitation*, vol 132, 2018, pp 41-46.

[6] Chang MP, Lu Y, Leroux B, et al. Association of ventilation with outcomes from out-of-hospital cardiac arrest. *Resuscitation*, vol 141, 2019, pp 174-181.

[7] Risdal M, Aase SO, Stavland M, et al. Impedance-based ventilation detection during cardiopulmonary resuscitation. *IEEE Transactions on Biomedical Engineering*, vol 54, 2007, pp 2237-2245.

[8] Jaureguibeitia X, Irusta U, Aramendi E, et al. Automatic detection of ventilations during mechanical cardiopulmonary resuscitation. *IEEE JBHI*, vol 24, 2020, pp 2580-2588.

[9] Wang HE, Schmicker RH, Daya MR, et al. Effect of a strategy of initial laryngeal tube insertion vs endotracheal intubation on 72-hour survival in adults with out-of-hospital cardiac arrest: a randomized clinical trial. *JAMA*, vol 320, 2018, pp 769-778.

[10] Irusta U, Ruiz J, de Gauna SR, et al. A least mean-square filter for the estimation of the cardiopulmonary resuscitation artifact based on the frequency of compressions. *IEEE Transactions on Biomedical Engineering*, vol 56, 2009, pp 1052-1062.

[11] Xiao Y, Tadokoro Y. Lms-based notch filter for the estimation of sinusoidal signals in noise. *Signal Processing*, vol 46, 1995, pp 223-231.

[12] Cortes C, Vapnik V. Support-vector networks. *Machine Learning*, vol 20, 1995, pp 273-297.

[13] Alonso E, Ruiz J, Aramendi E, et al. Reliability and accuracy of the thoracic impedance for measuring cardiopulmonary resuscitation quality metrics. *Resuscitation*, vol 88, 2015, pp 28-34.

Caracterización de la morfología foveal: parametrización, diferencias de sexo y efectos de la edad

D. Romero-Bascones[1], I. Gabilondo Cuellar[1,2,3], M. Barrenechea Carrasco[1], U. Ayala Fernández[1]

[1] Mondragon Unibertsitatea, Mondragón, España, {dromero, mbarrenetxea, uayala}@mondragon.edu

[2] Instituto de Investigación Sanitaria Biocruces Bizkaia, Barakaldo, España

[3] IkerBasque, Basque Foundation for Science, Bilbao, España

Resumen

Frente a los análisis convencionales de grosores de la retina, el análisis morfológico de la fóvea presenta una metodología alternativa de gran potencial pero poco explorada. En este estudio se implementa un pipeline completo de análisis morfológico basado en imágenes OCT (Optical Coherence Tomography), comparando los modelos matemáticos existentes para estudiar la fóvea y demostrando su capacidad de ajuste y mejora de la fiabilidad test-retest. Asimismo, se analiza la influencia de la edad y el sexo en parámetros morfológicos como la altura, anchura, pendiente o área de la fóvea en una cohorte de 272 sujetos sanos. Los resultados muestran un claro dimorfismo sexual en la fóvea junto con relevantes cambios experimentados durante el envejecimiento.

1. Introducción

La Tomografía por Coherencia Óptica (OCT, Optical Coherence Tomography) es una técnica de imagen in-vivo no invasiva capaz de obtener imágenes de alta resolución de la retina [1]. Además de facilitar el diagnóstico visual, las imágenes de OCT permiten analizar de forma cuantitativa y sencilla las capas que conforman la retina. Una de las aplicaciones más prometedoras de la OCT es su uso en neurología, donde se evalúa su potencial como biomarcador en enfermedades neurodegenerativas como la Enfermedad de Parkinson. Los resultados de numerosos estudios basados en OCT apuntan a una degeneración específica de las capas interiores de la retina derivada de dichas patologías [2]. No obstante, pese a la gran cantidad de estudios, los resultados obtenidos hasta ahora presentan importantes discrepancias [3] que evidencian la necesidad de plantear metodologías y estudios más robustos.

En lo referente a la metodología, la gran mayoría de estudios en la literatura se han limitado a analizar el grosor promedio en un número limitado de sectores, lo cual no permite caracterizar de forma precisa la fóvea [4], situada en el centro de la visión y de gran interés en enfermedades con un impedimento visual asociado. Como posible alternativa se ha propuesto el uso de modelos matemáticos para describir y estudiar la fóvea en base a parámetros morfológicos como la anchura, la pendiente máxima o el grado de concavidad.

Pese a que estos parámetros ya están comenzando a ser explorados para aplicaciones de diagnóstico [5], para emplear dichos parámetros como biomarcadores es crucial comprender previamente su distribución en sujetos sanos. Si bien los efectos de la edad y el sexo en el grosor de la retina han sido ampliamente documentados, sus efectos sobre la morfología foveal no han sido estudiados en profundidad. El presente estudio aborda esta problemática. Para ello, se busca en primer lugar determinar el modelo más adecuado mediante una comparación cuantitativa de cinco modelos de la bibliografía. A partir de estos resultados, se emplea el modelo escogido para estudiar los efectos de la edad y el sexo sobre la morfología de la fóvea en una cohorte de sujetos sanos.

2. Materiales y métodos

2.1. Base de datos

Un total de 272 sujetos sanos con edades entre 22 y 88 años dieron su consentimiento expreso para participar en el presente estudio. Todas las imágenes fueron revisadas por un optometrista excluyendo aquellas que no cumplían con los requerimientos mínimos de calidad OSCAR-IB [6] o que mostraban lesiones oftalmológicas. La Tabla 1 resume las características de los datos finalmente empleados.

Sexo	Sujetos	Ojos	Edad*°
Mujeres	169	327	54.4 ± 12.2
Hombres	103	197	56.1 ± 13.6
Total	272	524	55.0 ± 12.8

* Media ± desviación estándar

° No hay diferencias estadísticamente significativas en edad de hombres y mujeres (p=0.29, prueba t de Welch)

Tabla 1 *Características demográficas de la base de datos*

Mediante un escáner OCT de dominio espectral Spectralis (Heidelberg Engineering, Heidelberg, Alemania) se adquirieron imágenes OCT maculares de ambos ojos tanto de tipo ráster como estrella. Los volúmenes tipo ráster están formados por 25 cortes transversales (B-scans) de 512x469 píxeles con una resolución espacial de 243µm x11.4µm x 3.9µm. Por otro lado, los volúmenes de tipo estrella constan de 12 B-scans radiales de 251x469 píxeles con una resolución de 15º x 7µm x 3.9µm. Cada B-scan fue obtenido promediando 49 B-scans.

2.2. Procesado de imágenes y análisis morfológico

La Figura 1 muestra las etapas del pipeline empleado para el procesado de las imágenes. En primer lugar, el software del equipo segmenta las capas que delimitan la retina: la Membrana Limitante Interna (MLI) y la Membrana de Bruch (MB).

ISBN: 978-84-09-25491-0, pags. 356- 359

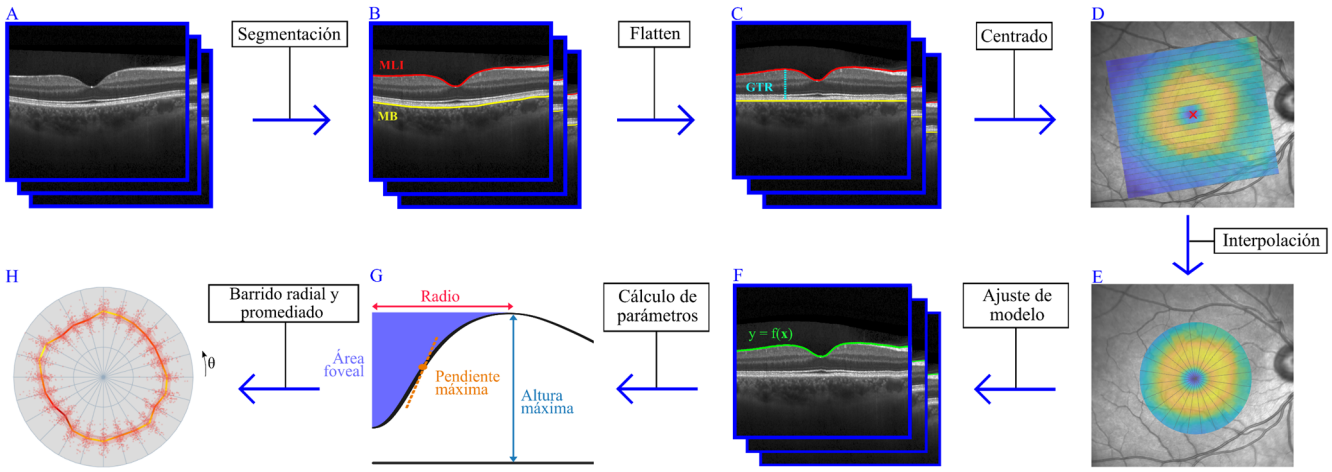


Figura 1 *Etapas de procesado de imagen y cálculo de parámetros morfológico*

A continuación, se obtiene el grosor total de la retina (GTR) como la diferencia entre las capas MLI y MB. Este paso, conocido como "flatten", es equivalente a recolocar los A-scans tomando la MB como referencia, reduciendo así posibles distorsiones debidas a la curvatura del ojo. Combinando los valores de GTR de los diferentes B-Scans se obtiene un mapa de grosor 2D como el de la Figura 1, D.

Para corregir posibles errores de centrado es importante garantizar que el origen de coordenadas se sitúa en el centro de la fóvea. Para ello, se determina primero el centro real de la fóvea como el mínimo del mapa de espesor que ha sido previamente suavizado mediante un filtro gaussiano tal y como está implementado en el software AURA tools [7]. La posición estimada es empleada después para centrar el mapa de espesor.

Finalmente, se transforman los puntos de muestreo tipo ráster en un patrón de muestreo radial (estrella) de 2.5 mm de radio con un total de 24 direcciones angulares mediante interpolación lineal para poder caracterizar la fóvea radialmente.

Para el análisis morfológico, en la literatura existen diversos modelos matemáticos que permiten modelar la fóvea y derivar parámetros morfológicos. La Tabla 2 muestra un resumen de los modelos implementados y comparados en este estudio. Las diferencias entre ellos radican en las ecuaciones matemáticas subyacentes y en la parte de los datos que modelan. Así, modelos como el de Dubis et al. [8] modelan cada B-scan por separado mientras que otros como el propuesto por Ding et al. [9] ajustan la superficie 2D completa.

Los cinco modelos se emplearon para modelar los valores de GTR de forma radial. El ajuste se llevó a cabo mediante el método de mínimos cuadrados no lineal tal y como está implementado en Matlab R2019b. Los valores iniciales, así como los rangos de cada parámetro, fueron ajustados de forma manual para cada modelo.

Modelo	Número de parámetros	Datos modelados
Dubis et al., 2010 [8]	6/B-scan	B-Scan
Ding et al., 2014 [9]	8	Mapa de grosor 2D
Scheibe et al., 2014 [10]	8/B-scan	½ B-scan
Yadav, 2017 [11]	8/B-scan	¼ B-scan
Breher, 2019 [12]	9/B-scan	B-Scan

Tabla 2 *Características de los modelos comparados*

A partir de la curva obtenida se derivaron la altura máxima, el radio, la pendiente máxima y el área foveal (ver Figura 1, G). El valor final de cada parámetro se obtuvo promediando todas las direcciones angulares por cada sujeto. Este proceso fue realizado tanto para las imágenes tipo ráster como para las de estrella.

2.3. Análisis de datos

El error de ajuste de cada modelo se evaluó mediante el Error Absoluto Medio (EAM) entre la curva original (GTR) y la ajustada (GTR_fit) para los N puntos dispuestos en un patrón radial con coordenadas polares r y θ.

$$EAM = \frac{1}{N}\sum_{r}\sum_{\theta}|GTR[r,\theta] - GTR_fit[r,\theta]|$$

Por otra parte, para evaluar la fiabilidad test-retest de cada modelo se empleó el coeficiente de reproducibilidad (CR) definido como el percentil 95 de las diferencias absolutas entre la estimación ráster y estrella de cada parámetro p:

$$CR_p = P_{95}|p_{ráster} - p_{estrella}|$$

$$p \in \{altura\ max, radio, pendiente\ max, área\}$$

En base al CR de los datos en bruto se definió a su vez la mejora de CR (MCR) como la mejora porcentual de cada modelo promediando los 4 parámetros estudiados:

$$MCR\ (\%) = \frac{1}{4}\sum_{p} 100\left(1 - \frac{CR_p^{bruto}}{CR_p^{modelo}}\right)$$

A efectos de comparación se consideró como modelo óptimo aquel que maximice el ratio MCR/EAM, es decir, que mejore la repetibilidad de los datos en bruto introduciendo el menor error de ajuste posible.

Tras comparar los modelos en base a estas métricas se estudió la relación entre el sexo, la edad y los parámetros morfológicos calculados. En el análisis se emplearon las imágenes tipo ráster de todos los ojos. Dado que los ojos de cada sujeto no son estadísticamente independientes, se optó por un modelo mixto basado en regresión lineal múltiple:

$$Parámetro = \beta_o + \beta_1 Sexo + \beta_2 Edad + \gamma_{sujeto},$$

con β_o, β_1 y β_2 como efectos fijos y γ_{sujeto} como intercepto aleatorio tomando el mismo valor en ambos ojos de cada sujeto. Para cada uno de los parámetros morfológicos escogidos se calcularon los intervalos de confianza del 95% de los coeficientes de los efectos fijos mediante bootstrapping. El análisis de datos fue llevado a cabo mediante los paquetes *lme4* [13] y *boot* [14] de R.

3. Resultados

3.1. Comparación de modelos

La Figura 2 compara los modelos en base a los valores de EAM y MCR obtenidos. Las líneas de contorno indican el ratio MCR/EAM.

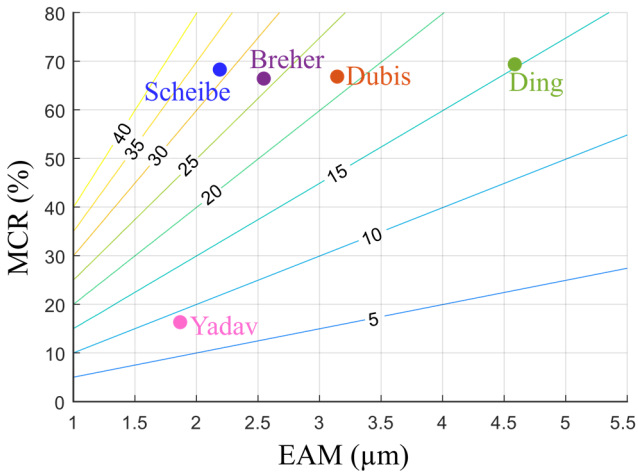


Figura 2 *Comparación de modelos en base al error de ajuste (EAM) y la mejora de la reproducibilidad test-retest (MCR) visualizando su ratio (MCR/EAM) mediante líneas de contorno.*

Como se puede observar, todos los modelos logran ajustar los datos con un EAM inferior a 5 µm. Los modelos propuestos por Dubis et al. [8] y Ding et al. [9] son los que peor ajuste presentan. Por otra parte, el modelo propuesto por Yadav et al. [11], pese a ser el que mejor ajusta los datos, presenta una mejoría de tan sólo un 17% en cuanto a valores de CR. Los modelos con mejores prestaciones son los propuestos por Scheibe et al. [10] y Breher et al. [12], con errores inferiores a 3µm y valores de MCR superiores al 60%. Finalmente, para los siguientes análisis se optó por el modelo de Scheibe et al. [10] por ser éste el modelo con el máximo ratio MCR/EAM.

3.2. Estudio poblacional

Los resultados del estudio de población se resumen en la Tabla 3. En ella, se muestra la estimación de cada uno de los coeficientes fijos junto con su intervalo de confianza del 95 por ciento para cada uno de los parámetros morfológicos estudiados.

Los resultados del efecto del sexo muestran que los hombres tienen en promedio una fóvea 9.2 µm más alta y con una pendiente máxima 0.87º mayor. Las mujeres, sin embargo, presentan una fóvea 60 µm más ancha con un área foveal 2501 µm^2 mayor. Estas diferencias se observan de forma más clara en la Figura 3, donde se superponen la fóvea promedio de hombres y mujeres en un B-scan horizontal.

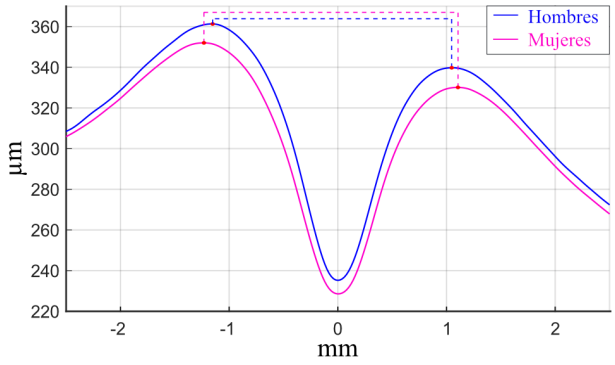


Figura 3 *Grosor total de la retina (GTR) promedio en hombres y mujeres en el B-scan horizontal central.*

Respecto al efecto de la edad, se observa una disminución de 0.22 µm/año en la altura y de 163 µm^2 en el área foveal. Los resultados también apuntan a una disminución del radio y de la pendiente máxima de la fóvea con la edad, aunque la incertidumbre de las estimaciones es mayor. Esto es posiblemente indicador de que el tamaño del efecto es pequeño considerando la alta variabilidad interindividual.

Parámetro	Intercepto*	Sexo (Hombre)*	Edad (x/año)*
Altura máxima (µm)	359.7 [352.5, 367.5]	9.21 [5.58, 12.69]	-0.22 [-0.36, -0.09]
Radio (µm)	1145 [1096, 1192]	-62.6 [-85.1,-39.1]	-0.70 [-1.52, 0.18]
Pendiente máxima (º)	10.43 [9.49, 11.47]	0.87 [0.39, 1.36]	-0.009 [-0.027, 0.007]
Área foveal (µm^2)	59556 [52383, 66501]	-2501 [-5666, 464]	-163 [-281, -35]

* Coeficiente estimado (β) e [intervalo de confianza del 95%]

Tabla 3 *Resultados del análisis poblacional*

4. Discusión

Tal y como muestra la Figura 2, el uso de modelos matemáticos permite mejorar la fiabilidad test-retest en más de un 60% a costa de introducir cierto error de ajuste. Esta mejora está probablemente relacionada con una reducción del ruido intrínseco de las medidas de OCT. Este ruido, debido a errores de cuantización, segmentación e interpolación, hace que los valores de grosor varíen de forma abrupta y que medidas como el punto de máximo grosor fluctúen sustancialmente entre adquisiciones. Ante esto, el ajuste de modelos, que considera las medidas de GTR en su conjunto, permite obtener una estimación suavizada de la fóvea mejorando la repetibilidad de los parámetros morfológicos derivados. Interesantemente, la relación entre esta mejora y el error de ajuste no es lineal. De hecho, la mayoría de los modelos no consiguen mejorar los valores MCR del modelo de Scheibe et al. [10], pese a introducir un error mayor. Por ello, éste fue el modelo empleado en el estudio poblacional. Su superioridad puede atribuirse a su diseño radial y a que, a diferencia del modelo propuesto por Yadav et al. [11], no utiliza el punto de grosor máximo como un punto prefijado.

Las diferencias morfológicas observadas entre hombres y mujeres son similares a las reportadas por un estudio previo [15]. Los efectos de la edad, menos estudiados en la literatura, son menos notables que los efectos del sexo, siendo la altura de la fóvea y su área los parámetros más afectados. La pérdida de grosor relacionada con el envejecimiento normal es probablemente la causa subyacente. Por el contrario, el efecto de la edad sobre la anchura o la pendiente de la fóvea no es tan directo. Aun así, los efectos observados podrían ser explicados en base a una disminución de grosor más acentuada en la región cercana al punto de altura máxima que redujera su anchura.

Dentro de las limitaciones del estudio se encuentran los errores de interpolación, segmentación o ajuste de modelos. Por otra parte, la longitud axial del globo ocular es un factor con influencia sobre las medidas de grosor que no ha sido controlado explícitamente. Aun así, el escáner Spectralis estima internamente la longitud axial de cada sujeto y ajusta la resolución transversal y longitudinal en función de dicha longitud. Finalmente, para la evaluación de la fiabilidad test-retest se han empleado imágenes con patrones de adquisición distintos (ráster y estrella), lo cual no es ideal considerando la posible existencia de diferencias sistemáticas entre ambas.

5. Conclusiones

De los cinco modelos comparados el propuesto por Scheibe et al. [10] es el que presenta un mejor rendimiento. Los resultados del estudio poblacional evidencian que los hombres tienen una fóvea más estrecha, alta y de mayor pendiente que las mujeres. Además, parámetros como la altura y el área se ven especialmente afectados por la edad. Todo ello remarca la necesidad de considerar el sexo y la edad como covariables a la hora de evaluar el potencial diagnóstico de la morfología foveal.

Parte del software empleado puede consultarse en el repositorio https://github.com/DRomBas/CASEIB_20.

Agradecimientos

Los autores agradecen a los sujetos su participación en el estudio. Este trabajo ha recibido ayuda financiera del Departamento de Salud del Gobierno Vasco, proyecto Exp. 2020333033.

Referencias

[1] Huang D et al., Optical Coherence Tomography, *Science*, vol. 254, sup 5035, pp 1178-1181, 1991

[2] Chrysou A, Jansonius N M, van Laar T, Retinal layers in Parkinson's disease: A meta-analysis of spectral-domain optical coherence tomography studies, *Parkinsonism and Related Disorders*, vol. 64, pp 40–49, 2019

[3] Mailankody P, Lenka A, Pal P, The role of Optical Coherence Tomography in Parkinsonism: A critical review, *Journal of the Neurological Sciences*, vol 403, pp 67-74, 2019

[4] Wagner-Schumann M et al., Race-and sex-related differences in retinal thickness and foveal pit morphology, *Investigative Ophthalmology and Visual Science*, vol 52, sup 1, pp 625-634, 2011

[5] Motamedi S et al., Altered fovea in AQP4-IgG–seropositive neuromyelitis optica spectrum disorders, *Neurology-Neuroimmunology & Neuroinflammation,* vol 7, sup 5, pp 1-11, 2020

[6] Tewarie P et al., The OSCAR-IB Consensus Criteria for Retinal OCT Quality Assessment, *PLoS One*, vol7, sup 4, pp 1-7, 2012

[7] Lang A, et al., Retinal layer segmentation of macular OCT images using boundary classification. *Biomedical Optics Express*, vol 4, sup 7, pp 1133-1152, 2013.

[8] Dubis A, McAllister J, Carroll J, Reconstructing foveal pit morphology from optical coherence tomography imaging, *British Journal of Ophthalmology*, vol 93, sup 9, pp 1223-1227, 2009

[9] Ding Y et al., Application of an OCT data-based mathematical model of the foveal pit in Parkinson disease. *Journal of Neural Transmission*, vol 121, sup 11, pp 1367-1376, 2014

[10] Scheibe P et al., Parametric model for the 3D reconstruction of individual fovea shape from OCT data, *Experimental Eye Research*, vol. 119, pp. 19–26, 2014

[11] Yadav S et al., CuBe: parametric modeling of 3D foveal shape using cubic Bézier, *Biomedical Optics Express,* vol 8, sup 9, pp 4181-4199, 2017

[12] Breher K, Agarwala R, Leube A, Wahl S, Direct modeling of foveal pit morphology from distortion-corrected OCT images, *Biomedical Optics Express*, vol 10, sup 9, pp 4815-4824, 2019

[13] Bates D, Mächler M, Bolker B, Walker S, Fitting Linear Mixed-Effects Models Using lme4, *Journal of Statistical Software,* vol 67, sup 1, pp 1-48, 2015

[14] Canty A, Ripley B, boot: Bootstrap R (S-Plus) Functions, 2020

[15] Scheibe P, Zocher M, Francke M, Rauscher F, Analysis of foveal characteristics and their asymmetries in the normal population, *Experimental Eye Research*, vol 148, pp 1-11, 2016

Determinación del efecto de la ablación en tejido porcino mediante un láser Nd:YAG en aplicaciones quirúrgicas

L. Santamaría-Bustamante, S. Pampín-Suárez, J. L. Arce-Diego, F. Fanjul-Vélez

Grupo de Ingeniería Biomédica, Departamento TEISA, Universidad de Cantabria, Av. de los Castros s/n, 39005 Santander (Cantabria), España

arcedj@unican.es, fanjulf@unican.es

Resumen

La interacción de la radiación óptica con los medios biológicos puede producir diversos efectos sobre los mismos, en función de los parámetros correspondientes a la fuente de radiación, y aquéllos propios del tejido biológico de interés. Uno de ellos son los efectos ablativos inducidos por la radiación láser, de creciente empleo en cirugía mínimamente invasiva. Estos efectos son fuertemente dependientes de parámetros como la longitud de onda de emisión del láser, el tiempo de exposición del tejido a la radiación láser, su energía y el tamaño del spot de iluminación Uno de los láseres más versátiles y que pueden presentar mejores resultados quirúrgicos es el láser de Nd: YAG. En este trabajo se presentan los resultados experimentales de la aplicación de dicho láser de estado sólido, con diferentes configuraciones paramétricas, a diferentes tipos de tejidos biológicos básicos como el epitelial, muscular y adiposo. Los efectos termoablativos resultantes de dicha interacción son determinados y cuantificados en función de los parámetros del láser. Esta cuantificación se puede realizar mediante algorítmica específica de tipo machine learning.

1. Introducción

El uso del láser en la cirugía del aparato digestivo ha experimentado un crecimiento importante. Gracias a las grandes ventajas que estos procedimientos ofrecen frente a la cirugía convencional, podemos tratar y diagnosticar las enfermedades del aparato digestivo de una manera no invasiva.

El uso del láser para la ablación biológica de tejidos es una técnica muy empleada hoy en día, ya que ofrece múltiples ventajas como la precisión, la minimización de cicatrices y una recuperación más corta [1]. Uno de los láseres más utilizados para realizar este tipo de procedimientos es el Nd: YAG, aunque existen también otros con aplicaciones específicas. El empleo de fuentes ópticas en aplicaciones de cirugía proporciona ventajas frente a otros tipos de radiaciones, e incluso frente a procedimientos de corte de tipo mecánico, tales como la ausencia de contacto, el carácter no ionizante o la posibilidad de ser integradas en sistemas no invasivos o mínimamente invasivos.

Uno de los aspectos críticos en el diseño y aplicación de las técnicas de cirugía basadas en fuentes ópticas, consiste en el adecuado control del volumen ablacionado. Este aspecto es particularmente interesante en la aplicación a intervenciones localizadas en zonas con posible daño a tejido sano adyacente de especial sensibilidad, tales como intervenciones en la zona del cuello o colonoscópicas, en este último caso con el riesgo de perforación. La adecuada planificación quirúrgica requiere, por lo tanto, de un conocimiento apropiado del efecto de la interacción, de tal forma que pueda estimarse el volumen ablacionado a partir de los parámetros del sistema de generación de radiación óptica, así como de los propios del tejido biológico afectado.

En este trabajo se establecen, en primer lugar, los mecanismos de interacción entre la radiación óptica y los tejidos biológicos, con particular atención a los de tipo ablativo. Posteriormente, se realizan experimentalmente ablaciones sobre diversas muestras de tejidos biológicos, con diferentes parámetros de la fuente óptica Nd: YAG, con el fin de analizar el volumen ablacionado, así como los posibles efectos de desnaturalización o carbonización, en su caso.

2. Mecanismos de interacción luz-tejido

Los mecanismos que pueden ocurrir al aplicar luz láser en un tejido biológico son múltiples debido a sus características específicas y a los parámetros del láser [2]. En estos mecanismos de interacción, son importantes las propiedades ópticas del tejido, concretamente los coeficientes de reflexión, absorción y esparcimiento o scattering, que determinan la transmisión total del tejido a una determinada longitud de onda. Además, debemos tener en cuenta las propiedades térmicas del tejido, como la conducción del calor y la capacidad calorífica, así como la radiación del láser, que nos proporciona parámetros como la longitud de onda, tiempo de exposición, energía aplicada, tamaño del punto focal y la densidad de energía y potencia.

En general, se pueden establecer cinco categorías de tipos de interacción con efecto sobre el tejido biológico:

- Interacciones fotoquímicas
- Interacciones térmicas
- Fotoablación
- Ablación inducida por plasma
- Fotorruptura

Todos los tipos de interacción comparten un dato común: la densidad de energía puede variar entre 1J/ cm^2 a 100 J/cm^2. Además, estos procesos se distinguen y pueden controlarse mediante un único parámetro: la duración de la exposición láser.

ISBN: 978-84-09-25491-0, pags. 360- 363

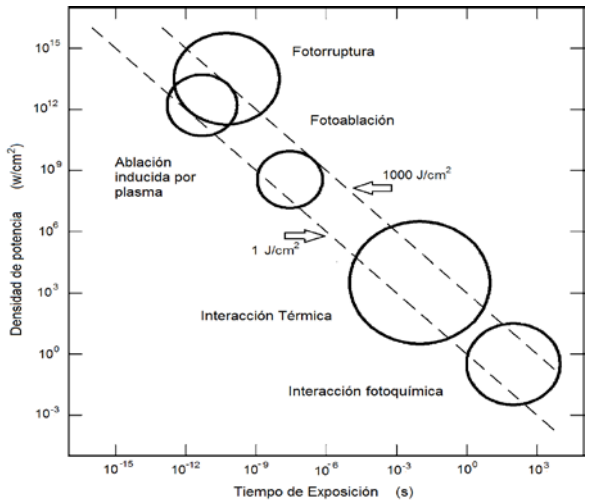


Figura 1. *Mapa de interacciones láser-tejido biológico; adaptado de [2].*

En el diagrama de la Figura 1 están representados los cinco tipos básicos de interacción. En el eje de ordenadas está expresada la densidad de potencia aplicada (W/cm^2) y en el eje de abscisas tenemos el tiempo de exposición (s). Las dos diagonales muestran los flujos de energía constantes a 1 J/cm^2 y a 100 J/cm^2. La escala de tiempo se puede dividir en cinco secciones: ondas continuas o tiempos de exposición > 1s para interacciones fotoquímicas, 1 ms hasta 1 µs para interacciones térmicas, 1 µs hasta 1 ns para fotoablación y < 1 ns para ablación inducida por plasma y fotorruptura [3], [4].

2.1. Interacción fotoquímica

La luz puede inducir efectos y reacciones químicas dentro de macromoléculas o tejidos dando lugar a las interacciones fotoquímicas, que se producen a bajas densidades de potencia y en tiempos de exposición que van desde segundos hasta ondas continuas.

2.2. Interacción térmica

La interacción térmica representa un grupo de interacción donde el aumento de temperatura local es el cambio significativo de los parámetros. Los efectos térmicos pueden ser inducidos por radiación láser y, dependiendo de la duración y del valor máximo de temperatura del tejido alcanzado, se pueden distinguir diferentes efectos:

- Coagulación
- Vaporización
- Carbonización
- Fusión

Durante la coagulación, las temperaturas alcanzan al menos 60ºC y el tejido se vuelve necrótico. En el proceso de vaporización se eliminan capas que contienen un número elevado de moléculas de agua, que absorben fuertemente la longitud de onda del láser logrando así la vaporización. Cuando ya aplicamos demasiada energía y logramos aumentar la temperatura del tejido más de 100ºC conseguimos la carbonización. Y, por último, si aumentamos aún más la temperatura se produce la fusión debida al estrés térmico inducido por ese gradiente de temperatura local.

2.3. Fotoablación

La fotodescomposición ablativa significa que el material se descompone cuando se expone a una irradiación láser de alta intensidad. Los valores de umbral típicos de este tipo de interacción son $10^7 - 10^8$ W/cm^2 con duraciones de pulso láser en el rango de nanosegundos. La profundidad de la ablación, es decir, la profundidad de eliminación del tejido por pulsos está determinada por la energía del pulso hasta un cierto límite de saturación.

2.4. Ablación inducida por plasma

Mediante la ablación inducida por plasma podemos lograr una extracción muy limpia y bien definida eligiendo los parámetros láser apropiados.

El parámetro más importante es la intensidad del campo eléctrico local, E, que cuando supera un cierto umbral se logra la ruptura óptica. Para pulsos de picosegundos, las intensidades de umbral típicas son 10^{11} W/cm^2 , por lo tanto, el campo eléctrico correspondiente es de 10^7 V/cm, que proporciona la condición necesaria para la ionización del plasma, denominándose este proceso, ruptura dieléctrica. Este término enfatiza que los rayos UV, visible y la luz infrarroja son absorbidos fuertemente por el plasma. Además, la característica más importante de la ruptura óptica es que hace posible una descomposición de energía no solo en el tejido pigmentado, sino también en medios de absorción débil, debido al aumento del coeficiente de absorción del plasma inducido. Por lo que el campo de aplicaciones médicas del láser se amplía [5], [6].

2.5. Fotorruptura

Cuando hablamos de la ablación inducida por plasma descuidamos cualquier efecto secundario del plasma. A energías de pulsos más altas y por lo tanto a energías de plasma más altas, las ondas de choque y otros efectos secundarios mecánicos se vuelven más significativos y pueden incluso determinar el efecto global sobre el tejido.

Para duraciones de pulso de nanosegundos, la extensión espacial de los efectos mecánicos es del orden de milímetros, asociando la ruptura óptica con la formación de ondas de choque, siendo a menudo un síntoma indeseado ya que el tejido adyacente puede ser dañado. Esto se puede evitar con duraciones de pulso de picosegundos o femtosegundos donde se puede lograr todavía la ruptura óptica reduciéndose significativamente la energía del plasma.

Como podemos ver no es fácil distinguir este mecanismo de la ablación inducida por plasma. Pero en esta última, se limita a un rango bastante estrecho de duraciones de pulso de hasta 500 ps ya que con duraciones de pulso más largas ya inducimos los efectos secundarios mecánicos.

Entonces, podemos considerar la fotorruptura como un efecto mecánico de múltiples causas-efectos a partir de la ruptura óptica. Siendo los mecanismos primarios la formación de ondas de choque, la cavitación y la formación de chorros, que tienen lugar cada uno de ellos en una escala de tiempo diferente.

3. Procedimiento experimental

El láser que vamos a utilizar para estudiar los efectos causados en diferentes tipos de tejidos ante la radiación óptica es el Nd: YAG de clase cuatro. Este tipo de láser tiene dos modos de funcionamiento, el FR y el Q-SWITCH, en nuestro trabajo experimental hemos utilizado este último debido a que tiene pulsos de haz y energías mayores que el FR. Trabajaremos a la longitud de onda de 532 nm (luz verde, fundamental del Nd: YAG doblado).

Los valores máximos a los que puede trabajar el láser son 20 J de energía de bombeo y 15 Hz de frecuencia de repetición de pulsos. Por lo que, teniendo en cuenta estos datos máximos, los valores de los parámetros que variaremos para tomar cada medida son de 15J, 17J y 20J para la energía de bombeo, que corresponden a 50 mJ, 85 mJ y 120 mJ de energía de los pulsos respectivamente. El tiempo de exposición será de 120 seg, 180 seg y 720 seg, y, por último, la frecuencia de repetición de pulsos la variaremos entre 10 Hz y 15 Hz.

Para obtener los diferentes tipos de tejidos utilizamos muestras de carne animal *in vitro*, concretamente carne porcina. Cortamos varias muestras de las diferentes partes. En primer lugar, donde se encuentra el tejido muscular, cortamos de manera transversal y longitudinal cuatro muestras diferentes dos a dos. Después, en la parte donde se encuentra la piel, sacaremos dos muestras y, por último, donde está el tejido con más grasa extraeremos dos muestras de tejido con grasa y otras dos únicamente de grasa.

Después de obtener todas las muestras de tejido, las colocamos en placas que limitan la zona de acción donde vamos a aplicar la radiación, por lo que podemos tomar un total de nueve medidas por muestra (Figura 2).

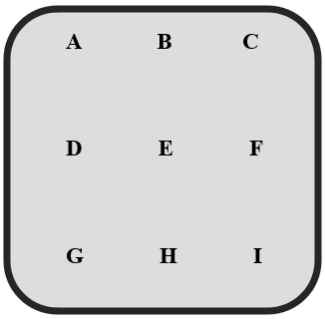


Figura 2.*Diagrama de ubicación de los diferentes puntos de acción*

Una vez colocadas todas las muestras en las diferentes placas procedemos a aplicar la radiación láser en cada tejido variando los diferentes parámetros. El montaje experimental básico se muestra en la Figura 3. Por último, cuando ya hemos obtenido todos los resultados tomamos imágenes correspondientes usando un microscopio, con lentes de 10 mm y 4 mm, en el que está incluido una cámara para poder visualizarlas en el ordenador.

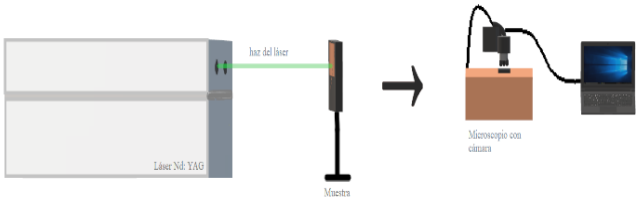


Figura 3. *Esquema del procedimiento para la obtención de los resultados.*

4. Resultados y discusión

Después de aplicar la radiación láser en los nueve puntos de acción de cada tejido, podemos observar varios efectos dependiendo del tiempo de exposición, energía de los pulsos y la frecuencia de repetición de pulsos. Algunos de ellos son la desnaturalización de las células, cambio de color y textura del tejido, la carbonización y la ablación. Además, dependiendo del tipo de tejido utilizado y dentro del mismo tejido al ser medios no homogéneos observaremos efectos muy diferentes.

Un ejemplo de estos efectos lo podemos ver en el tejido muscular cortado de manera transversal, Figura 4 y Tabla 1.

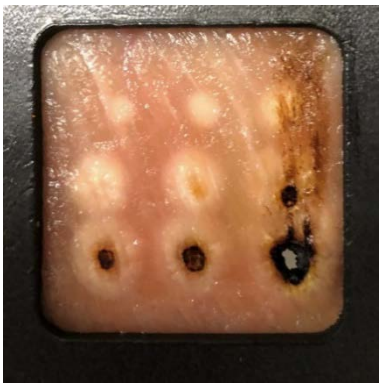


Figura 4.*Muestra de tejido muscular cortado de manera transversal.*

Tabla 1. *Resultados de la muestra tejido muscular cortado de manera transversal, frecuencia de repetición de 15 Hz.*

Punto	Energía pulsos (mJ)	Tiempo exposición (seg)	Efecto
Punto A	50 mJ	120 seg	Desnaturalización
Punto B	50 mJ	180 seg	Desnaturalización
Punto C	50 mJ	720 seg	Desnaturalización
Punto D	85 mJ	120 seg	Desnaturalización
Punto E	85 mJ	180 seg	Desnaturalización
Punto F	85 mJ	720 seg	Carbonización
Punto G	120 mJ	120 seg	Carbonización
Punto H	120 mJ	180 seg	Carbonización
Punto I	120 mJ	720 seg	Ablación

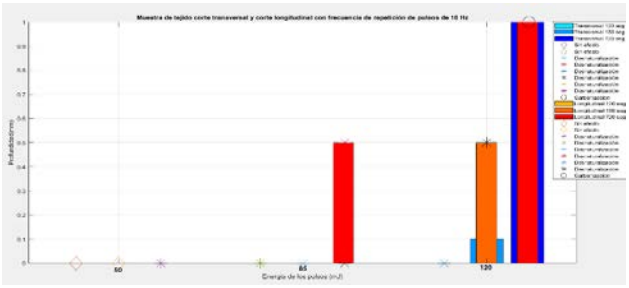


***Figura 5**. Diagrama de barras comparando las medidas de las profundidades (mm) en función de la energía de pulsos (mJ) de la muestra de tejido muscular transversal y longitudinal.*

Como podemos ver en los resultados obtenidos en la Figura 5, la profundidad de ablación provocada por la exposición láser aumenta a medida que aumentamos la energía de los pulsos (mJ) y el tiempo de exposición láser (seg).

Si comparamos este tejido muscular cortado de manera transversal con el cortado de manera longitudinal (Figura 5) vemos que este último tiene menor resistencia. Esto es debido a que cuando lo cortamos de manera transversal estamos rompiendo el conjunto de fibras haciendo que el tejido libere agua. Este agua va a actuar como "capa de protección" frente a la radiación láser debido a que va a mantener el tejido a una temperatura de 100ºC hasta que se evapore, por ello en vez de asarse y aumentar la temperatura, nuestro tejido se está cociendo. Así que, los efectos aparecerán más tarde y con menor intensidad.

Además del tejido muscular, hemos obtenido resultados de las profundidades de ablación y de los efectos provocados en la epidermis (Figura 6) y en el tejido adiposo (Figura 7). En ambos logramos la ablación del tejido, pero tanto en la epidermis como en el tejido adiposo los efectos térmicos se muestran más tarde que en el caso del tejido muscular.

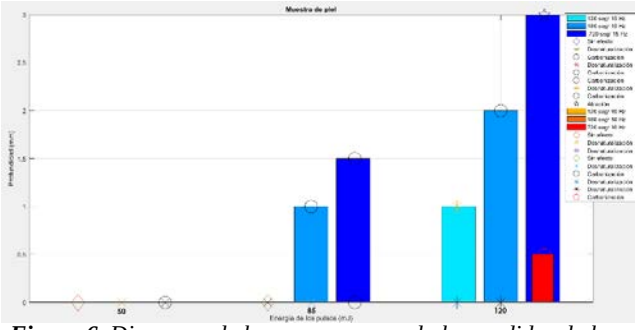


***Figura 6**. Diagrama de barras comparando las medidas de las profundidades (mm) en función de la energía de pulsos (mJ) de la muestra de tejido epitelial.*

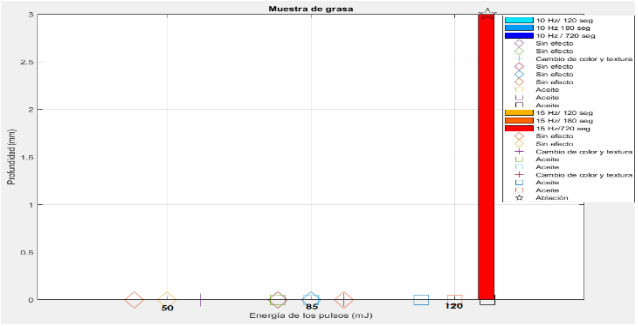


***Figura 7:** Diagrama de barras comparando las medidas de las profundidades (mm) en función de la energía de pulsos (mJ) de la muestra de tejido adiposo.*

5. Conclusión

A través de los resultados obtenidos, sabemos que el tejido adiposo es el tejido que más resistencia ofrece al aumento de temperatura (Figura 7) y que no obtenemos muestras físicas de desnaturalización de las células, ni carbonización, pasamos directamente a la ablación del tejido cuando usamos los valores máximos de los parámetros. El tejido epitelial (Figura 6) ofrece mayor resistencia que el tejido muscular, pero una vez se rompe la capa superficial (la epidermis), actúan de una manera similar logrando fácilmente en los dos tipos de tejido la ablación. Por lo que, además del tipo de tejido utilizado y la forma de corte del tejido, los valores de la energía de pulsos y el tiempo de exposición son determinantes para conseguir la ablación, ya que debemos llegar a un umbral de temperatura concreto para lograrlo.

La valoración del efecto de la ablación se puede automatizar mediante algoritmos de *machine learning* aplicados sobre el análisis espectral de las diferentes muestras. Esto permitiría valorar de forma automatizada el resultado de la intervención, incluso en tiempo real y de forma integrada en un dispositivo quirúrgico.

Agradecimientos

Este Trabajo ha sido parcialmente financiado por el proyecto del Plan Nacional de I+D+i "Modificación mediante alta presión de las propiedades plasmónicas y luminiscentes de nanopartículas de metales y óxidos desnudas y recubiertas" (PGC2018-101464-B-I00), del Ministerio de Ciencia, Innovación y Universidades, cofinanciado con fondos FEDER y por la Fundación San Cándido.

Referencias

[1] Félix Fanjul-Vélez, Irene Salas-García and José Luis Arce-Diego. "Analysis of laser surgery in non-melanoma skin cancer for optimal tissue removal", Laser Physics 25, 025606 (8pp) (2015)

[2] Markolf H. Niemz. Laser-Tissue Interactions. Third, Enlarged Edition. Springer-Verlag Berlin Heidelberg 2004, 2007 (ISBN: 3-540-40553-4).

[3] F. H. Loesel, M. H. Niemz, J. F. Bille, T. Juhasz, "Laser-Induced Optical Breakdown on Hard and Soft Tissues and Its Dependence on the Pulse Duration: Experiment and Model," IEEE Journal of Quantum Electronics 32, 1717-1722 (1996).

[4] A. Vogel, V. Venugopalan, "Mechanisms of Pulsed Laser Ablation of Biological Tissues," Chem. Rev. 103, 577-644 (2003).

[5] American Society for Gastrointestinal Endoscopy, "Mucosal ablation devices," Gastrointestinal Endoscopy 68, 1031-1042 (2008).

[6] N. Muguruma, K. Okamoto, T. Kimura, K. Kishi, T. Okahisa, S. Okamura, T. Takayama, "Endoscopic ablation therapy for gastrointestinal superficial neoplasia," Digestive Endoscopy 24, 139–149 (2012).

Simulación y Planificación Quirúrgica

Jueves, 26 de noviembre de 2020

ARHealth: App que combina Realidad Aumentada e Impresión 3D para Mejorar el Flujo Quirúrgico

A. Pose-Díez-de-la-Lastra[1,2], R. Moreta-Martínez[1,2], J. Calvo-Haro[2,3], L. Mediavilla-Santos[2,3], R. Pérez-Mañanes[2,3], J. Pascau[1,2]

[1] Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid

[2] Instituto de Investigación Sanitaria Gregorio Marañón, Madrid

[3] Servicio de Cirugía Ortopédica y Traumatología, Hospital General Universitario Gregorio Marañón, Madrid

Resumen

La realidad aumentada (AR) y la impresión 3D (I3D) son tecnologías introducidas recientemente en el flujo de trabajo de algunos hospitales. La I3D se emplea para crear biomodelos anatómicos de los pacientes que mejoran la planificación quirúrgica y sirven como entrenamiento para afrontar cirugías complejas. La AR, aunque no está tan consolidada en este ámbito, ha mostrado recientemente interesantes aportaciones. En este proyecto se plantean soluciones innovadoras combinando I3D y AR en el área clínica de la cirugía ortopédica oncológica. Se propone la creación de una aplicación para teléfonos móviles (app) que emplea la AR para mostrar las estructuras anatómicas de los pacientes, en tamaño real. Para su desarrollo, se proponen una serie de pasos, incluyendo: adquisición de imágenes médicas, segmentación, creación de biomodelos, I3D de modelos y herramientas, e integración de toda la información en la app. Esta app será utilizada durante la planificación quirúrgica, la comunicación con el paciente y la cirugía en sí con el objetivo de asistir a lo largo de todo el flujo de quirúrgico.

1. Introducción

Tanto la impresión 3D (I3D) como la realidad aumentada (AR) son tecnologías que están mostrando actualmente su utilidad creciente en el ámbito de la medicina. Principalmente se utilizan para caracterizar la anatomía de los pacientes, dado que en algunos casos ésta es difícil de interpretar con imágenes planas [1]. La I3D consiste en crear objetos tridimensionales a partir de un archivo digital superponiendo capas sucesivas de material hasta fabricar el diseño virtual [2]. Por otro lado, la AR superpone objetos virtuales tridimensionales a objetos físicos en el espacio real, permitiendo interactuar con ambos simultáneamente [3].

Un ejemplo de utilización de I3D en este ámbito sería el estudio de [4], en el que se emplean biomodelos de la anatomía de los pacientes como complemento para explicarles su condición y los posibles tratamientos. Por otro lado, en el proyecto publicado en [5], se resalta que la AR ofrece una visualización intuitiva de las estructuras internas de los pacientes durante los procedimientos, en su caso, de cirugía espinal. Ambas tecnologías se han evaluado también en combinación de cara a la planificación y navegación quirúrgica, con la intención de reducir las dosis de contraste y los riegos para el paciente [1].

Los buenos resultados obtenidos en los estudios anteriores garantizan la utilidad de la AR y la I3D en el ámbito de la salud, dada la precisión, interpretabilidad y fiabilidad que ofrecen. No obstante, no están exentas de desventajas, y todos los estudios coinciden en las mismas limitaciones. En el caso de la I3D, la más importante es la cantidad de tiempo y material necesario para imprimir grandes modelos anatómicos. Por su parte, la AR tiene como mayor inconveniente el registro entre el mundo real y el mundo virtual [6].

Nuestro equipo de investigación ha explorado con anterioridad la utilización de ambas tecnologías en el ámbito clínico [7, 8]. En este proyecto se pretende ir un paso más allá y crear un nuevo sistema basado en AR e I3D para su incorporación definitiva en los distintos pasos de un flujo quirúrgico. Específicamente, se ha diseñado una novedosa aplicación para móviles (app) que utiliza la AR para proyectar, o aumentar, las estructuras anatómicas internas de los pacientes con ayuda de la I3D. De este modo pretendemos afrontar las limitaciones mencionadas. Primero, mediante la impresión de porciones anatómicas más pequeñas aumentando lo restante gracias a la app, y segundo, registrando los datos reales con los virtuales de manera automática utilizando un marcador de AR fabricado usando la I3D. De esta forma los beneficios de cada técnica se aprovechan para solventar las desventajas de la otra.

El sistema propuesto es fácilmente adaptable a cada paciente para favorecer su universalidad. Se pretende alcanzar un triple objetivo dentro del flujo quirúrgico: servir de soporte durante la planificación quirúrgica, mostrando la anatomía del paciente en 3D y en tamaño real; asistir durante la comunicación con el paciente en la explicación de la patología y el abordaje quirúrgico; y, por último, ayudar durante las intervenciones, proyectando las estructuras anatómicas internas de los pacientes, *in situ*, para guiar a los cirujanos y aumentar su confianza.

2. Materiales y Métodos

2.1. Selección de pacientes

Para el diseño de la app, se ha utilizado la información anatómica de 4 pacientes de cirugía ortopédica oncológica. Con el objetivo de mantener su anonimato, se

ISBN: 978-84-09-25491-0, pags. 365- 368

le asignó a cada uno un código alfanumérico. Para evaluar la variabilidad del sistema con la patología, los tumores de los pacientes seleccionados tenían diferente tamaño, afectaban a un tipo de hueso distinto y estaban localizados en regiones anatómicas dispares: uno en el tórax (Paciente AR3DP0003), otro en una extremidad (Paciente AR3DP0005) y dos en la zona pélvica (Pacientes AR3DP0002 y AR3DP0004).

2.2. Adquisición de Imágenes Médicas

La tomografía axial computarizada (TAC) es la mejor modalidad para el diseño de estructuras anatómicas en el ámbito de la ortopedia y la traumatología, dada la buena resolución y contraste que ofrecen en las estructuras óseas [9]. Por esta razón, en esta aplicación se usa la imagen TAC adquirida para cada paciente, que servirá como imagen de referencia para la segmentación.

2.3. Segmentación

A partir de cada una de esas imágenes, se realiza la segmentación para seleccionar, de toda la imagen, sólo las regiones de interés. Este proceso se lleva a cabo de forma semiautomática combinando segmentación por umbral y edición manual con el software especializado de *3D Slicer* [10].

2.4. Modelado 3D

Tras la segmentación, se renderizan modelos 3D tanto del tumor como del hueso circundante, en formato de estereolitografía (STL) y se exportan al programa *MeshMixer*. En este software se diseñan las guías quirúrgicas, así como otros modelos 3D que servirán para los pasos sucesivos.

Las guías quirúrgicas son piezas específicas para cada paciente diseñadas para ser introducidas y fijadas en el interior del paciente durante las cirugías, a través de una incisión. Su superficie se crea como un negativo del relieve de los huesos, por lo que se ajustan perfectamente a la curvatura de los mismos. Hasta ahora, se han utilizado en algunos hospitales para delimitar regiones de interés o marcar los planos de corte propuestos durante la planificación quirúrgica [8]. En este proyecto hemos ido un paso más allá e incluimos en el diseño un pequeño soporte para poder fijar sobre ellas un marcador de realidad aumentada que se empleará durante la cirugía.

Los cirujanos son los encargados de seleccionar el área del hueso a partir de la cual se crea la guía, teniendo en cuenta la posición que tendrá el paciente durante la intervención, su línea visual y el tejido circundante. Las guías tienen que ser suficientemente grandes como para que su superficie sea reconocible y no pueda ajustarse a otra región del hueso y, a la vez, suficientemente pequeñas para minimizar el tiempo de impresión y además no ocupar un espacio excesivo en el paciente.

Además de la guía quirúrgica, *MeshMixer* se utiliza para crear otros modelos 3D, como un pequeño fragmento de hueso del área en la que se localizará la guía quirúrgica para practicar su posicionamiento antes de la cirugía. También se crean otros fragmentos anatómicos para la app de AR. En la Figura 1 se muestran los biomodelos de los 4 pacientes del estudio, con los huesos en blanco y los tumores en rojo. Las guías quirúrgicas aparecen en su posición correspondiente, en verde, y unidos a ellas, están los marcadores de realidad aumentada.

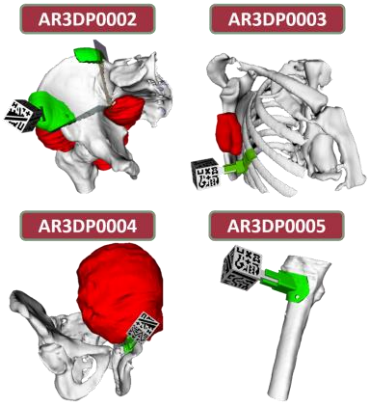


Figura 1. Modelos 3D de los pacientes del experimento.

2.5. Impresión *3D*

Una vez se definieron los modelos 3D, las guías quirúrgicas se imprimieron en una resina transparente (material biocompatible de tipo I) que les permite estar en contacto con el interior del paciente durante las intervenciones quirúrgicas. Se imprimieron con la impresora *Formlabs Form 2 SLA 3D printer* (Formlabs Inc., USA), que usa la técnica de impresión de estereolitografía.



Figura 2. Marcador tridimensional de AR

El resto de biomodelos se imprimieron mediante la técnica de modelado por deposición fundida (FDM) en ácido poliláctico (PLA) con las impresoras *Ultimaker 3 Extended* y *Ultimaker S5* (Ultimaker B.V. Países Bajos).

2.6. App de AR

La aplicación de móviles creada para este proyecto se llama *ARHealth* y fue desarrollada en *Unity 2019.3*. Es compatible con dispositivos Android y iOS y utiliza la cámara del móvil para detectar y rastrear la posición de un marcador de realidad aumentada en el espacio real. En este proyecto se utilizó el marcador tridimensional presentado en [7] (Ver Figura 2). Su tamaño es de *3x3x3 cm*, está impreso en PLA y contiene patrones distintos en cada cara. Para detectar el marcador en el campo visual de la cámara se utilizó el kit de desarrollo de software (SDK) de reconocimiento de patrones *Vuforia*. Importando los modelos virtuales y el marcador de AR a *3D Slicer* es posible calcular sus coordenadas relativas y almacenarlas en la app. De este modo, cuando *Vuforia* detecta un patrón del marcador, procesa y proyecta con precisión y en tiempo real los elementos virtuales en la localización y orientación especificada, permitiendo un registro automático de las estructuras.

La interfaz de la app consiste en un menú principal en el que se selecciona el paciente de interés. Una vez elegido, se despliega un segundo menú que ofrece tres modos de visualización: *Demo*, *Clinic* (Clínica) y *Surgery* (Cirugía). Ambos menús se muestran en la Figura 3.

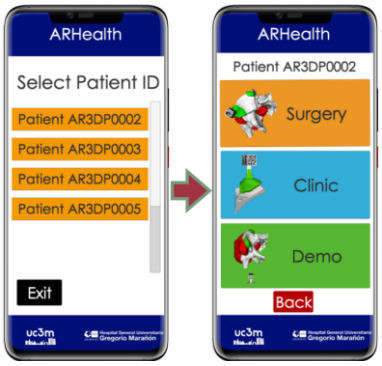


Figura 3. Capturas de pantalla de los diferentes menús de la app de AR.

- *Demo*: En este modo, todos los modelos se muestran alrededor del marcador de AR sin registrar con el paciente. El marcador se puede rotar para ver los modelos virtuales desde cualquier perspectiva. Este modo ha sido creado para que el cirujano eche un primer vistazo al caso sin la necesidad de tener ningún modelo impreso, ahorrando así tiempo y material.
- *Clinic*: Para utilizar este modo, es necesario colocar el marcador del AR en su soporte de la guía quirúrgica. Simulando que la guía está colocada en el hueso objetivo, los modelos virtuales se proyectan referenciados a la misma, en su posición correspondiente y relativa a la anatomía del paciente. Este modo está pensado para ser utilizado en la clínica, sin el hueso real del paciente. Por ello, se imprime en PLA un fragmento de hueso que representa la región del mismo sobre la que se acoplará la guía quirúrgica. Todos los modelos se proyectan a partir de él.
- *Surgery:* Este modo fue el creado para ser utilizado durante las cirugías. En este caso, de nuevo, el marcador de AR debe ser introducido en el soporte de la guía quirúrgica y ésta, en lugar de ajustarse a un fragmento de PLA, se coloca en el hueso real del paciente. En este modo únicamente se muestran los biomodelos que no son evidentes, como el tumor o los planos de corte propuestos. El hueso virtual del paciente se mantiene oculto en este caso, dado que el hueso real será ya visible para el cirujano durante la operación.

  La Figura 4 contiene capturas de pantalla de los modos de visualización en el paciente AR3DP0003 (*Demo*, tumor escapular), AR3DP0005 (*Clinic*, tumor femoral) y AR3DP0004 (*Surgery*, tumor en la cadera). En todos los modos, se añadió un menú desplegable en el que poder controlar la visualización y transparencia de cada uno de los modos propuestos.

2.7. Usos de *ARHealth*

A) Planificación quirúrgica: Tanto el modo *Demo* como el modo *Clinic* podrían utilizarse para facilitar la planificación quirúrgica. El primero serviría para que el cirujano pudiera echar un primer vistazo al caso sin la necesidad de tener ningún biomodelo impreso en 3D. El segundo podría también ser útil para practicar con el posicionamiento de la guía quirúrgica antes de la intervención. Asimismo, podría servir para analizar y discutir estrategias para los procedimientos quirúrgicos, comparar posibles perspectivas e incluso, instruir a cirujanos inexpertos.

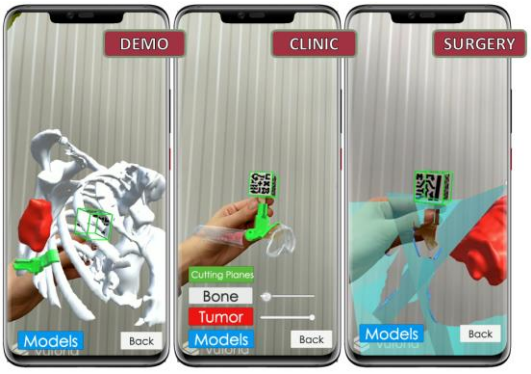


Figura 4. De izquierda a derecha, los modos de visualización *Demo*, *Clinic* y *Surgery* para los pacientes AR3DP0003 (tumor escapular), AR3DP0005 (tumor femoral) y AR3DP0004 (tumor en la cadera).

B) Comunicación con el paciente: Antes de las cirugías, los pacientes podrían utilizar la app como refuerzo para entender su situación y el posible abordaje quirúrgico. Para ello utilizarían el modo *Clinic* y en él, podrían rotar el marcador y modificar la transparencia de los biomodelos para comprender la magnitud de sus tumores y la parte de su hueso que sería reseccionado, en caso de ser necesario. Además, también podrían practicar para situar la guía quirúrgica en el fragmento de hueso de PLA, aprendiendo así cómo funciona la realidad aumentada y cómo podría ser utilizada en su caso.

C) Cirugías: Durante las operaciones, esta tecnología podría ser de gran utilidad para mostrar, *in situ*, la localización del tumor, y así aumentar la confianza de los cirujanos durante los procedimientos. Para poder utilizar este modo en los quirófanos, tanto la guía quirúrgica como el marcador de AR deben esterilizarse. La guía se esteriliza con una técnica basada en óxido de etileno (EtO) a 55ºC. Con respecto al marcador de AR, puesto que la temperatura de transición del PLA está en el rango 55-65 ºC, se lleva a cabo el mismo proceso que con la guía pero esta vez a sólo 37 ºC, para evitar la deformación. A continuación, se lleva a cabo un desgaseado extensivo para eliminar el EtO residual [11].

Una vez en la intervención, el cirujano inicialmente realiza una incisión en el paciente hasta alcanzar el hueso afectado y fija en él la guía quirúrgica (con el marcador de AR). A continuación, introduce el smartphone con la aplicación de *ARHealth* en una bolsa estéril y selecciona el modo *Surgery*. Después apunta hacia el marcador de AR y, una vez reconocido, las estructuras virtuales se proyectan automáticamente,

registradas con la anatomía real. Así, antes de empezar con la resección tumoral, los cirujanos pueden ver el modelo virtual del tumor superpuesto en su posición verdadera, así como los planos de corte previstos, si procede, para saber exactamente dónde cortar.

Para cuantificar la valoración de los cirujanos respecto a esta tecnología, se desarrolló una encuesta de 16 preguntas en las que los cirujanos evaluaron en una escala del 1 al 5, aspectos como la interpretabilidad de la aplicación o su utilidad en los diferentes pasos del flujo quirúrgico.

3. Resultados

Una vez desarrollado el sistema, se probó con dispositivos Android y iOS de diferentes gamas y generaciones y se pudo comprobar que la detección del marcador de AR por la app no sólo era factible sino también, prácticamente inmediata en todos ellos. Esto sugiere que no es necesaria una capacidad computacional muy potente, sino que cualquier smartphone con cámara podrá utilizar la app. Para la detección inicial, el marcador debe estar a unos 20 cm (aprox.). Después, el dispositivo se puede alejar y la visualización se mantiene mientras el patrón sea visible. Por otro lado, 5 cirujanos especializados probaron la app a lo largo del flujo quirúrgico y contestaron a la encuesta desarrollada, obteniendo una nota media total de 4.54. Esto nos lleva a pensar que el potencial del sistema es muy alto tanto para entender más fácilmente la patología en un comienzo, planear el abordaje y hablar con el paciente, como para complementar la información de la que disponen durante las propias intervenciones.

4. Conclusión

En los últimos años, el uso de AR e I3D se ha probado en varias áreas del campo de la medicina. Específicamente, en este proyecto se ha propuesto desarrollar una novedosa aplicación de smartphones llamada *ARHealth* que tiene como objetivo asistir durante la planificación quirúrgica, la comunicación con el paciente y las cirugías del área de ortopedia oncológica. Esta app emplea la AR para proyectar las estructuras anatómicas internas del paciente, como los huesos o los tumores, y añade una serie de modelos extra como los planos de corte para guiar la resección del hueso afectado. El sistema se ha validado con la información de 4 pacientes de cirugía ortopédica oncológica. Los buenos resultados en la visualización de la app en relación al registro de los elementos reales y virtuales, la inmediatez de la detección del marcador, y la buena acogida de los médicos, resultan prometedores para las tareas propuestas. Como trabajo futuro, se propone la implementación del sistema a lo largo de todo el flujo quirúrgico en nuevos pacientes para cuantificar su impacto en más casos reales.

Agradecimientos

Este trabajo ha sido subvencionado por el Ministerio de Ciencia, Innovación y Universidades, Instituto de Salud Carlos III, proyecto PI18/01625, cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER), “Una manera de hacer Europa”, y por la Comunidad de Madrid, proyecto IND2018/TIC-9753. La primera autora disfruta de un contrato de innovación asistencial financiado por el Instituto de Investigación Sanitaria Gregorio Marañón.

Quería agradecer a todas las entidades colaboradoras el haber hecho posible el desarrollo de este proyecto. Asimismo, agradecer al CASEIB la oportunidad de difundir este trabajo a nivel nacional y proporcionarnos guías de referencia para la redacción de este documento.

Referencias

[1] Witowski J, Darocha S, Kownacki K, Pietrasik A, Pietura R, Banaszkiewicz M, Kaminski J, Biederman A, Torbicki A, Kurzyna M. Augmented reality and three-dimensional printing in percutaneous interventions on pulmonary arteries. *Quantitative Imaging in Medicine and Surgery*, vol 9,2018,pp. 3-29, https://doi.org/10.21037/qims.2018.09.08.

[2] Ventola C. Medical applications for 3D printing: Current and projected uses. PT: a peer reviewed journal for formulary management, vol 39, 2014, pp.704-711. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4189697/pdf/ptj4910704.pdf.

[3] Shuhaiber J. Augmented reality in surgery. *Archives of surgery (Chicago, III: 1960)*, vol 139, 2004, pp.170-174.

[4] Sander I, Liepert T, Doney E, Leevy W, Liepert D. Patient education for endoscopic sinus surgery: Preliminary experience using 3D-printed xlinical imaging data. *Journal of Functional Biomaterials*, vol 8, 2017, pp. 13-19, https://doi.org/10.3390/jfb8020013

[5] Carl B, Bopp M, Sass B, Pojskic M, Nimsky C. Augmented reality in intradural spinal tumor surgery. *Acta Neurochirurgica*, vol 161, 2019, pp.2181-2193, https://doi.org/10.1007/s00701-019-04005-0.

[6] Monsky WL, James R, Seslar SS. Virtual and augmented reality applications in medicine and surgery – the fantastic voyage is here. *Anatomy Physiology: Current Research*, 2019, pp. 1-5, https://doi.org/10.4172/2161-0940.1000313.

[7] Moreta-Martínez R, García-Mato D, García-Sevilla M, Perez-Mañanes R, Calvo-Haro J, Pascau J. Combining augmented reality and 3D printing to display patient models on a smartphone. *Journal of Visualized experiments*, vol 155, 2020, pp. 1-10, https://doi.org/10.3791/60618.

[8] Perez-Mañanes R, Arnal-Burró J, Manaute J, Rodriguez F, Vaquero, J. 3D surgical printing cutting guides for open-wedge high tibial osteotomy: Do it yourself. *Journal of Knee Surgery*,vol29,2016,p.S131, https://doi.org/10.1055/s-0036-1572412.

[9] Abhari K, Baxter J, Chen E, Khan A, Peters T, De Ribaupierre S, Eagleson R. Training for planning tumour resection: Augmented reality and human factors. *IEEE Transactions on Biomedical Engineering*, vol 62, 2014, pp.1466-1477, https://doi.org/10.1109/TBME.2014.2385874.

[10] Pieper S, Halle M, Kikinis R. 3D Slicer. *2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821)*, vol 1, 2004, pp. 632-635, https://doi.org/10.1109/ISBI.2004.1398617

[11] Zislis T, Martin SA, Cerbas EL, Heath JR, Mansfield JL, Hollinger JO. A scanning electron microscopic study of in vitro toxicity of ethylene-oxide-sterilized bone repair materials. *The Journal of Oral Implantology*, vol 15, 1989, pp.41-46, https://doi.org/10.1177/2473974X18804492.

Diseño e implementación de un sistema electrónico para el control de un simulador de trauma

B. Larraga García[1], R.H. Mejía Blanco[1], J. Rubio Bolívar[2], M. Quintana Díaz[2], A. Gutiérrez Martín[1]

[1] Departamento de Fotónica y Bioingeniería, ETSI de Telecomunicación, Universidad Politécnica de Madrid, Madrid. España, {rafaelmejia.0495}@gmail.com, {blanca.larraga, a.gutierrez}@upm.es

[2] Instituto de Investigación IdiPAZ, Hospital Universitario La Paz, Madrid. España. {javier84.rubio, mquintanadiaz}@gmail.com

Resumen

El trauma es la principal causa de muerte en personas menores de 45 años y una de las principales causas de muerte en el mundo. Sin embargo, las muertes causadas por shock hipovolémico como consecuencia de un trauma son potencialmente prevenibles. Se han identificado que las lesiones que se presentan en la pelvis y en las extremidades inferiores provocadas por accidentes de tráfico producen una mayor probabilidad de muerte cuando el sangrado no se localiza rápidamente. Es por esto que entrenar mediante la simulación clínica a profesionales médicos para la atención de pacientes que sufren un trauma de esta naturaleza es clave. El objetivo de este trabajo es diseñar un sistema electrónico que controle un simulador de trauma creando diferentes escenarios clínicos que ayuden a los equipos médicos a entrenar y mejorar las técnicas y habilidades necesarias para detener eficazmente las situaciones de sangrado como consecuencia de un trauma.

1. Introducción

El shock hipovolémico es una consecuencia de un trauma que, aunque potencialmente prevenible, aumenta el riesgo de mortalidad durante el tratamiento del paciente por la importante pérdida de sangre que irriga los tejidos [1]. En situaciones críticas, los mecanismos de compensación por la pérdida de sangre no son suficientes y las transfusiones de sangre son necesarias para salvar las vidas de los pacientes [2,3,4].

El volumen de sangre en un adulto representa el 7 por ciento del peso del cuerpo (70 ml/kg de peso). Así pues, el shock hipovolémico puede clasificarse en cuatro niveles, tal y como se muestra en la Tabla 1 [5].

Parámetros	Nivel 1	Nivel 2	Nivel 3	Nivel 4
Pérdida de sangre (mL)	750	750-1500	1500-2000	>2000
% sangre perdida	15%	15-30%	30-40%	>40%
Ritmo cardíaco (latidos/min)	<100	100-120	120-140	>140
Frecuencia respiratoria (respiraciones/min)	14-20	20-30	30-40	>40
Síntomas del SNC	Normal	Ansiedad	Confusión	Letargo

Tabla 1. *Niveles de shock hemorrágico.*

El primer objetivo de los clínicos en situaciones de trauma es restablecer el volumen de sangre, parando las hemorragias que pudieran existir y aplicando protocolos frente a hemorragias masivas.

Así pues, para poder ofrecer una mejor atención a pacientes en estas situaciones, se han diseñado programas de simulación que permitan reproducir escenarios lo más realistas posible. En este artículo se diseña un sistema electrónico que controlará esos escenarios para traumas pélvicos y traumas en extremidades inferiores. Adicionalmente, este sistema recabará información acerca de las actuaciones realizadas por los clínicos que están practicando los escenarios simulados. De esta manera se podrá obtener feedback a tiempo real de las actuaciones llevadas a cabo por los clínicos.

El artículo se organiza de la siguiente manera: en la Sección 2 se presenta el diseño hardware realizado. La Sección 3 explica el diseño firmware para el correcto funcionamiento del sistema electrónico. En la Sección 4 se presentan los resultados y la validación del sistema desarrollado. Por último, la Sección 5 presenta las conclusiones.

2. Diseño hardware

Teniendo en cuenta el importante riesgo y las consecuencias de las lesiones causadas por un trauma, la propuesta de crear un sistema electrónico basado en módulos capaces de simular diferentes escenarios se centrará en la automatización de los principales tipos de lesiones. El principal objetivo del diseño hardware es que el sistema electrónico envíe y reciba información de un simulador a tiempo real. Este simulador utilizará actuadores para producir los escenarios de trauma y sensores para evaluar las acciones llevadas a cabo por el personal clínico.

El diseño hardware presentado en esta Sección se instalará en un simulador que no es el foco de este trabajo. No obstante, este sistema electrónico se integrará dentro de ese simulador para poder recrear los diferentes escenarios de trauma.

2.1. Trauma pélvico

La pelvis es una estructura ósea que se compone de tres partes principales: el ilion, el pubis y el isquion [6] y cuyos huesos principales se muestran en la Figura 1.

ISBN: 978-84-09-25491-0, pags. 369- 372

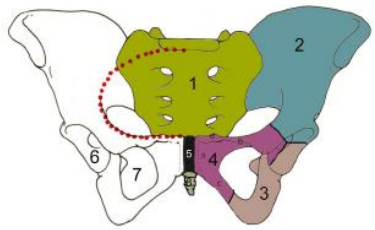


Figura 1.*Huesos de la pelvis:1. Sacro, 2. Ilion, 3. Isquion, 4. Pubis, 5. Sínfisis del pubis, 6. Acetábulo, 7. Anillo pélvico. 8. Coxis. [7]*

Cuando existe una fractura pélvica, la primera maniobra a realizar es la colocación de un cinturón pélvico. Es importante conocer el impacto que esta colocación tiene en la fractura, así como su ubicación. Para que esta maniobra sea eficiente, el cinturón pélvico debe colocarse a la altura del trocánter mayor [8].

En el simulador se definen diferentes tipos de fracturas según el sistema de clasificación Tile [9], que clasifica las fracturas de pelvis en:

- Fracturas tipo A: fracturas estables.
- Fracturas tipo B: fracturas que causan inestabilidad rotacional siendo parcialmente inestables.
- Fracturas tipo C: fracturas que producen inestabilidad rotacional y vertical. Estas fracturas son inestables.

Mediante actuadores electromecánicos se crearán estos diferentes tipos de fracturas insertándolos entre las fisuras de la pelvis. Además, se instalarán sensores ópticos junto con los actuadores para obtener información acerca de la unión de los huesos fracturados una vez se coloque un cinturón pélvico. Finalmente, se colocarán también sensores de contacto que proporcionarán una señal de salida binaria con respecto a la posición del cinturón pélvico colocado, determinando si es correcta su colocación o no. Así pues, nuestro sistema electrónico tendrá que enviar y recibir información de estos dispositivos en tiempo real.

2.2. Trauma extremidades inferiores

Para el diseño de una lesión de trauma en extremidades inferiores se han creado escenarios de simulación para que los clínicos puedan practicar maniobras de control de hemorragias: presionando manualmente o mediante la colocación de un torniquete. Hay que tener en cuenta los vasos sanguíneos principales que irrigan la extremidad inferior y producir diferentes tipos de hemorragias en función de si el sangrado es arterial o venoso o si ambos ocurren a la vez siendo pulsátil el arterial y continuo el venoso. Para ello, se utilizarán bombas hidráulicas. Se colocarán sensores de contacto para dar información sobre la colocación del torniquete, puesto que debe situarse entre 5 y 10 cm por encima de la herida [10]. Además, se colocarán sensores de presión para evaluar la presión que está ejerciendo el torniquete o la presión manual que el clínico está aplicando sobre la herida. De nuevo, el sistema electrónico tendrá que enviar y recibir información de todos estos dispositivos produciendo diferentes escenarios y evaluando la respuesta frente a los mismos en tiempo real.

2.3. Microcontrolador y conexión electrónica

El microcontrolador seleccionado es el STM32F446 de ST microelectronics. Este microcontrolador es un Arm Cortex-M4 de 32-bits que puede operar a una frecuencia de 180 MHz.

En la Figura 2 se presenta la conexión de los sensores y actuadores con el microcontrolador. Adicionalmente se utiliza un sistema de monitorización basado en Labview de National Instruments, que se comunica con el microcontrolador a través del puerto serie.

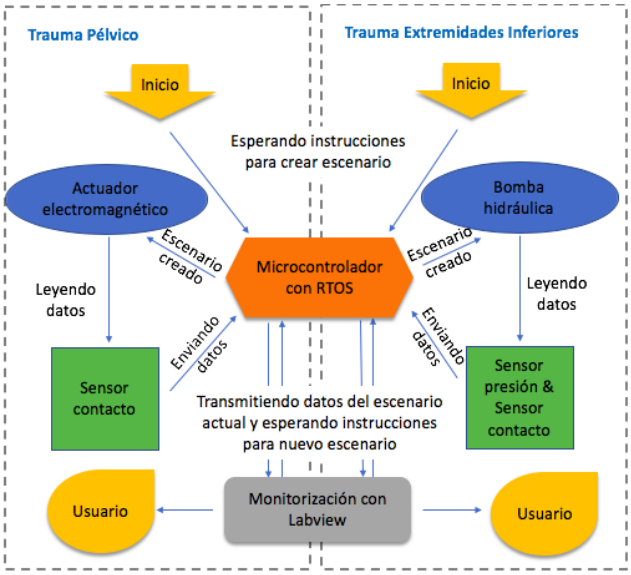


Figura 2.*Esquema de funcionamiento del sistema electrónico para simulación de trauma.*

Dependiendo del escenario que los clínicos quieran simular, el microcontrolador envía instrucciones para crear los escenarios y recibe información de las actuaciones que se están llevando a cabo en tiempo real de los diferentes sensores. Además, toda esta información en tiempo real se mostrará en una interfaz gráfica que los clínicos verán durante la simulación.

3. Diseño firmware

El firmware ha sido desarrollado utilizando el STM32CubeIDE de ST, que permite configurar los diferentes periféricos y tareas en tiempo real.

3.1. Configuración del microcontrolador

Esta plataforma está basada en C/C++ y permite generar el código, compilar y depurar. Adicionalmente, para desarrollar este sistema electrónico se necesitan realizar tareas programadas en tiempo real por lo que se configurará un sistema operativo a tiempo real, en concreto FreeRTOS.

Adicionalmente, tal y como se puede ver en la Figura 3, se utilizará un conversor analógico digital (ADC) para los sensores de presión, acceso directo a memoria (Direct Memory Access, DMA) para optimizar la configuración del ADC y una modulación de anchura de pulsos (Pulse Width Modulation, PWM) para controlar las bombas hidráulicas que permitirán los diferentes sangrados. Los sensores de

contacto se configuran como pines digitales con una resistencia pull-down.

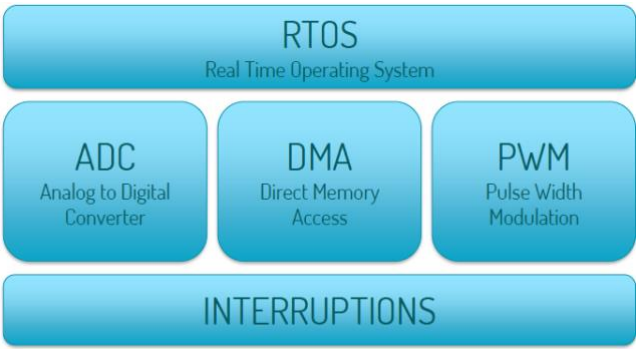


Figura 3.*Esquema de componentes principales utilizados para el desarrollo del software.*

3.2. Programación del software

Tras la configuración de los pines de entrada y de salida, el ADC para la lectura de las entradas analógicas, el PWM para el control de las bombas y las interrupciones para manejar los actuadores, se inicializa la declaración de las diferentes variables y las tareas del sistema operativo en tiempo real; en concreto:

- Una tarea que generará una rutina de interrupción (Interruption Service Routine, ISR) a través de UART.
- Una tarea en FreeRTOS para la simulación del trauma pélvico. Esta tarea generará una fractura pélvica a seleccionar entre los diferentes tipos presentados y evaluará la posición del cinturón pélvico así como el impacto del mismo en la unión de las fracturas pélvicas simuladas.
- Una tarea en FreeRTOS para la simulación del trauma de extremidades inferiores. Esta tarea generará un escenario hemorrágico a seleccionar entre diferentes tipos de hemorragias que se pueden dar y evaluará la presión ejercida para cortar la misma así como la maniobra realizada: presión manual o aplicación de torniquete. Adicionalmente se evaluará la colocación del torniquete en caso de que se haya colocado.

4. Protocolos de simulación

4.1. Implementación en Labview

Para mostrar los datos que se trasmiten por UART a los clínicos que están llevando a cabo las simulaciones, se necesita de un programa que permita el control y el análisis de los mismos para mostrarlos.

Labview es un software para el control y procesamiento de datos. La característica más importante es que permite integrar muy fácilmente las medidas realizadas y representar diagramas lógicos complejos incluyendo el desarrollo de algoritmos de análisis de datos y el diseño de interfaces gráficas a medida. Así pues, se establecerá el flujo de datos e información que se muestra en la Figura 4. En la parte superior se muestra el flujo para el envío de datos desde Labview y a través del VISA Serial Port al microcontrolador, que ejecutará las acciones de los diferentes escenarios de trauma. De esta manera llegará la información a los actuadores que generarán los diferentes escenarios de trauma, ya sea pélvico o de las extremidades inferiores.

En la parte inferior de la Figura 4, se muestra el flujo de datos cuando se recibe información de los sensores, la cual se leerá mostrando en Labview los indicadores pertinentes. Esto permite que el entrenamiento sea muy completo proporcionando información a los clínicos del impacto que tienen sus actuaciones en tiempo real.

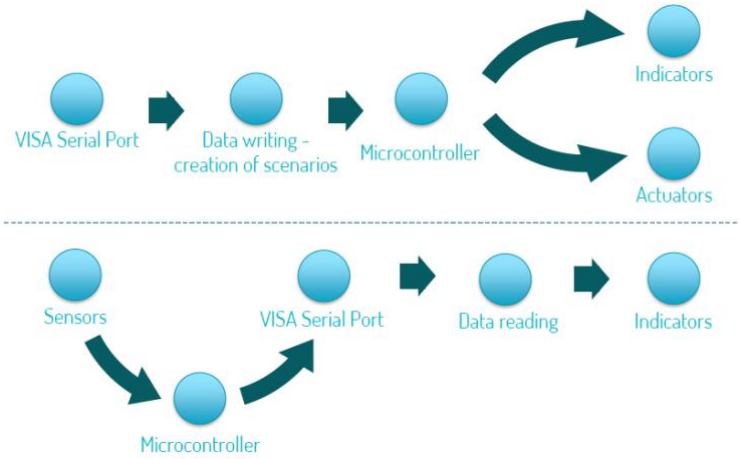


Figura 4.*Esquema de flujo de datos e información entre microcotrolador y Labview.*

4.2. Diseño software

Una vez realizada esta implementación, se ha desarrollado una interfaz gráfica que permite mostrar las diferentes simulaciones así como las acciones llevadas a cabo por los clínicos en formación a lo largo de las mismas.

Primero es necesario seleccionar qué tipo de trauma se va a emular, un trauma pélvico o un trauma en extremidades inferiores. Posteriormente, para poder validar tanto el hardware como el software desarrollado, se desarrollan diferentes escenarios para ambos tipos de trauma. En la Figura 5 se presenta una fractura de pelvis mostrando visualmente los actuadores que se activan para producir la rotura, así como si el cinturón pélvico se ha colocado o no. En caso de colocarse, se indicará si la posición ha sido la adecuada y el efecto que ha tenido el cinturón en la unión de las estructuras óseas fracturadas.

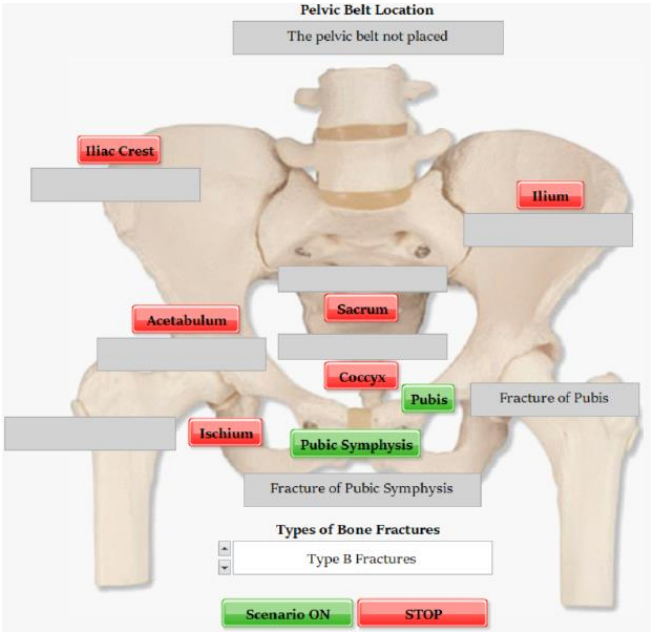


Figura 5.*Interfaz gráfica de Labview para la simulación de un trauma pélvico.*

Igualmente se desarrollarán escenarios para trauma de extremidades inferiores (ver Figura 6). En este caso se muestra el tipo de hemorragia así como la presión ejercida sobre la herida y la posición del torniquete en caso de que haya sido colocado. También se incluye información acerca de la frecuencia cardíaca y del porcentaje se sangre perdido que se obtendrá del escenario creado y la evolución del mismo.

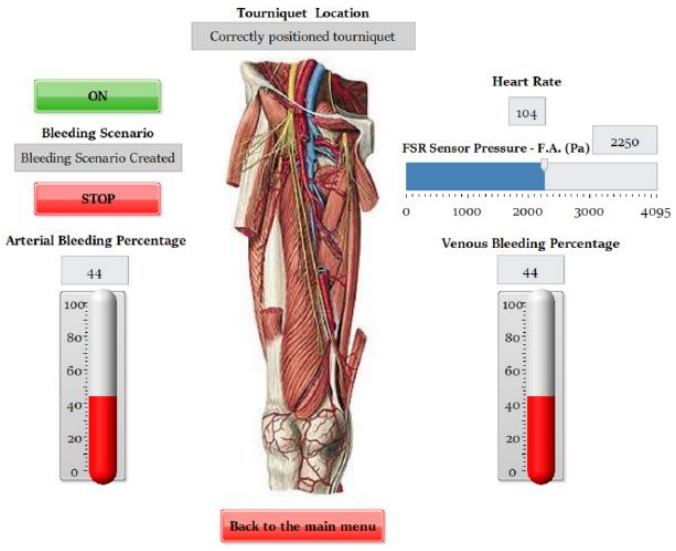


Figura 6.*Interfaz gráfica de Labview para la simulación de una trauma en extremidades inferiores.*

5. Conclusiones

El principal objetivo de este trabajo ha sido el diseño e implementación de un sistema electrónico que permita realizar diferentes escenarios de simulación para el entrenamiento de habilidades y técnicas necesarias para atender a pacientes de trauma y recabe información acerca de las actuaciones de los clínicos en formación.

El entrenamiento continuo a través de este sistema electrónico ayuda a mejorar las técnicas y habilidades necesarias permitiendo un aprendizaje más rápido y automatizado por lo que permitirá tener integradas las técnicas aprendidas de una manera muy eficiente.

Este sistema mejora la innovación en sistemas de simulación al permitir una realimentación continua entre un emulador físico y una interfaz gráfica. En esta última se muestra toda la información necesaria para que los clínicos puedan llevar a cabo su formación en este tipo de escenarios.

Referencias

[1] David F Gaieski and ME Mikkelsen. Definition, classification, etiology, and pathophysiology of shock in adults. UpToDate, Waltham, MA. Accessed, 8:17, 2016.

