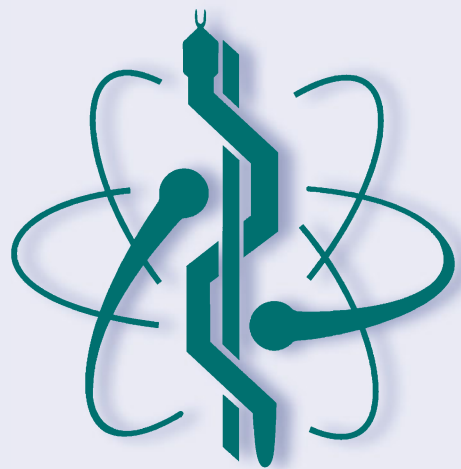


IFMBE Proceedings 96

José de Jesús Agustín Flores-Cuautle · Balam Benítez-Mata ·
Ricardo Antonio Salido-Ruiz ·
Gustavo-Adolfo Alonso-Silverio ·
Guadalupe Dorantes-Méndez ·
Esmeralda Zuñiga-Aguilar · Hugo A. Vélez-Pérez ·
Edgar Del Hierro-Gutierrez ·
Aldo Rodrigo Mejía-Rodríguez *Editors*

XLVI Mexican Conference on Biomedical Engineering

Proceedings of CNIB 2023, November 2–4,
2023, Villahermosa Tabasco, México -
Volume 1: Signal Processing And
Bioinformatics



 Springer

Series Editor

Ratko Magjarević, *Faculty of Electrical Engineering and Computing, ZESOI, University of Zagreb, Zagreb, Croatia*

Associate Editors

Piotr Ładyżyński, *Warsaw, Poland*

Fatimah Ibrahim, *Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia*

Igor Lackovic, *Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia*

Emilio Sacristan Rock, *Mexico DF, Mexico*

The IFMBE Proceedings Book Series is an official publication of *the International Federation for Medical and Biological Engineering* (IFMBE). The series gathers the proceedings of various international conferences, which are either organized or endorsed by the Federation. Books published in this series report on cutting-edge findings and provide an informative survey on the most challenging topics and advances in the fields of medicine, biology, clinical engineering, and biophysics.

The series aims at disseminating high quality scientific information, encouraging both basic and applied research, and promoting world-wide collaboration between researchers and practitioners in the field of Medical and Biological Engineering.

Topics include, but are not limited to:

- Diagnostic Imaging, Image Processing, Biomedical Signal Processing
- Modeling and Simulation, Biomechanics
- Biomaterials, Cellular and Tissue Engineering
- Information and Communication in Medicine, Telemedicine and e-Health
- Instrumentation and Clinical Engineering
- Surgery, Minimal Invasive Interventions, Endoscopy and Image Guided Therapy
- Audiology, Ophthalmology, Emergency and Dental Medicine Applications
- Radiology, Radiation Oncology and Biological Effects of Radiation
- Drug Delivery and Pharmaceutical Engineering
- Neuroengineering, and Artificial Intelligence in Healthcare

IFMBE proceedings are indexed by SCOPUS, EI Compendex, Japanese Science and Technology Agency (JST), SCImago. They are also submitted for consideration by WoS.


Proposals can be submitted by contacting the Springer responsible editor shown on the series webpage (see “Contacts”), or by getting in touch with the series editor Ratko Magjarevic.