[2] David S Kauvar and Charles E Wade. The epidemiology and modern management of traumatic hemorrhage: Us and international perspectives. Critical Care, 9(5):S1, 2005.

[3] D Thomas, M Wee, P Clyburn, I Walker, K Brohi, P Collins, H Doughty, J Isaac, PM Mahoney, and L Shewry. Association of anaesthetists of great britain and ireland. Blood transfusion and the anaesthetist: management of massive haemorrhage. Anaesthesia, 65(11):1153-1161, 2010.

[4] M Víctor Parra. Shock hemorrágico. Revista Médica Clínica Las Condes, 22(3):255-264, 2011.

[5] Guillermo Gutierrez, David Reines, and Marian E Wulf-Gutierrez. Clinical review: hemorrhagic shock. Critical care, 8(5):373-381, 2004.

[6] Merckmanuals. Fracturas de la pelvis. URL: https://www.merckmanuals.com/es-us/hogar/traumatismos-y-envenenamientos/fracturas/fracturas-de-la-pelvis, 2017.

[7] Fractura de pelvis. URL: https://www.wikiwand.com/es/Fracturadepelvis; 2019.

[8] TJ Bonner, WGP Eardley, N Newell, S Masouros, JJ Matthews, I Gibb, and JC Clasper. Accurate placement of a pelvic binder improves reduction of unstable fractures of the pelvic ring. The Journal of bone and joint surgery. British volume, 93(11):1524-1528, 2011.

[9] Medigraphic. Clasicación de las fracturas de pelvis. PDF: https://www.medigraphic.com/pdfs/orthotips/ot-2008/ot084e.pdf, 2008.

[10] Carlos de Benito Álvarez. El torniquete: Una revisión de sus indicaciones actuales y una propuesta para su empleo por socorristas de emergencias en la atención extrahospitalaria. Asociación Española de Socorristas de Emergencias, 1(4):1-11, 2008.

Fabricación de moldes mediante impresión 3D: Viabilidad de machos solubles para órganos artificiales huecos.

J.C. Gómez Blanco[1], J.B. Pagador Carrasco[1], A. Rodríguez[2], A. González-González[2], M. Olivares-Marín[2], J. Usón[1], F.M. Sánchez Margallo[1],

[1] Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España
{jcgomez,juson,msanchez,jbpagador}@ccmijesususon.com

[2] Centro Universitario de Mérida, Universidad de Extremadura, Mérida, España, {arodrigulu}@alumnos.unex.es,{agg, maraom}@unex.es

Resumen

Los órganos artificiales para entrenamiento quirúrgico son una necesidad actual para que los cirujanos noveles puedan entrenar todas las patologías posibles. La creación de estos órganos artificiales se suele realizar mediante coladas de silicona en moldes. En este sentido, la tecnología de fabricación 3D ha permitido que estos moldes sean más complejos a bajo coste. No obstante, la creación de órganos huecos sigue siendo un problema en el moldeo de estructuras orgánicas. Para poder fabricar un órgano hueco es necesario sacar el macho una vez la silicona ha secado. En la mayoría de los casos esta extracción requiere un corte en el modelo ya fabricado con su posterior reparación. Por lo tanto, en este trabajo se propone una metodología de creación de machos solubles de gelatina para poder superar este problema. Los resultados demuestran que es posible obtener vejigas artificiales mediante moldeo de silicona con machos fabricados con un 30% de gelatina.

1. Introducción

El cáncer de vejiga es uno de los cánceres más frecuentes después de los de pulmón, próstata y colon. Su tasa de incidencia es de las más elevadas del mundo y en España se presenta en 30 casos con una tasa de muerte de 5 personas por 100000 habitantes/año [1]. Como todos los tipos de cáncer la forma de abordar el tratamiento se decide en función del grado y del estadio del tumor, siendo la cirugía uno de los procedimientos más utilizados en este tipo de cáncer.

En este sentido los profesionales clínicos utilizan cada vez con mayor frecuencia la impresión 3D para obtener modelos preoperatorios de las cirugías a realizar [2]. Utilizando estos modelos físicos se pueden aumentar las tasas de supervivencia de los pacientes a la vez que reducir el tiempo de operación. No obstante, los modelos preoperatorios no sólo se utilizan para planificar una cirugía, también se utilizan como modelos de entrenamiento [2]. En este último caso, los modelos suelen ser de materiales que imiten la textura de los órganos reales siendo la silicona el material más común. La impresión 3D se utiliza para realizar los moldes donde se conforma el modelo [3]. Pero esta práctica tiene un problema, los órganos huecos. En estos tipos de modelos los huecos se suelen hacer mediante machos rígidos que posteriormente hay que extraer dañando en muchos casos los modelos fabricados.

Con esto en mente, el objetivo de este trabajo es diseñar una metodología para la fabricación de vejigas con tumores para entrenamiento quirúrgico que permita extraer el macho sin dañar el modelo.

2. Métodos y materiales

Para la fabricación de los moldes se parte de un escaneado 3D de una vejiga de entrenamiento, anteriormente validada por cirujanos expertos. Una vez obtenido el modelo 3D de la vejiga se precede a realizar los dos moldes necesarios. El primer molde es el de la vejiga con los huecos para los insertos de los machos. El segundo molde es el del macho del hueco interno de la vejiga. Ambos moldes se imprimieron en Ácido Poliláctico (PLA) con una Ender 3 con la menor altura de capa posible (0.15 mm).

El modelo de vejiga se fabricó en silicona con un agente de curado rápido. El mayor reto está en obtener una manera de extraer el macho del interior de la vejiga sin dañar el modelo de vejiga. Para ello se seleccionó un material que permite el cambio de fase de solido a liquido siendo solido a temperatura ambiente (25 °C) y que al calentarlo a unos 80 °C sea líquido. Se seleccionó la gelatina como material y para asegurar la consistencia en estado sólido se probaron diferentes concentraciones (10, 20 y 30 %). Una vez seleccionada la concentración final se hizo un modelo completo de la vejiga. Se introdujo el modelo en un baño de agua a 80 °C para fundir la gelatina durante 1h. Posteriormente se abrió la vejiga para comprobar la viabilidad del tipo de macho.

3. Resultados y discusión

De los resultados de las diferentes mezclas de gelatina se obtiene que concentraciones menores del 30% no son factibles para usarse como macho. Las concentraciones de 10 y 20% eran muy deformables. Al introducirlas en el molde posterior el macho tocaba la base del molde por lo que no son válidos como macho de moldes. En cambio, la concentración de 30% de gelatina era lo suficientemente rígida para poder aguantar su forma dentro del molde sin deformarse. Es necesario tener en cuenta que la concentración de gelatina a utilizar dependerá del volumen del hueco interior. Para el volumen de una vejiga, una disolución de gelatina del 30% es suficiente, pero ello no hace que volúmenes superiores o inferiores requieran

ISBN: 978-84-09-25491-0, pags. 373- 375

mayor o menor porcentaje de gelatina en la mezcla. No obstante, para asegurar el macho se utilizó una concentración del 35% de gelatina en el macho final (Figura 1). En el moldeo del macho se introdujo un tumor de silicona previamente fabricado para que se uniese al modelo de vejiga final.

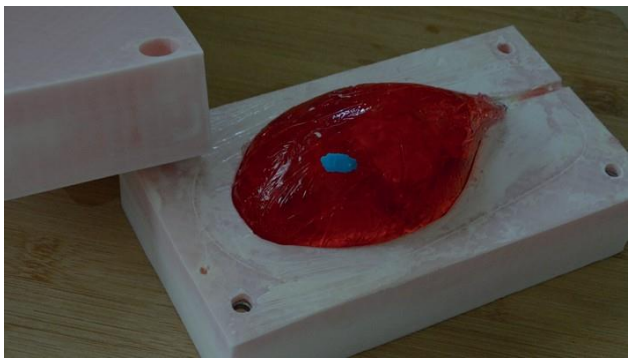


Figura 1.*Macho de 35% de gelatina (rojo) con inserto de tumor de silicona (azul)*

Una vez fabricado el macho con el tumor se introdujo en el molde de la vejiga y se fabricó el modelo final de vejiga (figura 2).



Figura 2.*Modelo de vejiga de silicona*

El modelo de vejiga se introdujo en un baño de agua caliente a 80°C durante una hora para disolver la gelatina. Una vez pasada la hora se realizó un corte en el modelo (figura 3) para comprobar si el macho se había disuelto y el tumor estaba adherido.



Figura 3.*Corte en el modelo de vejiga vaciado con tumor (círculo rojo).*

Como se ve en la figura 3 después del baño en agua caliente no quedan restos del macho y el tumor se ha fusionado con el resto de la vejiga.

A la luz de los resultados podemos concluir que la utilización de machos de gelatina para la creación de huecos internos en vejigas artificiales es viable. Sería necesario comprobar el límite de validez de este tipo de creación de machos para estructuras huecas. Dada la diferente complejidad de los huecos internos de órganos se prevé que sean posibles hacer machos de gelatina de modelos cardiacos, gástricos o similares. No obstante, es necesario realizar pruebas para poder asegurar la viabilidad y las concentraciones de soluto necesarias. La generación de machos mediante gelatina para estructuras internas más complejas como pulmones (red alveolar), con una gran cantidad de ramificaciones y huecos muy pequeños, puede no ser la mejor opción. En estos casos haría falta una gran concentración de gelatina y el rellenado de los moldes de los machos difícilmente sería correcto.

Existen otros tipos de tecnología de fabricación, como la impresión 3D, capaz de generar machos solubles. Utilizando los materiales solubles que existen en la actualidad para impresión 3D se pueden conseguir machos para tiradas de fabricación muy cortas o de machos muy complejos. No obstante, para machos como el validado la técnica de la gelatina puede ser más conveniente por los tiempos de fabricación. Mientras que en una impresora 3D estándar, dado el tamaño del macho, entraría solo un modelo tardando una gran cantidad de tiempo, la utilización del proceso de gelatina tardaría una menor cantidad de tiempo para hacerse. Además, para fabricar en paralelo serían necesarias varias impresoras 3D mientras que para la gelatina sólo se necesitan más moldes, lo que implica un coste asociado a la fabricación mucho menor. También hay que tener en cuenta el coste y el tiempo de mantenimiento, mientras que para los moldes es cero, en la impresora 3D es necesario actuar de manera periódica para asegurar la calidad del producto.

Todo ello nos lleva a concluir que la utilización de la técnica de los machos de gelatina por moldeo es viable para la realización de órganos huecos a bajo coste. No obstante, como trabajo futuro hay que asegurar la viabilidad de está técnica para cada tipo de órgano artificial que se quiera fabricar.

Agradecimientos

Los autores desean agradecer el apoyo de este trabajo de investigación al VI Plan Regional de Investigación y a la Junta de Extremadura y al Fondo Europeo de Desarrollo Regional (FEDER) vinculado a la financiación de los grupos de investigación del Centro de Cirugía Mínimamente Invasiva Jesús Usón (GR18199) y del grupo de Ingeniería de los Procesos de Fabricación (GR18029).

Referencias

[1] 1. Ríos González, E.; Jiménez Gordo, A.M. Cáncer de vejiga. Med. 2017, 12, 1955–1965.

[2] 2. Ganguli, A.; Pagan-Diaz, G.J.; Grant, L.; Cvetkovic, C.; Bramlet, M.; Vozenilek, J.; Kesavadas, T.; Bashir, R. 3D printing for preoperative planning and surgical training: a review. Biomed. Microdevices 2018, 20.

[3] 3. Nisar, H.; Moore, J.; Piazza, R.; Maneas, E.; Chen, E.C.S.; Peters, T.M. A simple, realistic walled phantom for

intravascular and intracardiac applications. Int. J. Comput. Assist. Radiol. Surg. 2020, 15, 1513–1523.

Estudio del efecto combinado de la Oxacina 4 Perclorato y la irradiación en la citotoxicidad de células U87

S. Pampín Suárez [1], F. M. Pérez-Campo [2], J. C. Rodríguez-Rey [2], J. L. Arce-Diego [1], F. Fanjul-Vélez [1]

[1] Grupo de Ingeniería Biomédica, Departamento TEISA, Universidad de Cantabria, Santander, ESPAÑA

[2] Grupo de Ingeniería de Tejidos, Departamento de Biología Molecular, Universidad de Cantabria, Santander, ESPAÑA

pampins@unican.es, f.perezcampo@unican.es, josecarlos.rodriguez@unican.es, arcedj@unican.es, fanjulf@unican.es

Resumen

Uno de los campos de aplicación con mayor potencial de la Ingeniería Biomédica lo constituyen las técnicas que emplean radiación en el rango óptico en medicina. Dentro de ellas, el diagnóstico de imagen por fluorescencia es uno de los abordajes más prometedores, ya que permite llevar a cabo una caracterización diagnóstica en tiempo real, con una interacción mínima con los procesos vitales. Sin embargo, la composición heterogénea de los tejidos y sus características ópticas, limitan el uso de esta tecnología. El aprovechamiento de la ventana de transmisión para tejido, con longitudes de onda comprendidas entre 650 y 950 nm, ofrece nuevas oportunidades, ya que incrementa la capacidad de penetración de la luz en los tejidos y disminuye la absorbancia del fotón, el esparcimiento o scattering y la autofluorescencia. Por este motivo, se está trabajando intensamente en la síntesis de nuevos fluoróforos para la implementación de imagen por fluorescencia en este rango de longitudes de onda y sus futuras aplicaciones en medicina. Uno de estos compuestos de nueva generación es la Oxacina 4 perclorato (O4P). Esta molécula se une de manera específica al tejido nervioso. En este trabajo se ha estudiado su citotoxicidad en células U87 mediante un ensayo colorimétrico. Los resultados muestran que concentraciones elevadas de O4P afectan a la viabilidad celular de manera dosis dependiente, tanto en oscuridad, como tras la irradiación a 630 nm. A concentraciones inferiores a 100 µM, la 04P parece ser inocua, dato importante a tener en cuenta para su uso en futuras aplicaciones médicas.

1. Introducción

La aplicación de las técnicas de la Ingeniería Biomédica a las Ciencias Biológicas, y en particular de aquéllas que emplean radiación óptica, ha permitido el análisis de estructuras subcelulares en 3-D, el estudio en profundidad de las rutas de transducción de señales y los procesos fisiológicos, así como la observación de células enteras e incluso de tejidos y de organismos pluricelulares completos [1], todo ello en tiempo real y con la mínima interacción con los procesos vitales.

En las últimas décadas, varios equipos de investigación de carácter multidisciplinar formados por ingenieros, biólogos, químicos o físicos, entre otros, han acometido el reto de diseñar y desarrollar nuevos sistemas de imagen in vivo, que han podido ser utilizados en diversos campos de la medicina.

La tecnología de imagen por fluorescencia es una de las que más desarrollo ha tenido en biomedicina debido a su alta especificidad, gran sensibilidad, a la posibilidad de obtención de señales múltiples, a la ausencia de radiación ionizante y la posibilidad de ver imágenes en tiempo real interfiriendo lo menos posible con los procesos biológicos [2].

Tradicionalmente, en fluorescencia, se han empleado longitudes de onda situadas en el espectro visible (400–650 nm), tanto para la excitación, como para la emisión fluorescente. Debido a la propia composición de los tejidos biológicos, la adquisición de señales lumínicas a estas longitudes de onda, se ve perjudicada por la dificultad de penetración del haz de luz en los tejidos (< 3 mm), pero también por fenómenos de absorbancia, reflexión, esparcimiento o scattering, o autofluorescencia de algunas moléculas biológicas [3]. El empleo de radiación óptica cercana al infrarrojo, puede solucionar muchos de estos problemas, pero para ello es necesaria la síntesis de fluoróforos cuyos picos de emisión se encuentren en el rango de estas longitudes de onda. Hasta la fecha, sólo dos de estos colorantes fluorescentes han sido aprobados por la US Food And Drug Administration (FDA), el Verde de Indocianina y el Azul de Metileno, los cuales han demostrado su utilidad en diagnóstico clínico y cirugía [4], [5].

La Oxacina 4 Perclorato (O4P) es un fluoróforo de pequeño tamaño con alta especificidad para tejido nervioso [6], que emite fluorescencia a longitudes de onda cercanas al infrarrojo [7]. Por sus características, podría ser un buen candidato para futuras aplicaciones en el ámbito de la medicina [8]. Un aspecto que es preciso tener en cuenta en la aplicación de este y otros fluoróforos, tiene que ver, además de con su especificidad por el tejido de interés, con su posible citotoxicidad en las condiciones de uso clínico. Los ensayos con pequeños animales, en primer lugar, y su traslación a humanos, requieren de estudios previos sobre citotoxicidad, que relacionen al menos las características de concentración y respuesta a radiación de la substancia inoculada. Por esta razón, en este trabajo se analiza la citotoxicidad de la O4P en la línea celular humana U87, que representa un buen ejemplo de tejido nervioso.

2. Materiales y Métodos

2.1. Cultivo celular

La línea de células de glioblastoma U87 de origen humano (ATCC HBT-14) fue cultivada en medio

ISBN: 978-84-09-25491-0, pags. 376- 379

Dulbecco´s modified Eagle´s medium (DMEM), suplementado con suero fetal bovino (FBS) al 10%, 100 U/mL de penicilina, 100 µg/mL de estreptomicina y L-Glutamina al 2 mM, todo de Thermo Fisher Scientific (Waltham, Massachusetts, USA), y fueron mantenidas en condiciones ambientales estándar para su crecimiento, a 37°C y con el 5% de CO_2.

Se escogió este modelo celular porque, además de su origen humano, la O4P marca específicamente células de tejido nervioso productoras de mielina, como es el caso de las células U87.

2.2. Preparación del fluoróforo

La O4P fue proporcionada por BOC Science (Shirley, NY, USA). Se preparó una solución stock a una concentración 15 mM en tampón fosfato salino (PBS), pH 7,4. La O4P posee un pico de absorbancia a 616 nm y un pico de emisión a 635 nm en este solvente [9]. La Figura 1 muestra la estructura molecular del compuesto.



Figura 1. *Estructura molecular de la Oxacina 4 Perclorato.*

2.3. Sistema de iluminación

El sistema de iluminación LED, cuyo esquema aparece en la Figura 2, fue diseñado para irradiar placas de Petri de 60 mm de diámetro con hasta 6 tipos de radiación. El sistema está compuesto por una fuente LED blanca de alta potencia con una anchura espectral de 200 nm. El sistema de control de la fuente permite la modulación de la misma, de tal forma que es posible controlar la irradiancia de emisión. Mediante una rueda de filtros se configuran los 6 modos de irradiación, de los que en este trabajo se emplea el modo de irradiación con longitud de onda central 630 nm. La radiación resultante se uniformiza mediante un sistema óptico para lograr una exposición homogénea a lo largo de la superficie de la placa de Petri.

Las células se iluminaron desde la parte inferior, durante 15 minutos a 10 mW/cm^2 y a una λ = 630 nm. Las células U87 crecen adheridas al fondo de los pocillos de fondo plano, por lo que se asumió que el espesor de la placa no interfiere a efectos de la irradiación. La dosis total calculada está en torno a los 9 J/cm^2.

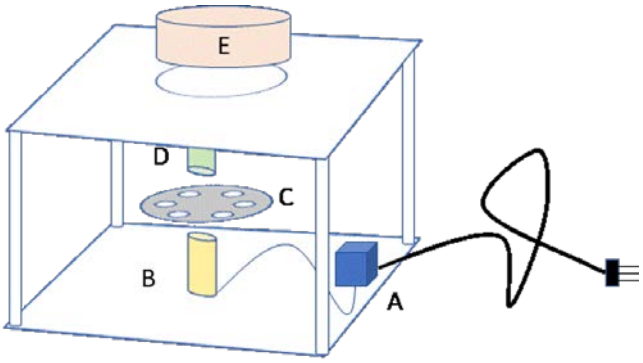


Figura 2. *Esquema del diseño del sistema de iluminación de las células. Este dispositivo está compuesto por: A) Un mecanismo de control de la intensidad de la fuente. B) Una fuente LED. C) Una rueda de filtros. D) Un sistema óptico de colimación y homogeneización. En el momento de la irradiación, la placa de Petri se coloca en la abertura creada para ello (E).*

2.4. Ensayo de citotoxicidad

La viabilidad de las células U87 fue evaluada in vitro con el método CellTiter 96® AQueous One Solution Cell Proliferation Assay de PROMEGA (Madison, Wisconsin, USA), basado en el ensayo colorimétrico MTT. Las células fueron cultivadas en placas p60 de CORNING (Corning, NY, USA), con 2 mL de medio. En el momento de realizar el experimento, se lavaron 2 veces con PBS y se les cambió el medio. Se añadieron concentraciones crecientes de O4P (50, 100, 200 y 400 µM), utilizando una solución stock a 15 mM en PBS. Tras 4h de incubación, las células se irradiaron desde abajo por medio de un dispositivo de iluminación controlada creado ad hoc en nuestro laboratorio**.** Fueron expuestas a 10 mW/cm^2, durante 15 min a una longitud de onda central de λ = 630 nm. En cada experimento se utilizó como control positivo una placa en la que no se ha añadido fluoróforo y un control negativo de células que no se han sometido a la acción de la luz. Después de la irradiación, las placas se lavaron dos veces con PBS y se tripsinizaron con Tripsina-EDTA de Thermo Fisher Scientific (Waltham, Massachusetts, USA). Se inocularon 10^4 células/100 µL en una placa de 96 pocillos (CORNING) y se incubaron a 37ºC durante 24h. Para cuantificar el número de células viables se añadieron 20 µL de reactivo CellTiter 96® AQueous One Solution Reagent siguiendo las instrucciones del fabricante, y se incubaron durante otras 4h en las mismas condiciones, pasadas las cuales se midió la absorbancia a 490 nm.

Los experimentos se repitieron 4 veces de manera independiente. En cada uno de ellos, se han realizado 6 replicados por muestra.

2.5. Análisis estadístico

Para evaluar la respuesta de las células U87 a las distintas dosis de O4P en presencia y ausencia de radiación, se llevó a cabo un análisis ANOVA de dos variantes con varias muestras por grupo, utilizando el paquete Excel de Microsoft Office. Se consideraron estadísticamente significativos valores con $p < 0,05$.

3. Resultados y discusión

En este estudio se ha analizado el efecto citotóxico de la O4P con el fin de establecer la dosis de compuesto óptima para la realización de ensayos biológicos, y descartar la posible inducción de hormesis, fenómeno de respuesta adaptativa por el que, a distintas dosis, una sustancia se comporta de manera contraria según los estímulos que recibe [10], [11], en este caso la radiación con una fuente de luz.

La Viabilidad celular se calculó según la siguiente ecuación: Viabilidad celular (%) = (A muestra/ A control) x 100.

Como se puede ver en la Figura 3, el comportamiento de las células a las dosis más bajas de O4P testadas (50 y 100 µM), tanto en ausencia como en presencia de luz, es similar, y no se observan diferencias significativas en la proliferación con respecto a las células control. Sin embargo, en oscuridad, las concentraciones más altas de O4P (200 y 400 µM), provocan descensos significativos en la viabilidad celular, del 28,6% y 71,4 %, respectivamente, mientras que a dosis menores no se aprecian efectos significativos. Es decir, el efecto citóxico de la O4P observado, es dosis dependiente. Este efecto letal parece independiente de la irradiación utilizada, ya que las células que no han sido irradiadas presentan la misma tendencia que las tratadas durante 15 min con luz a una λ=630 nm, longitud de onda a la que absorbe el fluoróforo de forma máxima. En este caso, también se produce una disminución significativa de la supervivencia celular con las dosis más elevadas, del 36% y 74,5% para 200 y 400 µM de O4P respectivamente, porcentajes muy similares a los observados en ausencia de luz, por lo que podemos descartar que este compuesto produzca respuesta hormética por el efecto combinado de la concentración y la irradiación a 630 nm.

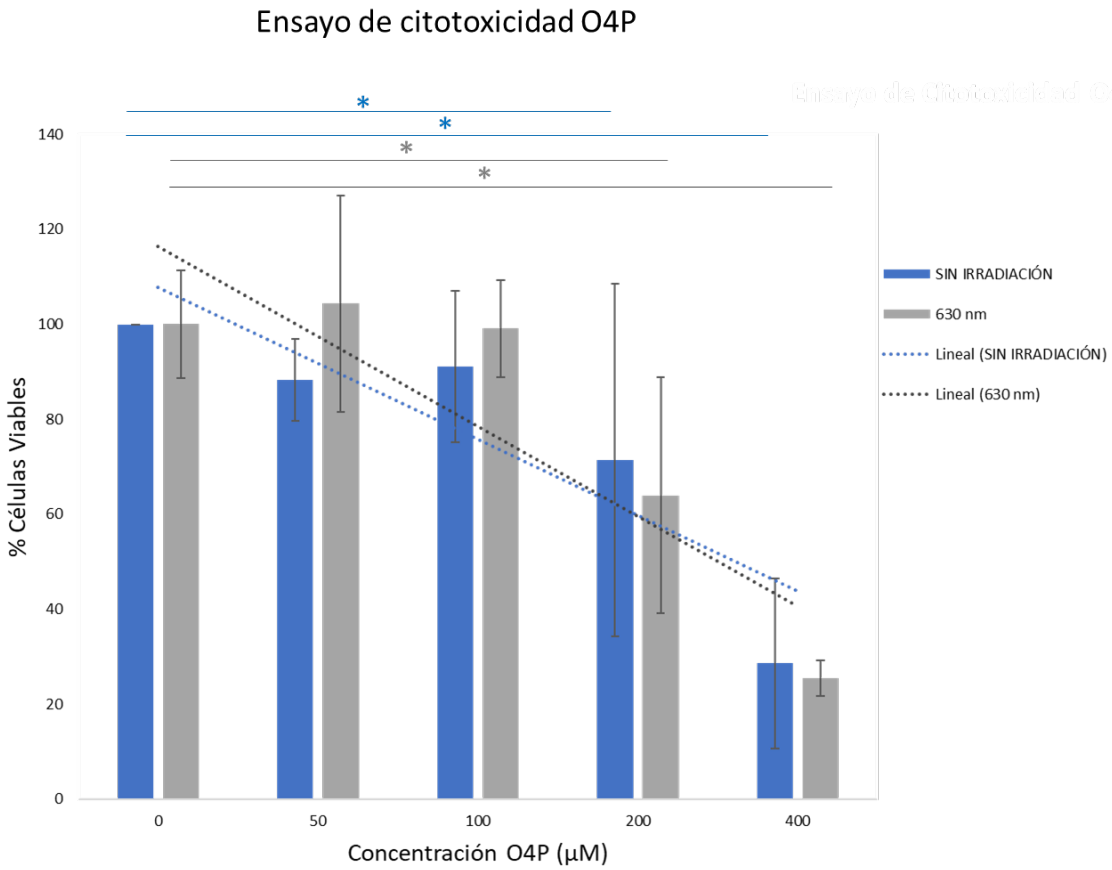


Figura 3. *Ensayo de citotoxicidad. Porcentaje de viabilidad de las células U87 en oscuridad y tras 15 min de irradiación a una dosis de 10 mW.cm^{-2}, y una ƛ=630 nm. Las células se han tratado a diferentes concentraciones de O4P (0, 50, 100, 200 y 400 µM), durante 4h. Estos datos representan la viabilidad celular relativa respecto al control negativo en porcentaje, de 4 experimentos independientes. Los resultados se analizaron usando un test ANOVA de dos factores con varias muestras por grupo. La significación estadística se indica con un asterisco *p<0,05.*

4. Conclusiones

Tras la realización de este estudio podemos concluir que la O4P presenta citotoxidad en células U87, si bien este efecto es dependiente de la dosis, siendo observada a concentraciones superiores a 100 μM. Además, en las condiciones del ensayo, no se detecta acción combinada entre concentración de O4P tras la irradiación durante 15 min a λ=630 nm y 10 mW/cm^2, lo que convierte a este fluoróforo en un buen candidato para posibles aplicaciones futuras en el campo de la medicina.

Agradecimientos

Este Trabajo ha sido parcialmente financiado por el proyecto del Plan Nacional de I+D+i "Modificación mediante alta presión de las propiedades plasmónicas y luminiscentes de nanopartículas de metales y óxidos desnudas y recubiertas" (PGC2018-101464-B-I00), del Ministerio de Ciencia, Innovación y Universidades, cofinanciado con fondos FEDER y por la Fundación San Cándido.

Bibliografía

[1] Weissleder R, Pittet MJ. "Imaging in the era of molecular oncology", 7187, 2008, Nature, Vol. 452, págs. 580-589.

[2] Hong g, Antaris A, Dai H.Near-Infrared fluorophores for biomedical imaging. 0010, 2017, Nature Biomedical Engineering.

[3] Li C, Chen G, Zhang Y, et al. "Advanced Fluorescence Imaging Technology in the Near-Infrared II Window for Biomedical Aplications", 2020, Journal of the American Chemical Society, Vol. 142, págs. 14789-14804.

[4] Troyan SL, Kianzad V, Gibbs-Strauss SL et al. "The FLARE intraoperative near-infrared fluorescence imaging system: a first-in-human clinical trial in breast cancer sentinel lymoh node mapping", 10, 2009, Annals of surgical oncology, Vol. 16, págs. 2943-2952.

[5] Tummers QR, Schepers A, Hamming JF et al. "Intraoperative guidance in parathyroid surgery using near-infrared fluorescence imaging and low-dose Methylene Blue", 5, 2015, Surgery, Vol. 158, págs. 1323-1330.

[6] "Visualizing Oxazine 4 nerve-specific fluorescence ex vivo in frozen tissue sections", 2016, Proceedings of SPIE - The International Society for Optical Engineering, Vol. 9696.

[7] Barth CW, Gibbs SL.Park MH, Hyun H, Ashitate Y et al. "Prototype nerve-specific near-infrared fluorophores", 8, 2014, Theranostics, Vol. 4, págs. 823-833.

[8] F. Fanjul-Vélez, A. M. Díaz-Martínez, E. Garro-Martínez, J. L. Arce-Diego, "Fluorescence imaging contrast in guided surgery on nerves measured in rats in vivo", Proceedings of SPIE - The International Society for Optical Engineering, Vol. 11222, 2020, 112220K.

[9] Park MH et al., "Prototype nerve-specific near-infrared fluorophores", Theranostics, Vol. 4, 2014, págs. 823-833.

[10] EJ, Calabrese, "Paradigm found: The re-emergence of hormesis as a fundamental dose-response model in the toxicological sciences", 2005, Environmental Pollution, Vol. 138, págs. 379-412.

[11] EJ, Cook RC and Calabrese. "The importance of hormesis to public health", 2006, Environment Health Perspective, Vol. 114, págs. 1631-1635.

Simulador de Examen Clínico Objetivo Estructurado para la evaluación de la cadena de supervivencia basado en realidad virtual

P. Sánchez-González [1,2], A. Marín[1], M. Rodríguez [1], I. Oropesa [1], E.J. Gómez [1,2]

[1] Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica, de Madrid, Madrid, España; {p.sanchez, manuel.rodriguez, i.oropesa, enriquejavier.gomez }@ upm.es

[2] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Madrid, España

Resumen

El examen clínico objetivo estructurado, ECOE, es una herramienta de evaluación multipropósito empleada para evaluar a los profesionales de la salud en un entorno clínico. Evalúa la competencia, fundamentada en pruebas objetivas a través de la observación directa en diferentes estaciones. Estas estaciones deben ser capaces de reflejar la realidad de una manera fidedigna, lo que en ocasiones se dificulta por los entornos en los que se llevan a cabo. La realidad virtual puede ser una herramienta útil para dotar de realismo e inmersión a los estudiantes. Este trabajo desarrolla un entorno de realidad virtual para la valoración de la cadena de supervivencia capaz de llevar a cabo la evaluación objetiva de la tarea. El entorno se centra en un andén de metro donde uno de los protagonistas sufre una parada cardiaca súbita y el alumno ha de ser capaz de recuperar las funciones básicas del paciente. Los resultados obtenidos animan a continuar con una validación más exhaustiva para su integración dentro de los ECOE que actualmente se llevan a cabo.

1. Introducción

El Examen Clínico Objetivo Estructurado (ECOE) es una herramienta de evaluación de la competencia clínica del estudiante de medicina que tradicionalmente se lleva a cabo mediante la observación directa de distintas situaciones clínicas [1]. El ECOE establece una serie de competencias y habilidades a evaluar a través de distintos escenarios de simulación clínica, denominados estaciones [2].

Entre las distintas estaciones que componen un ECOE, se encuentra la dedicada a la evaluación de las competencias de exploración, manejo y habilidades técnicas en la cadena de supervivencia. En esta estación el alumno debe realizar una serie de acciones en un orden secuencial determinado para que la probabilidad de supervivencia de la víctima sea la mayor posible [3]. La reanimación cardiopulmonar (RCP) constituye un paso fundamental en el desarrollo de esta cadena y se define como el proceso de sustitución de las funciones vitales de respiración y circulación espontáneas para una posterior restauración. La técnica empleada en la RCP es fundamental para el éxito de la maniobra [3].

La RCP ha sido tradicionalmente enseñada y entrenada mediante el uso de maniquís mecánicos [4]. Este método tiene limitaciones en cuanto a fidelidad ya que, a pesar de ofrecer una sensación háptica semejante, no aporta un entorno inmersivo y dinámico. Existen enfoques más novedosos que han señalado el potencial de la Realidad Virtual (RV) como una forma de obtener buenos resultados en el entrenamiento de habilidades técnicas [5] ya que los sistemas de RV son flexibles en usabilidad, simples y fáciles de configurar y están diseñados y probados para un uso fácil y seguro [6].

En este trabajo se presenta un simulador de RV como estación para la evaluación de la cadena de supervivencia del ECOE. En concreto se ha desarrollado una herramienta de evaluación empleada en la estación "RCP en vía pública", que forma parte del ECOE realizado en el sexto curso de Medicina de la Universidad Autónoma de Madrid.

2. Diseño del simulador ECOE en RV

El diseño se lleva a cabo siguiendo el modelo de estación de "RCP en vía pública" del ECOE de la Universidad Autónoma de Madrid del año 2018 [7] y las pautas del Consejo Europeo de Resucitación (CER) para el Soporte Vital básico en adultos y la desfibrilación automática externa [8]. Este diseño selecciona un área pública con desfibriladores como el tema del entorno, así como la creación de todos los gráficos 3D necesarios para la realización de eventos e interacción con objetos y personajes dentro de la simulación.

2.1. Escenario

Tradicionalmente, los examinadores entregan al alumno una ficha con la denominada "situación de partida". En ella se le proporciona una descripción del escenario en el que se sitúa la prueba: dónde ha ocurrido la emergencia, edad del paciente, etc. Al realizar la prueba mediante RV, esta descripción ha de quedar implícita y el alumno deberá de ser capaz de obtener esta información prestando atención al entorno e interactuando con los objetos del escenario como si de una situación real se tratara.

El escenario escogido es una estación de tren modelada tomando como base el Metro de Madrid. Se encuentra dividido en dos situaciones, también denominadas escenas:

- *Escena 1: Aprendizaje y familiarización con el entorno*: Esta escena proporciona instrucciones al alumno para que practique los movimientos, e interaccione con los objetos del entorno, la

ISBN: 978-84-09-25491-0, pags. 380- 383

interfaz de usuario y las funciones de la RV. El alumno es guiado mediante mensajes auditivos y visuales.

- *Escena 2: Prueba de Evaluación*: El estudiante está inmerso en un escenario donde sucede la emergencia. Puede moverse e interactuar de manera libre con el entorno y los personajes. Mientras tanto, el sistema monitoriza de forma continua sus acciones y movimientos. En esta escena, el usuario se encuentra en una estación donde espera al tren. En el andén hay más personajes controlados por la inteligencia artificial del sistema. Después de un tiempo, uno de ellos cae por parada cardiaca súbita (PCS). El usuario debe llamar al servicio de emergencias, usar el Desfibrilador Externo Automático (DEA) y salvar al paciente en menos de 10 minutos.

El sistema cuenta con un dispositivo háptico que simula la resistencia del pecho del paciente: un maniquí de reanimación cuya posición en el espacio se corresponde con la posición del torso del paciente virtual.

2.2. Evaluación de la realización

El estudiante debe prestar atención no solo a la precisión de las maniobras de RCP, sino también a los objetos y personajes presentes en la escena y a cómo reaccionan con el entorno para completar la tarea por completo. Para la evaluación del ejercicio, el simulador recoge un conjunto de métricas basadas en la secuencia de acciones relacionadas con identificar la emergencia y el estado del paciente, y dar un diagnóstico a los servicios de emergencia (Tabla 1).

Métrica	Valor
Comprobación del estado del paciente	Sí / No
Aviso a emergencias	Sí / No
Desviste al paciente	Sí / No
Tiempo de reacción	Min:Sg
Usa DEA	Sí / No
Intentos de colocación del electrodo derecho	Número
Intentos de colocación del electrodo izquierdo	Número
Tiempo de configuración del DEA	Min:Sg
Realización de RCP	Sí / No
Localización de las compresiones	T. acierto (%)
Inclinación de las compresiones	T. acierto (%)
Frecuencia de las compresiones	T. acierto (%)
Frecuencia instantánea	Compresiones / minuto
Frecuencia media	Compresiones / minuto
Tiempo de pausa	Min:Seg
Tiempo de pausa medio	Min:Seg
Profundidad de las compresiones	T. acierto (%)
Profundidad media	Cm

Tabla 1: *Relación de métricas incorporadas en el simulador.*

Algunas métricas se guardan únicamente si el usuario utiliza el DEA. En caso contrario, se descartan y aparecen en la ficha final con un signo "-". Cuando el DEA está encendido y cerca del paciente, aparecen sobre su pecho los dos electrodos. El alumno tendrá que colocarlos en la posición correcta. Para ello ha de cogerlos y soltarlos en una zona cercana a la posición correcta. Si se hace de forma correcta, los electrodos se quedan pegados al paciente y el DEA comenzará a analizar su pulso. Si no, vuelven a su posición inicial. El entorno contabiliza el número de veces que el alumno ha intentado colocar cada electrodo. Además, se mide el tiempo transcurrido desde que el alumno interacciona con el botón de la vitrina para coger el DEA, hasta que éste inicia su funcionamiento.

Las colisiones de las manos del usuario con el pecho del paciente permiten analizar una serie de métricas respecto a su ejecución en la RCP. Así, además de la frecuencia y compresión instantáneas se monitorizan la localización e inclinación de las compresiones, la profundidad y frecuencia de estas y el tiempo de pausa entre ellas. Para que sean efectivas, la profundidad de las compresiones torácicas se debe situar entre 5 y 6 cm y la frecuencia entre 100 y 120 compresiones/minuto, según los protocolos establecidos por el CER [8].

2.3. Arquitectura

La arquitectura del simulador desarrollado (Figura 1) está basada en capas sobre el motor de desarrollo de aplicaciones Unity Engine [9], donde se provee una abstracción sobre el emplazamiento de objetos, las físicas, el renderizado y el audio en una escena tridimensional. En términos de secuenciación de eventos y comunicación entre los componentes de la aplicación se han desarrollado *scripts* que actúan de controlador de los diferentes datos. Finalmente, los datos procesados se envían al *driver* de SteamVR que permite la muestra de los gráficos y la reproducción del audio en un casco de HTC Vive.

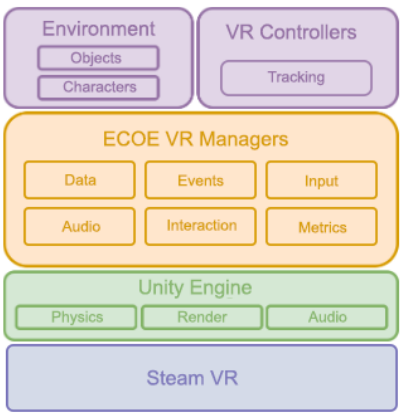


Figura 1. *Arquitectura del simulador ECOE VR*

3. Implementación

Los modelos 3D del escenario se han obtenido a través de la tienda de Assets de Unity y se han dispuesto sobre el propio editor de Unity Engine. Todos los personajes incorporados se han animado mediante la web Mixamo [10], que permite hacer *Auto-Rigging* y animado de modelos 3D bípedos, siguiendo una estructura de modelo humanoide.

El usuario utiliza un avatar 3D dentro del simulador. De este avatar se renderizan las manos en la misma posición relativa a las muñecas del alumno gracias a los *trackers* de HTC Vive. Esta información se emplea también para realizar los cálculos de profundidad y ángulo en las compresiones, así como la interacción con el entorno mediante movimientos corporales. Por otro lado, la posición en el espacio del propio avatar se sigue mediante el seguimiento de la cabeza del usuario mediante el casco.

Los objetos y personajes interactivos se cargan en las dos escenas. La primera escena, al ser de adaptación, es un recinto cerrado de 7x7 metros, donde el usuario tiene tanto libre movimiento como un teletransporte (Figura 2). La escena de adaptación está compuesta por los tipos de objetos que el usuario puede encontrar en la escena de evaluación: personajes interactivos, objetos que puede agarrar y lanzar, objetos con los que puede interactuar para que sucedan eventos, y un cuerpo con físicas que reacciona al movimiento y la presión con los controladores de RV.

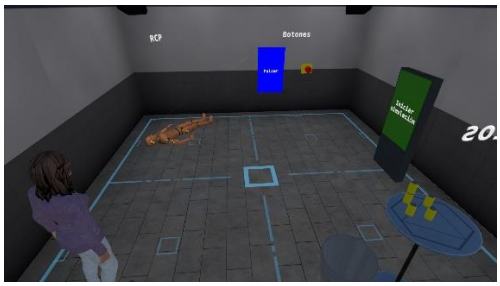


Figura 2.*Escena de adaptación*

Por otro lado, la escena de evaluación cuenta con la estación de metro y los personajes propios de la simulación en sí. Los personajes no solo están para dar una mayor sensación de realismo con su presencia, sino que además disponen de una inteligencia artificial basada en árboles de comportamiento que cambian en función de cómo se comporte el usuario dentro de unas reglas. Por ejemplo, cuando el paciente cae en el andén, los personajes se sorprenden, pero al ir el usuario a reanimar al paciente, algunos tratarán de ofrecer ayuda. Las paredes y los objetos no interactivos son estáticos y no podrán moverse, aunque otros, podrán ser agarrados con las propias manos virtuales del usuario, como por ejemplo el desfibrilador, que deberá extraer después de abrir manualmente la caja, o el teléfono móvil que le ofrece uno de los personajes. (Figura 3).

La validación preliminar del prototipo desarrollado se ha llevado a cabo mediante entrevistas con dos expertos en formación médica del Hospital Universitario La Paz. Estas entrevistas se centraron en la usabilidad del prototipo desarrollado y su potencial uso como estación ECOE.

4. Discusión

El simulador implementado proporciona un escenario con una situación de emergencia en la que el alumno puede demostrar sus conocimientos y habilidades respecto a la ejecución del soporte vital básico (SVB). El simulador monitoriza la ejecución de la tarea en todo momento y proporciona una evaluación objetiva al finalizar, que refleja si la ejecución del alumno ha sido correcta según las indicaciones del CER. La evaluación cubre todas las etapas de la cadena de supervivencia: identificación y aviso a emergencias, uso del desfibrilador y, de manera más exhaustiva, maniobras de RCP, proporcionando más métricas que las que tradicionalmente ofrecen los dispositivos de retroalimentación empleados en la actualidad, siendo éstas únicamente la frecuencia y la compresión instantánea.

.

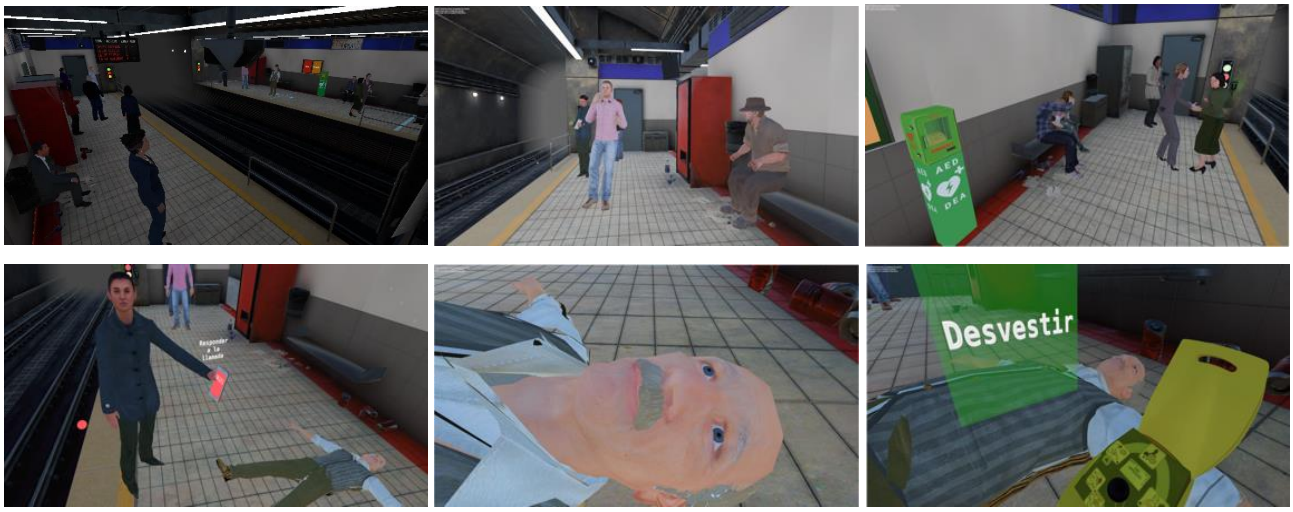


Figura 3.*Escena de evaluación: (Arriba) Vistas de los andenes al comienzo de la simulación; (Abajo) Ejemplos intervención*

El escenario de emergencia que proporciona la aplicación se desarrolla en un lugar público (estación de metro), en el cual es posible que una persona sufra una PCS. El entorno cuenta con una primera fase de aprendizaje para que el alumno aprenda los controles y gestos que debe usar en prueba sin desvelarle información sobre la situación de emergencia a la que se enfrenta. La realización de estas tareas de adaptación previa reduce la curva de aprendizaje del alumno respecto al uso de la simulación. De esta manera la evaluación del alumno no se ve condicionada por su nivel de experiencia con la RV. Además, esta primera escena permite evaluar la experiencia de usuario respecto a la RV y detectar posibles inconvenientes tales como mareos o desorientaciones.

Una exhaustiva validación del sistema puede determinar con eficacia la usabilidad y eficacia real del simulador. No obstante, los resultados obtenidos y las primeras evaluaciones llevadas a cabo con personal clínico mediante entrevistas arrojan luz sobre los prometedores resultados que presenta el uso de la RV en las estaciones ECOE, no solo como parte de la evaluación, sino también como herramienta de entrenamiento.

5. Conclusión

Este trabajo ha presentado un simulador virtual para la evaluación de la cadena de supervivencia en el que el alumno ha de demostrar sus habilidades y conocimientos en la identificación de la emergencia, el aviso a emergencias, el uso del desfibrilador y la realización de las maniobras de RCP.

Todas las mejoras están supeditadas a los resultados que se obtengan de la validación del simulador, y de ello dependerá su integración dentro del ECOE realizado en el sexto curso de Medicina de la Universidad Autónoma de Madrid.

Referencias

[1] M. Zayyan, Objective structured clinical examination: The assessment of choice, Oman Medical Journal, 2011 vol. 26, no. 4. pp. 219–222.

[2] R. M. Harden, P. Lilley, and M. Patricio, The definitive guide to the OSCE: the Objective Structured Clinical Examination as a performance assessment. Elsevier Health Sciences, 2015

[3] R. W. Neumar et al., American Heart Association guidelines update for CPR and emergency cardiovascular care, 2015, vol. 132, no. 18. Lippincott Williams and Wilkins.

[4] C. loconsole, A. Frisoli, N. Mastronicola, An experimental comparison with a standard CPR training mannequin, 2015, vol 93, E1-E2.

[5] G S Ruthenbeck, K J Reynolds. Virtual reality for medical training: the state-of-the-art. Journal of Simulation 2015, Vol 9, pp 16-26.

[6] J Pottle. Virtual reality and the transformation of medical education. Future Healthcare Journal. 2019. Vol 6. no 3. pp 181-185.

[7] M. Quintana, N. Cancellere, E. de Vega. Preparación de estación con maniquí: RCP en la vía pública, Madrid, 2018.

[8] G. D. Perkins et al., European Resuscitation Council Guidelines for Resuscitation 2015. Section 2. Adult basic life support and automated external defibrillation., Resuscitation, Oct. 2015. vol. 95, pp. 81–99

[9] Página de documentación de Unity Engine. https://unity.com/es (Consultada: septiembre 2020)

[10] Página de documentación de Mixamo. https://www.mixamo.com/#/ (Consultada: septiembre 2020)

Real-Time Workflow Detection using Video Streams in Craniosynostosis Surgery

L. García-Duarte Sáenz[1], D. García-Mato[1,2], S. Ochandiano[2,3], J. Pascau[1,2]

[1] Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid, España, {dgmato,jpascau}@ing.uc3m.es

[2] Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España

[3] Servicio de Cirugía Oral y Maxilofacial, Hospital General Universitario Gregorio Marañón, Madrid, España

Abstract

Real-time automatic surgical tool detection is a demanding task in deep learning applications aiming to recognize and optimize the surgical workflow. This work focuses on applying such technology in a navigation-based craniosynostosis surgical procedure, which presents an elevated complexity and surgical outcomes are highly dependent on the surgeon's expertise. A total of four neural networks were trained for automatic tool detection and phase estimation. Three state-of-the-art networks (VGG-16, MobileNetV2 and InceptionV3) were trained with transfer learning, and a convolutional neural network, named CranioNet, was developed. These four networks were tested for the automatic classification of tools and surgical phase estimation from video streams recorded in two different scenarios: a surgical simulation performed on a 3D printed phantom, and a real surgical procedure for the correction of craniosynostosis. Classification models were integrated into a fully customized application software, developed to perform the tool-recognition task in real-time and to provide information to estimate the surgical phase duration. Results of this study demonstrate that deep learning can accurately classify surgical tools and estimate surgical phases in video streams of craniosynostosis surgeries. VGG-16 and CranioNet networks showed optimal performance for this task. The integration of this technology into the clinical practice could provide objective feedback and performance evaluation of surgical skills and could lead to a reduction of surgical errors and complications in craniosynostosis surgery.

1. Introduction

Craniosynostosis is a congenital defect that implies the partial or total premature fusion of one or more cranial vault sutures that separate the skull bones [1]. This condition causes cranial deformities which may limit the infant's brain growth, leading to many functional and head shape alterations. Surgical correction is typically carried out to correct the child's cranium shape and relieve intracranial hypertension. This surgery is based on open cranial vault remodeling, which consists of three steps: (1) removal of the affected bone tissue, (2) remodeling of the resected bone, and (3) placement and fixation of the remodeled bone fragments [2]. Nowadays, the surgical procedure relies mostly on the subjective judgement of the surgeon, who must determine the best approach to restore normal cranial shape based on their knowledge and practice. Therefore, this complex surgery is highly dependent on experience and is more prone to error in the less experienced surgeons [3].

Novel technological advancements, such as 3D printed patient-specific tools or intraoperative navigation systems, have recently been integrated into these surgical workflows to ensure optimal outcomes [2], at the cost of an increased complexity. Hence, new methods must be developed to evaluate surgical performance and to reduce surgical errors and complications.

Automatic surgical tool detection has been a challenging task in medical image processing research. Efforts have been made to develop algorithms that are both accurate and sufficiently fast for optimal deployment. Recognizing the surgical workflow is key to support surgeons with feedback and real-time warnings, enhancing patient's safety throughout the surgical intervention. Additionally, automatic tool recognition has potential uses in simulation-based training of medical students to provide quantitative performance metrics and ensure the proper acquisition of required surgical skills. Other important utilities of such technology include post-operative clinician self-assessment, which is useful to review the process and easily detect errors or correct performances of the different surgical steps. Also, statistical information related to performance and duration can be generated and analyzed to optimize the surgical workflow [4].

Many algorithms have been proposed for real-time object recognition in medical applications. Deep learning-based approaches are preferred for this purpose as they have proven to be very effective when performing the tool recognition task during different types of surgeries [5]. State-of-the-art networks include VGG, Inception and MobileNet architectures, which have demonstrated high-quality performance in this area. However, optimal detection is difficult to achieve due to motion, presence of occlusions, illumination changes, or different camera angles. Therefore, image variability during training must be maximized to overcome these issues.

In this work, we present a method for surgical workflow estimation based on deep learning for phase identification, post-operative analysis and medical training in craniosynostosis reconstruction surgeries. The proposed technique was evaluated on a phantom-based simulation and tested during real surgery. To our knowledge, this is the first study to apply automatic tool detection in the context of craniosynostosis surgery.

ISBN: 978-84-09-25491-0, pags. 384- 387

2. Materials and methods

2.1. Data acquisition

(A) Phantom simulation

A 3D printed patient-specific phantom was designed and manufactured for the simulation of craniosynostosis reconstruction surgery. This phantom is based on data from patient presenting metopic craniosynostosis and suffering from severe trigonocephaly malformation. The phantom was designed from the anatomical 3D models of the patient obtained from the preoperative computed tomography (CT) scan and manufactured using a desktop 3D printer (Ultimaker B.V., NL). Polylactic acid (PLA) material was used to simulate bone tissue and silicone material (modeled using PLA-based moulds) was used to simulate soft tissue: brain, dura mater and skin. In addition, several tools were designed and 3D printed in PLA material to simulate all main phases of the surgical procedure: a piezoelectric tool for trepanation of bone tissue, a scalpel to simulate incision in the skin tissue, and two surgical cutting guides to aid during osteotomy. Also, we used a stylus tool with optical markers that is tracked during intraoperative navigation (**Figure 1**).

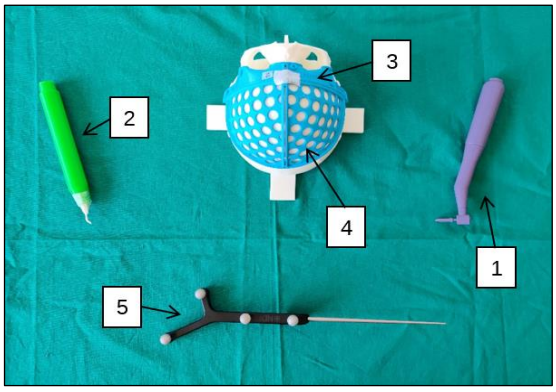


Figure 1*. 3D printed phantom and tools used to simulate the craniosynostosis surgical procedure: piezoelectric tool (1), scalpel (2), supraorbital bar guide (3) and frontal bone guide (4) and tracked stylus tool (5).*

The surgical simulation included the main phases of navigation-based open cranial vault remodeling: (1) *preparation* of the surgery, where no tools are present in the surgical field of view (FOV); (2) *skin incision*, using the scalpel tool; (3) *cranial surface exposure*, where no tools are present in the surgical FOV; (4) *supraorbital bar cutting guide placemen*t over the supraorbital region of the skull; (5) *frontal cutting guide placement* over the frontal region of the cranial vault; (6) *registration*, where the stylus tracked tool is used to record landmarks on the surgical guides for intraoperative registration; (7) *osteotomy*, where the piezoelectric tool is used to cut the bone; (8) *remodeling*, where the remodeled bone fragments are fixed to the patient's anatomy; and (9) *intraoperative navigation*, where the stylus tool is used to correct the positions of the remodeled fragments to match the preoperative virtual plan.

We used a Logitech C905 camera to record video streams during the simulation with a resolution of 800x600 pixels and an acquisition rate of 15fps. The camera was positioned on top of the simulated surgical field to enable a complete visualization of the phantom and its surroundings from a superior view. Shadows were created during video recordings to increase image complexity. In addition, video recordings of individual tools were collected for training, as well as a free-of-tool case.

(B) Surgical deployment

Video images were recorded at Hospital General Universitario Gregorio Marañón (HGUGM) during a real surgical procedure on a 10-month-old infant presenting metopic craniosynostosis. The parent or legal guardian of the patient signed an informed consent for study participation, providing necessary permission for the recording of images. The study was performed following the principles of the 1964 Declaration of Helsinki as revised in 2013 and approved by the Research Ethics Committee at HGUGM.

A camera, GoPro Hero7, was used to record the videos with a resolution of 1920 x 1080 pixels and a frame rate of 60 fps. The camera was placed outside the surgical field to avoid interference with the surgeon's tasks and was controlled through a mobile phone application to prevent disturbances. Due to illumination issues, occlusions, and non-ideal FOV, only three surgical phases were correctly recorded for training and testing: (1) *cranial surface exposure*, (2) *remodeling* and (3) *intraoperative navigation*.

2.2. Training of the networks

Frames were extracted from the videos and labeled for training. A total of 6692 images were acquired during phantom simulation, from which a 67% (4473 images) was used for training, a 17% (1118 images) for validation and a 16% (1101 images) for testing. On the other hand, a total of 600 images were selected from the real surgery. 70% (420 images) of frames were used as a training set, whereas 15% (90 images) were marked as the validation set and the remaining 15% (90 images) as the testing set.

A total of four neural networks were trained for automatic tool detection and phase estimation. Three state-of-the-art networks (VGG-16, MobileNetV2 and InceptionV3) were trained with transfer learning. In addition, a customized convolutional neural network, named CranioNet, was developed for classification following a simple architecture. It consists of a sequence of blocks made up of two 3x3 convolutions with a stride of 1 followed by 2x2 max pooling operations with a stride of 2. No padding was used except for the convolutions of the first block (P = 1). Next, a series of three fully-connected layers is performed. Dropout is introduced after the first two dense layers to control overfitting, with probabilities 0.4 and 0.1, respectively. There is a reLU activation function after each layer. A weighted L1 regularization was introduced. The output layer contains a Softmax classifier to compute the probability of each class.

The networks were trained using a Lenovo G50-70 computer. In all cases, input images were resized to 224x224 spatially and data augmentation was performed to prevent overfitting by increasing image variability.

RMSProp optimization was used to minimize the cross-entropy loss function (learning rate = 0.001, decay constant = 0.9, and $\epsilon=10^{-7}$). In addition, Xavier initialization was used for the weights and zero initialization for the biases.

2.3. Software application

A customized software application, named *Tool Detector*, was developed to perform real-time tool classification. The application enables the users to select the desired model to be used for classification and to automatically detect tools from live or recorded video streams. It also allows to select a minimum probability threshold for prediction and to choose whether to display only the tool with maximum probability or all tools with a probability higher than the selected threshold. This software was integrated as a module in the open-source platform 3D Slicer. Live or recorded videos can be streamed into the software application using the OpenIGTLink communication protocol. Classification data can be exported into CSV files for further analysis.

2.4. Phase estimation

A specific algorithm was developed to estimate surgical phases from the tool detection probabilities provided by the software application. Each surgical phase is defined by selecting those surgical tools that are normally present in the surgical field during that phase (right-phase tools) and those tools which are not usually present (wrong-phase tools). Then, the algorithm generates a binary signal from the data where 1s correspond to the detection of right-phase tools, and 0s to wrong-phase tools. Next, a moving average filter is applied to reduce high frequencies in the binary signal. Finally, thresholding is performed to estimate the time intervals for each surgical phase, and the final boundaries between phases are computed by taking the average between the end-point of a phase and the start-point of the following.

2.5. Performance evaluation

Trained networks were evaluated using the test dataset and computing different metrics to assess their performance: precision, recall, f1-score, top-1 accuracy (conventional accuracy) and top-2 accuracy (true class is included in the two most probable classes). Regarding the surgical phase estimation, the classification accuracy (CA) of each predicted phase was obtained for performance evaluation.

3. Results

3.1. Phantom simulation

The evaluation metrics calculated for the different networks can be observed in **Table 1**. The highest precision values for tool classification are achieved by VGG-16 (86.24%) and MobileNetV2 (83.56%). However, the training evolution plots of MobileNetV2 show overfitting (data not shown). Although the model adjusts well to this specific type of data, it cannot be generalized, meaning that it will perform worse with more diverse data (e.g. images with a different background or illumination conditions). In the case of VGG-16, no overfitting is observed, and the evaluation metrics indicate optimal performance for tool classification. Results show that this network is able to determine the correct class of the 94.01% of the images, considering the two most probable classes (top-2 accuracy). Additionally, the developed model, CranioNet, performs similarly as VGG-16 since its architecture is resemblant, but with a reduced number of layers.

Model	P	R	F1-S	Top-1 A	Top-2 A
VGG-16	**86,24 %**	70,12 %	70,69 %	70,12 %	**94,01 %**
MobileNetV2	83,56 %	**78,66 %**	**78,67 %**	**78,66 %**	87,38 %
InceptionV3	65,56 %	48,32 %	46,05 %	48,31 %	77,38 %
CranioNet	85,41 %	64,67 %	63,91 %	64,67 %	78,75 %

***Table 1**. Evaluation metrics of the models trained with phantom data: precision (P), recall (R), f1-score (F1-S), and top-1 and top-2 accuracies (A).*

Regarding the phase estimation in the phantom simulation scenario, MobileNetV2 and InceptionV3 presented suboptimal results probably caused by overfitting. However, VGG-16 and CranioNet show adequate performance in the estimation of surgical phases from the video frames in comparison with the ground truth (**Figure 2**). Results showed that VGG-16 performs better than CranioNet, achieving an overall CA of 92.62 %, while CranioNet presents a CA of 73.56 %. Lower accuracy values were found for the estimation of phases 4, 5 and 6, which correspond to the placement of the surgical cutting guides and registration. The surgical field during these three phases is fairly similar and contains common tools (i.e. cutting guides are visible in phases 4,5, and 6).

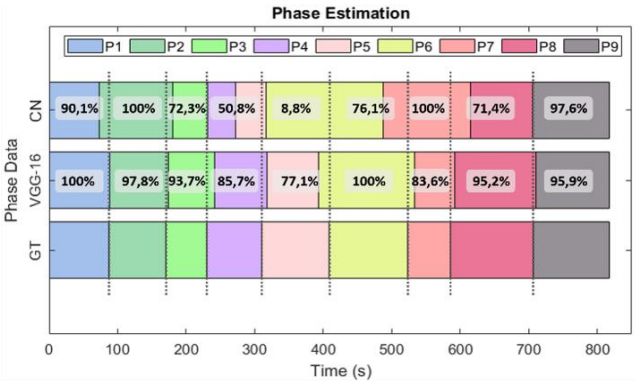


***Figure 2**. Phase estimation using the CranioNet (CN) and VGG-16 models trained with phantom data. Classification accuracy is shown for each phase and classification model.*

3.2. Surgical deployment

During surgical deployment, no overfitting occurred in any of the four models under study and optimal performance was achieved. Measured evaluation metrics in CranioNet, MobileNetV2 and VGG-16 showed similar performance for tool classification with an average precision above 95%. Results for InceptionV3 model indicate inferior performance in classification.

Regarding the surgical phase estimation, a suboptimal performance was obtained using InceptionV3 model with an overall CA of 78.61% (**Figure 3**). However, CranioNet, MobileNetV2 and VGG-16 models showed higher overall CA values of 97.14%, 94.89% and 92.01%, respectively.

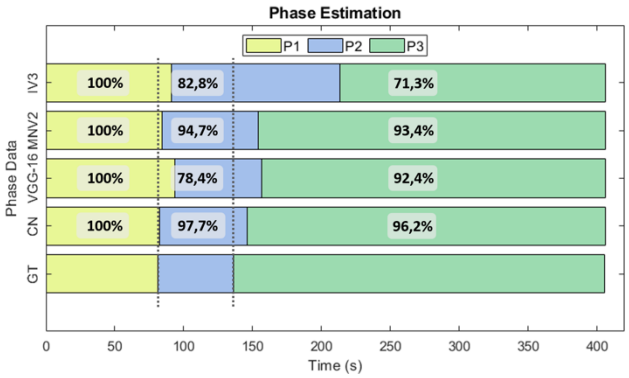


***Figure 3**. Phase estimation using the CranioNet (CN), VGG-16, MobileNetV2 (MNV2), and InceptionV3 (IV3) models trained with real surgery data. Classification accuracy is shown for each phase and classification model.*

4. Discussion

The results of this study demonstrate the feasibility of applying deep learning for the automatic detection of surgical tools and the estimation of surgical phases during craniosynostosis reconstruction surgery. Tool detection and phase estimation have been successfully tested in phantom simulation data and real surgery showing optimal performance. VGG-16 and CranioNet networks provided the most optimal classification results. Although VGG-16 obtained the best overall performance, CranioNet could also be used for accurate classification in applications where computational resources are limited and a faster frame processing and classification is needed.

Additionally, automatic tool detection has been integrated into a software application. This application was developed to display helpful guidance and generate data to provide further assessment and statistics related to phase duration estimation. The application is easy to use and provides the user with different options to adjust the predictions and export the classification data. The software is suitable for its use in real-time tool detection from live video streams, or prerecorded videos.

The implementation of a context-aware software in the operating room would be key in providing accurate feedback and warnings to the clinical staff when unexpected variations are identified in the workflow. Moreover, performing real-time phase detection would allow the tailoring of the intraoperative navigation software to optimize data visualization on the surgical screen depending on the current surgical phase, reducing the surgeon's subjectivity when performing bone reconstruction. It could also be used as a new training method based on a phantom simulator, which would be helpful not only for residents to practice before operating surgery for the first time but also for experienced surgeons to become confident with the workflow. Overall, the inclusion of this novel technology in the current surgical workflow will contribute to an improvement in the efficiency and reproducibility of this intervention and may lead to a reduction of surgical errors and complications. As a result, surgery could be more efficient, and cranial deformities could be corrected with better outcomes for the patient.

The main limitation of this study was the position of the camera for video recording during surgery. Camera placement was not suitable for acquisition of images of the surgical field. Therefore, the camera should be positioned closer to the surgical field (e.g. in the surgical lamps) to minimize occlusions by surgical staff and to ensure optimal illumination conditions. Modifying the position of the camera will improve the quality of the acquired images and will lead to better classification performance. In addition, the networks used in this study are only able to detect the presence of tools in the images, but they do not provide an estimation of their position within the images. Future research must focus on integrating new neural networks to consider tool positions in the estimation of surgical phases [6]. Additionally, it will be important to extend the study to more cases to increase the dataset and its variability, leading to an improvement of the results.

5. Conclusions

This work has demonstrated the feasibility of deep learning for real-time tool detection and phase estimation during craniosynostosis surgery. This technology has enabled the development of context-aware software for surgical workflow analysis and performance evaluation in the operating room. The integration of automatic tool detection in the current surgical workflow could benefit intraoperative assessment, postoperative analysis and medical training.

Acknowledgements

Supported by projects PI18/01625 (Ministerio de Ciencia, Innovación y Universidades, Instituto de Salud Carlos III and European Regional Development Fund "Una manera de hacer Europa") and IND2018/TIC-9753 (Comunidad de Madrid).

References

[1] D. Johnson and A. O. M. Wilkie, "Craniosynostosis," *Eur. J. Hum. Genet.*, vol. 19, no. 4, pp. 369–376, 2011, doi: 10.1038/ejhg.2010.235.

[2] D. García-Mato *et al.*, "Craniosynostosis surgery: workflow based on virtual surgical planning, intraoperative navigation and 3D printed patient-specific guides and templates," *Sci. Rep.*, vol. 9, no. 1, Dec. 2019, doi: 10.1038/s41598-019-54148-4.

[3] J. Burge *et al.*, "Application of CAD/CAM prefabricated age-matched templates in cranio-orbital remodeling in craniosynostosis." *J. Craniofac. Surg.*, vol. 22, no. 5, pp. 1810–1813, 2011, doi: 10.1097/scs.0b013e31822e8045.

[4] Y. Jin *et al.*, "Multi-task recurrent convolutional network with correlation loss for surgical video analysis," *Med. Image Anal.*, vol. 59, Jan. 2020, doi: 10.1016/j.media.2019.101572.

[5] S. Wang, A. Raju, and J. Huang, "Deep learning based multi-label classification for surgical tool presence detection in laparoscopic videos," in *Proceedings - International Symposium on Biomedical Imaging*, Jun. 2017, pp. 620–623, doi: 10.1109/ISBI.2017.7950597.

[6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection". arXiv preprint arXiv:1506.02640, 2015.

Premio José María Ferrero Corral

Jueves, 26 de noviembre de 2020

Characterization of polycaprolactone based electrospun scaffold towards in vitro human trabecular meshwork model

M. Bikuña Izagirre[1], E. Carnero González[2], L. Extramiana Esquisabel[2], J. Aldazabal[1], J. Moreno Montañés[2], J. Paredes Puente[1]

[1] Departamento de Ingeniería Biomédica, Universidad de Navarra TECNUN, San Sebastián, España, mbikunai,jaldazabal,jparedes@tecnun.es

[2] Departamento de Oftalmología, Clínica Universidad de Navarra, Pamplona, España, ecarnero,lextramiana@unav.es

Abstract

Glaucoma is the world´s second leading cause of irreversible blindness according to the World Health Organization. Age, race, family history of glaucoma, myopia and elevated intraocular pressure (IOP) are the most important risk factors to develop this pathology. The human trabecular meshwork (HTM), which is responsible for the regulation of the IOP by draining the aqueous humor (AH), is usually dysregulated during Glaucoma development. A decrease of draining capacity would render an increase of the IOP, which causes the stiffening of the trabecular meshwork (TM). However, the biophysical mechanisms by which these pathologic changes contribute to IOP elevation remain unclear. Moreover, the lack of appropriate in vitro TM models makes it challenging to understand the specifics of outflow physiology as well as to derive new anti-glaucoma therapies. For this reason, we propose a TM-emulating model based on nanofibrous composite scaffold fabricated by means of electrospinning technologies. In this work, randomly oriented polycaprolactone (PCL) nanofibers were used as scaffolds of the TM. Scanning electron microscopy and mechanical testing were performed in order to quantify the microstructure porosity and stiffness of both the scaffolds and the native tissue. Permeability studies were also undertaken by passing out a specific flow rate through the electrospun mesh, which simulates the AH. Our results show cellular viability of human trabecular meshwork cells onto the scaffolds. Finally, this in vitro platform will allow further research on the effects of TM rigidity on regulation of the IOP and glaucoma progression. In all likelihood, this research will pave the way for the development of new treatment strategies for this pathology.

1. Introduction

Glaucoma is an optic neuropathy that is characterized by the progressive functional degeneration of the optic nerve. It is a major public health concern, as it represents the second leading cause of irreversible blindness in the world[1]. There are several types of glaucoma, out of which primary open angle glaucoma (POAG) is the most extended one. A recent review estimated the global number of diagnosed POAG cases in 2013 to be around 44 million, while also foreshadowing an increase to 53 million by 2020 due to population aging[2]. Some uncertainty regarding the actual number of people with POAG remains, since as it is an asymptomatic disease, the number of affected individuals could be substantially larger than the number of confirmed cases[1,3]. The American Academy of Opthalmology recommends lowering the intraocular pressure (IOP) of the eye toward specific values that are believed to lower the progression of the disease, avoiding in this manner, the major retinal damage[3] for which glaucoma is notorious. IOP is generated by aqueous humor (AH) circulation system in the anterior eye segment. AH passes from the posterior chamber, through the pupil, into the anterior chamber, and then exits the eye via two pathways: (1) the conventional or trabecular outflow pathways, or (2) the unconventional or uveoscleral outflow pathway[4]. The conventional outflow pathway is more relevant, as it performs 80-90% of the AH drainage and provides resistance against the flow of AH, which inadvertently generates the IOP[4]. The human trabecular meshwork (TM) is composed of a connective tissue that contains a core of collagenous and elastic fibers, which are covered by flat cells that rest on basal lamina. These beams attach to one another in several layers and form a porous filter-like structure, which essentially divide the TM into three main regions: the uveoscleral meshwork, the corneoscleral meshwork and the juxtacanalicular tissue, localized directly adjacent to the inner wall endothelium of the Schlemm canal [4]. Changes in HTM structure and functionality are likely to affect tissue rigidity and biomechanical properties, modifying its resistance to AH flow and augmenting the probability of developing glaucoma.

Despite the pressing issue that this malady poses, Glaucoma research is severely limited by the use of animal models that do not translate well to human forms of the disease. Due to the wide anatomical inter-species variation in the TM and aqueous humour outflow pathway, using an animal in vivo model of the TM represents a substantial problem. Although advances in glaucoma research have been facilitated by in vitro 2D culture models, they are not ideal to gain a good understanding of the AH outflow process and TM physiology[5]. The uniqueness of this tissue, along with the lack of good animal models, makes the study of Glaucoma a challenging and complex task.

In consequence, developing a 3D artificial TM has become a justified endeavor whose necessity is reaching unprecedented heights. Until now, very few bioengineered in vitro HTM models have been developed. Torrejon et al. fabricated a SU-8 based scaffold, coated with 1% gelatin and with pores of 12 µm, which turned out to be the optimum combination for cell adhesion and growth[6]. More recently, a collagen and sulfate chondroitin tridimensional matrix has been developed. The authors of this work observed that the TM porcine cells were able to proliferate while maintaining their functionality[7].

ISBN: 978-84-09-25491-0, pags. 389- 392

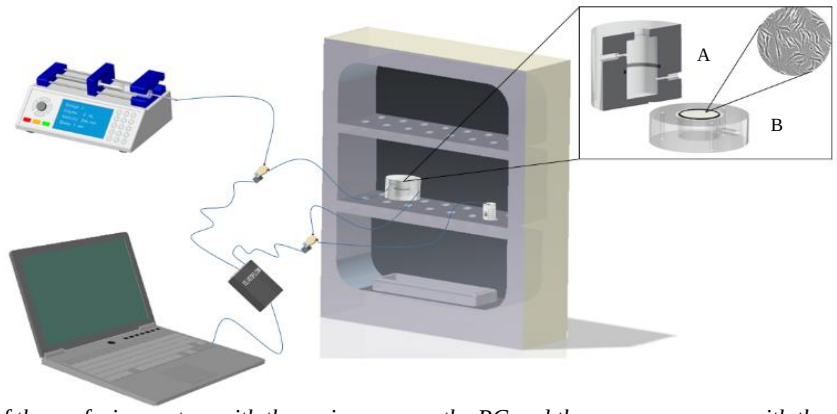


Figure 1. *Complete scheme of the perfusion system with the syringe pump, the PC and the pressure sensor with the reader.The camera is located inside the incubator. In the right-hand side of the figure, the inner structure of the camera can be seen. (A) Represents the way the liquid follows, entering from the upper part to exit from the bottommost hole. (B) The lowest piece of the camera, which contains the PCL mesh with the HTMC*

Building upon current knowledge of HTM morphology and physiology, in this study we fabricate an electrospun based porous membrane of polycaprolactone (PCL) to generate an in vitro HTM model. The PCL is a biocompatible and biodegradable polyester that allows cell adhesion and growth, as well as over time progressive degradation and elasticity to accommodate contractile functions[8]. The electrospinning technology allows the fabrication of porous microstructures to support HTM cells and to simulate similar outflow characteristics to in vivo TM. The fiber sizes as well as the stiffness were analyzed by scanning electron microscopy (SEM) and traction trials, respectively, with the goal of ensuring the native mechanical properties. These properties will allow the physiological proliferation of human trabecular meshwork cell (HTMC) on PCL fibers.

Following this, human trabecular meshwork cells (HTMC) were seeded onto PCL scaffolds to study the expression of their markers and characterize the secretion of extracellular matrix proteins. Subsequently, the permeability was studied introducing the bioengineered scaffold in a watertight pressure camera as part of a flow and pressure-controlled system, specifically designed to mimic the AH outflow. These studies established that higher permeability rates lead to lower scaffold stiffness because of lower pressure values.

2. Materials and Method

2.1. Materials & reagents

PCL pellets were purchased from Sigma-Aldrich (St. Louis, MO). Chloroform and methanol were purchased from Fisher Scientific. Dulbecco´s modified eagle medium (DMEM) was obtained from Gibco. The HTMC were obtained from donors.

2.2. Nanofibrous scaffold fabrication

Firstly, PCL (10 wt. %) has been dissolved in the chloroform/methanol solvent mixture with a volume rate of 4:1, v/v and by agitating the mixture with a magnetic stirrer at 600rpm overnight at room temperature (21 ±1 °C). Similarly, PCL (20 wt. %) was dissolved in a chloroform/methanol (3:1, v/v) solvent mixture by agitating the mixture with a magnetic stirrer at 600 rpm overnight at room temperature. For the electrospinning process, all the solutions were placed in a 1 mL plastic syringe exiting the solution with a speed of 1 mL/h through a 20G metallic needle. For PCL (10 wt. %) a voltage of 10kV was applied, with a distance of 15cm between the needle and the collector. For the higher concentration polymer, the voltage was raised to 15kV. The nanofibers were collected on a 9x9 cm^2 aluminum foils.

2.3. Characterization of nanofibrous scaffolds

The morphology of the composite nanofibrous scaffolds was studied with SEM. The fiber diameters were calculated based on SEM images by using image analysis software (ImageJ, NIH, USA).

The elasticity studies were carried out using the Zwik Roell (Ulm, Germany) traction instrument. All the samples were 6x1cm large, with a grip separation of 10mm and a test speed of 10mm/min.

2.4. Perfusion studies

The constructed circuit is equipped with a programmable syringe pump (Chemyx® Fusion 100, Stafford, USA) and two pressure sensors (Elveflow®, Paris, France) as shown in Figure 1. The system follows a simple working principle. A control program designed using the graphical programming environment LabView (National Instruments® Austin, Texas) is used to monitor and ensure the correct functioning of the apparatus. The interface portrays the real time values of the pressure sensors and generates a scenario-representative graph. The nanofibrous scaffolds were secured inside the camera. This system allows the control of the flow and the measurement of the transmembrane pressure, enabling the study of the permeability characteristics of our HTM model. In our experiment, we first introduced the different electropsun meshes inside the camera and perfused them with a flow rate of 20 µL/min for 20 min. On forward steps, HTMC seeded on PCL scaffolds will be introduced in the camera with the same flow rate to observe pressure differences. For future lines, the effect of pressure on these cells will be studied and different drugs will be perfused to observe the cell´s response.

2.5. HTMC culture

The primary HTMC were obtained from donor eyes (University of Navarra) or corneal ring donors (Barcelona tissue bank). HTMC isolated from the trabecular meshwork, were firstly cultured in a 6 well plate with 0.1% gelatin coating and DMEM, 20% FBS and 1% P/S. Fresh culture medium was supplied every 48h. Cells were maintained at 37°C in a humidified atmosphere with 5% carbon dioxide until 80–90% confluence was achieved at which point the cells were trypsinized using 0.25% Trypsin/0.5mM EDTA and subcultured. HTMC were subcultured on gelatin-coated 75 cm2 cell culture flasks. All studies were conducted using cells before the 5th passage.

2.6. Culture of HTM cells on PCL scaffolds

PCL 10% and PCL 20% scaffolds were sterilized under a UV hood. The electrospun meshes were suspended on cellular crowns (Figure 2) for 12 well plate (Sigma Aldrich, St. Louis, MO) and coated with gelatin 0.1 % to promote HTM cell attachment by soaking them in sterile solution for 30 min. HTM cells were seeded on the scaffolds in a density of 1x10^4 cells/cm^2 and cultured with 10 % FBS. Cell growth was monitored every 48h for 15 days, to finally treat them with 300nM of dexamethasone for 5 days to observe changes on F-actin distribution and to measure MYOC expression levels to corroborate the good proliferation of the cells over the nanofibrous scaffold.



Figure 2. *Cell crown with PCL 10% mesh where HTMC are seeded*

2.7. Inmunocytochemistry

Cells were fixed and stained for characteristic F-actin cytoskeleton and the nucleus. After 15 days in culture, the cells grown on the scaffolds were fixed with 4% paraformaldehyde, permeabilized with 0.2 % Triton X-100, and blocked with 5 % goat serum in PBS. HTM cells were subsequently incubated with Phalloidin ifluor 594 reagent (Abcam, Netherlands) to visualize the F-actin and stained with DAPI to reveal cell nuclei.

3. Results and Discussion

3.1. Mechanical characterization of PCL nanofibrous scaffolds

The 10 wt. % and 20wt. % PCL scaffolds were prepared by means of electrospinning methods and mechanically characterized. The fiber structure was observed under SEM as shown in Figure 3. The PCL 10% scaffolds had a mean fiber diameter of 1.85±0.21µm (mean ± sd) and mean porous area of 143,141±181,043 µm^2. For the PCL 20% scaffolds the mean fiber diameter value was 2.97 µm with a standard deviation of 0.38 µm, and the mean porous area was 366,501±507,185 µm^2. Pore dimensions among these sizes have shown successful results in cellular migration through scaffolds[9], with inadequate proliferation when pores are too small.

With regard to the traction studies, PCL 10% meshes show a considerably lower stiffness than the PCL 20% ones (Figure 4), the former type mesh having a Young Modulus of 3.9MPa versus the 20MPa Young Modulus of the latter type. Some studies showed the Young Modulus of healthy HTM states to be around 4kPa, whereas for the glaucomatous one it rises up to 80kPa [10]. Even despite the evident stiffening of the tissue due to the disease, the PCL scaffolds remain distant from those values.

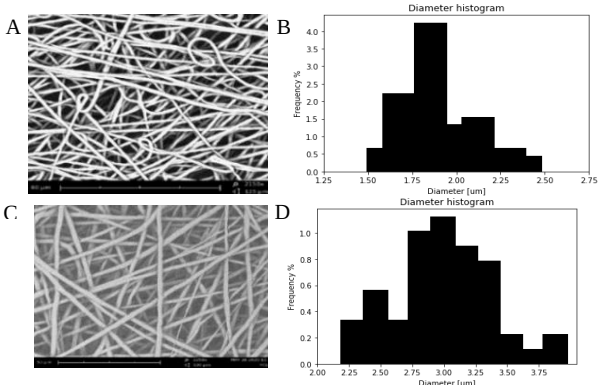


Figure 3. *Fiber analysis under SEM. (A) PCL 10% fibers (B) Fiber diameter histogram showing the prevalence of a diameter value of 1.85 µm. (C) PCL 20% fibers. (D) Fiber diameter histogram showing common sizes between 2.59 and 3.35 µm.*

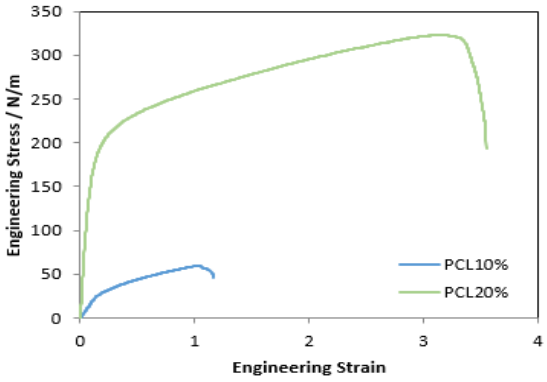


Figure 4. *Strain stress graphic comparing the both PCL scaffolds*

3.2. Permeability studies

Permeability studies were carried out in order to study the outflow resistance of the nanofibrous scaffolds. Thus, PCL 10% and PCL 20% meshes were introduced in the chamber and were exposed to different flow rates (10, 20 and 40 µL/min). The observed outcome is shown in Figure 5 which portrays how the pressure values increase as the flow also grows. Pressure measurements at different flow rates allow for the calculation of the outflow facility of the bioengineered HTM. The slope of the transmembrane pressure (P) versus flow rate (F) was 0.304 mBar/ µL/min for the PCL 10% and 0.209 mBar/µL/min for the PCL 20%, and the outflow facility (ΔF/ ΔP), calculated by taking the inverse of the slope, was 3.289 µL/min/mBar and 4.784 µL/min/mBar, respectively. These results concur with what has been shown in the literature, where

the conventional AH drainage resistance moves around 3.99-5.33mBar/ µL/min [11]. Therefore, PCL 10% meshes show higher permeability values in consequence of lower values in pressure.

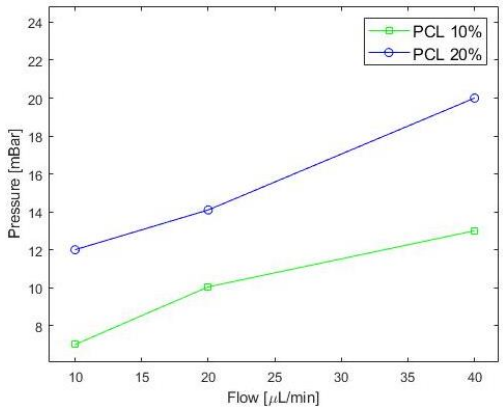


Figure 5. *Pressure values depending on the flow rate for PCL 10% and PCL 20%*

3.3. HYM cell proliferation over PCL 10% mesh

The cellular viability and proliferation was evaluated over PCL 10% scaffolds (Figure 6). There was not any significant difference between dexamethasone treated and non-treated cultures. Both cultures show a successful cellular proliferation obtaining 30-100ng/µL of RNA, together with MYOC expression, concluding a good cytoskeletal functioning.

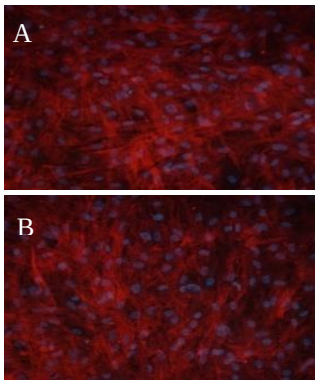


Figure 6. *PCL 10% cultures stained with DAPI for the nucleus and in red for F-actin. (A) Culture without dexamethasone (B) culture with dexamethasone. Scale bar: 100µm*

4. Conclusion

In this study, we confirm the feasibility of using an electrospinning based PCL nanofibrous scaffold to construct an in vitro HTM model for the purpose of studying the HTM outflow physiology and its response to biological agents. Despite the fact that PCL 10% meshes displayed a good mechanical and cellular response, we would like to study the PCL scaffolds with different natural polymers such as collagen, gelatin or elastin to study their mechanical characteristics and cellular viability. Moreover, our HTM perfusion platform and nanofibrous scaffold are ready for perfusion studies with HTM cells on the scaffolds to bring them under different drugs in order to study outflow facility and the cellular response.

References

[1] E. M. Hoffmann and V. Prokosch-Willing, “Primary Open Angle Glaucoma,” *Klin. Monbl. Augenheilkd.*, vol. 234, no. 11, pp. 1407–1422, 2017, doi: 10.1055/s-0043-118611.

[2] V. V. Kapetanakis, M. P. Y. Chan, P. J. Foster, D. G. Cook, C. G. Owen, and A. R. Rudnicka, “Global variations and time trends in the prevalence of primary open angle glaucoma (POAG): A systematic review and meta-analysis,” *Br. J. Ophthalmol.*, vol. 100, no. 1, pp. 86–93, 2016, doi: 10.1136/bjophthalmol-2015-307223.

[3] P. Robert N. Weinreb, MD, Tin Aung, MD, PhD, and Felipe A. Medeiros “The Pathophysiology and Treatment of Glaucoma,” *Texto Livre*, vol. 11, no. 2, pp. 192–205, 2018, doi: 10.17851/1983-3652.11.2.192-205.

[4] B. M. Braunger, R. Fuchshofer, and E. R. Tamm, “The aqueous humor outflow pathways in glaucoma: A unifying concept of disease mechanisms and causative treatment,” *Eur. J. Pharm. Biopharm.*, vol. 95, pp. 173–181, 2015, doi: 10.1016/j.ejpb.2015.04.029.

[5] J. Z. Gasiorowski and P. Russell, “Biological properties of trabecular meshwork cells,” *Exp. Eye Res.*, vol. 88, no. 4, pp. 671–675, 2015, doi: 10.1016/j.exer.2008.08.006.

[6] K. Y. Torrejon, D. Pu, M. Bergkvist, J. Danias, S. T. Sharfstein, and Y. Xie, “Recreating a human trabecular meshwork outflow system on microfabricated porous structures,” *Biotechnol. Bioeng.*, vol. 110, no. 12, pp. 3205–3218, 2013, doi: 10.1002/bit.24977.

[7] M. Osmond, S. M. Bernier, M. B. Pantcheva, and M. D. Krebs, “Collagen and collagen-chondroitin sulfate scaffolds with uniaxially aligned pores for the biomimetic, three dimensional culture of trabecular meshwork cells,” *Biotechnol. Bioeng.*, vol. 114, no. 4, pp. 915–923, 2017, doi: 10.1002/bit.26206.

[8] J. S. Choi, S. J. Lee, G. J. Christ, A. Atala, and J. J. Yoo, “The influence of electrospun aligned poly(ε-caprolactone)/collagen nanofiber meshes on the formation of self-aligned skeletal muscle myotubes,” *Biomaterials*, vol. 29, no. 19, pp. 2899–2906, Jul. 2008, doi: 10.1016/j.biomaterials.2008.03.031.

[9] C. M. Murphy and F. J. O’Brien, “Understanding the effect of mean pore size on cell activity in collagen-glycosaminoglycan scaffolds,” *Cell Adhes. Migr.*, vol. 4, no. 3, pp. 377–381, 2010, doi: 10.4161/cam.4.3.11747.

[10] J. A. Last *et al.*, “Elastic modulus determination of normal and glaucomatous human trabecular meshwork,” *Investig. Ophthalmol. Vis. Sci.*, vol. 52, no. 5, pp. 2147–2152, 2011, doi: 10.1167/iovs.10-6342.

[11] M. Goel, “Aqueous Humor Dynamics: A Review~!2010-03-03~!2010-06-17~!2010-09-02~!,” *Open Ophthalmol. J.*, vol. 4, no. 1, pp. 52–59, 2010, doi: 10.2174/1874364101004010052.

Parametrización del Vectorcardiograma Auricular Para la Caracterización de Distintos Tipos De Flutter

S. Ruipérez-Campillo[*] [1,3], S. Castrejón, M. Martínez[2], R. Cervigón[4], O. Meste[5], J.L. Merino[2], J. Millet[1] y F. Castells[1]

[1] Grupo de Investigación BIO-ITACA, Universitat Politècnica de València, Valencia, España
[2] Unidad de Arritmias y Electrofisiología Robotizada, Hospital Universitario La Paz, Madrid, España
[3] Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid, España
[4] Universidad de Castilla la Mancha, Cuenca, España
[5] Université Cote d'Azul, Niza, Francia

*Correspondencia: samuel.ruiperez@alumnos.uc3m.es

Resumen

El objetivo de este estudio es caracterizar de forma no invasiva distintas subclases del flutter auricular (FLA), un tipo de taquiarritmia definida por un circuito de macrorreentrada en la cavidad auricular. Una metodología basada en el vectorcardiograma (VCG) es propuesta con el objetivo de comparar dichos circuitos. Para ello, se definen una serie de arquetipos de modo que se puedan calcular parámetros como el de la similitud. La metodología se aplica en ambos, un grupo de VCGs sintéticos creados a partir de una simulación basada en un modelo matemático, y en señales de pacientes. En estos últimos, se empleó adenosina para bloquear la actividad ventricular en la adquisición de la señal de ECG, que posteriormente se transforma en VCG. Además, las zonas de conducción lenta aparecen de forma clara cuando se representa esta señal discretizada en tiempo, de acuerdo con la definición de los mecanismos fisiológicos del FLA y su importancia a la hora de caracterizarlo. En conclusión, la metodología diseñada en este artículo permite diferenciar de forma no invasiva entre los tipos más recurrentes de FLA a partir del análisis de la representación del VCG, prediciendo la presencia de zonas de conducción lenta en la macrorreentrada. Esto puede ser muy útil a la hora de planear y modelar la intervención quirúrgica en el paciente.

1. Introducción

La prevalencia del flutter auricular (FLA) está en aumento [1], por lo que su diagnóstico es un tema esencial de estudio hoy en día. El fundamento fisiológico del FLA es un circuito auricular de macrorreentrada alrededor de un obstáculo en sentido horario (H) o antihorario (AH). La perpetuación del circuito de macrorreentrada está asociada a zonas de conducción lenta que permiten la repolarización total de los miocitos terminado el periodo de refracción. En función de este recorrido, se distinguen distintos tipos de FLA, siendo el FLA común el más prevalente. En las últimas décadas, sistemas de mapeo electroanatómico han facilitado mapear la aurícula y el proceso de ablación. Sin embargo, estos requieren tiempo y experiencia, por lo que nuevas técnicas no invasivas de mapeo como las de imagen electrocardiográfica (ECGI) están siendo desarrolladas (aunque sus resultados aún suponen controversia hoy en día [2]). Por ello, obtener información previa a la ablación sería una potente herramienta si permitiese identificar el tipo de FLA y el circuito de macrorreentrada más probable. Esto permitiría simplificar el mapeo electroanatómico dado que la región de interés ya habría sido identificada. También supondría por tanto una mejora de la intervención quirúrgica, dado que se podría identificar si se trata un FLA de aurícula izquierda o derecha.

El diagnóstico de FLA apenas a cambiado a lo largo de la historia, teniendo como referencia las ondas auriculares del electrocardiograma (ECG). Sin embargo, es muy común encontrar falsos positivos y negativos incluso en el FLA común. Además, la ratio aurícula/ventrículo suele ser 2:1, complicando la diferenciación de la onda T. Algunos estudios tratan de cancelar el componente ventricular en la señal [3,4], pero esto no es posible con el FLA dado que la contribución ventricular está acoplada. Tampoco sería suficiente la separación ciega de fuentes [5] pues se requerirían tres componentes auriculares independientes para crear vectorcardiograma (VCG). Por tanto, obtener la señal auricular en FLA es todavía un reto, por lo que administrar adenosina para bloquear el nodo atrio ventricular (AV) por un periodo corto de tiempo es una posible solución.

La interpretación más común del VCG consiste en bucles auriculares en 3D. Algunos parámetros han sido definidos para caracterizar y cuantificar rasgos del circuito QRS [6]. Los VCGs han sido descritos en FLA, pero no se han establecido correlaciones claras con sus características fisiológicas [7]. La hipótesis de este estudio establece que hay una correlación entre el tipo de FLA y su representación en forma de VCG, basada fundamentalmente en las regiones de conducción lenta. Por tanto, caracterizando los VCGs se podrían determinar y diferenciar distintos tipos de FLA.

2. Diseño Experimental

Con el objetivo de demostrar la hipótesis y llevar a cabo la caracterización del FLA, se diseña una metodología basada en un punto de vista innovador de la señal de VCG, donde las zonas de conducción lenta toman un papel fundamental. Además, se diseñan simulaciones computacionales basadas en un modelo geométrico que permiten evaluar esta metodología. Sobre esta base surge un generador de VCGs sintéticos en base al modelo matemático. Por otro lado, las señales ECG de pacientes que padecen FLA provienen de la sección de Arritmias y Electrofisiología Robotizada del Hospital Universitario La Paz (Madrid, España). Todos los procedimientos llevados a cabo están en concordancia con el Protocolo de Helsinki. La caracterización de los datos clínicos de FLA viene dada por

ISBN: 978-84-09-25491-0, pags. 393- 396

cardiólogos y electrofisiólogos. Este diagnóstico es nuestro estándar de referencia. Las señales ECG fueron transformadas a VCGs a través de la inversa transformada de Dower para extraer parámetros cuantificables sobre los que se han llevado a cabo las pruebas estadísticas.

3. Materiales y Métodos

3.1. Generador de VCGs Sintéticos

La creación de un modelo geométrico para la creación de trayectorias con características similares a los bucles de VCGs auriculares de pacientes, será muy conveniente para demostrar la hipótesis y evaluar la metodología. Este modelo genera 8 grupos diferentes de VCGs sintéticos, diferenciados por la localización de la zona de conducción lenta. Varios parámetros se aplican para tener una extensa variabilidad entre los diferentes casos pertenecientes a un mismo grupo: desde los grados de libertad del plano del VCG hasta la complejidad de la curvatura de la trayectoria.

Los bucles auriculares sintéticos se generan a partir de proyecciones de figuras elipsoidales aleatorizadas que provienen de la expresión de Euler implementada. Además de las zonas de conducción semilenta que aparecen debido a la naturaleza de la elipse al discretizar angularmente de forma continua cerca de los semiejes de la misma (Figura 1.A), una zona lenta característica del bucle se define a través de una discretización angular no uniforme:

$$\Delta\theta_n = \lambda \left|\cos\left(n\frac{\pi}{N} + \phi\right)\right| + \Delta\theta_{min} \qquad (1)$$

donde $\Delta\theta_n$ representa el incremento angular en cada muestra n, siendo el mínimo $\Delta\theta_{min}$ (que define la zona lenta) y el máximo $\lambda + \Delta\theta_{min}$ (dando lugar a las zonas de conducción rápida), con N como número total de muestras y ϕ, el argumento de la fase, determinando la localización de las zonas de conducción lenta. La suma de todos los incrementos $\Delta\theta_n$ debe ser 2π (Figura 1.B-C y ecuación (2)).

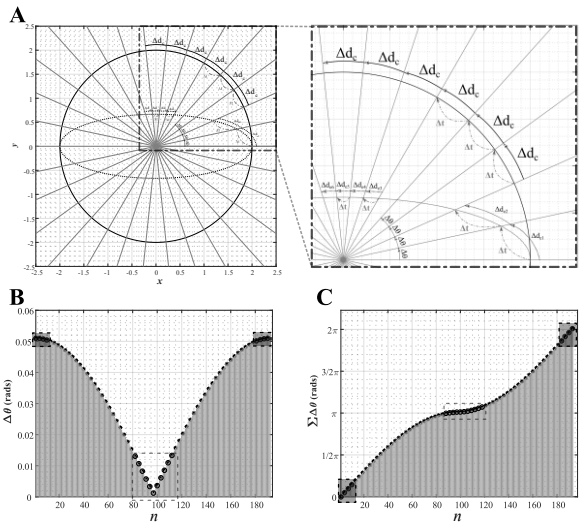


Figura 1. **A.** *Discretización angular constante en 2D sobre un círculo y una elipse para mostrar la concentración de muestras en las zonas cercanas a los semiejes de la elipse (semilentas).* **B.** *Representación del incremento angular (ecuación (1)) de la elipse.* **C.** *Sumatorio del incremento angular de la elipse a lo largo de la trayectoria en un plano 2D: zonas rápidas ($\lambda+\Delta\theta_{min}$) en gris oscuro y lentas ($\Delta\theta_{min}$) en gris claro (ecuación (2)).*

$$\sum_{n=1}^{N}(\Delta\theta_n) = 2\left[\lambda + \frac{N}{2}\Delta\theta_{min} + \lambda\frac{\cos\frac{\pi}{4}\sin\frac{\pi}{4}}{\sin\left(\frac{\pi}{2N}\right)}\right] = 2\pi \qquad (2)$$

Para evitar que los bucles estén restringidos a un solo plano, se introducen variaciones en la curvatura de las proyecciones de los diferentes planos con funciones (F_1, F_2 y F_3) ponderadas (por P1, P2 y P3) que modulan la frecuencia:

$$r(\theta) = \frac{x \cdot y}{\sqrt{(x \cdot \cos(\theta) + P_1F_1)^2 + (y \cdot \sin(\theta) + P_2F_2)^2 + P_3F_3}} \qquad (3)$$

Finalmente, algunos comportamientos antinaturales en términos fisiológicos pueden aparecer, por ejemplo, en el perfil de la velocidad. Por tanto, la señal es tratada con una implementación de un filtro adaptado de paso bajo Savitzky-Golay en los tres ejes (Figura 2 B-C). En conclusión, este modelo define 4 tipos de VCGs sintéticos dependiendo de la localización de la zona lenta. Además, se tienen en cuenta ambos sentidos, H y AH, dando lugar a 8 grupos independientes: del 1 al 4 y del 5 al 8 (H y AH, respectivamente). La figura 2 muestra un ejemplo VCG sintético.

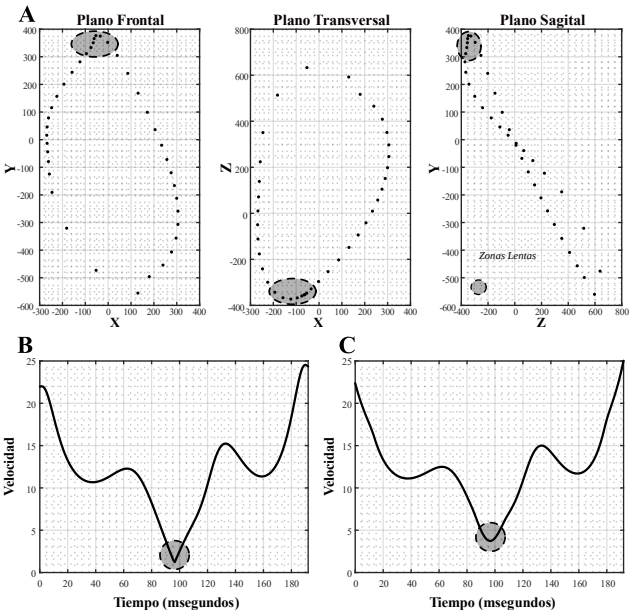


Figura 2. *Ejemplo de VCG sintético.* **A**. *Vistas de los planos frontal, transversal y sagital.* **B**. *Perfil de velocidad sin filtrar.* **C**. *Perfil de velocidad tras aplicar la implementación de un Savitzky-Golay adaptado.*

3.2. Datos Clínicos: ECGs de los Pacientes

En el estudio participaron 30 pacientes de FLA. Se registraron las señales de ECG y electromiograma (EGM) con una frecuencia de muestreo de 1kHz usando un polígrafo (Labsystem pro, Bard, Boston Scientific) sincronizado con un electronavegador. Durante la adquisición de la señal de ECG, se le subministró adenosina a los pacientes para descartar la contribución ventricular. Los pacientes pertenecientes a los siguientes grupos han sido considerados para el estudio: 5 de FLA común H, 6 de FLA común AH, 3 de FLA perimitral H, 8 de FLA perimitral AH y 8 misceláneos, de grupos como FLA de las perivenas cavas o pulmonares.

3.3. Procesado y Creación de Arquetipos Clínicos

Los bucles de VCG se calculan a tomando segmentos de señales ECG bajo el efecto de adenosina. Estas señales se filtran entre 1Hz y 30Hz bidireccionalmente para minimizar la distorsión transitoria. Los VCGs se calculan usando la Inversa Transformada de Dower, para estimar las derivaciones ortogonales de Frank a partir de las 12 derivaciones del ECG. La reconstrucción de un VCG junto a sus zonas de conducción lenta se muestra en la Figura 3. La consistencia entre los distintos bucles se calcula para observar la repetibilidad de la onda auricular, es decir, la calidad de la señal. Tomando 10 ondas auriculares, se calcula la consistencia por descomposición de autovalores, obteniendo cerca de 1 cuando el bucle de FLA es repetible.

Para cada grupo de FLA considerado (en ambos casos, clínicos y sintéticos), se crea un arquetipo a partir de la media normalizada de los diferentes VCGs, alineados temporalmente, utilizando la técnica de mínimos cuadrados.

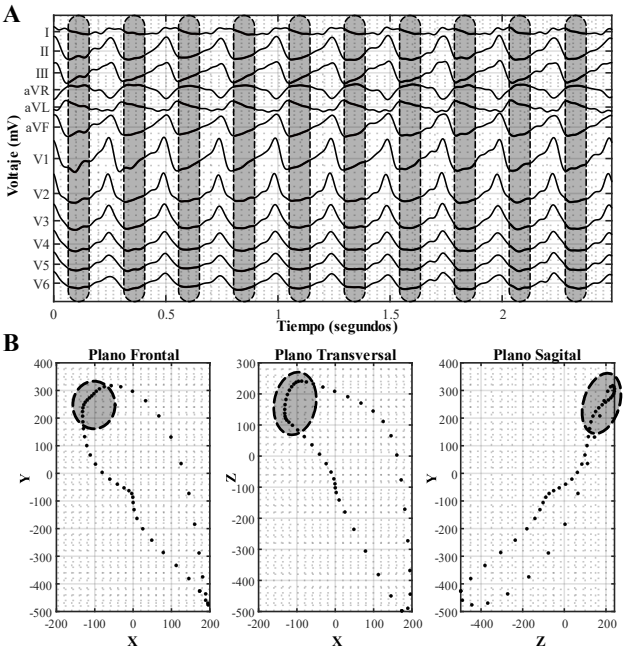


Figura 3. *Señal de un paciente enfermo de FLA perimitral AH.* **A.** *ECG con las zonas de conducción lenta marcadas.* **B.** *Reconstrucción del VCG del paciente con las zonas lentas sombreadas.*

3.4. Caracterización de los Bucles de VCG

Con el objetivo de caracterizar los bucles de VCG se han definido una serie de parámetros. Trataremos algunos en esta sección. Para comenzar, la similitud entre bucles de VCG se calcula como una correlación punto a punto:

$$S = \frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{x}_i^{\mathrm{T}}\mathbf{y}_i}{\|\mathbf{x}_i\|\|\mathbf{y}_i\|}, \qquad (4)$$

donde $\mathbf{x}_i$ e $\mathbf{y}_i$ son vectores en 3D de cada bucle en la muestra i, siendo el número total de estas, N. Es importante tener en cuenta que las zonas de conducción lenta se pueden entender como una acumulación de puntos, dando lugar a una ponderación mayor en este parámetro que se basa en la dirección de los vectores (obsérvese la Figura 3). Para identificar el arquetipo al que más se asemeja un bucle, se calcula el parámetro de similitud con respecto a todos ellos, por lo que el mayor coeficiente se corresponderá con el más parecido. En los casos clínicos se utiliza el algoritmo LOO ('Leave One Out'), para no sesgar el resultado.

Otro parámetro que se ha definido es el perfil de velocidad a lo largo del circuito del VCG. Las unidades se expresan en unidad de incremento de voltaje por tiempo [V/ciclo]. Para cuantificar las zonas lentas en el recorrido, los siguientes parámetros se han implementado: porcentaje de la distancia en la macrorreentrada donde la velocidad era lenta (VL_D), el porcentaje temporal (VL_T) y el ratio entre ambos (VL_R). Además, el perfil de velocidad angular también puede aportar información complementaria para entender mejor el comportamiento de la señal:

$$\omega_i = \frac{1}{T_s}\arccos\left(\frac{u_i^{\mathrm{T}}\,u_{i+1}}{\|u_i\|\|u_{i+1}\|}\right), \qquad (5)$$

donde u es el vector 3D en la muestra i, con periodo T_s.

Finalmente, un parámetro que resulta interesante en este estudio es el de la complejidad del recorrido (C), que evalúa los cambios en la dirección de la trayectoria a lo largo del bucle. El sumatorio de estos cambios tiene un límite inferior de 2π, por lo que el parámetro es normalizado como se muestra en la ecuación (6). El recorrido más simple supone $C = 0$ mientras que uno muy complejo significaría que C tiende a 1.

$$C = 1 - \frac{2\pi}{\sum_i \arccos\left(\frac{(\mathbf{x}_i - \mathbf{x}_{i-1})^{\mathrm{T}}(\mathbf{x}_{i+1} - \mathbf{x}_i)}{\|\mathbf{x}_i - \mathbf{x}_{i-1}\|\|\mathbf{x}_{i+1} - \mathbf{x}_i\|}\right)} \qquad (6)$$

3.5. Análisis Estadístico

Para analizar los resultados de los casos generados por el modelo matemático se ha usado el análisis de varianza ANOVA, siguiendo la distribución de F de Snedecor. La hipótesis nula se descarta para p-valores menores a 0.05. En este caso se utiliza la corrección de Bonferroni para tener en cuenta la aleatoriedad debido a las comparaciones múltiples. También se ha utilizado el área bajo la curva ROC al comparar la correlación con cada arquetipo de los grupos sintéticos. Para los datos clínicos, la prueba de Kruskal-Wallis se llevó a cabo, con la misma condición de p-valor para descartar la hipótesis nula.

4. Resultados

Los datos sintéticos muestran resultados estadísticamente significativos al comparar los diferentes casos con los arquetipos de su propio tipo: desde F (3,3996) = 959,20 con $p < 0{,}001$ hasta F (3,3996) con $p < 0{,}001$. Al llevar a cabo comparaciones múltiples en el parámetro de la correlación o similitud, se utilizó la corrección de Bonferroni que dio lugar a significación estadística para todas las posibilidades ($p < 0{,}001$). Evaluando las curvas ROC a través del área bajo la curva (ABC), se obtuvieron valores entre 0,895 y 0,923 (Figura 4.A). Además, en un entorno controlado, se evaluaron las funciones moduladoras de la frecuencia dando lugar a un rango de C entre 0 y 0.8. También se evaluó el movimiento de las zonas lentas rotando con respecto a ambos semiejes, sobre la zona lenta (S_{min} = 0,8) y sobre el semieje mayor, suponiendo el movimiento espacial de la zona lenta (S_{min} = 0,3).

Para los datos clínicos, la consistencia intra-paciente se calculó obteniendo un coeficiente de 0,95 ± 0,04. La similitud entre los distintos pacientes y los arquetipos (calculados acorde al algoritmo LOO) dieron lugar a coeficientes mayores a 0,84 para todos los grupos con su propio arquetipo, con desviaciones estándar por debajo de 0,05 para estos (Figura 4.B). La tabla 1 muestra estos resultados.

	AFL C. AH	AFL C. H	AFL P. AH	AFL P. H	AFL Miscel.
Arq. C.AH	**0,85±0,03**	0,67±0,11	0,77±0,05	0,73±0,12	0,68±0,14
Arq. C.H	0,68±0,05	**0,95±0,03**	0,66±0,13	0,75±0,04	0,58±0,13
Arq. P.AH	0,78±0,09	0,63±0,11	**0,87±0,04**	0,68±0,12	0,62±0,18
Arq. P.H	0,76±0,08	0,75±0,01	0,70±0,12	**0,91±0,02**	0,65±0,17
p-valor	0.014	<0.01	<0.01	0.62	N.A.

Tabla 1. *Correlación entre los pacientes de los diferentes tipos de FLA (C. para común, P. para perimitral) y los arquetipos (arq.).*

El análisis del perfil de la velocidad usando la prueba de Kruskal-Wallis dio lugar a resultados estadísticamente significativos para VL_T (p = 0,047), y no significativos para el resto: VL_D (p = 0,146) y VL_R (p = 0,658). El parámetro de la complejidad es mayor para los casos perimitrales, para los cuales se encuentran también mayores periodos. El grupo misceláneo, que incluye pacientes con malformaciones, presenta el mayor coeficiente *C*. Además, para este parámetro, se obtienen resultados estadísticamente significativos (p = 0,035).

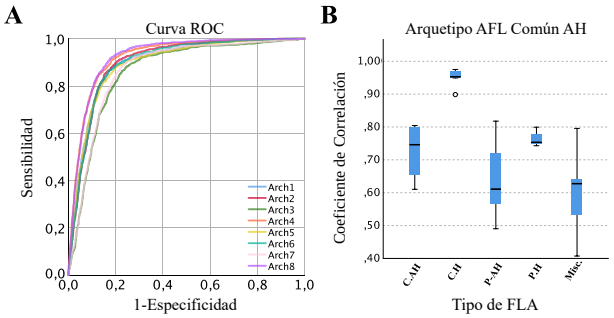


Figura 4. **A.** *Curva ROC de los grupos de VCG sintéticos.* **B.** *Gráfica de la correlación entre los distintos grupos de AFL clínico y el arquetipo común AH.*

5. Discusión y Conclusiones

La causa fisiológica del FLA son circuitos de macrorrentrada en la aurícula que requieren distintas intervenciones de ablación. Actualmente no es posible identificar el tipo de FLA antes de realizar la intervención, por lo que en este estudio se propone una metodología no invasiva cuya base es el VCG para obtener información útil antes de llevar a cabo el procedimiento quirúrgico.

El modelo matemático creado para obtener una base de datos de VCGs sintéticos, respalda la hipótesis de la importancia de las regiones de conducción lenta a la hora de caracterizar el FLA. La capacidad discriminatoria de este modelo se observa, por ejemplo, en los altos valores del ABC al comparar los distintos grupos con todos los arquetipos.

En el análisis clínico, los coeficientes de correlación más altos y la desviación estándar más baja, se obtuvieron al comparar los pacientes de un tipo con el arquetipo del mismo grupo. Estos resultados revelan una capacidad muy prometedora para identificar nuevos casos de FLA que no hayan sido previamente caracterizados. Además, las regiones lentas aparecen en áreas similares para todos los VCGs de pacientes de un mismo grupo (Figura 5).

En conclusión, este estudio propone una metodología no invasiva para caracterizar el FLA a partir de señales de ECG, aportando valiosa información para planificar y llevar a cabo las intervenciones quirúrgicas de forma más eficiente.

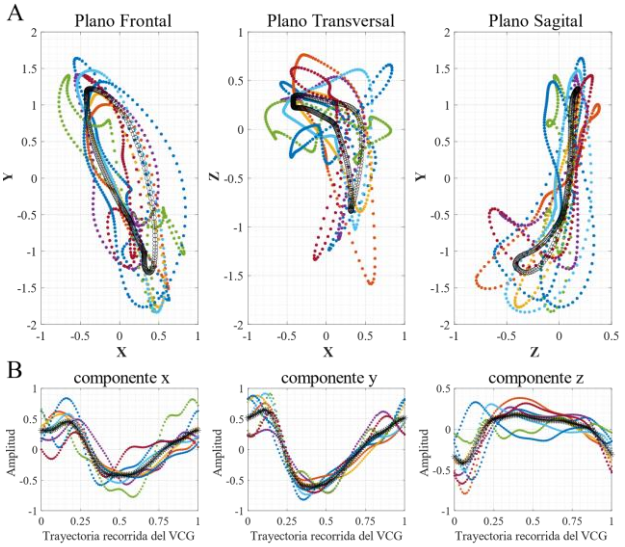


Figura 5. **A**. *Representación de los VCGs de pacientes de FLA primitral AH y su arquetipo (negro).* **B**. *Trayectoria de los VCG en A, en cada una de las componentes cartesianas.*

Agradecimientos

Este proyecto ha sido financiado por el Ministerio de Ciencia e Innovación del Gobierno de España (PID2019-109547RB), el Instituto de Salud Carlos III (CIBERCV CB16/11/00486) y la Generalitat Valenciana (PROMETEOII/2014/037). Gracias también a la Fundación Rafael del Pino por el apoyo proporcionado para la formación del primer autor del artículo.

Referencias

[1] William Dresen and Pamela K Mason. “Atrial flutter after surgical maze: incidence, diagnosis, and management”. In: Current opinion in cardiology 31.1 (2016), pp. 57–63.

[2] Charulatha Ramanathan et al. “Noninvasive electrocardiographic imaging for cardiac electrophysiology and arrhythmia”. In: Nature medicine 10.4 (2004), pp. 422–428.

[3] Castells F, Mora C, Rieta J, Millet J. Estimation of atrial fibrillatory wave from single-lead atrial fibrillation electrocardiograms using principal component analysis concepts. Medical Biological Engineering Computing 2005;43:557–560.

[4] Stridh M, Sornmo L. Spatiotemporal qrst cancellation techniques for analysis of atrial fibrillation. IEEE Transactions on Biomedical Engineering 2001;48(1):105–111.

[5] Castells F, Rieta JJ, Millet J, Zarzoso V. Spatiotemporal blind source separation approach to atrial activity estimation in atrial tachyarrhythmias. IEEE Transactions on Biomedical Engineering 2005;52(2):258–267.

[6] Choudhuri S, Ghosal T, Goswami DP, Sengupta A. Planarity of the spatial qrs loop of vectorcardiogram is a crucial diagnostic and prognostic parameter in acute myocardial infarction. Medical hypotheses 2019;130:109–251.

[7] Ng J, Sahakian AV, Fisher W, Swiryn S. A vectorcardiographic approach to understanding the 12-lead electrocardiogram of atrial flutter. In Computers in Cardiology, 2004. IEEE, 2004; 629–632.

Validación de prototipo con sensores vestibles para captura y análisis de movimientos del cuerpo humano con aplicación en medicina

J. González-Alonso[1], M. Martínez-Zarzuela[1], H. J. Aguado[2], F.J. Díaz- Pernas[1],

[1] Grupo de Telemática e Imagen, Universidad de Valladolid, España,

javier.galonso@alumnos.uva.es, {marmar, pacper}@tel.uva.es

[2] Unidad de Traumatología, Hospital Clínico Universitario de Valladolid, España, hjaguado@gmail.com

Resumen

Estudios recientes confirman la precisión y posibilidades de aplicación de sistemas de sensores inerciales para la captura de movimientos en investigación biomédica. En este trabajo se presenta un prototipo de sistema de captura y análisis de movimiento con sensores inerciales de desarrollo propio. El sistema incluye el hardware y el software necesario para permitir la adquisición de movimiento en tiempo real con hasta 11 sensores de forma inalámbrica. Las medidas de precisión obtenidas son similares a las alcanzadas con un sistema comercial Xsens Dot, con diferencias en la captura inferiores a 1º en pruebas de laboratorio con un goniómetro digital e inferiores al margen de error de un goniómetro estándar en pruebas médicas de medición de rango articular en codo y hombro.

1. Motivación

El análisis del movimiento humano (Human Motion Analysis) tiene relevancia en el contexto de la bioingeniería, ya que permite realizar valoraciones clínicas a partir del cálculo de parámetros asociados a la movilidad física. La medición precisa de los movimientos del cuerpo humano es crucial para identificar desórdenes de carácter biomecánico y patrones atípicos de control neuromuscular, así como prevenir la aparición de lesiones musculoesqueléticas, entre otras aplicaciones [1].

En los últimos años, se ha despertado un interés creciente en el desarrollo y estudio de factibilidad de sistemas de captura de movimiento basados en la utilización de vestibles con sensores inerciales ó IMU (Inertial Measurement Unit) [2][3]. La utilidad de estas soluciones está avalada a través de estudios comparativos de precisión con sistemas optoelectrónicos, considerados el estándar de referencia [4]. Una gran ventaja frente a estos últimos es la de que no requieren de un espacio específico dedicado para la adquisición del movimiento. Esta circunstancia hace que en el futuro tengan bastantes posibilidades de extenderse en el ámbito clínico. Adicionalmente, es importante señalar su posible utilización en el entorno natural del paciente, ya sea para obtener mediciones relevantes para la personalización de los tratamientos o para actuar como dispositivos de interacción en sistemas de telerehabilitación [5].

Los sistemas comerciales con sensores IMU como Xsens o Physilog, tienen un coste bastante elevado, al que por lo general es necesario sumar licencias de software. Estudios recientes, utilizan versiones de vestibles no comerciales desarrolladas por equipos de investigación [6].
De manera general, en estos estudios se emplea un número reducido de sensores. En este sentido, las versiones comerciales más asequibles también presentan esta limitación. El sistema Xsens Dot empleado en nuestro estudio comparativo sólo permite 5 sensores simultáneos. Una limitación adicional, es la de que estos sistemas utilizan protocolos de comunicación sobre bandas de uso muy extendido por dispositivos WiFi y Bluetooth, lo que causa interferencias que se traducen en la pérdida de datos.

En este artículo presentamos un prototipo de sistema de captura de movimiento con sensores vestibles, creado desde cero por nuestro equipo de investigación, así como un primer estudio de pruebas de funcionamiento en laboratorio y en clínica. Nuestro sistema proporciona una medida fiable análoga a la de sensores comerciales a una fracción de su coste y con hasta 11 sensores, lo que permite realizar capturas en las que se recogen las principales articulaciones del cuerpo (1 en cabeza, 2 en espalda, 4 en brazos y 4 en piernas). Así mismo, han sido diseñados para comunicarse de forma inalámbrica con un ordenador y ser inmunes a interferencias en la banda de 2,4 GHz, mediante un protocolo desarrollado por el grupo de investigación, cuya descripción será objetivo de un futuro trabajo. Aunque nuestros sensores permiten otros tipos de comunicación sobre 2,4 GHz como Bluetooth, nuestro protocolo permite elevar el número de sensores de estos sistemas y facilita su utilización en todo tipo de contextos, incluso en los más saturados.

La sección 2 describe el sistema propuesto, la sección 3 el procedimiento utilizado durante las pruebas y la sección 4 los resultados obtenidos. En la sección 5 se discute la validez de nuestros sensores en base a los experimentos llevados a cabo en otros estudios. Finalmente, en la sección 6 se plantean las principales conclusiones obtenidas y posibles aplicaciones y modificaciones futuras del sistema.

2. Sistema Propuesto

El objetivo final es la implementación de una solución para el análisis de movimientos corporales con sensores vestibles y con una fiabilidad y estabilidad de mediciones comparable con sistemas comerciales, empleados en un gran número de estudios clínicos [3]. Además, el prototipo del sistema debe

ISBN: 978-84-09-25491-0, pags. 397- 400

responder al compromiso de proporcionar una tasa de datos alta y estable, para una buena captura de movimientos a cierta velocidad [6], además de robustez a factores ambientales e inmunidad a interferencias [7].



(a) Comparación tamaño de los vestibles.



(b) Colocación conjunta con Xsens Dot.

Figura 1. *Sensores utilizados para el estudio.*

A nivel hardware, nuestro sistema consta de 11 vestibles, conteniendo cada uno, una unidad de medida inercial (IMU), un núcleo de procesamiento y transmisión de datos y una batería de pila de botón recargable (LIR2032 3,7V y 45mAh en ambos casos), todo en un mismo encapsulado, con unas dimensiones totales de 45 × 30 × 10 mm similares a las de los Xsens Dot de 36 x 30 x 11 mm, como podemos apreciar en la Figura 1(a). Para la construcción de nuestros prototipos elegimos procesadores ARM Cortex M4 con transceptor de 2,4 GHz (alcance de comunicación aproximado de 7-8 m) y sensores inerciales BNO080, que permiten tasas de datos superiores a 100Hz y buena respuesta a perturbaciones ambientales [7]. Cada uno de los sensores contiene un acelerómetro, un giroscopio y un magnetómetro triaxiales. La duración aproximada de la batería en el caso de los Xsens Dot (5 sensores) es de 5 horas, y nuestros prototipos alcanzan una autonomía de 4 horas sobre Bluetooth y 2 horas con nuestro protocolo de comunicación y 11 sensores.

El sistema propuesto cuenta con varias capas de software. En el primer nivel se ha programado el firmware de los sensores empleando C/C++ y el sistema operativo de Tiempo Real *FreeRTOS*. Para permitir un sistema funcional con un elevado número de sensores simultáneos y reducir las interferencias con otras señales, se ha desarrollado un protocolo propio que utiliza frecuencias portadoras distintas a las de los canales WiFi y Bluetooth. Los datos enviados por los sensores se recogen en un dispositivo conectado al puerto USB del ordenador, permitiendo comunicar todos los vestibles en tiempo real. Cabe aclarar, que en ambos casos el firmware de los sensores ya cuenta con técnicas de fusión de datos y de filtrado, eliminando la necesidad de realizar cálculos posteriores.

La segunda capa de software consiste en una aplicación de escritorio desarrollada con el motor de juegos Unity 3D, en la que un avatar sirve de ayuda para la correcta colocación de los sensores en el cuerpo del paciente o trabajador. Antes de la captura de movimientos es necesaria una rápida calibración. Una vez iniciada la captura, el avatar representa en tiempo real y de manera precisa todos los movimientos realizados por el usuario generando al finalizar un archivo de datos separados por comas (CSV, comma-separated values).

La última capa de software está compuesta de varios notebooks en Python que permiten realizar cálculos a partir de los datos capturados por los sensores, así como generar gráficas e informes de manera semiautomática. Ésta capa será compartida tanto por nuestros sensores como por los Xsens Dot, analizando para ello los archivos CSV de salida de la aplicación oficial de Android. Sirve así, como capa unificadora, permitiendo una comparación directa de los resultados obtenidos por ambos sistemas.

Finalmente, la información resultante se le ofrece al profesional sanitario de manera que pueda obtener rápidamente una valoración precisa y objetiva de los movimientos del sujeto.

3. Procedimiento

Para llevar a cabo la medición comparativa, se utilizó un goniómetro digital comercial genérico con una resolución de una décima de grado y error de hasta dos décimas, para simular el movimiento de la articulación deseada, y dos sensores inalámbricos de cada tipo colocados a ambos ejes del goniómetro. Los valores obtenidos con el goniómetro servirán como *ground truth* de las variaciones angulares del movimiento a simular en cada experimento.

Los datos se recopilaron utilizando los sensores inalámbricos en cada uno de los posibles ejes de movimiento de la articulación, para comparar los resultados de ambos sistemas entre sí y con el goniómetro. Se utilizó cinta de doble cara para sujetar los vestibles.

3.1 Primer y segundo experimento

El objetivo del primer experimento fue determinar la precisión de las medidas registradas por los sensores inalámbricos de desarrollo propio en comparación con los dispositivos comerciales Xsens Dot, comparando un sensor en posición estática con otro en movimiento (uno de cada tipo) sobre una superficie plana. Las pruebas se realizaron mediante la separación progresiva de uno de los sensores de cada par a pasos de 10º en 10º, del otro extremo del goniómetro en que se colocó el otro sensor. Buscando así incrementos de 10º en cada medida, guiados por la medición en la pantalla del goniómetro digital. Se realizó dicho experimento una vez por cada uno de los ejes de la articulación, sumando un total de 30 mediciones en plano horizontal, llegando a un ángulo máximo en cada una, determinado por el rango articular fisiológico de la articulación a medir.

El objetivo del segundo experimento fue determinar la precisión de la medición sobre una superficie inclinada variable. El propósito era comprobar si el hecho de que los sensores no se encontrasen totalmente paralelos a los ejes de giro afecta a la precisión. Para ello tomamos como base unas inclinaciones de 45º, 60º y 85º del plano en que se encuentra apoyado el goniómetro, en diferentes ejes con respecto al sistema de coordenadas del sensor y se llevaron a cabo las mismas medidas que en el experimento anterior. Así podremos verificar la precisión del sistema en un caso más cercano a la colocación en sujetos de prueba, comprobando si es un método útil para el cálculo del ángulo de una articulación apoyada en un plano no ideal.

3.2 Tercer experimento

En el tercer y último experimento, examinamos la fiabilidad del sistema para la valoración angular de los movimientos articulares, colocando los sensores sobre un sujeto de prueba como vemos en la Figura 1(b) y comparando las mediciones obtenidas con las realizadas por un profesional sanitario cualificado.

Para ello, se evaluaron distintos ángulos de las articulaciones de codo y hombro en los diferentes ejes de movimiento. De este modo, es posible comparar los valores estimados en el ejercicio de goniometría con los detectados por los sensores en articulaciones como el codo, en la que el sensor del brazo se toma como referencia y el del antebrazo realiza un movimiento en cualquiera de los ejes; o la articulación del hombro, comparando en este caso el sensor colocado en el tronco (espalda) con el situado en el brazo.

4. Resultados

4.1 Primer y segundo experimento

El error de los sensores con respecto al valor mostrado por el goniómetro digital, manteniendo la medición del goniómetro por varios segundos, fue 0,62 ° (± 0,06) en el caso de los sensores Xsens Dot y 0,25° (± 0,06) para nuestro prototipo de sensores (***Gráficas 1 y 2***). Observamos una desviación del valor aceptable en el rango de la prueba y muy similar en ambos sensores, ya que según sus hojas de especificaciones ambos aseguran un error de rotación estática menor de dos grados y dinámica menor de tres.



*Precisión del sistema propuesto y sensores comerciales en configuración de dos sensores sobre plano horizontal, **(Gráfica 1)** y sobre planos inclinados 45º, 60º y 85º **(Gráfica 2)**, respecto a las tomas de medidas de 10º a 160º de un goniómetro digital.*

En el segundo experimento, comparando los resultados de ambos sensores sobre una superficie inclinada, se observaron errores de cálculo de ángulo ligeramente superiores que para el caso del plano horizontal. No obstante, no se observaron diferencias significativas entre los errores de medida obtenidos en los diferentes planos inclinados. El error medio calculado para todas las medidas realizadas en los tres planos de inclinación fue de 1,25º (± 0,26) para los Xsens Dot y 0,92° (± 0,17) para nuestros sensores con respecto al valor mostrado por el goniómetro digital.

4.3 Tercer experimento

El error medio entre las mediciones de los sensores y los resultados de goniometría apuntados por el experto ergónomo, llegó hasta los 5° en ambos casos, con variaciones de 2 ° y 2,27° de media con respecto al valor estimado por el médico (***Tabla 1***).

	Médico	Sensores comerciales		Sensores sistema	
	Medida (°)	Medida (°) redondeada	Diferencia (°)	Medida (°) redondeada	Diferencia (°)
Flexión codo (°)	85	85	0	85	0
	100	103	3	98	-2
	120	120	0	116	-4
	130	125	-5	127	-3
	135	131	-4	134	-1
	150	147	-3	147	-3
	165	165	0	164	-1
Flexión hombro(°)	55	56	1	54	-1
	80	80	0	77	-3
	115	116	1	117	2
	160	156	-4	156	-4
Abducción hombro(°)	40	39	-1	42	2
	90	94	4	94	4
	125	129	4	126	1
	140	140	0	137	-3
			2 °		***2,27 °***

Tabla 1. *Variaciones en mediciones realizadas para articulaciones de codo y hombro respecto a la medición realizada por un médico con un goniómetro.*

Los resultados de este experimento muestran una buena precisión potencial, siendo una correcta aproximación para determinar la concordancia entre los sensores inalámbricos y el goniómetro, la herramienta que normalmente se usa clínicamente.

5. Discusión

Las mediciones con los distintos sensores proporcionan una medida a la obtenida con un goniómetro digital o bajo la supervisión de un experto.

Los resultados de nuestro estudio están en consonancia con los obtenidos por sistemas análogos no comerciales, que obtuvieron errores inferiores a tres grados, si bien en pruebas con distinta metodología [6]. Respecto a nuestro caso, los errores resultan muy pequeños en el primer experimento, encontrándonos bajo las circunstancias ideales de funcionamiento de los IMU. En el segundo experimento, observamos desviaciones en las medidas ligeramente más grandes, pero aun así poco significativas respecto al tercer experimento, que se asemeja más a las condiciones reales de funcionamiento de los sensores, colocados sobre un sujeto en movimiento. Esta última prueba comparaba la medida obtenida por los sensores con una evaluación clínica común

que realizan los terapeutas con ayuda de un goniómetro. Se obtuvo una desviación respecto al valor estimado dentro de lo esperable en este tipo de análisis y dada la resolución del goniómetro empleado por el especialista. Una limitación de nuestro estudio es debida a la participación de un único profesional en las medidas. En futuros trabajos para validar nuestro sistema plantearemos estudios comparativos de precisión con mediciones realizadas por un mayor número de profesionales. No obstante, estudios previos demuestran la alta variabilidad en las medidas tomadas por distintos especialistas y los beneficios de utilizar un instrumento como el que proponemos para objetivar y homogeneizar las mediciones [8].

Cabe comentar que nuestro estudio no se ha centrado en demostrar la influencia de la velocidad de los movimientos en la precisión de los sensores. No obstante, varios estudios como el de Zhou y Hu [9] o el de Walmsley y Xu [6] no percibieron ningún cambio significativo en el error asociado con variaciones de velocidad en las medidas angulares, obtenidas durante la utilización de estos sensores en participantes humanos. En concreto, se observó variabilidad en la magnitud y la dirección del error a través de los diferentes ángulos de medición, sin un error sistemático o constante aparente, como en otros estudios que comparaban sensores vestibles con diferentes dispositivos [10]. En el caso de los experimentos que observaban alguna variación pequeña en el error debido a la velocidad, el error máximo alcanzado fue inferior de 3,8° [11], por lo que sigue estando dentro de los valores aceptables.

6. Conclusiones

El propósito de este estudio es realizar una primera validación de un prototipo de sensores vestibles para la adquisición de movimientos corporales. Se realizaron experimentos comparativos de nuestros sensores y sensores comerciales en distintas condiciones experimentales, incluyendo una simulación de valoración goniométrica. Las medidas obtenidas por los sensores comerciales y los desarrollados, eran muy cercanas a los valores determinados por un goniómetro digital y por el especialista, permitiendo constatar la estabilidad de ambos sistemas comparados con los dispositivos generalmente utilizados en el ámbito clínico para este tipo de evaluaciones.

Aún a falta de una validación más exhaustiva, cabe prever que los sensores inalámbricos desarrollados ofrecen suficiente precisión y constituyen una herramienta válida para el análisis de los movimientos corporales y el estudio ergonómico en tiempo real con aplicación en entornos médicos o industriales, entre otros.

No obstante, para completar el análisis del error del presente artículo, se buscará desarrollar un experimento más extenso con un mayor número de sujetos de prueba, que proporcionen una mejor evidencia de la validez del sistema en distintas articulaciones del cuerpo humano y diferentes individuos. Además, estos experimentos servirán de base para la utilización de los sensores desarrollados en otros ámbitos de estudio, para la valoración ergonómica en puestos de trabajo en el ámbito de la medicina preventiva, o para actividades que fomenten el acceso universal a la telerehabilitación.

Referencias

[1] N. Özkaya, D. Leger, D. Goldsheyder, y M. Nordin, *Fundamentals of Biomechanics: Equilibrium, Motion, and Deformation*. Cham: Springer International Publishing, 2017.

[2] I. H. Lopez-Nava y A. Munoz-Melendez, «Wearable Inertial Sensors for Human Motion Analysis: A Review», *IEEE Sens. J.*, vol. 16, n.º 22, pp. 7821-7834, nov. 2016, doi: 10.1109/JSEN.2016.2609392.

[3] F. Porciuncula *et al.*, «Wearable Movement Sensors for Rehabilitation: A Focused Review of Technological and Clinical Advances», *PM&R*, vol. 10, pp. S220-S232, sep. 2018, doi: 10.1016/j.pmrj.2018.06.013.

[4] X. Robert-Lachaine, H. Mecheri, C. Larue, y A. Plamondon, «Validation of inertial measurement units with an optoelectronic system for whole-body motion analysis», *Med. Biol. Eng. Comput.*, vol. 55, n.º 4, pp. 609-619, abr. 2017, doi: 10.1007/s11517-016-1537-2.

[5] M. Adcock *et al.*, «Effects of an In-home Multicomponent Exergame Training on Physical Functions, Cognition, and Brain Volume of Older Adults: A Randomized Controlled Trial», *Front. Med.*, vol. 6, p. 321, ene. 2020, doi: 10.3389/fmed.2019.00321.

[6] C. P. Walmsley *et al.*, «Validation of custom wearable sensors to measure angle kinematics: A technical report», *Health Technol.*, vol. 9, n.º 5, pp. 887-892, nov. 2019, doi: 10.1007/s12553-019-00360-1.

[7] R. Stanzani, P. Dondero, A. Mantero, y M. Testa, «Measurement Accuracy of an Upper Limb Tracking System Based on Two Hillcrest Labs BNO080 IMU Sensors: An Environmental Assessment», *IEEE Sens. J.*, vol. 20, n.º 17, pp. 10267-10274, sep. 2020, doi: 10.1109/JSEN.2020.2992733

[8] C. M. dos Santos, G. Ferreira, P. L. Malacco, G. S. Sabino, G. F. de S. Moraes, y D. C. Felício, «Intra and inter examiner reliability and measurement error of goniometer and digital inclinometer use», *Rev. Bras. Med. Esporte*, vol. 18, n.º 1, pp. 38-41, feb. 2012, doi: 10.1590/S1517-86922012000100008.

[9] Huiyu Zhou y Huosheng Hu, «Reducing Drifts in the Inertial Measurements of Wrist and Elbow Positions», *IEEE Trans. Instrum. Meas.*, vol. 59, n.º 3, pp. 575-585, mar. 2010, doi: 10.1109/TIM.2009.2025065.

[10] D. Álvarez, J. C. Alvarez, R. C. González, y A. M. López, «Upper limb joint angle measurement in occupational health», *Comput. Methods Biomech. Biomed. Engin.*, vol. 19, n.º 2, pp. 159-170, ene. 2016, doi: 10.1080/10255842.2014.997718.

[11] M. El-Gohary y J. McNames, «Human Joint Angle Estimation with Inertial Sensors and Validation with A Robot Arm», *IEEE Trans. Biomed. Eng.*, vol. 62, n.º 7, pp. 1759-1767, jul. 2015, doi: 10.1109/TBME.2015.2403368.

Predicting 3D Photon Interaction in a Hexagonal Positron Emission Tomography Detector: A Deep Learning Approach

Leandro A. Hidalgo-Torres[1], David Pérez-Benito[1], Rigoberto Chil[1], Pedro Macias[1], Juan José Vaquero[1,2]

[1] Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid.
[2] Instituto de Investigación Sanitaria Gregorio Marañón, Madrid.

Abstract

Spatial resolution is considered one of the main limitations of Positron Emission Tomography (PET). Most of the commercial scanners use pixelated array crystal scintillators on their detectors, and its pixel size determines the detector's intrinsic spatial resolution. Alternative developments have exploited the characteristics of monolithic crystal scintillators to further improve spatial, energy and timing resolution. The major challenge of monolithic crystals is the estimation of the gamma photon position of interaction due to the broad scintillation light distribution. In this work, we explore the use of artificial neural networks to estimate the 3D position of the γ-photon interaction in a non-conventional hexagonal detector with a side length of 17.6 mm, 10 mm thick scintillator and 61 hexagonal silicon photomultipliers channels with a side length of 2.25 mm. The selected model was successfully implemented and trained with Monte Carlo simulated data including optical interactions in GATE v8.0. The best results were achieved using a multilayer perceptron architecture with 5 hidden layers of 128 hidden nodes. The obtained 2D and 3D average spatial resolution were 0.65 mm and 0.70 mm. The obtained results with the proposed model confirm that ANN approaches could provide spatial resolution comparable to state of art methods tested on simulated data. Hence, this study supports the suitability of artificial neural networks as a positioning algorithm for PET detectors with monolithic crystal scintillators.

Keywords: Positron Emission Tomography, Detectors, Monolithic Crystal Scintillator, Artificial Neural Network, Spatial Resolution.

1. Introduction

Positron Emission Tomography (PET) is an imaging technique well-known for its use in both clinical diagnosis and research. This modality is mainly used for studying *in-vivo* physiological processes by measuring the concentration of a radiotracer within an organism. In the clinical field, PET is mainly used for tumor detection, tumor progression and study treatment response. Preclinical PET imaging is used as well in the development of new radiotracers and drugs, which are validated using animal studies [1]. One of the main drawbacks of PET is the low spatial resolution (SR) [2]. The actual PET SR performance is lower than other medical imaging modalities like Magnetic Resonance Imaging (MRI) or Computed Tomography (CT). In small-animal PET systems, an SR lower than 1 mm is required to obtain images of similar quality and achieve the same relative SR of a clinical scanner, due to the size differences between small animals and humans. Attempting to improve the SR, researchers have opted for non-conventional geometries [3]. These new geometries must have high sensitivity performance, which depends on geometric and intrinsic detector efficiency.

Regarding PET detectors, their role is to detect efficiently the 511 keV annihilation γ-photons with accurate energy, time and spatial resolutions. They consist of scintillator crystal(s) and photomultipliers. There are two major scintillator arrangements: an array of crystals and monolithic crystal (Figure 1). The use of the former brings some limitations to SR because of crystal and photodetector sizes.

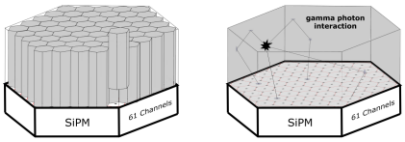


Figure 1. *Differences between PET detector with pixelated scintillator crystals (left) and with monolithic crystal (right)*

Monolithic scintillator crystals present potential advantages, such as intrinsic information about depth-of-interaction (DOI) in the scintillator and higher sensitivity [4–6]. Additionally, monolithic crystals are cheaper than pixelated arrays. Due to previous facts and the availability of new event positioning algorithms, monolithic scintillator crystals have attracted researchers' interest as an alternative to achieve a significant SR improvement in PET systems. Nevertheless, the mapping function between the light distribution measured by the detector and the γ-photon position-of-interaction can be complex and non-linear [4].

To face this problem, researchers have used different algorithms such as support vector machines (SVM), gradient tree boosting (GTB), k-nearest neighbors (k-NN), multilayer perceptron (MLP) or convolutional neural networks (CNN) [4]. The main objective of using these algorithms is to improve the position-of-interaction estimation error of the conventional linear algorithms, such as center-of-gravity (CoG).

In this work, we used Pytorch to implement an MLP model that performs the estimation of the γ-photon position-of-interaction inside a new non-conventional detector geometry based on a monolithic LYSO scintillator crystal and SiPMs. The calibration data for training our model is provided by GATE v8.0, a Monte Carlo simulation modeling tool that has been used and validated in a wide range of PET studies [6]. The prediction model performance is measured using the Euclidean distance (ED) and mean absolute error (MAE) values between estimated and ground-truth position information. The DOI spatial-resolution analysis is performed by determining 2D spatial-resolution in terms of FWHM through 10 "virtual layers" separated by 1mm in the crystal depth.

2. Methods

The detector used in this work consists of a hexagonal back-side SiPM readout and an LYSO scintillator crystal with a side length of 17.61mm and 10mm of thickness, see Figure 2 (right). This topology allows us to have a spherical spatial arrangement of detectors surrounding the subject to be imaged providing a high solid angle which translates into higher geometric efficiency, sensitivity and SR.

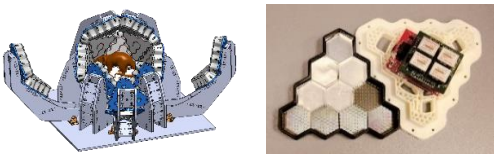


Figure 2. *Face with five hexagonal scintillators and SiPMs (right), parts of a new PET scanner design (left)*

The SiPM array in Figure 2 (left) consists of 61 hexagonal channels with a side length of 2.25 mm. The number of optical-photons detected by each channel per γ-photon deposition and its position-of-interaction is the training data for our position estimation algorithm.

2.1. Monte Carlo simulation data

With GATE v8.0 software we simulated a hexagonal-shaped monolithic LYSO crystal. Surface reflections were modeled with LUT Davis. The upper crystal surface is covered with a Lambertian reflector to bounce photons towards the detector while the sides are wrapped with a polished specular reflector. The source is a pure gamma emitter at 511 keV and it is a fine pencil beam of 0.2 mm width. It was set in the perpendicular direction of the crystal and placed at 181 positions (Figure 3).

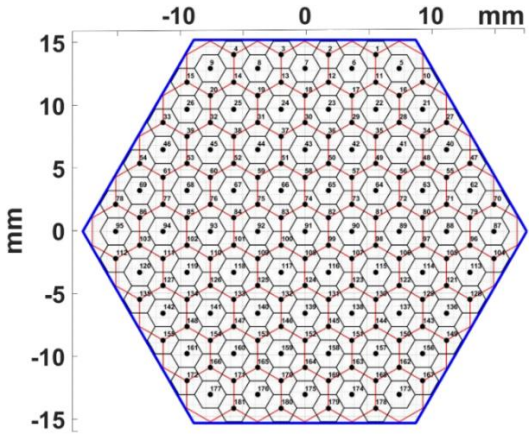


Figure 3. *Calibration positions. Location of the source (black dots), photosensitive pixels (red hexagons) and virtual delimitations (black hexagons)*

2.2. MLP architecture

The selection of the definite MLP number of hidden layers and the number of hidden nodes were selected by a grid-search process. The input layer is composed of 61 nodes. Each of these nodes represents the photon count performed by the corresponding SiPM channel. The output layer has 3 nodes which correspond to the target values: the Cartesian coordinates of the γ-photon position-of-interaction (x,y,z).

ReLU was the activation function used because it facilitates the optimizer convergence and it is one of the most used activation functions in several machine learning applications.

The input (light distribution) and output (position-of-interaction coordinates) data sets are scaled using z-score normalization. It mainly consists of rescaling feature values to have a zero-mean and unitary standard deviation.

2.3. MLP training

For training the proposed model, data was shuffled and were selected 5000 and 3000 events per crystal position for training and validation, respectively. For testing the model performance and generalization, were created 5 different test sets of 5000 events per position. The used optimizers were stochastic gradient descent (SGD) with momentum and adaptive moment estimation (Adam), both with an initial learning rate of 0.001 and *"ReduceLROnPlateau"* learning rate scheduler. Batch normalization with a batch size of 32, dropout and early stopping criteria were the methods employed to ensure a good fit for our model. Mean square error (MSE) was the error or loss function used to measure the training process quality.

The metrics used to measure our model performance in position-of-interaction estimation were ED, MAE, and FWHM, for 2D (x,y) and 3D (x,y,z) coordinates.

3. Results and discussion

The best results were obtained with 5 hidden layers and 128 hidden nodes in terms of prediction error and FWHM. Additionally, there was no difference regarding the used optimizers, both provided positive outcomes.

For our selected MLP model, we obtained a mean ED (bias) of 0.83 mm and 1.10 mm for 2D and 3D respectively. These values represent the average errors made by our algorithm when predicting the position-of-interaction for all the events in the data set.

For 2D, Figure 4 (top) provides the distribution of the Euclidean distances between predicted and real position-of-interaction in the crystal center. For this area, distance median values vary between 0.26 mm and 0.50 mm.

It is important to remark that Figure 3 represents the enumerated irradiation positions over the monolithic crystal and the divisions are only for visualization purposes.

Figure 4 (bottom) shows the distribution of the ED values for the border positions. In contrast to the central region, the median distances range from 0.45 mm to 0.57 mm.

For the case of 3D, Figure 5 shows the same behavior explained previously. There is also more variability in the distance distributions for central positions than for those in the border. The only difference is that median distance intervals increased to 0.54 mm - 0.68 mm and 0.60 mm - 0.73 mm for central and border positions respectively. In this case, 75 percent of the predictions are below 1.4 mm.

Figure 6 (top) shows a profile from position 87 to position 95, where is clear that FHWM values have a decreasing trend toward edges. This can be attributed to the previously analyzed reflection phenomenon on crystal side surfaces which produce higher outliers' rates. In contrast, irradiation positions 87 and 95 located at extremes of the profile, show an increment of FWHM. As both positions

are located exactly at the crystal edge, the SR increment can be attributed to the accumulation of events in those areas that present truncated normal distributions.

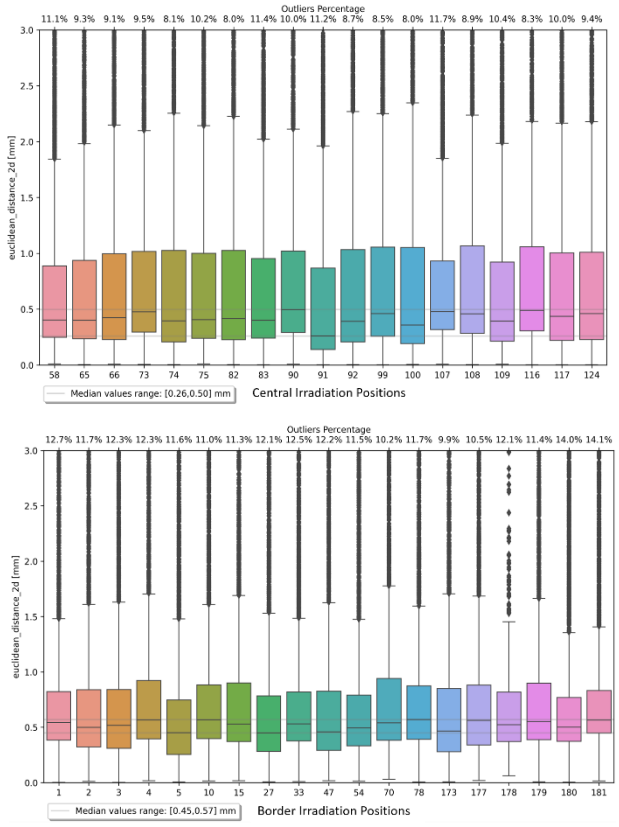


Figure 4. *2D Euclidean distances distribution for central (top) and border (bottom) irradiation positions in the crystal*

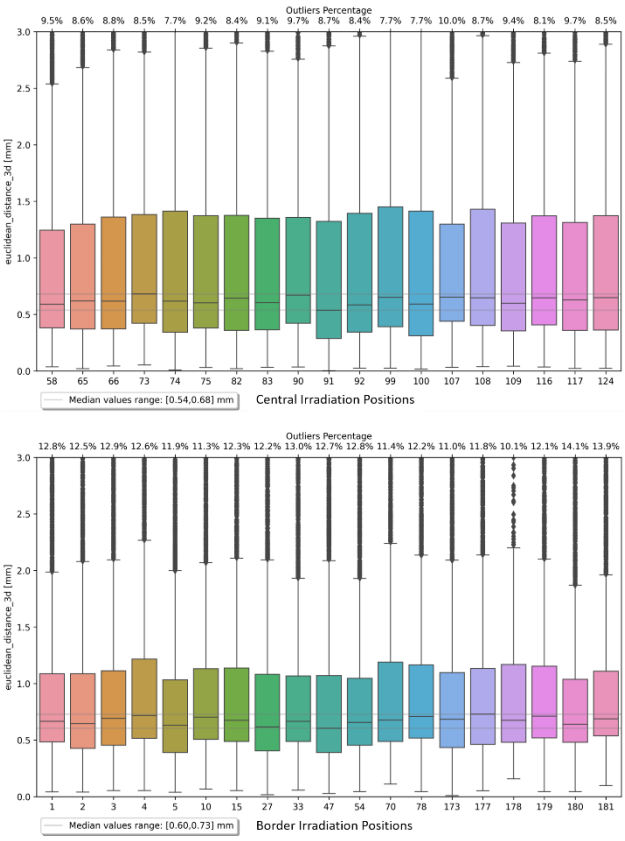


Figure 5. *3D Euclidean distances distribution for central (top) and border (bottom) irradiation positions in the crystal*

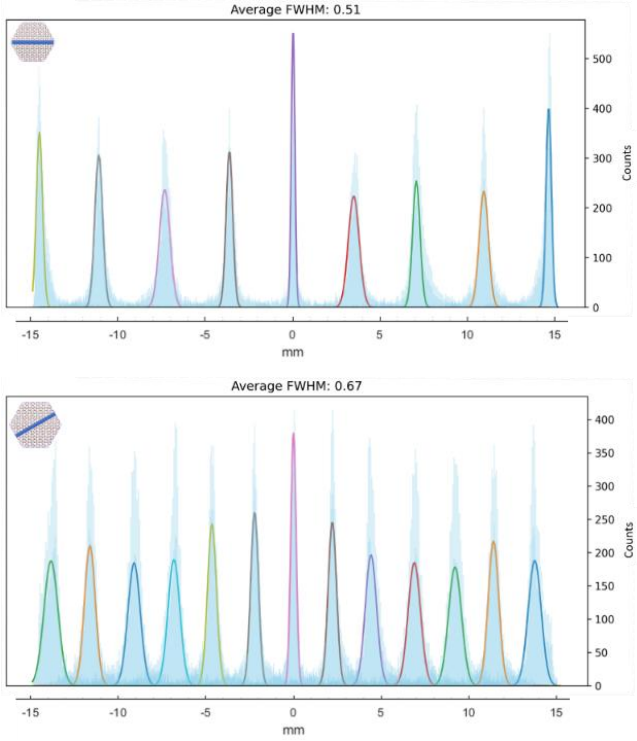


Figure 6. *Profiles for FWHM analysis: (top) from irradiation position 87 to 95, (bottom) from irradiation position 40 to 142*

Figure 6 (bottom) shows the profile from irradiation position 40 to 142, where we can see the same trend of SR degradation toward crystal edges. Moreover, in this profile, extreme positions 40 and 142 do not show the increment of FWHM shown by positions 87 and 95. This is because they are not in the crystal edges. Besides the SR degradation trend, the FWHM values obtained are below 1 mm. The heatmap in Figure 7 represents the 2D FWHM values over the whole crystal. In this image, it is quite clear that SR degradation occurs almost regularly over the entire crystal surface and it is also evident that those positions at crystal edges present higher FWHM values due to the Normal distribution's truncation.

Finally, it is important to remark that the SR degradation toward edges is a common phenomenon present in previous works related to position-of-interaction estimation in monolithic detectors [7,8]. Besides our predictions also present this SR degradation phenomenon, the results obtained in terms of SR are quite relevant. The achieved FWHM values are ranging from 0.30 mm to 1.00 mm over the whole detector, with a mean 2D FWHM of 0.65 mm. Hence, the degradation phenomenon presented in our model predictions should not affect the 2D SR performance of our detector.

Figure 8 shows the evolution of the SR in the whole crystal space. For this purpose, the FWHM was computed at every position from layer_10 (crystal top) to layer_0 (close to photodetectors). From this figure is evident that our algorithm performs correctly for all the layers, except for the layer closer to the photodetector.

Regarding the limitations of this work, Monte Carlo simulations only considered perpendicular irradiation. Adding different incident angles make Monte Carlo simulations more realistic. Additionally, an experimental setup is required to measure the real artificial neural

network (ANN) performance as positioning algorithms for PET detectors.

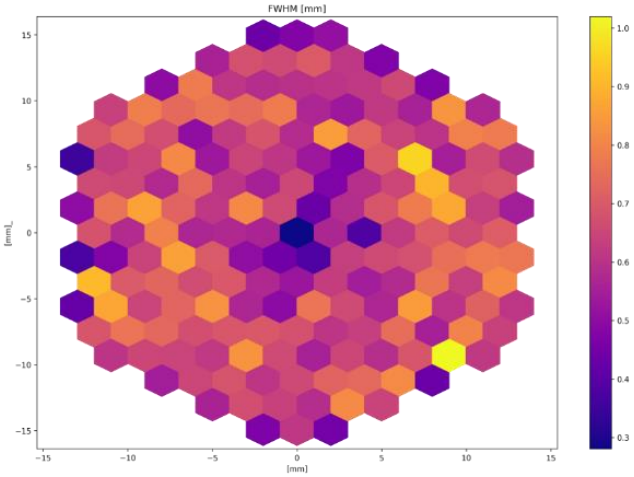


Figure 7. *2D FWHM on crystal top*

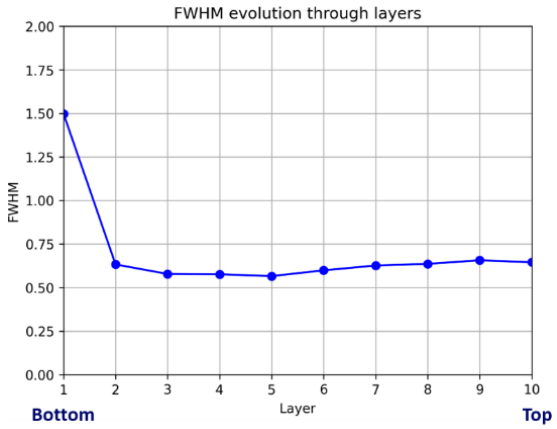


Figure 8. *FWHM evolution through crystal depth*

4. Conclusions

We implemented and trained an ANN to learn the relationship between the light distribution on the photodetectors and the gamma photon position-of-interaction in a non-conventional hexagonal PET detector with a monolithic crystal scintillator. The definite model was obtained through a grid-search process that allowed us to select the best parameters and architectures for this specific problem. The used model for position-of-interaction estimation was an MLP with 5 hidden layers of 128 nodes regularized using dropout and early stopping to avoid overfitting.

With this model, we achieved an average 2D SR of 0.65 mm FWHM and a mean ED of 0.87 mm over the entire crystal. In 3D estimation, we obtained an average 3D SR of 0.70 mm of FWHM and a mean ED of 1.10 mm. Our model provided a nearly uniform positioning performance over the whole crystal surface (2D) and depth (3D). The obtained results confirm that ANN approaches could provide comparable intrinsic SR values to state of art models tested over simulated data. This is mainly due to ANN model predictions being more accurate and precise, which in turn produces a reduction of positioning bias at the crystal edges where the mapping relation between a light distribution and position-of-interaction is more complex.

The performance obtained with our model will be verified with experimental data from a real setup that is being developed. Furthermore, future research should also include additional parameters like different incident angles both at the simulation and at experimental data to push to its limits the performance of the proposed ANN model.

Acknowledgment

Ministerio de Ciencia, Innovación, Proyecto "PID2019-109820RB-I00"; Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación, proyecto "TEC2016-78052-R / MINECO/FEDER, UE".

This project has received funding from 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 organization, Bill & Melinda Gates Foundation and University of Dundee.

Authors would also like to thank Vlaams Supercomputer Centrum and the Medical Image and Signal Processing (MEDISIP) at Ghent University for granting access to run the simulations.

References

[1] S. R. Cherry, J. Sorenson, M. E. Phelps, and B. M. Methé, *Physics in Nuclear Medicine*, vol. 31, no. 8. 2004.

[2] W. W. Moses, "Fundamental limits of spatial resolution in PET," *Nucl. Instruments Methods Phys. Res. Sect. A Accel. Spectrometers, Detect. Assoc. Equip.*, vol. 648, no. SUPPL. 1, pp. S236–S240, 2011.

[3] K. Doroud, Z. Liu, and M. C. S. Williams, "A new geometry of scintillating crystals with strip SiPMs: A PET detector with precise position and time determination," *J. Instrum.*, vol. 14, no. 1, pp. P01016–P01016, Jan. 2019.

[4] K. Gong, E. Berg, S. R. Cherry, and J. Qi, "Machine Learning in PET: From Photon Detection to Quantitative Image Reconstruction," *Proc. IEEE*, vol. 108, no. 1, pp. 51–68, 2020.

[5] J. J. Vaquero and P. Kinahan, "Positron Emission Tomography: Current Challenges and Opportunities for Technological Advances in Clinical and Preclinical Imaging Systems," *Annu. Rev. Biomed. Eng.*, vol. 17, no. 1, pp. 385–414, 2015.

[6] M. Cañadas, P. Arce, and P. Rato Mendes, "Validation of a small-animal PET simulation using GAMOS: A GEANT4-based framework," *Phys. Med. Biol.*, vol. 56, no. 1, pp. 273–288, 2011.

[7] Y. Wang, W. Zhu, X. Cheng, and D. Li, "3D position estimation using an artificial neural network for a continuous scintillator PET detector.," *Phys. Med. Biol.*, vol. 58, no. 5, pp. 1375–1390, 2013.

[8] P. Fan, T. Xu, Z. Lyu, S. Wang, Y. Liu, and T. Ma, "3D positioning and readout channel number compression methods for monolithic PET detector," in *2016 IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop*, 2017.pp.1–4.

Redes de asociación: una nueva forma de explorar la huella fisiopatológica de la enfermedad de Alzheimer

V. Gutiérrez-de Pablo[1], C. Gómez[1,2], J. Poza[1,2,3], A. Maturana-Candelas[1], J. Gómez-Pilar[1,2], M. Revilla-Vallejo[1], R. Hornero[1,2,3]

[1] Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, carlos.gomez@tel.uva.es

[2] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Madrid, España

[3] IMUVA, Instituto de Investigación en Matemáticas de la Universidad de Valladolid, Valladolid, España

Resumen

La enfermedad de Alzheimer (EA) es el tipo de demencia más común y con mayor impacto en los países occidentales. Esta patología presenta diversos factores que condicionan su evolución, como son la edad, el sexo o la genética. Con el objetivo de explorar las complejas relaciones entre los diversos factores que inciden en la progresión de la EA, se ha construido una red de asociación a partir de datos sociodemográficos, variables clínicas, información genética y medidas obtenidas a partir de la actividad neurofisiológica de 190 sujetos, divididos en 5 grupos en el continuo de la enfermedad: 43 controles, 45 sujetos con deterioro cognitivo leve, 43 con EA leve, 37 con EA moderada y 22 con EA severa. Los resultados muestran una asociación entre las medidas obtenidas a partir de la actividad eléctrica cerebral y la severidad de la enfermedad, indicando su utilidad para cuantificar la progresión de la EA. Además, se observa también asociación entre la severidad, la puntuación del test MMSE y el genotipo del gen ApoE, cuya influencia en el deterioro cognitivo característico de la enfermedad no es del todo claro. En conclusión, este nuevo tipo de redes prueban ser de utilidad, pudiendo ayudar a comprender la huella fisiopatológica de la EA.

1. Introducción

La enfermedad de Alzheimer (EA) se caracteriza por la aparición de alteraciones cognitivas, funcionales y conductuales de forma progresiva [1]. La progresión del deterioro cognitivo debido a la EA parece estar asociado a diversos factores, como son la edad, el sexo y la genética, así como riesgos modificables, como el tabaquismo, alcoholismo y contaminación, entre otras [2]. No obstante, el principal factor de riesgo de la EA es la edad, con una prevalencia del 32% en mayores de 85 años. Debido al envejecimiento progresivo de la población mundial, se estima que en 2050 la EA afectará a 152 millones [2]. Por otro lado, la enfermedad afecta a más mujeres que hombres, por lo que el sexo también parece condicionar la aparición de la patología [3].

Diferentes estudios de asociación del genoma completo (GWAS) han identificado diferentes genes relacionados con el riesgo de desarrollar EA, como son *apolipoproteina E* (*APOE*), *phosphatidylinositol-binding clathrin assembly protein* (*PICALM*) y *clusterina* (*CLU*) [4], [5]. Estos genes se relacionan con el riesgo de desarrollar la enfermedad por diferentes caminos, como el metabolismo del colesterol y el sistema inmunitario, entre otros [4], [5]. El efecto de estos genes y su influencia sobre la aparición de la EA ha sido objeto de análisis. Algunos estudios han empleado imagen de resonancia magnética para observar su influencia a nivel fisiológico [6], [7]. Otros estudios han usado técnicas neurofisiológicas, como la electroencefalografía (EEG), para analizar la importancia de estos genes en la progresión de la demencia debida a la EA [8]–[10].

Diversas investigaciones sugieren que los anteriores factores pueden afectar al deterioro de la actividad cerebral [3]–[5], por lo que deben tenerse en cuenta para analizar cómo la demencia debida a la EA altera la dinámica neuronal. En relación con las ideas anteriormente comentadas, en este trabajo se han utilizado redes de asociación para estudiar la relación de todos estos factores de riesgo de la enfermedad con la actividad EEG. Para medir los cambios de la actividad eléctrica neuronal se han utilizado distintas medidas espectrales y no lineales [11], que permiten obtener información sobre las alteraciones en la dinámica neuronal bajo diferentes puntos de vista.

Los objetivos de este trabajo son por tanto: (i) analizar el grado de asociación existente entre datos sociodemográficos, variables clínicas, información genética y medidas espectrales y no lineales de la actividad EEG; y (ii) estudiar si estas asociaciones son capaces de caracterizar la huella fisiopatológica de la EA, tanto en fases previas de la enfermedad, como en su estadio más severo.

2. Materiales

2.1. Sujetos

La base de datos que se ha empleado en el presente estudio está compuesta por 190 sujetos: 43 controles cognitivamente sanos, 45 sujetos con deterioro cognitivo leve (DCL) por EA, 43 pacientes con demencia leve por EA, 37 pacientes con demencia moderada por EA y 22 pacientes con demencia severa por EA. Para el diagnóstico de DCL y de demencia debido a EA se han seguido los criterios del *National Institute of Aging and Alzheimer's Association* (NIA-AA) [12]. Los sujetos de control eran personas cognitivamente sanas, de edad superior a los 65 años y sin patologías neurológicas en su historial clínico. El deterioro cognitivo de cada sujeto se evaluó con el test *Mini-Mental State Examination* (MMSE) [8]. La Tabla 1 presenta los datos sociodemográficos y clínicos de los sujetos participantes.

ISBN: 978-84-09-25491-0, pags. 405- 408

Todos los sujetos, sus cuidadores y/o familiares dieron su consentimiento por escrito para el estudio, que ha sido aprobado por el Comité Ético de la Universidad de Oporto, Portugal.

2.2. Análisis genético

El análisis del genoma se ha llevado a cabo extrayendo muestras de saliva de todos los sujetos con el kit *Oragene DNA* (*OG-500*). Por otra parte, para los sujetos en una fase más avanzada de la enfermedad se emplearon bastoncillos de algodón esterilizado. El ADN contenido se extrajo de las muestras con el kit de extracción *prepIT DNA*, mientras que para los bastoncillos se empleó el kit de extracción *Citogene*. Los alelos del gen *ApoE* se extrajeron de los polimorfismos de nucleótido único (SNP) rs74389 y rs7412, mientras que para los genes *PICALM* y *CLU*, los alelos se extrajeron de los SNPs rs3851179 y rs11136000, respectivamente.

2.3. Registros electroencefalográficos

Las señales EEG se registraron empleando un sistema de 19 canales (Nihon Kohden Neurofax JE-921A), utilizando una frecuencia de muestreo de 500 Hz. Los electrodos fueron posicionados siguiendo el sistema internacional 10-20: Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz y Pz. Se registraron 5 minutos de actividad en estado de reposo mientras los sujetos permanecían despiertos, relajados y con los ojos cerrados en un ambiente libre de ruidos para minimizar la presencia de artefactos. Posteriormente, se llevó a cabo un preprocesado de las señales siguiendo los siguientes pasos [11]: (i) eliminación de la media; (ii) filtrado a 50 Hz para eliminar la componente eléctrica; (iii) filtrado paso banda de 0.5 Hz a 70 Hz con ventana de Hamming; (iv) análisis de componentes independientes (ICA) para eliminar componentes ruidosas; y (v) segmentación de las señales en épocas de 5 s y rechazo visual de aquellas artefactuadas.

3. Métodos

Se calcularon dos parámetros espectrales y dos no lineales a partir de las señales EEG para, posteriormente, calcular la red de asociación de todos los datos existentes. A continuación, se detallan las características de la metodología empleada.

3.1. Procesado de la señal EEG

Se han estimado diferentes parámetros espectrales y no lineales a partir de los registros EEG obtenidos [11]. En concreto, se calculó un valor de estos parámetros por canal y época para, posteriormente, promediar estos valores hasta obtener un único valor por sujeto.

1) Frecuencia mediana (MF). La MF se define como la frecuencia a la que se alcanza el 50 % de la potencia de la densidad espectral de potencia (DEP). Permite resumir todo el contenido espectral de la DEP en un único valor, lo cual es muy útil a la hora de caracterizar la lentificación progresiva de la actividad cerebral en la EA [11].

2) Frecuencia alfa individual (IAF). La IAF se define como la frecuencia en la que se observa el máximo de potencia en la banda alfa. Este máximo aparece debido a las oscilaciones alfa dominantes que presentan los adultos en estado de reposo. A medida que la demencia debida a EA progresa, estas oscilaciones alfa van desapareciendo y ralentizándose, fenómeno que esta medida permite cuantificar [11]. En este trabajo, se ha calculado como la frecuencia en la que se obtiene el valor mediano de la DEP en la banda alfa extendida (4-15 Hz) [11].

3) Complejidad de Lempel-Ziv (LZC). Esta medida se emplea para estimar la complejidad de una serie temporal, proporcionando valores mayores a aquellas señales que muestran un comportamiento más complejo [11]. Para el cálculo de la LZC se emplea un umbral T_d, definido como la mediana de la amplitud de la señal, debido a su robustez frente a valores espurios [11]. La LZC resulta de utilidad para medir la pérdida de complejidad que sufre la actividad cerebral con el progreso de la demencia por EA [11].

4) Entropía muestral (SampEn). Esta medida se utiliza para medir la irregularidad de una serie temporal, proporcionando valores mayores a señales más irregulares. Para poder estimarla, se requiere configurar dos parámetros: la longitud de la secuencia *m* y la ventana de tolerancia *r*. Se mide la probabilidad de que dos secuencias similares en *m* observaciones lo sigan siendo en la observaciones sucesivas, con una diferencia menor que la tolerancia *r*, sin tener en cuenta autocomparaciones [11]. La SampEn ha resultado útil para estimar la pérdida de irregularidad que sufre la actividad neuronal con la progresión de la demencia por EA [11].

3.2. Redes de asociación

Las redes de asociación se construyen en base en las relaciones directas entre diferentes variables [13]. Para este estudio se analizaron todos los sujetos de forma conjunta, relacionando variables sociodemográficas (edad y sexo), información clínica (severidad de la enfermedad y puntuación del test MMSE), información genética (genes

Datos	Controles	DCL	EA Leve	EA Moderada	EA Severa
Sujetos (M:H)	43 (22:21)	45 (31:14)	43 (26:17)	37 (31:6)	22 (19:3)
Edad (años)	79,7 (7,4)	84,8 (7,1)	80,1 (6,7)	80,9 (8,1)	80,3 (6,6)
MMSE	28,8 (1,1)	23,2 (2,9)	22,7 (2,3)	13,7 (2,9)	3,9 (4,2)
***ApoE* (A:B:C)**	0:6:37	0:9:36	2:15:26	2:13:22	4:9:9
***PICALM* (A:B:C)**	3:21:19	8:22:15	7:19:17	2:17:18	2:13:7
***CLU* (A:B:C)**	5:14:24	5:21:19	7:22:14	5:13:19	2:10:10

Tabla 1. *Datos sociodemográficos y clínicos de la base de datos utilizada. Los valores se representan como: media (desviación típica). M: mujer; H: hombre; MMSE: Mini-Mental State Examination (rango: [0 30]); A: sujetos homocigóticos de riesgo o con protección; B: sujetos heterocigóticos; C: sujetos homocigóticos normales.*

ApoE, *PICALM* y *CLU*), y parámetros neurofisiológicos espectrales (IAF, MF) y no lineales (LZC, SampEn) calculados a partir del EEG. Se aplicó el test Chi-Cuadrado para comparar variables categóricas. Aquellas variables que inicialmente no eran categóricas se transformaron mediante deciles. Para la generación de las redes, se utilizaron tanto los valores del estadístico como los *p*-valores correspondientes.

Para la representación de las redes, se utilizó el software *Gephi,* mediante el algoritmo *Force Atlas 2*, que utiliza el grado de nodo como efecto repulsor; es decir, los nodos que posean mayor número de conexiones ejercerán una menor repulsión hacia el resto de nodos que aquellos que tengan menor número de conexiones. Por otra parte, la anchura del enlace representa el grado de asociación entre los dos nodos que une, de forma que los nodos que tengan mayor asociación estarán más cerca y su enlace será más grueso.

4. Resultados y discusión

La Figura 1 muestra la red de asociación, en la que se muestran las relaciones significativas de todos los grupos involucrados en el estudio. Se puede apreciar como la LZC y la SampEn muestran la asociación más significativa, puesto que los conceptos de complejidad e irregularidad son muy similares. Estos dos parámetros de actividad cerebral aparecen también asociados con la IAF y la MF. A su vez, las cuatro medidas de EEG están asociadas a la severidad. Esto indica que su uso podría ser de utilidad para estudiar la progresión de la EA caracterizando el enlentecimiento progresivo del espectro y la pérdida de complejidad e irregularidad de la actividad EEG provocado por la misma [11]. Por otra parte, se observa una estrecha asociación entre el grado de severidad y la puntuación del test MMSE, indicando la utilidad que presenta este test a la hora de evaluar el deterioro cognitivo que sufren los pacientes durante la progresión de la enfermedad. Sin embargo, también aparecen asociados a estas dos variables el sexo y la edad. Aunque tienen un bajo nivel de asociación, esto podría deberse a que los datos no están totalmente pareados en base a esas dos variables.

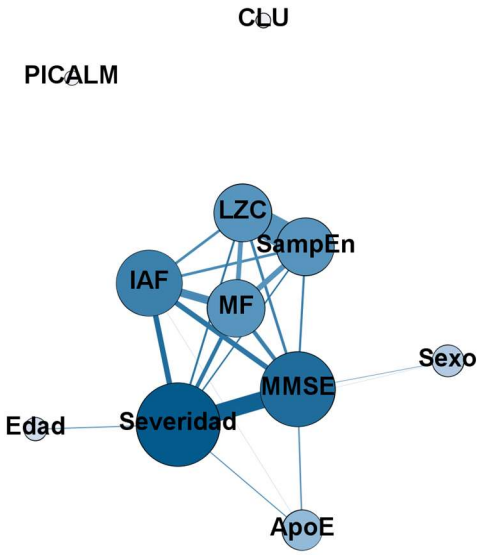


Figura 1. *Red de asociación, en la que aparecen las relaciones significativas entre variables sociodemográficas, datos clínicos, genes y parámetros neurofisiológicos espectrales y no lineales de todos los grupos de estudio (controles, pacientes con DCL y pacientes con EA en los diferentes estadios de la enfermedad).*

Por otro lado, otro aspecto reseñable en la Figura 1 es la asociación de la severidad y la puntuación del test MMSE con el gen *ApoE*. Estudios previos sugieren que el gen *ApoE* posee un mayor impacto en etapas previas a la enfermedad; sin embargo, esta influencia se va perdiendo una vez que la enfermedad se ha desarrollado [6]–[8]. También puede observarse una ligera asociación entre el gen *ApoE* y la IAF. A este respecto, Babiloni *et ál.* [14] analizaron la influencia del gen *ApoE* en la actividad eléctrica cerebral a nivel de fuente, y observaron que la actividad alfa en los pacientes con DCL y demencia por EA es menor en los portadores del alelo de riesgo que en los no portadores, si bien no encontraron asociación ninguna con la IAF. Es por este motivo que se sugiere que esta asociación encontrada se deba a que el efecto del gen en la IAF es muy leve. Por el contrario, los genes *PICALM* y *CLU* son las únicas variables que no poseen asociación ninguna con el resto de los parámetros. Trabajos previos estudiaron la influencia de estos dos genes en la actividad eléctrica cerebral. En concreto, Ponomareva *et ál.* [9] analizaron la asociación de la actividad EEG de sujetos sanos con el genotipo del gen *PICALM*, encontrando que los sujetos con el alelo *A* protector presentaban una menor latencia en el P300 que los sujetos no portadores, acentuándose este efecto con la edad. En otro estudio, Ponomareva *et ál.* [10] investigaron la influencia del gen *CLU* y encontraron que aquellos sujetos sanos con el alelo protector *T* presentaban una menor amplitud en los ritmos alfa superiores (11-13 Hz). Cabe destacar que en ambos estudios se eliminó la influencia del genotipo *ApoE* en los resultados. En nuestro caso, al querer analizar las asociaciones de los 3 genes, es posible que la influencia del gen *ApoE* sea más relevante que la de los genes *PICALM* y *CLU*. Esto podría justificar que en nuestros resultados los genes anteriores no posean asociaciones significativas con el resto de variables ya que, aunque la influencia de estos genes en la patogénesis de la EA ha sido estudiada, esta es más leve que la proporcionada por el gen *ApoE* [4].