José de Jesús Agustín Flores Cuautle ·
Balam Benítez-Mata ·
Ricardo Antonio Salido-Ruiz ·
Gustavo Adolfo Alonso-Silverio ·
Guadalupe Dorantes-Méndez ·
Esmeralda Zúñiga-Aguilar ·
Hugo A. Vélez-Pérez ·
Edgar Del Hierro-Gutiérrez ·
Aldo Rodrigo Mejía-Rodríguez
Editors


XLVI Mexican Conference on Biomedical Engineering


Proceedings of CNIB 2023, November 2–4, 2023,
Villahermosa Tabasco, México - Volume 1:
Signal Processing And Bioinformatics


Editors


José de Jesús Agustín Flores Cuautle 
División de Estudios de Posgrado e
Investigación
Instituto Tecnológico de Orizaba
Orizaba, Mexico


Ricardo Antonio Salido-Ruiz 
Departamento de Bioingeniería Traslacional
Universidad de Guadalajara
Guadalajara, Jalisco, Mexico


Guadalupe Dorantes-Méndez 
Facultad de Ciencias
Universidad Autónoma de San Luis Potosí
San Luis Potosí, Mexico

Hugo A. Vélez-Pérez 
Departamento de Bioingeniería Traslacional
Universidad de Guadalajara
Guadalajara, Mexico

Aldo Rodrigo Mejía-Rodríguez 
Facultad de Ciencias
Universidad Autónoma de San Luis Potosí
San Luis Potosí, Mexico

Balam Benítez-Mata 
Department of Biomedical Engineering
University of California, Irvine
Irvine, CA, USA

Gustavo Adolfo Alonso-Silverio 
Facultad de Ingeniería
Universidad Autónoma de Guerrero
Chilpancingo de los Bravo, Guerrero, Mexico

Esmeralda Zúñiga-Aguilar 
Departamento de Ingeniería Eléctrica y
Computación
Universidad Autónoma de Ciudad Juárez
Ciudad Juárez, Chihuahua, Mexico

Edgar Del Hierro-Gutiérrez
Facultad de Medicina
Universidad Autónoma del Estado de México
Toluca de Lerdo, Mexico

ISSN 1680-0737

IFMBE Proceedings

ISBN 978-3-031-46932-9

<https://doi.org/10.1007/978-3-031-46933-6>

ISSN 1433-9277 (electronic)

ISBN 978-3-031-46933-6 (eBook)

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

Preface

The XLVI Mexican Conference on Biomedical Engineering (CNIB 2023), organized by the Mexican Society of Biomedical Engineering (SOMIB), took place at Villahermosa Tabasco, Mexico, on November 2–4. The CNIB 2023 aims to close the gap among academia, industry, and clinics by sharing the latest advances in biomedical engineering and related fields. The conference joins researchers, academics, students, health professionals, and industry partners in an enriching environment to discuss and foster the advancements in biomedical engineering.

This book condenses the advances related to signal and image processing. Within this book, it is possible to find the novelties regarding health monitoring using the various modalities of bioinstrumentation. Several artificial intelligence applications are presented throughout this book, impacting all biomedical engineering branches. The topics covered by this book are related to, but not limited to, artificial intelligence, signal and image processing, biosensors, biomaterials, data science, biosensors, and computer modeling.

In this book, the scientific committee presents high-quality papers resulting from a peer review. More than 120 referees performed the screening process under the guidance of the scientific committee. As a result of the strict review process, 42 papers conform to this volume set, representing 60% of the total number of submissions.

All plenary and scientific presenters and distinguished guests from academia, government, health care, industry, and society made this conference the leading event in the field in the country. We want to express our gratitude to the scientific committee and the organizing committee for their essential labor, diligence, and optimism leading up to the conference. We would also like to express our gratitude to the session chairpersons, authors, and reviewers for their involvement and essential contributions to the biomedical engineering field.

We hope to inspire new generations to continue growing and developing this discipline. We thank you on behalf of all those involved in organizing this CNIB 2023.

Balam Benítez-Mata
Co-chair Program CNIB 2023

José de Jesús Agustín Flores Cuautle
Chair Program CNIB 2023

President of the Mexican Society of Biomedical Engineering Message

Dear readers and collaborators in the field of biomedical engineering,

It is an honor to address you on the occasion of the XLVI Mexican Conference on Biomedical Engineering, an event that symbolizes our collective commitment to the advancement of biomedical engineering. At a time when technology and innovation play an essential role in the development of our country, this congress represents an invaluable opportunity to bring together the brightest and most dedicated minds in this field.

Biomedical engineering plays a fundamental role in merging technology with the health sector. This congress is a showcase for the latest research and developments in this area. The scientific papers that are presented here reflect the hard work and dedication of experts who seek to innovate and contribute to improving health through engineering; by coming together at this event, we create a space to share ideas and establish fruitful collaborations that further advance our field.

This congress is not only a platform for the dissemination of results, but also a space for personal and professional growth. I hope that you find inspiration in the work and that you take the opportunity to establish collaborations that allow you to develop new research.

I appreciate your participation and hope you enjoy and benefit from this event as much as we did organizing it. Together, we are shaping the future of biomedical engineering in Mexico.

Sincerely,

Francisco Javier Aceves Aldrete
President of the Mexican Society of
Biomedical Engineering 2022–2023,
Organizing Committee of the XLVI
Mexican Conference on Biomedical
Engineering

Organization

CNIB 2023 was organized by the Mexican Society of Biomedical Engineering (SOMIB).

Executive Committee

Conference Chair

Francisco Javier Aceves Aldrete

Program Chair

José Jesús Agustín Flores-Cuautle

Scientific Committee

Program Chair

José Jesús Agustín Flores-Cuautle Instituto Tecnológico de Orizaba, México

Program Co-chair

Balam Benítez-Mata University of California Irvine, USA

Session Coordinators

Ricardo Antonio Salido-Ruiz	Universidad Guadalajara, México
Gustavo-Adolfo Alonso-Silverio	Universidad Autónoma de Guerrero, México
Guadalupe Dorantes-Méndez	Universidad Autónoma de San Luis Potosí, México
Esmeralda Zuñiga-Aguilar	Universidad Autónoma de Ciudad Juárez, México
Hugo A. Vélez-Pérez	Universidad de Guadalajara, México
Edgar Del Hierro-Gutierrez	Universidad Autónoma del Estado de México, México
Aldo Rodrigo Mejía-Rodríguez	Universidad Autónoma de San Luis Potosí, México

Awards Committee

Student Competition

Hugo A. Vélez-Pérez

Universidad de Guadalajara, México

Organizing Committee

Organizing Chair

Francisco Javier Aceves Aldrete

Program and Guest

Ximena Ruiz

Workshop

Manuel Nuñez Meza

Exposition Chair

Verónica Castillo

Special Session

Ana Luz Portillo

Students Coordinator

Rodrigo García

Education Committee

Chair

Eduardo Mendez Palos

Industry Committee

Chair

Christopher Bricio

Innovation Committee

Chair

Luis Fernández

Logistics Committee

Chair

Zayra Resendiz

Administrative Support

Nicolle Ramírez
Gabriela Sámano

Local Committee

Chair

Alejandra Casanova

Sponsoring Institution

Mexican Society of Biomedical Engineering (SOMIB)



Referees

Albino Martinez-Sibaja
Aldo Rodrigo Mejia-Rodriguez
Alejandro Miranda Cid
Alejandro Santos Díaz
Alicia Agueda CondeIslas
Andrea Cuevas-Flores
Antonio Alarcón-Paredes
Avinash L.
Braniff de la Torre
Brenda Guadalupe Muñoz Mata
Camilo Caraveo
Carlos Alberto Vizcaino
Carlos Daniel Sandoval Puentes
Carlos Omar Gonzalez Moran
Catalina Arenas Huertero
Citlalli Tujillo Romero
Daniel Ulises Campos-Delgado
Daniela Martínez
Diomar Rodriguez-Obregon
Dora-Luz Flores
Edgar Roman Arce-Santana
Eduardo Antonio Murillo Bracamontes
Edwin Christian Becerra Alvarez
Eleazar Samuel Kolosovas-Machuca
Elías Jesus Ventura-Molina
Erick Gutierrez Zeferino
Erik Bojorges-Valdez
Fabian Vazquez de la Rosa
Francisco Alvarado
Francisco Javier Alvarez Padilla
Gaspar Gonzalez Briceño
Gemima Lara Hernández
Gerardo Mendizabal-Ruiz
Griselda Quiroz
Humiko Yahaira Hernandez Acosta
Iris Paola Gúzman-Gúzman
Irma Uriarte Ramírez
Israel Román-Godínez
Joaquin Azpiroz Leehan
Jorge Alberto Roacho Pérez
Jorge Luis Perez-Gonzalez
José Israel Hernández-Oropeza
Jose Javier Reyes-Lagos