Con el objetivo de analizar más detalladamente la asociación entre la puntuación en el MMSE, el gen ApoE y la severidad de la enfermedad, en la Figura 2 se han representado los valores del test MMSE para sujetos que no posean el alelo ε4 del gen *ApoE* y para sujetos que posean al menos 1 alelo ε4. Se puede observar, para los 4 grupos patológicos, la existencia de diferencias en los valores medios del MMSE entre sujetos no portadores y sujetos portadores del alelo de riesgo del gen *ApoE*. Aunque este hecho sea llamativo, se requieren de nuevas herramientas y más investigación para poder esclarecerlo.

Finalmente, existen algunas limitaciones en el presente estudio que deben considerarse. En primer lugar, la base de datos es limitada; por ello, sería necesario aumentar el número de sujetos. Por otra parte, se puede observar en la Figura 1 que la edad y el sexo están levemente asociadas a

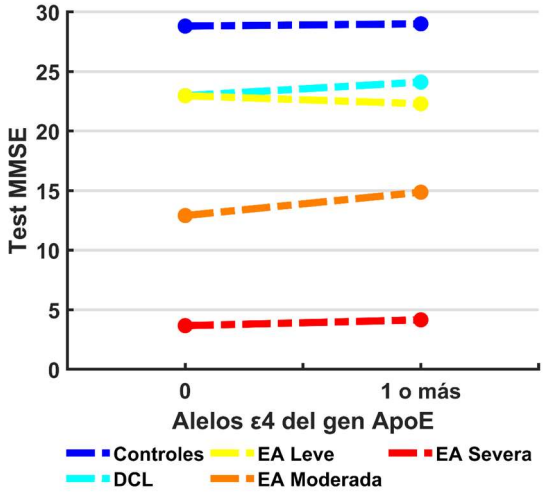


Figura 2. *Puntuación media del test MMSE en función del grupo y genotipo de ApoE.*

la severidad y al MMSE, lo que podría estar sesgando los resultados. Para solucionarlo, sería necesario balancear las cinco poblaciones bajo estudio. Por último, este estudio constituye un primer paso para demostrar la utilidad de las redes de asociaciones o correlacionales. Sin embargo, con el objetivo de proporcionar una huella más detallada de la evolución de la EA, es necesario introducir nuevas variables, como medidas de red obtenidas a partir del EEG, nuevos datos sociodemográficos y clínicos (como hábitos alimenticios, patologías previas, consumo de tabaco y alcohol, etc.), así como otros genes relacionados con la EA analizados en la literatura.

5. Conclusiones

Los resultados obtenidos muestran la utilidad de las medidas espectrales y no lineales para cuantificar el deterioro cognitivo que provoca la evolución de la EA, dada su importante asociación con la variable de severidad de la enfermedad y la puntuación del MMSE. Asimismo, se ha observado la asociación existente entre el gen *ApoE* y la severidad de la enfermedad, lo que podría indicar la influencia de este gen en el deterioro producido.

Agradecimientos

Este estudio ha sido financiado por los proyectos 'Análisis y correlación entre el genoma completo y la actividad cerebral para la ayuda en el diagnóstico de la enfermedad de Alzheimer' y 'Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres' (Programa de Cooperación Transfronteriza INTERREG V-A España-Portugal 2014-2020) de la Comisión Europea y del Fondo Europeo de Desarrollo Regional (FEDER), por el proyecto PGC2018-098214-A-I00 del 'Ministerio de Ciencia e Innovación – Agencia Estatal de Investigación' y FEDER, y por el CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto de Salud Carlos III cofinanciado con fondos de FEDER.

Referencias

[1] J. L. Cummings *et ál.*, *La neuropsiquiatría de la enfermedad de Alzheimer y demencias relacionadas*. 2004.

[2] G. Livingston *et ál.*, "Dementia prevention, intervention, and care: 2020 report of the Lancet Commission," *Lancet*, vol. 396, no. 10248, pp. 413–446, 2020.

[3] B. C. Riedel *et ál.*, "Age, APOE and sex: Triad of risk of Alzheimer's disease," *J. Steroid Biochem. Mol. Biol.*, vol. 160, no. 1, pp. 134–147, Jun. 2016.

[4] C. M. Karch *et ál.*, "Alzheimer's disease risk genes and mechanisms of disease pathogenesis," *Biol. Psychiatry*, vol. 77, no. 1, pp. 43–51, 2015.

[5] K. Bettens *et ál.*, "Genetic insights in Alzheimer's disease," *Lancet Neurol.*, vol. 12, no. 1, pp. 92–104, 2013.

[6] S. Y. Bookheimer *et ál.*, "Patterns of brain activation in people at risk for Alzheimer's disease," *N. Engl. J. Med.*, vol. 343, no. 7, pp. 450–456, 2000.

[7] S. Duke Han *et ál.*, "Verbal paired-associate learning by APOE genotype in non- demented older adults: fMRI evidence of a right hemispheric compensatory response," *Neurobiol. Aging*, vol. 28, no. 2, pp. 238–247, 2007.

[8] V. Gutiérrez-De Pablo *et ál.*, "Relationship between the presence of the ApoE ε4 allele and EEG complexity along the Alzheimer's disease continuum," *Sensors (Switzerland)*, vol. 20, no. 14, pp. 1–17, 2020.

[9] N. V. Ponomareva *et ál.*, "Genetic Association between Alzheimer's Disease Risk Variant of the PICALM Gene and Auditory Event-Related Potentials in Aging," *Biochem.*, vol. 83, no. 9, pp. 1075–1082, 2018.

[10] N. Ponomareva *et ál.*, "Age-dependent effect of Alzheimer's risk variant of CLU on EEG alpha rhythm in non-demented adults," *Front. Aging Neurosci.*, vol. 5, pp. 1–10, 2013.

[11] S. J. Ruiz-Gómez *et ál.*, "Automated multiclass classification of spontaneous EEG activity in Alzheimer's disease and mild cognitive impairment," *Entropy*, vol. 20, no. 1, pp. 1–15, 2018.

[12] C. R. Jack *et ál.*, "NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease," *Alzheimer's Dement.*, vol. 14, no. 4, pp. 535–562, Apr. 2018.

[13] D. Borsboom *et ál.*, "The Small World of Psychopathology," *PLoS One*, vol. 6, no. 11, 2011.

[14] C. Babiloni *et ál.*, "Apolipoprotein E and alpha brain rhythms in mild cognitive impairment: A multicentric electroencephalogram study," *Ann. Neurol.*, vol. 59, no. 2, pp. 323–334, 2006.

Comparación de Redes Neuronales Preentrenadas para la Evaluación de la Calidad en Registros Electrocardiográficos

A. Huerta Herraiz [1], A. Martínez Rodrigo[1], A. Puchol Calderón[2], M.I. Pachón Iglesias[2], J.J. Rieta Ibáñez[3], R. Alcaraz Martínez[1]

[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] Departamento de Arrítmias Cardíacas, Hospital Virgen de la Salud, Toledo, España, apucholcal@gmail.com, mpachon@secardiologia.es
[3] BioMIT.org, Dep. de Ingeniería Electrónica, Universitat Politecnica de Valencia, Valencia, España, jjrieta@upv.es

Resumen

Actualmente, es habitual el uso de las redes neuronales convolucionales para el análisis del registro electrocardiográfico (ECG). Sin embargo, su capacidad para evaluar la calidad con la que se registra esta señal fisiológica todavía no se ha examinado a fondo. Esta difícil tarea, cuyo objetivo es identificar los intervalos de ECG de baja calidad, podría ayudar a un diagnóstico más preciso de enfermedades cardíacas, especialmente si el ECG se ha registrado con un sistema de monitorización a largo plazo portátil o vestible. Así pues, en este trabajo se introduce un estudio comparativo en el que se analizan varios modelos de redes neuronales en la tarea de discernir entre intervalos de ECG de alta y de baja calidad. Al no existir bases de datos suficientemente extensas, anotadas, y con acceso público, se comparan tres redes neuronales preentrenadas, las cuales son AlexNet, GoogLeNet y VGGNet. Para ello, se han utilizado 2.000 segmentos de ECG de 5 segundos de duración, y clasificados en dos grupos de 1.000 segmentos cada uno. Se ha obtenido una validación robusta de los algoritmos realizando cinco procesos de entrenamiento y dividiendo el conjunto de datos de manera aleatoria, con un 80 % para entrenamiento y un 20 % para test. Los resultados muestran una exactitud generalizada en todos los modelos en torno a un 90 %. Sin embargo, en cuanto al tiempo de cálculo y memoria utilizada se han observado importantes diferencias. AlexNet ha mostrado ser el método más rápido para clasificar los segmentos de ECG, mientras que GoogLeNet ha requerido menos memoria. Por lo tanto, ambos algoritmos han demostrado un buen equilibrio entre la identificación de intervalos de ECG y carga computacional.

1. Introducción

La reciente evolución de las redes neuronales convolucionales ha permitido su uso en una amplia variedad de aplicaciones en campos tan diversos como la salud, el deporte, la robótica, etc. [1]. En el ámbito de la medicina se han utilizado en gran medida para la interpretación y el análisis de las señales electrocardiográficas (ECG) [2]. De hecho, algunas de las aplicaciones novedosas en las que ya se están utilizando son la clasificación de arritmias, la detección automática de fibrilación auricular y la identificación de personas a través de biometría, entre otras [2]. Sin embargo, una aplicación muy interesante en el cual apenas se han utilizado hasta el momento es en la evaluación de la calidad con la que se registra el ECG.

Lamentablemente, hasta el momento la presencia de niveles de ruido elevados en el ECG dificulta su interpretación, limitando así el posterior diagnóstico clínico [3]. Además, la presencia de artefactos de movimiento y otras interferencias provocan muchas falsas alarmas en sistemas de monitorización de ECG en tiempo real [4]. La presencia de ruido y artefactos también es especialmente problemática para los dispositivos de ECG vestibles o portátiles, los cuales tienen la difícil tarea de monitorizar y almacenar señales de ECG durante varias semanas e incluso meses, en los que el paciente sigue realizando sus tareas cotidianas [5]. Así pues, estos dispositivos tendrán que trabajar en entornos muy ruidosos y cambiantes, de forma que las señales captadas estarán ampliamente corrompidas por artefactos de movimiento, interferencias de la red eléctrica, desplazamiento de la línea de base, y ruido de alta frecuencia [5]. Por lo tanto, en estos dispositivos la identificación automática de los tramos de ECG de alta calidad no contaminados por ruido es muy importante, con el fin de extraer de manera fiable la información clínica [3].

Para acometer este objetivo se han presentado diversos algoritmos durante los últimos años [3]. La mayoría de ellos están basados en la detección de puntos fiduciales y eventos morfológicos de la señal de ECG, y en el posterior análisis de parámetros como la media entre intervalos RR, la consistencia en el tiempo de las formas de onda PQRST, la coherencia de los complejos QRS, etc. [3]. Sin embargo, en segmentos de ECG muy ruidosos, no siempre es posible una detección exacta de los puntos fiduciales, así como tampoco lo es en señales limpias que presentan formas de onda muy cambiantes u ondas T muy puntiagudas y de gran amplitud [6]. Estos aspectos son comunes en registros de ECG de larga duración, por lo que en este contexto

ISBN: 978-84-09-25491-0, pags. 409- 412

los métodos presentados hasta la fecha reducen mucho su rendimiento [3]. Por lo tanto, existe una fuerte demanda de algoritmos novedosos que no estén basados en características morfológicas de la señal y que puedan trabajar satisfactoriamente con los registros de larga duración adquiridos por dispositivos vestibles [3].

Como respuesta a esta necesidad, la alternativa más viable en la actualidad son las redes neuronales convolucionales. Éstas son capaces de extraer las características más importantes del ECG sin la necesidad de delinear puntos fiduciales, y sin requerir además ningún tipo de intervención externa [7]. Sin embargo, hasta la fecha solamente existen unos pocos trabajos que han utilizado este tipo de algoritmos para evaluar la calidad del registro de ECG [8]. Desafortunadamente, estos han entrenado modelos desde cero con un número muy reducido de muestras y, por lo tanto, podrían estar sobreentrenados. Además, teniendo en cuenta que no hay muchas bases de datos de ECG públicas con anotaciones que sean suficientemente largas para entrenar una red desde cero [9], la utilización de redes preentrenadas es una opción muy interesante. En este caso, se emplea el concepto de *aprendizaje de transferencia*, para aprovechar el conocimiento adquirido por una red y en una determinada tarea, ahorrando tiempo de aprendizaje al solo requerirse un pequeño entrenamiento específico. A este respecto, en un reciente estudio piloto ya se ha evaluado la calidad del ECG empleando una red preentrenada, la cual ha obtenido un mayor rendimiento que otros modelos entrenados desde cero [10]. Por lo tanto, el presente trabajo tiene como objetivo comparar el rendimiento de varias redes neuronales convolucionales preentrenadas para discernir entre segmentos de ECG de alta y baja calidad.

2. Base de datos

La base de datos empleada en el presente trabajo fue del conjunto de registros de ECG propuesto para entrenamiento en el PhyisoNet/CINC Challenge 2017 [11]. Se trata de un conjunto público de 8.528 registros de ECG de una sola derivación, y una duración de entre 9 y 60 segundos. Se adquirieron utilizando un dispositivo portátil de la marca AliveCor ^{TM} conectado a un teléfono móvil, con una frecuencia de muestreo de 300 Hz y 16 bits de resolución sobre un rango dinámico de 5 mV.

Para entrenar y validar cada uno de los modelos preentrenados que se analizan, las señales de ECG se dividieron en segmentos de 5 segundos, y se agruparon en dos categorías. En primer lugar, los episodios de fibrilación auricular, ritmo sinusal y otros ritmos se consideraron señales de alta calidad, mientras que los segmentos de ruido se consideraron de baja calidad. Debido al gran desequilibrio entre ambos grupos (47.439 segmentos de alta calidad frente a 1.168 intervalos de baja calidad, ver Tabla 1), se creó un conjunto de datos específico tomando 1.000 fragmentos de ECG de cada clase. En el grupo de alta calidad se mantuvo la representatividad de todos los ritmos cardíacos, escogiendo 340 fragmentos de ritmo sinusal, 330 fragmentos de fibrilación auricular y 330 fragmentos de otros ritmos.

Ritmo	Registros	Intervalos de 5 s.
Ritmo sinusal	5.154	28.413
Fibrilación auricular	771	4.329
Otros ritmos	2.557	14.697
Ruido	46	1.168

Tabla 1. *Número de registros y de intervalos de 5 segundos para cada ritmo pertenecientes a la base de datos de PhysioNet/CinC Challenge 2017 [11].*

3. Métodos

3.1. Transformación de los segmentos de ECG

La mayoría de redes preentrenadas se caracterizan por recibir una señal en dos dimensiones como entrada. Por lo tanto, cada segmento de ECG de 5 segundos se convirtió a una representación tiempo-frecuencia, empleando para ello la transformada wavelet continua (TWC). Esta transformación es especialmente adecuada para señales no estacionarias, como lo es el ECG [12]. La representación gráfica de los coeficientes wavelet obtenidos mediante esta herramienta es conocido como escalograma y se empleó como entrada para las redes analizadas. En este caso, la TWC se realizó con una wavelet madre llamada *Morlet* y se emplearon 48 voces por octava para determinar el número de escalas.

3.2. Arquitecturas preentrenadas

En el presente trabajo se evaluaron tres arquitecturas preentrenadas muy conocidas, tales como son AlexNet, VGGNet y GoogLeNet. AlexNet se hizo muy popular por ganar una competición de identificación automática de imágenes en el año 2012 [13], y es probablemente la arquitectura más conocida dentro de las redes neuronales convolucionales. De hecho, este algoritmo se ha utilizado en una gran variedad de aplicaciones [1]. Se compone de ocho capas con capacidad de aprendizaje, de las cuales cinco son convolucionales y tres son totalmente conectadas. Además, también incluye dos capas de agrupación o *pooling* para reducir las dimensiones espaciales del mapa de características, sin perder información relevante. Como función de activación utiliza una función lineal en todas las capas convolucionales y en las totalmente conectadas. También incluye dos funciones de *dropout* para reducir el sobreentrenamiento después de las dos primeras capas totalmente conectadas.

Por otro lado, VGGNet se ha hecho muy popular debido a su sencillez y al empleo de convoluciones de pequeño tamaño [14]. La principal aportación de esta red es el uso de muchas capas convolucionales con filtros de pequeño tamaño para poder lograr redes muy profundas, las cuales suelen obtener un buen rendimiento. Existen diferentes variantes dentro de VGGNet con diferentes niveles de profundidad, pero la alternativa escogida en este trabajo es VGG16 por ser una de las más comunes. Este modelo emplea filtros de convolución de tamaño 3×3 en una arquitectura que se compone de trece capas convolucionales, obteniendo una eficiencia muy alta [14]. Todas las capas de convolución van seguidas por funciones de activación lineales. Además, en el centro de la red se emplean cinco capas de agrupación para extraer características. Por

último, el modelo se completa con tres capas totalmente conectadas, las cuales van seguidas de funciones de activación y *dropout*.

Por último, GoogLeNet [15] está basado en un diseño totalmente diferente a los dos modelos anteriores. Mientras que AlexNet y VGG16 se basan en una estructura de capas colocadas de manera secuencial, GoogLeNet está diseñado mediante una compleja estructura con diferentes ramas en paralelo. Concretamente, esta red utiliza el *Módulo Inception* como base de su modelo. Este módulo se compone de una serie de capas de convolución y de agrupación, las cuales trabajan en paralelo combinando las características extraídas. Conectando nueve *Módulos Inception* en cascada junto con otras capas y funciones necesarias, GoogLeNet consigue una arquitectura con una profundidad de veintidós capas y más de cien bloques computacionales.

3.3. Evaluación del rendimiento

Para el entrenamiento y validación de cada uno de los modelos preentrenados, se utilizó un enfoque basado en repeticiones. Así pues, en 5 iteraciones consecutivas la base de datos se dividió, empleándose el 80 % para entrenamiento y el 20 % para test. El proceso de re-entrenamiento de la red se realizó con un algoritmo de descenso de gradiente estocástico con un momento de 0,9 y una tasa de aprendizaje de 0,0001, fijados de manera experimental. En todos los casos se consideraron 10 épocas con un tamaño de lote de 10 muestras. Además, en todos los modelos se adaptó la capa de salida de la red para distinguir solamente dos clases: alta y baja calidad.

El rendimiento de las diferentes redes discerniendo entre intervalos de ECG de alta y baja calidad pertenecientes al conjunto de test se evaluó en términos de sensibilidad (Se), especificidad (Es) y exactitud (Ex). Además, para las cinco iteraciones se obtuvo la media, desviación estándar (std), y los valores máximo y mínimo de estos índices. La sensibilidad se definió como la tasa de segmentos de ECG de alta calidad correctamente clasificados, la especificidad como el porcentaje de intervalos de baja calidad correctamente identificados, y por último, la exactitud, como el número total de fragmentos de ECG correctamente detectados. Además, para comparar el comportamiento de los algoritmos, también se analizó el tiempo, el uso de la CPU y la memoria requerida durante la fase de validación. Los experimentos se desarrollaron en una estación de trabajo all-in-one HP ProOne 600 G2 con un procesador Intel i7 a 3,41 GHz, utilizando Matlab R2019a. Cabe mencionar que también se obtuvieron los valores medio, estándar, máximo y mínimo de estas medidas para las cinco iteraciones conducidas.

4. Resultados

En la Tabla 2 se observa una exactitud media similar en todos los modelos, oscilando ligeramente entre el 89,65 % para VGG16 y el 90,75 % para GoogLeNet. Además, los resultados medios de sensibilidad y especificidad se encuentran bastante balanceados para los tres modelos, existiendo diferencias de tan solo un 4–8 % entre ambas métricas. Sin embargo, en los tres casos se observa una mayor capacidad para identificar segmentos de ECG de baja calidad, siendo los valores de especificidad y sensibilidad alrededor del 92 % y el 88 %, respectivamente. Por último, no se aprecia una gran dispersión en los tres índices para las cinco iteraciones realizadas, ya que los valores de std son inferiores al 4 % y las diferencias entre los valores máximos y mínimos inferiores al 10 % para la mayoría de los casos.

Modelo	Valor	Se. (%)	Es. (%)	Ex. (%)
AlexNet	Media	88,90	92,50	90,70
	Std	3,27	4,00	0,96
	Máximo	93,50	95,50	91,75
	Mínimo	84,50	85,50	89,50
VGG16	Media	85,60	93,70	89,65
	Std	8,98	3,87	3,42
	Máximo	94,00	97,00	91,50
	Mínimo	71,00	89,00	84,00
GoogLeNet	Media	88,8	92,7	90,75
	Std	2,77	1,72	0,75
	Máximo	92,00	94,00	91,50
	Mínimo	85,00	90,00	89,50

Tabla 2. *Media, std, valores máximo y mínimo obtenidos de las métricas de Sensibilidad, Especificidad y Exactitud para cada uno de los tres modelos en la fase de test.*

Modelo	Valor	Tiempo (s)	CPU (%)	Mem. (MB)
AlexNet	Media	12,76	35,42	2.814,8
	Std	0,21	0,52	75,08
	Máximo	12,95	36,20	2.909
	Mínimo	12,43	34,80	2,724
VGG16	Media	78,84	39,27	4.854,2
	Std	0,57	0,29	58,87
	Máximo	79,75	39,62	4.935
	Mínimo	78,05	38,86	4.784
GoogLeNet	Media	25,33	37,03	1.589,8
	Std	0,45	0,31	8,87
	Máximo	25,86	37,40	1.596
	Mínimo	24,78	36,72	1.575

Tabla 3. *Media, std, valores máximo y mínimo obtenidos para el tiempo, consumo de CPU y utilización de memoria para cada uno de los tres modelos durante la fase de test.*

En cuanto al coste computacional de los algoritmos, la Tabla 3 muestra los resultados obtenidos. Respecto al tiempo y memoria consumida, se observan diferencias considerables para los tres modelos. Así pues, el tiempo medio de cálculo requerido varía entre 12,76 y 78,84 segundos por iteración, por lo que AlexNet es 6,5 veces más rápida que VGG16. La memoria empleada por cada uno de los modelos también es muy variable, necesitando GoogLeNet 1.589,8 MB y VGG16 4.854,2 MB. Finalmente, en cuanto al consumo de CPU, se observa una menor dispersión entre valores medios, con una diferencia inferior al 4 % entre las tres redes. Al igual que en los resultados de clasificación, en este caso no se observa gran dispersión entre los valores obtenidos para las cinco iteraciones conducidas.

5. Discusión y conclusiones

En el presente trabajo se han comparado tres conocidos modelos de redes neuronales convolucionales preentrenadas para la tarea específica de evaluar la calidad del ECG. Aunque recientemente se han realizado algunas comparaciones sobre cómo funcionan diferentes modelos de redes preentrenadas en aplicaciones relacionadas con el ECG, ta-

les como la detección de morfologías atípicas [16], la clasificación de arritmias [17], o la identificación de personas [18], sus resultados no pueden compararse directamente con la tarea de discernir entre segmentos de ECG de alta y de baja calidad. Además, en estos trabajos los algoritmos se compararon principalmente en función de los resultados de clasificación obtenidos en un conjunto de test, sin tener en cuenta otra información sobre su eficacia, como el porcentaje de uso de la CPU, el consumo de memoria y el tiempo de cálculo. Sin embargo, estas métricas de rendimiento pueden servir de referencia y mostrar los recursos necesarios para embeber cada algoritmo en sistemas de monitorización de ECG portátiles o vestibles.

En cuanto a los resultados de rendimiento obtenidos, se observa que VGG16 requirió más tiempo que AlexNet y GoogLeNet para clasificar todos los segmentos de ECG incluidos en el conjunto de test. A pesar de tener una estructura simple, el diseño secuencial hace que este modelo sea menos productivo que otras redes basadas en una arquitectura paralela [15]. Además, para un uso similar de CPU que el resto de redes, VGG16 requirió mucha más memoria. Esto podría ser debido al enorme número de hiperparámetros que emplea. Así pues, mientras que VGG16 tiene más de 138 millones de hiperparámetros, AlexNet utiliza 65 millones y GoogLeNet sólo 6 millones [14]. Por lo tanto, VGG16 no parece ser una opción adecuada para ser incorporado en dispositivos portátiles para la monitorización de ECG a largo plazo, ya que requiere una fuerte carga computacional y, por lo tanto, un consumo elevado de batería. En cuanto a los otros dos modelos, AlexNet necesitó menos tiempo para clasificar todos los intervalos de ECG, mientras que GoogLeNet utilizó menos memoria para ejecutarse. Por lo tanto, la implementación de ambos algoritmos para la monitorización de ECG en tiempo real podría implicar requerimientos computacionales similares.

Por otro lado, a pesar de las diferencias entre las tres arquitecturas, han obtenido valores muy similares de exactitud, sensibilidad y especificidad. De hecho, solamente se han registrado diferencias entre 1 y 3 % en estas tres métricas a la hora de analizar los segmentos de alta y baja calidad del conjunto de test. Este resultado sugiere que modelos no muy complejos, tales como AlexNet y GoogLeNet, presentan suficiente capacidad de aprendizaje para obtener un buen rendimiento en la evaluación la calidad del ECG.

En vista de todas estas observaciones, AlexNet y GoogLeNet son los modelos que muestran un mejor compromiso entre exactitud, discerniendo entre segmentos de ECG de alta y baja calidad, y carga computacional. Por lo tanto, estos modelos parecen más adecuados que VGG16 para evaluar la calidad del ECG, así como para ser embebidos en sistemas de monitorización de ECG vestibles o portátiles.

Agradecimientos

Este trabajo ha sido cofinanciado por los proyectos de investigación competitiva DPI2017-83952-C3 de MINECO-AEI-FEDER-UE, SBPLY/17/180501/000411 de la JCCLM, AICO/2019/036 de la GVA y FEDER 2018/11744.

Referencias

[1] Patil T, Pandey S, Visrani K. A review on basic deep learning technologies and applications. In Kotecha K, Piuri V, Shah H, Patel R (eds.), Data Science and Intelligent Applications, volume 52. Springer, 2021; 565–573.

[2] Hong S, Zhou Y, Shang J, Xiao C, Sun J. Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review. Comput Biol Med Jul 2020; 122:103801.

[3] Satija U, Ramkumar B, Manikandan MS. A review of signal processing techniques for electrocardiogram signal quality assessment. IEEE Rev Biomed Eng 2018;11:36–52.

[4] Bashar SK, Ding E, Walkey AJ, McManus DD, Chon KH. Noise detection in electrocardiogram signals for intensive care unit patients. IEEE Access 2019;7:88357–88368. ISSN 2169-3536.

[5] Vavrinsky E, Subjak J, Donoval M, Wagner A, Zavodnik T, Svobodova H. Application of modern multi-sensor holter in diagnosis and treatment. Sensors Basel May 2020;20(9).

[6] Martínez A, Alcaraz R, Rieta JJ. Application of the phasor transform for automatic delineation of single-lead ECG fiducial points. Physiol Meas Nov 2010;31(11):1467–85.

[7] Xu X, Wei S, Ma C, Luo K, Zhang L, Liu C. Atrial fibrillation beat identification using the combination of modified frequency slice wavelet transform and convolutional neural networks. J Healthc Eng 2018;2018:2102918.

[8] Zhang Q, Fu L, Gu L. A cascaded convolutional neural network for assessing signal quality of dynamic ECG. Computational and Mathematical Methods in Medicine 2019;2019.

[9] Mincholé A, Rodriguez B. Artificial intelligence for the electrocardiogram. Nat Med 01 2019;25(1):22–23.

[10] Huerta A, Martínez-Rodrigo A, González VB, Quesada A, Rieta J, Alcaraz R. Quality assessment of very long-term ECG recordings using a convolutional neural network. In 2019 E-Health and Bioengineering Conference (EHB). 2019; 1–4.

[11] Clifford GD, Liu C, Moody B, Li-wei HL, Silva I, Li Q, Johnson A, Mark RG. AF classification from a short single lead ECG recording: the PhysioNet/Computing in Cardiology Challenge 2017. In 2017 Computing in Cardiology (CinC). 2017; 1–4.

[12] Shoeb A, Cliord G. Chapter 16—wavelets; multiscale activity in physiological signals. Biomedical Signal and Image Processing 2005;–.

[13] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems. 2012; 1097–1105.

[14] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv14091556 2014;1–14.

[15] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In IEEE Conference on Computer Vision and Pattern Recognition. 2015; 1–9.

[16] Diker A. Cömert Z, Avcı E, Toğaçar M, Ergen B. A novel application based on spectrogram and convolutional neural network for ECG classification. In 1st IEE International Informatics and Software Engineering Conf. 2019; 1–6.

[17] Alquran H, Alqudah A, Abu-Qasmieh I, Al-Badarneh A, Almashaqbeh S. ECG classification using higher order spectral estimation and deep learning techniques. Neural Network World 2019;29(4):207–219.

[18] Byeon YH, Pan SB, Kwak KC. Intelligent deep models based on scalograms of electrocardiogram signals for biometrics. Sensors Basel Feb 2019;19(4).

Señales Biomédicas IV

Viernes, 27 de noviembre de 2020

RHRVEasy: Heart Rate Variability Analysis Made Easy

S. Bardají Ortiz, C.A. García Martínez, J. Tejedor Noguerales, A. Otero Quintana

Department of Information Technology, Escuela Politécnica Superior, Universidad San Pablo-CEU,

Campus Montepríncipe, Boadilla del Monte, 28668 Madrid, Spain s.bardaji@usp.ceu.es, {constantino.garciama, javier.tejedornoguerales, aotero}@ceu.es

Abstract

Heart rate variability analysis is a simple non-invasive tool that can be used to find markers of cardiac and neurological pathologies based on an electrocardiogram recording. RHRV is an opensource R package that carries out HRV analysis; it currently has around 1000 monthly downloads. Many of the users of this package are not familiar with programming in R and/or do not have a strong statistical training, which hinders the usage of this package. We have created an RHRV script that takes as inputs case and control recordings of a given experiment and calculates the HRV temporal and frequency domain indexes. Then, it applies a proper statistical test, depending on the normality of the data, to find differences between those indexes. The final goal of the script is making the usage of RHRV easier for non-technical users.

1. Introduction

Heart Rate Variability (HRV) analysis studies the constant variation between successive heartbeats. The periods of time between successive heartbeats are called RR intervals. HRV analyzes the cardiac activity of the autonomic nervous system (ANS), which controls the involuntary aspects of the body's physiology. The ANS is divided into two main branches: the sympathetic and parasympathetic nervous systems. The sympathetic system tends to activate the body, for instance, by increasing the heart rate (HR). Conversely, the primary function of the parasympathetic is relaxing the body, thus decreasing the HR. Were the subject a healthy one, these two systems would be in a constant state of interaction, thus leading to everlasting variations between the measured HR. The sympathetic and parasympathetic have different speed of response, which can be subsequently utilized to categorize their contributions to HRV [1-2].

HRV analysis is currently considered a simple non-invasive tool for the study of ANS alterations. The aim of this analysis is typically to establish new markers of diseases. The analysis of HRV requires only an electrocardiogram (ECG) recording, or a recording containing the beat positions. Overall, HRV has proven to be a useful tool to discover markers of cardiac and neurological pathologies [3].

There are different types of HRV analysis, which can be broadly classified in time, frequency, and non-linear domain methods. Time domain methods are based on statistical approaches, whereas frequency domain methods use the Fourier Transform or the Wavelet Transform to obtain the power spectrum of the RR series in certain frequency bands [3].

The typical HRV analysis is generally comprised of two different populations groups ("control" and "case"), as well as different HRV indexes, which are calculated for each population group. After this calculation, a statistical test is usually conducted with the aim of finding statistically significant differences in these indexes. Were a significant difference be found in a specific index between both population groups, that index is highlighted as a potential marker of the condition of the case population. If that index has a clear meaning, it is possible to search for physiological interpretations associated with the condition in question.

RHRV is an opensource R package to carry out HRV analysis, which currently has about 1,000 monthly downloads. Its user base is both diverse and widespread, comprising, among others, engineers, statisticians, doctors, biologists, and psychologists. Some of these users are not familiar with programming in R, and may not have statistical training. This hampers the usage of RHRV.

The purpose of this project is to create a script that takes as input both case and control recordings, calculates the HRV indexes for the temporal and frequency domains, and uses statistical tests to find differences between them. The final objective of this project is to be able to carry out a full HRV analysis with a single line of R code. The rest of the article is structured as follows. The second section presents the structure of the script. Thereinafter, the third section presents an illustrative example that will allow to show the results of the script execution. Finally, the fourth section draws conclusions and suggests some plausible improvements.

2. Methods and Materials

The main function of the program (*RHRVEasy*), which is the one that the final user will use, only takes two mandatory parameters as input: the paths of the folders where the RR recordings to analyze are located; they are named case and control. The function also takes parameters that permit configuring the time and frequency domain analyses, although default values are provided for all of them; therefore, it is not necessary to specify them. All these parameters use the same name as in the respective RHRV functions. There also is a parameter that lets the user decide whether to apply (default option) or not significance level correction to the statistical tests. The following sections describe the operations performed by the *RHRVEasy* function.

ISBN: 978-84-09-25491-0, pags. 414- 417

2.1. Verification of files

The first thing that *RHRVEasy* function does is to validate its first two parameters, which should be the paths of two folders; one containing the records corresponding with the "control" subjects and another with the records of the "case" subjects. Subsequently, it creates a complete list with the records found in each folder. These records will be the ones that will be analyzed through time and frequency analyses.

2.2. Time and Frequency Analyses

After listing the records, the calculation of the different time and frequency indexes is carried out. This starts by calling the function *time_analysis*, which calculates in a loop the indexes of the temporal analysis. For this, a function is called, which is in charge of preparing the data structures for the different analyses that are going to be used afterwards (*preparing_analysis*). This function creates an *hrv.data* object, which is a data structure that contains the information of the imported files and the results of the HRV analysis. This structure is created with the function *CreateHRVData*. Afterwards, the data to be analyzed is loaded into this structure with *LoadBeatRR*. This constructs the RR and the HR series, and stores them in *hrv.data*. To eliminate spurious beats, a filter operation is applied: *FilterNIHR*. This *preparing_analysis* function returns the already mentioned *hrv.data*, which will be used by both time and frequency analyses. It is important to note that by default all calls to RHRV functions are always made using the default parameter values specified in RHRV.

Afterwards, time domain parameters are calculated with the *CreateTimeAnalysis* function. They include the Standard Deviation of the NN (RR) Interval (SDNN), Standard Deviation of the 5 minute Average NN intervals (SDANN), Standard Deviation of Successive Differences (SDSD), and the Triangular Interpolation of NN interval histogram (TINN).

All these steps mentioned are applied to control files, and to case files. Two different *dataframes* are created with their respective data. Finally, all of the data is put together in a single *dataframe*. In this *dataframe* both case and control information is stored, with a column that indicates the group it belongs to.

Then, the script proceeds with the frequency analysis. First, the *InterpolateNIHR* function is applied in a loop, with a default interpolation frequency of 4Hz. This is done in order to obtain a uniformly sampled HR series, which is needed to perform spectral analysis in the frequency domain. Then, the *preparing_analysis* function is again applied, obtaining the *hrv.data* with all the calculations that will be necessary later.

Once interpolated, it was observed that some zeroes appeared in the interpolated registers. RHRV has some criteria that if there is a gap too large in the HR series, it is zeroed when interpolating. These zeroes of the HR were removed because they produced errors. This was achieved by finding them in the HR series, then computing the median of HR once those zeroes were already removed, and finally substituting the zeroes by the median.

Then, the spectral analysis structure is created by using *CreateFrequencyAnalysis*. Afterwards, the spectrogram of the HR series in the Ultra Low Frequency (ULF), Very Low Frequency (VLF), Low frequency (LF) and High Frequency (HF) bands are obtained by applying the *CalculatePowerBand* function. All of this is again applied for both case and control separately, creating two different *dataframes*. Subsequently, all the information obtained is put together in a unique *dataframe* for both case and control, with a column that indicates the group. For the Wavelet analysis, broadly the same procedure is done. The main difference is that when applying *CalculatePowerBand* function, we specify that the analysis type is Wavelet. The mother Wavelet used by default is Daubechies's extremal phase ("d4") [8].

All this information will be returned by the main function on a list, which will contain on one hand the time analysis *dataframe* and on the other hand the frequency/Wavelet analysis *dataframe,* depending on what frequency analysis was used.

2.3. Normality Test

After calculating the HRV indexes and storing them in *dataframes,* the following step is the statistical analysis. This is done by first analyzing if the results follow a normal distribution. To achieve this, the Shapiro-Wilk test is applied, which is a test that allows checking if a random sample comes from a normal distribution or not. If the value returned is lower than 0.05, it indicates that the sample is not normal. On the other hand, if the value obtained is greater than 0.05, we can assume that it does follow a normal distribution [4].

The *dataframes* created previously contained both case and control information together. This information is later split into case and control by separating it through the group column. This is done in order to apply Shapiro-Wilk test for each one separately. The different *p-values* obtained for control and case are both compared with 0.05 to check if the distributions are normal or not, and subsequently choose the proper statistical test to compare the groups.

2.4. Statistical Hypothesis Testing

Given that multiple statistical tests are going to be carried out on several HRV indexes, a correction of the significance level should be applied, a fact that is sometimes over-looked in the literature. A Boolean parameter named "correction" is passed to the *RHRVEasy* function in order to decide if the correction is applied or not. By default, it will be applied using the Bonferroni correction. This type of correction consists on dividing the raw *p-values* by the number of tests. This way, a vector with adjusted *p-values* is obtained [5].

The following step is carrying out the statistical test. If the data was previously observed to be normal, the ANOVA test is applied. It was decided to apply this test because, although the script does not support it right now, there

could be more than one factor that influences the results (more than two study groups).

On the other hand, if the data did not follow a normal distribution, the Kruskal-Wallis test is applied. This is a non-parametric test, which is recommended when the assumptions of one-way ANOVA are not met [7].

All the results of the statistical test are stored in a list which contains the ANOVA and Kruskal-Wallis results. This list is returned by the main function, as it was mentioned previously. This object before being returned to the main it is converted into an instance of the class *RHRVEasyResult*. A print method was created for the objects of this class that shows for all the tests whose *p-value* is lower than 0.05 (both ANOVA and Kruskal-Wallis) the name of the HRV index, the *p-value* and the mean and standard deviation of the index for both groups (see Figure 1).

```
> result
Result of the analysis of the variability of the he
art rate of the group normal versus the group chf

There is a statistically significant difference in
 SDNN; pvalue:  1.993373e-07
SDNN for the group  normal is 139.9524 +- 31.94841
 and for the group chf  is 78.98597 +- 45.94111

There is a statistically significant difference in
 SDANN; pvalue:  2.714068e-08
SDANN for the group  normal is 130.1846 +- 32.3014
 and for the group chf  is 65.77019 +- 42.68555

There is a statistically significant difference in
 SDNNIDX; pvalue:  0.001018641
SDNNIDX for the group  normal is 50.74575 +- 14.148
 and for the group chf  is 39.43364 +- 23.92845

There is a statistically significant difference in
 IRRR; pvalue:  2.268287e-07
IRRR for the group  normal is 202.6743 +- 52.56548
 and for the group chf  is 103.4375 +- 67.74388
```

***Figure 1**. Output of the print method of the object (named result in the example) returned by the RHRVEasy function*

3. Results

The data used to check the correct functioning of the code contained recordings from two different types of patients: a group of healthy patients, and a group of patients with a severe cardiac pathology. This was done in order to compare data from groups where we knew there should be significant differences in the HRV indexes. On one hand, we used recordings from people who suffered Congestive Heart Failure (CHF) [9]. They are 29 long-term ECG recordings of subjects aged from 34 to 79, including 8 men and 2 women. On the other hand, there are files from Normal Sinus Rhythm (NSR) ECGs [9], with no cardiac issues. They include 54 long-term ECG recordings of 30 men aged from 28 to 76, and 24 women aged from 58 to 73. Both original ECG recordings, CHF and NSR, were digitalized at 128 samples per second, and the beat annotations were obtained by automated analysis with a subsequent manual review and correction [9].

When running the main function *RHRVEasy* with *verbose=TRUE*, the first output consists of some lines which indicate whether the statistical analysis obtained follow a normal distribution or not, and depending on this, whether Kruskal-Wallis or ANOVA tests will be employed. This information is shown for each time and frequency index. The *p-values* obtained are also shown.

The object that the *RHRVEasy* function returns is a list that contains four different objects. The first one is a *dataframe* with the results obtained for the time analysis. This includes a column with the name of the file, all the indexes for the time analysis obtained, and a final column that indicates the group it belongs to. The second object in the list returned by the *RHRVEasy* function is another list that contains the results obtained for the statistical test of the time indexes. It does this for the ANOVA and Kruskal-Wallis results. If the data is not normal, and, hence, Kruskal-Wallis test is employed, the ANOVA contents will be set to NA (which is a logical constant that contains a missing value indicator). The same is done if ANOVA is used for the Kruskal-Wallis results. In Figure 2 an example of the information in this list for the Kruskal-Wallis test is shown, with the SDNN test expanded.

kruskal	list [10]	List of length 10
SDNN	list [5] (S3: htest)	List of length 5
statistic	double [1]	27.03953
Kruskal-Wallis ...	double [1]	27.03953
parameter	integer [1]	1
df	integer [1]	1
p.value	double [1]	1.993373e-07
method	character [1]	'Kruskal-Wallis rank sum test'
data.name	character [1]	'SDNN by clase'
SDANN	list [5] (S3: htest)	List of length 5
SDNNIDX	list [5] (S3: htest)	List of length 5
pNN50	list [5] (S3: htest)	List of length 5
SDSD	list [5] (S3: htest)	List of length 5
rMSSD	list [5] (S3: htest)	List of length 5
IRRR	double [1]	0
MADRR	list [5] (S3: htest)	List of length 5
TINN	list [5] (S3: htest)	List of length 5
HRVi	list [5] (S3: htest)	List of length 5

***Figure 2.** List with the statistical analysis results for the time domain indexes*

filename	ULF	VLF	LF	HF	clase
nsr048_rr_secs.txt	5372.0903	352.738443	716.255789	77.761762	normal
nsr049_rr_secs.txt	6540.5686	238.957131	676.925845	141.708776	normal
nsr050_rr_secs.txt	4927.9108	115.856421	325.821860	83.014728	normal
nsr051_rr_secs.txt	17193.7587	251.449689	674.912727	315.763901	normal
nsr053_rr_secs.txt	8398.0708	230.341170	564.899568	191.253031	normal
nsr054_rr_secs.txt	4769.5648	161.716198	434.284976	153.829750	normal
chf201_rr_secs.txt	2528.0667	122.009599	161.677226	58.796723	chf
chf202_rr_secs.txt	1855.7710	104.542314	281.247979	140.849164	chf
chf203_rr_secs.txt	620.7563	6.047730	11.650019	19.153828	chf
chf204_rr_secs.txt	1085.5905	29.484504	77.443316	80.861202	chf

***Figure 3.** Frequency Analysis Data Frame*

The third object is another *dataframe*; in this case, it contains the results obtained in frequency analysis. It again has a column with the name of the file, the results for each frequency band for each file, and a column indicating the group it belongs to (see Figure 3). The last object is a list with all the statistical test results in the frequency domain. Just like before, a different list is obtained for ANOVA and Kruskal-Wallis (depending on which test was employed). This is illustrated in Figure 4, where only the Kruskal-Wallis analysis is shown, and only the ULF is extended so it can be appreciated.

kruskal	list [4]	List of length 4
ULF	list [5] (S3: htest)	List of length 5
statistic	double [1]	30.54012
parameter	integer [1]	1
p.value	double [1]	3.270333e-08
method	character [1]	'Kruskal-Wallis rank sum te
data.name	character [1]	'ULF by clase'
VLF	list [5] (S3: htest)	List of length 5
LF	list [5] (S3: htest)	List of length 5
HF	list [5] (S3: htest)	List of length 5

Figure 4. *List with the statistical analysis results for the frequency domain indexes*

This information, contained in an *RHRVEasy* object, allows the user to access to detailed information of all the statistical tests carried out, as well as to all the HRV indexes calculated for each of the recordings. When printed, this object shows the result of the HRV analysis of the group control versus the group case if there is statistically significant difference in each time and frequency index, followed by the mean and standard deviation values of those indexes (see Figure 1).

4. Conclusions

RHRVEasy allows the user to carry out a complete heart rate variability analysis in time and frequency domains, including the statistical contrast of hypotheses for each index, with a single line of R code. This script may be very useful for all the diverse RHRV users (doctors, veterinarians, psychologists…) who are not used to working with R. It makes its usage effortless because the only parameters needed for the function are the paths of the folders that contain the files wanted to be analyzed and, optionally, the different parameters that configure the analyses in time and frequency domains can be overwritten.

The main disadvantage that this script has is that, as mentioned previously, it employs the default parameters for every calculation. This is a reasonable compromise for our goals, but in some cases, it could be necessary to change them. For example, if we are working with non-adult patients or individuals performing serious physical effort, as it could be the case with children, athlete and animal recordings, it would be desirable to review the default values of some of the parameters. It is also important to take into account that the ventricular heartbeat filtering employed is automatic one. No ventricular heartbeat filtering algorithm is perfect. An ideal HRV analysis would include the manual inspection and a possibly removal of ventricular beats.

Regarding future work, it would be necessary to extend the script for cases where there are more than two workgroups, which may be a limitation in this script. We will also try to automate the generation of the HRV indexes based on non-linear techniques. Additionally, we would also like to validate the usability of the tool with end users.

Acknowledgments

This research was funded by the Ministry of Science, Innovation, and Universities of Spain and by the European Regional Development Fund of the European Commission under grant number RTI2018-095324-B-I00.

References

[1] Acharya UR, Joseph KP, Kannathal N, Lim CM, Suri JS. Heart rate variability: a review. *Medical and biological engineering and computing.* 2006 Dec 1;44(12):1031-51.

[2] Thayer JF, Åhs F, Fredrikson M, Sollers III JJ, Wager TD. A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. *Neuroscience & Biobehavioral Reviews*. 2012 Feb 1;36(2):747-56.

[3] Martínez CA, Quintana AO, Vila XA, Touriño MJ, Rodríguez-Liñares L, Presedo JM, Penín AJ. *Heart rate variability analysis with the R package RHRV*. Cham, Switzerland: Springer International Publishing; 2017 Sep 18.

[4] Razali NM, Wah YB. Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics*. 2011 Jan 1;2(1):21-33.

[5] Jafari M, Ansari-Pour N. Why, when and how to adjust your P values?. *Cell Journal* (Yakhteh). 2019;20(4):604.

[6] Hogg RV, McKean J, Craig AT. *Introduction to mathematical statistics.* Pearson Education; 2005.

[7] Hecke TV. Power study of anova versus Kruskal-Wallis test. *Journal of Statistics and Management Systems*. 2012 May 1;15(2-3):241-7.

[8] García, C.A., Otero, A., Vila, X. and Márquez, D.G., 2013. A new algorithm for wavelet-based heart rate variability analysis. *Biomedical Signal Processing and Control*, 8(6), pp.542-550.

[9] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*. 2000 Jun 13;101(23):e215-20.

Aprendizaje máquina para la predicción del éxito de la desfibrilación mediante el análisis de la forma de onda de la fibrilación ventricular

G. Zubia Garea[1], X. Jaureguiabeitia Lara[1], U. Irusta Zarandona[1], E. Aramendi Ecenarro[1]

[1] BioRes, Ingeniería de comunicaciones, Euskal Herriko Unibertsitatea UPV/EHU, Bilbo, España,
gzubia004@ikasle.ehu.eus, {unai.irusta, xabier.jaureguibeitia, elisabete.arabemendi}@ehu.eus

Resumen

Predecir el éxito de la terapia de desfibrilación eléctrica para revertir la fibrilación ventricular (FV) permitiría mejorar la supervivencia de los pacientes en parada cardiorrespiratoria extra-hospitalaria (PCREH). Los métodos de predicción se basan en el análisis de la forma de onda del electrocardiograma (ECG) de la FV usando predictores individuales. El objetivo del trabajo es desarrollar modelos de aprendizaje máquina para mejorar la precisión de los métodos. Se dispuso de una base de 1630 casos de PCREH anotados por clínicos, que incluían 3836 intentos de desfibrilación (1081 exitosos). Se diseñaron modelos predictores del éxito de la desfibrilación caracterizando con 27 parámetros el intervalo ECG pre-shock de 2.05s de duración. Se evaluaron los modelos en términos de sensibilidad (SE, éxito), especificidad (SP, no éxito) y la media de ambas (precisión balanceada, PB). Los modelos se ajustaron y evaluaron mediante validación cruzada en particiones por paciente y estratificadas, repitiendo el proceso 100 veces para caracterizar estadísticamente las métricas. Se comparó un modelo de regresión logística monoparamétrico con varios modelos de aprendizaje máquina: regresión logística multiparamétrica, bosques de árboles y máquinas de vectores soporte (SVM) con kernel gaussiano. El mejor modelo monoparamétrico produjo una PB mediana (rango interdecilo, IDR) de 79.5 (79.4-79.6)%. El mejor modelo multiparamétrico (SVM con 6 parámetros) resultó en una PB mediana de 81.4 (81.2-81.6)%, y SE, SP de 83.4 (83.1-84.0)% y 79.3 (79.1-80.0)%, respectivamente. Los modelos de aprendizaje máquina permiten mejorar la predicción monoparamétrica hasta en 2 puntos de PB. En el futuro deberán desarrollarse nuevos predictores, para lo que la extracción exhaustiva de caracterísi-cas mediante redes convolucionales (CNN) sería una buena alternativa.

1. Motivación

Las enfermedades cardiovasculares (ECV) son la principal causa de mortalidad en Europa, representando el 39% de los fallecimientos. En el estado español, las ECV son responsables del 29% de las defunciones, es decir, de más de 122.000 personas-año. La parada cardiorrespiratoria extrahospitalaria (PCREH) es un evento cardiovascular que constituye un problema de salud pública, con una supervivencia inferior al 10% y causando más de 30.000 muertes al año en este país [1].

Entre las arritmias, la fibrilación ventricular (FV) es causante del 85% de PCREHs [2,3], y su único tratamiento efectivo es la desfibrilación por descarga eléctrica. Desafortunadamente no todas las descargas son exitosas, y las que no lo son pueden deteriorar el miocardio por lo que los métodos para la predicción del éxito en la desfibrilación son de gran interés. Estos métodos se basan en el análisis del ECG de la FV. El ECG durante la FV es irregular, caracterizándose por presentar latidos ventriculares de amplitud y duración dispares, tal y como se aprecia en la Figura 1. Durante la FV no hay contracción ventricular efectiva y los ventrículos se contraen desordenadamente.

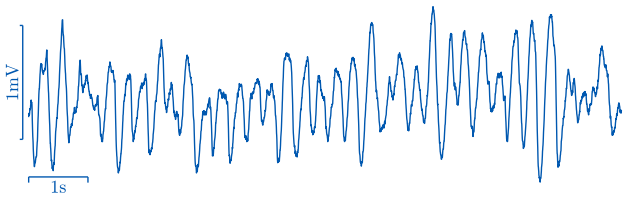


Figura 1. *Ejemplo del ECG de una FV. La forma de la onda es irregular, con latidos ventriculares desorganizados*

En las últimas décadas se ha avanzado hacia el cálculo de distintas características de la FV a partir de la señal de ECG que permitan predecir el éxito de la terapia eléctrica aplicada por un desfibrilador externo automático (DEA) al paciente en PCREH. Una de la más relevantes es AMSA (del inglés Amplitude Spectrum Area), que es una suma ponderada de las amplitudes en el dominio espectral [4,5]. Recientemente, Chicote et. al han desarrollado nuevos indicadores de características no lineales de las FV, como la entropía difusa y muestral [6,7], obteniendo resultados que mejoran los de los predictores clásicos como AMSA.

Aunque se han propuesto numerosos predictores [8–14], éstos se han empleado mayoritariamente de forma individual y excepcionalmente combinados en modelos de aprendizaje máquina (AM) [8]. Es por ello que el objetivo de este trabajo es analizar el poder predictivo de estos métodos y valorar la posibilidad de combinarlos de manera eficiente en modelos AM, empleando para ello una base de datos extensa con anotaciones clínicas contrastadas.

2. Materiales

La base de datos empleada está compuesta por los registros ECG de 1630 casos de PCREH de un estudio multicéntrico realizado entre 2008 y 2010 por la agencia italiana de la emergencia, Azienda Regionale Emergenza Urgenza. Los registros proceden de 3 marcas de DEA: Philips HeartStart FR2, Physio-Control Lifepak y ZOLL AED plus. El objetivo de este estudio era doble, por una parte, conocer el éxito en la desfibrilación con las técnicas actuales; y, por

ISBN: 978-84-09-25491-0, pags. 418- 421

otra, poder utilizar dichos registros para desarrollar nuevos métodos predictores del éxito en la desfibrilación.

Todos los registros contenían eventos de desfibrilación para los cuales además se dispuso de anotaciones clínicas sobre su éxito. Se definió como desfibrilación exitosa aquella en la que en el minuto posterior a la descarga se observaron latidos con complejos QRS visibles con frecuencia superior a 30 latidos por minuto [9]. Los registros se exportaron y convirtieron a un formato común en Matlab (MathWorks, Natick, MA), y se remuestrearon a 250 Hz. Los segmentos de análisis empleados fueron de 2.05s (512 muestras) tomados 0.5s antes del momento de la descarga. En total, se extrajeron 3836 segmentos de análisis, cuyo desglose puede observarse en la Tabla 1.

		Éxito	
DEA	**Shocks (pacientes)**	**Sí**	**No**
Philips	2066 (814)	591	1475
Physio	1581 (767)	405	1176
Zoll	189 (49)	85	104
Total	**3836 (1628)**	**1081**	**2755**

Tabla 1. Desglose de los registros por cada DEA en segmentos de análisis y pacientes, así como en descargas con y sin éxito.

3. Métodos

Para cada segmento se calcularon los 27 predictores principales del éxito de la desfibrilación descritos en la literatura [5–14], para los que existe una implementación de código abierto en formato Matlab accesible a través de https://github.com/BChicote/shockOutcome. Los métodos incluyen características de la FV calculadas en el dominio temporal, espectral, tiempo-frecuencia, de entropía, y basadas en el análisis de dinámica no lineal de la señal.

Para los análisis se tomaron las descargas exitosas como clase positiva ($y = 1$) en modelos de decisión binaria. Se emplearon como métricas de desempeño la sensibilidad (SE, clase positiva), la especificidad (SP, clase negativa) y la precisión balanceada (PB), que es la media de SE y SP.

3.1. Análisis monoparamétrico

Inicialmente se evaluó la capacidad discriminativa de cada característica mediante un análisis de Característica Operativa del Receptor (ROC), calculando el área bajo la curva (AUC en inglés). A su vez se ajustó un clasificador monoparamétrico basado en regresión logística (RL), como modelo base de clasificación.

3.2. Análisis multiparamétrico

A continuación, se analizaron varios modelos de AM para combinar los predictores en modelos de decisión más elaborados. Para ello cada segmento i se representó mediante la tupla $(\boldsymbol{x}_i,\ y_i)$, en la que el vector $\boldsymbol{x}_i$ agrupa las N características del segmento e y_i representa el diagnóstico clínico. Se analizaron los siguientes modelos:

- **Regresión logística multiparamétrica (RLM).** Se empleó como clasificador base por ser un modelo lineal sencillo de decisión. En este caso se realizó un análisis de todas las combinaciones posibles de predictores, es decir se las combinaciones de N características de las 27 características. Dado que el número de combinaciones es muy elevado se realizó el análisis hasta $N = 5$ (y $N = 27 - 5 = 22$) y a partir de ahí se fijaron las 2, 3,… características que aparecían en las combinaciones con mayor PB.
- **Bosques Aleatorios (RF en inglés).** Estos algoritmos se basan en la idea de que es posible ajustar muchas clasificadores débiles (individualmente imprecisos) que deciden independientemente. Cuando las decisiones de estos clasificadores se agregan (por mayoría de votos) el clasificador se vuelve robusto y muy preciso. Por lo tanto, un RF es un conjunto de B árboles de decisión $\{T_1(\boldsymbol{x}), \dots, T_B(\boldsymbol{x})\}$ que produce predicciones B cuasi-incorreladas $\{\hat{y}_1 = T_1(\boldsymbol{x}), \dots, \hat{y}_B = T_B(\boldsymbol{x})\}$ sobre el éxito en la desfibrilación de cada segmento de análisis del ECG [15], siendo la predicción del modelo $\hat{y}$ la decisión mayoritaria de los $\hat{y}_i$. Para diseñar el clasificador RF se tuvieron en cuenta 3 hiperparámetros: número de predictores por rama, el número de árboles y el número mínimo de observaciones en la última división. Finalmente, se asignaron probabilidades previas uniformes para cada clase durante el entrenamiento para abordar el desequilibrio de clase.
- **Maquinas de vectores de soporte (SVM en inglés).** Se empleó una SVM con kernel gaussiano para modelar relaciones no lineales entre parámetros [16], mediante la siguiente transformación métrica:

$$K(\boldsymbol{x}, \boldsymbol{x}_i) = \exp(-\gamma \|\boldsymbol{x} - \boldsymbol{x}_i\|^2)$$

en la que $\|\cdot\|^2$ es la norma L_2. Se ajustaron dos hiperparámetros, el ancho del kernel, γ, y el margen o la flexibilidad del límite de decisión, C. Se buscaron los valores de C y γ que maximizaban la PB, para ello se realizó una búsqueda logarítmica exhaustiva previa a cada entrenamiento, de 40×40 en los rangos $10^{-1} \leq C \leq 10^{1.5}$ y $10^{-3} \leq \gamma \leq 10^{1}$ respectivamente.

3.3. Evaluación de los análisis

Los 3 modelos propuestos fueron evaluados usando una estrategia de validación cruzada de 10 iteraciones (10-fold CV). La partición de los datos se realizó por paciente (patient-wise) y cuasiestratificada, de modo que los segmentos de análisis de cada iteración presentaran una distribución similar a la del conjunto de datos. Como los resultados podían depender de cómo se dividieran los datos, se realizaron 100 particiones aleatorias utilizando el Servidor de Números Cuánticos de la Universidad Nacional de Australia (https://qrng.anu.edu.au). Esto permitió la caracterización estadística de las métricas de desempeño de los modelos.

4. Resultados

Los resultados del análisis monoparamétrico se muestran en la Tabla 2. Dicha tabla resume por un lado los resultados del análisis de curva ROC para todos los datos mediante el AUC, y se muestra a su vez la PB obtenida para el modelo LR monoparamétrico. En este último caso, se muestra el

valor mediano para las 100 particiones 10-fold CV, así como el rango inderdecilo (IDR, entre 10%-90%). Se observa que los valores de AUC y PB son muy similares para las mejores características (el resto no se muestra). Éstas se corresponden con las que mejores resultados han reportado en la literatura [7,8], e incluyen predictores basados en la pendiente (MS, MdS), entropías muestral y difusa (SmpEn, FuzzyEn), AMSA, y el análisis de la forma de onda basados en diagramas de Poincaré (MSI).

La Figura 2 muestra el análisis de curva ROC para los mejores y peores predictores. Se observa que predictores propuestos en la literatura como la entropía wavelet (WE), el análisis de fluctuación sin tendencia (DFA), o la medida de planicie espectral (SFM) y entropía espectral (SpecEn), que miden características asociadas exclusivamente a la no linealidad y desorden de la forma de onda FV, no producen resultados satisfactorios. Es importante indicar que las implementaciones de entropía muestral y difusa con altos valores de AUC y PB (Tabla 2) se modificaron para incluir información de amplitud de la FV [7,8].

Característica	AUC (%)	PB (%)
MS	87.0	79.5 (79.4-79.6)
SmpEn	86.4	79.2 (79.1-79.3)
FuzzyEn	86.6	79.1 (79.0-79.2)
AMSA	86.5	79.0 (78.9-79.0)
MdS	86.5	78.8 (78.7-78.9)
MSI	86.6	78.5 (78.4-78.6)

Tabla 2. *Ranking de las mejores características para el análisis de curva ROC (AUC), y para el clasificador RL monoparamétrico (PB). En la PB se muestra el valor mediano (IDR) para las 100 particiones.*

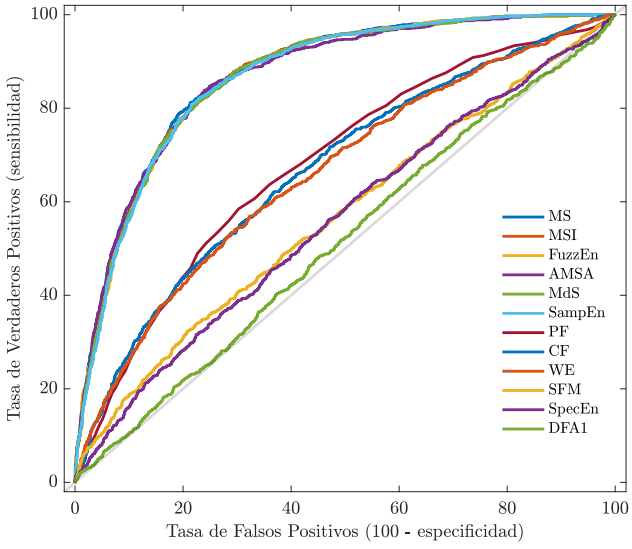


Figura 2. *Curvas ROC para los 6 parámetros con mejor y peor poder discriminativo.*

Al combinar los parámetros en los tres modelos de AM se obtuvieron los resultados mostrados en la Figura 3 donde se muestra el valor mediano (IDR) para las 100 particiones 10-fold CV, y se compara la PB de los tres modelos de AM estudiados en función del número de características de los modelos. Para los modelos basados en RF se emplearon 150 árboles, $\sqrt{N}$ predictores por rama, y 40 observaciones en la última división, ya que en los experimentos preliminares produjeron valores de PB más elevados. Tal y como se observa en la figura los mejores resultados se obtuvieron para los modelos basados en SVM, que superaron en más de 1.2-puntos de PB al resto de modelos, y en 1.9-puntos al mejor modelo monoparamétrico (Tabla 2). Además, se observa que los modelos no mejoran con la inclusión de nuevas características a partir de un número bajo de características ($N = 6$ para el caso de la SVM). Por ejemplo, los mejores resultados se obtienen para la SVM con 6 características, que resultan en valores de 81.4 (81.2-81.6)%, 83.4 (83.1-84.0)% y 79.3 (79.1-80.0)% para la PB, SE y SP, respectivamente. Los mejores resultados para LRM y RF se obtuvieron para $N = 7$ y $N = 9$, con valores de PB, SE y SP de 80.2 (80.0-80.4)%, 81.4 (81.0-81.6)% y 79.0 (78.9-79.2)% para LRM y de 80.0 (79.7-80.3)%, 81.6 (80.9-82.2)% y 78.4 (78.1-78.8)% para RF.

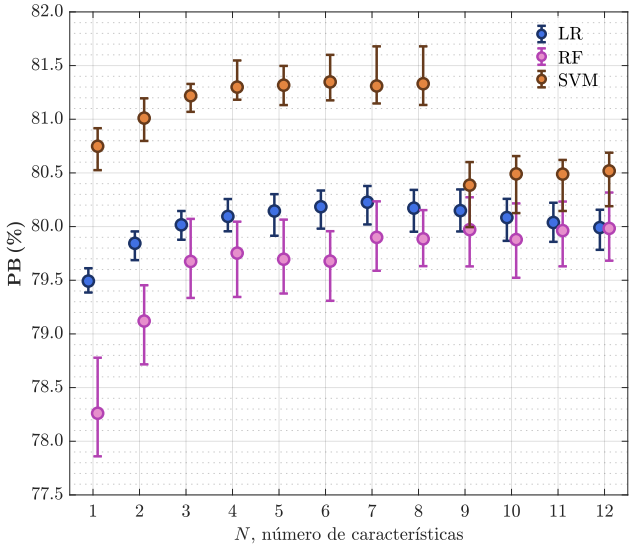


Figura 3. *Mediana (IDR) de las puntuaciones de rendimiento para los modelos RLM, RF y SVM en función del número de características utilizadas.*

5. Discusión

En este estudio se ha analizado el poder discriminativo de los predictores de éxito de la desfibrilación basados en análisis de la forma de onda FV pre-descarga, empleando para ello la base de datos con anotaciones clínicas más extensa reportada en la literatura para este problema. Se ha analizado también la mejora en la predicción asociada al uso de modelos avanzados de AM.

Los mejores predictores descritos en la literatura muestran valores similares de AUC y PB, corroborando los resultados reportados en estudios previos [7,8], aunque algunos de ellos no demostraron poder discriminativo sobre la base de datos de estudio. Además, los mejores predictores (Tabla 2) contienen información asociada a la amplitud, bien directamente (AMSA, y pendiente), o mediante versiones modificadas de los predictores con información de amplitud (SmpEn, FuzzyEn) [6,7]. Es bien conocido que el valor de amplitud de la FV está asociado a la duración de la misma, y predice en gran medida el éxito de la desfibrila-

ción [17]. La combinación de parámetros en modelos de AM (Figura 3) demuestra que es posible obtener mejoras de la PB, por ejemplo, el mejor modelo multiparamétrico mejora en 1.9-puntos la PB del mejor modelo LR monoparamétrico. Sin embargo, esta mejora es limitada, ya que supone solo un recorte de un 10% sobre los 20-puntos de margen para la PB de 80% del modelo monoparamétrico. Se concluye así que los parámetros de discriminación proporcionan información muy parecida de la forma de onda ECG de la FV. Dado el tamaño de la base de datos disponible, es posible que en el futuro se puedan desarrollar métodos eficientes de extracción de características de la FV basadas en aprendizaje profundo (CNN). Este tipo de soluciones ya han sido demostradas para el análisis del ECG y la detección de FV [18]. Además, otra línea de trabajo para la mejora de la predicción son los modelos multimodales que integren información de varias señales, en particular información sobre el CO_2 expirado que proporciona una medida subrogada de perfusión. Estos modelos han demostrado ya mejoras en la predicción del éxito de la desfibrilación [19], pero la disponibilidad de las señales es limitada y no se podrían analizar con los datos del presente estudio.

Finalmente, estos nuevos métodos se podrían implementar en los DEAs para mejorar la supervivencia en la PCREH. Actualmente, los DEA analizan segmentos cortos del ECG para decidir si es necesario desfibrilar; sin embargo, si se predice que la desfibrilación será no exitosa conviene posponer el shock y continuar con la reanimación.

6. Conclusiones

Como conclusión, los mejores predictores de la literatura reportan valores muy parecidos de AUC y PB. Además, las técnicas de AM multiparamétrico mejoran la PB, siendo la SVM con kernel gaussiano y 8 predictores la que mejores resultados reporta. Sin embargo, la mejora es leve respecto de modelos monoparamétricos porque muchos de los predictores cuantifican características similares de la FV. Por lo tanto, sería interesante introducir la extracción exhaustiva de características usando CNNs.

Agradecimientos

Este trabajo ha recibido ayuda financiera del Ministerio de Ciencia, Innovación y Universidades, proyecto RTI-2018-101475-BI00, junto con el Fondo Europeo de Desarrollo Regional (FEDER), y del Gobierno Vasco por medio de la subvención a grupos de investigación IT-1229-19 y la beca del programa Ikasiker IkasC_2019_1_0275.

Referencias

[1] Consejo Español de Resucitación Cardiopulmonar (CERCP), «Día Europeo de Concienciación del Paro Cardíaco, 16 de octubre». Disponible en: https://www.cercp.org/images/stories/recursos/2015/varios/NOTA_PRENSA_DIA_PARO_CARDIACO_2015.pdf.

[2] M. Rubart, «Mechanisms of sudden cardiac death», *J. Clin. Invest.*, vol. 115, n.º 9, pp. 2305-2315, ago. 2005.

[3] N. Rodríguez de Viguri, J. López Mesa, y M. Ruano Campos M, *Manual de soporte vital avanzado.*, 4.ª ed. Madrid, Espana: Masson, 2007.

[4] G. Ristagno, Y. Li, F. Fumagalli, A. Finzi, y W. Quan, «Amplitude spectrum area to guide resuscitation—A retrospective analysis during out-of-hospital cardiopulmonary resuscitation in 609 patients with ventricular fibrillation cardiac arrest», *Resuscitation*, vol. 84, n.º 12, pp. 1697-1703, dic. 2013.

[5] G. Ristagno *et al.*, «Amplitude Spectrum Area to Guide Defibrillation: A Validation on 1617 Patients With Ventricular Fibrillation», *Circulation*, vol. 131, n.º 5, pp. 478-487, feb. 2015.

[6] B. Chicote *et al.*, «Application of Entropy-Based Features to Predict Defibrillation Outcome in Cardiac Arrest», *Entropy*, vol. 18, n.º 9, p. 313, ago. 2016.

[7] B. Chicote *et al.*, «Fuzzy and Sample Entropies as Predictors of Patient Survival Using Short Ventricular Fibrillation Recordings during out of Hospital Cardiac Arrest», *Entropy*, vol. 20, n.º 8, p. 591, ago. 2018.

[8] M. He *et al.*, «Combining multiple ECG features does not improve prediction of defibrillation outcome compared to single features in a large population of out-of-hospital cardiac arrests», *Crit. Care*, vol. 19, n.º 1, p. 425, dic. 2015.

[9] C. G. Brown y R. Dzwonczyk, «Signal Analysis of the Human Electrocardiogram During Ventricular Fibrillation: Frequency and Amplitude Parameters as Predictors of Successful Countershock», *Ann. Emerg. Med.*, vol. 27, n.º 2, pp. 184-188, feb. 1996.

[10] I. Jekova, F. Mougeolle, y A. Valance, «Defibrillation shock success estimation by a set of six parameters derived from the electrocardiogram», *Physiol. Meas.*, vol. 25, n.º 5, pp. 1179-1188, oct. 2004.

[11] A. Neurauter *et al.*, «Prediction of countershock success using single features from multiple ventricular fibrillation frequency bands and feature combinations using neural networks», *Resuscitation*, vol. 73, n.º 2, pp. 253-263, may 2007.

[12] T. Eftestøl, K. Sunde, S. Ole Aase, J. H. Husøy, y P. A. Steen, «Predicting Outcome of Defibrillation by Spectral Characterization and Nonparametric Classification of Ventricular Fibrillation in Patients With Out-of-Hospital Cardiac Arrest», *Circulation*, vol. 102, n.º 13, pp. 1523-1529, sep. 2000.

[13] L. D. Sherman, T. D. Rea, J. D. Waters, J. J. Menegazzi, y C. W. Callaway, «Logarithm of the absolute correlations of the ECG waveform estimates duration of ventricular fibrillation and predicts successful defibrillation», *Resuscitation*, vol. 78, n.º 3, pp. 346-354, sep. 2008.

[14] Y. Gong, Y. Lu, L. Zhang, H. Zhang, y Y. Li, «Predict Defibrillation Outcome Using Stepping Increment of Poincare Plot for Out-of-Hospital Ventricular Fibrillation Cardiac Arrest», *BioMed Res. Int.*, vol. 2015, pp. 1-7, 2015.

[15] T. Hastie, J. Friedman, y R. Tibshirani, *The Elements of Statistical Learning*. New York, NY: Springer New York, 2001.

[16] C. A. González, «SVM: Máquinas de Vectores Soporte», presentado en Escuela de Ingeniería Informática de Valladolid.

[17] W.D. Weaver, L.A. Cobb, D. Dennis, R. Ray, A.P. Hallstrom, M.K. Copass, «Amplitude of ventricular fibrillation waveform and outcome after cardiac arrest», Ann Intern Med 1985;102(1):53–5.

[18] X. Jaureguibeitia, G Zubia, U Irusta, E Aramendi, et al. «Shock Decision Algorithms for Automated External Defibrillators Based on Convolutional Networks», *IEEE Access*, vol. 8, pp. 154746-58, 2020.

[19] B. Chicote, E Aramendi, P Owens, M Daya, A Idris, «Value of capnography to predict defibrillation success in out-of-hospital cardiac arrest», *Resuscitation*, vol. 138, pp. 74-81, 2019

Verificación Sobre la Estabilidad de Electrogramas Fragmentados para la Caracterización del Sustrato Auricular en Pacientes con Fibrilación Auricular

E. Finotti[1], F. Hornero Sos[2], V. Bertomeu-González[3],
J. Osca Asensi[4], R. Alcaraz Martínez[5], J.J. Rieta Ibáñez[1]

[1] BioMIT.org, Departamento de Ingeniería Electrónica, Universitat Politecnica de Valencia, España, {aivra, jjrieta}@upv.es
[2] Servicio de Cirugía Cardiovascular, Hospital Clínico Universitario de Valencia, España, hornero_fer@gva.es
[3] Departamento de Medicina Clínica, Universidad Miguel Hernández, Alicante, España, vbertomeu@umh.es
[4] Unidad de Arritmias, Serv. de Cardiología, Hospital Universitari I Politecnic La Fe, Valencia, España, osca_joa@gva.es
[5] Grupo de Inv. en Electrónica, Telecom. y Bioingeniería, Univ. de Castilla-La Mancha, España, raul.alcaraz@uclm.es

Resumen

Para caracterizar el sustrato auricular en fibrilación auricular (FA) se ha investigado un amplio número de índices aplicados a electrogramas auriculares fragmentados complejos (CFAE). Sin embargo, se ha obviado la evaluación de su calidad, así como de su estabilidad interna (intra-canal) e intra-paciente. Este trabajo evalúa la estabilidad intra-canal e intra-paciente de 204 CFAEs bipolares de FA registrados en pacientes con FA paroxística ($n = 15$) y persistente ($n = 19$) aplicando indices no lineales. Para estimar las diferencias entre los CFAEs se ha utilizado el coeficiente de variación (CV) de la entropía muestral (SE) y el determinismo (DET) del análisis de cuantificación de recurrencias (RQA) de los datos. Además, también se analizaron las variaciones introducidas al descartar los segmentos de CFAEs ruidosos y con artefactos. El análisis intra-canal reportó una variación significativa del CV(%) para cualquier longitud de segmento analizado tanto para DET como para SE al descartar segmentos ruidosos, habiendo mayores disminuciones para segmentos más largos. Además, se observaron grandes variaciones de CV(%) para DET y SE en cualquier longitud de segmento en el análisis intra-paciente, pero en este caso, el descarte de segmentos ruidosos no mejoró resultados. La prueba de Kruskal-Wallis reportó diferencias estadísticamente significativas para DET y SE entre canales, independientemente del proceso de descarte. Por tanto, la alta variabilidad observada de los CFAEs ha demostrado que promediar los datos en un canal o entre diferentes canales puede conducir a una simplificación excesiva de la caracterización del sustrato auricular basada en CFAEs.

1. Introducción

La fibrilación auricular (FA) es la arritmia más comúnmente diagnosticada en la práctica clínica, con una prevalencia estimada del 1-2 % de la población en todo el mundo [1]. Gracias su éxito a largo plazo [2], la ablación cardíaca de las venas pulmonares (VP), procedimiento conocido como ablación con catéter (AC), se considera la mejor alternativa a los fármacos para el control de la frecuencia cardíaca [3]. Por ello, muchos esfuerzos de investigación en FA se centran en el mapeo del sustrato electrofisiológico auricular para identificar lugares ectópicos responsables de generar impulsos eléctricos anormales [4]. De hecho, las estadísticas han demostrado que la tasa de éxito de la AC depende estrictamente del área de ablación y del tipo de FA [5].

En el intento de personalizar el tratamiento de la FA a menudo se aplican índices no lineales a electrogramas auriculares fraccionados complejos (CFAE) para cuantificar el remodelado auricular y el sustrato electrofisiológico auricular, respaldar las decisiones de manejo clínico y sugerir el enfoque más apropiado para la ablación [6].

Recientemente se han publicado trabajos sobre la aplicación de índices no lineales a los electrogramas de FA. Con tal propósito, Ciaccio *et al.* midieron la repetitividad de los CFAEs [7] y cuantificaron el grado de heterogeneidad morfológica en sus deflexiones [8]; Acharya *et al.* utilizaron gráficos de recurrencia, medidas de RQA y de entropía para probar que el proceso de generación de los CFAEs es, de alguna manera, repetitivo incluso para secuencias tan cortas como un segundo [9]; Ndrepepa *et al.* y Ravi *et al.* usaron la longitud del ciclo auricular para demostrar que los pacientes con FA persistente tenían ciclos de duración más corta y un mayor grado de actividad desorganizada que los pacientes con FA paroxística [10, 11]; Finalmente, Sanders *et al.* emplearon análisis espectral para identificar puntos de alta frecuencia, encontrado distribuciones diferentes en FA paroxística frente a permanente [12].

Sin embargo, la mayoría de los trabajos mencionados no han considerado la estabilidad intra-canal e intra-paciente de los CFAEs, ni tampoco la calidad de señal. Con este tipo de omisiones pueden surgir dos problemas principales: primero, promediar valores entre canales y entre pacientes sin haber verificado previamente la estabilidad intra-canal e intra-paciente, puede conducir a una simplificación excesiva de los procesos fibrilatorios que acontecen en diferentes regiones de la aurícula izquierda; en segundo lugar, la

ISBN: 978-84-09-25491-0, pags. 422- 425

inclusión de ruido y artefactos no vinculados al mecanismo de FA, por ejemplo, tramos de amplitud saturada y muy pequeña o grabaciones distorsionadas que son debidas al mal contacto del electrodo en las paredes auriculares, puede producir resultados erróneos y poco fiables.

En este trabajo se ha evaluado la estabilidad intra-canal e intra-paciente de los CFAEs con dos índices no lineales: el determinismo (DET) del Análisis de Cuantificación de Recurrencias (RQA) y la Entropía Muestral (SE). La estabilidad se cuantificó con el coeficiente de variación (CV), expresado en %, y se ha considerado la presencia de segmentos ruidosos o con artefactos en los CFAEs, evaluando cómo afecta su descarte en los resultados finales. La idea consiste en que eliminar segmentos de mala calidad favorecerá la evaluación veraz de la estabilidad intra-canal e intra-paciente. Además, esto debería disminuir las diferencias en DET y SE en canales intra-paciente, ayudando así a cuantificar más verazmente el remodelado del sustrato auricular con valores fiables y representativos.

2. Materiales

2.1. Base de datos

Se analizaron 204 electrogramas de 16 s muestreados a 977 Hz extraídos de pacientes con FA paroxística (ParAF) y con FA persistente (PerAF) sin terapia de fármacos arrítmicos y que iban a ser sometidos a ablación por catéter de radiofrecuencia en el laboratorio de electrofisiología cardíaca. Para ParAF y PerAF se registraron respectivamente 90 y 114 CFAEs, que se identificaron observando criterios publicados [7], y que se registraron de las cuatro venas pulmonares (VP): superior izquierda (VPSI), inferior izquierda (VPII), superior derecha (VPSD), inferior derecha (VPID), la pared libre anterior (ANT) y posterior (POS) de la aurícula izquierda.

El análisis incorporó registros que tuvieran FA durante más de 10 minutos, sometiéndolos a enventanado rectangular de 1, 2 y 4 segundos, creando tres conjuntos de datos. Cada uno estuvo compuesto por 204 secuencias no superpuestas de 1, 2 y 4 segundos, respectivamente. En caso de ParAF en ritmo sinusal, la FA se indujo por estimulación rápida en el seno coronario o en la pared lateral de la aurícula derecha (intervalo de acoplamiento 250-200ms).

Antes de la digitalización, todas las señales se filtraron paso-banda con el sistema de adquisición (0,5-450 Hz, CardioLab, GE Healthcare, Waukesha, WI), que eliminó la deriva de la línea de base y el ruido de alta frecuencia. Después se remuestrearon a 1 KHz y se filtraron con un método basado en wavelets para la reducción de ruido [13].

2.2. Proceso de descarte de CFAEs

Los CFAEs fueron inspeccionados visualmente para evaluar la calidad de su señal. Inicialmente, ésta se evaluó en secuencias de 1 s, ya que era la longitud de ventana más pequeña. Luego se obtuvo la calidad de los segmentos de 2 y 4 s, tomando los subconjuntos mutuamente excluyentes de dos y cuatro secuencias de 1 s, y se estableció la calidad de cada grupo dependiendo de si contenía solo segmentos de calidad de 1 s o, al menos, un elemento para descartar. Es inevitable que el método descarte también algunas secuencias de buena calidad, pudiendo provocar una pérdida leve de información. Los subconjuntos se crearon siguiendo el orden temporal de los segmentos, de modo que hubiera un alineamiento temporal entre los tres conjuntos de datos de diferente longitud analizados.

La exploración de los CFAEs tuvo dos enfoques principales: distinguir los artefactos de las activaciones auriculares e identificar aquellos segmentos de CFAEs muy ruidosos. En principio, para encontrar los artefactos de baja frecuencia se visualizó el espectro de Welch de los segmentos de 1 s de longitud y se observaron aquellos segmentos que presentaban picos en el rango de 0-2 Hz. Aquellos que tenían, al menos, una activación auricular con anormalidad en su amplitud y/o su forma fueron descartados. En segundo lugar, para reconocer las partes de los CFAEs sin activaciones auriculares, se extrajeron los rangos de amplitud de cada secuencia de 1 s de longitud y se descartaron los segmentos que presentaban rangos alrededor de cero mV. A parte, se evaluó el porcentaje de descartes para cada longitud de ventana, el lugar de registro y el tipo de FA.

3. Métodos

Las tres bases de datos que incluyen segmentos de 1, 2 y 4 s de CFAEs se procesaron posteriormente con métodos no lineales para obtener SE y DET. A continuación se presenta una breve descripción de este procesamiento.

3.1. Entropía muestral

El índice SE asignó a cada secuencia un valor no negativo que refleja su complejidad, con valores más grandes correspondientes a una mayor irregularidad en los datos [14]. La longitud máxima de la plantilla m se fijó en 2 muestras y la tolerancia r en 0,35 veces la desviación estándar del segmento, como en trabajos anteriores [15]. SE se aplicó a los conjuntos de datos de las secuencias de 1, 2 y 4 s de longitud. Los resultados serán presentados indicando el rango, la media y la desviación estándar.

3.2. Análisis de cuantificación de recurrencias

Se utilizaron gráficos de recurrencia (GR) para encontrar patrones dentro de las secuencias de 1, 2 y 4 s. Para reconstruir los GRs a partir de los CFAEs, considerados como series de tiempo, sus trayectorias se proyectaron en un espacio fásico d-dimensional. El espacio se obtuvo por el teorema de Takens, con una dimensión de reconstrucción d y un retardo en el tiempo τ, definidos con los métodos de vecinos más próximos e información mutua, respectivamente [16]. A continuación, se produjeron los GR representando cada par de estados de trayectoria como un punto negro, si su distancia estaba por debajo de un umbral especificado, y blanco en caso contrario. El umbral se fijó en el 10 % de la media del diámetro del espacio fásico [17].

La medida RQA seleccionada para el estudio ha sido el determinismo (DET), que cuantifica la predictibilidad de la secuencia, midiendo el porcentaje de puntos de recurren-

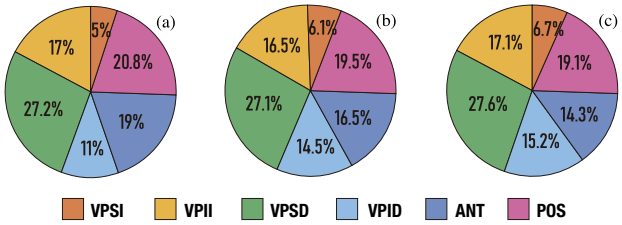


Figura 1. *Distribución de porcentajes de segmentos ruidosos o artefactados descartados de las seis zonas de registro analizadas para los datos de 1 s de longitud (a), los de 2 s (b) los de 4 s (c).*

	SE sin descarte		SE con descarte	
	rango	media ± std	rango	media ± std
1s	[0.004-0.362]	0.137 ± 0.033	[0.013-0.362]	0.138 ± 0.032
2s	[0.007-0.312]	0.134 ± 0.025	[0.018-0.312]	0.135 ± 0.023
4s	[0.008-0.3]	0.132 ± 0.017	[0.02-0.3]	0.134 ± 0.014

Tabla 1. *Resultados de la análisis de los descriptores estadísticos de entropía muestral para conjuntos de datos de los segmentos de 1, 2 y 4 s de longitud, con y sin descartes. Los valores informados se obtuvieron promediando entre los CFAEs.*

cia que pertenecen a líneas diagonales del GR con longitud mínima $l_{min} > 50$. Los resultados se reportan con el rango, la media y la desviación estándar.

3.3. Estabilidad intra-canal e intra-paciente

La estabilidad de los parámetros, tanto dentro del canal como dentro del paciente, se verificó para los conjuntos de datos de 1, 2 y 4 s de longitud, utilizando el coeficiente de variación (CV) para medir la dispersión. El coeficiente de variación, expresado en %, se calculó como CV $= \sigma/\mu$, siendo σ la desviación estándar y μ la media del índice en estudio. Los análisis se repitieron descartando los segmentos de CFAEs ruidosos o con artefactos y se evaluó cualquier aumento y disminución del CV.

Además, se seleccionó la prueba de Kruskal-Wallis para evaluar si el índice analizado en cada zona auricular registrada procedía de la misma distribución. Previamente, se verificó la hipótesis de homocedasticidad con la prueba de Breush-Pagan. Para facilitar el proceso, se compararon los valores medianos de los grupos numéricamente y gráficamente, de modo que cualquier inexactitud podría haber sido detectada. La hipótesis nula (H_0) probada fue que los rangos medios de DET (o SE) entre las zonas de registro intra-paciente eran los mismos. La aceptación de H_0 demuestra que el remodelado auricular es similar en diferentes zonas de la aurícula izquierda. Al revés, el rechazo de H_0 lleva a la conclusión de que el promedio entre los canales provoca una simplificación excesiva de la condición del sustrato auricular en la aurícula izquierda. La prueba de Kruskal-Wallis se realizó en los conjuntos de datos de 1, 2 y 4 segundos, con y sin descartes.

4. Resultados y discusión

De la inspección de los CFAEs, los segmentos de 1 s descartados fueron el 5.6 %, mientras que en los segmentos de 2 y 4 s los porcentajes fueron levemente mayores 8.1 % y 12.9 %, respectivamente, debido a la pérdida adicional

	DET sin descarte		DET con descarte	
	rango	media ± std	rango	media ± std
1s	[0.075-0.998]	0.561 ± 0.120	[0.075-0.963]	0.559 ± 0.113
2s	[0.122-1]	0.600 ± 0.101	[0.122-0.958]	0.595 ± 0.091
4s	[0.092-1]	0.631 ± 0.072	[0.124-0.942]	0.621 ± 0.058

Tabla 2. *Resultados de determinismo para los segmentos de 1, 2 y 4 s de longitud, con y sin descartes. Los valores informados se obtuvieron promediando entre CFAEs.*

	1-s	ΔD1-s	2-s	ΔD2-s	4-s	ΔD4-s
DET	23.3 %	−15.6 %	19.1 %	−22.8 %	13.3 %	−47.9 %
SE	26.6 %	−16.1 %	20.5 %	−20.1 %	13.9 %	−42.2 %

Tabla 3. *Valores intra-canal del CV para DET y SE para segmentos de 1, 2 y 4 s y variación de CV (ΔD) introducida descartando segmentos de baja calidad.*

de información. La Figura 1 muestra la distribución de los segmentos descartados para cada zona auricular de registro y longitud de segmento. En particular, el número de descartes en la VPSI es bastante bajo en comparación con los demás canales, contrastando así con la VPSD, en la que los descartes fueron más frecuentes. La proporción de segmentos descartados en pacientes con FA paroxística y persistente fue similar: para los segmentos de 1 s de longitud, el 46,4 % de los segmentos descartados eran de ParAF y el 53,6 % de PerAF; para segmentos de 2 s de longitud, el 44,4 % eran de ParAF y 55,6 % de PerAF y para segmentos de 4 s, el 43,8 % era de ParAF y el 56,2 % de PerAF.

El rango, la media y la desviación estándar de los descriptores estadísticos promediados resultantes de la aplicación de SE a los segmentos de 1, 2 y 4 s, con y sin descartes, se presentan en la Tabla 1. Con descartes, los rangos se redujeron conforme aumentaba el límite inferior debido, principalmente, a la eliminación de la deriva de señales que generalmente presentaban una alta amplitud y bajos valores de SE. También las desviaciones estándar disminuyeron con los descartes.

Los descriptores estadísticos de DET se calcularon igual que para SE y sus resultados se indican en la Tabla 2. Aquí los rangos también se redujeron con los descartes, pero esta vez debido al límite superior, que asumía valores más bajos, disminuyendo así también las desviaciones estándar. De hecho, DET y SE son, hasta cierto punto, medidas complementarias, ya que una determina la previsibilidad y la otra la complejidad de una serie temporal.

El análisis intra-canal mostró una variación significativa de CV(%) en cualquier longitud de segmento tanto para DET como para SE (ver Tabla 3). Los segmentos descartados mejoraron la estabilidad disminuyendo el CV con disminuciones más profundas para segmentos más largos. Estas variaciones fueron mayores en promedio para PerAF con DET = 29.1 % y SE = 37.6 % que para ParAF, con DET = 19.6 % y SE = 31.8 % respectivamente.

La estabilidad intra-paciente también proporcionó grandes variaciones en CV (%) para DET e incluso mayores para SE en cualquier longitud de segmento, como se muestra en

	1-s	ΔD1-s	2-s	ΔD2-s	4-s	ΔD4-s
DET	23.9 %	+2.0 %	24.8 %	+3.5 %	24.1 %	+7.7 %
SE	34.2 %	+0.5 %	34.8 %	−0.1 %	35.9 %	−0.3 %

Tabla 4. *Valores intra-paciente del CV para DET y SE para segmentos de 1, 2 y 4 s y variación de CV (ΔD) introducida descartando segmentos de baja calidad..*

la Tabla 4. En este caso, descartar segmentos fue inútil y el CV proporcionó variaciones limitadas.

La prueba de Kruskal-Wallis reveló que las características subyacentes de los CFAEs eran en su mayoría diferentes en cada zona de registro analizada, mostrando una gran variabilidad de los índices intra-paciente. Para conjuntos de datos de 1 s de longitud, la H_0 siempre se rechazó para SE, mientras que para DET se aceptó solo una vez sin descartes. Para 2 s de longitud, H_0 todavía se rechazó siempre para SE, mientras que para DET se aceptó en dos casos sin descartes y en un caso con descartes. Para 4 s de longitud, se aceptó la hipótesis nula en siete casos, uno para SE con descartes, uno para DET sin descartes y cinco para DET con descartes. Sin embargo, una vez visualizada la distribución correspondiente para verificar la exactitud de los resultados, los valores medios se revelaron diferentes. La inexactitud encontrada para los conjuntos de datos de 4 s de longitud se justifica por el hecho de que el tamaño de la muestra es bastante limitado, por lo que la prueba de la muestra no puede seguir una distribución χ^2.

5. Conclusiones

El proceso de descarte favoreció más la estabilidad intra-canal que la intra-paciente. Sin embargo, la alta variabilidad del CV ha revelado que promediar los datos en un canal, así como entre los sitios de registro, puede conducir a una simplificación excesiva de la caracterización del sustrato auricular basado en CFAEs. En muchos casos, los análisis intra-paciente revelaron que solo una parte de las zonas auriculares registradas presentaba características de sustrato auricular similares, identificadas por valores de DET y SE similares. Esto refuerza la conclusión de que promediar, tal como hacen la mayoría de trabajos previos, deteriora la singularidad del sustrato auricular por zonas de registro, la cual es la base para el desarrollo de tratamientos personalizados en ablación con catéter.

Agradecimientos

El presente trabajo ha sido cofinanciado por los proyectos de investigación competitiva DPI2017-83952-C3 de MINECO-AEI-FEDER-UE, SBPLY/17/180501/000411 de la JCCLM y AICO/2019/036 de la GVA.

Referencias

[1] January CT, Wann LS, et al. 2019 AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation. Circulation 07 2019;140(2):e125–e151.

[2] Turagam MK, Garg J, Whang W, Sartori S, Koruth JS, Miller MA, Langan N, Sofi A, Gomes A, Choudry S, Dukkipati SR, Reddy VY. Catheter ablation of atrial fibrillation in patients with heart failure: A meta-analysis of randomized controlled trials. Ann Intern Med 01 2019;170(1):41–50.

[3] Hesselson AB. Catheter ablation in the treatment of atrial fibrillation. Int J Angiol Jun 2020;29(2):108–112.

[4] Kottkamp H, Bender R, Berg J. Catheter ablation of atrial fibrillation: how to modify the substrate? J Am Coll Cardiol Jan 2015;65(2):196–206.

[5] Latchamsetty R, Morady F. Catheter ablation of atrial fibrillation. Heart Fail Clin Apr 2016;12(2):223–33.

[6] Cirugeda-Roldán E, Novak D, Kremen V, Cuesta-Frau D, Keller M, Luik A, Srutova M. Characterization of complex fractionated atrial electrograms by sample entropy: An international multi-center study. Entropy Oct 2015; 17(12):7493–7509. ISSN 1099-4300.

[7] Ciaccio EJ, Biviano AB, Whang W, Vest JA, Gambhir A, Einstein AJ, Garan H. Differences in repeating patterns of complex fractionated left atrial electrograms in longstanding persistent atrial fibrillation as compared with paroxysmal atrial fibrillation. Circ Arrhythm Electrophysiol Aug 2011;4(4):470–7.

[8] Ciaccio EJ, Biviano AB, Whang W, Gambhir A, Garan H. Different characteristics of complex fractionated atrial electrograms in acute paroxysmal versus long-standing persistent atrial fibrillation. Heart Rhythm Sep 2010;7(9):1207–15.

[9] Acharya UR, Faust O, Ciaccio EJ, Koh JEW, Oh SL, Tan RS, Garan H. Application of nonlinear methods to discriminate fractionated electrograms in paroxysmal versus persistent atrial fibrillation. Comput Methods Programs Biomed Jul 2019;175:163–178.

[10] Ndrepepa G, Karch MR, Schneider MAE, Weyerbrock S, Schreieck J, Deisenhofer I, Zrenner B, Schömig A, Schmitt C. Characterization of paroxysmal and persistent atrial fibrillation in the human left atrium during initiation and sustained episodes. J Cardiovasc Electrophysiol Jun 2002; 13(6):525–32.

[11] Ravi KC, Krummen DE, Tran AJ, Bullinga JR, Narayan SM. Electrocardiographic measurements of regional atrial fibrillation cycle length. Pacing Clin Electrophysiol Mar 2009;32 Suppl 1:S66–71.

[12] Sanders P, Berenfeld O, Hocini M, Jaïs P, Vaidyanathan R, Hsu LF, Garrigue S, Takahashi Y, Rotter M, Sacher F, Scavée C, Ploutz-Snyder R, Jalife J, Haïssaguerre M. Spectral analysis identifies sites of high-frequency activity maintaining atrial fibrillation in humans. Circulation Aug 2005; 112(6):789–97.

[13] Martínez-Iniesta M, Ródenas J, Rieta JJ, Alcaraz R. The stationary wavelet transform as an efficient reductor of powerline interference for atrial bipolar electrograms in cardiac electrophysiology. Physiol Meas 07 2019; 40(7):075003.

[14] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol Jun 2000;278(6):H2039–49.

[15] Alcaraz R, Abásolo D, Hornero R, Rieta JJ. Optimal parameters study for sample entropy-based atrial fibrillation organization analysis. Comput Methods Programs Biomed Jul 2010;99(1):124–32.

[16] Trulla L, Giuliani A, Zbilut J, Webber C. Recurrence quantification analysis of the logistic equation with transients. Physics Letters A 1996;223(4):255 – 260. ISSN 0375-9601.

[17] Schinkel S, Dimigen O, Marwan N. Selection of recurrence threshold for signal detection. The European Physical Journal Special Topics 10 2008;164:15–53.

Modelo predictivo del retorno de circulación espontánea en la parada cardiorrespiratoria utilizando el ECG y la impedancia torácica

J. Urteaga Urizarbarrena[1], E. Aramendi Ecenarro[1], A. Elola Artano[1], U. Irusta Zarandona[1], A. Idris[2]

[1] Departamento de Ingeniería de Comunicaciones, Universidad del País Vasco, 48013 Bilbao, España

[2] Department of Emergency Medicine, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

Resumen

El análisis de los diferentes tipos de ritmo cardíaco durante la parada cardiorrespiratoria y la predicción de su evolución permitiría ajustar la terapia de resucitación a cada paciente. El ritmo con actividad eléctrica sin pulso (AESP) es el ritmo inicial predominante durante la parada cardiorrespiratoria extrahospitalaria, y es de gran interés disponer de modelos que predigan el retorno espontáneo de circulación (RCE). En este trabajo se propone un método automático que discrimina los casos de AESP que evolucionan a RCE de los que no recuperan el pulso. El modelo combina parámetros de las señales de electrocardiograma (ECG) e impedancia torácica (IT) adquiridas con los parches del desfibrilador. La base de datos consiste en 185 pacientes (73 con RCE) de los que se extrajeron 1600 segmentos (432 con RCE). Aplicando una validación cruzada de 10 particiones y un clasificador de máquinas de vectores de soporte (SVM), se demuestra que la IT añade valor discriminativo al modelo basado en ECG. Para un clasificador SVM con un núcleo polinómico de orden 2 se obtuvo una sensibilidad del 79.8%, una especificidad del 85.5% y un área bajo la curva ROC de 0.91.

1. Introducción

La parada cardiorrespiratoria se define como el cese repentino e inesperado de las funciones cardiacas y respiratorias del paciente, y sigue siendo una de las mayores causas de muerte en los países industrializados. La terapia de resucitación cardiopulmonar (RCP) básica incluye compresiones torácicas junto a las ventilaciones, para mantener el flujo necesario de oxigenación en los órganos vitales, y la desfibrilación eléctrica, con el objetivo de restaurar la funcionalidad del sistema cardiaco. El objetivo principal de la terapia RCP es la recuperación de la circulación espontánea (RCE).

El ritmo regular con complejos QRS pero sin pulso (AESP) es uno de los ritmos más frecuentes durante la parada cardiorrespiratoria extrahospitalaria. Predecir la evolución de dicho ritmo es de gran interés ya que permite adecuar la terapia RCP y la terapia farmacológica, y mejorar así la probabilidad de supervivencia del paciente. Aunque hay estudios clínicos que analizan características morfológicas del electrocardiograma (ECG) para diferenciar los casos de AESP que evolucionan a RCE de los que no, los resultados son contradictorios y el volumen de casos escaso [1]–[3].

Por otro lado, la impedancia torácica (IT) es una señal que se adquiere típicamente durante la parada cardiaca extrahospitalaria junto con el ECG a través de los parches de desfibrilación. Cuando el corazón presenta cierta actividad mecánica (aunque no sea suficiente para mantener la circulación del paciente), la IT muestra una componente de fluctuaciones que se ha demostrado que está relacionada con valores elevados de presión sistólica y diastólica y con el RCE [4]. Sin embargo, obtener una señal continua de presión arterial invasiva en el ámbito extrahospitalario es difícil.

El objetivo de este estudio es proponer el primer sistema automático para la predicción de evolución de los ritmos AESP en parada cardiorrespiratoria extrahospitalaria. Se compararán modelos combinados de ECG e IT con modelos basados exclusivamente en el ECG y se analizará su poder para discriminar los casos que evolucionan a RCE de los que no recuperan el pulso.

2. Materiales

En este estudio se han considerado pacientes tratados por los servicios de emergencia *Dallas FortWorth Center for Resuscitation Research* (Dallas, Texas), utilizando el monitor/desfibrilador *Philips HeartStart MRx*. Dicho dispositivo adquiere la señal de ECG e IT a través de los parches de desfibrilación a 250 Hz y 200 Hz, respectivamente. Se dispone además de la información clínica de cada paciente, que incluye si el paciente tuvo RCE y en qué instante, anotado por el médico en escena.

Se seleccionaron, por un lado, los casos sin RCE anotada (noRCE) que no presentan intervalos sospechosos de RCE (2 min con ritmo cardíaco organizado y niveles de presión de CO_2 expirado superiores a 25 mmHg, o pacientes transportados al hospital). Se consideraron los casos RCE que fueron anotados por los clínicos, siempre que no hubiera terapia de desfibrilación o RCP posterior. El detalle del proceso de anotación se puede encontrar en [5, 6].

La base de datos final contiene los casos de 185 pacientes (73 con RCE). Siguiendo el criterio propuesto en [3], para aquellos pacientes con RCE se consideran los últimos 12 minutos antes del instante de RCE, y para pacientes sin RCE, se consideran los últimos 12 minutos del episodio. Se extrajeron un total de 1600 segmentos de 5 s de las señales ECG e IT sin artefactos (432 pertenecen al grupo RCE), una mediana (percentil 25-percentil 75) de 6 (2-11) segmentos por paciente.

ISBN: 978-84-09-25491-0, pags. 426- 429

3. Métodos

Se propone un método que permite discriminar los segmentos del grupo noRCE, clase 0, de los segmentos del grupo con RCE, clase 1. Para ello se aplica un preprocesado a la señal ECG e IT, se extraen características morfológicas de ambas y se aplica un clasificador basado en algoritmos de aprendizaje de máquina.

3.1. Extracción de características

El segmento de ECG se preprocesa con un filtro Butterworth de orden 4 entre 0.5 y 35 Hz utilizando filtrado de fase cero para eliminar las variaciones de la línea base y el ruido de alta frecuencia. La señal es enventanada con una función Tuckey y se calcula el módulo de su FFT de 4096 puntos, S[f], para obtener la característica v_1:

$$v_1 = \frac{2}{4096}\sum_{f=2}^{48} S[f] \times f$$

Dicha característica, también conocida como AMSA, definida en el intervalo de 2-48 Hz ha demostrado ser muy efectiva para discriminar los intervalos con RCE [8].

Los segmentos de la señal IT, previamente remuestreada a 250 Hz, se filtran 1-10 Hz utilizando un filtro Butterworth (orden 4) y filtrado de fase cero. Se calcula la energía logarítmica de la señal a partir de la señal preprocesada *z*[n]:

$$v_2 = \sum_i \ln(z^2[i])$$

Ambas características fueron normalizadas a media cero y varianza uno.

3.2. Clasificación

En este estudio se consideran varios clasificadores: regresión logística (LR), LR con un núcleo gaussiano (KLR), *máquina de vectores de* soporte (del inglés Support Vector Machine, SVM) con un núcleo polinómico de orden 2 (SVM2), SVM con un núcleo polinómico de orden 3 (SVM3), y SVM con un núcleo gaussiano (SVMG).

Para ajustar el desequilibro en el número de segmentos por paciente, se ajustan los pesos de la función de coste de cada clasificador de manera que todos los pacientes tengan la misma incidencia en el entrenamiento del clasificador. El desbalanceo de clases se compensa multiplicando por 1.5 los pesos de los segmentos del grupo RCE.

Para la optimización de los hiperparámetros de cada clasificador se aplicó una validación cruzada en los datos de entrenamiento y se aplicó una optimización bayesiana con procesos gaussianos. Los hiperparámetros optimizados fueron la anchura del núcleo y el termino de regularización para el KLR, y la anchura del núcleo y el *soft margin* para SVMG.

3.3. Evaluación y métricas de desempeño

En este estudio se utiliza la validación cruzada con 10 particiones. En cada iteración, los hiperparámetros necesarios se optimizan utilizando las 9 particiones de entrenamiento y aplicando una validación cruzada entre dichos grupos.

Las métricas de desempeño incluyen la probabilidad de clasificar correctamente los segmentos RCE (Se, sensibilidad), clasificar correctamente los segmentos noRCE (Sp, especificidad) y el área bajo la curva ROC (AUC). Para su cálculo se pesan todos los pacientes equitativamente. Todas las métricas se representan como mediana (percentil 25-percentil 75).

4. Resultados

En primer lugar, se realizó un análisis univariable utilizando LR. El modelo basado exclusivamente en v_1 mostró un AUC de 0.84 (0.80-0.91) y el modelo basado exclusivamente en v_2 0.83 (0.79-0.86). Al combinar ambas características, el AUC obtenido ascendió a 0.90 (0.84-0.91). En la Figura 1 se puede observar la mediana de la probabilidad a posteriori obtenida mediante el clasificador LR antes del instante 0, el cual se corresponde con el instante de RCE en el grupo RCE o con el final del episodio en el grupo noRCE.

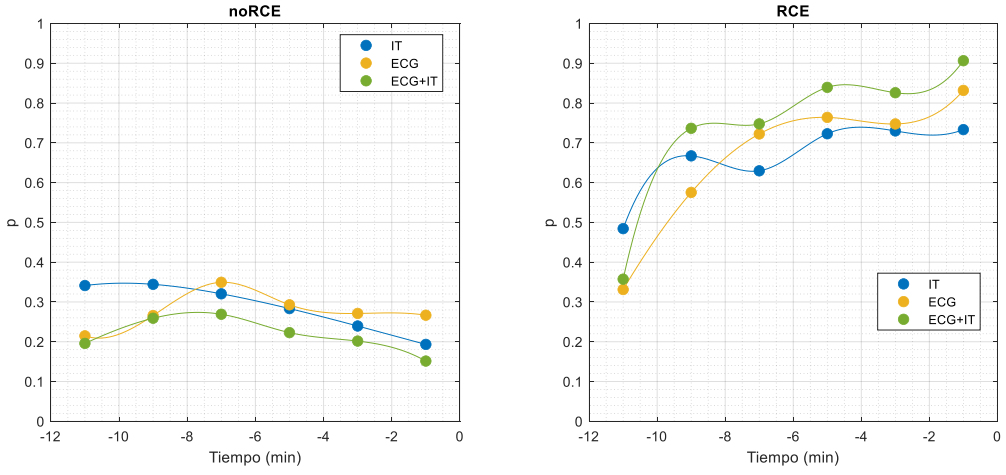


Figura 1. *Evolución de la probabilidad a posteriori obtenida mediante el clasificador LR utilizando solamente la IT, el ECG o la combinación de ambos. El instante 0 representa el final del episodio en pacientes sin RCE y el instante de RCE para aquellos pacientes con RCE.*

Los pacientes sin RCE mantienen una mediana de probabilidad constante por debajo de p=0.5, mientras que en los pacientes con RCE el modelo muestra una evolución de probabilidad ascendente por encima de p=0.5 a medida que se acerca el instante de RCE. Para pacientes con RCE, la mediana aumenta cuando se acerca el instante de RCE y se puede observar mayor margen desde p=0.5 cuando se utilizan ambas señales.

En la Tabla 1 se muestran las métricas de desempeño para los diferentes clasificadores. Se observa que el modelo KLR muestra el mejor valor de Se, mientras que el modelo SVM2 presenta la mejor Sp y AUC.

	Se	Sp	AUC
LR	76.3(64.9-86.7)	81.2(79.9-86.2)	0.90(0.84-0.91)
KLR	**84.7(72.0-90.9)**	82.6(75.2-84.5)	0.88 (0.86-0.92)
SVM2	79.8(70.2-87.4)	**85.5(82.2-87.7)**	**0.91(0.89-0.92)**
SVM3	83.2(70.0-86.9)	80.1(74.1-84.1)	0.90(0.88-0.93)
SVMG	80.9(64.9-86.9)	83.0(77.0-86.7)	0.91(0.85-0.91)

Tabla 1. *Métricas de desempeño para diferentes clasificadores, representados como mediana (percentil 25-percentil 75)*

En la Figura 2 se muestran los diagramas de dispersión de los parámetros v_1 y v_2 para el clasificador SVM2 considerando la base de entrenamiento y la base de test para aquella partición más cercana al AUC mediano. Puede observase como los valores más elevados de ambos parámetros se asocian a los casos positivos, RCE, y que la región de delimitación del clasificador permite discriminar con éxito ambos grupos.

5. Discusión

La predicción del pronóstico favorable durante la parada cardíaca extrahospitalaria es un aspecto altamente interesante en el tratamiento de la parada cardiorrespiratoria extrahospitalaria, ya que permitiría ajustar el tratamiento al paciente aumentando sus probabilidades de supervivencia. El ritmo AESP es el ritmo inicial más frecuente en la parada cardiorrespiratoria extrahospitalaria, y aunque se han dedicado esfuerzos a predecir su evolución a RCE, las contribuciones hasta el momento han sido fundamentalmente clínicas, con parámetros medidos manualmente y con bases de datos limitadas. En este trabajo se propone el primer método automático que incluye parámetros del ECG y de la IT, que son las señales habitualmente adquiridas mediante los parches de desfibrilación de cualquier monitor/desfibrilador comercial.

Contribuciones previas han propuesto parámetros del ECG asociadas a la frecuencia cardíaca o la duración del complejo QRS como predictores de la evolución de un ritmo AESP a un ritmo con pulso. Sin embargo, las conclusiones obtenidas con parámetros medidos manualmente y con bases de datos limitadas no han sido concluyentes.

Alonso et al. propusieron un modelo de regresión logística con parámetros del ECG para predecir la transición inmediata (en un intervalo de 3 s) de AESP a ritmo con pulso [7]. Aunque los resultados fueron prometedores con un AUC de 0.79, estos métodos no son comparables con los propuestos en este estudio para la predicción con varios minutos de antelación.

En este estudio se demuestra que añadir una característica de la IT mejora los resultados. Durante PEA, el corazón puede presentar cierta actividad mecánica, pero puede no ser suficiente para mantener la circulación del paciente. Sin

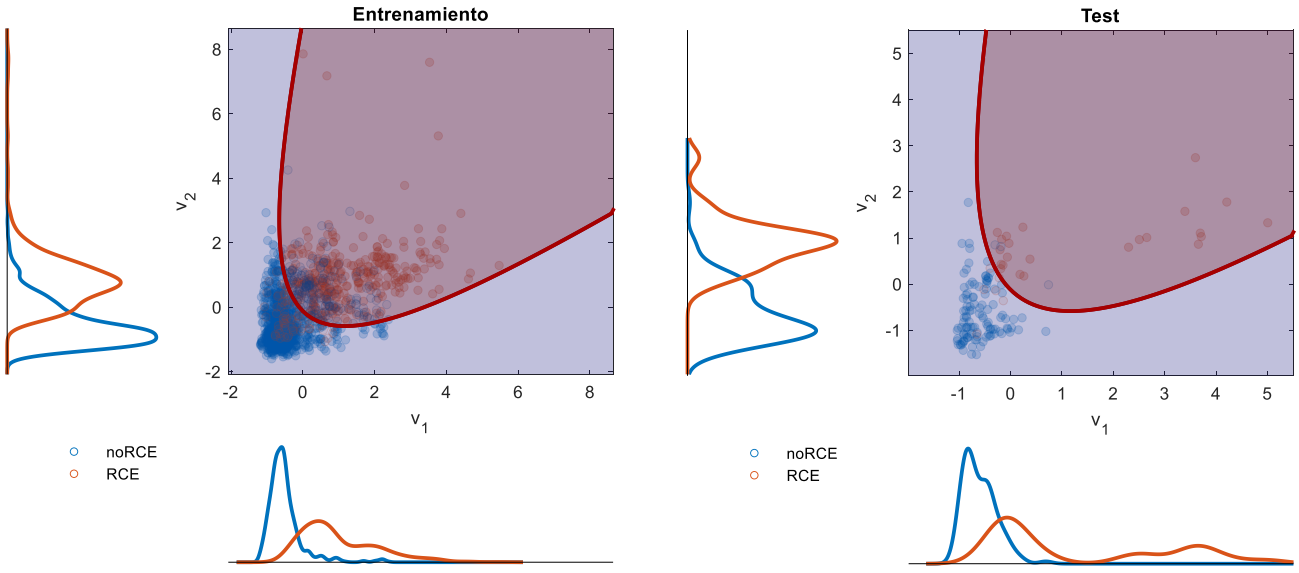


Figura 2. *Diagramas de dispersión y distribuciones de ambas características para el conjunto de entrenamiento y test en aquella partición más cercana al AUC mediano. La región de delimitación que se muestra es la obtenida con el clasificador SVM2.*

embargo, este estado circulatorio también conocida como pseudo AESP, se asocia con mejores pronósticos [9].

Una actividad mecánica mínima causa un flujo mínimo de sangre y esto provoca pequeños cambios en la IT. En cambio, cuando el corazón presenta una actividad eléctrica pero no mecánica, las fluctuaciones observadas en la IT son menores. Por ello, los valores de v_2 para aquellos pacientes que presentan una RCE son mayores.

En la Figura 1 se puede observar que el modelo muestra una mayor capacidad predictiva según se acerca el instante de RCE. Esta conclusión es coherente con las conclusiones propuestas en [3], ya que demostraron que el ritmo cardiaco y la duración del complejo QRS cambian según se acerca el instante de RCE. En este estudio, se demuestra que la variación de la energía logarítmica de la IT refleja esa tendencia.

Un análisis más detallado del poder discriminativo del modelo en diferentes periodos del episodio de resucitación permitiría un análisis temporal del estado hemodinámico del paciente. Es necesario disponer de bases de datos más extensas para llevar a cabo este tipo de estudio, que sin duda contribuirá a mejorar el tratamiento del paciente en parada.

6. Conclusiones

Se ha propuesto un modelo que predice el RCE en pacientes con AESP. Mediante métodos de aprendizaje máquina y utilizando las señales de ECG e IT adquiridas por los parches de desfibrilación se ha diseñado un algoritmo que podría implementarse en cualquier desfibrilador comercial. Este tipo de sistema permitiría optimizar el tratamiento del paciente en parada cardiorrespiratoria extrahospitalaria.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el Ministerio de Ciencia, Innovación y Universidades a través del proyecto RTI2018-101475-BI00, en conjunto con el Fondo Europeo de Desarrollo Regional (FEDER), y en parte por el Gobierno Vasco por medio del proyecto IT-1229-19.

Referencias

[1] C. Weiser *et al.*, «Initial electrical frequency predicts survival and neurological outcome in out of hospital cardiac arrest patients with pulseless electrical activity», *Resuscitation*, vol. 125, pp. 34-38, abr. 2018.

[2] D. Bergum *et al.*, «ECG patterns in early pulseless electrical activity-Associations with aetiology and survival of in-hospital cardiac arrest», *Resuscitation*, vol. 104, pp. 34-39, jul. 2016.

[3] G. W. Skjeflo, T. Nordseth, J. P. Loennechen, D. Bergum, y E. Skogvoll, «ECG changes during resuscitation of patients with initial pulseless electrical activity are associated with return of spontaneous circulation», *Resuscitation*, vol. 127, pp. 31-36, 2018.

[4] A. Elola *et al.*, «Abstract 237: Using the Thoracic Impedance to Predict Measures from Invasive Arterial Blood Pressure in Out-Of-Hospital Cardiac Arrest», *Circulation*, vol. 140, n.º Suppl_2, pp. A237-A237, nov. 2019.

[5] A. Elola *et al.*, «Capnography: A support tool for the detection of return of spontaneous circulation in out-of-hospital cardiac arrest», *Resuscitation*, vol. 142, pp. 153-161.

[6] A. Elola *et al.*, «Deep Neural Networks for ECG-Based Pulse Detection during Out-of-Hospital Cardiac Arrest», *Entropy*, vol. 21, n.º 3, p. 305, mar. 2019.

[7] E. Alonso *et al.*, «Beyond ventricular fibrillation analysis: Comprehensive waveform analysis for all cardiac rhythms occurring during resuscitation», *Resuscitation*, vol. 85, n.º 11, p. 1541-1548. 2014.

[8] A. Elola *et al.*, «ECG characteristics of Pulseless Electrical Activity associated with Return of Spontaneous Circulation in Out-of-Hospital Cardiac Arrest», *Resuscitation*, vol. 130, p. e54, sep. 2018.

[9] A. Elola *et al.*, «Multimodal algorithms for the classification of circulation states during out-of-hospital cardiac arrest», *IEEE transactions on bio-medical engineering*.

Confiabilidad de las Ondas de Activación Local para Evaluar el Sustrato Auricular de Pacientes con Fibrilación Auricular Paroxística

A. Vraka[1], F. Hornero Sos[2], V. Bertomeu-González[3],
J. Osca Asensi[4], R. Alcaraz Martínez[5], J.J. Rieta Ibáñez[1]

[1] BioMIT.org, Departamento de Ingeniería Electrónica, Universitat Politecnica de Valencia, España, {aivra, jjrieta}@upv.es
[2] Servicio de Cirugía Cardiovascular, Hospital Clínico Universitario de Valencia, España, hornero_fer@gva.es
[3] Departamento de Medicina Clínica, Universidad Miguel Hernández, España, vbertomeu@umh.es
[4] Unidad de Arritmias, Serv. de Cardiología, Hospital Universitari I Politecnic La Fe, España, osca_joa@gva.es
[5] Grupo de Inv. en Electrónica, Telecom. y Bioingeniería, Univ. de Castilla-La Mancha, España, raul.alcaraz@uclm.es

Resumen

El análisis de electrogramas (EGMs) del seno coronario (SC) durante la ablación por catérer (AC) de fibrilación auricular (FA) es clave para la evaluación del sustrato auricular. No obstante, los canales del catéter de registro suelen verse afectados por mal contacto y por el movimiento cardíaco. Este trabajo estudia los canales más fiables para preservar las dinámicas de la FA durante el ritmo sinusal (RS). Se han empleado 44 registros bipolares multicanal del SC de 60-300 segundos de 28 pacientes con FA paroxística sometidos a AC. Cada registro contenía cinco canales: distal, medio-distal, medio, medio-proximal, proximal. En cada canal se calcularon características de las ondas de activación local (OALs): duración, área y correlación entre morfologías dominantes. Se han utilizado los tests de Kruskal-Wallis, Mann-Whitney y Mann-Whitney con corrección de Bonferroni para la comparación multicanal, por pares y uno frente a los demás, respectivamente. El canal distal presentó resultados menos fiables, con OALs de duración larga ($p = 0{,}047$), amplitud baja ($p_{max} < 0{,}003$) y menos área ($p_{max} < 0{,}020$) y con correlación entre sus OALs y las OALs del canal proximal relativamente baja ($85{,}57\,\%$). Las OALs entre los canales medio y medio-próximal fueron más consistentes, con correlación entre sus OALs muy altas ($95{,}02\,\%$), presentando además duración más corta ($p = 0{,}084$) y amplitud ($p_{max} < 0{,}050$) y áreas mayores ($p_{max} < 0{,}070$). En consecuencia, los canales medios y medio-proximales son más robustos, preservando al máximo las dinámicas de la FA, por ello se aconseja su empleo para evaluar el sustrato auricular de pacientes con FA durante el RS.

1. Introducción

La fibrilación auricular (FA) es actualmente la arritmia cardiaca más predominante, adicionalmente acompañada con riesgos significativos para la salud [1]. Por tanto, su tratamiento exitoso y la cartografía precisa del sustrato auricular son de importancia primordial. Actualmente, la ablación por catéter (AC) de las venas pulmonares (VPs) logra un alto porcentaje de éxito en pacientes con FA paroxística (FAP). No obstante, tanto para una evaluación precisa del proceso de AC, como para obtener una perspectiva más en detalle de las dinámicas de la propia FA, se requiere un conocimiento más profundo sobre el sustrato auricular. Para ello, se suele registrar la actividad cardiaca de varios sitios de las aurículas durante la AC [2].

En las aurículas, aparte de las propias VPs, el seno coronario (SC) es uno de los sitios más analizados y estudiados [2–4]. Al ser un puente interauricular, ubicado entre la auriculas, el SC puede proporcionar información muy valiosa [5]. Puede desencadenar taquicardia auricular, que conduce a FA [6] o contribuir a la perpetuación de la propia FA [3, 7]. El análisis de electrogramas (EGMs) del SC, junto con EGMs de las VPs, se emplea para la evaluación del sustrato auricular durante la propia AC. Así, valores de amplitud bajos o presencia de EGMs complejos y fragmentados (CFAEs) indican fibrosis y remodelado auricular [7]. Recientemente, un estudio ha reportado la terminación de FA persistente después de ablación del SC en un paciente que se había sometido anteriormente a AC de VPs y de la pared posterior auricular izquierda sin éxito [8]. Otro estudio reciente ha reportado que el análisis de complejidad de Kolmogórov en EGMs del SC antes de la AC puede predecir la terminación de la FA [9].

La colocación del catéter en el SC es una tarea complicada [10]. Así, la anatomía del SC no solo varía para cada paciente, además las áreas extremas del SC se encuentran cerca del anillo mitral [11], que cambia en forma y tamaño dinámicamente durante el ciclo cardíaco [12], afectando a la actividad eléctrica registrada en los extremos del SC. Finalmente, la dilatación ventricular o auricular derecha puede también tener impacto sobre la anatomía del SC [13].

Obviamente, éstas y otras circunstancias fisiológicas pueden causar mal contacto o desplazamiento del catéter de registro durante la AC. Por otro lado, la contracción miocárdica puede afectar aún más los registros del SC. Como el catéter del SC contiene diez electrodos que forman cinco

ISBN: 978-84-09-25491-0, pags. 430- 433

canales de EGMs bipolares, el canal distal, medio-distal, medio, medio-proximal y proximal, su impacto en los antedichos problemas puede ser de mayor o menor importancia, dependiendo de la posición de cada canal. El presente estudio tiene como objeto la evaluación de los canales del SC más fiables y robustos para la evaluación del sustrato auricular en pacientes con FAP durante RS.

2. Materiales

La base de datos se compuso con registros del SC bipolares multicanal de 28 pacientes sometidos por primera vez a AC con radiofrecuéncia (RF) de FAP . Los registros se obtuvieron mediante un sistema LabSystem® Pro (Boston Scientific, Marlborough, MA, USA) con frecuencia de muestreo de 1 kHz y duraciones entre 60 y 300 segundos antes y después de la AC. En algunos casos, ciertos canales contenían casi exclusivamente ruido y fueron excluidos. La base de datos final contenía 44 registros de cinco canales.

3. Métodos

3.1. Preprocesado y análisis de registros

El preprocesado consistió en la cancelación de actividad ventricular [14], seguida de la eliminación de ruido con una técnica basada en descomposición wavelet [15]. Después, se detectaron las ondas de activación local (OALs) [16] y se calcularon las siguientes características para cada OAL: duración, amplitud máxima, mínima y pico a pico (PP), valor cuadrático medio (RMS), área positiva y negativa y número de tramos crecientes y decrecientes (NTCD) de la OAL principal. Además, también se calculó la correlación entre las morfologías dominantes de los canales.

3.2. Duración, amplitud y área de las OALs

Para facilitar la detección de puntos fiduciales, un filtro de media móvil de 5 puntos se aplicó para suavizar la OAL original, que luego fue transformada en una curva normalizada de 0 a 1 mediante una técnica alternativa al preprocesado de Botteron [16]. Los puntos de inicio y fin para calcular la duración de la OAL original quedaron definidos por la primera y última muestra de la señal con amplitud mayor que $0{,}1$ mV.

El pico positivo y negativo más alto correspondió, respectivamente, a la amplitud máxima y mínima de la OAL. La amplitud PP se calculó mediante la distancia entre los dos picos y, finalmente, el valor RMS fue también calculado.

Los puntos de cruce por cero marcaron los límites entre las áreas positivas y negativas, integrando después la señal correspondiente. El área positiva y negativa total de cada OAL se calculó mediante la suma de las áreas positivas y negativas individuales, respectivamente.

3.3. Número de tramos crecientes y decrecientes

Para tener una referencia común, cada OAL se normalizó entre ± 1 y se trazaron dos líneas de base auxiliares a alturas de ± 0.25 respecto de la señal normalizada. Los tramos crecientes quedaron definidos por los puntos que cruzaron las lineas de base auxiliares con pendiente positiva, mientras que los tramos decrecientes se definieron de la misma manera, pero con pendiente negativa. Para evitar recuentos dobles, se contabilizó un solo tramo en el caso de que dos tramos crecientes o dos decrecientes consecutivos fueran marcados por causa de la morfología de la OAL.

Tabla 1. *Resultados del test KW (primera columna) y MWU (columnas restantes) de la comparación entre el grupo con los canales medio y medio-proximal (juntos) y los canales restantes, analizando la base de datos entera. D: distal; Md: medio-distal; M: medio; Mp: medio-proximal; P:proximal.*

	KW	**MWU**		
		D-M\Mp	Md-M\Mp	P-M\Mp
Amp $(+)$	$< 0{,}001$	$< 0{,}001$	$0{,}005$	$0{,}050$
Amp $(-)$	$< 0{,}001$	$0{,}002$	$0{,}059$	$0{,}004$
PP	$< 0{,}001$	$< 0{,}001$	$0{,}018$	$0{,}049$
RMS	$< 0{,}001$	$< 0{,}001$	$0{,}041$	$0{,}068$
Área $(+)$	$< 0{,}001$	$0{,}002$	$0{,}049$	$0{,}095$
Área $(-)$	$0{,}002$	$0{,}003$	$0{,}035$	$0{,}076$

3.4. Correlación entre morfologías dominantes

El objeto de esta etapa es la comparación del grado de similitud de morfologías dominantes entre dos canales del mismo registro. La OAL más representativa de cada canal fue definida por la media de sus 20 señales más correlacionadas. La correlación entre las OALs más representativas de dos canales se calculó usando un índice adaptativo de correlación con signo con $12\,\%$ de tolerancia [17].

3.5. Evaluación del método

El análisis estadístico se aplicó inicialmente a toda la base de datos y, después, se formaron dos grupos con los registros de antes y de después de la AC, respectivamente. Estos grupos constaron de los 16 pacientes cuyos registros eran válidos tanto antes como después de la AC.

Para cada característica, menos la correlación, se exploraron las diferencias estadísticas comparando todos los canales, utilizando el test de Kruskal-Wallis (KW) ya que ninguno de los datos proceden de una distribución normal. A continuación, se buscaron las parejas de canales que presentaban diferencias estadísticamente significativas entre sus características, con la ayuda del test U de Mann-Whitney (MWU). Finalmente, se detectaron aquellos canales con valores estadísticamente diferentes a los restantes mediante un test de 1-vs-todos, aplicado mediante MWU con corrección de Bonferroni. Para todas las características, incluyendo la correlación entre morfologías dominantes, se calcularon los valores medianos, permitiendo así la detección de los canales que presentaron las morfologías más parecidas y características más consistentes.

4. Resultados

4.1. Todos los registros

La Tabla 1 contiene los resultados estadísticos para los tests KW y MWU para las parejas de canales que mostra-

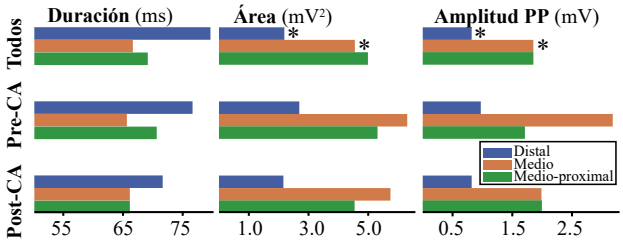


Figura 1. *Valores medianos de duración, área y amplitud PP en los canales distal, medio y medio-proximal para los tres grupos. El asterisco indica valores estadísticamente significativos, con el umbral de significancia modificado a* 0,01*, debido a la corrección de Bonferroni.*

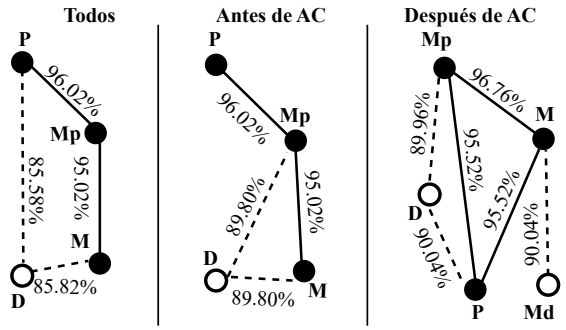


Figura 2. *Ilustración de correlaciones más fuertes (línea continua) y más débiles (linea discontinua) entre OALs. Las correlaciones altas se representan el símbolo ● y las débiles con ○.*

ron diferencias estadísticamente significativas. Como los canales medio y medio-proximal mostraron siempre diferencias estadísticamente significativas con los canales restantes, se describen en la tabla agrupados y se ilustra su comparación con los demás canales. Por tanto, el valor p de los resultados de los tests KW corresponde al más alto de ambos. El NTCD no ofreció valores variables en ninguno de los canales analizados y, por tanto, no se presentan aquí sus resultados.

En las Figuras 1 y 2 se pueden observar los valores medianos de las características analizadas más importantes y las correlaciones más fuertes y más débiles entre los canales, respectivamente. El canal distal mostró la duración más larga, así como la amplitud y área más grandes. Este canal presentó además las correlaciones morfológicas más débiles con los canales restantes ($85{,}82-93{,}03\,\%$). Al contrario, las OALs de los canales medio y medio-proximal mostraron la duración más corta y valores de amplitud y áreas más elevados. El canal medio-proximal presentó adicionalmente correlación muy alta con los canales adyacentes ($95{,}02-96{,}02\,\%$).

4.2. Registros antes de la AC

El análisis estadístico mostró en este caso resultados similares a los del análisis de todos los registros. Sin embargo, las relaciones estadísticas fueron más débiles y solo llegaron a tener importancia en la amplitud de los canales medio y distal. El test de KW solo encontró significancia estadística en la amplitud máxima en varios canales. La Tabla 2 muestra un sumario de los resultados más relevantes.

Los valores medianos de duración, amplitud PP y área positiva se muestran en la Figura 1, donde siguen el mismo comportamiento que para todos los registros. Por la corrección de Bonferroni, estos valores no fueron estadísticamente significativos para los registros antes de AC, pero se observó una tendencia común en los canales distal y medio de todas las características. La fila central de la Figura 2 demuestra que la correlación sigue también el mismo comportamiento que en el caso de todos los registros.

Tabla 2. *Test de KW y MWU para duración, amplitud PP y área positiva de los grupos antes y después de la AC. D: distal; Md: medio-distal; M: medio; Mp: medio-proximal; P:proximal.*

	Caract.	KW	MWU	
			D-M\Mp	Md-M\Mp
Pre-AC	Duración	0,651	0,418	1,000
	Amp. PP	0,159	0,078	0,194
	Amp. Max.	0,015	0,017	0,069
	Área Pos.	0,163	0,069	0,214
Post-AC	Duración	0,865	0,326	0,817
	Amp. PP	0,032	0,023	0,046
	Amp. Max.	0,042	0,023	0,069
	Área Pos.	0,056	0,035	0,141

4.3. Registros después de la AC

Los resultados en este grupo fueron también similares a los del caso de todos los registros. Aunque las relaciones estadísticas no fueron tan fuertes como en dicho caso, sí que superaron el caso de los registros antes de la AC. Estas relaciones aparecen en la Tabla 2.

En las Figuras 1 y 2 se muestran los valores medianos y la correlación entre canales, respectivamente. Como puede observarse en ambos casos, los resultados están en concordancia con los otros grupos. En este caso también, no se obtuvieron valores medianos estadísticamente significativos, pero se observó una tendencia entre los canales medio y distal de todas las características.

5. Discusión

En el estudio presente se han evaluado las variaciones en las características de OALs del SC entre los diferentes canales del catéter. Estas diferencias se encuentran principalmente en las características de amplitud, área y duración entre dos partes del catéter: la parte central (canales medio y medio-proximal) y la parte extrema (canales distal y medio-distal por un lado y el canal proximal). El área extrema presenta OALs de duración más larga y amplitud más baja, mientras las OALs del área central son de duración más corta y de amplitud más elevada. No obstante, las OALs del canal proximal estuvieron más correlacionadas con las del área central y menos con las de la parte distal.

Comparando los grupos antes y después de la AC, se observó mayor capacidad discriminatoria en el grupo post-AC para el área y la amplitud de OALs. No obstante, teniendo en cuenta que los registros del canal distal contenían se-

ñales de amplitud más baja, es posible que la presencia de ruido en estos canales impida observar más claramente el efecto de la AC.

Por otro lado, aunque no fuera estadísticamente significativa, la duración en los canales distal y medio-proximal se ha reducido notablemente, aunque esto no se cumple para OALs del canal medio, como se puede ver en la Figura 1. Esto se puede explicar por la gran diferencia en la duración entre los canales distal y medio\medio-proximal, que permite considerar la posibilidad de que antes de la ablación los canales distal y medio-proximal contuvieran cierto porcentaje de ruido, que después de la AC ha sido eliminado, sobre todo en el canal medio-proximal. Se tiene que destacar que la duración de OALs fue la misma entre los canales medio y medio-proximal, después de AC. En todo caso, serían necesario estudios con más pacientes y más exhaustivos sobre la variación de las características de las OALs después de AC para poder profundizar en este asunto.

Los resultados presentados en este estudio indican que los canales medio y medio-proximal presentan mayor calidad de señal y ofrecen una evaluación más fiel y robusta de las dinámicas de la FA en RS durante la AC. Estos resultados están de acuerdo con trabajos previos, que también han observado capacidad discriminatoria de características obtenidas de los canales medio y medio-proximal, mientras que se ha demostrado la insuficiencia del canal distal para el mismo fin [4, 18, 19].

6. Conclusiones

Este trabajo ha demostrado que las características de las OALs del SC varían según el canal de registro. Por lo tanto, su selección adecuada es un paso fundamental. El canal distal ha resultado ser el menos fiable, con variaciones significativas que podrían alterar la correcta estimación de las dinámicas de las OALs y, por ello, se desaconseja su uso. Por otro lado, los canales medio y medio-proximal preservan al máximo la información sobre las dinámicas de las OALs y se aconseja su empleo para una evaluación del sustrato auricular más precisa y fiable en pacientes con FAP.

Agradecimientos

Trabajo cofinanciado por los proyectos de investigación competitiva DPI2017-83952-C3 de MINECO-AEI-FEDER-UE, SBPLY/17/180501/000411 de la JCCLM y AICO/2019/036 de la GVA.

Referencias

[1] Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, Castella M, Diener HCea. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. European Heart Journal 08 2016; 37(38):2893–2962. ISSN 0195-668X.

[2] Ghzally Y, Gerasimon G. Catheter Ablation. StatPearls Publishing, 2020 Jan.

[3] Yoshida K, Ulfarsson M, Tada H, Chugh A, Good E, Kuhne M, Crawford T, Sarrazin JF, Chalfoun Nea. Complex electrograms within the coronary sinus: Time- and frequency-domain characteristics, effects of antral pulmonary vein isolation, and relationship to clinical outcome in patients with paroxysmal and persistent atrial fibrillation. Journal of Cardiovascular Electrophysiology 2008;19(10):1017–023.

[4] Boles U, Gul EE, Enriquez A, Starr N, Haseeb S, Abdollah H, Simpson C, Baranchuk A, Redfearn Dea. Coronary sinus electrograms may predict new-onset atrial fibrillation after typical atrial flutter radiofrequency ablation (cse-af). Journal of atrial fibrillation Jun 2018;11(1):1809–1809.

[5] Antz M, Otomo K, Arruda M, Scherlag BJ, Pitha J, Tondo C, Lazzara R, Jackman WM. Electrical conduction between the right atrium and the left atrium via the musculature of the coronary sinus. Circulation 1998;98(17):1790–1795.

[6] Tritto M, Zardini M, De Ponti R, Salerno-Uriarte JA. Iterative atrial tachycardia originating from the coronary sinus musculature. Journal of Cardiovascular Electrophysiology 2001;12(10):1187–1189.

[7] Ahmed N, Perveen S, Mehmood A, Rani GF, Molon G. Coronary sinus ablation is a key player substrate in recurrence of persistent atrial fibrillation. Cardiology 2019;143(3-4):107–113.

[8] Della Rocca DG, Gianni C, Gedikli O, Chen Q, Natale A, Al-Ahmad A. Provocation and ablation of non–pulmonary vein triggers in nonparoxysmal atrial fibrillation: Role of the coronary sinus. HeartRhythm Case Reports 2020; 6(5):231–236.

[9] Stępień K, Kuklik P, Żebrowski JJ, Sanders P, Derejko P, Podziemski P. Kolmogorov complexity of coronary sinus atrial electrograms before ablation predicts termination of atrial fibrillation after pulmonary vein isolation. Entropy 2019;21(10):970.

[10] Gianni C, Mohanty S, Trivedi C, Di Biase L, Natale A. Novel concepts and approaches in ablation of atrial fibrillation: the role of non-pulmonary vein triggers. Ep Europace 2018; 20(10):1566–1576.

[11] El-Maasarany S, Ferrett CG, Firth A, Sheppard M, Henein MY. The coronary sinus conduit function: anatomical study (relationship to adjacent structures). EP Europace 2005; 7(5):475–481.

[12] Pai RG, Varadarajan P, Tanimoto M. Effect of atrial fibrillation on the dynamics of mitral annular area. The Journal of heart valve disease 2003;12(1):31–37.

[13] Mahmud E, Kerumati S, Raisinghani A. Dilation of the coronary sinus on echocardiogram: prevalence and significance. J Am Soc Echocardiography January 2001;1(14):44–49.

[14] Alcaraz R, Rieta JJ. Adaptive singular value cancelation of ventricular activity in single-lead atrial fibrillation electrocardiograms. Physiol Meas Dec 2008;29(12):1351–69.

[15] Martínez-Iniesta M, Ródenas J, Rieta JJ, Alcaraz R. The stationary wavelet transform as an efficient reductor of powerline interference for atrial bipolar electrograms in cardiac electrophysiology. Physiol Meas 07 2019; 40(7):075003.

[16] Osorio D, Alcaraz R, Rieta JJ. A fractionation-based local activation wave detector for atrial electrograms of atrial fibrillation. In 2017 Computing in Cardiology (CinC). IEEE, 2017; 1–4.

[17] Alcaraz R, Hornero F, Martínez A, Rieta JJ. Short-time regularity assessment of fibrillatory waves from the surface ecg in atrial fibrillation. Physiological measurement 2012; 33(6):969.

[18] Teh AW, Kalman JM, Kistler PM, Lee G, Sutherland F, Morton JB, et al. Prevalence of fractionated electrograms in the coronary sinus: comparison between patients with persistent or paroxysmal atrial fibrillation and a control population. Heart Rhythm 2010;7(9):1200–1204.

[19] Di Marco LY, Raine D, Bourke JP, Langley P. Characteristics of atrial fibrillation cycle length predict restoration of sinus rhythm by catheter ablation. Heart Rhythm 2013; 10(9):1303–1310.

Monitorización en diálisis de la Concentración de Potasio en Sangre mediante los cambios en la morfología multi-lead de la onda T: Comparación Entre usar la transformación en Componentes Periódicas y Principales

F. Palmieri[1,2,4], P. Gomis[1,2], J.E. Ruiz[3], D. Ferreira[4], A. Martín-Yebra[5], E. Pueyo[2,5], P. Laguna[2,5], J.P. Martínez[2,5], J. Ramírez[6],

[1] Centre de Recerca en Enginyeria Biomèdica, Universitat Politècnica de Catalunya, Barcelona, España.

[2] CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Zaragoza, España.

[3] Servicio de Nefrología, Hospital Clínico Universitario Lozano Blesa, Zaragoza, España.

[4] Laboratorios Rubio, Castellbisbal, Barcelona, España.

[5] BSICoS Group, I3A, IIS Aragón, Universidad de Zaragoza, Zaragoza, España.

[6] William Harvey Research Institute, Queen Mary University of London, London, United Kingdom

Resumen

Antecedentes Los pacientes con enfermedad renal en etapa terminal (ESRD) sometidos a terapia de hemodiálisis (HD) experimentan variaciones de potasio en sangre ($[K^+]$) que se reflejan en la morfología de la onda T (TW) del ECG. Métodos: en este trabajo, evaluamos el comportamiento de diferentes métodos de reducción del número de derivaciones (LSR): análisis de componentes principales (PCA), maximizando la energía de las TW, y dos versiones derivadas del análisis de las componentes periódicas (πCA) llamadas πC^B y πC^T, maximizando la periodicidad entre latidos del QRST o la TW, respectivamente. Aplicamos estos métodos en 24 pacientes con ESRD-HD, derivando tres marcadores de cambios de morfología TW (d_w^u, d_w y $\hat{d}_{w,c}$), que comparan la morfología de una TW promedia calculada cada 30 minutos con la del final de la HD, en aquellas PCA, πC^B y πC^T con el mayor contenido energético de las TW. Las similitudes entre estos tres métodos se evaluaron mediante el uso de gráficos de Bland-Altman y el error de ajuste lineal (ϵ) evaluado desde la hora 12 hasta la hora 44 después del inicio de la HD. Resultados: Todos los marcadores (d_w^u, d_w y $\hat{d}_{w,c}$) mostraron bajos valores medios de deferencia entre ellos (sesgo $\leq$ 0.5ms) y pequeños valores de error en el ajuste lineal (mediana $\epsilon \leq$ 3.3ms). Conclusiones: PCA y πCA pueden usarse indistintamente para monitorizar los cambios TW en pacientes con ESRD-HD.

1. Introducción

Las enfermedades cardiovasculares son las principales causas de morbilidad y mortalidad en pacientes con enfermedad renal en etapa terminal (ESRD) sometidos a hemodiálisis (HD). Además, las variaciones en la concentración de potasio en la sangre ($\Delta[\boldsymbol{K^+}]$), entre otros efectos, pueden alterar la secuencia de repolarización ventricular espacio-temporal, reflejada como cambios en la morfología de la onda T en el electrocardiograma (ECG) [1]. En un estudio previo [2], hemos investigado la capacidad de los marcadores de cambios en morfología de la onda T resultantes del análisis time-warping [3] para monitorear $\Delta[\boldsymbol{K^+}]$, en pacientes con ESRD-HD. Tres de estos marcadores mostraron la mayor correlación con $\Delta[\boldsymbol{K^+}]$: $\boldsymbol{d_w^u}$ (implementado como en [3]), $\boldsymbol{d_w}$ (una versión con signo de $\boldsymbol{d_w^u}$) y $\widehat{\boldsymbol{d}}_{\boldsymbol{w,c}}$ una versión corregida por la frecuencia cardíaca de $\boldsymbol{d_w^u}$. En [2,3] el análisis de componentes principales (PCA) se aplicó para concentrar la energía de las ondas T en una única derivación, la primera componente principal (PC1), asumiendo que es la que tenía la máxima energía en las ondas T. En [4], el análisis de componentes periódicas (πCA), una transformación cuyo objetivo es enfatizar la periodicidad en señales multicanal, se utilizó como una alternativa a PCA para enfatizar la alternancia de las ondas T embebidas en una alta contaminación de ruido. En ese estudio se encontró un rendimiento superior de πCA en comparación con PCA para la detección de la alternancia de las ondas T cuando la relación señal a ruido es grande. El objetivo de este trabajo es comparar el comportamiento de PCA y πCA como reducción del número de derivaciones (LSR) previo al análisis time-warping para monitorizar $\Delta[\boldsymbol{K^+}]$ en HD. En concreto, en este trabajo se modificó la versión πCA propuesta en [4] para el análisis de alternancia de las ondas T (periodicidad cada dos latidos) para estudiar la morfología de la onda T (periodicidad cada latido). Luego, el análisis time-warping se realizó sobre los derivaciones que se obtuvieron por PCA y πCA, extrayendo y comparando los valores de $\boldsymbol{d_w^u}$, $\boldsymbol{d_w}$ y $\widehat{\boldsymbol{d}}_{\boldsymbol{w,c}}$.

2. Población y protocolo de estudio

En este estudio se han evaluado 24 pacientes en ESRD sometidos a HD en el Hospital Clínico Universitario Lozano Blesa de Zaragoza, España. Para cada uno de ellos se recogieron muestras de sangre, de donde se obtuvieron las concentraciones de $[\boldsymbol{K^+}]$, y un registro Holter ECG de 12 derivaciones durante 48 horas (H12+, Mortara Instruments, Milwaukee, WI, EE. UU.) a una frecuencia de muestreo de 1000 Hz y resolución de amplitud 3.75 μV. El protocolo de estudio fue aprobado por el comité de ética (*CEICA*, ref. PI18 / 003) y todos los pacientes firmaron un consentimiento informado.

ISBN: 978-84-09-25491-0, pags. 434- 437

3. Métodos

Pre-procesado

El pre procesamiento del ECG incluyó filtrado paso bajo Butterworth de fase 0 y orden 6 (frecuencia de corte 40 Hz) para eliminar el ruido muscular sin modificar la morfología de las ondas T, y de la línea de alimentación y filtrado paso alto Butterworth de fase 0 y orden 6 (frecuencia de corte 0,5 Hz) para atenuar el desplazamiento de la línea base. Los complejos QRS se detectaron y delinearon en cada derivación utilizando [5] (más informaciones en [2,3]).

Reducción del número de derivaciones

Se calculó una matriz de transformación Ψ para cada método de reducción del número de derivaciones (LSR) y luego se aplicó a las 8 derivaciones independientes para obtener sus derivaciones transformadas correspondientes.

Análisis de las componentes principales: Los coeficientes que definen Ψ_{PCA} se obtuvieron de los vectores propios de la matriz 8×8 de auto-correlación de las derivaciones calculada utilizando las muestras en las ondas T definidas a partir de T-on hasta el T-off (desde línea discontinua naranja a roja en la Fig. 1 (a)) [3] en una ventana de 10 minutos al final de la sesión de HD. Esta ventana de 10 minutos solo se utilizó para aprender los vectores principales de la PCA. Como resultado, PC1 fue la derivación transformada que maximizó la energía de la onda T, Fig. 1(b).

Análisis de las componentes periódicas: Los complejos QRST, definidos como una ventana de 450 ms desde el inicio de QRS (línea azul en la Fig. 1(a)), se tomaron de la ventana de 10 minutos mencionada anteriormente para el aprendizaje de la transformación, lo que resultó en πC^Bs (Fig. 1 (c)). Alternativamente, teniendo en cuenta que el complejo QRS y la onda T, aunque ambos son periódicos de un solo latido, no necesariamente tienen la misma distribución espacial, solo se consideró la zona de la onda T para calcular los vectores propios usando una ventana desde el inicio del QRS más 130 ms a 450 ms (línea discontinua verde a roja en la Fig. 1 (a)), lo que resulta en πC^Ts (Fig. 1 (d)).

Sea $m = 1$ la periodicidad de un latido, K el número de latidos en la ventana de análisis, N el número de muestras en el intervalo de latido seleccionado, L el número de derivaciones y $x_{k,l}\,(n)$ el n-ésima muestra en el latido k-ésimo para la l-ésima derivación. En notación vectorial, cada complejo QRST, u onda T, se definió como $\boldsymbol{x}_{k,l} = [x_{k.l}\,(0), \dots, x_{k.l}\,(N-1)\,]^T$. Para evitar variabilidades adicionales debidos a la delineación, se realizó una prealineación de los intervalos seleccionados en cada derivación calculando un complejo QRST de referencia (u onda T) como la media de todos los complejos QRST (u ondas T) en esa ventana. Luego, se calculó la correlación cruzada entre la referencia y cada complejo QRST (u onda T) y éstos se desplazaron de acuerdo con el retraso de la correlación cruzada máxima. Para cada latido, todas las derivaciones se juntaron en la matriz $L \times N$:

$$\boldsymbol{X}_k = \left[\boldsymbol{x}_{k,1}, \dots, \boldsymbol{x}_{k,L}\right]^T \qquad (1)$$

Se construyeron dos matrices $\mathbb{X}$ y $\mathbb{X}^{(m)}$ concatenando $\boldsymbol{X}_k$ de 1 a K y de $1+m$ a $K+m$, respectivamente:

$$\begin{aligned} \mathbb{X} &= [\boldsymbol{X}_1, \dots, \boldsymbol{X}_K\,]^T; \\ \mathbb{X}^{(m)} &= [\boldsymbol{X}_{1+m}, \dots, \boldsymbol{X}_{K+m}\,]^T \end{aligned} \qquad (2)$$

Luego, las matrices de correlación espacial $R_{\mathbb{X}}$ por $\mathbb{X}$ y $A_{\mathbb{X}^{(m)}}$ por $\left(\mathbb{X}^{(m)} - \mathbb{X}\right)$ se calcularon como:

$$R_{\mathbb{X}} = \frac{1}{KN}\,\mathbb{X}\,\mathbb{X}^T \qquad (3)$$

$$A_{\mathbb{X}^{(m)}} = \frac{1}{KN}\left(\mathbb{X}^{(m)} - \mathbb{X}\right)\left(\mathbb{X}^{(m)} - \mathbb{X}\right)^T \qquad (4)$$

Finalmente, $\Psi_{\pi CA}$ se eligió como la matriz de vectores propios generalizada del par $\left(A_{\mathbb{X}^{(m)}}, R_{\mathbb{X}}\right)$[4]; con los vectores propios ordenados según los valores propios correspondientes en orden de magnitud ascendente; siendo $\pi C1$ la dirección espacial en la que se ve mejor la componente más periódica de señal.

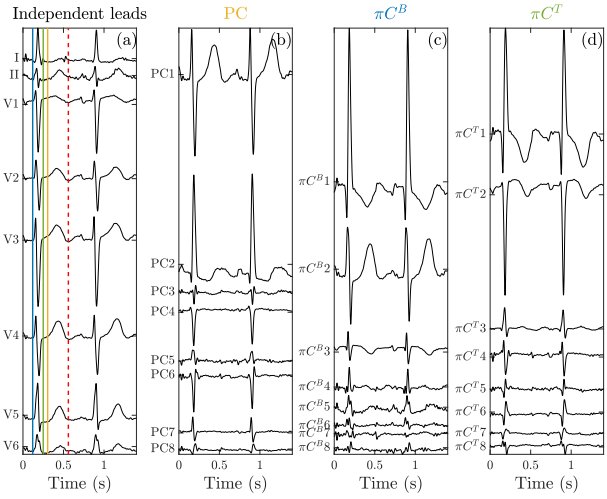


Figura 1. *El panel (a) muestra las 8 derivaciones independientes de uno de los pacientes. Cada línea continua vertical representa el punto de partida de la ventana que contiene las muestras utilizadas para construir los vectores propios en PCA (naranja, denotando Ton), πC^B (azul, denotando QRSon) y πC^T (verde, que indica QRSon + 130ms). La línea discontinua roja vertical marca el final de la ventana (final de la onda T). Las PC resultantes, πC^B y πC^T se representan en el panel (b) (c) y (d) respectivamente. En este ejemplo, PC1, $\pi C^B 2$, $\pi C^T 1$ son las elegidas para el análisis, siendo las que maximizan la energía de las ondas T.*

Cuantificación de los cambios en la morfología de las ondas T a través del time-warping

Para cada paciente y transformación, todas las ondas T en la primera y segunda derivaciones transformadas se delinearon [5]. En el caso de PCA y $\boldsymbol{\pi C^T}$, se eligió el primer componente para el análisis, ya que, por construcción, siempre tiene la energía máxima para la onda T. Para $\boldsymbol{\pi C^B}$, por el contrario, no hay garantía de que la onda T en la primera derivación transformada tenga la energía máxima. Por lo tanto, para seleccionar la derivación de $\boldsymbol{\pi C^B}$ se calculó la energía de la onda T en las primeras dos componentes de $\boldsymbol{\pi C^B}$en la ventana de aprendizaje de 10 minutos al final del HD, y se eligió aquella con el contenido energético de ondas T más alto.

Luego, las ondas T en las derivaciones transformadas elegidas se seleccionaron en una ventana de 2 minutos centrada en el minuto 5º y 35º de cada hora disponible y se usaron para calcular una onda T deformada media (MWTW) como en [3]. Esa ventana es lo suficientemente estrecha como para permitir la asunción de estabilidad en los niveles de frecuencia cardíaca (FC) y $[\boldsymbol{K^+}]$. Cada ventana de dos minutos ha sido inspeccionada visualmente para confirmar que no había ruido residual y de todas las MWTW calculadas se tomó, como referencia, la MWTW al final del tratamiento de HD y se comparó la morfología de cada MWTW con respecto a esta referencia, a través del time-warping análisis que se describe en [2,3]. El marcador $\boldsymbol{d_w^u}$, obtenido de este procedimiento, cuantifica el nivel de deformación necesario para alinear de manera óptima dos ondas T como el promedio del valor de diferencia absoluta entre $\boldsymbol{\gamma^*(t^r)}$ y $\boldsymbol{t^r}$:

$$d_w^u = \frac{1}{N_r}\sum_{n=1}^{N_r}\left|\gamma^*\left(\boldsymbol{t}^r(n)\right) - \boldsymbol{t}^r(n)\right| \qquad (5)$$

Siendo $\gamma^*(\boldsymbol{t}^r)$ la función óptima de warping que minimiza la diferencia de amplitud entre la *square-root slope function* (SRSF) [6] de la MWTW de referencia $\boldsymbol{f}^r(\boldsymbol{t}^r) = \left[f^r\left(t^r(1)\right),\cdots,f^r\left(t^r(N_r)\right)\right]^T$ y cada MWTW $\boldsymbol{f}^s(\boldsymbol{t}^s) = \left[f^s\left(t^s(1)\right),\cdots,f^s\left(t^s(N_s)\right)\right]^T$, $\boldsymbol{t}^r = [t^r(1),\cdots,t^r(N_r)]^T$ y $\boldsymbol{t}^s = [t^s(1),\cdots,t^s(N_s)]^T$ y N_r y N_s son la duración total de $\boldsymbol{t}^r$ y $\boldsymbol{t}^s$. La definición original de d_w^u [3] se modificó [2] para permitir que llevase signo, distinguiendo así la ampliación de la onda T del estrechamiento. Esta nueva versión, $\mathrm{d_w}$, y su signo, s_d, se definen como:

$$d_w = \left(\frac{s_d}{s_d}\right)\frac{1}{N_r}\sum_{n=1}^{N_r}\left|\gamma^*\left(\boldsymbol{t}^r(n)\right) - \boldsymbol{t}^r(n)\right| \qquad (6)$$

$$s_d = \sum_{n\in N_r^u}\left(\gamma^*\left(\boldsymbol{t}^r(n)\right) - \boldsymbol{t}^r(n)\right) + \sum_{n\notin N_r^u}\left(\boldsymbol{t}^r(n) - \gamma^*\left(\boldsymbol{t}^r(n)\right)\right) \qquad (7)$$

con N_r^u siendo el conjunto de muestras de pendiente ascendente de onda T. Un signo positivo significa que $\boldsymbol{f}^s(\boldsymbol{t}^s)$ debe ampliarse para que se ajuste la referencia $\boldsymbol{f}^r(\boldsymbol{t}^r)$ y viceversa para un signo negativo. Finalmente, para compensar las variaciones concomitantes de la onda T debido a la FC, se modeló $\boldsymbol{d_w}$ como la suma de dos componentes $\boldsymbol{d_w} = d_{w,c} + d_{w,HR}$ donde $d_{w,HR}$ representa los cambios en la morfología de la onda T relacionados con la FC y $\boldsymbol{d_{w,c}}$ representa otros cambios, incluidos los relacionados con $\Delta[K^+]$**.** Para estimar $d_{w,c}$ utilizamos la siguiente fórmula de corrección, basada en una aproximación lineal utilizada para explicar la dependencia de la FC con el intervalo QT [7]

$$\hat{d}_{w,c} = d_w - \hat{c}\,(RR_s - RR_r) \qquad (8)$$

Siendo $\widehat{\boldsymbol{d}}_{w,c}$ el valor estimado de $\boldsymbol{d_{w,c}}$, RR_s y RR_r el promedio de los intervalos RR de los segmentos estudiado y de referencian respectivamente, y $\hat{c}$ estimado para cada paciente.

Evaluación de la dinámica de los marcadores

Para evaluar la concordancia entre cada marcador en cada LSR, utilizamos gráficos de Bland-Altman (B-A) evaluados para cada paciente y cada par de transformaciones, extrayendo el sesgo y los límites de confianza (CL). Además, utilizamos el error de ajuste (ϵ) para cuantificar la dispersión de cada marcador específico de LSR con respecto a una línea de regresión entre las horas 12 y 44 después del inicio de la HD. Esto proporciona información sobre la desviación del marcador de una tendencia lineal gradual a lo largo del tiempo, que podría estar relacionada con la tendencia seguida por $[\boldsymbol{K^+}]$**.** Los resultados para sesgo, CL y ϵ en todo el conjunto de datos se dieron como mediana y rango intercuartil (IQR).

4. Resultados y Discusión

La duración promedio de los registros ECG solo llego hasta las 44 horas, debido al desprendimiento de electrodos o al agotamiento temprano de la batería. $\boldsymbol{\pi C^B}1$ se eligió en 17 pacientes (71%) como la que tiene la energía de onda T máxima. Esto sugeriría que, para el 29 % de los pacientes, la proyección espacial que maximiza la periodicidad QRST no fue la misma que maximiza la periodicidad de la onda T (Fig.1).

La evolución de los marcadores $\boldsymbol{d_w^u}, \boldsymbol{d_w}$ y $\widehat{\boldsymbol{d}}_{w,c}$ basados en PCA y πCA a lo largo de la grabación del ECG se muestra en la Fig. 2 paneles (a)- (c), respectivamente. En cada panel se presenta una vista detallada de las primeras 10 horas (solo valores medianos). En este intervalo específico, todos los marcadores muestran una curva en forma de V que tiene el vértice al final de la HD (Fig. 2 (d)), de manera similar a las curvas de recuperación de $[\boldsymbol{K^+}]$ reportadas [8, 9]. Esto destaca una analogía interesante que necesita ser investigada más a fondo.

Mirando los resultados en la Tabla 1, el sesgo y los CL fueron, en promedio, consistentemente más pequeños al comparar la PC con $\boldsymbol{\pi C^T}$ (-0.2 ≤ B-A bias ≤ 0.0 y 7.5 ≤ CL ≤ 9.3), y en ambos casos menor ϵ (siendo $1.8 \le \epsilon \le 2.1$ para PC y $2.0 \le \epsilon \le 2.4$ por πC^T) que $\boldsymbol{\pi C^B}$. Estos resultados indican una evolución similar de $\boldsymbol{\pi C^T}$ y PCA. Sin embargo, dado que PCA separa los componentes de señal de acuerdo con su varianza, proyectar el componente con la energía más alta en la primera componente transformada podría ser una limitación en caso de ECG altamente ruidosa [4].

A partir de la décima hora, los cambios bruscos de los valores medianos de los marcadores se reducen gradualmente, asentándose como una tendencia que podría calificarse de lineal. Suponiendo eso, una pequeña mediana de ϵ para una técnica de LSR dada sugeriría que dio como resultado mediciones más coherentes en los pacientes para un marcador específico. Sin embargo, incluso si los valores de la Tabla 1 pueden indicar PCA como la técnica que conduce a los marcadores menos dispersos, la mediana y el IQR son muy similares entre sí para concluir definitivamente cuál entre PCA, $\boldsymbol{\pi C^T}$ y $\boldsymbol{\pi C^B}$ fue la más consistente en términos de mediciones.

5. Conclusiones

En este trabajo, probamos el comportamiento de tres métodos de LSR, $\boldsymbol{\pi C^T}$ y $\boldsymbol{\pi C^B}$y PCA, comparando tres

marcadores basados del análisis time-warping onda T calculados a partir de registros de ECG de larga duración de 24 pacientes con ESRD-HD. Nuestros resultados apoyarían el uso intercambiable de PCA y πCA como trasformación previa al "warping" para seguir las variaciones de la morfología de la onda T durante y después de la HD. Los estudios futuros en condiciones de ruido más pesado aclararán si $\boldsymbol{\pi C^T}$ o $\boldsymbol{\pi C^B}$ se prefiere sobre PCA, como previamente mostrado en la detección de alternancia de onda T [4].

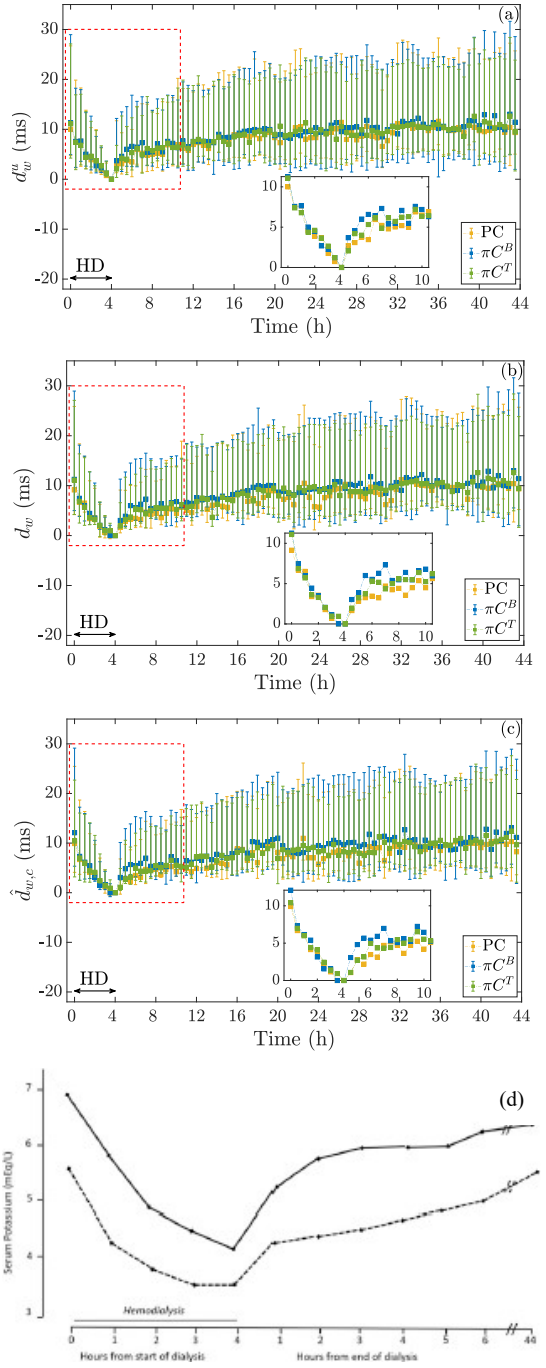


Figura 2. *Evolución temporal de cada marcador d_w^u, d_w y $\hat{d}_{w,c}$ (a) - (c), expresada como mediana e IQR. Se muestra una vista detallada de las primeras 10 horas (rectángulo discontinuo rojo) dentro de cada panel para permitir su comparación con la tendencia de $[\boldsymbol{K^+}]$ disponible en la literatura (panel (d)) en Pun et al. [9] Fig.1, p.3442.*

	B-A bias/CL (ms/ms)			ϵ/IQR (ms/ms)		
	PC-πC^B	PC-πC^T	πC^B-πC^T	PC1	$\pi C^B s$	$\pi C^T 1$
d_w^u	0.0/ 11.2	0.0/ 7.5	0.3/ 9.2	1.8/ 1.8	2.4/ 2.0	2.0/ 1.3
d_w	0.1/ 13.3	-0.1/ 9.5	-0.2/ 14.5	2.2/ 2.6	3.1/ 3.8	2.4/ 2.1
$\hat{d}_{w,c}$	0.5/ 12.8	-0.2/ 9.3	-0.3/ 14.2	2.1/ 1.7	3.3/ 3.4	2.2/ 2.0

Tabla 1. *Valores medianos de bias y CL de gráficos B-A y mediana/ IQR evaluados en todos los pacientes.*

Agradecimientos

Este trabajo fue financiado por Products & Technology S.L. (Castellbisbal) y por la Generalitat de Catalunya (España) DI001-2018. El trabajo también fue apoyado por el Gobierno de Aragón (proyecto LMP124-18 y el Grupo de referencia Interpretación de señales biomédicas y simulación computacional (BSICoS) T39-20R) cofinanciado por FEDER 2014-2020 "Construyendo Europa desde Aragón ", El ministerio de Ciencia: PID2019-105674RB-I00 y PID2019-104881RB-I00, y por el Consejo Europeo de Investigación (ERC) a través del proyecto ERC-2014-StG 638284. J. R. desea agradecer el apoyo del acuerdo de subvención Marie Sklodowska-Curie no 786833.

Referencias

[1] Secemsky E, *et al*. High prevalence of cardiac autonomic dysfunction and T-wave alternans in dialysis patients. *Heart Rhythm*, vol 8, sup 4, 2011, pp 592–98.

[2] Palmieri F, *et al*. Monitoring blood potassium concentration in hemodialysis patients by quantifying T-wave morphology dynamics. *Scientific Reports* 2020 bajo revisión del editor.

[3] Ramírez J, *et al*. Variability of ventricular repolarization dispersion quantified by time-warping the morphology of the T-wave. IEEE TBME, vol 64, sup 7, 2017, pp 1619–30.

[4] Monasterio V, *et al*. A multilead scheme based on periodic component analysis for T-wave alternans analysis in the ECG. *Ann of Biom Eng*, vol 38, sup 8, 2010, pp.2532-41.

[5] Martínez JP, *et al.* A wavelet-based ECG delineator: Evaluation on standard databases. *IEEE TBME*, vol 4, sup51, 2004, pp 570–81.

[6] Tucker JD, *et al*. Generative models for functional data using phase and amplitude separation, vol 61, 2013, pp 50–60.

[7] Pueyo E, *et al*. Characterization 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 TBME*, vol 51, sup 9, 2004, pp 1511–20.

[8] Blumberg A, *et al*. Plasma potassium in patients with terminal renal failure during and after haemodialysis; relationship with dialytic potassium removal and total body potassium. Nephrol Dial Transplant 1997;12(8):1629–34.

[9] Pun PH, *et al.* Dialysate potassium, dialysate magnesium, and hemodialysis risk. J Am Soc Nephrol, vol 28, sup 12, 2017, pp 3441-51.

Contacto: Flavio Palmieri, UPC CREB, Av. Diagonal, 647, 08028, Barcelona (España). flavio.palmieri@upc.edu

Imágenes Biomédicas II

Viernes, 27 de noviembre de 2020

Nuevo método para la obtención de imágenes TAC libres de endurecimiento de haz vía aprendizaje automático

C. Martínez [1,2], C. F. Del Cerro [1,2], M. Desco [1,2,3,4], M. Abella [1,2,3]

[1] Depto. Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, España {crismart, carlosfe}@pa.uc3m.es, {mdesco, mabella}@ing.uc3m.es

[2] Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España,

[3] Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, España

[4] Centro de investigación en red en salud mental (CIBERSAM), Madrid, España

Resumen

El endurecimiento de haz es un efecto mediante el cual la energía media de un haz de rayos X se incrementa a medida que atraviesa un material. Este efecto provoca dos tipos de artefactos en la imagen reconstruida: cupping, en muestras homogéneas y bandas negras, entre zonas densas de muestras heterogéneas.

Los métodos de corrección propuestos en la literatura se pueden dividir en métodos de post-proceso, los cuales necesitan de varias reconstrucciones y proyecciones extra basadas en una segmentación de una reconstrucción preliminar. Esta segmentación puede no ser óptima en adquisiciones con baja dosis, limitando la calidad de la imagen final. Se han propuesto métodos de reconstrucción iterativos como alternativa para corregir los artefactos de endurecimiento de haz en adquisiciones de baja dosis, incluyendo modelos de ruido y de fuente policrómatica, pero suponen un aumento considerable del tiempo de ejecución.

En este trabajo se propone un nuevo método para la obtención de imágenes sin artefactos de endurecimiento de haz en tomografía axial computarizada basado en técnicas de aprendizaje profundo, usando la red neuronal U-Net. El método se ha evaluado en dos escenarios, de alta y baja dosis, resultando en una corrección óptima en ambos.

1. Introducción

El origen del endurecimiento de haz en tomografía computarizada reside en la naturaleza policromática de la fuente de rayos X. Se define como el proceso por el cual la energía media del haz de rayos X aumenta al atravesar un material debido a que los fotones menos energéticos son absorbidos con mayor facilidad que los de mayor energía. El endurecimiento de haz produce principalmente dos tipos de artefactos en la imagen reconstruida: *cupping* en regiones homogéneas y bandas entre zonas densas de regiones heterogéneas [1].

En la literatura se pueden encontrar múltiples esquemas de corrección. El más sencillo, consistente en pre-endurecer el rayo a través de un filtrado físico para descartar la mayoría de fotones de baja energía que emite la fuente, es insuficiente para eliminar completamente los artefactos [1]. El método más implementado en los escáneres comerciales es la linealización, en el que se asume que la muestra es homogénea, corrigiendo el artefacto de *cupping* pero no eliminando las bandas negras [2]. Para corregir ambos artefactos, Nalcioglu et al. [3] extienden la linealización incluyendo dos factores de corrección, uno para el hueso y otro para el tejido blando. Para ello necesitan la estimación de la cantidad de tejido blando y hueso atravesado, obtenida mediante una reconstrucción preliminar, y el conocimiento del espectro de la fuente. Joseph et al. [4] eliminan la necesidad de conocer el espectro, corrigiendo los datos de proyección con un modelo no lineal de dos parámetros, en el que también es necesaria una reconstrucción preliminar para estimar la cantidad de hueso atravesado. La principal dificultad de este método reside en el hecho de que el valor óptimo de los parámetros varía según el tamaño y composición del hueso, dificultando la obtención de un valor óptimo global. En [5] se extiende el modelo anterior buscando los parámetros de forma iterativa por medio de la maximización de la uniformidad en el tejido blando de la imagen. Para eliminar el paso de segmentación, en [6] se desarrolla el mismo algoritmo basándose en la combinación lineal de las imágenes pero extraídas a partir la proyección y reconstrucción de la imagen inicial, a la cual se le ha realizado una transformación del histograma. Ambos métodos están basados en la suposición de que el objeto está compuesto por zonas homogéneas, lo que puede resultar en una reducción de resolución de contraste en tejido blando.

Recientemente se han presentado dos métodos para la corrección tanto de bandas como de *cupping* extendiendo la idea de linealización a 2D [7, 8]. El método sBHCM [7] es el más robusto de los dos frente al ruido, ya que obtiene los factores de corrección a través de un paso de calibración con un maniquí formado por dos materiales que emulan tejido blando y hueso. Sin embargo, al igual que los anteriores, necesita de la segmentación de un tejido que puede fallar en adquisiciones con baja dosis.

Para compensar el efecto de endurecimiento de haz en adquisiciones con baja dosis, podemos recurrir a los algoritmos iterativos. En [9], los autores presentan un método iterativo que incluye un modelo policrómatico de la fuente, sin embargo, requiere del conocimiento del espectro para introducir el efecto de la energía en la matriz de proyección. Este requerimiento se elimina en el método propuesto en [10], que llamaremos bhSIR, con una simplificación del modelo de fuente policromático basado en dos parámetros y un paso de calibración Sin embargo, el uso de métodos iterativos supone un aumento considerable del tiempo de ejecución.

ISBN: 978-84-09-25491-0, pags. 439- 442

En este trabajo se propone un nuevo método para la obtención de imágenes sin artefactos de endurecimiento de haz en tomografía axial computarizada, compensando tanto los artefactos de endurecimiento de haz como el ruido asociado a las imágenes adquiridas en escenarios de baja dosis en FDK, a través de técnicas de aprendizaje profundo.

2. Algoritmo propuesto

El algoritmo propuesto se basa en la red U-net [11], principalmente utilizada para segmentación de imágenes médicas, con el Error Cuadrático Medio (ECM) como función de coste. Para evitar la normalización de las imágenes que salen de la red se ha eliminado la capa sigmoidea de la red original. La Figura 1 muestra un esquema de la red.

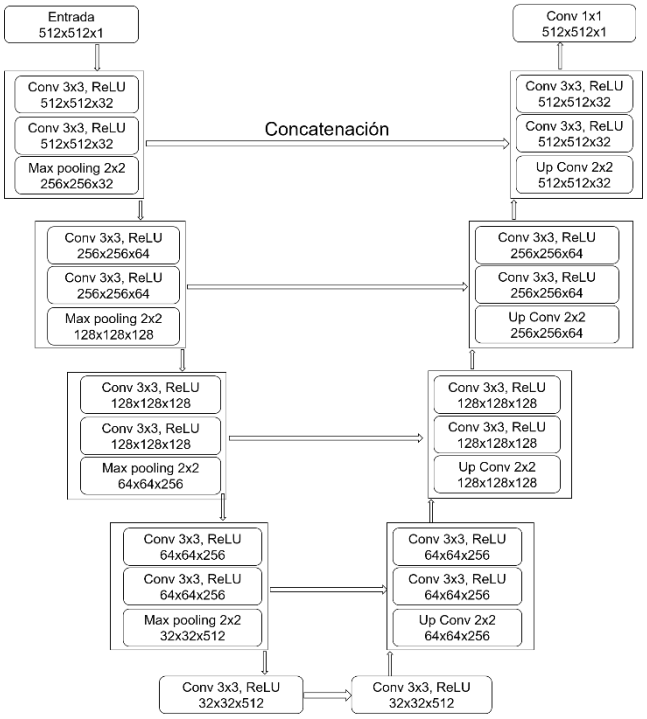


Figura 1: *Arquitectura modificada de la red U-Net*

La red se ha entrenado dos veces de forma independiente para escenarios de dosis estándar (360 proyecciones obtenidas en un rango angular de 360 grados) y escenarios de baja dosis (180 proyecciones obtenidas en un rango angular de 360 grados). El entrenamiento se ha realizado a partir de cortes axiales de tres estudios de roedor adquiridos con el escáner de micro-CT ARGUS/CT (SEDECAL) [12]. Los dos escenarios se han simulado a con el software de simulación y reconstrucción FUX-SIM [13], obteniéndose proyecciones de 512×218 píxeles. La reconstrucción se hace con el algoritmo FDK [14], dando lugar a volúmenes de 512×512×218 vóxeles.

Para ambos entrenamientos se ha considerado como referencia la imagen obtenida con bhSIR [10]. Los parámetros de entrenamiento han sido una tasa de aprendizaje de 10^{-5} durante 100 épocas y el optimizador Adam [15]. La Figura 2 muestra ejemplos de imágenes para cada escenario de baja dosis y dosis estándar, así como su reconstrucción ideal.



Figura 2: *Corte axial de un estudio de roedor adquirida con baja dosis (izquierda), dosis estándar (centro) y reconstruida con el método iterativo (derecha)*

3. Evaluación y Resultados

Se ha aplicado la red neuronal a un estudio de roedor independiente de los estudios entrenados, en escenarios de dosis estándar y baja dosis, comparando el resultado con el obtenido con FDK, con FDK+sBHCM [7] y bhSIR [10]. Para la evaluación cuantitativa se ha calculado el ECM de los volúmenes anteriores con respecto a la reconstrucción de referencia.

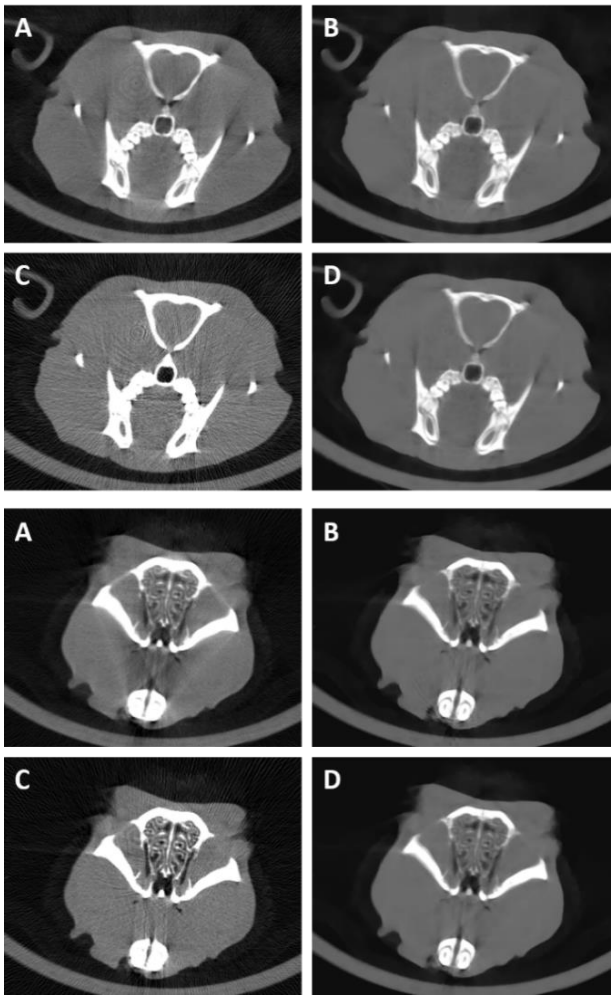


Figura 3: *Escenario de dosis estandar para el corte axial 1 (arriba) y 2 (abajo) obtenido con FDK (A), bhSIR (B), FDK+sBHCM (C) y el esquema propuesto (D)*

En la Figura 3 podemos observar una reducción de las bandas negras en todos los métodos de corrección para el escenario de dosis estándar. Sin embargo, existe un pequeño aumento de ruido en el caso de usar FDK+sBHCM que no se observa en el método propuesto.

La imagen corregida con la red neuronal se asemeja a la imagen referencia obtenida con el método iterativo pero reduciendo el tiempo de ejecución a 0.26 segundos por corte.

En la Figura 4 también se observa una reducción significativa de las bandas negras para todos los métodos en el escenario de baja dosis. El FDK+sBHCM resulta en artefactos debido al bajo muestreo de las proyecciones, que, sin embargo, se ven reducidos por el método propuesto. Al igual que en el escenario de dosis estándar, el método propuesto genera una imagen muy similar a bhSIR pero con un tiempo de ejecución reducido.

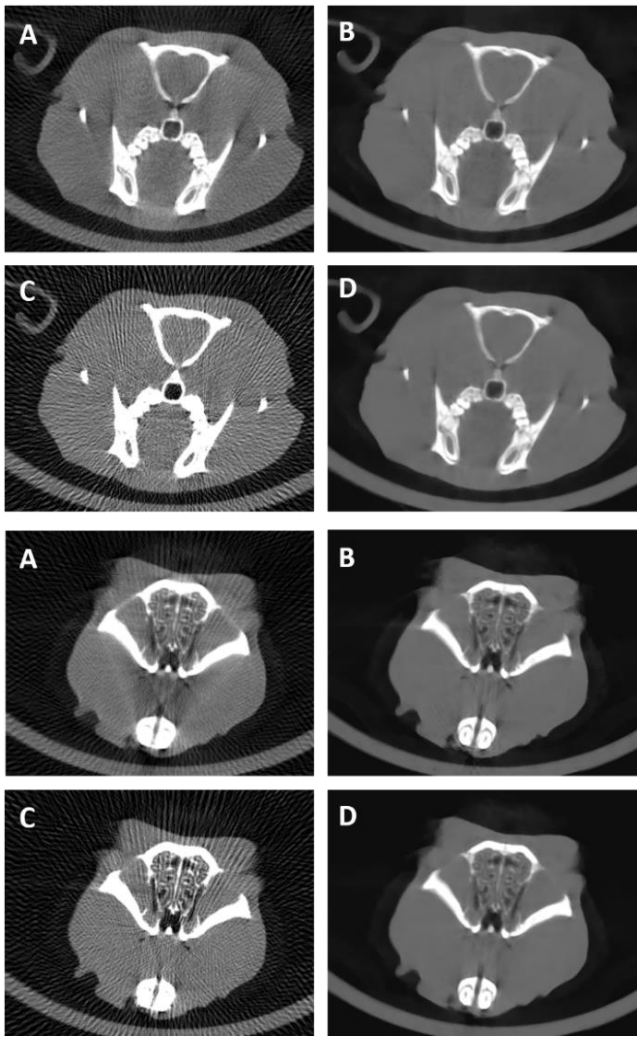


Figura 4: *Escenario de baja dosis para el corte axial 1 (arriba) y 2 (abajo) obtenido con FDK (A), bhSIR (B), FDK+sBCHM (C) y el esquema propuesto (D)*

La Tabla 1 muestra el ECM del volumen completo para los distintos métodos. Se observa como el menor error corresponde al método propuesto.

TABLA I
ECM DEL ESTUDIO DE ROEDOR

	FDK	FDK+sBHCM	Método propuesto
Dosis estandar	0.03	0.23	0.0007
Baja dosis	0.06	0.3	0.003

4. Conclusiones y discusión

En este trabajo se propone un nuevo método para la obtención de imágenes libres de los artefactos producidos por el endurecimiento de haz en imágenes de rayos X. El método se basa en una reconstrucción preliminar con FDK, con artefactos, seguida de un paso de post-proceso basado en técnicas de aprendizaje profundo.

La red utilizada es una modificación de la arquitectura U-net eliminando la función de activación sigmoide final para evitar la normalización de la imagen de salida.

La cantidad de ruido de la imagen obtenida con el método propuesto es inferior a la obtenida con FDK+sBHCM, debido a que el procesamiento de las proyecciones introduce ruido en la imagen reconstruida.

El método propuesto alcanza una reducción de ruido similar al método iterativo con bhSIR [10], pero con una reducción muy considerable del tiempo de ejecución lo que permite su uso en aplicaciones en las que es necesaria la condición de tiempo real, como imagen intraoperatoria. Este método puede ser incluido en los sistemas TAC clínicos o preclínicos sin necesidad de realizar cambios en el hardware o en el método de reconstrucción.

Los resultados muestran una pequeña perdida de resolución espacial comparada con la imagen reconstruida con bhSIR. Esta pérdida de resolución es algo común en la utilización de redes neuronales y puede ser reducida a través de la optimización de la función de coste. En un trabajo futuro, se estudiará la posibilidad de usar diferentes funciones de coste, así como la posibilidad de utilizar redes basadas en arquitecturas GAN (*Generative Adversarial Networks*). También se realizará un estudio sobre la viabilidad de entrenar la red una única vez, combinando los estudios de dosis estándar y baja dosis. Para ello será necesario estudiar si se requiere un aumento del número de imágenes de entrenamiento, incluyendo estudios de diferentes partes anatómicas.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia, Innovación y Universidades (Instituto de Salud Carlos III, proyecto DTS17/00122; Agencia Estatal de Investigación, proyecto DPI2016-79075-R-AEI/FEDER, UE), cofinanciado por Fondos de la Unión Europea (FEDER), "A way of making Europe". Además, ha sido financiado por el Programa de apoyo a la realización de proyectos interdisciplinares de I+D para jóvenes investigadores de la Universidad Carlos III de Madrid 2019-2020 en el marco del Convenio Plurianual Comunidad de Madrid- Universidad Carlos III de Madrid (proyecto DEEPCT-CM-UC3M). El CNIC está financiado por el Ministerio de Ciencia, Innovación y Universidades y la fundación PRO-CNIC y es un centro de excelencia Severo Ochoa (SEV-2015-0505).

Referencias

[1] J. F. Barrett and N. Keat, "Artifacts in CT: recognition and avoidance," *Radiographics*, vol. 24, pp. 1679-91, Nov-Dec 2004.

[2] R. A. Brooks and G. Di Chiro, "Beam hardening in x-ray reconstructive tomography," *Phys Med Biol*, vol. 21, pp. 390-8, May 1976.

[3] O. Nalcioglu and R. Y. Lou, "Post-reconstruction method for beam hardening in computerised tomography," *Phys Med Biol*, vol. 24, pp. 330-40, Mar 1979.

[4] P. M. Joseph and R. D. Spital, "A method for correcting bone induced artifacts in computed tomography scanners," *J Comput Assist Tomogr*, vol. 2, pp. 100-8, Jan 1978.

[5] Y. Kyriakou, E. Meyer, D. Prell, and M. Kachelriess, "Empirical beam hardening correction (EBHC) for CT," *Med Phys*, vol. 37, pp. 5179-87, Oct 2010.

[6] S. Schuller, S. Sawall, K. Stannigel, M. Hulsbusch, J. Ulrici, E. Hell, *et al.*, "Segmentation-free empirical beam hardening correction for CT," *Med Phys*, vol. 42, pp. 794-803, Feb 2015.

[7] C. Martínez, C. De Molina, M. Desco, and M. Abella, "Simple method for beam-hardening correction based on a 2D linearization function," in *The 4th International Meeting on Image Formation in X-Ray Computed Tomography (CTMeeting 2016)*, Bamberg, 2016.

[8] C. Martìnez, C. de Molina, M. Desco, and M. Abella, "Calibration-free method for beam-hardening compensation: preliminary results," in *2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, 2017, pp. 1-3.

[9] I. A. Elbakri and J. A. Fessler, "Statistical image reconstruction for polyenergetic X-ray computed tomography," *IEEE transactions on medical imaging*, vol. 21, pp. 89-99, 2002.

[10] M. Abella, C. Martínez, M. Desco, J. J. Vaquero, and J. A. Fessler, "Simplified Statistical Image Reconstruction for X-ray CT with Beam-Hardening Artifact Compensation," *IEEE transactions on medical imaging*, 2019.

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015, pp. 234-241.

[12] J. J. Vaquero, S. Redondo, E. Lage, M. Abella, A. Sisniega, G. Tapias, *et al.*, "Assessment of a new high-performance small-animal x-ray tomograph," *IEEE Transactions on Nuclear Science*, vol. 55, pp. 898-905, 2008.

[13] M. Abella, E. Serrano, J. Garcia- Blas, I. García, C. de Molina, J. Carretero, *et al.*, "FUX-Sim: Implementation of a fast universal simulation/reconstruction framework for X-ray systems," *PLOS ONE*, vol. 12, p. e0180363, 2017.

[14] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *Josa a*, vol. 1, pp. 612-619, 1984.

[15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

Deep convolutional encoder-decoder network for semantic segmentation of atrial cavity

M. Saiz Vivó[1], A. Colomer [2], V. Naranjo [2]

[1] Escuela Técnica Superior Ingenieros Industriales (ETSII), Universitat Politècnica de València, Valencia, España, marsaivi@etsii.upv.es

[2] Instituto de Investigación e Innovación en Bioingeniería, Universitat Politècnica de València, Valencia, España, adcogra@i3b.upv.es , vnaranjo@i3b.upv.es

Abstract

Atrial Fibrillation (AF) is the most common cardiac arrythmia and its prevalence increases with the ageing population. It is estimated that there are currently 4.5 million cases in Europe. At present, cardiac ablation is the main treatment procedure for AF. To guide and plan this procedure it is essential for clinicians to obtain 3D anatomical reconstruction models of the atria which are patient specific. The aim of this work is to develop computational models using Deep Learning (DL) techniques for left and right atrium (LA and RA respectively) automatic segmentation from MRI volumetric images. For this, a 3D Dual U-Net algorithm was employed. Multiple models were trained with two different databases; the first database had a large number of training samples (80) and the second database had a small number of training samples (19) and was considered of high variability. The model trained with the second database was capable of accurately segmenting the RA with a Dice coefficient of 0.9160 when fine tuning techniques are implemented. The trained models were also evaluated with a third database for LA segmentation. In the third database evaluation experiment, the network trained with the second database yielded a higher Dice coefficient (0.8515) than the network trained with the first database (0.7715) even though it contained a larger number of training samples. These results suggest that a network trained with a high variability database could improve its generalisation capability and yield good segmentation results when evaluated, without retraining, with an external database.

1. Introduction

Atrial Fibrillation (AF) is the most common sustained cardiac arrythmia and has become one of the most important public health problems in the last 20 years [1]. At this moment, the most common method for treating AF patients is radiofrequency catheter ablation to produce scars and electrically isolate the pulmonary veins [2].

Accurate segmentation of LA is highly desirable for patient specific scar characterization to select the appropriate ablation strategy. Furthermore, to plan the procedure 3D geometrical models of both the left and right atrium (LA and RA) are required. 3D patient-specific models of the atria are obtained from the segmentation of MRI and CT images. In clinical routine this segmentation is normally performed manually by experts. However, manual delimitation is a time-consuming, labour-intensive and error-prone method. Therefore, there is a need for computational driven methods for automatic image segmentation. Besides that, the automatic segmentation of atrial structures, specially the LA, remains a challenging task due to the complex geometry of the LA and the morphological variations between patients [3].

In recent years, Deep Learning (DL) techniques have gained popularity in the medical field for image segmentation. However, the need to retrain the networks each time they are provided with an external database is a significant limitation for their adoption in the clinical practice.

In the field of automated atrial segmentation deep neural networks based on Convolutional Neural Networks (CNNs) have recently emerged and shown promising results. For example, Bai et al. [4] applied Fully Convolutional Networks (FCN) for LA and RA segmentation from 2D MRI images.

In addition, automatic segmentation of 3D volumes by 3D neural networks enable the full use of spatial information and their feature extraction may better reflect the anatomical structures in cardiac segmentation. 3D U-Nets were first proposed by Çiçek et al. [5] by expanding the operation filter in 3D of the U-Net, originally introduced by Ronneberger et al. [6]. Several variants of the 3D U-Net have been proposed for 3D LGE-MRI atrial segmentation achieving very promising results [7]

In this work the network proposed by Jia et al. [8] for 3D LGE-MRI LA segmentation will be employed. This network consists of a two-stage 3D U-Net. The 'first' U-Net is used to coarsely segment and locate the left atrium, based on this output the 'second' U-Net accurately segments the left atrium under higher resolution after the region of interest has been cropped.

The aim of this work is to introduce fine tuning techniques and retrain the network proposed by [8] to obtain high generalisation models capable of obtaining good accuracy of LA and RA segmentation with 3D MRI images from an external database without the need of retraining.

2. Materials

For the development of this work three different databases were used to train and test the generalisation capability of the network.

ISBN: 978-84-09-25491-0, pags. 443- 446

Database 1 was provided by the STACOM 2018 Atrial Segmentation Challenge. It contains 100 3D LGE-MRIs images from patients with AF with their corresponding LA binary segmentations. The data resolution is 0.625x0.625x0.625mm^3. All volumes had the same number of slices (88) and there was little variation in terms of image contrast. [9]

Database 2 was provided by the Multi-Modality Whole Heart Segmentation (MM-WHS) challenge, in conjunction with MICCAI 2017. It contains 20 volumetric MRI images from human patients with a data resolution of 0.8-1x0.8-1x1-1.6mm^3. The data also contained the corresponding binary mask of different cardiac structures including RA and LA. [10]

This database can be considered of high variability due to the different sizes of cardiac structures in subjects with different pathologies and the variability in image quality due to motion artifacts, with variable contrast-to-noise ratio (CNR) and signal-to-noise ratio (SNR). Furthermore, there was also high variability in terms of slice number and spatial dimensions of the volumes. [10]

Database 3 was provided by STACOM 2013 Left Atrial Segmentation Challenge in conjunction with MICCAI13. It contains 30 volumetric MRI images acquired with a voxel resolution of 1.25x1.25x2.7mm^3. The data also contained the corresponding binary segmentation of LA cavity. [11]

3. Method

3.1. Dual 3D U-Net

The dual 3D U-Net framework proposed by Jia et al. is illustrated in Figure 1 . It consists of one 3D U-Net that is trained twice with different inputs. The 3D U-Net is a specialization of the net proposed by Çiçek et al. [5] and its implementation follows the work of Isensee et al. [12].

The 3D U-Net has an encoder-decoder architecture with a 5-level depth. Each level in the encoder consists of a convolutional block with 16 filters per convolution and a pooling layer with a strided convolution. To reduce stochasticity induced by small batch size the instance normalization layer instead of batch normalization layer is used. Each convolutional block acts as a pre-activation residual blocks and skip connections propagate same level features from encoder to decoder. In the decoder, segmentation layers at different levels are integrated via concatenation to reduce the coarseness of the final segmentation output. The activation function used throughout the network is the leaky ReLU with negative slope of 0.01.

In the first training, the net is tasked with coarsely locating the region of interest, whether it is LA or RA. The input of the first net is volumetric MR images and its output is preliminary prediction masks of the atrium. A pre-processing function was externally designed in MATLAB which keeps the largest connected components of the prediction masks and computes the spatial location of the atrium. The MR images and ground-truth masks are then cropped with a cuboid centred around the atrium.

The secondary training is performed with the cropped images at full resolution. Cropping the images removes background noise and enables for a more precise segmentation. Specifically, the input of the secondary training is volumetric MR images cropped around the predicted atrium without resampling. The output is binary masks with the final prediction segmentation of the atrium.

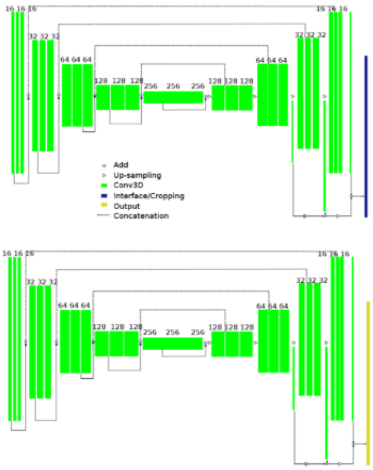


Figure 1. *3D Dual U-Net structure proposed by* [8]*. Green blocks represent 3D features; Dark blue refers to the cropping interface to crop the region of interest of the first U-Net prediction.*

3.2. Training and evaluation

The network was trained several times with fixed experimental settings. The input of the network consisted of volumetric images resized to (224, 144, 96) with a batch size equal to 1. Training was performed using the Adam optimizer with an initial learning rate equal to 5e^{-4} that reduces by half after 10 epochs if the validation loss is not improving. The convergence is defined as not improving after 50 epochs and the maximum number of epochs is 500 with 200 steps per epoch. Training-validation split ratio equals to 0.8. In addition, data augmentation techniques were applied during training to reduce overfitting. Furthermore, to avoid the issue of class imbalance the Dice loss function introduced by Millietari et al. [13] was used. The evaluation of the predicted segmentations is performed with the Dice coefficient. Deep fine-tuning techniques were also used to improve segmentation accuracy by initialising the weights of all convolutional layers with pre-trained weights and retraining the network.

The training and evaluation processes performed throughout this project can be divided into 3 experiments.

In Experiment 1 the network was trained with images from Database 1 (80 volumes) and evaluated with testing images also from Database 1 (20 volumes). The aim of this experiment was to obtain a *high-performance* model for LA segmentation. The trained model obtained in this experiment will be referred to as N1.

In Experiment 2 the network was trained and tested with Database 2. This database consists of 16 volumes for training and 3 volumes for testing. The network was trained and tested twice, once with LA ground-truth masks and the other with RA. The trained model obtained in this experiment will be referred to as N2.

In the second part of Experiment 2 deep fine-tuning techniques were implemented to increase segmentation accuracy. This was done by initialising the weights from all convolutional layers with the pre-trained weights obtained in the model N1 and retraining the network. As the model N1 can be considered of *high-performance* this served as a positive inference in the learning of the network for LA and RA segmentation. The trained model obtained in this experiment will be referred to as N3.

In Experiment 3 the trained models obtained in the previous experiments: N1, N2 and N3 were evaluated with images from Database 3. Database 3 contains ground-truth for only the LA, therefore the N2 and N3 models for RA were excluded from this experiment. 30 volumes were passed as testing images and the prediction output was post processed and 3D reconstructed. The aim was to test the generalisation capacity of the previously trained networks when they were evaluated, without retraining, with an external database.

4. Results and Discussion

4.1. Experiment 1

In Experiment 1 the network was trained and tested with patients from Database 1. Database 1 contains 20 testing volumes each of them composed of 88 slices. The mean Dice Coefficient and standard deviation obtained during evaluation was 0.9155 ± 0.0270. Experiment 1 yields very accurate segmentation results. This was to be expected as Database 1 corresponds to the data provided by the 2018 LA segmentation challenge and the network proposed by [8] was specifically designed to enter this challenge. Therefore, by keeping the same learning parameters, the dice obtained coincides with the dice reported in [8] (0.91-0.92). Furthermore, the network was trained with a relatively large number of patients which contributed to the good performance of the network.

4.2. Experiment 2

Table 1 and Table 2 show the Dice coefficient obtained for the 3 testing patients with and without fine tuning implementation for LA and RA segmentation, respectively.

Technique	Patient 1.	Patient 2.	Patient 3.	Mean ±σ
w/o f.t	0.8693	0.8756	0.8736	0.8729 ± 0.0032
f.t	0.8781	0.8948	0.8707	0.8813 ± 0.0124

Table 1. *Dice coefficients in LA prediction segmentation*

Technique	Patient 1.	Patient 2.	Patient 3.	Mean ±σ
w/o f.t	0.9162	0.8773	0.8892	0.8942 ± 0.0199
f.t	0.9273	0.9043	0.9166	0.9160 ± 0.0115

Table 2. *Dice coefficients in RA prediction segmentation*

Each testing volume contains between 100-200 image slices. However, to obtain more representative results an increased number of testing patients would be needed.

From the results obtained in Table 1 and Table 2 it can be observed that the application of deep fine-tuning does increase the dice coefficient for both atrium segmentations. Furthermore, the results shown in Table 2 suggest that the network could successfully segment the RA which attests for the robustness of the algorithm.

The Dice obtained in both LA and RA were higher than the mean Dice obtained by other research groups [10] which suggests that fine tuning technique is an effective method for LA and RA segmentation.

4.3. Experiment 3.

This experiment consisted in an evaluation of the previously trained models using an external database (Database 3) to test their generalization ability. Table 3 shows the average Dice coefficient obtained with each trained network for the 30 patients in Database 3.

Network N	Mean ±σ
N1	0.7773 ± 0.0643
N2	0.8230 ± 0.0943
N3	0.8515 ± 0.0796

Table 1*. Average Dice for each trained network N*

Figure 2 shows the axial view of predictions obtained for 3 patients with the 3 different trained networks.

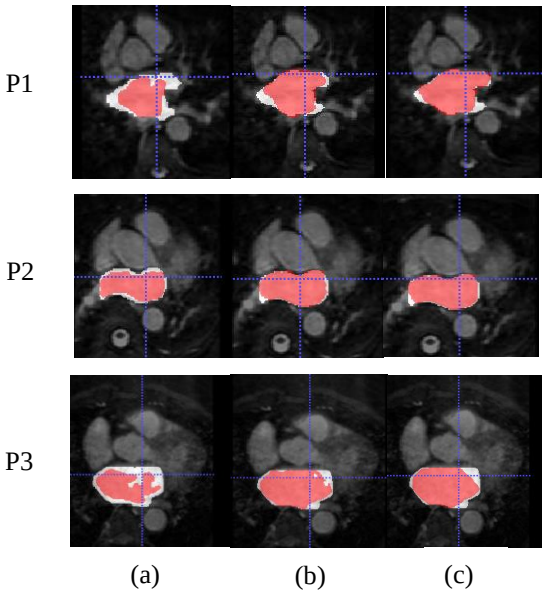


Figure 2*. Axial slice of LA segmentations for 3 patients (P1, P2, P3) in networks (a) N1, (b) N2 and (c) N3. ground-truth mask in white, prediction mask in red and intersection in pink.*

Figure 3 shows the 3D reconstruction of the predictions obtained for 2 patients.

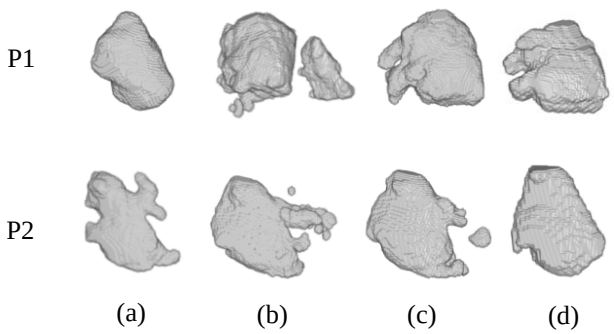


Figure 3. *LA 3D reconstructions of prediction for 2 patients (P1 and P2) obtained with trained networks (a) N1, (b) N2 and (c) N3. (d) 3D reconstruction of ground-truth.*

As can be observed in Table 3 the highest Dice coefficient was obtained with the trained model from Experiment 2 with fine tuning. This suggests that training a network with highly variable images and fine tuning the weights with a high-performance network could increase the generalisation capacity of the network and result in higher accuracy segmentations when evaluated with external databases. This can be further observed in Figure 2. It should be remarked upon that a network trained with 16 volumes (N2/N3) could provide segmentations of a higher accuracy than a network trained with 80 volumes (N1).

In Figure 3, when comparing the 3D reconstruction with the 3D ground-truth, it can be seen that the network trained with Database 1 (N1) was not capable of capturing the geometrical anatomy of the atria whilst N2 and N3 remained closer. Furthermore, with fine tuning the atrial geometry further approximates to the ground-truth in both patients.

5. Conclusions

In conclusion, after introducing deep fine-tuning techniques and training the network with databases of different image variability the results obtained show that a network trained with a highly variable database is more capable of accurately segmenting an external database without retraining. Furthermore, with the additional use of fine tuning the segmentation accuracy can be further improved. These results may be a first step towards the integration of automatic segmentation algorithms in the clinical practice for AF management. For future work the trained networks should also be evaluated with databases containing RA ground-truths.

Acknowledgements

I would like to thank Dr. Maxime Sermesant and Dr. Shuman Jia from the INRIA Institute in France for allowing me access to their code and their help with any doubts.

References

[1] M. Zoni-Berisso, F. Lercari, T. Carazza, and S. Domenicucci, "Epidemiology of atrial fbrillation: European perspective," *Clin. Epidemiol.*, vol. 6, no. 1, pp. 213–220, 2014.

[2] L. Li *et al.*, "Atrial scar quantification via multi-scale CNN in the graph-cuts framework," *Med. Image Anal.*, vol. 60, p. 101595, 2020.

[3] T. Kurzendorfer, C. Forman, M. Schmidt, C. Tillmanns, A. Maier, and A. Brost, "Fully automatic segmentation of left ventricular anatomy in 3-D LGE-MRI," *Comput. Med. Imaging Graph.*, vol. 59, pp. 13–27, 2017.

[4] W. Bai *et al.*, "Automated cardiovascular magnetic resonance image analysis with fully convolutional networks," *J. Cardiovasc. Magn. Reson.*, vol. 20, no. 1, pp. 1–12, 2018.

[5] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-net: Learning dense volumetric segmentation from sparse annotation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9901 LNCS, no. June 2016, pp. 424–432, 2016.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9351, pp. 234–241, 2015.

[7] S. Vesal, N. Ravikumar, and A. Maier, "Dilated Convolutions in Neural Networks for Left Atrial Segmentation in 3D Gadolinium Enhanced-MRI," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11395 LNCS, pp. 319–328, 2019.

[8] S. Jia *et al.*, "Automatically Segmenting the Left Atrium from Cardiac Images Using Successive 3D U-Nets and a Contour Loss," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11395 LNCS, pp. 221–229, 2019.

[9] "http://atriaseg2018.cardiacatlas.org/." .

[10] X. Zhuang *et al.*, "Evaluation of algorithms for Multi-Modality Whole Heart Segmentation: An open-access grand challenge," *Med. Image Anal.*, vol. 58, p. 101537, 2019.

[11] "https://www.cardiacatlas.org/challenges/left-atrium-segmentation-challenge/." .

[12] F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, "Brain tumor segmentation and radiomics survival prediction: Contribution to the BRATS 2017 challenge," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10670 LNCS, pp. 287–297, 2018.

[13] F. Milletari, N. Navab, and S. A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," *Proc. - 2016 4th Int. Conf. 3D Vision, 3DV 2016*, pp. 565–571, 2016.

Estimación del desplazamiento horizontal del detector en un sistema de rayos X utilizando Aprendizaje por Transferencia

C. F. Del Cerro[1,2], R. C. Gimenez[2], P. M. Olmos[1,2], A. Piol[1,2], M. Desco[1,2,3,4], M. Abella[1,2,3]

[1] Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España

[2] Dept. Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid, España

[3] Centro Nacional Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, España

[4] Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid, España

Resumen

En las imágenes de tomografía axial computarizada por rayos X (TAC) pueden aparecer artefactos causados por errores en la calibración geométrica del sistema de rayos X. Aunque existen muchos estudios que tratan de resolver el problema de la calibración geométrica de los sistemas de TAC, a menudo son diseñados específicamente para una configuración concreta y necesitan la adquisición previa de un maniquí. En este trabajo se propone un método basado en aprendizaje profundo (Deep Learning) para estimar la calibración geométrica de un sistema de TAC por rayos X directamente sobre una reconstrucción preliminar con artefactos. Los resultados preliminares muestran la viabilidad de la propuesta que abre la puerta a eliminar la necesidad de una fase de calibración previa con un maniquí específico para la calibración de sistemas de TAC de rayos X, sin suponer un extra de tiempo de cómputo.

1. Introducción

La tomografía axial computarizada por rayos X (TAC) combina una serie de imágenes 2D, llamadas proyecciones, tomadas desde diferentes ángulos alrededor del cuerpo para crear un volumen 3D conocido como imagen reconstruida. Para que estas imágenes sean de alta calidad se debe conocer con mucha precisión las posiciones de la fuente y del detector en cada proyección. Errores en estas posiciones resultan en artefactos de desalineamiento en la imagen reconstruida. Los sistemas TAC suelen incluir un paso de calibración geométrica para estimar los errores de la posición de los distintos elementos respecto a su posición ideal de diseño, fruto de tolerancias en los procesos de fabricación. En la Figura 1 se pueden ver algunos de estos errores en la posición del detector.

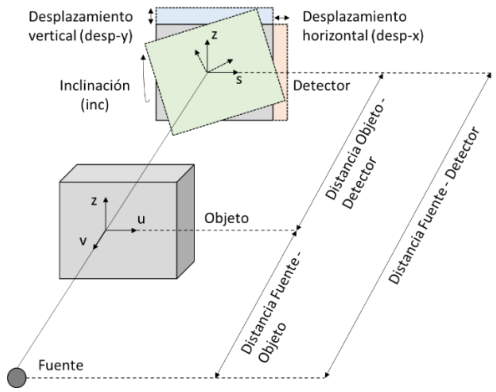


Figura 1. *Esquema de un sistema TAC con diferentes errores en la posición del detector.*

La aproximación tradicional al problema de la calibración geométrica hace uso de un maniquí de calibración que debe ser específicamente diseñado con marcadores que son esenciales para guiar la estimación de los parámetros. En [1] se propuso un método que estima el centro de rotación en SPECT con geometría de haz cónico a partir de las coordenadas del centro de una fuente puntual proyectada sobre 360 grados. En [2, 3] se presentaron alternativas a este enfoque para lograr una calibración más precisa. En [4] se presentó un enfoque similar para sistemas de rayos X. Más tarde, en [5], los autores desarrollaron un método analítico basado en un maniquí sencillo con dos marcadores esféricos. En [6], se mejora este método incorporando la calibración de cada proyección individualmente, sin asumir una trayectoria fuente-detector perfectamente circular, utilizando un maniquí con mayor número de marcadores. En la actualidad, han surgido alternativas de bajo coste, rápidas y flexibles para diversos sistemas de TAC de haz cónico que se basan en maniquís construidos con ladrillos de LEGO [7-9]. El principal inconveniente de las técnicas mencionadas es que precisan la adquisición previa de un maniquí de calibración, que normalmente está limitado a una configuración concreta.

Alternativamente, se han desarrollado métodos de auto-calibración que tratan de determinar los parámetros de desalineamientos a partir de los datos adquiridos, sin necesidad de maniquí de calibración. En el espacio de la proyección se propuso en [10, 11] un enfoque que trata de estimar el error aprovechando las redundancias de las proyecciones obtenidas en sistemas de TAC de haz cónico. Otros estudios han intentado optimizar la correlación cruzada en el espacio de las proyecciones para determinar desplazamientos verticales y horizontales del detector [12, 13]. Por otra parte, se han desarrollado métodos que trabajan en el espacio de la imagen reconstruida para estimar los parámetros de calibración. En [14, 15] se propuso una corrección iterativa utilizando la minimización de la entropía de la imagen reconstruida. Con un enfoque similar, [16-18] se inspiran en los métodos de autoenfoque de sistemas ópticos, en los que se maximiza la nitidez de la imagen reconstruida para estimar los desalineamientos desconocidos. En [19] se propuso un método que explota la propiedad de simetría de la suma de proyecciones para obtener los parámetros geométricos. El principal problema de los métodos de auto-calibración es su alto coste computacional.

ISBN: 978-84-09-25491-0, pags. 447- 450

Este trabajo propone un método de calibración para sistemas TAC con geometría de haz cónico a partir del volumen reconstruido utilizando técnicas de aprendizaje profundo. El estudio preliminar se centra en la estimación del parámetro de calibración que más impacto tiene en la calidad de la imagen, el desplazamiento horizontal del detector (*desp-x* en la Figura 1).

2. Algoritmo Propuesto

El algoritmo comienza con una reconstrucción preliminar considerando geometría ideal (*desp-x* = 0.0 mm), que se utiliza como entrada a la red. La red devuelve un valor estimado de *desp-x* para cada corte y después se realiza la media.

2.1. Red Neuronal Convolucional

El algoritmo se basa en el flujo de trabajo típico de aprendizaje por transferencia. Se probaron las redes ResNet-34, ResNet-50 y ResNet-101 [20], seleccionándose ésta última al ser la que consiguió un error menor. Se han mantenido todas las capas convolucionales con los pesos generados tras entrenar la red con el conjunto de datos conocido como *Imagenet* [21], pero se ha definido un predictor nuevo inicializado aleatoriamente, tal y como se muestra en la Figura 2. Este nuevo predictor está compuesto por dos capas lineales con una función de activación (*ReLU*) entre ambas para obtener como salida la estimación del valor de *desp-x*.

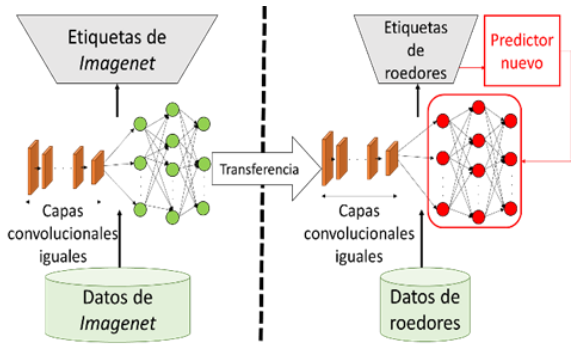


Figura 2. *Estrategia de aprendizaje por transferencia implementada en este trabajo.*

2.2. Generación de la base de datos

Para crear la base de datos se utilizaron seis estudios de roedor adquiridos con el micro-CT Add-On de SEDECAL [22] mostrados en la Figura 3.

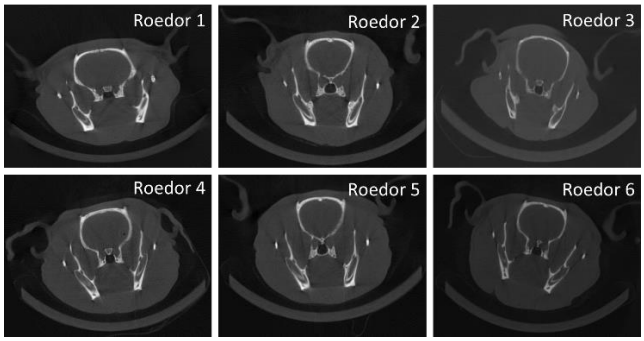


Figura 3. *Corte axial central de los seis estudios de roedor utilizados.*

Para cada roedor, se simularon 25 conjuntos de proyecciones con distintos valores de *desp-x* siguiendo una distribución uniforme con media 0 y desviación estándar usando la herramienta de simulación FuxSim [23]. Cada conjunto de proyecciones se reconstruyó tomando *desp-x* = 0.0 mm, resultando en un volumen con artefactos de dimensiones 336×436×200 y tamaño de píxel 0.12 mm. De cada volumen se extrajeron los 200 cortes axiales generándose una base de datos de imágenes 2D. Los 200 cortes axiales correspondientes a los cinco primeros roedores se utilizaron para entrenar la red y los del sexto para validar el modelo. La Figura 4 muestra ejemplos con distintos valores de *desp-x* para el corte axial central del roedor 1.

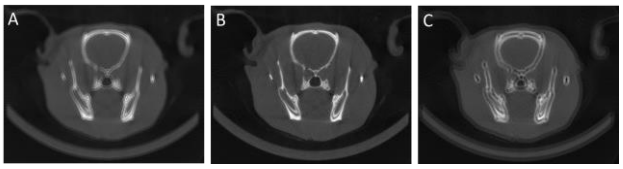


Figura 4. *Corte axial del roedor 1 reconstruido con diferentes valores de desp-x: 1.9 mm (A), -0.5 mm (B), -3.8 mm (C).*

2.3. Entrenamiento y validación de la red

Se utilizaron 25000 imágenes para el entrenamiento y 5000 para la validación. El entrenamiento de la red realizó en dos pasos: en un primer paso se entrenó el predictor, manteniendo intactos los pesos de las capas convolucionales y en un segundo paso entrenó la red de principio a fin para aumentar la precisión del modelo. Como función de coste usó el error cuadrático medio entre el valor real del desalineamiento y el valor la salida de la red. El optimizador elegido fue Adam [24]. Para seleccionar la tasa de aprendizaje más adecuada se utilizó la prueba diseñada por Leslie N. Smith en [25], cuyo resultado se puede ver en la Figura 5 para el entrenamiento del predictor.

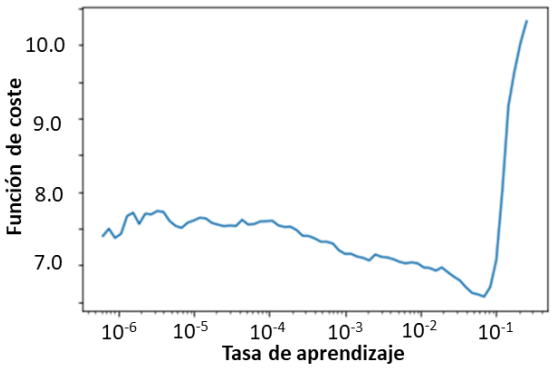


Figura 5. *Resultado de la prueba de Leslie N. Smith para determinar el valor óptimo de la tasa de aprendizaje para el predictor.*

Observando el resultado de la Figura 5, se decidió un valor para la tasa de aprendizaje de 10^{-3} (zona de máxima pendiente). Una vez alcanzado un error lo suficientemente pequeño, se descongelaron las capas convolucionales y se volvió a realizar la prueba de Leslie N. Smith para determinar la tasa de aprendizaje óptima para entrenar la red de principio a fin, obteniendo el resultado mostrado en la Figura 6.

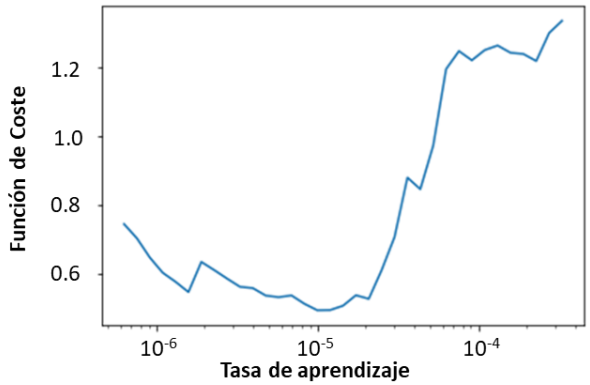


Figura 6. *Resultado de la prueba de Leslie N. Smith para determinar el valor óptimo de la tasa de aprendizaje para la red completa.*

En este caso se seleccionó una tasa de aprendizaje discriminativa entre 10^{-6} y 10^{-5} correspondiente de nuevo con la zona de máxima pendiente.

Una vez definidos todos los parámetros, el error cuadrático medio conseguido por la red es de 0.3 mm sobre el conjunto de validación.

3. Evaluación del método y resultados

Para evaluar el método se utilizó el roedor 6, que no se incluyó en el conjunto de entrenamiento. Se generaron proyecciones con *desp-x* igual a 1.2 mm, que es un valor representativo de este tipo de desalineamiento en sistemas TAC de pequeño animal. El valor de *desp-x* estimado fue de 1.1 mm. En la Figura 7 se puede observar el corte axial central de los volúmenes reconstruidos con *desp-x* igual a 1.2 mm (ideal), 0.0 mm (sin calibrar) y 1.1 mm (estimado) y el zoom de una región de interés donde se puede observar mejor las diferencias. Se puede comprobar como la reconstrucción con el valor estimado mediante el método propuesto consigue recuperar los detalles presentes en el caso ideal.

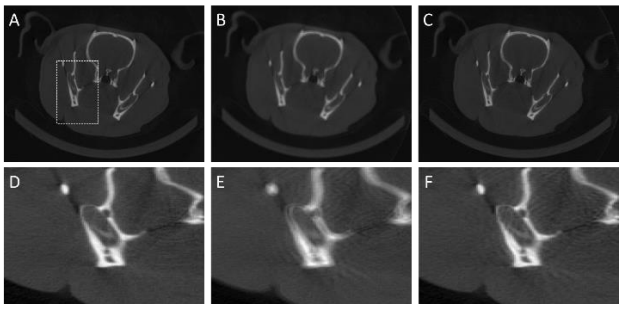


Figura 7. *Corte axial central de los volúmenes reconstruidos con desp-x igual a: 1.2mm (A), 0.0 mm (B), 1.1 mm (C). Aumento de: (A) en (D), (B) en (E), (C) en (F).*

El error cuadrático medio con el volumen ideal se reduce de 0.0038 para el volumen reconstruido con *desp-x* = 0.0 mm a 0.0012 cuando se reconstruye con la estimación obtenida con el método propuesto. Se aprecia como el método propuesto reduce los dobles bordes debido a los errores de calibración. Sin embargo, aún se aprecian ligeros artefactos probablemente debidos al error de 0.1 mm de la estimación.

4. Discusión y Conclusiones

Este trabajo propone una nueva metodología de calibración geométrica en imágenes de TAC mediante el uso de aprendizaje profundo, basada en el análisis directo de una reconstrucción preliminar con artefactos. Esta idea abre la puerta a eliminar la necesidad de una fase de calibración previa con un maniquí específico para la calibración de sistemas de TAC de rayos X, sin suponer un extra de tiempo de cómputo. En este trabajo se presenta una primera aproximación, considerando solo errores debidos al desplazamiento horizontal del detector, que es el parámetro que más afecta en la calidad de imagen reconstruida.

Los resultados muestran una recuperación de los detalles dobles bordes, pero mantienen ligeros artefactos debido al error en la estimación de *desp-x*. Esto puede deberse a que para el entrenamiento se han utilizado imágenes en las que las zonas con artefactos eran pequeñas respecto a la imagen completa. Por otra parte, la base de datos actualmente compuesta por seis estudios de roedor, lo que limita tanto el número de muestras como la variabilidad de estas. Como trabajo futuro, se ampliará esta base de datos para intentar reducir el error y obtener un modelo más robusto.

El paso siguiente será la incorporación de errores en otros parámetros geométricos. Por último, se investigará la posibilidad de trasladar este modelo a sistemas de adquisición más complejos donde cada una de las proyecciones adquiridas tiene una calibración diferente, lo que supondría un avance significativo para los sistemas de tomografía con geometrías no estándar.

5. Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia, Innovación y Universidades (Instituto de Salud Carlos III, proyecto DTS17/00122; Agencia Estatal de Investigación, proyecto DPI2016-79075-R-AEI/FEDER, RTI2018-099655-B-100, UE), cofinanciado por Fondos de la Unión Europea (FEDER), “A way of making Europe”. Además, ha sido financiado por el Programa de apoyo a la realización de proyectos interdisciplinares de I+D para jóvenes investigadores de la Universidad Carlos III de Madrid 2019-2020 en el marco del Convenio Plurianual Comunidad de Madrid- Universidad Carlos III de Madrid (proyecto DEEPCT-CM-UC3M). El CNIC está financiado por el Ministerio de Ciencia, Innovación y Universidades y la fundación PRO-CNIC y es un centro de excelencia Severo Ochoa (SEV-2015-0505).

Referencias

[1] G. T. Gullberg, B. M. W. Tsui, C. R. Crawford, J. G. Ballard, and J. T. Hagius, "Estimation of geometrical parameters and collimator evaluation for cone beam tomography," *Med. Phys.*, vol. 17, pp. 264–72, 1990.

[2] J. Li, R. Jaszczak, H. Wang, K. Greer, and R. Coleman, "Determination of both mechanical and electronic shifts in cone beam SPECT," *Physics in Medicine & Biology*, vol. 38, p. 743, 1993.

[3] P. Rizo, P. Grangeat, and R. Guillemaud, "Geometric calibration method for multiple-head

cone-beam SPECT system," *IEEE transactions on nuclear science*, vol. 41, pp. 2748-2757, 1994.

[4] A. Rougée, C. Picard, C. Ponchut, and Y. Trousset, "Geometrical calibration of X-ray imaging chains for three-dimensional reconstruction," *Computerized Medical Imaging and Graphics*, vol. 17, pp. 295-300, 1993.

[5] F. Noo, R. Clackdoyle, C. Mennessier, T. A. White, and T. J. Roney, "Analytic method based on identification of ellipse parameters for scanner calibration in cone-beam tomography," *Physics in Medicine & Biology*, vol. 45, p. 3489, 2000.

[6] Y. Cho, D. J. Moseley, J. H. Siewerdsen, and D. A. Jaffray, "Accurate technique for complete geometric calibration of cone-beam computed tomography systems," *Medical physics*, vol. 32, pp. 968-983, 2005.

[7] F. Quercioli, B. Tiribilli, A. Mannoni, and S. Acciai, "Optomechanics with LEGO," *Applied optics*, vol. 37, pp. 3408-3416, 1998.

[8] Z. H. Levine, S. Grantham, D. S. Sawyer IV, A. P. Reeves, and D. F. Yankelevitz, "A low-cost fiducial reference phantom for computed tomography," *Journal of research of the National Institute of Standards and Technology*, vol. 113, p. 335, 2008.

[9] V. Nguyen, J. De Beenhouwer, J. Sanctorum, S. Van Wassenbergh, S. Bazrafkan, J. Dirckx, *et al.*, "A low-cost geometry calibration procedure for a modular cone-beam X-ray CT system," *Nondestructive Testing and Evaluation*, pp. 1-14, 2020.

[10] V. Patel, R. Chityala, K. Hoffmann, C. Ionita, D. Bednarek, and S. Rudin, "Self-calibration of a cone-beam micro-CT system," *Medical physics*, vol. 36, pp. 48-58, 2009.

[11] D. Panetta, N. Belcari, A. Del Guerra, and S. Moehrs, "An optimization-based method for geometrical calibration in cone-beam CT without dedicated phantoms," *Physics in Medicine & Biology*, vol. 53, p. 3841, 2008.

[12] S. Mayo, P. Miller, D. Gao, and J. Sheffield-Parker, "Software image alignment for X-ray microtomography with submicrometre resolution using a SEM-based X-ray microscope," *Journal of microscopy*, vol. 228, pp. 257-263, 2007.

[13] A. Sasov, X. Liu, and P. L. Salmon, "Compensation of mechanical inaccuracies in micro-CT and nano-CT," in *Developments in X-ray Tomography VI*, 2008, p. 70781C.

[14] Y. Kyriakou, R. Lapp, L. Hillebrand, D. Ertel, and W. Kalender, "Simultaneous misalignment correction for approximate circular cone-beam computed tomography," *Physics in Medicine & Biology*, vol. 53, p. 6267, 2008.

[15] Y. Kyriakou, R. Lapp, L. Hillebrand, D. Ertel, and W. Kalender, "Image-based online correction of misalignment artifacts in cone-beam CT," in *Medical Imaging 2009: Physics of Medical Imaging*, 2009, p. 72581V.

[16] A. Kingston, A. Sakellariou, A. Sheppard, T. Varslot, and S. Latham, "An auto-focus method for generating sharp 3D tomographic images," in *Developments in X-Ray Tomography VII*, 2010, p. 78040J.

[17] A. Kingston, A. Sakellariou, T. Varslot, G. Myers, and A. Sheppard, "Reliable automatic alignment of tomographic projection data by passive auto-focus," *Medical physics*, vol. 38, pp. 4934-4945, 2011.

[18] T. Varslot, A. Kingston, G. Myers, and A. Sheppard, "High-resolution helical cone-beam micro-CT with theoretically-exact reconstruction from experimental data," *Medical physics*, vol. 38, pp. 5459-5476, 2011.

[19] Y. Meng, H. Gong, and X. Yang, "Online geometric calibration of cone-beam computed tomography for arbitrary imaging objects," *IEEE transactions on medical imaging*, vol. 32, pp. 278-288, 2012.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.

[21] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, pp. 211-252, 2015.

[22] J. Vaquero, S. Redondo, E. Lage, M. Abella, A. Sisniega, G. Tapias, *et al.*, "Assessment of a new high-performance small-animal X-ray tomograph," *IEEE transactions on nuclear science*, vol. 55, pp. 898-905, 2008.

[23] M. Abella, E. Serrano, J. Garcia-Blas, I. García, C. De Molina, J. Carretero, *et al.*, "FUX-Sim: Implementation of a fast universal simulation/reconstruction framework for X-ray systems," *PloS one*, vol. 12, p. e0180363, 2017.

[24] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[25] L. N. Smith, "Cyclical learning rates for training neural networks," in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2017, pp. 464-472.

Diferenciación entre variantes de afasia progresiva primaria mediante el análisis cuantitativo de neuroimágenes de ^{18}F-FDG PET

N. Pérez-Jiménez[1], A.P. Seiffert[1], A. Gómez-Grande[2], A. Villarejo-Galende[3-6], E.J. Gómez[1,7], P. Sánchez-González[1,7]

[1] Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, España; perezjimenez.natalia@gmail.com, {ap.seiffert, enriquejavier.gomez, p.sanchez}@upm.es

[2] Servicio de Medicina Nuclear, Hospital Universitario 12 de Octubre, Madrid, España; adolfogomez@gmail.com

[3] Servicio de Neurología, Hospital Universitario 12 de Octubre, Madrid, España; avgalende@yahoo.es

[4] Facultad de Medicina, Universidad Complutense de Madrid, Madrid, España

[5] Grupo de Enfermedades Neurodegenerativas, Instituto de Investigación Sanitaria (i+12), Hospital Universitario 12 de Octubre, Madrid, España

[6] Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Madrid, España

[7] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Madrid, España

Resumen

La afasia progresiva primaria (APP) es un síndrome clínico que causa un deterioro progresivo del lenguaje, pero en el que se mantienen relativamente preservadas el resto de las funciones cognitivas, instrumentales o conductuales. Se definen tres variantes principales de APP (no-fluente, semántica y logopénica) que presentan cursos clínicos y características distintos. Además, cada una de dichas variantes están relacionadas con enfermedades neurodegenerativas subyacentes y presentan patrones de metabolismo alterado característicos y visibles en neuroimágenes de PET con ^{18}F-fluorodesoxiglucosa (FDG). En este trabajo se estudian de manera semicuantitativa los patrones de metabolismo en una muestra de 19 pacientes con APP. Se extraen ratios del valor estandarizado de captación (SUVR) de varias regiones corticales y subcorticales y se realizan análisis tanto intravariante como intervariante. Se detectan diferencias entre los valores del metabolismo cerebral entre las distintas variantes de APP, siendo de especial interés el lóbulo temporal. Los resultados obtenidos indican la utilidad del estudio del metabolismo cerebral mediante imágenes de ^{18}F-FDG PET para la diferenciación de las variantes de APP.

1. Introducción

La afasia es un trastorno neurológico que se caracteriza por la imposibilidad parcial o total de emitir palabras o comprender su significado. La afasia progresiva primaria (APP) o Afasia de Mesulam es un tipo concreto de afasia [1]. Se trata de un síndrome clínico heterogéneo caracterizado por la degeneración de las regiones del lenguaje en el hemisferio dominante. Además, dicho síndrome causa un deterioro progresivo del lenguaje en el que se mantienen relativamente preservadas el resto de las funciones cognitivas, instrumentales o conductuales y que puede ser el modo de presentación de diversas enfermedades neurodegenerativas [1, 2].

En los últimos años, la clasificación de la APP en varias formas clínicas y las patologías subyacentes que están relacionadas con ellas es una línea de investigación de gran interés. Está demostrado que según las zonas del cerebro que están afectadas, los síntomas de la afasia varían y hay varias clasificaciones que describen subtipos clínicos de APP. La más extendida es la clasificación de Gorno-Tempini et al. [3], que define 3 subtipos o variantes principales de APP y 3 patologías subyacentes con las que están estrechamente relacionadas [2, 3]:

- Variante agramatical o no-fluente (APPnf): Las patologías subyacentes principales son las tauopatías. Los pacientes presentan un lenguaje hipofluente y laborioso, con gramática deficiente, y tienen dificultad en la comprensión de estructuras gramaticales complejas. Se relaciona con alteraciones en el lóbulo frontal.
- Variante semántica (APPvs): Está altamente relacionada con la demencia frontotemporal (DFT). Los pacientes tienen dificultades a la hora de nombrar objetivos o imágenes y al intentar evocar y comprender el significado de palabras sencillas, pero presentan lenguaje fluente y gramática correcta. Se relaciona con alteraciones en el lóbulo temporal.
- Variante logopénica (APPvl): Se asocia mayoritariamente a pacientes con enfermedad de Alzheimer (EA). Los pacientes tienen un lenguaje lento, anómico y con muchas pausas para la búsqueda de palabras. Presentan también dificultad para repetir frases y evocar palabras. Sin embargo, la gramática y la comprensión de palabras sencillas están preservadas. Se relaciona con alteraciones en el lóbulo parietal y temporal.

Existen, sin embargo, una minoría de pacientes que presentan APP no clasificable (APPnc) que no pueden incluirse en ninguno de los grupos previamente mencionados. Esto

ISBN: 978-84-09-25491-0, pags. 451- 454

puede deberse a que presentan síntomas lingüísticos aislados o bien porque tienen características de más de una variante [4].

El diagnóstico clínico de la APP consiste, básicamente, en comprobar que, a pesar de existir una disfunción progresiva del lenguaje del paciente, existe un correcto funcionamiento del resto de facultades mentales como la memoria, las habilidades visuoespaciales y el comportamiento. Para descartar otras posibles enfermedades y detectar las regiones afectadas se suelen realizar pruebas neurocognitivas y del lenguaje y pruebas de neuroimagen [4]. Sin embargo, cabe mencionar que el correcto diagnóstico y clasificación de APP es complejo y la diferenciación entre las variantes de APP y otras enfermedades neurodegenerativas en base a la presentación clínica no siempre es clara.

Como está mencionado anteriormente, cada variante de APP presenta unas alteraciones corticales diferentes y, por lo tanto, algunas modalidades de neuroimagen son una herramienta de ayuda en el diagnóstico. De esta manera, las alteraciones pueden ser tanto atrofia, reduciendo el grosor cortical y evaluado mediante resonancia magnética, como del metabolismo cortical que se visualiza mediante imágenes PET con ^{18}F-fluorodesoxiglucosa (FDG) [4–6].

En el presente trabajo se estudia la capacidad de las imágenes ^{18}F-FDG PET para diferenciar entre las variantes de APP definidos por Gorno-Tempini et al. [3]. El análisis visual de estas imágenes implica una variabilidad interoperador en el diagnóstico. Por lo tanto, se realizan análisis inter e intravariante de valores cuantitativos del metabolismo regional para detectar regiones cerebrales que presentan alteraciones características de cada variante de APP.

2. Material y métodos

2.1. Materiales

En el trabajo se emplean imágenes ^{18}F-FDG PET de 19 pacientes diagnosticados con APP y clasificados según Gorno-Tempini et al [3]. Los datos demográficos se resumen en la Tabla 1.

Variante de APP	N	Sexo (F/M)	Edad (años ± desviación típica)
APPnf	9	6/3	68,5 ± 11,5
APPvs	1	0/1	68
APPvl	4	1/3	67,5 ± 15,5
APPnc	5	1/4	69,6 ± 6,6
Total	19	8/11	68,4 ± 16,4

Tabla 1. *Datos demográficos.*

Las imágenes fueron adquiridas con un equipo PET/CT Siemens Biograph 6 TruePoint (Siemens Healthineers, Erlangen, Alemania). Se inyectó una dosis total media (± desviación típica) de 264,73 ± 180 MBq de radiofármaco de manera intravenosa y el tiempo medio (± desviación típica) entre inyección y adquisición fue de 46 ± 20 minutos. Las imágenes PET fueron corregidas por atenuación mediante la imagen CT. El tamaño de las imágenes reconstruida es de 336×336 vóxeles y 1,01821×1,01821 mm. El grosor de corte es de 3mm.

Se emplea MATLAB R2019b (The MathWorkds, Inc., Natick, Massachussetts, EE. UU.) para la extracción de características y el análisis estadístico. Las imágenes son preprocesadas en Statistical Parametric Mapping 12 (SPM12) (Wellcome Centre for Human Neuroimaging, University College London, Londres, Reino Unido) [7][1]. Para el análisis regional se emplea el atlas *Automatic Anatomical Labelling* (AAL2) que divide el cerebro en 120 regiones corticales y subcorticales [8, 9]. Se usan las regiones que se corresponden a la materia gris del cerebelo como región de referencia debido a que no se suele afectar en pacientes con EA y otras enfermedades neurodegenerativas relacionadas con la APP [10].

2.2. Métodos

La metodología de este trabajo se divide en 3 pasos: (1) Preprocesamiento, (2) cuantificación y (3) análisis estadístico.

Preprocesamiento. Todas las imágenes son preprocesadas mediante SPM12. En primer lugar, se coloca manualmente el origen del espacio nativo de las imágenes en la comisura anterior de manera que esté alineada horizontalmente en el plano sagital con la comisura posterior. A continuación, las imágenes son normalizadas espacialmente al espacio estándar definido por el *Montreal Neurological Institute* (MNI). Las imágenes finales tienen una resolución de 91×109×91 y un tamaño de vóxel de 2×2×2 mm. De esta manera se reducen las variabilidades anatómicas interpacientes y se preparan para la segmentación de las distintas regiones anatómicas.

Cuantificación. Después de la normalización espacial de las imágenes se generan los mapas representando el valor de captación estandarizado (SUV, *standardized uptake value*) en cada vóxel. Para ello, se convierte la concentración de actividad [Bq/ml] del vóxel a SUV [g/ml] en función del peso del paciente usando la ecuación en [11].

Posteriormente, se segmenta el cerebro en las regiones anatómicas definidas por el AAL2. Además, se juntan aquellas correspondientes a la materia gris del cerebelo y se excluyen las regiones del vermis. A continuación, se extraen los valores medios de SUV (SUVmean) de cada región y se normalizan por el SUVmean de la región de referencia para obtener las ratios regionales de SUV (SUVR).

Análisis estadístico. Se realiza un análisis estadístico para encontrar aquellas regiones cerebrales características de cada variante de APP con alteraciones del metabolismo. Para ello, se lleva a cabo primero un análisis intravariante con todas las regiones cerebrales y después un análisis intervariante para cada una de las regiones por separado.

Se estudia la normalidad de las variables con la prueba de Shapiro-Wilk [12]. El análisis multivariante se realiza mediante pruebas ANOVA o Kruskal-Wallis. Se estudian las diferencias entre pares de variantes mediante un análisis *post hoc* con correcciones de Bonferroni. Un valor de $p <$

[1] https://www.fil.ion.ucl.ac.uk/spm/ Último acceso: Julio 2020

0,05 se considera estadísticamente significativo y el análisis estadístico se realiza en MATLAB R2019b.

3. Resultados y discusión

En la Figura 1 se muestran ejemplos de los hipometabolismos característicos (en comparación con el hemisferio opuesto y/o regiones adyacentes, y marcadas con flechas) y enfermedades subyacentes descritos en Gorno-Tempini et al. [3]. Las imágenes mostradas se corresponden a un hombre de 65 años con APPnf (izquierda), un hombre de 68 años con APPvs (centro) y una mujer de 68 años con APPvl (derecha). En el primer caso (APPnf) se puede observar hipometabolismo en la zona frontal-temporal, y con un diagnóstico probable de DFT. La imagen central (APPvs) muestra hipometabolismo temporal y el paciente fue diagnosticada con una probable DFT variante demencia semántica en forma de APP. Por último, en la imagen derecha (APPvl) se observa hipometabolismo parietal y la paciente fue diagnosticada con una demencia leve relacionada con EA.

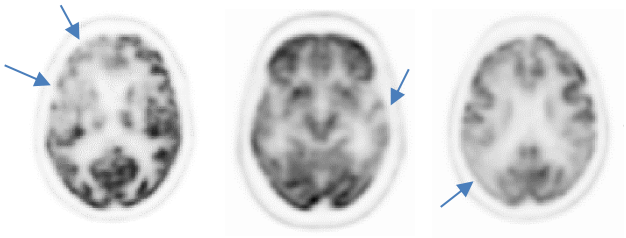


Figura 1. *Ejemplos de imágenes ^{18}F-FDG PET con patrones característicos de las variantes de APP. Las regiones de hipometabolismo están señaladas con flechas. Izquierda: APPnf; Centro: APPvs; Derecha: APPvl.*

Primero, cabe destacar que la muestra en estudio tiene un sólo paciente con APPvs y no es posible hacer un análisis estadístico de una muestra única. Por tanto, se agrupa éste con los pacientes de APPnc en un grupo de otras variantes de APP. Los otros dos grupos empleados en el análisis estadístico se corresponden con los pacientes con APPnf y APPvl. En la Tabla 2 se muestran los resultados del análisis intravariante. Como se puede observar, hay regiones en todas las variantes que presentan un metabolismo que es significativamente distinto. Además, debido a las patologías subyacentes de las variantes se puede asumir que presentan un hipometabolismo.

Grupo	APPnf	APPvl	APP-Otras
Valor de *p*	< 0,05	< 0,05	< 0,05

Tabla 2. *Resultados del análisis intravariante.*

Observando la distribución de los SUVRs regionales para cada variante se pueden identificar ciertas regiones con valores bajos, indicando posibles hipometabolismos. Cabe destacar que tanto en el grupo de APPnf como APPvl los lóbulos parietal y temporal, el hipocampo y giro hipocampal presentan valores moderadamente bajos. Un solapamiento entre variantes también fue detectado por Matías-Guiu et al. [5]. Además, el metabolismo en el lóbulo frontal en el grupo de APPnf es moderadamente menor, como fue descrito por Gorno-Tempini et al. [3]. Todos estos datos concuerdan con aquellos mostrados tanto por Matías-Guiu et al. [5] como por Rabinovici et al. [6]. En cuanto al tercer grupo de otras variantes de APP, se observa una mayor dispersión de los SUVRs, lo cual se debe a la mayor heterogeneidad del grupo, pero generalmente son las mismas regiones que en los otros dos grupos.

Por otro lado, se estudian las diferencias entre los SUVRs de cada variante para cada región por separado. De las 94 regiones estudiadas, se obtienen 15 regiones estadísticamente significativas en el análisis intervariante. Entre estas regiones se pueden destacar:

- Cíngulo posterior derecho
- Giro parahipocampal
- Giro fusiforme
- Giro temporal superior derecho
- Giro temporal medio e inferior
- Polo temporal

A modo de ejemplo, se muestra en la Figura 2 el diagrama de cajas de los SUVRs del giro temporal inferior derecho. Como se puede observar, tanto el grupo de pacientes con APPnf como el grupo de las otras variantes de APP muestra hipometabolismo en esta región.

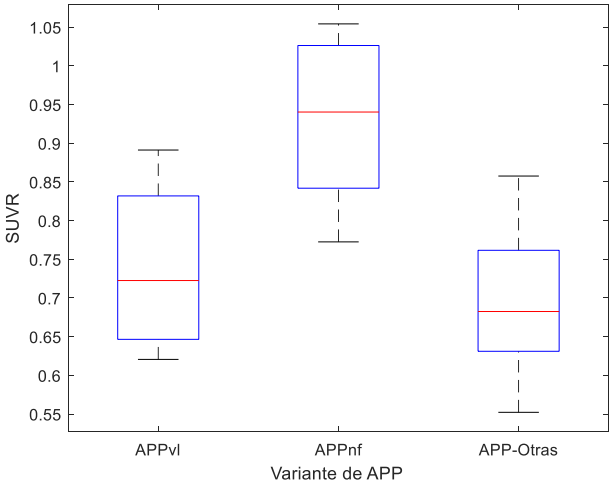


Figura 2. *Diagrama de cajas de SUVRs del giro temporal inferior derecho de los grupos APPvl, APPnf y APP-Otras.*

Estos resultados concuerdan con las observaciones anteriores de los SUVRs de todas las regiones para las tres variantes. Se puede afirmar, por tanto, que el lóbulo temporal es una región de alto interés en el diagnóstico y la clasificación de APPs. Esto también está reflejado en la descripción de las distintas variantes y la literatura científica [3, 5, 6].

El análisis *post hoc* muestra diferencias estadísticamente significativas entre el grupo APPnf y APP-Otras en la mayoría de las regiones anteriormente mencionadas. Sin embargo, no se encuentra diferencias estadísticamente significativas entre el grupo APPvl y los otros dos grupos. Este comportamiento también se puede observar en la Figura 2 que muestra la dispersión de los SUVRs en el giro temporal inferior derecho de los pacientes de los tres grupos.

Además, en todos los casos se observa que la variante APPnf presenta SUVRs en las regiones anteriormente significativas que son más altos que en los demás grupos. Esto se

puede explicar por la presentación típica de hipometabolismo en la región frontal en la APPnf mientras que en la APPvl y APPvs la región parietal y temporal son aquellas con mayor afectación [2, 3]. Los casos de APPnc también presentan mayoritariamente hipometabolismo en la región parieto-temproral.

Con respecto a las limitaciones de este estudio, se debe mencionar en primer lugar que se dispone de una muestra de pacientes reducida. Se cuenta únicamente con 19 pacientes diagnosticados con APP. Además, solo 1 de estos 19 casos se corresponde con un paciente con APPvs, mientras que gran parte (5 de 19) presentan una APPnc. Por lo tanto, no se pueden estudiar los patrones típicos de esta variante con la base de datos disponible.

Por otro lado, se debe tener en cuenta que, a la hora de realizar el preprocesamiento de las imágenes, se trabaja únicamente con las imágenes PET de los pacientes, pues no se cuenta con una imagen de referencia anatómica del cerebro del paciente. Esto puede provocar que los resultados tras la alineación y normalización espacial sean más inexactos que si se contara con dicha imagen como referencia.

Con una muestra mayor y más equilibrada, compuesta especialmente de un mayor número de casos con APPvs, se espera poder refinar los resultados y la detección de regiones significativas en la diferenciación de las variantes. Sin embargo, los resultados obtenidos en este trabajo están acordes con la literatura científica y muestra el potencial de las imágenes de metabolismo cerebral ^{18}F-FDG PET en la diferenciación de las variantes de APP.

4. Conclusiones

La afasia progresiva primaria es un síndrome clínico que se caracteriza por un deterioro del lenguaje y la presencia de enfermedades neurodegenerativas subyacentes. La APP se clasifica en tres variantes con características distintas. En este trabajo se detectan diferencias en el metabolismo cerebral medido por imágenes de ^{18}F-FDG PET entre variantes de APP. Se concluye, por lo tanto, que las imágenes de metabolismo cerebral de ^{18}F-FDG PET pueden ser una ayuda en el diagnóstico y la clasificación de APPs.

Agradecimientos

Este trabajo está parcialmente financiado por una Beca de Formación de Profesorado Universitario del Ministerio de Ciencia, Innovación y Universidades (FPU16/06487).

Referencias

[1] Mesulam MM. Slowly progressive aphasia without generalized dementia. *Annals of Neurology*, vol 11, no 6, 1982, pp 592–598, (ISSN: 15318249).

[2] Mesulam MM, Rogalski EJ, Wieneke C, Hurley RS, Geula C, Bigio EH, Thompson CK, Weintraub S. Primary progressive aphasia and the evolving neurology of the language network. *Nature Reviews Neurology*, vol 10, no 10, 2014, pp 554–569, (ISSN: 17594766).

[3] Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, Ogar JM, Rohrer JD, Black S, Boeve BF, Manes F, Dronkers NF, Vandenberghe R, Rascovsky K, Patterson K, Miller BL, Knopman DS, Hodges JR, Mesulam MM, Grossman M. Classification of primary progressive aphasia and its variants. *Neurology*, vol 76, no 11, 2011, pp 1006–1014, (ISSN: 1526632X).

[4] Matías-Guiu JA, García-Ramos R. Primary progressive aphasia: From syndrome to disease. *Neurología (English Edition)*, vol 28, no 6, 2013, pp 366–374, (ISSN: 21735808).

[5] Matías-Guiu JA, Cabrera-Martín MN, Pérez-Castejón MJ, Moreno-Ramos T, Rodríguez-Rey C, García-Ramos R, Ortega-Candil A, Fernandez-Matarrubia M, Oreja-Guevara C, Matías-Guiu J, Carreras JL. Visual and statistical analysis of 18F-FDG PET in primary progressive aphasia. *European Journal of Nuclear Medicine and Molecular Imaging*, vol 42, no 6, 2015, pp 916–927, (ISSN: 16197089).

[6] Rabinovici GD, Jagust WJ, Furst AJ, Ogar JM, Racine CA, Mormino EC, O'Neil JP, Lal RA, Dronkers NF, Miller BL, Gorno-Tempini ML. Aβ amyloid and glucose metabolism in three variants of primary progressive aphasia. *Annals of Neurology*, vol 64, no 4, 2008, pp 388–401, (ISSN: 03645134).

[7] Friston KJ, Ashburner J, Kiebel S, Nichols T, Penny WD. Statistical parametric mapping: the analysis of funtional brain images. Elsevier/Academic Press, 2007, (ISBN: 9780080466507).

[8] Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, vol 15, no 1, 2002, pp 273–289, (ISSN: 10538119).

[9] Rolls ET, Joliot M, Tzourio-Mazoyer N. Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. *NeuroImage*, vol 122, 2015, pp 1–5, (ISSN: 10959572).

[10] Soonawala D, Amin T, Ebmeier KP, Steele JD, Dougall NJ, Best J, Migneco O, Nobili F, Scheidhauer K. Statistical parametric mapping of 99mTc-HMPAO-SPECT images for the diagnosis of Alzheimer's disease: Normalizing to cerebellar tracer uptake. *NeuroImage*, vol 17, no 3, 2002, pp 1193–1202, (ISSN: 10538119).

[11] Kinahan PE, Fletcher JW. Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy. *Seminars in ultrasound, CT, and MR*, vol 31, no 6, 2010, pp 496–505, (ISSN: 0887-2171).

[12] BenSaïda A. Shapiro-Wilk and Shapiro-Francia normality tests. 2014. https://www.mathworks.com/matlabcentral/fileexchange/13964-shapiro-wilk-and-shapiro-francia-normality-tests (Consultada: Jul 2020).

Sistema para el reconocimiento automático de los cambios en el tamaño de la pupila en pequeños roedores

Alejandro Lara-Doña[1,2], Sonia Torres-Sánchez[1,3,4], Daniel Sánchez-Morillo[1,2] Esther Berrocoso[1,3,4]

[1] Instituto para la Investigación e Innovación Biomédica de Cádiz (INiBICA), Hospital Universitario Puerta del Mar, Cádiz, España.

[2] Grupo de Investigación de Ingeniería Biomédica y Telemedicina. Departamento de Ingeniería Automática, Electrónica y Arquitectura y Redes de Computadores. Universidad de Cádiz, Puerto Real (Cádiz), España. {alejandro.lara, daniel.morillo}@uca.es

[3] CIBER de Salud Mental (CIBERSAM), Instituto de Salud Carlos III, Madrid, España.

[4] Grupo de Investigación de Neuropsicofarmacología y Psicobiología, Departamento de Psicología, Área de Psicobiología, Universidad de Cádiz, Puerto Real (Cádiz), España, {sonia.torres, esther.berrocoso}@uca.es

Resumen

El estudio de la pupila está recibiendo en los últimos años una mayor atención por su utilidad en diversas áreas del conocimiento desde la medicina hasta el marketing. Cada vez existen más soluciones para utilizar esta técnica en humanos, pero para su aplicación en otros animales, como roedores, se dispone de soluciones estandarizadas o comerciales limitadas. Estudios recientes sugieren que el diámetro de la pupila se correlaciona con la activación neuronal del Locus Coeruleus, área cerebral involucrada en la respuesta al dolor. La utilización de esta técnica permite avanzar en el conocimiento de los mecanismos involucrados.

Un porcentaje significativo de los estudios científicos más recientes que hacen uso de esta técnica proporcionan una información incompleta sobre los componentes utilizados o el tratamiento aplicado a las señales registradas para la extracción del diámetro de la pupila. Este trabajo describe los resultados preliminares de una plataforma desarrollada ad-hoc y los métodos asociados de procesamiento para sistematizar los registros de imagen y el reconocimiento automático en tiempo real de los cambios en el tamaño de la pupila en pequeños roedores bajo sedación leve.

1. Introducción

El estudio de los cambios en el diámetro de la pupila se ha extendido en los últimos años a diferentes ámbitos del conocimiento. Mientras que el área del marketing se emplea, por ejemplo, para estudiar dónde centra la atención el usuario de un determinado producto o servicio, en el campo de la medicina se utiliza, por ejemplo, como herramienta de monitorización no invasiva gracias a la cual se han establecido correlatos con el nivel de actividad neuronal de áreas cerebrales específicas [1].

Existen múltiples pupilómetros para ser utilizados en humanos como el presentado por Kassner [2] o Biopac Systems [3]. Ambos se enfocan al seguimiento ocular para el estudio del estado cognitivo del sujeto o a la generación de mapas de calor con las zonas donde el usuario mira con más frecuencia, y por tanto centra su atención visual, aunque también permiten extraer el tamaño de la pupila.

En el caso de pupilómetros para animales, en particular para pequeños roedores, es común encontrar sistemas diseñados *ad-hoc*, implementados para satisfacer las demandas propias de las investigaciones realizadas dentro de cada grupo de investigación [1, 4–8].

El objetivo de este trabajo es el desarrollo de una plataforma para la realización de registros de pupilometría y de los métodos de procesamiento de la imagen para el cálculo automático de la evolución del diámetro de la pupila durante los experimentos con pequeños roedores sometidos a sedación leve. El sistema propuesto permitirá correlacionar la información obtenida con la actividad neuronal registrada en ciertas áreas cerebrales de animales sometidos a determinadas condiciones experimentales.

2. Materiales y métodos

2.1. Plataforma para el registro de datos

A partir de las especificaciones funcionales emanadas de las necesidades experimentales de nuestro grupo de trabajo, se diseñó una plataforma que permite colocar al sujeto del experimento, una cámara de video para la adquisición de los datos y una fuente de luz. En la bibliografía existente no existe un consenso sobre la disposición de estos elementos, su forma o características, por lo que se priorizaron parámetros como la flexibilidad de la plataforma, de forma que se disponga del mayor grado de libertad posible para disponer cómodamente los diferentes elementos del sistema. Esta flexibilidad permitirá definir y posteriormente depurar el método para la toma de datos experimentales.

La Figura 1 ilustra la plataforma desarrollada, consistente en una placa de acero base sobre la que se colocaron dos elementos: a) una plataforma regulable en altura para colocar el roedor sedado para evitar su movimiento; y b) dos cámaras para enfocar los ojos del roedor, acopladas a una estructura que permite colocarlas en cualquier posición dentro de la plataforma. La estructura de sujeción de las cámaras se hace necesaria para permitir la utilización en diferentes tamaños de roedores (ratas y ratones).

ISBN: 978-84-09-25491-0, pags. 455- 458

El soporte base sobre el que se coloca el roedor se diseñó específicamente para el uso deseado. Se fabricó mediante impresión 3D. Se utilizó PETG como material, primero por su biocompatibilidad, y segundo por la posibilidad que ofrece de ser esterilizado tras una prueba empleando desinfectantes utilizados comúnmente en laboratorios de experimentación animal, como la acetona o el alcohol isopropílico.

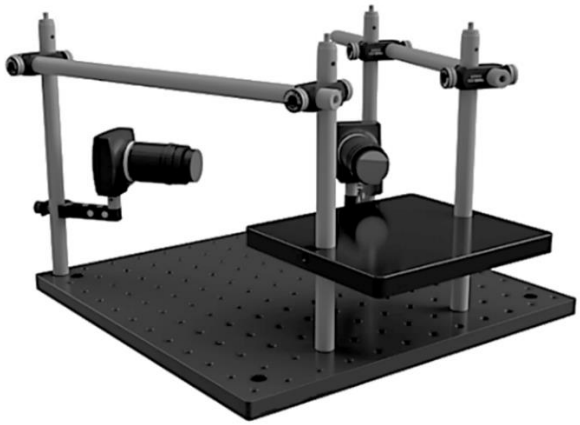


Figura 1. *Plataforma desarrollada para la monitorización de los cambios en la pupila de pequeños roedores.*

Se seleccionó para el registro de las imágenes la cámara Thorlabs modelo DCC1545M. Se trata de una cámara CMOS de 1.3 Megapíxeles monocromo que ha sido empleada en estudios muy recientes [6]. Este modelo permite la captura de imágenes con una tasa de hasta 25 fotogramas por segundo (fps), con una resolución de 1280x1024 pixeles. Se empleó una lente Edmund Optics de 25.0mm de distancia focal, sin filtro infrarrojo y una apertura de f/2.5. Esta lente permite enfocar el ojo del animal a una distancia aproximada de 2 centímetros, de forma que ocupe el campo completo de visión de la cámara. Este posicionamiento permite una mayor resolución en la cuantificación del diámetro de la pupila. Para la realización de los experimentos se utilizó una sola cámara y una iluminación ambiental baja (10 lúmenes en la posición del roedor) para evitar reacciones en la pupila. Para la iluminación del ojo se utilizó una bombilla Philips BR125 IR 50W pues la luz roja no afecta a la conducta de los roedores por no tener receptores para ello [9].

Una vez realizado el montaje de la plataforma, se realizaron diferentes experimentos para definir la metodología y la configuración del software.

2.2. Software de control de la cámara

Para el control de la cámara DCC1545M de Thorlabs se ha utilizado el software uEye (IDS Imaging Development Systems GmbH, Germany) y los drivers proporcionados por el fabricante. uEye permite la selección de los parámetros óptimos de la cámara, y habilita la compatibilidad completa con paquetes de procesamiento matemático ampliamente extendidos como MATLAB o LabVIEW.

Los parámetros de configuración de la cámara se muestran la Tabla 1. La selección de estos parámetros se ha realizado para una comprobación funcional del prototipo en la que se estudiarán dinámicas de cambio del diámetro de la pupila en la respuesta a un estímulo nociceptivo (compresión de las patas traseras, ipsilateral o contralateral al ojo registrado[4]) siendo la respuesta de una duración de 15 segundos). Los parámetros de tiempo de exposición y ganancia son ajustables por el usuario pero se mantienen constantes durante el experimento.

Parámetro	**Valor**
Reloj del pixel	9 MHz
Fotogramas por segundo	5 fps
Tamaño de la imagen	1280x1024px
Formato	Mono 8

Tabla 1. *Configuración del software uEye de IDS*

2.3. Tratamiento de datos

Las pruebas iniciales se realizaron sobre fotogramas individuales con objeto de diseñar algoritmos eficaces extrapolables a una secuencia de video. En una primera etapa de preprocesado, se realizó un tratamiento del contraste de la imagen para aumentar su rango dinámico. Se persigue modificar el histograma de la imagen para asemejarlo a un histograma plano.

A continuación, se aplicó un filtro de desenfoque circular para promediar la intensidad de los pixeles y suavizar las texturas del iris y la córnea. Posteriormente se realizó una nueva normalización del histograma.

Para localizar una primera aproximación de la región de interés se aplicó una segmentación adaptativa local. Esta segmentación adaptativa puede generar huecos por lo que se hace necesario aplicar un proceso posterior de llenado de huecos menores a 200px [10]. Una etapa adicional de filtrado eliminó zonas oscuras no relacionadas con la pupila. Estas zonas oscuras se deben principalmente a residuos generados por el lagrimal y la forma del globo ocular. Para la selección de la región de interés se cribaron todas las zonas cuyo centro se encontraba a una distancia inferior a 100px del centro de la pupila del primer fotograma del video. La región pupilar se corresponde con la zona de mayor área, para la que se calcula el diámetro del circulo equivalente (circulo de área igual a la de la zona seleccionada).

Empleando la metodología anterior, es posible realizar el análisis de la secuencia completa de video almacenando los datos correspondientes a los centros detectados, los ejes de la elipse que delimita la zona correspondiente a la pupila y el área del polígono convexo que contiene la zona seleccionada. Esta última característica permite una mejor aproximación del diámetro de la pupila en el caso de que existan artefactos en la segmentación.

Para comprobar la precisión del método propuesto se analizaron registros de 5 animales, submuestreados a 1fps. La precisión del algoritmo queda definida como el número de fotogramas en los que se ha detectado correctamente la pupila y su contorno frente al número total de fotogramas tratados. El procesamiento de la imagen se realizó empleando el software MATLAB (Mathworks Inc.).

3. Resultados

La Figura 2 ilustra un fotograma correspondiente a uno de los registros realizados.

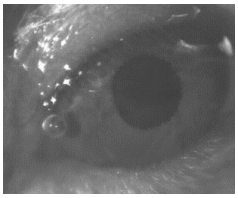


Figura 2. *Imagen registrada del ojo de un ratón utilizando la plataforma diseñada y la cámara y lente seleccionadas.*

Como se puede observar en la Figura 3, el preprocesamiento descrito mejora la diferenciación de la pupila del contorno del iris lo que supone una mejora de los resultados en la segmentación posterior.

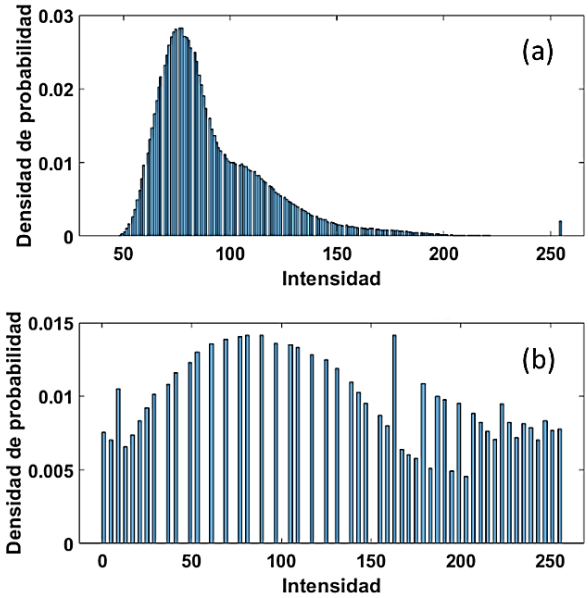


Figura 3. *Mejora del contraste mediante la ecualización del histograma. (a) Histograma de: (a) imagen original; (b) imagen con histograma ecualizado.*

La Figura 4 ilustra de forma cualitativa los resultados obtenidos tras las diferentes etapa de filtrado.

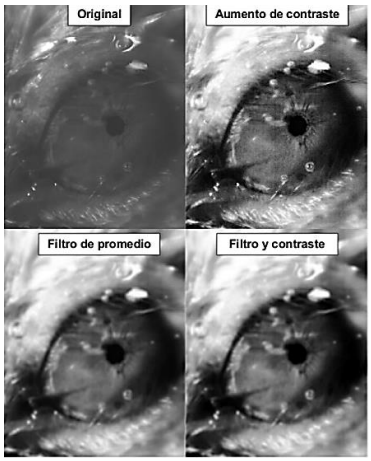


Figura 4. *Detalle de los resultados obtenidos en las diferentes etapas de procesamiento.*

La Figura 5 ilustra el procedimiento de selección de la región de interés tras el cribado de áreas espurias en el primer fotograma. Se aprecia marcada la zona de mayor área, para la que se calcula el diámetro del circulo equivalente.

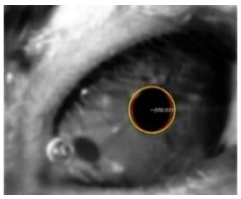


Figura 5. *Zona de interés para la localización del centro de la pupila.*

La Figura 6 ilustra un artefacto aparecido tras el proceso de segmentación, causado por una vibrisa del animal.

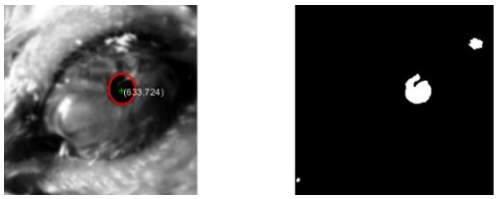


Figura 6. *Artefacto en la detección de la zona correspondiente a la pupila por una vibrisa.*

La Figura 7 muestra el desplazamiento de la posición del centro de la pupila durante el experimento. Este desplazamiento está causado por los movimientos respiratorios del animal. Estos movimientos pueden provocar un desenfoque en la imagen induciendo a errores en la detección de la región correspondiente a la pupila.

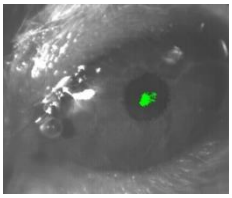


Figura 7. *Artefacto en la detección de la zona correspondiente a la pupila por una vibrisa.*

En la Figura 8 se representan los resultados obtenidos a partir de un registro realizado con el equipamiento y la metodología descrita. Se aprecia el aumento gradual en el tamaño de la pupila tras la aplicación de un estímulo nociceptivo (t=215s y t=800s). Como se aprecia en la Figura 8, la utilización del área que queda definida por el polígono convexo (trazo azul) proporciona un diámetro mayor que el resultante de emplear una aproximación elíptica de la pupila (trazo negro) y superior al diámetro obtenido tras aplicar el criterio de cálculo del área resultante tras proceso de segmentación (trazo verde).

Finalmente, la Tabla 2 sintetiza los resultados obtenidos a partir de los registros realizados en 5 animales, con un total de 10.066 fotogramas analizados. La precisión media alcanzada fue del 82.1%.

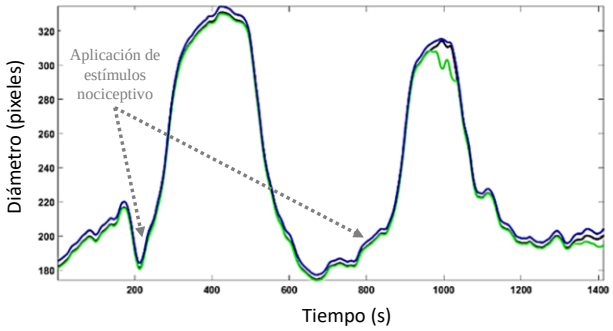


Figura 8. *Ilustración de la evolución temporal del diámetro de la pupila en uno de los registros realizados. Trazo azul: Trazo verde: Área resultante del proceso de segmentación. Área definida por el polígono convexo circunscrito. Trazo negro: Aproximación elíptica.*

4. Discusión y Conclusiones

En este trabajo se ha presentado una plataforma y un método preliminar para el cálculo automático y el seguimiento temporal de los cambios en el diámetro de la pupila de pequeños animales anestesiados y sometidos a estímulos sensoriales.

Experimento	FC	TF	Precisión (%)
1	1414	1508	93,77
2	2126	2359	90,12
3	1393	2339	59,56
4	1856	2193	84,63
5	1372	1667	82,30

Tabla 2. *Precisión alcanzada por el sistema y método propuesto. FC: Fotogramas con resultado correcto. TF: Número total de fotogramas analizados.*

Aunque los resultados son prometedores, se han identificado áreas de mejora como la iluminación del ojo (se contempla incluir iluminación LED alrededor del objetivo), la mejora en el tratamiento de artefactos que pueden aparecer en la zona de interés como las vibrisas del ratón u opacidades causadas por el gel que se aplica sobre los ojos de los sujetos del experimento para evitar daños en la córnea. Una mejora en el algoritmo que permita unir las áreas cercanas al centro de la pupila o una mejora en la iluminación para disminuir las sombras provocadas por el globo ocular podría reducir errores y mejorar el rendimiento. Los errores producidos en el tratamiento de los experimentos 3, 4 y 5 fueron debidos a la opacidad aparecida en la pupila al final del experimento y causada por el gel lubricante.

El modelo de cámara utilizado tiene una conexión USB 2.0, limitando el ancho de banda a 38MB/s. En la práctica, la tasa de muestreo está limitada a 22.11 fps, lo que genera una pérdida de fotogramas. Esta situación supone una pérdida de la coherencia temporal en la prueba. Para solventar esta situación, se trabaja en el registro de imágenes mediante MATLAB o LabVIEW, asignando un código temporal a cada fotograma. Se espera alcanzar una tasa de 10 fps para permitir detectar cambios más rápidos en el diámetro de la pupila [8].

Una vez optimizada la metodología, se procederá con la extensión de las pruebas a un grupo significativo de animales con objeto de obtener resultados estadísticamente relevantes y a la realización de registros EEG simultáneos con el objetivo de correlacionar la respuesta pupilar con la actividad en determinadas áreas cerebrales. Se estudiarán además los tiempos de procesamiento a nivel de fotograma y de registros completos.

Financiación

Este estudio ha sido financiado a través de la convocatoria 2018 de proyectos de I+D+i «Retos Investigación», del programa estatal de I+D+i orientada a los retos de la sociedad, en el marco del plan estatal de investigación científica y técnica y de innovación 2017-2020, proyecto PAINAX (RTI2018-099778-B-I00).

Referencias

[1] Joshi S, Li Y, Kalwani RM, Gold JI. 2016. Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. Neuron 89(1):221–234.

[2] Kassner M, Patera W, Bulling A. 2014. Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction. In: UbiComp 2014 - Adjun. Proc. 2014 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput. pp 1151–1160.

[3] Biopac Systems Inc. Eye Tracking System for Binoc | BIOPAC. Available from: https://www.biopac.com/product/eye-tracking-systems/, [10/09/2020].

[4] Liu Y, Rodenkirch C, Moskowitz N, Schriver B, Wang Q. 2017. Dynamic Lateralization of Pupil Dilation Evoked by Locus Coeruleus Activation Results from Sympathetic, Not Parasympathetic, Contributions. Cell Rep 20(13):3099–3112.

[5] Yüzgeç Ö, Prsa M, Zimmermann R, Huber D. 2018. Pupil Size Coupling to Cortical States Protects the Stability of Deep Sleep via Parasympathetic Modulation. Curr Biol 28(3):392-400.e3.

[6] Breton-Provencher V, Sur M. 2019. Active control of arousal by a locus coeruleus GABAergic circuit. Nat Neurosci 22(2):218–228.

[7] Zerbi V, Floriou-Servou A, Markicevic M, Vermeiren Y, Sturman O, Privitera M, von Ziegler L, Ferrari KD, Weber B, De Deyn PP, Wenderoth N, Bohacek J. 2019. Rapid Reconfiguration of the Functional Connectome after Chemogenetic Locus Coeruleus Activation. Neuron 103(4):702-718.e5.

[8] Reimer J, McGinley MJ, Liu Y, Rodenkirch C, Wang Q, McCormick DA, Tolias AS. 2016. Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nat Commun 7(1):1–7.

[9] Peirson SN, Brown LA, Pothecary CA, Benson LA, Fisk AS. 2018. Light and the laboratory mouse. J Neurosci Methods 300:26–36.

[10] Bradley D, Roth G. 2007. Adaptive Thresholding using the Integral Image. J Graph Tools 12(2):13–21.

Método de calibración geométrica para tomosíntesis de tórax

A. Galán González[1,2], A. Del Valle Hidalgo[1,2], J. García Blas[3], M. Desco Menéndez[1,2,4,5], M. Abella García[1,2,4]

[1] Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, Madrid, España

[2] Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España

[3] Departamento de Informática, Universidad Carlos III de Madrid, Madrid, España

[4] Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, España

[5] Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid, España

Resumen

La tomosíntesis de tórax permite un diagnóstico más preciso de las alteraciones de las estructuras torácicas que la radiografía, manteniendo una dosis de radiación inferior a la involucrada en Tomografía Axial Computerizada (TAC). Al igual que en TAC, la calibración geométrica del sistema es fundamental en tomosíntesis para evitar artefactos en la imagen reconstruida. Sin embargo, la literatura sobre métodos de calibración para tomosíntesis es mucho menor que para TAC. Este trabajo propone un método de calibración geométrica para tomosíntesis a partir de la adaptación de un método previamente desarrollado para TAC. La validación en simulación de un estudio de tórax muestra que el método es capaz de reducir los artefactos de desalineamiento en la imagen reconstruida.

1. Introducción

La tomosíntesis de tórax ha demostrado tener utilidad clínica, especialmente en la detección de cáncer de pulmón [1, 2]. En esta modalidad, la fuente de rayos X se mueve en un eje lineal y el detector se mantiene fijo, obteniéndose una serie de proyecciones que permiten reconstruir una imagen pseudo-tomográfica. Al igual que en Tomografía Axial Computerizada con geometría cónica (CBCT), la calibración geométrica del sistema es fundamental para evitar artefactos en la imagen reconstruida. Existen numerosos métodos de calibración para CBCT en la literatura. Noo et al. [3] desarrollaron un método analítico basado en un maniquí sencillo con dos marcadores esféricos. Cho et al. [4] mejoraron este método incorporando la calibración de cada proyección individualmente, sin asumir una trayectoria fuente-detector perfectamente circular. Este último método es teóricamente extrapolable a otras geometrías [4] y ha sido aplicado en tomosíntesis de mama [5].

Los métodos propuestos para la calibración geométrica en tomosíntesis se basan en el cálculo de la matriz de proyección, que relaciona las coordenadas espaciales de los marcadores de un maniquí con sus proyecciones en la imagen [6-10]. Dichos métodos requieren conocer con exactitud la posición espacial de los marcadores, siendo esta una posible fuente de error [11]. Además, el maniquí debe incluir un número alto de marcadores para compensar posibles inexactitudes [6]. Frente a estos métodos, el método de Cho et al., diseñado para sistemas CBCT, no requiere conocer las coordenadas de los marcadores en el espacio, lo que lo hace deseable por su facilidad de implementación.

En este trabajo se propone un método de calibración geométrica para tomosíntesis de tórax basado en el método de Cho et al. y se comprueba su validez en simulaciones.

2. Método

El método de Cho et al. [4] emplea un maniquí cilíndrico hueco con una serie de marcadores distribuidos formando dos círculos. Las coordenadas de sus centros en las proyecciones se ajustan a elipses que se utilizan para calcular los parámetros de calibración geométrica de forma analítica. En la implementación del método, dichas coordenadas se han obtenido con una función basada en la transformada circular de Hough.

2.1. Traducción de la geometría de CBCT a la de tomosíntesis de tórax

En tomosíntesis de tórax, la fuente se mueve siguiendo una trayectoria lineal, mientras el detector se mantiene fijo (Figura 1).

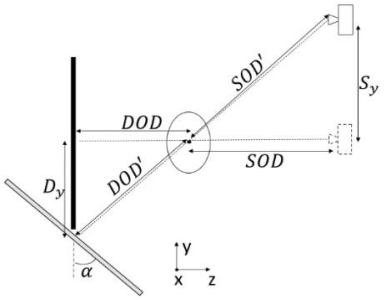


Figura 1.*Representación esquemática de la posición del detector en tomosíntesis (en negro) y su posición de partida equivalente en CBCT (en gris).*

Exceptuando la proyección central (sin desplazamiento de la fuente), el rayo central no es perpendicular al detector. Estas diferencias geométricas respecto a CBCT deben ser consideradas para traducir los parámetros geométricos obtenidos con las fórmulas descritas en [4].

ISBN: 978-84-09-25491-0, pags. 459- 462

La posición del detector en tomosíntesis de tórax puede expresarse como una posición del detector en modo CBCT ideal más un desplazamiento en el eje vertical (S_y) y una inclinación del detector alrededor del eje horizontal (α), como se muestra en la Figura 1. La inclinación del detector se obtiene con la posición angular fuente-detector en modo CBCT como:

$$\alpha = \text{atan}\left(\frac{S_y + D_y}{SOD+DOD}\right) \quad (1)$$

donde *SOD* y *DOD* son las distancias fuente-objeto y objeto-detector, respectivamente, en la proyección central (sin desplazamiento de la fuente), S_y es el desplazamiento de la fuente y D_y la posición del detector en CBCT, que coincide con el desplazamiento vertical que hay que incorporar, y que se calcula:

$$D_y = S_y \times \frac{DOD}{SOD} \quad (2)$$

SOD y *DOD* de las proyecciones distintas a la central se calculan en (3) y (4) respectivamente:

$$SOD' = \sqrt{SOD^2 + S_y^2} \quad (3)$$

$$DOD' = \sqrt{DOD^2 + D_y^2} \quad (4)$$

El método de calibración propuesto obtiene los desalineamientos del detector reflejados en la Figura 2: desplazamiento horizontal y vertical del detector (*Uoffset* y *Voffset*, respectivamente) y los ángulos (*Zeta*, *Fi* y *Eta*) alrededor de los tres ejes de coordenadas. Junto a estos, también se calcula la posición de la fuente, así como *SOD* y *DOD* de cada proyección.

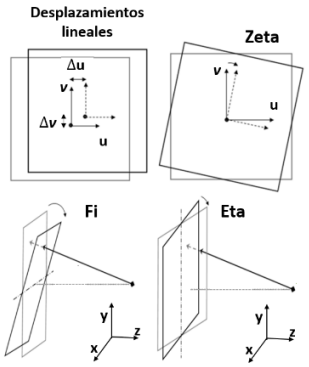


Figura 2.*Parámetros de calibración geométrica utilizados, correspondientes a los posibles desalineamientos del detector.*

El método propuesto de calibración geométrica incluye tres pasos: (1) se obtienen los parámetros de calibración geométrica para CBCT siguiendo las ecuaciones descritas en [4]; (2) se calculan D_y y α para cada proyección usando las fórmulas anteriores y (3) se restan los valores de D_y y α a los obtenidos en el primer paso para *Voffset* y *Fi* respectivamente.

2.2. Características y colocación del maniquí

El maniquí de calibración es un cilindro hueco de 150 mm de diámetro, con 24 bolas de 1.5 mm de diámetro situadas formando dos círculos de 12 bolas equidistantes separados 170 mm entre sí (Figura 3). Las bolas son de un material mucho más denso (101 g/cm^3) que el cilindro (1 g/cm^3) para que el contraste sea alto. Este maniquí de calibración se ha generado con un tamaño de vóxel de 0.14 mm y tamaño de matriz de 1250 en las tres dimensiones.

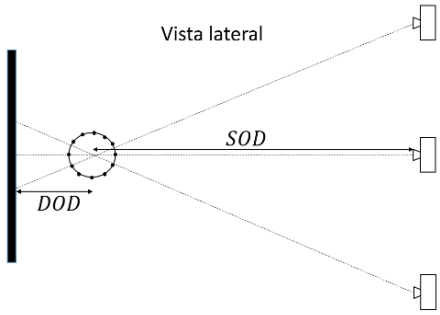


Figura 3.*Geometría de adquisición del maniquí de calibración.*

Para comprobar la posición óptima del maniquí, se han simulado 31 proyecciones con una posición angular del par fuente-detector de 0 a 15 grados, con un paso de 0.5 grados, variando la distancia fuente-objeto y manteniendo la distancia fuente-detector en 1800 mm. La Tabla 1 muestra que el error de calibración, medido como el error cuadrático medio (RMSE) en las 31 proyecciones, es menor según aumenta la *DOD*. Al incorporar los desalineamientos del detector (ver siguiente sección) no es posible usar la *DOD* de 346 mm debido a la limitación del campo de visión Por este motivo, se ha usado la *DOD* de 270 mm, de forma que todos los marcadores aparezcan en las proyecciones.

	RMSE (mm)		RMSE (grados)		
DOD (mm)	**Uoffset**	**Voffset**	**Eta**	**Zeta**	**Fi**
175	0.09	0.23	5.3 e-3	0.84	9.0e-2
270	0.06	0.23	5.1 e-3	0.47	9.0e-2
346	0.04	0.22	3.8 e-3	0.35	6.8 e-2

Tabla 1. *Resultados de la calibración para cada una de las distancias evaluadas.*

Por otra parte, se ha examinado la dependencia del error con la altura del maniquí, su diámetro y el diámetro de los marcadores. Para ello se han simulado 21 proyecciones del maniquí descrito, variando sólo una de las características anteriores en cada caso y se han calculado los valores de RMSE. Se ha observado que el incremento de la altura del maniquí reduce considerablemente los errores. El incremento del diámetro de las bolas también mejora ligeramente los resultados, mientras que la influencia del diámetro del maniquí es despreciable.

2.3. Evaluación

Para simular las proyecciones y realizar las reconstrucciones se ha usado FuxSim [12], un software que permite simular diferentes geometrías de adquisición a partir de un mapa de densidad y unos parámetros del sistema.

Para realizar las pruebas hemos usado el TAC de tórax del maniquí antropomórfico PBU-60 (KYOTO KAGAKU, Kyoto, JP), con un tamaño de matriz de 512×259×600 y un tamaño de vóxel de 0.843×0.843×0.8 mm. Se ha introducido en dicho volumen un nódulo esférico de 6 mm de diámetro (densidad = 1.6 g/cm^3) para analizar el efecto de los desalineamientos en la imagen reconstruida.

La simulación se ha basado en un protocolo estándar de tomosíntesis de tórax: 61 proyecciones con una distancia total *SOD* + *DOD* de 1800 mm y un ángulo de barrido de 30 grados. Para las simulaciones se ha usado *DOD* = 109.17 mm. Las proyecciones tienen un tamaño de 2144×2144 píxeles con un tamaño de píxel de 0.2×0.2 mm. La Tabla 2 muestra los desalineamientos evaluados.

Uoffset (mm)	Voffset (mm)	Eta (grados)	Zeta (grados)	Fi (grados)
5	5	5	0	5

Tabla 2. *Desalineamientos del detector para las proyecciones simuladas.*

Se han reconstruido usando FDK tres volúmenes del mismo tamaño del CT original: uno sin calibración, otro incorporando la calibración obtenida con el método propuesto y un tercero incorporando la calibración ideal descrita en la Tabla 2. Se han trazado los perfiles horizontales y verticales sobre cortes coronales que se muestran en la Figura 4. Asimismo, se han trazado perfiles a lo largo del nódulo de 6 mm en las dos direcciones.

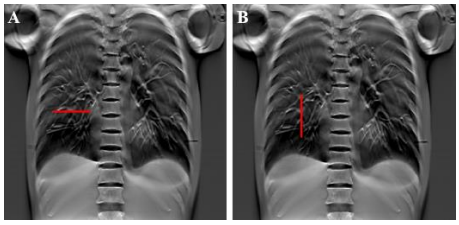


Figura 4. *Corte coronal de la reconstrucción con calibración ideal, mostrando los perfiles horizontal y vertical.*

Para cada parámetro de calibración se ha calculado la media de todas las proyecciones, salvo de *SOD* y *DOD* ya que varían según el ángulo de proyección, así como la desviación media absoluta (MAD).

3. Resultados

La Tabla 3 muestra los resultados de la calibración. Los valores de desviación media absoluta en *SOD* y *DOD* son muy pequeños en términos relativos (0.15% y 0.18% con respecto a los valores mínimos) y se ha comprobado que su efecto en las reconstrucciones es despreciable.

Parámetro	Simulado	Calibrado (Media/MAD)
SOD (mm)	Variable	Variable/2.29
DOD (mm)	Variable	Variable/0.48
Uoffset (mm)	5.00	5.04/0.06
Voffset (mm)	5.00	4.80/0.07
Eta (grados)	5.00	5.00/0.01
Zeta (grados)	0.00	0.27/0.64
Fi (grados)	5.00	5.03/0.13

Tabla 3. *Resultado de la calibración geométrica.*

La Figura 5 muestra los cortes coronales de los tres volúmenes, donde se aprecia una mejora en la definición de la imagen con calibración con respecto a la imagen sin calibrar.

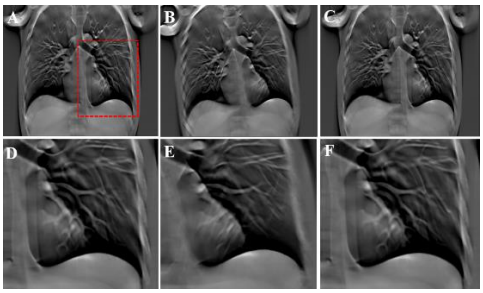


Figura 5.*Cortes coronales de los volúmenes reconstruidos con calibración ideal (A y D), sin calibrar (B y E), y con el método de calibración propuesto (C y F). Delineada en rojo en A la región amplificada mostrada en el panel inferior.*

La Figura 6 muestra los perfiles de intensidad, mostrando una similitud de los perfiles en la imagen calibrada y en la de calibración ideal.

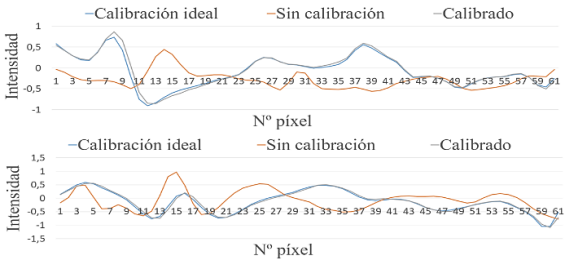


Figura 6. *Perfiles de intensidad en dirección horizontal (arriba) y vertical (abajo).*

Los cortes coronales en la rodaja correspondiente al centro del nódulo para los tres volúmenes se muestran en la Figura 7, mostrando la recuperación del nódulo con el método propuesto.

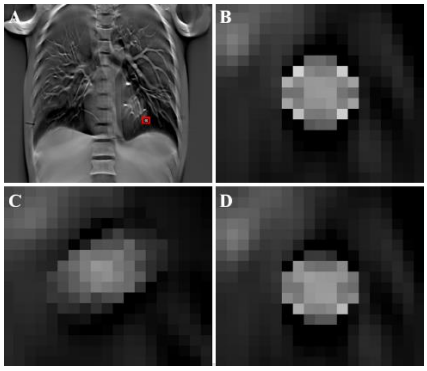


Figura 7.*Corte coronal del volumen completo reconstruido con calibración ideal (A, nódulo en rojo). Ampliación del nódulo en la reconstrucción con calibración ideal (B), sin calibración (C) y con el método de calibración propuesto (D).*

La Figura 8 muestra los perfiles a lo largo del nódulo, con valores en la imagen con calibración muy cercanos a los de la imagen con calibración ideal en la dirección horizontal. En la dirección vertical son también similares, aunque se observa un ligero suavizado del perfil en la imagen calibrada.

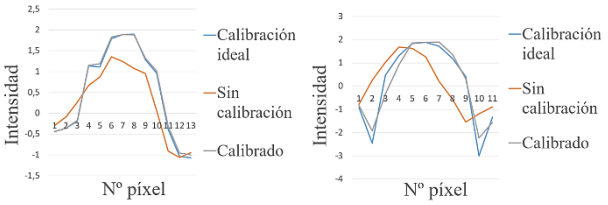


Figura 8.*Perfiles de intensidad del nódulo en dirección horizontal (izquierda) y vertical (derecha).*

4. Discusión y conclusiones

Este trabajo propone un método de calibración con maniquí para obtener los parámetros geométricos de un sistema de tomosíntesis de tórax basado en la adaptación del método de Cho et al. [4], diseñado para CBCT. Una dificultad para la calibración en tomosíntesis es la limitación en la distancia objeto detector y en consecuencia de la magnificación, ya que por las limitaciones del campo de visión el objeto debe estar cerca del detector. Los resultados muestran que, al aumentar la DOD, los errores obtenidos se reducen, por lo que es conveniente que ésta sea lo mayor posible. Por otra parte, se ha observado que los errores son parecidos para todas las posiciones de fuente. En cuanto a las dimensiones del maniquí, se ha observado que los errores se reducen cuando aumenta la altura del maniquí y el diámetro de las bolas, mientras que el diámetro del maniquí no tiene apenas influencia. Los resultados muestran un ligero suavizado en el perfil en dirección vertical de la imagen calibrada con el método propuesto respecto al de la imagen con calibración ideal. Esto se debe a que el filtro rampa para las reconstrucciones con FDK se ha aplicado en dirección vertical, por lo que las pequeñas diferencias de contraste entre ambas imágenes, debidas al artefacto de emborronamiento provocado por los desalineamientos del detector, se ven amplificadas en esta dirección. La severidad de dicho artefacto es proporcional a la magnitud de los errores de calibración geométrica, los cuales podrían reducirse empleando un método más preciso para la detección de los centros de los marcadores en las proyecciones, así como optimizando la geometría del maniquí de calibración. Hay que tener en cuenta que la magnitud de los desalineamientos ha sido exagerada para que se apreciara claramente su efecto. En el futuro se explorará en detalle la tolerancia del sistema a los distintos desalineamientos.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia, Innovación y Universidades (Instituto de Salud Carlos III, proyecto DTS17/00122; Agencia Estatal de Investigación, proyecto DPI2016-79075-R-AEI/FEDER, UE), cofinanciado por Fondos de la Unión Europea (FEDER), "A way of making Europe". Además, ha sido financiado por el Programa de apoyo a la realización de proyectos interdisciplinares de I+D para jóvenes investigadores de la Universidad Carlos III de Madrid 2019-2020 en el marco del Convenio Plurianual Comunidad de Madrid- Universidad Carlos III de Madrid (proyecto DEEPCT-CM-UC3M). El CNIC está financiado por el Ministerio de Ciencia, Innovación y Universidades y la fundación PRO-CNIC y es un centro de excelencia Severo Ochoa (SEV-2015-0505).

Referencias

[1] Dobbins JT, McAdams HP, Sabol JM. Multi-institutional evaluation of Digital Tomosynthesis, Dual-energy radiography, and conventional chest radiography for the Detection and Management of Pulmonary nodules. *Radiology*, vol 282, sup 1, 2017.

[2] Söderman C, Johnsson ÅA, Vikgren J. Detection of Pulmonary Nodule Growth with Chest Tomosynthesis: A Human Observer Study Using Simulated Nodules. *Academic Radiology*, vol 26, sup 4, 2019, pp 508-18.

[3] Noo F, Clackdoyle R, Mennessier C. Analytic method based on identification of ellipse parameters for scanner calibration in cone-beam tomograph. *Physics in Medicine and Biology*, vol 45, 2000, pp 3489–508.

[4] Cho Y, Moseley DJ, Siewerdsen JH. Accurate technique for complete geometric calibration of cone-beam computed tomography systems. *Medical physics*, vol 32, sup 4, 2005, pp 968-83.

[5] Wang X, Mainprize JG, Kempston MP. Digital breast tomosynthesis geometry calibration. *Physics of Medical Imaging*, vol 6510, 2007.

[6] Li X, Da Z, Liu B. A generic geometric calibration method for tomographic imaging systems with flat-panel detectors--a detailed implementation guide. *Medical physics*, vol 37, sup 7, 2010, pp 3844-54.

[7] Chang C-H, Ni Y-C, Huang S-Y. A geometric calibration method for the digital chest tomosynthesis with dual-axis scanning geometry. *PLoS ONE*, vol 14, sup 4, 2019.

[8] Miao H, Wu X, Zhao H. A phantom-based calibration method for digital x-ray tomosynthesis. *Journal of X-Ray Science and Technology*, vol 20, 2012, pp 17-29.

[9] Nett BE, Zambelli J, Riddell C. Circular tomosynthesis implemented with a clinical interventional flat-panel based C-arm: Initial performance study. *Physics of Medical Imaging*, vol 6510, 2007.

[10] Ferris WS, Vent TL, Maidment TD. Geometric Calibration for a Next-Generation Digital Breast Tomosynthesis System. *Physics of Medical Imaging*, vol 10132, 2017.

[11] Li X, Zhang D, Liua B. Sensitivity analysis of a geometric calibration method using projection matrices for digital tomosynthesis systems. *Medical physics*, vol 38, sup 1, 2011.

[12] Abella M, Serrano E, García Blas J. FUX-Sim: Implementation of a fast universal simulation/reconstruction framework for Xray systems. *PLoS ONE*, vol 12, sup 7, 2017.

Sistemas de Ayuda a la Decisión II

Viernes, 27 de noviembre de 2020

Feature Selection and Tree-based Models to Predict Multidrug-Resistance

Ó. Escudero Arnanz[1], I. Mora Jiménez[1], S. Martínez Agüero[1], J. Álvarez Rodríguez[2], C. Soguero Ruiz[1]

[1]Department of Signal Theory and Communications, Telematics and Computing Systems, Rey Juan Carlos University, Fuenlabrada, Madrid, Spain

[2]Intensive Care Department, University Hospital of Fuenlabrada, Madrid, Spain.

Abstract

Antimicrobial resistance is one of the greatest threats to the health system. Owing to the critical health condition of patients admitted in the Intensive Care Unit (ICU), time is critical to identify bacteria and their resistance to antibiotics. In this paper, the acquisition of the first multidrug-resistance (MDR) has been temporarily characterized by considering data from the University Hospital of Fuenlabrada for 16 years. We have considered different feature selection strategies and data-driven methods inspired in trees, to predict whether a patient will acquire the first MDR or not in the next 24 hours. The obtained results are quite promising, providing a 97% success rate for non-MDR and 83% for MDR cases. They can be considered as a first step to help the clinician in decision making, for example, to isolate the potential MDR patient to avoid the spread of multi-resistant strains in the ICU.

1. Introduction

The multidrug-resistance (MDR) is the phenomenon that occurs when a given germ develops acquired resistance to many families of antimicrobials to which it was previously sensitive [1]. Tackling the MDR germ becomes a very complicated process, as the number of treatments that can be applied by clinical staff is reduced, which can even make any antibiotic treatment ineffective. The absence of new treatments for MDR strains will mean that minor illnesses and injuries that are currently easy to treat will require more complex treatments [2].

Indeed, antimicrobial resistance is considered one of the greatest challenges to global health in the coming decades. This problem causes approximately 25,000 deaths and 4 million severe infections per year in Europe [3]. The predictions for 2050 about the impact of the MDR provide alarming data, with approximately 10 million deaths [4] and a loss of more than 100 trillion US dollars, equivalent to 7 per cent of the world Gross Domestic Product [5], causing an economic crisis with an impact similar to that of 2008 [6].

The Intensive Care Unit (ICU) is considered the epicentre of the MDR development due to the high rate of patients who are admitted in this unit and acquire infections (20-30 % of all ICU admissions [7]). During the ICU stay, an excessive use of antimicrobials is produced. In particular, some studies [7] reveal that more than 60 % of patients take antibiotics during their stay in the ICU.

The antibiogram, also know as antimicrobial susceptibility test, is used by the ICU staff to determine the sensitivity of germs to certain antibiotics. This test takes approximately 48 hours to give the results.

This paper proposes a procedure for a temporal characterization of the first MDR for a patient in the ICU. Towards that end, different methods for feature selection (FS) and machine learning (ML) models to predict the result of the antibiogram have been considered. Our approach focuses on predicting the emergence of multidrug-resistant bacteria for any ICU patient by modelling it as a classification task (MDR patients and non-MDR patients). For this purpose, we have used data from the electronic medical record of the University Hospital of Fuenlabrada (UHF) in Madrid (Spain). Anonymous data records were provided by means of a patient identifier (sequence of random numbers) generated by UHF staff, making it impossible to identify any patient.

The structure of this work is as follows. In Section 2 we describe the methodology. Data are presented in Section 3. Results when applying different ML approaches are provided in Section 4. The conclusions and future work are shown in Section 5.

2. Methodology

In this section we give an overview of the FS strategies and the ML models considered to predict the antibiogram results.

2.1. Feature Selection

The objective of FS is to select a subset of features that can improve the performance of ML model predictions and reduce the number of features, providing in most applications a better generalisation [8].

The methods for FS can be classified as filter, wrapper or embedded [9]. The filter methods perform the FS process independently of the model construction. In the wrapper methods, a subset of features is evaluated based on the model performance in the training phase. The embedded methods include the FS process as part of the model training process. In this paper, we focus on filter and embedded method, which are listed below:

- Confidence Interval (CI). We use the Bootstrap resampling method [10] to estimate the CI of the mean difference, for each feature, between MDR and non-MDR patients. Bootstrapping is useful when the actual distribution function is unknown since no assumptions are made about the distribution.

ISBN: 978-84-09-25491-0, pags. 464- 467

- Mutual Information (MI). This method is based on entropy, a key measure of information, because of its ability to quantify the uncertainty of random features and the amount of information shared by them [11]. The MI is calculated between each pair of available features.
- Logistic regression with least absolute shrinkage and selection operator (LASSO). To obtain the coefficients associated with each feature, a cost function with a Lasso regularization term has been optimized by Stochastic Gradient Descent (SGD) [12]. The use of a regularization term favours that some coefficients tend to 0, which allows us to determine irrelevant features.

2.2. Machine learning models for classification

ML techniques consider a set of *N* samples (also called examples or observations) for learning a model, where the *i*-th sample is defined through a set of *D* explanatory variables considering different times slots. Therefore, ML techniques allows us to address a task by learning from experience. The most appropriate learning scheme depends on the data and the task to be tackled. In general, there are two types of ML strategies [13]: supervised learning, where the examples used for learning the model have been previously labelled, and unsupervised learning, where no target is considered for learning.

In this paper, given that the target variable *y* is known in advance for a set of patients (historical medical records), an approach based on supervised learning has been followed. In particular, the target variable has been coded as a binary variable: '1' for MDR, and '0' for non-MDR. In this setting, we have addressed the objective as a classification task, based on predicting whether the antibiogram result will be positive ('1') or negative ('0'). The input data for learning the model can be represented as a $N \times D$ matrix named **X**, where $X_{i,j}$ represent the *j*-th variable of the *i*-th patient. To predict the patients' class, different supervised ML schemes have been applied: Logistic Regression (LR), Decision Trees (DT) and Random Forest (RF).

- *Logistic Regression:* LR is a parametric approach where the model is obtained as a linear combination of different features [14]. Despite its name, it is a classification approach since the result of the linear combination is the input to a logistic function. The Equation 1 presents the regularized cost function that has been optimized to find the weights **w** and *b* of the linear model. The first sum of this equation is weighted by the penalty coefficient *C*, which value is chosen by exploring a range of values through cross-validation.

$$C \sum_{i=1}^{N} \log(\exp(-y_i(\boldsymbol{x}_i \boldsymbol{w}^T + b) + 1)) + \sum_{j=1}^{D} \boldsymbol{w}_j^2 \qquad (1)$$

- *2) Decision Tree:* DT is a non-linear and non-parametric supervised learning model organized under a hierarchical structure [15]. The classification is obtained by following the conditions that are fulfilled from the root node to one of its leaves. In this paper, we have used the Gini index as a criterion for the branch division [16]. For generalization purposes, cross-validation has been applied to select the most appropriate value of two hyperparameters: maximum tree depth and minimum number of samples per leave.
- *3) Random Forest:* RF is a popular technique introduced by Breiman in 2001 [17], based on the implementation of various decision trees. RF implements a meta estimator, i.e., an estimator resulting from the combination of different decision trees unrelated to each other, built using different subsets from the original set and choosing a subset of explanatory variables at random in each node. Therefore, some samples can be recurrently used in training while others may never be used. This makes the classifier more robust to slight variations in input observations and, at the same time, increases the classification performance.

2.3. Machine learning models for classification

Most classification tasks in the medical context do not have a similar number of samples to represent each class. Thus, it is usual that the number of ill patients is lower than the number of healthy patients. In this work, the number of MDR patients is lower than the number of non-MDR patients. This problem is known as learning with imbalanced classes.

In this paper, different strategies for dealing with imbalanced classes have been considered [18]: (1) an undersampling approach to reduce the number of samples of the majority class before training the model; (2) an oversampling strategy named SMOTE to increase the number of the minority class samples before training the model; and (3) a combination of oversampling and undersampling, known as Pipeline.

3. Database description

Anonymized data from heterogeneous sources for 3178 patients (575 MDR and 2603 non-MDR) admitted to the ICU over 16 years from January 2004 to June 2019 have been provided by the UHF. Specifically, the following patient information is available: patient demographic information (age and gender), cultures performed by microbiology laboratory, antibiotics administered, and mechanical ventilation (MV). Also, with the information provided, we have created the following environmental features: total number of patients (NPIT), number of patients who take some antibiotics (NPIt_antibiotic), and number of MDR patients (NPI_MDR) in the ICU at a particular time. The families of the antibiotics the patients can take are the following: Aminoglycosides (AMG), Antifungals (ATF), Carbapenemes (CAR), 1st generation Cephalosporins (CF1), 2nd generation Cephalosporins (CF2), 3rd generation Cephalosporins (CF3), 4th generation Cephalosporins (CF4), unclassified antibiotics (Others), Glycyclines (GCC), Glycopeptides (GLI), Lincosamides (LIN), Lipopeptides (LIP), Macrolides (MAC), Monobactamas (MON), Nitroimidazolics (NTI), Miscellaneous (OTR), Oxazolidinones (OXA), Broad Spectrum Penicillins (PAP), Penicillins (PEN), Polypeptides (POL),

Quinolones (QUI), Sulfamides (SUL) and Tetracyclines (TTC).

Figure 1 shows the boxplot of the ICU stay length until the emergence of the first MDR germ, in comparison with the total stay for non-MDR patients. Since the average of days for MDR patients is 6.90 days versus 7.08 days for non-MDR, we decided to consider a time window of 7-days for each patient, from t_0 to t_6.

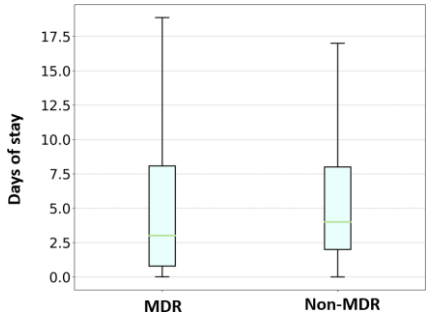


Figure 1. *Boxplot of the stay length (in days), both for MDR patients (until the acquisition of their first MDR) and for non-MDR patients (total stay).*

As it is shown in Figure 2, t_0 corresponds to the slot of 24-hours which is nearest to the ICU admission. The construction of a matrix of temporal features with a time window composed of 7 consecutive 24-hour time instants is also sketched in Figure 2. The upper panel represents the time window for MDR patients, while the lower panel is for non-MDR patients. For each patient, the mask feature represents a value different from 0 if the patient was in the ICU at that time slot.

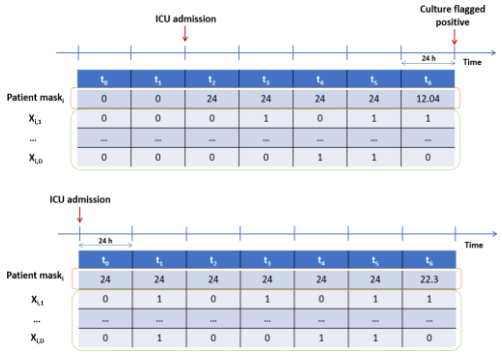


Figure 2. *Sketch with a 7-days time window for MDR patients (upper panel) and non-MDR patients (lower panel).*

Note that the maximum length of 7 days for the window is considered for all patients. However, if a given patient has stayed less days in the ICU, the beginning of the time window is completed with zero values. The time slot t_6 is considered as the day on which the culture was flagged a positive result. After this consideration, the time series is completed in reverse order. For non-MDR patients, we also consider a similar order of time in matrix **X**, but the first-time interval corresponds to the day the patient was admitted to the ICU.

4. Results

The model is designed with a set of observations (training set, 70% of the observations) independent from that used to evaluate the generalization capabilities (test set, 30% of the observations). We have carried out this process 25 times.

A variable is considered as relevant if it was selected with any of the three methods presented in Subsection 2.1. After that, and taking into account that the antibiogram takes 48 hours to provide the results, we have considered those characteristics that satisfy the following conditions: the feature is relevant during the first 48 hours of the time window (t_0 and t_1), from 48 hours (t_2, t_3, t_4, t_5 and t_6) or along the time window (from t_0 to t_6). Table I shows, marked with a "*X*", the variables corresponding to the patient and to the environment that have been finally selected. These variables are used to train the ML models, along with age, gender and the mask.

Patient features							
	t_0	t_1	t_2	t_3	t_4	t_5	t_6
MV(Hours)			X	X	X	X	X
LIN	X	X					
AMG	X	X					
CAR	X	X	X	X	X	X	X
OXA	X	X	X	X	X	X	X
PEN	X	X					
QUI	X	X	X	X	X	X	X
PAP	X	X					
ATF	X	X	X	X	X	X	X
Others	X	X	X	X	X	X	X
CF3	X	X					
CF4	X	X	X	X	X	X	X
SUL			X	X	X	X	X
GLI	X	X	X	X	X	X	X
Environmental features							
	t_0	t_1	t_2	t_3	t_4	t_5	t_6
NPIt MON	X	X	X	X	X	X	X
NPI MDR	X	X	X	X	X	X	X
NPIT			X	X	X	X	X

Table I. *Features selected with any of the three FS methods considered: CI, MI, LASSO.*

The subset of selected features has been used to train the models, making use of different strategies to balance the classes. The learning process using different training strategies has been carried out with a total of 88 features (see Table I), plus age, gender and mask. Table II presents the results for each ML approach and training strategy. The best results obtained for accuracy, specificity and F1-score have been obtained with RF using SMOTE. On the other hand, the best results obtained for Area Under the Curve (AUC) and sensitivity, very import in the clinical field, have been provided by LR using undersampling as a training strategy.

Approach	Training strategy	Accuracy	Specificity	Sensitivity	F1-score	AUC
LR	*Undersampling*	83.43 ± 1.25	83.43 ± 1.73	**83.38 ± 2.56**	64.46 ± 1.74	**83.41 ± 1.13**
	SMOTE	85.26 ± 0.99	86.33 ± 1.46	80.39 ± 2.98	66.27 ± 1.62	83.36 ± 1.23
	Pipeline	84.39 ± 1.12	84.96 ± 1.52	81.77 ± 3.03	65.36 ± 1.93	83.36 ± 1.37
DT	*Undersampling*	79.66 ± 2.23	80.2 ± 2.93	77.12 ± 4.53	57.79 ± 2.8	78.66 ± 2.18
	SMOTE	88.32 ± 1.24	92.48 ± 1.29	69.43 ± 4.03	68.16 ± 3.18	80.96 ± 2.07
	Pipeline	83.49 ± 2.85	85.41 ± 3.75	74.73 ± 4.14	62.23 ± 3.99	80.07 ± 2.11
RF	*Undersampling*	84.35 ± 1.51	85.33 ± 1.9	79.82 ± 3.05	64.79 ± 2.41	82.58 ± 1.54
	SMOTE	**91.21 ± 0.86**	**97.28 ± 0.71**	63.55 ± 3.25	**72.23 ± 2.72**	80.42 ± 1.65
	Pipeline	88.07 ± 2.06	91.04 ± 2.85	74.46 ± 3.64	69.41 ± 3.04	82.75 ± 1.49

Table II. *Mean ± standard deviation for several merit figures (accuracy, specificity, sensitivity, F1-score and AUC) in 25 test sets, considering different classification techniques (first column) and training strategies (second column).*

5. Conclusions and future work

In this work, we have evaluated a set of temporal features making use of MI, CI, and using LR with SGD and loop-based training as a regularization term. In this way, those features that are not relevant to the temporal characterization of MDR acquisition are discarded for training.

Under our studied scenario, the use of an LR model allows classifying a patient as MDR with an 83% success rate (sensitivity). Also, non-MDR patients can be classified with a percentage of 97% (specificity). These results are a first step to accelerate the workflow of the clinical centers anticipating the result of the antibiogram. A fast prediction of MDR would allow the clinical staff to isolate the patient so that the time in which the patient can cause the spread of multiresistant strains to other patients is reduced. Also, the use of the ML to predict MDR would allow effective treatment to be applied, avoiding aggravating the patient's clinical situation.

We can conclude that the acquisition of MDR is a real problem with serious economic and social consequences worldwide. There is a need for joint decision-making at the global level, to acquire greater knowledge of the real extent of the problem and to promote new research that will succeed in mitigating the effects of this problem, the short-term forecasts for which may be catastrophic. A first step to address this problem, as demonstrated in this paper, is the use of ML methods, which can help clinical staff to identify and isolate patients with MDR.

Despite obtaining acceptable results, the analyse of more clinical features such as vital signs or blood test could enhance the prediction of MDR. Furthermore, the use of ML schemes in daily clinical practice can be improved. For this reason, future work is proposed: change the number of time slots that make up the time window, use more complex time models, such as those based on graphs or deep learning, and employ more sophisticated balancing techniques, such as Generative Adversarial Networks.

Acknowledgements

This work has been partly supported by the Spanish Re-search projects PID2019-106623RB-C41 (Beyond), TEC2016-75361-R (Klinilycs), PID2019-107768RA-I00 (AAVis-BMR), DTS17/00158, by the Project Ref. F656 (Proyectos Puente URJC) financed by Rey Juan Carlos University and by the Young Researchers R&D Project Ref. 2020-661 financed by Rey Juan Carlos University and Community of Madrid (Spain), and by the Youth Employment Initiative (YEI) R&D Project Ref. TIC-11649 financed by the Community of Madrid (Spain).

References

[1] Magiorakos, A-P and Srinivasan, A and Carey and others. "Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance". *Clinical microbiology and infection* [Wiley Online Library], 2012.

[2] Infectious Diseases Society of America (IDSA). 'Combating antimicrobial resistance: policy recommendations to save lives". *Clinical Infectious Diseases* [Oxford University Press], 2011.

[3] Freire-Moran, Laura and Aronsson, Bo and others. "Critical shortage of new antibiotics in development against multidrug-resistant bacteria—Time to react is now". Drug resistance updates [Elsevier], 2011.

[4] Bassetti, Matteo and Poulakou, Garyphallia and Ruppe and others. "Antimicrobial resistance in the next 30 years, humankind, bugs and drugs: a visionary approach". Intensive care medicine [Springer], 2017.

[5] Bassetti, Matteo and Poulakou, Garyphallia and Ruppe and others. "The World Health Organization global action plan for antimicrobial resistance". Intensive care medicine [Health & Medical Publishing Group], 2015.

[6] World Bank Group. "Drug-resistant Infections: A Threat to Our Eco- nomic Future". 2017.

[7] Brusselaers, Nele and Vogelaers, Dirk and Blot, Stijn. "The rising problem of antimicrobial resistance in the intensive care unit". Intensive care medicine [Springer], 2011.

[8] Brusselaers, Nele and Vogelaers, Dirk and Blot, Stijn. "Application of high-dimensional feature selection: evaluation for genomic prediction in man". Scientific reports [Nature Publishing Group], 2015.

[9] Guyon, Isabelle and Elisseeff, André. "An introduction to variable and feature selection". Journal of machine learning research, 2003.

[10] Efron, Bradley and Tibshirani, Robert J. "An introduction to the boot- strap". CRC press, 1994.

[11] Liu, Huawen and Sun, Jigui and Liu, Lei and Zhang, Huijie. "Feature selection with dynamic mutual information". Pattern Recognition [Elsevier], 2009.

[12] Meinshausen, Nicolai. "Relaxed lasso". Computational Statistics & Data Analysis [Elsevier], 2007.

[13] Alsheikh, Mohammad Abu and Lin, Shaowei and Niyato, Dusit and Tan, Hwee-Pink. "Machine learning in wireless sensor networks: Algorithms, strategies, and applications". IEEE Communications Surveys & Tutorials [IEEE], 2014.

[14] Martínez-Agüero, Sergio and Mora-Jiménez, Inmaculada and Lérida- García, Jon and Álvarez-Rodríguez, Joaquín and Soguero-Ruiz, Cristina. "Machine Learning Techniques to Identify Antimicrobial Resistance in the Intensive Care Unit". Multidisciplinary Digital Publishing Institute [Entropy], 2019.

[15] Kotsiantis, Sotiris B. "Decision trees: a recent overview". Artificial Intelligence Review [Springer], 2013.

[16] James, Gareth and Witten, Daniela and Hastie, Trevor and Tibshirani, Robert. "An introduction to statistical learning". [Springer], 2013.

[17] Genuer, Robin and Poggi, Jean-Michel and Tuleau-Malot, Christine. "Variable selection using random forests". Pattern recognition letters [Elsevier], 2010.

[18] García, Vicente and Sánchez, José Salvador and Mollineda, Ramón Alberto. "On the effectiveness of preprocessing methods when dealing with different levels of class imbalance". Knowledge-Based Systems [Elsevier], 2012.

Detección de criterios dermatoscópicos del carcinoma basocelular mediante aprendizaje profundo

M. Lazo[1], B. Acha[1] , T. Toledo[2], A. Serrano[2], R. Barros[2], C. Serrano[1]

[1] Departamento de Teoría de la Señal y Comunicaciones, Universidad de Sevilla, Sevilla, España, manu11235@gmail.com {bacha,cserrano}@us.es

[2] Servicio de Dermatología, Hospital Universitario Virgen Macarena, Sevilla, España

Resumen

En este artículo se presenta un sistema de ayuda al diagnóstico que permite dar soporte a la decisión con una clasificación de lesiones de piel por medio del uso de imágenes de dermatoscopia en color. En este caso se trabaja para ayudar a detectar el carcinoma basocelular (BCC, Basal Cell Carcinoma) con el uso de procesamiento de imágenes y redes neuronales convolucionales. Para ello se detectan automáticamente los patrones dermatoscópicos, que los dermatólogos emplean para la discriminación del BCC. Como entradas a la red neuronal profunda (VGG16) se combinan las imágenes originales con imágenes preprocesadas mediante un análisis del color. Se obtienen precisiones entre 0.62 y 0.83, dependiendo del patrón detectado.

1 Introducción

En este trabajo se desarrolla una herramienta para ayudar al diagnóstico de lesiones de piel, concretamente del BCC, que es el cáncer de piel más frecuente [1,2]. Éste presenta unas estructuras dermatoscópicas típicas, por lo que la detección de esas estructuras puede servir para identificarlo. Éste es, por tanto, el objetivo de este trabajo.

La técnica de la dermatoscopia es esencial en el estudio de lesiones de piel [1,3,4], y permite al dermatólogo analizar la lesión con enorme precisión mediante inspección visual no invasiva [1,4] e incluso tomar capturas fotográficas para un análisis posterior. Así, gran cantidad de herramientas que asisten en la detección de lesiones de piel se diseñan para imágenes de dermatoscopia.

Desde hace algunos años se está haciendo cada vez más popular el uso de tecnologías como el aprendizaje máquina o las redes neuronales, para la investigación en dermatología, porque son herramientas que han experimentado una notable evolución y pueden resultar de gran ayuda en el diagnóstico temprano [2]. Son muy numerosos los trabajos centrados en la detección de melanoma mediante redes neuronales profundas [6,7]. Estas investigaciones han sido posibles gracias a las extensas bases de datos, que se han puesto a disposición de la comunidad científica a través de concursos organizados por ISIC (International Skin Imaging Collaboration) [8]. Sin embargo, el número de trabajos centrados en la detección del BCC es mucho más limitado [1,2,9]. Y en concreto, hasta donde los autores conocen, no se han encontrado trabajos de detección automática de los criterios o patrones dermatoscópicos en BCCs.

En este trabajo se busca realizar un diagnóstico emulando el de los dermatólogos, de forma que el clasificador encuentre los patrones dermatoscópicos presentes en la lesión, que son identificativos del BCC.

2 Metodología

El trabajo de programación se realiza en lenguaje Python, utilizando las librerías TensorFlow y Keras así como numpy, os, PIL, scikit, opencv, etc.

Se dispone de un conjunto de 524 imágenes originales obtenidas con pacientes residentes en el área de asistencia sanitaria que corresponde al Hospital Virgen Macarena de Sevilla, atendidos mediante teledermatología con dermatoscopia desde octubre de 2019 a marzo de 2020. Además, todos los pacientes son adultos de raza caucásica y presentan una lesión que ha motivado la teleconsulta, a la que se le supone, por su naturaleza, al menos cuatro meses de evolución.

2.1 Generación de la base de datos

Para la generación de la base de datos se ha desarrollado un programa con una interfaz gráfica sencilla de visualización de imágenes que permite hacer recortes y generar nuevas imágenes etiquetadas (Figura 1). Los especialistas pueden visualizar la base de datos y buscar patrones relevantes para tomar recortes de los mismos y etiquetarlos. De esta forma se generan las verdades de referencia o *ground truth*. La base de datos únicamente contiene imágenes que presentan las estructuras que tienen que identificarse.

Las etiquetas corresponden a los patrones que representan los criterios dermatoscópicos que los dermatólogos buscan para establecer su diagnóstico [1,3]: múltiples glóbulos azul-gris, nidos grandes ovoides, áreas en hoja de arce, estructuras en rueda de carro, telangiectasias ramificadas y ulceraciones. Además, tenemos en cuenta el patrón de red pigmentada, cuya presencia denota que la lesión no es carcinoma basocelular (criterio negativo). Se muestran ejemplos en la Figura 2.

La aplicación informática diseñada está implementada con la librería tkinter en Python y consta de dos visores de

ISBN: 978-84-09-25491-0, pags. 468- 471

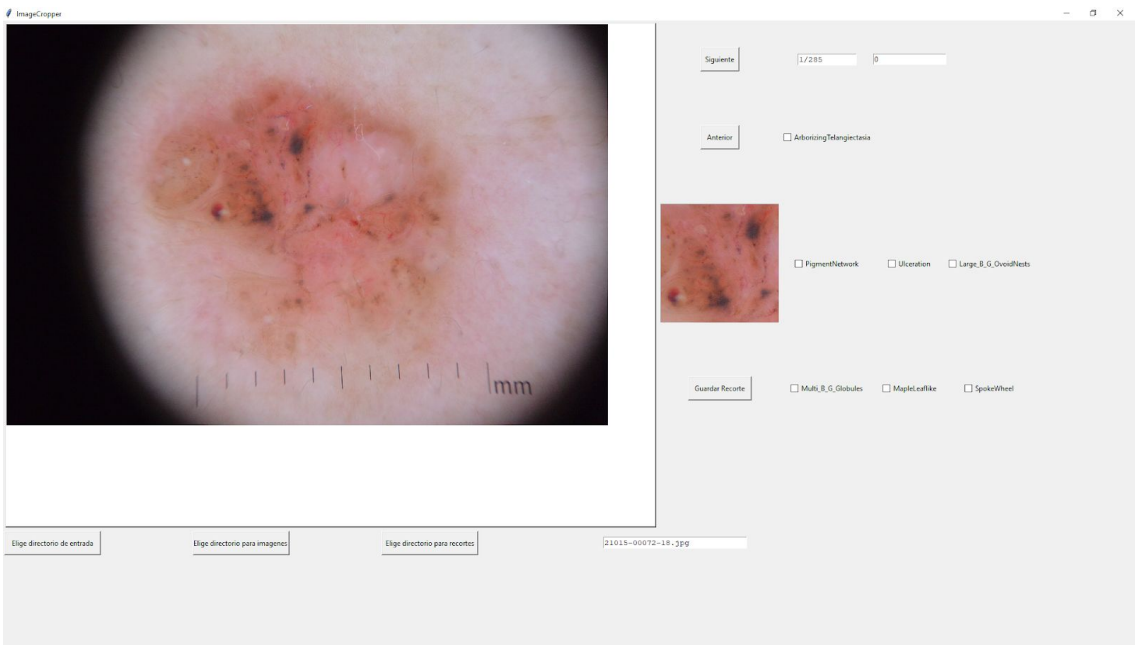


Figura 1. *Aspecto de la aplicación imageCropper para el recorte y etiquetado de las imágenes.*

imágenes y algunas funcionalidades para desplazarse a lo largo de una lista o directorio con ficheros. Cuando el médico selecciona con el ratón una zona en la que un patrón está presente, se muestra ésta en el visor secundario. Seguidamente puede marcar como activas las etiquetas que designan los patrones presentes en la zona y podrá finalmente generar un recorte etiquetado

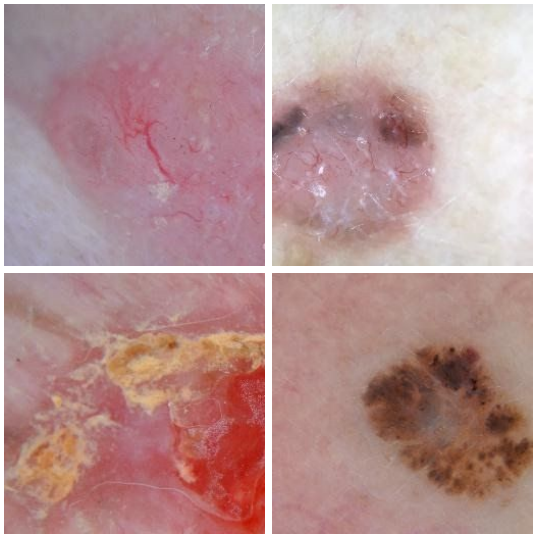


Figura 2. *Muestra de imágenes originales. Telangiectasia (superior izquierda), ovoides (superior derecha), ulceración (inferior izquierda) y red pigmentada (inferior derecha).*

Como en este caso se tienen siete patrones que, en un principio, podrían coexistir en el mismo área, el número de posibles etiquetas será 128. Esto provoca que se deba realizar una codificación de las mismas. Cada imagen de recorte es examinada y se anotan los patrones que están presentes para finalmente pasar esta información al codificador. De esta forma, se puede ahorrar algo de tiempo y coste computacional al pasar de usar vectores binarios de 7 elementos a enteros con valores entre 0 y 127. Así, el problema a abordar es multi-etiqueta y multi-clase.

La herramienta ha sido mejorada para permitir además hacer selección de regiones de forma arbitraria mediante recortes a mano alzada.

2.2 Preprocesamiento

El preprocesamiento consiste en la cuantificación en los colores principales presentes en cada uno de los patrones o estructuras dermatoscópicas típicas.

Inicialmente se aplica un algoritmo de búsqueda de agrupaciones (*clustering*) al conjunto de imágenes de entrenamiento pertenecientes a un mismo patrón. De esta manera se identifican los colores principales que típicamente presenta ese patrón. Para esto, se utiliza el algoritmo *K-Means con K=18,* valor hallado experimentalmente. En la Figura 3 se muestran las paletas de colores obtenidas para los siete patrones dermatoscópicos.

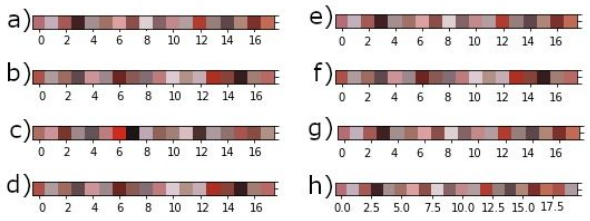


Figura 3. *Paletas de colores cuantificados con el algoritmo K-means para los siete patrones dermatoscópicos. La relación es: a) red pigmentada, b) ulceración, c) nidos ovoides azul-gris, d) múltiples glóbulos azul-gris, e) hoja de arce, f) telangiectasia y h) paleta generada cribando las demás.*

A continuación, para reducir el número total de colores cuantificados, se fusionan aquellos colores cuya distancia Euclídea es menor que un umbral. En la Figura 3.h se muestra la paleta final obtenida.

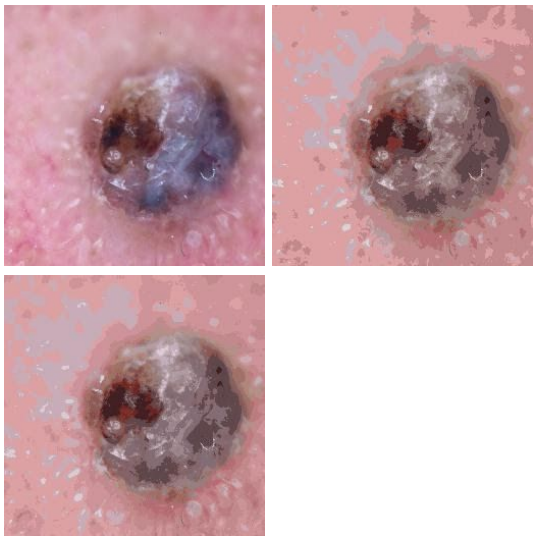


Figura 4. *Comparación de una imagen original del patrón ulceración (izquierda), una codificada (derecha) en 20 colores y otra en 32 colores (abajo).*

Además, se emplea la técnica de *data augmentation* para aumentar el tamaño de la base de datos. Las transformaciones que se han practicado son: rotaciones, volteo horizontal, volteo vertical y cambio de escala. Se realizan cinco versiones de cada imagen (recorte), con lo que, incluyendo las originales se logra un total de 5802 imágenes. Antes de realizar este paso se tienen 967 imágenes. La distribución de patrones puede verse en la Tabla 1.

Tabla 1. *Distribución de patrones en la base de datos.*

	Número de ocurrencias
Red pigmentada	187
Ulceración	408
Nidos ovoides azul-gris	300
Múltiples glóbulos azul-gris	410
Hoja de arce	258
Rueda de carro	178
Telangiectasia	490

2.3 Clasificación

La arquitectura de red neuronal de la que se parte para realizar los experimentos de este proyecto está basada en la de una red VGG16. Esta red es modificada para clasificar las imágenes. Al cargar el modelo, se bloquea el aprendizaje en las capas, salvo en las de salida (las capas *fully connected)*, que son adaptadas al problema. De esta forma, se cargan los pesos preentrenados con la base de datos ImageNet para la red, y las últimas capas van modificándose durante el entrenamiento. Así, aplicamos *transfer learning* para paliar el tamaño reducido de la base de datos.

En este trabajo se presentan los resultados de tres escenarios distintos, que se detallan en las siguientes subsecciones.

Como se ha mencionado, se está trabajando con siete patrones y, a pesar de que existen enlaces que relacionan la presencia de unos con otros, están atados a criterios médicos y dermatológicos muy difíciles de controlar. Por eso se trata la presencia de cada patrón con independencia de los demás. Por ello, la última capa de la VGG16 ha sido sustituida por una función sigmoidal a la salida de cada neurona, de forma que se permite la activación de varias salidas a la vez, correspondientes a la presencia de varios patrones.

2.3.1 Clasificación de imágenes originales

Inicialmente, la entrada al clasificador son los recortes originales, a los que se les ha aplicado *data augmentation,* según se ha descrito en la sección anterior. Se emplea, como se ha mencionado, una estructura basada en la red VGG16, con pesos preentrenados en ImageNet.

2.3.2 Clasificación de imágenes procesadas

Para comprobar el efecto del procesado que se ha explicado anteriormente, también se prueba a entrenar el mismo clasificador con los recortes cuantificados según el método explicado en la sección 2.2 .

2.3.3 Clasificación simultánea de imágenes procesadas y no procesadas

Como prueba final, se diseña una arquitectura con dos redes VGG16 modificadas de forma similar a las que se utilizan en los escenarios descritos en las secciones 2.3.1 y 2.3.2. En esta ocasión, las redes tienen un punto de unión que concatena los resultados de las últimas capas de ambas redes.

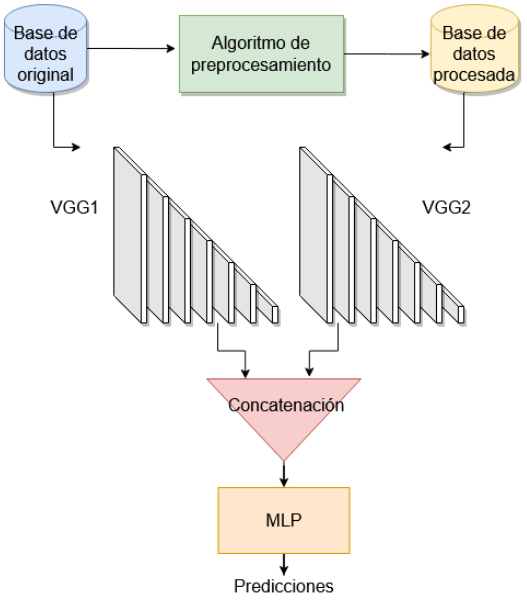


Figura 5. *Diagrama con la estructura del último experimento.*

Se muestra en la Figura 5 un diagrama con la arquitectura propuesta. Se ha trabajado con dos redes VGG16. Se ha bloqueado el aprendizaje de las capas convolucionales de éstas. La salida de las redes VGG16, se concatena y se utiliza como entrada de otro modelo, esta vez *fully connected* (MLP: Multilayer Perceptron) Así, las complejas redes VGG16 tendrán que identificar los patrones visuales, texturas, detalles de color..., y la red MLP deberá aprender a combinar esa información para finalmente determinar la existencia o no de los patrones dermatológicos.

3 Resultados

En la Tabla 2 se muestran los resultados de las pruebas realizadas para cada patrón.

Tabla 2. *Resultados de la precisión para cada patrón y cada experimento.*

	Imagen Original	Imagen Preprocesada	Entrada dual
Red pigmentada	0.81	0.79	**0.83**
Ulceración	0.69	0.62	**0.71**
Nidos ovoides	0.71	0.71	**0.75**
Múltiples glóbulos	**0.65**	0.64	0.62
Hoja de arce	0.73	**0.75**	0.74
Rueda de carro	0.81	0.83	**0.83**
Telangiectasia	0.61	0.64	**0.69**

Como puede verse, hay patrones que en la imagen original resultan más fáciles de clasificar que en la imagen preprocesada (red pigmentada, ulceración y múltiples glóbulos azul-gris), pero también ocurre lo inverso, así que el preprocesamiento facilita la tarea en algunas clases (hoja de arce, rueda de carro y telangiectasia). Como el preprocesamiento mejora la clasificación de algunos patrones, se intenta mantener esa mejora y a la vez no perder los buenos números que da la imagen original con la red de entrada dual. Esta última es la que presenta mejores resultados en casi todos los patrones. Es notable cómo el último experimento mejora el resultado, por ejemplo, en la telangiectasia, donde las redes simples dan resultados moderadamente distantes al de la red dual.

4 Conclusiones

Con este trabajo se han establecido las bases para un sistema de ayuda al diagnóstico para la detección del carcinoma basocelular, y estos resultados corroboran que el método es factible, con lo cual hay que llevarlo a la versión más eficiente y refinada posible para que el sistema sea de veras una herramienta útil en la práctica clínica.

A continuación, se presentan las líneas de mejora del trabajo:

- El trabajo tiene margen de mejora en cuanto a la precisión de clasificación. El primer paso para ello es el aumento de la base de datos.
- Las herramientas que se han comentado que sirven de asistencia para la generación de la base de datos podrían tener una versión renovada, más cómoda para el usuario e, incluso, con una versión para teléfonos móviles.
- Un uso importante de la herramienta sería para la formación de especialistas. Por ejemplo, se presenta una imagen en un visor y se pregunta por los patrones presentes.
- Otra oportunidad con este método es la de localizar el recorte dentro de la imagen completa para así comenzar también a ubicar espacialmente los patrones. Eventualmente podría tenerse disponible un clasificador que barra completamente una imagen de entrada y señale los patrones que ha clasificado correctamente así como su posición.
- Si se logra generar una base de datos de mayor tamaño, sería deseable analizar las relaciones inter-clase de patrones.

Agradecimientos

Este trabajo se ha financiado con el proyecto DPI2016-81103-R del Plan Nacional.

Referencias

[1] Zaballos DP. Criterios dermatoscópicos de los tumores no melanocíticos. *Med Cutan Iber Lat Am*, vol 43, sup 2, 2015, pp 93-103.

[2] Marka A, Carter JB, Toto E, Hassanpour S. Automated detection of nonmelanoma skin cancer using digital images: a systematic review. *BioMedCentral Medical Imaging*, vol 19, 21, 2019.

[3] Menzies SW et al. Surface Microscopy of Pigmented Basal Cell Carcinoma, *Archives of Dermatology*, vol 136, pp 1012-1016, 2000.

[4] Barata C, Celebi ME, Marques JS. A survey of Feature Extraction in Dermoscopy Image Analysis of Skin Cancer, *IEEE Journal of Biomedical and Health Informatics*, vol 23, sup 3, 2019.

[5] Madooei A y Drew MS, Incorporating Colour Information for Computer-Aided Diagnosis of Melanoma from Dermoscopy Images: A Retrospective Survey and Critical Analysis, *International Journal of Biomedical Imaging*, vol 2016, ID 4868305.

[6] Esteva A, Kuprel B, Novoa R, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, vol, 542, pp, 115–118, 2017. https://doi.org/10.1038/nature21056

[7] Ammara M, Adel AAJ. Computer Aided Diagnostic Support System for Skin Cancer: A Review of Techniques and Algorithms, *International Journal of Biomedical Imaging*, vol. 2013, ID 323268, 2013. https://doi.org/10.1155/2013/323268

[8] Página web del archivo de la *International Skin Imaging Collaboration*, https://www.isic-archive.com/

[9] Kaymak S, Esmaili P y Serener A, Deep Learning for Two-Step Classification of Malignant Pigmented Skin Lesions, *2018 14th Symposium on Neural Networks and Applications (NEUREL)*,vol 2018, pp, 1-6, doi: 10.1109/NEUREL.2018.8587019.

Diagnóstico de esclerosis múltiple mediante el análisis de imágenes de tomografía de coherencia óptica con redes neuronales convolucionales

R. Barea[1], M. Ortiz[2], E. Garcia-Martin[3], M.J. Rodrigo[3], E. Vilades[3], A. López Dorado[1], R. Gutiérrez[1], C. Cavaliere[1], L. Boquete[1]

[1] Grupo de Ingeniería Biomédica, Departamento de Electrónica, Universidad de Alcalá, España

[2] School of Physics, University of Melbourne, Australia

[3] Servicio de Oftalmología del Hospital Universitario Miguel Servet de Zaragoza. Universidad de Zaragoza. Instituto de Investigación Sanitaria de Aragón

Resumen

El objetivo de este trabajo es investigar la capacidad diagnóstica de la enfermedad de esclerosis múltiple (EM) utilizando datos de los espesores de diferentes capas de la retina, obtenidos mediante Tomografía de Coherencia Óptica (OCT) y clasificados mediante una red neuronal convolucional (RNC). Se dispone de imágenes de OCT de 48 sujetos de control y 48 pacientes EM (recurrente-remitente) de diagnóstico reciente, con los espesores (45x60 puntos) de las siguientes estructuras retinianas: retina completa, capa de fibras nerviosas (RNFL), dos capas de células ganglionares (GCL+, GCL++) y coroides. Mediante el análisis del área bajo la curva, se obtienen las estructuras con mayor capacidad discriminante controles vs EM. Se implementa una RNC formada por un módulo de extracción de características y un módulo de clasificación. Se utilizan como entradas a la RNC las imágenes de las estructuras más discriminantes. El entrenamiento y test se realiza mediante validación cruzada dejando uno fuera. En la base de datos disponible, las tres estructuras más discriminantes son GCL+, GCL++ y las imágenes de la retina completa. Utilizando las mismas en una RNC con un solo módulo de extracción de características (número de filtros: 64, dimensiones de cada filtro: 7x7, función de transformación: f(x)=max(0,x), reducción de datos en ventanas 2x2), se obtiene una precisión en la clasificación del 94,79%. Estos resultados sugieren que la utilización de técnicas de aprendizaje profundo para la clasificación de espesores de diferentes estructuras de la retina, obtenidas por OCT, puede contribuir al diagnóstico de EM.

1. Motivación

La esclerosis múltiple (EM) es una enfermedad crónica, neurodegenerativa y desmielinizante que afecta al sistema nervioso central y provoca en los pacientes un elevado número de síntomas, siendo frecuentes las afectaciones al sistema cognitivo, motor y visual [1]. En el año 2015 el número de afectados a nivel mundial fue de 2,5 millones de personas y en un elevado porcentaje de casos se empieza a desarrollar la enfermedad entre los 20-40 años; tres de cada cuatro pacientes son mujeres.

El diagnóstico de EM se realiza mediante los criterios de McDonald [2], que principalmente se basan en el análisis de imágenes de resonancia magnética (MRI) y en la situación clínica del paciente. Sin embargo, su efectividad en las fases iniciales de la enfermedad no es plena y en consecuencia, es conveniente investigar en otros posibles biomarcadores.

La Tomografía de Coherencia Óptica (OCT) permite obtener in vivo, de forma no invasiva y con gran precisión espesores de diferentes estructuras de la retina. Mediante un sistema de interferometría óptica de baja coherencia se puede medir la distancia a partir de la luz reflejada por diferentes tejidos y con ello, se obtiene la morfología tisular de capas con diferentes índices de reflectividad. Numerosos estudios evalúan las ventajas de la técnica OCT como biomarcador en el diagnóstico de EM [3][4][5][6] y su capacidad diagnóstica sigue siendo objeto de investigación.

Las técnicas de aprendizaje profundo se utilizan ampliamente para el análisis de información en medicina, como pueden señales unidimensionales, por ejemplo electroencefalogramas [7]. Sin embargo son más numerosas las aplicaciones dedicadas al análisis de imágenes médicas [8]. En el campo de los sistemas OCT, se utilizan para la segmentación de las diferentes capas de la retina [9].

La estructura de las redes neuronales convolucionales (RNC) se pueden dividir en dos módulos conectados en cascada: módulo de extracción de características y módulo de clasificación [10]. El primero obtiene las características de la imagen mediante convoluciones con máscaras de diferentes coeficientes y dimensiones, que destacan los bordes, gradientes de intensidad, etc. de la imagen original. El módulo de clasificación tiene como entradas características obtenidas en el módulo previo y puede ser implementado con una red neuronal con conexiones hacia delante (perceptrón multicapa, red de base radial, etc); su entrenamiento puede realizarse con algoritmos basados en descenso por el gradiente, entre otras opciones.

ISBN: 978-84-09-25491-0, pags. 472- 475

1.1. Objetivo

El objetivo de este trabajo es evaluar un sistema de ayuda al diagnóstico de EM mediante el análisis de imágenes de OCT con redes neuronales convolucionales.

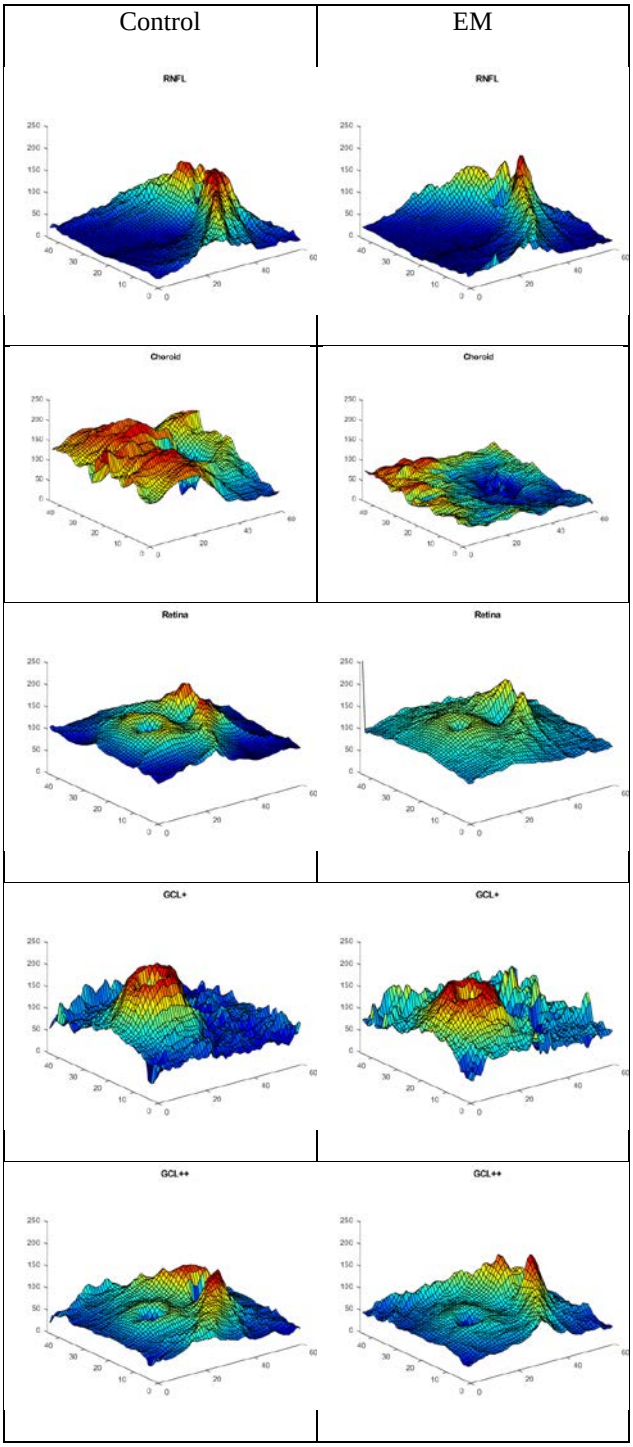


Figura 1. *Ejemplo de imágenes 3D de las 5 estructuras obtenidas con OCT en sujetos de control (izquierda) y pacientes EM (derecha). En orden descendente, espesores de: RNFL, coroides, retina, GCL+ y GCL++.*

2. Métodos

El estudio ha sido aprobado por el Comité Ético del Hospital Universitario Miguel Servet (HUMS), se han seguido los protocolos recogidos en la declaración de Helsinki y los participantes ratificaron su consentimiento por escrito.

Se dispone de una base de datos consistente en 48 registros OCT de ambos ojos de pacientes con diagnóstico precoz de EM (recurrente-remitente) sin neuritis óptica previa y 48 registros de sujetos sanos.

Se ha utilizado un equipo tomografía de coherencia óptica de barrido modelo Topcon DRI OCT Triton Swept source OCT: SS-OCT (Topcon, Japón). El valor medio de la longitud de onda utilizado es de 1050 nm, con velocidad de escaneo de 100K cortes/sg. El protocolo de exploración obtiene una imagen de 45x60 puntos de medida para cada una de las estructuras exploradas que abarca el área macular y peripapilar, cubriendo una superficie de 9 mm de alto x 12 mm de ancho. Se selecciona aleatoriamente un ojo de cada sujeto para incluirlo en el análisis.

Se obtienen espesores de las siguientes estructuras (Fig.1):

- Espesor de la retina completa: desde la membrana limitante interna (ILM) hasta el límite entre el epitelio pigmentario y la capa de fotorreceptores.
- Espesor de la RNFL (capa de fibras nerviosas): desde la ILM hasta el límite interno de la capa de células ganglionares.
- Espesor GCL+: desde el límite interno de la capa de células ganglionares hasta el límite externo de la IPL (capa plexiforme interna).
- Espesor GCL++: Desde la ILM hasta la línea IPL/INL, englobando así tanto la RNFL como la capa de células ganglionares.
- Espesor coroides: desde la membrana de Bruch hasta la superficie escleral de la coroides.

2.1. Selección de estructuras discriminantes

Mediante un análisis del área bajo la curva ROC (Receiver Operating Characteristic), se ha comprobado que no todas las estructuras disponibles de la retina tienen la misma capacidad discriminante para diferenciar imágenes de controles y pacientes EM. Las estructuras con mayor capacidad diagnóstica, de acuerdo a este criterio, son: GCL+, GCL++ y las imágenes de la retina completa; sus espesores serán utilizados en el proceso de clasificación.

2.2. Arquitectura de la RNC

Se muestra en la fig. 2. la estructura de la RNC implementada mediante Toolboox Deep Learning (Matlab). La entrada son 3 imágenes de dimensiones 45X60 de las tres estructuras más discriminantes. El módulo de extracción de características está formado por una unidad de convolución con N_F filtros de dimensiones

dxd. Los resultados de la convolución son transformados mediante la función $f(x) = \max(0, x)$ para mejorar el proceso de entrenamiento. El último paso dentro de este módulo consiste en la reducción de las dimensiones (reducción por muestreo), también para disminuir el número de pesos ajustables en la red neuronal y en consecuencia, beneficiar el entrenamiento.

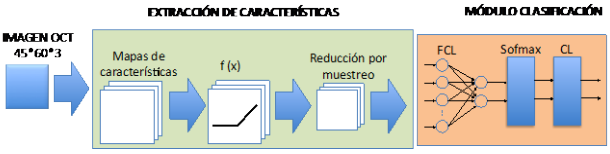


Figura 2. *Arquitectura de la red neuronal convolucional implementada.*

El módulo de clasificación está formado por una red totalmente conectada (FCL), constituida por la capa de entrada y la capa de salida (2 neuronas), estando todas las neuronas de entrada conectadas a las dos neuronas de salida con un peso y un bias, ambos ajustables. Las salidas de FCL se normalizan mediante una función exponencial normalizada (sofmax):

$$p(i) = \frac{e^{z_i}}{\sum_j e^{z_j}}$$

que asigna un valor de probabilidad a la salida de la FCL. La capa de clasificación asigna el valor de salida de la clasificación (control o EM).

Los resultados del clasificador se obtienen mediante validación cruzada dejando uno fuera: de los 96 datos de sujetos disponibles, se utilizan 95 en el entrenamiento y los datos del sujeto restante se utilizan en el test. El proceso se repite 96 veces, de forma que todos los sujetos son testeados. Los resultados de la clasificación se presentan mediante una matriz de confusión.

3. Resultados

La base de datos está formada por imágenes de OCT provenientes de 48 pacientes (hombres/mujeres: 9/39) con EM (43.79 ± 8.41 años) y 48 (hombres/mujeres: 10:38) sujetos de control (44.44 ± 7.18 años). No existe diferencia significativa en las medias de edad entre ambos grupos (p=0.107) ni en la distribución entre sexos (p=0.451). Los pacientes son de reciente diagnóstico (7.35 ± 1.95 meses).

Los parámetros que definen la arquitectura de la RNC que permite obtener los mejores resultados en la clasificación son: número de filtros N_F=64, dimensiones de los filtros=7x7. Por lo tanto, se obtienen un total de 64 imágenes que representan los rasgos o características de las imágenes OCT. En la capa de reducción se definen ventanas de dimensiones 2x2, sin solapamiento entre ventanas y se selecciona el pixel de mayor valor, destacando de este modo las características más abruptas de los mapas de características.

El ajuste del clasificador se realiza mediante descenso de gradiente estocástico [11].

En la tabla 1 se muestra la matriz de confusión con los resultados de la clasificación de las imágenes de los 96 sujetos participantes mediante la RNC.

		Casos Reales		**Precisión**
		Controles	EM	
Predicción RNC	Controles	46	3	
	EM	2	45	94,79%

Tabla 1. *Matriz de confusión*

4. Discusión

En este trabajo se ha propuesto un sistema de ayuda al diagnóstico de EM mediante el análisis de imágenes de OCT utilizando redes neuronales convolucionales, obteniendo unos resultados prometedores en la clasificación (precisión=94,79%).

Los resultados obtenidos en el presente trabajo suponen una mejora en relación a un trabajo previo, en el que se utilizaron redes neuronales modulares [12] y la misma base de datos y permiten abrir nuevas posibilidades en el diagnóstico temprano de pacientes con EM.

Una de las ventajas de las redes neuronales convolucionales es que pueden considerarse como una caja negra, en la que se establecen correspondencias entre sus entradas y salidas, no siendo necesario en muchos casos una extracción explícita de características. En el presente trabajo se han seleccionado las estructuras más discriminantes para ser utilizadas como entrada de la RNC. Sin embargo, como líneas futuras de investigación se considera la posibilidad de implementar métodos capaces de seleccionar aquellas partes de las imágenes con mayor capacidad discriminante, antes de ser utilizadas como entradas a la RNC. Además sería conveniente ampliar la base de datos, de tal modo que se incluyan datos de sujetos obtenidos en diferentes centros y también utilizando equipos OCT con diferentes tecnologías. También sería deseable complementar la información obtenida con OCT con otros tipos de imágenes médicas, que han demostrado eficacia en el diagnóstico de EM.

Si los trabajos futuros confirman unos resultados satisfactorios, sería posible implementar un sistema de cribado para la detección de EM de fácil uso por parte del personal médico, no invasivo y de coste razonable. Dentro de este posible sistema de ayuda al diagnóstico precoz deberían ser incluidos otros datos clínicos del paciente, que aportaran capacidad diagnóstica al proceso.

Agradecimientos

Financiado por el Ministerio de Ciencia, Innovación y Universidades mediante el proyecto: DPI2017-88438-R, "Investigación de la técnica de potenciales evocados visuales multifocales. Aplicación en estudios de evolución de esclerosis múltiple y evaluación de medicamentos".

PI17/01726 (Instituto de Salud Carlos III) (EGM, EV y MJR) RETICS: "Enfermedades oculares: prevención, detección precoz y tratamiento de la patología ocular prevalente degenerativa y crónica: OFTARED". (EGM: RD16/0008/029) (LB: RD16/0008/0020).

Proyecto "Diagnóstico de esclerosis múltiple mediante técnicas de aprendizaje profundo (Deep- learning) a partir de espesores de la neuro-retina obtenidos mediante OCT. Ref. CCG19/CCS-041, financiado por la Universidad de Alcalá.

Referencias

[1] C. A. Dendrou, L. Fugger, and M. A. Friese, "Immunopathology of multiple sclerosis," *Nat. Rev. Immunol.*, vol. 15, no. 9, pp. 545–558, Sep. 2015.

[2] A. J. Thompson *et al.*, "Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria," *Lancet Neurol.*, vol. 17, no. 2, pp. 162–173, Feb. 2018.

[3] A. Petzold *et al.*, "Retinal layer segmentation in multiple sclerosis: a systematic review and meta-analysis," *Lancet. Neurol.*, vol. 16, no. 10, pp. 797–812, 2017.

[4] J. Britze, G. Pihl-Jensen, and J. L. Frederiksen, "Retinal ganglion cell analysis in multiple sclerosis and optic neuritis: a systematic review and meta-analysis," *J. Neurol.*, vol. 264, no. 9, pp. 1837–1853, Sep. 2017.

[5] R. Alonso, D. Gonzalez-Moron, and O. Garcea, "Optical coherence tomography as a biomarker of neurodegeneration in multiple sclerosis: A review.," *Mult. Scler. Relat. Disord.*, vol. 22, pp. 77–82, May 2018.

[6] C. Cavaliere *et al.*, "Computer-Aided Diagnosis of Multiple Sclerosis Using a Support Vector Machine and Optical Coherence Tomography Features," *Sensors*, vol. 19, no. 23, p. 5323, Dec. 2019.

[7] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *J. Neural Eng.*, vol. 16, no. 5, p. 051001, Aug. 2019.

[8] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep Learning Applications in Medical Image Analysis," *IEEE Access*, vol. 6, pp. 9375–9389, 2018.

[9] M. Pekala, N. Joshi, T. Y. A. Liu, N. M. Bressler, D. C. DeBuc, and P. Burlina, "Deep learning based retinal OCT segmentation," *Comput. Biol. Med.*, vol. 114, p. 103445, Nov. 2019.

[10] A. Shrestha and A. Mahmood, "Review of Deep Learning Algorithms and Architectures," *IEEE Access*, vol. 7, pp. 53040–53065, 2019.

[11] L. Bottou, "Large-Scale Machine Learning with Stochastic Gradient Descent," en *Proceedings of COMPSTAT'2010*, Heidelberg: Physica-Verlag HD, 2010, pp. 177–186.

[12] R. Barea *et al.*, "Detección de pacientes con esclerosis múltiple mediante procesamiento de imágenes de tomografía de coherencia óptica utilizando redes neuronales," en *Actas del 27 Seminario Anual de Automática, Electrónica Industrial e Instrumentación (SAAEI 2020), Ciudad Real, (ISBN: 978-84-122260-2-7)*, 2020, pp. 432–436.

Predicción del Éxito de la Ablación por Catéter en Fibrilación Auricular Persistente Mediante Características Tiempo-Frecuencia de las Ondas Fibrilatorias

P. Escribano Cano[1], J. Ródenas García[1], M.A. Arias Palomares[2], P. Langley[3], J.J. Rieta Ibáñez[4], R. Alcaraz Martínez[1]

[1] Grupo de Investigación en Ingeniería Electrónica, Biomédica y de Telecomunicación, DIEEAC, Universidad de Castilla-La Mancha , Albacete, España; {pilar.escribano, juan.rodenas, raul.alcaraz}@uclm.es
[2] Unidad de Arritmias, Hospital Virgen de la Salud, Toledo, España
[3] Faculty of Science and Engineering, University of Hull, Hull, United Kingdom
[4] BioMIT.org, Departamento de Ingeniería Electrónica, Universidad Politécnica de Valencia, Valencia, España

Resumen

A pesar de su limitada tasa de éxito a largo plazo, la ablación por catéter (AC) es el tratamiento de primera línea de la fibrilación auricular (FA) persistente. La predicción preoperatoria del resultado de este procedimiento permitiría optimizar la selección de pacientes, entre otras ventajas clínicas y económicas. Con este fin, la frecuencia dominante (FD) y la amplitud de las ondas fibrilatorias (ondas f) normalizada (AOFn) reflejadas en el registro electrocardiográfico (ECG) ya han proporcionado resultados prometedores, pero insuficientes desde un punto de vista clínico. Este trabajo explora la capacidad de un conjunto de características frecuenciales y de amplitud de las ondas f para mejorar la predicción del resultado de la AC. Así pues, además de los dos índices anteriormente indicados, se han analizado: la entropía espectral (EE), la medida de la planicidad espectral (MPE) y la suma ponderada de la amplitud espectral (SPAE). Los resultados obtenidos para 204 intervalos de ECG, extraídos de 51 pacientes con FA persistente, muestran para todos los índices diferencias estadísticamente significativas entre los pacientes que mantuvieron ritmo sinusal (RS) y los que recayeron a FA tras un seguimiento de 9 meses, pero una capacidad discriminante limitada, entre un 55 y un 62 %. Este resultado se mejoró en un 15-23 % cuando se combinaron mediante un análisis lineal discriminante AOFn, EE y SPAE, obteniendo un valor del 77,2 %. En consecuencia, esta combinación de características de frecuencia y amplitud parece proporcionar nueva información sobre el remodelado del sustrato auricular, la cual podría ser útil para mejorar la predicción del resultado de la AC.

1. Introducción

La fibrilación auricular (FA) es una taquicardia supraventricular común que interrumpe el ritmo sinusal (RS) del corazón. De hecho, es la arritmia más frecuente en la práctica clínica, que afecta aproximadamente a 37,5 millones de personas en todo el mundo [1]. Más allá de reducir la calidad de vida de los pacientes, la FA también suele provocar síntomas como palpitaciones, aceleración de la frecuencia cardíaca, fatiga, dolor en el pecho [2]. Además, se trata del factor de riesgo más común de sufrir un accidente cerebrovascular isquémico [3], y puede agravar significativamente el estado de salud de los pacientes que sufren insuficiencia cardíaca, diabetes e hipertensión [2].

Según la duración y la recurrencia de los episodios arrítmicos, la FA se clasifica en varios grados que a menudo evolucionan a modos mantenidos en pocos años [4]. El punto de partida suele ser la FA paroxística, que se caracteriza por episodios que terminan automáticamente en una semana sin intervención externa. Cuando esto no es posible y se requiere intervención clínica para restaurar el RS, se clasifica como persistente. Si, a pesar de los intentos realizados para mantener el RS, la arritmia persiste después de un año, la FA se clasifica como persistente de larga permanencia. Finalmente, la etapa más avanzada de la enfermedad es la FA permanente, en la que el paciente y el médico deciden no hacer más esfuerzos para detener la arritmia debido a su fuerte permanencia [4]. Dado que esta arritmia conduce a cambios persistentes en la estructura y función auricular que promueven su perpetuación [5], es conveniente encontrar la mejor manera de restaurar el RS lo antes posible [6].

En este contexto, la ablación con catéter (AC) es hoy en día la estrategia de primera línea para el tratamiento de la FA persistente [7]. Aunque el procedimiento es inicialmente efectivo en la mayoría de pacientes con FA persistente, la arritmia recurre en aproximadamente el 40-50 % de ellos en el primer año [7]. Esta situación explica el interés clínico emergente en el desarrollo de predictores preoperatorios del resultado de la AC, ya que permitirían seleccionar a aquellos pacientes que podrían beneficiarse del tratamiento [8]. De esta manera, se podría reducir las tasas de hospitalización, limitar el número de procedimientos repetidos y reducir los costes del tratamiento [8].

Hasta el momento, se han explorado algunos predictores clínicos de la recurrencia de la FA después de la AC, co-

ISBN: 978-84-09-25491-0, pags. 476- 479

mo la diabetes, la duración de la arritmia y el tamaño de la aurícula izquierda, pero solo han proporcionado resultados controvertidos [9]. Como alternativa, también se han analizado algunos marcadores desde el registro electrocardiográfico (ECG). A este respecto, la frecuencia dominante (FD) calculada a partir de las ondas fibrilatorias (ondas f) reflejadas sobre dicho registro ha presentado una capacidad prometedora para anticipar el resultados de la AC, además de estar asociada con el grado de remodelado eléctrico presentado por las aurículas [10, 11]. Del mismo modo, la amplitud de las ondas f también ha demostrado ser un fuerte predictor del resultado de la AC, relacionándose con el nivel de remodelado estructural auricular [12, 13]. Sin embargo, todavía no hay ningún estudio que analice parámetros que combinen información de ambos dominios y, por lo tanto, el objetivo de este trabajo es explorar un conjunto novedoso de características de frecuencia y de amplitud de las ondas f destinadas a mejorar la predicción del resultado de la AC.

2. Método

2.1. Población de estudio

La población de estudio consistió en 51 pacientes (9 mujeres y 42 hombres) con FA persistente sometidos a ablación por catéter de radiofrecuencia en el Hospital Virgen de la Salud de Toledo. Toda terapia con medicamentos antiarrítmicos, excepto la amiodarona, se detuvo algunos días antes de la AC. Inicialmente, los pacientes fueron sedados con anestesia general o sedación consciente. A continuación, se utilizó el acceso venoso femoral para introducir catéteres y el acceso a la aurícula izquierda se logró mediante punción transeptal. La anticoagulación se mantuvo a través de un bolo de heparina inicial, y la administración adicional de heparina se basó en la monitorización del tiempo de coagulación activado durante todo el procedimiento. El aislamiento eléctrico de las venas pulmonares consiste en la creación de límites eléctricamente impenetrables que rodeen las venas pulmunares (VP) y se llevó a cabo administrando lesiones de ablación, posicionando el catéter y aplicando corriente de radiofrecuencia punto por punto durante al menos 30 segundos, cuya localización se determinó con un catéter de mapeo. El procedimineto finalizó cuando todas las VPs fueron aisladas con éxito o tras restablecer el RS mediante cardioversión eléctrica en el caso en el que la FA se mantuviese al final de la AC.

El procedimiento fue inicialmente exitoso en todos los pacientes, los cuales tampoco sufrieron ninguna complicación durante su monitorización de algunas horas tras el procedimiento y durante un seguimiento de nueve meses. Después de este tiempo, 30 pacientes mantuvieron RS y los 21 restantes recayeron a FA. Se debe notar que todos los pacientes recibieron anticoagulantes y fármacos antiarrítmicos de acuerdo al juicio clínico del experto siguiendo el caso.

2.2. Adquisición y preprocesado de la señal de ECG

La señal de ECG estándar de 12 derivaciones se grabó continuamente durante un tiempo de entre 6 y 1361 segundos,

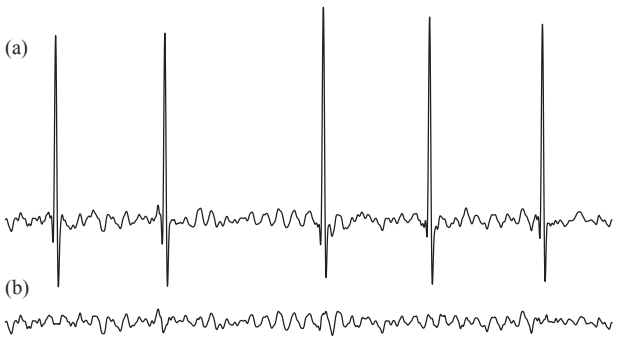


Figura 1. *Ejemplo de un intervalo de ECG (a) junto con las ondas f extraídas usando un método previamente publicado de cancelación del complejo QRST (b) [17].*

en función del paciente, justo antes de la AC. Dicha señal fue adquirida con una frecuencia de muestreo de 977 Hz y una resolución de 16 bits, pero se remuestreó a 1 kHz. Solo se analizó la derivación V1 porque refleja las ondas f más grandes en comparación con la actividad ventricular [14]. Esta derivación fue preprocesada para eliminar la deriva de línea base, la interferencia de la red eléctrica y el ruido de alta frecuencia. De forma más precisa, el desplazamiento de la línea de base se estimó usando un filtro paso bajo con una frecuencia de corte de 0,8 Hz y luego se substrajo de la señal original [15]. Seguidamente, la interferencia de la red eléctrica se eliminó por medio de un algoritmo basado en la transformada wavelet estacionaria [16]. Aunque este método también fue capaz de eliminar la mayor parte del ruido de alta frecuencia, se utilizó un filtro paso bajo con una frecuencia de corte de 70 Hz para obtener la señal lo más limpia posible [15].

Por otro lado, para analizar las ondas f, primeramente se extrajeron del ECG preprocesado mediante el uso de un método previamente publicado de cancelación QRST [17]. Como un ejemplo, la Figura 1 muestra las ondas f obtenidas de un intervalo de ECG típico.

2.3. Caracterización de las ondas f

La señal que contiene las ondas f para cada paciente se dividió en segmentos no superpuestos de 6 segundos de longitud, los cuales se caracterizaron en términos de los parámetros que se describen a continuación. Así pues, finalmente se analizaron un total de 204 segmentos, 84 de pacientes que recayeron a FA y 120 de aquellos que mantuvieron RS durante el seguimiento.

Como referencia, para cada intervalo de ECG, la FD se estimó como la frecuencia con la mayor amplitud de la densidad espectral de potencia (DEP) dentro del rango de 3–12 Hz [10, 11]. La DEP de las ondas f se calculó utilizando el Periodograma de Welch, con una ventana de Hamming de 3.000 puntos de longitud, una superposición del 50 % entre ventanas adyacentes y una transformada rápida de Fourier de 6.000 puntos. La amplitud normalizada de las ondas f (AOFn) se obtuvo expresando la amplitud de las ondas f (AOF) como un porcentaje de la magnitud del pico R para evitar el efecto de aspectos, tales como el uso de una diferente ganancia durante el registro, la

diferente impedancia de la interfaz piel-electrodo en cada paciente, la diferente conductividad de los tejidos, etc. Al igual que en trabajos anteriores [18], la AOF se estimó como el valor cuadrático medio de las ondas f.

Además de estos parámetros, se analizaron tres nuevos índices que combinan información de los dominios del tiempo y la frecuencia de la ondas f, los cuales son la entropía espectral (EE), la medida de la planicidad espectral (MPE) y la suma ponderada de la amplitud espectral (SPAE). Concretamente, la EE cuantifica la complejidad espectral de las ondas f calculando la dispersión de su distribución espectral [19]. Así, la DEP de las ondas f se obtuvo como se describió anteriormente y, después de su normalización para obtener una función de probabilidad con un área unitaria, se calculó la entropía de Shannon [19].

Por otro lado, la MPE proporciona una medida de cómo se distribuye el contenido espectral de una señal [20]. Se obtuvo dividiendo las medias geométrica y aritmética de la DEP de las ondas f. El índice oscila entre 0 y 1, de modo que los valores más bajos se asocian con una concentración espectral más alta en un número relativamente pequeño de bandas de frecuencia, mientras que los valores más altos sugieren distribuciones espectrales con una cantidad similar de potencia en todas las bandas de frecuencia [20].

Finalmente, el índice SPAE representa una suma ponderada de las amplitudes en el dominio espectral [21]. Se calculó como la suma de los productos de las frecuencias individuales y sus amplitudes, obteniendo la DEP de las ondas f, tal y como en las métricas anteriores [21].

2.4. Análisis estadístico

Para todos los índices analizados, los resultados se expresan en media $\pm$ desviación típica y las diferencias estadísticas entre los grupos de pacientes se analizan mediante un test t de Student. Un valor de significación estadística de p $< 0{,}05$ se consideró estadísticamente significativo.

Por otro lado, la capacidad de cada característica para discriminar entre los pacientes que mantuvieron RS y aquellos que recayeron a FA se evaluó mediante una curva característica operativa del receptor (ROC). Esta curva es el resultado de trazar la fracción de verdaderos positivos, que son los positivos correctamente clasificados como tales (sensibilidad), frente a la fracción de falsos positivos, que son los positivos clasificados como negativos (1- especificidad), para varios umbrales establecidos. Se consideró la sensibilidad (Se) como el porcentaje de pacientes que recayeron a FA correctamente clasificados. Del mismo modo, la tasa de pacientes que mantuvieron RS correctamente clasificados se consideró como la especificidad (Es). Se seleccionó el umbral óptimo como aquel que proporcionó el mejor equilibrio entre Se y Es, aunque esto implicase que no se lograra el porcentaje más alto de pacientes clasificados correctamente, es decir, la máxima exactitud (Ex). Finalmente, el área bajo la curva ROC (AROC) también se obtuvo como una medida global del rendimiento de una variable en todos los umbrales de clasificación posibles.

Para mejorar la clasificación entre los dos grupos de pacientes, también se llevó a cabo un análisis lineal discriminante (ALD). La selección de variables se realizó mediante un enfoque progresivo incluyendo en cada paso la característica que llevó a maximizar la traza de Lawley-Hotelling.

Tabla 1. *Valores medios y desviación típica obtenidos para las métricas analizadas en los segmentos de ambos grupos*

	Grupo de pacientes		
Índice	Mantiene RS	Recae a FA	p-valor
FD (Hz)	$6{,}01 \pm 1{,}47$	$6{,}44 \pm 1{,}25$	0,048
AOFn (%)	$6{,}47 \pm 4{,}45$	$5{,}14 \pm 2{,}98$	0,040
EE (sin unidades)	$0{,}77 \pm 0{,}08$	$0{,}74 \pm 0{,}08$	0,001
MPE (%)	$37{,}07 \pm 13{,}82$	$32{,}62 \pm 13{,}86$	0,012
SPAE (μV $\cdot$ Hz)	$189{,}34 \pm 82{,}74$	$231{,}83 \pm 85{,}71$	$< 0{,}001$

Tabla 2. *Resultados de clasificación obtenidos por los parámetros calculados a partir de las ondas f*

Índice	Se (%)	Es (%)	Ex (%)	AROC (%)
FD	57,14	57,27	57,21	57,38
AOFn	55,45	54,76	55,15	58,63
EE	61,82	61,9	61,85	63,76
MPE	60,00	60,71	60,31	60,55
SPAE	60,71	60,91	60,82	64,56

3. Resultados

En la Tabla 1 se muestran los valores medios y la desviación típica para los índices calculados a partir de los 204 segmentos del ECG de 6 segundos de longitud. Como se puede observar, todos ellos presentan diferencias estadísticamente significativas entre ambos grupos de pacientes, obteniendo p-valores inferiores a 0,05. Además, mientras que la FD y la SPAE proporcionaron valores medios más altos para los segmentos de los pacientes que recayeron a FA, los índices restantes devolvieron valores más bajos en comparación con aquellos segmentos de los pacientes que mantuvieron un RS durante el seguimiento.

Con respecto a la clasificación entre segmentos de ambos grupos de pacientes, la Tabla 2 muestra los valores de Se, Es, Ex y AROC logrados por todas las métricas, observándose valores limitados entre el 55 y el 65 %. No obstante, la combinación de AOFn, EE y SPAE a través de un ALD alcanzó mejoras del 15-23 % con respecto a los índices individuales, ya que se obtuvieron valores de Se, Es, Ex y AROC de 77,38 %, 77,27 %, 77,2 %, y 79,79 %, respectivamente.

4. Discusión

La alta tasa de ineficiencia a largo plazo de la AC en FA persistente ha aumentado el interés clínico en la predicción del resultado del procedimiento para optimizar la selección de pacientes y, así, permitir enfoques personalizados, reducir los costes del tratamiento y limitar los riesgos asociados con la intervención [8]. Con este fin, la caracterización avanzada tiempo-frecuencia de las ondas f en términos de EE, MPE y SPAE se ha abordado por primera vez en el presente trabajo. Los tres índices han mostrado una mejor correlación con el resultado de la AC que otras métricas comunes, tales como la FD y la AOFn. De hecho, han exhibido mejoras de entre el 2 y 8 % en los va-

lores de Ex y AROC. Además, la combinación basada en un ALD de algunas de estas métricas también aumentó la precisión diagnóstica en más del 15 % con respecto a todos los parámetros individuales.

Los valores medios obtenidos por estos índices para cada grupo de pacientes también concuerdan con hallazgos previos. De hecho, la EE y la MPE presentaron valores más altos para los pacientes que mantuvieron RS durante el seguimiento, lo que sugiere una distribución espectral más uniforme de las ondas f, debido a la presencia de varios armónicos de la FD. También, algunos trabajos previos han revelado que la presencia de ondas f organizadas es indicativa de una mayor probabilidad de terminación espontánea de la FA paroxística, así como de un resultado exitoso de la cardioversión eléctrica [18]. Por otro lado, la SPAE presentó valores más altos para los pacientes que recayeron a FA. Debido a que este índice se calcula como la suma del producto de las frecuencias individuales y sus amplitudes, la frecuencia parece predominar sobre la amplitud. Este resultado es consistente con el hallazgo de que la FD y su primer armónico fueron más predictivos del éxito de la AC a medio plazo que sus amplitudes [10].

5. Conclusiones

El análisis de algunas características tiempo-frecuencia de las ondas f reflejadas en el ECG antes de la AC ha proporcionado información clínicamente útil sobre el resultado del procedimiento. De hecho, la combinación basada en un ALD de varias de estas variables ha mejorado significativamente la predicción conseguida con otras medidas de tiempo o frecuencia bien establecidas, proporcionando así nuevos conocimientos sobre el remodelado del sustrato auricular antes del procedimiento. No obstante, se deberán realizar más estudios con bases de datos más amplias en el futuro para corroborar estos resultados.

Agradecimientos

Trabajo financiado por los proyectos DPI2017-83952-C3 MINECO/AEI/FEDER, EU, SBPLY/17/180501000411 de la Junta de Comunidades de Castilla-La Mancha y AICO/2019/036 de la Generalitat Valenciana. Además, Pilar Escribano Cano tiene una beca de investigación de posgrado y un posterior contrato predoctoral 2020-PREDUCLM-15540, ambos por la Universidad de Castilla-La Mancha.

Referencias

[1] Lippi G, Sanchis-Gomar F, Cervellin G. Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge. Int J Stroke Jan 2020; 1747493019897870.

[2] Platonov PG, Corino VD. A clinical perspective on atrial fibrillation. In Atrial Fibrillation from an Engineering Perspective. Springer, 2018; 1–24.

[3] Warmus P, Niedziela N, Huć M, Wierzbicki K, Adamczyk-Sowa M. Assessment of the manifestations of atrial fibrillation in patients with acute cerebral stroke - a single-center study based on 998 patients. Neurol Res Jun 2020; 42(6):471–476.

[4] Kirchhof P, Benussi S, Kotecha D, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. Eur Heart J 10 2016; 37(38):2893–2962.

[5] Schotten U, Dobrev D, Platonov PG, Kottkamp H, Hindricks G. Current controversies in determining the main mechanisms of atrial fibrillation. J Intern Med May 2016; 279(5):428–38.

[6] Nattel S, Guasch E, Savelieva I, et al. Early management of atrial fibrillation to prevent cardiovascular complications. Eur Heart J Jun 2014;35(22):1448–56.

[7] Schmidt B, Brugada J, Arbelo E, 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 Apr 2020;22(4):558–566.

[8] Walsh K, Marchlinski F. Catheter ablation for atrial fibrillation: current patient selection and outcomes. Expert Rev Cardiovasc Ther Sep 2018;16(9):679–692.

[9] Dretzke J, Chuchu N, Agarwal R, Herd C, Chua W, Fabritz L, Bayliss S, Kotecha D, Deeks JJ, Kirchhof P, Takwoingi Y. Predicting recurrent atrial fibrillation after catheter ablation: a systematic review of prognostic models. Europace May 2020;22(5):748–760.

[10] Alcaraz R, Hornero F, Rieta JJ. Electrocardiographic spectral features for long-term outcome prognosis of atrial fibrillation catheter ablation. Ann Biomed Eng Nov 2016; 44(11):3307–3318.

[11] Szilágyi J, Walters TE, Marcus GM, et al. Surface ECG and intracardiac spectral measures predict atrial fibrillation recurrence after catheter ablation. J Cardiovasc Electrophysiol 10 2018;29(10):1371–1378.

[12] Nault I, Lellouche N, Matsuo S, et al. Clinical value of fibrillatory wave amplitude on surface ECG in patients with persistent atrial fibrillation. J Interv Card Electrophysiol Oct 2009;26(1):11–9.

[13] Cheng Z, Deng H, Cheng K, Chen T, Gao P, Yu M, Fang Q. The amplitude of fibrillatory waves on leads aVF and V1 predicting the recurrence of persistent atrial fibrillation patients who underwent catheter ablation. Ann Noninvasive Electrocardiol Jul 2013;18(4):352–8.

[14] Petrėnas A, Marozas V, Sörnmo L. Lead systems and recording devices. In Atrial Fibrillation from an Engineering Perspective. Springer, 2018; 25–48.

[15] Sörnmo L, Laguna P. ECG Signal Processing in Bioelectrical Signal Processing in Cardiac and Neurological Applications. 1st edition. Elsevier Academic Press, 2005. ISBN 978-0-12-437552-9.

[16] García M, Martínez-Iniesta M, Ródenas J, Rieta JJ, Alcaraz R. A novel wavelet-based filtering strategy to remove powerline interference from electrocardiograms with atrial fibrillation. Physiol Meas 11 2018;39(11):115006.

[17] Alcaraz R, Rieta JJ. Adaptive singular value cancelation of ventricular activity in single-lead atrial fibrillation electrocardiograms. Physiol Meas Dec 2008;29(12):1351–69.

[18] Alcaraz R, Hornero F, Rieta JJ. Noninvasive time and frequency predictors of long-standing atrial fibrillation early recurrence after electrical cardioversion. Pacing Clin Electrophysiol Oct 2011;34(10):1241–50.

[19] Uldry L, Van Zaen J, Prudat Y, Kappenberger L, Vesin JM. Measures of spatiotemporal organization differentiate persistent from long-standing atrial fibrillation. Europace Aug 2012;14(8):1125–31.

[20] Eftestol T, Sunde K, Ole Aase S, Husoy JH, Steen PA. Predicting outcome of defibrillation by spectral characterization and nonparametric classification of ventricular fibrillation in patients with out-of-hospital cardiac arrest. Circulation Sep 2000;102(13):1523–9.

[21] Ristagno G, Mauri T, Cesana G, et al. Amplitude spectrum area to guide defibrillation: a validation on 1617 patients with ventricular fibrillation. Circulation Feb 2015; 131(5):478–87.

Transformada wavelet estacionaria y redes neuronales para la detección de pulso en episodios de parada cardiorrespiratoria

Erik Alonso González[1], Unai Irusta Zarandona[2], Elisabete Aramendi Ecenarro[2], Mohamud R. Daya[3]

[1] Departamento de Matemática Aplicada, Universidad del País Vasco (UPV/EHU), Bilbao, España, erik.alonso@ehu.eus

[2] Departamento de Ingeniería de Comunicaciones, Universidad del País Vasco (UPV/EHU), Bilbao, España, unai.irusta@ehu.eus, elisabete.aramendi@ehu.eus

[2] Departamento de Emergencias, Oregon Health & Science University (OHSU), Portland, OR, EEUU, dayam@ohsu.edu

Resumen

La disponibilidad de un detector automático de pulso durante un episodio de parada cardiorrespiratoria extrahospitalaria (PCREH) permitiría una rápida identificación de la parada cardiaca y una temprana detección del retorno de la circulación espontánea. El objetivo de este estudio fue desarrollar un algoritmo de detección de pulso basado en el electrocardiograma (ECG) y la impedancia torácica (IT), las señales que registra la mayoría de los desfibriladores comerciales. La base de datos de estudio se compuso de 456 segmentos de 5 s que contenían las señales ECG e IT. Un total de 281 segmentos correspondieron a ritmos con pulso (RP) mientras que 175 correspondieron a actividad eléctrica sin pulso (AESP). Primero, la componente circulatoria de la impedancia (CCI) fue extraída de la señal de IT a través de la transformada wavelet estacionaria. Después, se extrajeron características para la discriminación RP/AESP del ECG preprocesado (0.5–30 Hz), de la CCI y de su primera derivada. Estas características fueron introducidas en una red neuronal que clasificó el segmento como RP/AESP. Un procedimiento basado en validación cruzada de 5-subconjuntos quasi-estratificada en términos de paciente fue usada para la selección del mejor conjunto de características y para optimizar y testear la red neuronal. Este procedimiento se repitió 50 veces para estimar la distribución estadística de las métricas de rendimiento del método. La solución óptima consistió en una red neuronal basada en 3 características y con 20 neuronas en su capa oculta que obtuvo una media (desviación estándar) en términos de sensibilidad, especificidad, precisión balanceada y precisión de 90.2% (1.9), 90.9% (1.2), 90.6% (0.9) y 90.5% (1.1), respectivamente. Se ha desarrollado un detector de pulso para uso en PCREH que hace uso de las señales disponibles en los desfibriladores.

1. Introducción

La parada cardiorrespiratoria extra-hospitalaria (PCREH) es una de las principales causas de muerte en los países industrializados ya que de media es responsable de entre 83.7-95.9 muertes al año por cada 100000 habitantes [1]. Un reconocimiento precoz de la parada cardiorrespiratoria es crucial para la supervivencia del paciente ya que posibilita, por un lado, un rápido inicio del masaje de resucitación cardiopulmonar (RCP) y por otro, una desfibrilación temprana la cual se ha comprobado que mejora la probabilidad de supervivencia [2,3]. Las actuales guías de resucitación recomiendan evaluar la ausencia de 'respiración normal' como indicador de falta de circulación, y por tanto, parada cardiorrespiratoria [4]. Desafortunadamente, se ha comprobado que este procedimiento no es fiable [5] e introduce interrupciones en el masaje RCP que son perjudiciales para la supervivencia del paciente [6]. Por tanto, la implementación en los actuales desfibriladores de un detector fiable de detección de pulso contribuiría a una pronta y rápida identificación de la parada cardiorrespiratoria, y minimizaría las interrupciones en el masaje RCP para confirmar la ausencia de circulación. Además, también permitiría la rápida detección del retorno de la circulación espontánea (RCE) lo que a su vez conduciría a la inmediata iniciación de los cuidados pos-resucitación.

En esencia, la detección de pulso en la parada cardiorrespiratoria consiste en discriminar entre dos tipos de ritmos organizados: actividad eléctrica sin pulso (AESP) y ritmo generador de pulso (RP). Ambos ritmos presentan un ECG con complejos QRS y aparentemente normal. Sin embargo, mientras que los ritmos RP muestran una actividad eléctrica y mecánica normal, durante la AESP se produce una disociación electromecánica de manera que el corazón pierde su función contráctil y por lo tanto, no existe un flujo sanguíneo efectivo. La diferencia entre los ritmos RP y AESP no es siempre perceptible en el ECG. Sin embargo, la mayoría de los desfibriladores comerciales registran además del ECG, la señal de impedancia torácica (IT) a través de los parches de desfibrilación. Ésta última puede ayudar en la discriminación de estos dos ritmos ya que refleja una componente circulatoria de la impedancia (CCI) que consiste en pequeñas fluctuaciones (<100 mΩ) correlacionadas con los complejos QRS en los ritmos RP, pero no en la AESP. La CCI está oculta entre otras componentes de señal de IT.

El objetivo de este estudio fue proponer un método de detección de pulso. Para ello primero se desarrolló un método para aislar la CCI de la señal IT y después se entrenó un clasificador basado en aprendizaje automático que hiciese uso de características extraídas de la CCI y del ECG para discriminar entre RP y AESP.

2. Materiales y métodos

2.1. Base de datos de estudio

Los datos usados en este estudio fueron extraídos de 49 pacientes de PCREH tratados por la agencia Tualatin Valley Fire & Rescue (Tigard, OR, EEUU) haciendo uso del monitor/desfibrilador Philips HeartStart MRx (Philips

ISBN: 978-84-09-25491-0, pags. 480- 483

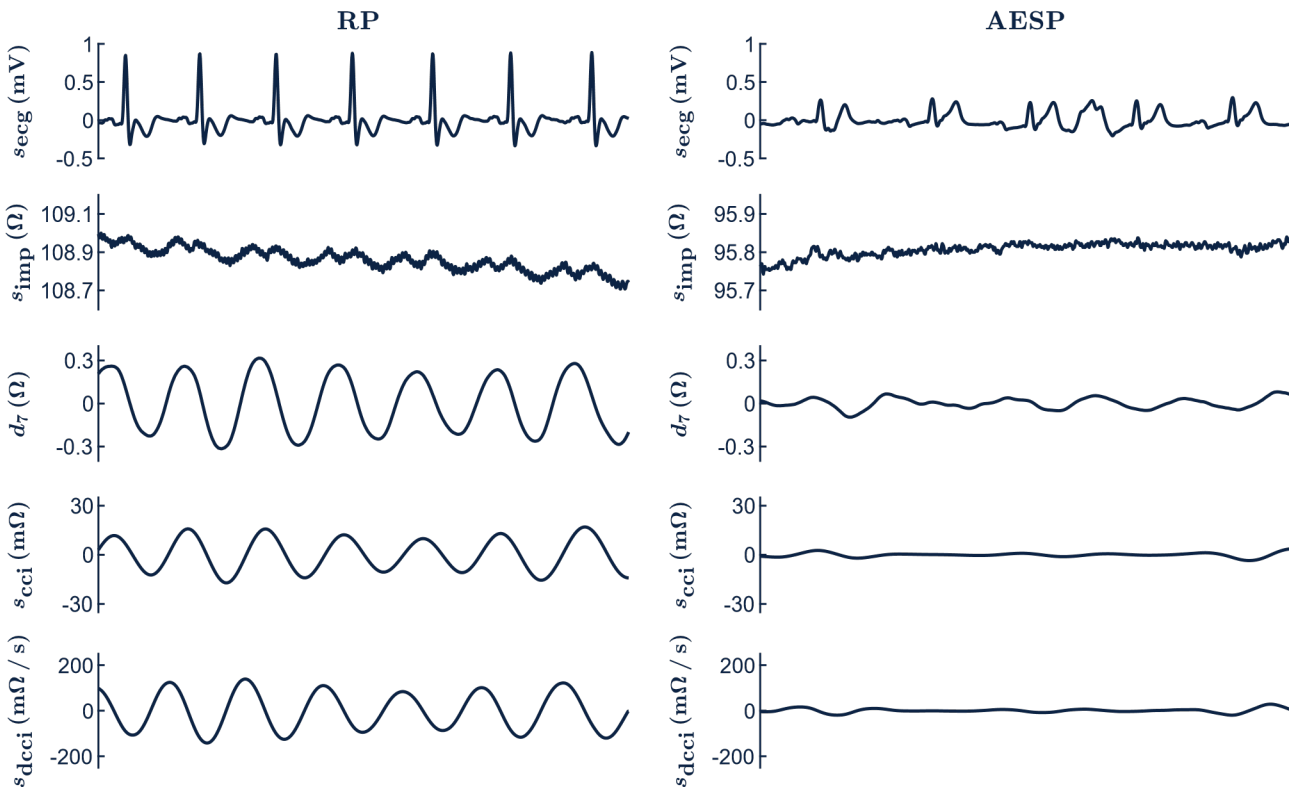


Figura 1. *Ejemplos de segmentos RP (izq.) y AESP (dcha.) de la base de datos de estudio en los que se ha llevado a cabo el análisis wavelet de la señal de IT. De arriba abajo, se muestran las señales ECG (s_{ecg}), la señal IT original (s_{imp}), una sub-banda relevante de la IT (d_7) para la estimación de la CCI, la CCI estimada (s_{cci}) y su primera derivada (s_{dcci}).*

HealthCare, MA, EEUU) en el año 2012. Cada episodio constó de las señales ECG, IT y capnografía muestreadas a 200 Hz. Los episodios fueron inspeccionados visualmente en busca de segmentos de 5 s que presentasen complejos QRS y ninguna interferencia debida a las compresiones torácicas o ventilaciones. Estos segmentos fueron anotados como RP o AESP por tres expertos en base a anotaciones RCE realizadas durante la actuación y a incrementos/decrementos abruptos en el capnograma que fueron indicadores de RCE y pérdida de RCE, respectivamente. La base de datos usada en este estudio se compuso de un total de 456 segmentos de los cuales 175 correspondieron a AESP y 281 a RP. Los paneles superiores de la Figura 1 muestran las señales ECG e IT para RP (izquierda) y AESP (derecha).

2.2. Extracción de la CCI

La CCI fue extraída de la señal IT, $s_{\mathrm{imp}}(n)$, a través de la transformada wavelet estacionaria (TWE) que consiste en la aplicación de un par de filtros paso-bajo, $h_{j-1}(n)$, y paso-alto, $g_{j-1}(n)$, para producir, en el siguiente nivel de descomposición, los coeficientes de aproximación, $a_j(n)$, y detalle, $d_j(n)$:

$$a_j(n) = h_{j-1}(n) * a_{j-1}(n) \qquad (1)$$

$$d_j(n) = g_{j-1}(n) * a_{j-1}(n) \qquad (2)$$

donde * denota la convolución, y $a_0(n) = s_{\mathrm{imp}}(n)$. 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.

Se hizo uso de una wavelet madre Biorthogonal 4.4 para descomponer $s_{\mathrm{imp}}(n)$ en $J - 7$ niveles de descomposición que produjeron los coeficientes de detalle $d_1 - d_7$ y el coeficiente de aproximación a_7. El ruido de los coeficientes de detalle $d_6 - d_7$ fue reducido haciendo uso de una función umbral suave. Para sintetizar la CCI, $s_{\mathrm{cci}}(n)$, se usaron los coeficientes de detalle d_6 y d_7 (banda 0.78-3.13 Hz) o solo d_7 (banda 0.78-1.56 Hz) para segmentos con frecuencia cardiaca superior e inferior a 93 latidos/min, respectivamente. El resto de coeficientes detalle y aproximación se igualaron a cero. La síntesis de $s_{\mathrm{cci}}(n)$ se llevó a cabo mediante la evaluación recursiva de la siguiente expresión:

$$a_{j-1}(n) = \frac{1}{2}(g'_j(n) * a_j(n) + h'_j(n) * d_j(n)) \qquad (3)$$

para $j = J, \ldots, 1$, donde $g'_j(n)$ y $h'_j(n)$ son los denominados filtros de síntesis y el coeficiente sintetizado $a_0(n)$ se corresponde con $s_{\mathrm{cci}}(n)$. La frecuencia cardíaca usada para la comparación con el umbral de 93 latidos/min en la síntesis de $s_{\mathrm{cci}}(n)$ fue calculada como la inversa de la duración media de los intervalos entre complejos QRS detectados a través de una versión modificada del detector QRS propuesto por Hamilton-Tompkins [7]. Los paneles inferiores de la Figura 1 muestran las señales más relevantes obtenidas a través del análisis wavelet de $s_{\mathrm{imp}}(n)$.

2.3. Extracción de características

Un total de 18 características fueron calculadas de $s_{\mathrm{ecg}}(n)$, la señal ECG filtrada paso-banda (0.5-30 Hz), de $s_{\mathrm{cci}}(n)$ y

de su primera derivada $s_{\mathrm{dcci}}(n)$. Las primeras 7 características fueron la amplitud pico-valle media, la desviación estándar de la amplitud pico-valle y el área media de $s_{\mathrm{cci}}(n)$ y $s_{\mathrm{dcci}}(n)$ [8], además de la fuzzy entropy (FuzzyEn) [9] de $s_{\mathrm{cci}}(n)$. Las siguientes 10 características fueron extraídas de $s_{\mathrm{ecg}}(n)$ y fueron el número de complejos QRS, la media y la desviación estándar del intervalo entre ondas R, la media y la desviación estándar de la amplitud de los complejos QRS [10], la mediana de la anchura de los complejos QRS [11], amplitude spectrum area (AMSA) [12], la energía en las bandas altas de frecuencia [13], FuzzyEn y signal length [11]. Finalmente, se calculó la potencia cruzada media entre $s_{\mathrm{ecg}}(n)$ y $s_{\mathrm{cci}}(n)$ [14]. Los detalles de estas características pueden encontrarse en las referencias detalladas arriba y su implementación Matlab en https://github.com/erik-alonso/PEA-PR-features.

2.4. Clasificación

El algoritmo de clasificación usado se basó en una red neuronal feed-forward. En la capa de entrada la red contó con K nodos, uno por cada característica usada para caracterizar los segmentos. La capa oculta estuvo compuesta por un número N de neuronas y la capa de salida por una única neurona. La función sigmoide fue la seleccionada como función de activación entre capas. Así, la neurona de la capa de salida proporcionó la probabilidad de que un determinado segmento correspondiera a un ritmo RP dado el vector de K características de dicho segmento. Los segmentos con probabilidades mayores a 0.5 fueron clasificados como RP, mientras que aquellos con probabilidad igual o inferior a 0.5 fueron clasificados como AESP.

2.5. Arquitectura del modelo y evaluación

La validación cruzada de 5 subconjuntos fue la elegida para entrenar el modelo y evaluar su rendimiento. Los subconjuntos fueron creados como particiones aleatorias quasi-estratificadas respecto a pacientes. Cada subconjunto mantuvo al menos el 90% de las proporciones RP/AESP de la base de datos original.

En cada una de las cinco iteraciones de la validación cruzada, 4 subconjuntos fueron dedicados primero a la selección de las K-características más discriminativas y después, a la optimización de la red neuronal basada en esas K-características seleccionadas. El último de los subconjuntos fue usado para evaluar el rendimiento del modelo en términos de sensibilidad (SE, capacidad para detectar correctamente RP), especificidad (ESP, capacidad para detectar correctamente AESP), precisión (P, proporción de ritmos correctamente clasificados) y precisión balanceada (PB, media de SE y ESP). Estas métricas de precisión se calcularon comparando las decisiones RP/AESP tomadas por el modelo con las anotaciones realizadas por los revisores expertos.

La selección de características se llevó a cabo mediante el algoritmo de selección incremental secuencial (SIS) y un clasificador basado en regresión logística multivariable (RLM) para evaluar el rendimiento del conjunto de características a la hora de discriminar entre RP y AESP. Tres subconjuntos fueron usados para entrenar el clasificador RLM y una para evaluar su rendimiento en términos de PB. El algoritmo SIS se ejecutó hasta que se seleccionó el mejor conjunto de K-características, usando la maximización de la PB como criterio para la inclusión de una nueva característica en el modelo.

La optimización de la red neuronal consistió por un lado en la actualización de los pesos y valores intercepto a través del algoritmo de propagación hacia atrás (*backpropagation*), y por otro, en encontrar el número óptimo de neuronas en la capa oculta que maximizase la PB de la red.

El procedimiento de validación cruzada de 5 subconjuntos se repitió 50 veces haciendo uso de diferentes particiones aleatorias para estimar la distribución estadística de las métricas de rendimiento. Todos los resultados se proporcionarán como media (desviación estándar, DE) de los valores obtenidos en las 50 repeticiones de los experimentos.

3. Resultados y discusión

La Figura 2 muestra la PB del algoritmo de detección de pulso en función del número de características incluidas en el modelo. Este análisis se llevó a cabo para una red neuronal con 10 neuronas en su capa oculta. La inclusión de nuevas características supuso una gran mejora en la PB hasta conseguir su máximo (90.5%) cuando tres características fueron incluidas en el modelo. Para modelos basados en cuatro o más características la PB se mantuvo alrededor del 90%

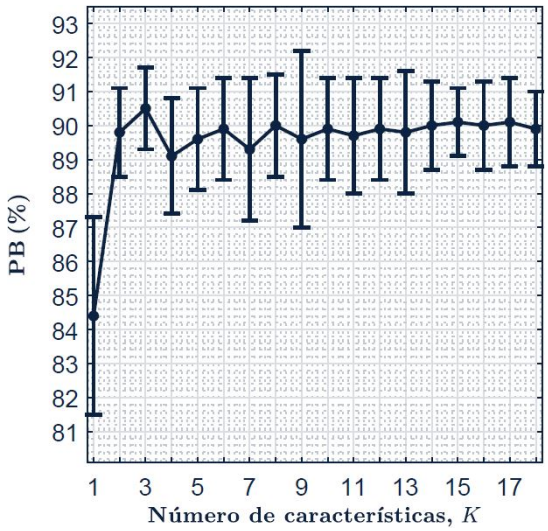


Figura 2. *Media (DE) de la PB del algoritmo de detección de pulso en las 50 réplicas del procedimiento de validación cruzada en función del número de características, K, incluidas en el modelo.*

La Figura 3 muestra la PB obtenida por el algoritmo de detección de pulso en función del número de neuronas, N, en capa oculta de la red neuronal. En este caso, el análisis se realizó para un algoritmo de detección de pulso basado en tres características. La máxima PB mostrada por el algoritmo de detección de pulso fue 90.6% (0.9) y fue alcanzada cuando la red neuronal contó con 20 neuronas en

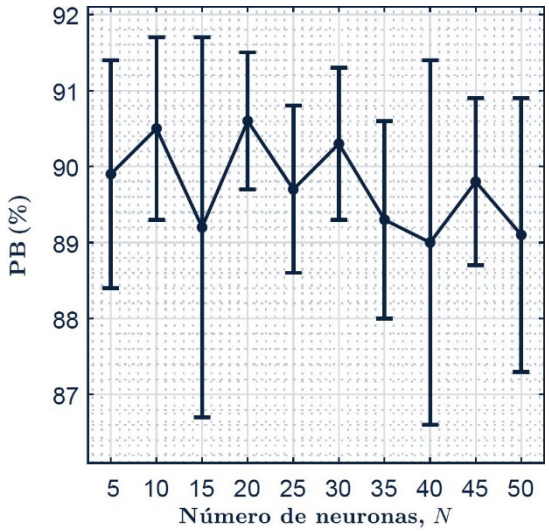


Figura 3.*Media (DE) de la PB del algoritmo de detección de pulso en las 50 réplicas del procedimiento de validación cruzada en función del número de neuronas, N, en la capa oculta de la red neuronal.*

su capa oculta. a partir de 30 se observa una tendencia descendente en la PB. Por lo tanto, el algoritmo de detección de pulso que mejor rendimiento en términos de PB mostró en las 50 réplicas del proceso de validación cruzada fue aquel basado en una red neuronal alimentada por tres características y con 20 neuronas en su capa oculta que presentó valores medios de SE, ESP, PB y P de 90.2% (1.9), 90.9% (1.2), 90.6% (0.9) y 90.5% (1.1), respectivamente.

4. Conclusiones

Este estudio ha propuesto un novedoso algoritmo de detección de pulso basado en las señales ECG e IT adquiridas por los desfibriladores comerciales. De esta manera, la incorporación del algoritmo en los desfibriladores únicamente requeriría de modificaciones software, sin necesidad de realizar cambios en su hardware. El algoritmo se basa en la TWE para extraer la CCI de la IT de manera que maximiza la capacidad de ésta para discriminar entre ritmos AESP y RP. Se extrajeron un total de 18 características de las señales ECG, CCI y su primera derivada, y se introgujeron en una red neuronal que se encargó de decidir si esas características correspondían a un ritmo AESP o RP. La versión optimizada de este modelo presenta resultados satisfactorios. Sin embargo, una base de datos más amplia permitiría corroborar estas conclusiones.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el Ministerio de Ciencia, Innovación y Universidades a través del proyecto RTI2018-101475-BI00, conjuntamente con el Fondo Europeo de Desarrollo Regional (FEDER), y en parte por el Gobierno Vasco por medio del proyecto IT-1229-19.

Referencias

[1] Berdowski J, Berg RA, Tijssen JGP, et al. Global incidences of out-of-hospital cardiac arrest and survival rates: Systematic review of 67 prospective studies. *Resuscitation*, vol. 81(11), 2010, pp. 1479–1487.

[2] Wissenberg M, Lippert FK, Folke F, et al. Association of National Initiatives to Improve Cardiac Arrest Management With Rates of Bystander Intervention and Patient Survival After Out-of-Hospital Cardiac Arrest. *JAMA*, vol. 310(13), 2013, pp.1377–1384.

[3] Blom MT, Beesems SG, Homma PCM, et al. Improved survival after out-of-hospital cardiac arrest and use of automated external defibrillators. Circulation, vol.130(21), 2014, pp. 1868–1875.

[4] Perkins GD, Handley AJ, Koster RW, et al, Adult basic life support and automated external defibrillation section Collaborators. European Resuscitation Council Guidelines fornResuscitation 2015: Section 2. Adult basic life support and automated external defibrillation. *Resuscitation*, vol. 95, 2015, pp. 81–99.

[5] Ruppert M, Reith MW, Widmann JH, et al. Checking for breathing: evaluation of the diagnostic capability of emergency medical services personnel, physicians, medical students, and medical laypersons. *Annals of Emergency Medicine* vol. 34(6), 1999, pp. 720–729.

[6] Vaillancourt C, Everson-Stewart S, Christenson J, et al, Resuscitation Outcomes Consortium Investigators. The impact of increased chest compression fraction on return of spontaneous circulation for out-of-hospital cardiac arrest patients not in ventricular fibrillation. *Resuscitation*, vol. 82(12), 2011, pp. 1501–1507.

[7] Hamilton PS and Tompkins WJ, Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. *IEEE Transactions on Biomedical Engineering*, vol. 33(12), 1986, pp. 1157–1165.

[8] Ruiz J, Alonso E, Aramendi E, et al. Reliable extraction of the circulation component in the thoracic impedance measured by defibrillation pads. *Resuscitation*, vol. 84(10), 2013, pp. 1345–1352.

[9] Chen W, Wang Z, Xie H, et al, Characterization of surface EMG signal based on fuzzy entropy, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15(2), 2007, pp. 266-272.

[10] Alonso E, Aramendi E, Daya M, et al, Circulation detection using the electrocardiogram and the thoracic impedance acquired by defibrillation pads. *Resuscitation*, vol. 99, 2016, pp. 56-62.

[11] Risdal M, Aase SO, Kramer-Johansen J, et al. Automatic identification of return of spontaneous circulation during cardiopulmonary resuscitation, *IEEE Transactions on Biomedical Engineering*, vol. 55(1), 2008, pp. 60-68.

[12] Ristagno G, Li Y, Fumagalli F, et al, Amplitude spectrum area to guide resuscitation-a retrospective analysis during out of-hospital cardiopulmonary resuscitation in 609 patients with ventricular fibrillation cardiac arrest, *Resuscitation*, vol. 84(12), 2013, pp. 1697-1703.

[13] Elola A, Aramendi E, Irusta U, et al. ECG-based pulse detection during cardiac arrest using random forest classifier, *Medical & Biological Engineering & Computing*, vol. 57(2), 2019, pp. 453-462.

[14] Ruiz JM, Ruiz de Gauna S, González-Otero DM, et al, Circulation assessment by automated external defibrillators during cardiopulmonary resuscitation, *Resuscitation*, vol. 128, 2018, pp. 158-163.

Detección automática de patología en imágenes de fondo de ojo utilizando técnicas de *deep learning*

R. Romero-Oraá[1,2], C. P. Muñoz-Zamarro[1], M. García[1,2], J. Oraá-Pérez[1,3], M. I. López Gálvez[1,2,3,4], R. Hornero Sánchez[1,2]

[1] Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, roberto.romero@gib.tel.uva.es, cristinapinarmz@outlook.es, margar@tel.uva.es, javier.oraa@gib.tel.uva.es, robhor@tel.uva.es

[2] Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN, España)

[3] IOBA (Instituto de Oftalmobiología Aplicada), Universidad de Valladolid, España

[4] Hospital Clínico Universitario de Valladolid, Valladolid, España, maribel@ioba.med.uva.es

Resumen

El análisis de retinografías permite diagnosticar importantes enfermedades oculares, tales como la retinopatía hipertensiva, el glaucoma, la retinopatía diabética y la degeneración macular asociada a la edad. Existen muy pocos trabajos previos que abarquen la detección general de signos patológicos y, en un contexto de cribado, es un desafío muy relevante. En este trabajo se propone un método basado en técnicas de deep learning para detectar automáticamente signos patológicos en imágenes de fondo de ojo. Para ello, se compararon 13 arquitecturas de redes neuronales convolucionales. Además, se utilizaron las técnicas data augmentation, transfer learning y fine-tuning. Se empleó una base de datos privada de 1000 retinografías, que se dividió en un conjunto de entrenamiento (800 imágenes), un conjunto de validación (100 imágenes) y un conjunto de test (100 imágenes). Los mejores resultados se obtuvieron con la arquitectura DenseNet-121, con la que se alcanzaron una precisión del 99%, una sensibilidad del 100% y una especificidad del 98%, con un tiempo medio de procesado de 7.72 segundos por imagen. Se puede considerar que el método propuesto podría ser de utilidad en un entorno clínico, facilitando el screening de diversas enfermedades oculares.

1. Introducción

Existe un gran número de patologías oculares que pueden dañar las estructuras normales del ojo y provocar la aparición de lesiones en la retina [1], [2]. Entre estas enfermedades, prevalecen la retinopatía diabética (RD), la retinopatía hipertensiva, el glaucoma, la degeneración macular asociada a la edad (DMAE), las cataratas, el edema macular diabético y la miopía, entre otras [3]. En estos casos, el diagnóstico precoz es clave para aplicar los tratamientos adecuados y evitar la pérdida de visión. La mejor forma de prevenir estas enfermedades es llevar a cabo exámenes periódicos del fondo de ojo. En este sentido, el método estándar de diagnóstico de las principales causas de ceguera es el análisis de retinografías. Sin embargo, el análisis manual de estas imágenes supone una elevada carga de trabajo para los especialistas en oftalmología. Los altos costes y la escasez de recursos humanos hacen inviable el control adecuado de la salud ocular de la población. Por ello, los sistemas automáticos de cribado de distintas patologías oculares reducirían mucho la carga de trabajo de los expertos, permitiendo el análisis rápido de grandes cantidades de retinografías, y mejorarían la atención sanitaria de los pacientes [3].

En la literatura pueden encontrarse varios métodos automáticos basados en el análisis de retinografías para la detección de alguna patología concreta. En particular, han recibido especial atención la RD, el glaucoma y la DMAE [3]. Sin embargo, existen muy pocos trabajos previos que abarquen la detección general de signos patológicos, independientemente de la patología de la que se trate. En un contexto de cribado, donde se pretende descartar de inmediato la gran mayoría de casos que suponen los controles, esta aproximación es la más interesante. Cualquier retinografía que presente signos patológicos de cualquier tipo requiere un seguimiento especializado y, en su caso, derivación a consulta. En este sentido, Koh et al. [4] propusieron un método de varias etapas. Utilizaron un histograma piramidal de gradientes orientados (PHOG) y el algoritmo *Speeded-Up Robust Features* (SURF) para describir la distribución de intensidad de la imagen. Después, utilizaron un clasificador *k-Nearest Neighbors* (*k*-NN) obteniendo un 96.21% de precisión. Este método se mejoró posteriormente utilizando un histograma piramidal de palabras visuales (PHOW) y un clasificador *Random Forest* [5] para obtener un 96.79% de precisión. Otros autores utilizaron *deep learning* [1], [6]. Choi et al [6] utilizaron una red neuronal convolucional (CNN) con arquitectura VGG-19 en combinación con la técnica *transfer learning* y un clasificador *Random Forest*, pero alcanzaron una precisión inferior al 85%. Quellec et al. [1] utilizaron *few-shot learning*, obteniendo los mejores resultados con la arquitectura CNN Inception-v3.

Nuestra hipótesis se basa en que las técnicas de *deep learning* son muy efectivas para detectar retinografías con signos patológicos, pudiéndose superar los resultados de los métodos convencionales. Por ello, en este trabajo se propone un nuevo método de clasificación de retinografías según la presencia o ausencia de patología. Para ello, se ha comparado el rendimiento de distintas arquitecturas CNN en combinación con el uso de las técnicas *transfer learning* y *data augmentation*. Hasta donde tenemos conocimiento, esta aproximación no se ha abordado previamente en la literatura con esta metodología.

ISBN: 978-84-09-25491-0, pags. 484- 487

2. Base de datos de retinografías

En este trabajo utilizamos una base de datos privada formada por 1000 retinografías procedentes del Programa de Cribado de RD del servicio público de Sanidad de Castilla y León (Sacyl). Estas imágenes se capturaron con el retinógrafo no midriático Topcon TRC-NW400, con un campo de visión de 45º y en formato DICOM convertido a JPEG de 24 bits. La mayoría de imágenes (782) tenían 1956×1934 píxeles de resolución y el resto (218) 4288×2848 píxeles. Un especialista determinó, para cada una, si presentaban signos patológicos. Entre las patologías encontradas, predominaba la RD y se incluían el glaucoma, la miopía, la DMAE, las cataratas y el edema macular. La base de datos se separó aleatoriamente en 3 conjuntos balanceados (mismo número de imágenes patológicas y de control). El conjunto de entrenamiento contenía 800 imágenes (400 patológicas y 400 de control) y los conjuntos de validación y test estaban formados por 100 imágenes cada uno (50 patológicas y 50 de control).

3. Métodos

El método propuesto parte de una etapa de preprocesado para normalizar las imágenes de entrada. Después, se construyeron 13 arquitecturas CNN que se utilizaron aplicando las técnicas *data augmentation*, *transfer learning* y *fine-tuning*.

3.1. Preprocesado

En primer lugar, con el objetivo de adaptar su resolución a la capa de entrada de cada modelo CNN construido, se redimensionaron todas las imágenes utilizando la resolución de entrada correspondiente a cada arquitectura. En segundo lugar, las imágenes se normalizaron en el intervalo [-1,1], lo que mejora el proceso de entrenamiento y da lugar a un modelo más estable [7].

3.2. Data augmentation

La técnica *data augmentation* permite generar imágenes sintéticas a partir de las imágenes reales disponibles. De esta manera, aumentamos el número de imágenes de entrenamiento, lo que resulta muy conveniente con el uso de redes profundas [8]. En este trabajo, se utilizó el aumento de datos en tiempo real, por lo que se fue generando un nuevo *batch* de imágenes para cada iteración. Por lo tanto, el número de imágenes obtenidas dependió del número de épocas. Para ello, se aplicaron trasformaciones simples sobre las 800 imágenes originales del conjunto de entrenamiento: rotaciones aleatorias en el rango [-50,+50] grados y volteos horizontales y verticales aleatorios [8].

3.3. Transfer learning y *fine-tuning*

La técnica *transfer learning* consiste en resolver un problema de *machine learning* en un dominio de interés particular con el conocimiento aprendido de los datos de entrenamiento de otro dominio de interés. De esta manera, no es necesario disponer de un conjunto de imágenes tan amplio como si se tuviera que entrenar la red desde cero. La forma más común de aplicar esta técnica es el uso de redes pre-entrenadas. Para todas las arquitecturas utilizadas en este trabajo, se partió de los pesos obtenidos con imágenes del proyecto ImageNet [9]. Esta base de datos se compone de más de 14 millones de imágenes pertenecientes a más de 20000 clases distintas y se ha utilizado con éxito en análisis de retinografías [6], [8], [10].

Además de partir de redes pre-entrenadas, se aplicó la técnica *fine-tuning*, que consiste en reentrenar todas o algunas de las capas finales de la CNN utilizando imágenes específicas del problema a resolver. De esta manera, los pesos se adaptan mejor a nuestro objetivo. En este trabajo, se reentrenaron todas las capas de nuestras CNNs [10].

3.4. Arquitecturas CNN

Las CNNs son redes neuronales con una estructura especial que es invariable tanto a la rotación como a la traslación. Normalmente se componen de capas convolucionales, capas de *pooling* y capas *fully-connected.* Son capaces de extraer de forma optimizada características representativas de un gran conjunto de imágenes en crudo [7], [8].

Con el objetivo de encontrar la arquitectura más adecuada para clasificar retinografías patológicas y no patológicas, se estableció una comparación entre 13 arquitecturas CNN distintas: Xception, VGG-16, VGG-19, ResNet-50, Inception-V3, Inception-ResNet-V2, MobileNet, MobileNet-V2, DenseNet-121, DenseNet-169, DenseNet-201, NASNet-Mobile y NASNet-Large [2]. Las arquitecturas seleccionadas han tenido éxito en problemas de clasificación de imagen y muchas de ellas se han utilizado con imágenes de fondo de ojo [2], [9].

Para adaptar todas estas arquitecturas a nuestro problema de clasificación binaria, reemplazamos las últimas 3 capas por 3 capas *fully connected* de 1024, 512 y 1 neurona, respectivamente [8]. Las dos primeras contaban con una función de activación ReLU [8]. Para la última, se empleó una función de activación sigmoide, que proporciona una salida entre 0 y 1 y puede interpretarse como una probabilidad a posteriori [7].

Se utilizaron los mismos hiperparámetros para todas las arquitecturas consideradas excepto para VGG-16. En la fase de *fine-tuning*, los modelos se entrenaron durante 100 épocas (*N*). Se aplicó la entropía cruzada binaria como función de pérdidas y el descenso de gradiente estocástico como algoritmo de optimización [8], [10]. La tasa de aprendizaje (*learning rate*, *lr*) se fijó en 0.001 y el valor de *momentum* en 0.9 [4,12]. Para evitar sobre-entrenamiento 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 [8]. Además, se estableció un tamaño de *batch* de 16 imágenes [8]. En el caso de VGG-16, fue necesario optimizar experimentalmente los valores de *lr* y *N* para que el modelo convergiera. De este modo, construimos 3 modelos distintos: (1) $lr = 0.0001$, $N = 100$; (2) $lr = 0.00005$, $N = 500$; (3) $lr = 0.0001$, $N = 500$.

El conjunto de entrenamiento se utilizó como datos de aprendizaje para ajustar los parámetros internos de las redes. El conjunto de validación nos permitió evaluar el rendimiento de cada modelo a cada época de entrenamiento. Esto nos permitió verificar el correcto aprendizaje y configurar el hiperparámetro *lr*, reduciendo su valor a medida que el modelo convergía. Por último, el

conjunto de test, formado por imágenes independientes, se utilizó para obtener los resultados de cada modelo una vez se había completado el entrenamiento.

4. Resultados y discusión

En este trabajo se han estudiado distintas arquitecturas CNN para detectar retinografías con signos patológicos. Adicionalmente, se han utilizado las técnicas *data augmentation*, *transfer learning* y *fine-tuning* para optimizar los resultados. Hasta donde sabemos, nunca se había comparado el rendimiento de estas arquitecturas con este objetivo.

Los modelos CNN entrenados se evaluaron en términos de precisión (PR), sensibilidad (SE) y especificidad (ES) utilizando las 100 imágenes del conjunto de test. Además, se calculó el tiempo medio que el algoritmo tardó en procesar cada imagen (*t*). Como se puede apreciar en la Tabla 1, se han obtenido resultados con alta precisión para todas las arquitecturas utilizadas. De hecho, no hay mucha diferencia entre los resultados de las distintas arquitecturas. Podemos observar que, para todos los casos, el valor de SE es del 100%, por lo que el sistema siempre tiende a clasificar correctamente las imágenes con signos patológicos. Este alto valor de SE es muy deseable en los sistemas de cribado, donde es mucho mayor el coste de fallar en la detección de un sujeto patológico que enviar a consulta a un sujeto sano. En términos de PR y ES, varias de las arquitecturas alcanzaron los valores más altos (PR=99.00% y ES=98.00%). Por lo tanto, debemos considerar el tiempo medio de procesado por imagen para elegir la arquitectura con mejores resultados. Este valor de tiempo estaba en el rango de 6 a 11 segundos, dependiendo de la arquitectura. Entre las arquitecturas con mayor PR, DenseNet-121 fue la más rápida de todas, con un tiempo medio de procesado de 7.72 segundos por imagen. Por lo tanto, se considera que esta sería la arquitectura óptima para nuestro problema.

Arquitectura CNN	PR (%)	SE (%)	ES (%)	*t* (s)
Xception	99.00	100	98.00	11.28
VGG-16				
lr =0.0001, *N* = 100	99.00	100	98.00	10.34
lr = 0.00005, *N* = 500	98.00	100	96.00	10.80
lr = 0.0001, *N* = 500	99.00	100	98.00	10.49
VGG-19	99.00	100	98.00	9.14
ResNet-50	99.00	100	98.00	8.50
Inception-V3	99.00	100	98.00	10.90
Inception-ResNet-V2	99.00	100	98.00	10.47
MobileNet	99.00	100	98.00	9.62
MobileNet-V2	96.00	100	92.00	6.82
DenseNet-121	99.00	100	98.00	7.72
DenseNet-169	99.00	100	98.00	9.15
DenseNet-201	96.00	100	92.00	9.08
NASNet-Mobile	98.00	100	96.00	9.12
NASNet-Large	99.00	100	98.00	11.17

Tabla 1. *Resultados sobre el conjunto de test*

Para evaluar la influencia de *transfer learning* en los resultados, se calcularon los valores de PR, SE y ES sobre el conjunto de test sin el uso de esta técnica. Para ello, se utilizó la arquitectura DenseNet-121 y se entrenó el modelo de nuevo. Igualmente, para evaluar la influencia de *fine-tuning*, hemos calculado los resultados reentrenando únicamente las 3 últimas capas agregadas, en lugar de todo el modelo. Estos resultados se recogen en la Tabla 2. Como se puede observar, el uso de las técnicas *transfer learning* y *fine-tuning* tiene un impacto muy positivo en los resultados. En el caso de la prueba sin *transfer learning*, los resultados son muy bajos porque el modelo ni siquiera llega a converger con *N*=100. Esto es debido a que los modelos de red profunda entrenados desde cero necesitan más épocas de entrenamiento y suelen alcanzar resultados inferiores. La prueba sin *fine-tuning* demuestra que esta técnica permite mejorar los resultados adaptando la red a cada problema concreto.

Los resultados obtenidos nos permiten proponer un método automático de diagnóstico general de patologías oculares basado en la arquitectura CNN DenseNet-121 y las técnicas *data augmentation*, *transfer learning* y *fine-tuning*. Aunque existen pocos métodos automáticos con el mismo objetivo, es posible establecer una comparación con algunos estudios, como se muestra en la Tabla 3. Cabe destacar que nuestros resultados superan a los obtenidos en otros estudios en términos de PR, SE y ES. No obstante, todos los trabajos han utilizado bases de datos distintas, por lo que tienen distinta representación de las patologías y la comparación debe hacerse con cautela. En el trabajo de Choi et ál. [6] también se utilizó una CNN junto con *transfer learning*. Sin embargo, se empleó la arquitectura VGG-19 encadenada a un clasificador *random forest* y no se aplicó *data augmentation*. Son destacables también los métodos de Koh et ál. [4], [5], que se basaron en técnicas

	PR (%)	SE (%)	ES (%)
Sin *transfer learning*	61.00	96.00	26.00
Sin *fine-tuning*	94.00	96.00	92.00
Con *transfer-learning* y *fine-tuning*	99.00	100	98.00

Tabla 2. *Resultados sobre el conjunto de test sin utilizar las técnicas transfer learning y fine-tuning*

Estudio	PR (%)	SE (%)	ES (%)
Koh et al., 2017 [4]	96.21	95.00	97.42
Choi et al., 2017 [6]	-	80.30	85.50
Koh et al., 2018 [5]	96.79	96.73	96.96
Método propuesto	**99.00**	**100**	**98.00**

Tabla 3. *Comparación de los resultados con otros estudios*

clásicas de procesado de imagen, lo que nos permite confirmar que las técnicas de *deep learning* pueden superar a los métodos convencionales.

A modo ilustrativo, en la Figura 1 se muestra la única imagen que falla el método propuesto. En la parte izquierda se puede apreciar una pequeña mancha oscura. En la parte derecha, se observa una pequeña mancha brillante junto a un vaso sanguíneo. Estas manchas pueden confundirse fácilmente con lesiones, por lo que pueden ser la causa de que el método automático haya detectado signos patológicos en la imagen.

Este estudio presenta algunas limitaciones. En primer lugar, únicamente se ha empleado una base de datos privada de retinografías evaluada por un experto. En futuros trabajos se tratará de aumentar el número de imágenes con los que entrenar la red y se utilizará la evaluación de varios expertos para minimizar sesgos y errores humanos sistemáticos. Además, se incluirán bases de datos adicionales para evaluar la robustez del método. En segundo lugar, se han utilizado los mismos hiperparámetros para todas las arquitecturas estudiadas. En estudios futuros puede ser interesante optimizar los hiperparámetros individualmente para cada arquitectura.

5. Conclusión

En este trabajo se ha propuesto un método automático basado en técnicas de *deep learning* para clasificar imágenes de fondo de ojo según la presencia o ausencia de patología. Se han comparado 13 arquitecturas CNN y la arquitectura DenseNet-121 ha alcanzado los mejores resultados. Asimismo, se ha demostrado que las técnicas *transfer learning* y *fine-tuning* permiten mejorar notablemente los resultados.

Se puede considerar que el método propuesto podría ser de utilidad en un entorno clínico, facilitando el *screening* y evaluación de diversas enfermedades oculares.

Agradecimientos

Este trabajo ha sido financiado por el ‘Ministerio de Ciencia, Innovación y Universidades’ y el ‘Fondo Social Europeo de Desarrollo Regional’ (FEDER) bajo los proyectos DPI2017-84280-R y PGC2018-098214-A-I00, por la ‘Comisión Europea’ y el FEDER bajo el proyecto 'Análisis y correlación entre la epigenética y la actividad cerebral para evaluar el riesgo de migraña crónica y episódica en mujeres' (Programa de cooperación Interreg V-A SpainPortugal POCTEP 2014-2020'), y por el CIBER en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) a través del Instituto Carlos III, cofinanciado por FEDER. R. Romero-Oraá disfruta de un contrato predoctoral cofinanciado por el Fondo Social Europeo, el Programa Operativo de Castilla y León, y la Junta de Castilla y León, a través de la Consejería de Educación.



Figura 1.*Retinografía sin signos patológicos que supone el único fallo de detección por el método propuesto*

Referencias

[1] G. Quellec, M. Lamard, P. H. Conze, P. Massin, and B. Cochener, “Automatic detection of rare pathologies in fundus photographs using few-shot learning,” *Med. Image Anal.*, 2020.

[2] S. Sengupta, A. Singh, H. A. Leopold, T. Gulati, and V. Lakshminarayanan, “Ophthalmic diagnosis using deep learning with fundus images – A critical review,” *Artificial Intelligence in Medicine*, vol. 102. Elsevier B.V., p. 101758, 01-Jan-2020.

[3] M. D. Abramoff, M. K. Garvin, and M. Sonka, “Retinal imaging and image analysis,” *IEEE Rev. Biomed. Eng.*, vol. 3, pp. 169–208, 2010.

[4] J. E. W. Koh, E. Y. K. Ng, S. V. Bhandary, A. Laude, and U. R. Acharya, “Automated detection of retinal health using PHOG and SURF features extracted from fundus images,” *Appl. Intell.*, 2018.

[5] J. E. W. Koh, E. Y. K. Ng, S. V. Bhandary, Y. Hagiwara, A. Laude, and U. R. Acharya, “Automated retinal health diagnosis using pyramid histogram of visual words and Fisher vector techniques,” *Comput. Biol. Med.*, 2018.

[6] J. Y. Choi, T. K. Yoo, J. G. Seo, J. Kwak, T. T. Um, and T. H. Rim, “Multi-categorical deep learning neural network to classify retinal images: A pilot study employing small database,” *PLoS One*, 2017.

[7] C. Bishop, *Neural Networks for Pattern Recognition*, 1st ed., vol. 1995. New York: Oxford University Press, 1995.

[8] G. T. Zago, R. V. Andreão, B. Dorizzi, and E. O. Teatini Salles, “Retinal image quality assessment using deep learning,” *Comput. Biol. Med.*, 2018.

[9] Jia Deng, Wei Dong, R. 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*, 2009, pp. 248–255.

[10] A. S. Coyner *et al.*, “Deep Learning for Image Quality Assessment of Fundus Images in Retinopathy of Prematurity.,” *AMIA ... Annu. Symp. proceedings. AMIA Symp.*, vol. 2018, pp. 1224–1232, 2018.

[11] R. J. Chalakkal, W. H. Abdulla, and S. S. Thulaseedharan, “Quality and content analysis of fundus images using deep learning,” *Comput. Biol. Med.*, 2019.

ÍNDICE DE AUTORES

A

Abella, M. 73, 85, 206, 273, 285, 439, 447, 459
Acha, B. 468
Aguado, H. J. 397
Aguiló, J. 256
Albaladejo-Belmonte, M. 240
Alberola-Rubio, J. 240
Albert-Smet, I. 265
Alcaraz Martínez, R. 2, 14, 105, 109, 117, 319, 327, 335, 409, 422, 430, 476
Aldazabal, J. 40, 389
Alegre, M. 57, 235
Aleixandre Soriano, I. 77
Alférez, S. 157
Alonso González, E. 480
Alonso López, J. F. 323
Álvarez González, D. 252, 315, 340
Álvarez Rodríguez, J. 145, 464
Álvarez Rodríguez, M. 190
Amado Forero, L. 65
Aragón Arqués, J. 310
Aramendi Ecenarro, E. 352, 418, 426, 480
Arce-Diego, J. L. 294, 360, 376
Arias Palomares, M. A. 319, 335, 476
Armada-Maresca, F. 36
Artieda, J. 57, 235
Artola, G. 40
Astrain, J. J. 235
Ayala Fernández, U. 356
Aymerich Moreno, C. 165
Aznar-Cervantes, S. D. 36

B

Bachiller, A. 248
Bailón, R. 256
Balaban, E. 61
Bardají Ortiz, S. 414
Barea, R. 472
Barquero-Pérez, O. 177
Barrachina Fernández, M. 97
Barrenechea Carrasco, M. 356
Barrera, K. 157
Barros, R. 468
Barroso García, V. 252, 340
Bayés-de-Luna, A. 10
Belletti, R. 141
Berdón, P. M. 285
Bernal Ruíz, M. L. 256
Berrocoso, E. 455
Bertomeu González, V. 105, 109, 422, 430
Besné, G. M. 57, 235
Bikuña Izaguirre, M. 389
Bizy, A. 298
Boquete, L. 472
Borràs Argemí, M. 269
Bosch Roig, I. 77
Bote, J. 53
Botija Yagüe, M. P. 214
Brox, P. 85
Buj-Corral, I. 48
Bujanda Cundin, X. 185
Buruaga Lamarain, L. 173

C

Caffarena Fernández, G. 93
Calero Núñez, S. 14, 117
Callejón, M. A. 198
Calvillo Arbizu, J. 227
Calvo, C. J. 298
Calvo-Haro, J. 365
Camacho Ramos, E. 77
Cano, M. 121, 260
Cano, O. 101
Cano Serrano, J. 2
Canós, M. A. 306
Carnero González, E. 389
Carretero, J. 85
Carvajal Ahumada, L. A. 19
Castells, F. 393
Castrejón, S. 393
Cavaliere, C. 472
Cenis, J. L. 36
Cepero Escribano, V. 27
Ceruelo Abajo, S. 190
Cervigón, R. 393
Chamorro Servent, J. 269
Cherry, E. 141
Chil, R. 81, 281, 401
Cichy, J. 36
Cirillo, G. 36
Cirugeda Roldán, E. M. 14, 117
Cobo, A. 89
Colly, A. 48

Colomer, A. 443
Corbett, B. 53
Corral Bolaños, M. 181
Corral, J. 181
Courtial, E. J. 48
Cubero Gutiérrez, L. 61
Cygankiewicz, I. 10

D

Daya, M. R. 480
De Andrés, J. 290, 306
de la Cámara, C. 256
de la Nava, A. S. 161
de los Reyes Guzmán, A. 190
del Campo, F. 252, 315, 340
Del Cerro, C. F. 73, 206, 285, 439, 447
del Valle Hidalgo, A. 273, 459
Desco, M. 73, 85, 206, 273, 285, 439, 447, 459
Díaz Carnicero, J. 214
Díaz Gómez, D. 65
Díaz-Pernas, F. J. 397
Diaz Zelaia, A. 173
Dómine Gómez, M. 181
Durá, J. L. 290, 306
Durán Rey, D. 129

E

Eguizabal Argaiz, C. 173
Elola Artano, A. 426
Enero Navajo, J. 14
Enríquez Medina, R. 169
Escobar, F. 137
Escribano Cano, P. 327, 476
Escudero Arnanz, O. 464
Espinosa-Salinas, I. 223
Extramiana Esquisabel, L. 389

F

Falcone, F. 235
Fanjul-Vélez, F. 294, 360, 376
Farina, B. 181
Farré Ventura, R. 165
Fenollosa-Artés, F. 48
Fenton, F. H. 141
Fernández Avilés, F. 161
Fernandez-Bes, J. 344
Fernández Cerezo, M. 89
Fernández Santos, M. E. 161
Ferreira, D. 434
Finotti, E. 109, 422
Francés Morcillo, L. 201
Franco-Rosado, P. 198

G

Gabilondo Cuellar, I. 356
Galán González, A. 273, 459
Gallardo Madueño, G. 181
Gallego-Ferrer, G. 153
Ganoza-Quintana, J. L. 294
García Blas, J. 85, 206, 273, 459
García-Briega, M. I. 153
García Carmona, R. 169
García-Carretero, R. 177
García-Casado, J. 240
García-Cazorla, M. A. 248
García-Duarte Sáenz, L. 384
García, E. 256
García Gadañón, M. 315, 484
Garcia-Martin, E. 472
García Martínez, C. A. 93, 414
García-Mato, D. 384
García Teruel, M. 327
Gil Agudo, A. 190
Gil-Bazo, I. 181
Gil, E. 256
Giménez, R. C. 285, 447
Gimeno Tébar, M. 125
Gómez Blanco, J. C. 44, 149, 373
Gómez, C. 121, 260, 331, 405
Gómez Cid, L. 161
Gómez-Cruz, C. 265
Gómez, E. J. 210, 223, 231, 277, 380, 451
Gómez-Grande, A. 277, 451
Gómez-Pilar, J. 405
Gómez-Ribelles, J. L. 153
Gomis, P. 434
Gomis-Tena, J. 137
González-Alonso, J. 397
González de Julián, S. 214
González-González, A. 373
Gonzalez-Nieto, D. 36
Gozal, D. 252, 315, 340
Grigorian Shamagian, L. 161
Guedan-Duran, A. 36

Guillot-Ferriols, M. 153
Guinea, G. 36
Gutiérrez-de Pablo, V. 121, 260, 331, 405
Gutiérrez Martín, A. 369
Gutiérrez, R. 472
Gutiérrez Tobal, G. C. 252, 315, 340
Guzmán-García, C. 210

H

Haro, J. M. 256
Hernando, M. E. 133, 223
Hidalgo Olivares, V. M. 117
Hidalgo-Torres, L. A. 61, 281, 401
Holgado-Cuadrado, R. 177
Hornero, R. 121, 244, 252, 260, 315, 331, 340, 405, 484
Hornero Sos, F. 105, 109, 327, 422, 430
Huerta Herraiz, A. 335, 409

I

Idris, A. 352, 426
Iniesta-Chamorro, J. M. 223
Invernon Garrido, D. 323
Irigoien, E. 235
Irusta Zarandona, U. 352, 418, 426, 480

J

Jaén-Vargas, M. 89, 194
Jaureguibeitia Lara, X. 352, 418
Jemni-Damer, N. 36
Jiménez García, J. 315
Jimenez-Hernando, M. 231
Jiménez, S. 81
Jiménez-Torres, C. 210
Jurado-Barba, R. 133

K

Kheirandish Gozal, L. 252, 315, 340
Kobler, R. J. 244
Kontaxis, S. 256
Kovalyk, O. 348

L

Laguna, P. 6, 256, 344, 434
Laguna, S. 265
Lanceros-Méndez, S. 153
Langley, P. 335, 476
Lara-Doña, A. 218, 455
Largo-Aramburu, C. 36
Larraga García, B. 369
Lazaro, J. 256
Lazo, M. 468
Ledesma-Carbayo, M. J. 181
León-Jiménez, A. 218
Lombardi, D. 344
López Dolado, E. 190
López Dorado, A. 472
López Gálvez, M. I. 484
Lopez-Martin, A. 235
López Plaza, B. 223
Lozano Berrio, V. 190
Lozano-Picazo, P. 36

M

Macias, P. 401
Malanda Trigueros, A. 113
Mañana, M. A. 248
Mancha Sánchez, E. 44
Marcos Romero, A. C. 44
Marín, A. 380
Marquette, C. A. 48
Martí-Bonmatí, L. 306
Martín Montero, A. 315
Martín Odriozola, A. 201
Martín-Rodríguez, J. F. 198
Martín-Yebra, A. 434
Martínez Agüero, S. 464
Martínez-Albero, E. 277
Martínez, C. 73, 439
Martínez-Cagigal, V. 244
Martínez Iniesta, M. 327
Martínez, J. P. 6, 10, 434
Martínez, Marta 256
Martínez, Marcel 393
Martínez-Maldonado, A. 133
Martinez Mateu, L. 141
Martínez Rodrigo, A. 335, 409
Martínez-Zarzuela, M. 397
Matamoros Pacheco, M. 44
Maturana-Candelas, A. 331, 405
Mediavilla-Santos, L. 365
Medina López, R. A. 31
Mejía Blanco, R. H. 369
Merino, A. 157

Merino, J. L. 393
Meste, O. 393
Migliorelli, C. 248
Millet, J. 393
Mir, P. 198
Miranda Martínez, A. 19
Mitxelena-Iribarren, O. 185
Mondini, V. 244
Montero-Cordal, L. 133
Mora Jiménez, I. 145, 464
Morales-González, M. 218
Morales Sánchez, J. 348
Moreno Blanco, D. 231
Moreno Montañés, J. 389
Moreta-Martínez, R. 365
Moro López, M. 161
Moyano-García, J. L. 149
Müller-Putz, G. R. 244
Muñoz-Barrutia, A. 181, 265
Muñoz-Zamarro, C. P. 484

N

Naranjo Hernández, D. 31
Naranjo, V. 443
Navallas Irujo, J. 113
Niño Merchán, J. 65
Nohales, F. J. 240
Núñez, P. 331
Nüzket, T. 113

O

Ochandiano, S. 384
Olivares-Marín, M. 373
Olmos, P. M. 447
Oraá-Pérez, J. 484
Oropesa, I. 133, 210, 380
Ortega Morán, J. F. 53
Ortigosa, N. 101, 125
Ortiz-Gómez, F. 89
Ortiz, M. 472
Osca Asensi, J. 422, 430
Otero, J. 165
Otero Quintana, A. 93, 414

P

Pachón Iglesias, M. I. 319, 409
Pagador Carrasco, J. B. 44, 149, 373
Palacios Miras, C. 181
Palacios, S. 10
Palmieri, F. 434
Pampín Suárez, S. 360, 376
Panetsos, F. 36
Papa, M. 36
Paredes, J. 23, 389
Pascau, J. 61, 81, 365, 384
Peces-Barba, G. R. 181
Pelliccioni Monrroy, O. 69
Pérez, A. 157
Pérez Benito, D. 81, 281, 401
Pérez-Campo, F. M. 376
Pérez-Domínguez, R. 277
Pérez Javato, R. 231
Pérez-Jiménez, N. 451
Pérez-Mañanes, R. 365
Pérez Martínez, C. 6
Pérez-Porro Segarra, A. 93
Perez-Rigueiro, J. 36
Pérez Valdivia, M. A. 31
Piol, A. 206, 447
Piras, F. 218
Plancha Burguera, E. 14
Pose-Díez-de-la-Lastra, A. 365
Poza, J. 121, 260, 331, 405
Puchol Calderón, A. 319, 409
Pueyo, E. 6, 10, 344, 434

Q

Quesada Dorador, A. 105, 109, 117
Quintana Díaz, M. 369

R

Ramírez, J. 434
Ramírez Palomino, L. 302
Ramos Guerra, A. D. 181
Ramos Rojas, F. 227
Raphel, F. 344
Reina Tosina, J. 31, 198, 227
Revilla-Vallejo, M. 121, 260, 331, 405
Reyes, C. 53
Reyes Leiva, K. 89, 194
Rieta Ibáñez, J. J. 2, 14, 105, 109, 117, 319, 327, 335, 409, 422, 430, 476
Ripoll, J. 265
Rivera González, M. X. 19

Roa, L. M. 31, 198
Rodellar, J. 157
Ródenas García, J. 327, 476
Rodrigo, M. J. 472
Rodríguez, A. 373
Rodriguez Aierbe, C. 173
Rodríguez Almeida, A. J. 298
Rodríguez-de-Pablo, C. 201
Rodríguez-González, V. 121
Rodríguez Lázaro, M. 165
Rodríguez, M. 380
Rodríguez-Rey, J. C. 376
Rojo, F. 36
Román Martínez, I. 227
Romero-Bascones, D. 356
Romero, L. 137, 141
Romero-Oraá, R. 484
Romero, S. 248
Roycroft, B. 53
Rubio Bolívar, J. 369
Ruipérez-Campillo, S. 393
Ruiz-Gómez, S. J. 331
Ruiz, J. E. 434
Ruiz Moreno, A. 319
Ruiz-Solís, S. 277

S

Saiz, J. 137, 141, 290, 306
Saiz Vivó, M. 443
Sakaltras, N. 73
Salgueria Lazo, M. 31
Sampedro-Puente, D. A. 344
Sánchez Ávila, C. 97
Sánchez García, J. 77
Sánchez-González, P. 210, 231, 277, 380, 451
Sánchez Lara, S. 194
Sánchez Margallo, F. M. 44, 129, 149, 373
Sánchez Margallo, J. A. 129
Sánchez-Morillo, D. 218, 455
Sánchez-Trancón, A. 149
Sandberg, F. 101
Santamaría-Bustamante, L. 360
Santamaría Vázquez, E. 340
Santos Cabrera, S. 173
Santos Paz, J. A. 169
Seiffert, A. P. 277, 451
Seijo, L. M. 181
Sellés Navarro, I. 348
Sereno Moyano, M. 223
Serra, P. 149
Serrano, A. 468
Serrano Arriezu, L. 235
Serrano, C. 468
Serrano Olmedo, J. 19, 89, 194
Siddi, S. 256
Soguero Ruiz, C. 145, 464
Solanes, C. 290, 306
Sorbet Santiago, S. 214

T

Talavera Díaz, F. 190
Tapia-Galisteo, J. 133
Tarancón-Rey, J. 145
Tarazona, M. 240
Tejedor Noguerales, J. 414
Tejo-Otero, A. 48
Tola-Arribas, M. A. 121, 260
Toledo, T. 468
Torrado, O. 149
Torres-Sánchez, S. 455
Tost, A. 248
Tovar, F. A. 73
Trigo, J. D. 235

U

Urteaga Urizarbarrena, J. 426
Usó Talamantes, R. 214
Usón, J. 373

V

Valencia, M. 57, 235, 310
Valenzuela López, L. 97
Vallejo-Illarramendi, A. 23
Vaquerizo Villar, F. 252, 340
Vaquero, J. J. 61, 81, 265, 281, 401
Vásquez López, J. 302
Vaz Pato, I. 69
Vega Beltrán, J. 65
Verdú Monedero, R. 348
Vesga-Castro, C. 23
Viik, J. 6
Vilades, E. 472
Villarejo-Galende, A. 451
Virtuoso, A. 36

Vraka, A. 105, 430

W

Wang, H. E. 352

Z

Zabala Eguren, A. 173

Zabaleta Rekondo, H. 201

Zarzoso-Muñoz, M. 298

Zeinoun, M. 19

Zubia Aranburu, J. 173

Zubia Garea, G. 418

CASEIB
2020
XXXVIII
Congreso Anual de la Sociedad Española de Ingeniería Biomédica

ISBN: 978-84-09-25491-0