José Manuel Mejía Muñoz
Jose Pastor Rodriguez Jarquin
Juan Miguel Colores Vargas
Kevin Morales Almanza
Laura Mendez Magdaleno
Leonel Gonzalez Corral
Luis Antonio Aguilar
Luis Tomás Calvario Velásquez
Luis Torres Treviño
Manuel Chávez
María Cristina Padilla-Becerra
Martin Oswaldo Mendez Garcia
Mauricio Comas-Garcia
Miguel Alejandro Díaz Hernández
Miguel Ramírez Elias
Natanael Hernandez Vazquez
Nidiyare Hevia-Montiel
Norma Alicia Barboza Tello
Omar Paredes
Omar Piña-Ramirez
Oscar Osvaldo Sandoval Gonzalez
Pablo Antonio Stack Sánchez
Paul Medina
Rafael Bayareh Mancilla
Rafael E. González-Landaeta
Raquel Delgado
Raquel Martinez Valdez
Rigoberto Martinez-Mendez
Roberto L. Avitia
Rodrigo Martin
Sebastian Roldan-Vasco
Sergio Alfonso Sanchez-Manzo
Sergio Eduardo Sánchez Hernández
Sonia Charleston Villalobos
Stewart Santos
Sulema Torres-Ramos
Tomás Aljama Corrales
Vijay Kumar S.
Wilfrido Gómez

Contents

Artificial Intelligence, Modeling and Simulation of Biological Systems, Bioinformatics and Computational Biology

Comparison of Machine Learning Algorithms for ADHD Detection with Eye Tracking	3
<i>Karen P. Rodríguez Rivera, Cynthia D. Márquez Pizarro, Astrid J. Ríos Dueñas, Jesús J. Martínez Rodríguez, Carlos E. Cañedo Figueroa, Ana P. Leyva Aizpuru, Abimael Guzmán Pando, and Natalia Gabriela Sámano Lira</i>	
SkinSight: A Melanoma Detection App Based on Deep Learning Models with On-Device Inference	14
<i>Adrian Chavez-Ramirez, Andrea Romero-Ramos, Monica Aguirre-Ortega, Samantha Aguilar-Gameros, and Graciela Ramirez-Alonso</i>	
Classification of Acid-Base Disorders Through Machine Learning	24
<i>Rodrigo Ruiz de la Peña Martinez, Eutzhel André Del Val Saucedo, Paola Aidee de la Cruz Gallardo, Carlos Eduardo Cañedo Figueroa, Abimael Guzmán Pando, and Natalia Gabriela Sámano Lira</i>	
Identification of Pneumonia with X-ray Images Using Deep Transfer Learning	32
<i>Zarah Campos-Lopez, Jose Diaz-Roman, Boris Mederos-Madrazo, Nelly Gordillo-Castillo, Juan Cota-Ruiz, and Jose Mejia-Muñoz</i>	
In Silico Coaxial Antenna Design Applicator Optimization for Microwave Ablation Therapy in Medium Adipose Tissue Density Breast with Ductal Carcinoma <i>In-Situ</i>	41
<i>Eduardo Pinto García, Mario Francisco Jesús Cepeda Rubio, Geshel David Guerrero López, Kristian Segura Felix, José Irving Hernández Jaquez, Sergio Israel Velázquez Chávez, and Francisco Flores García</i>	
Convolutional Neural Network for Classifying Thoracic Diseases in Adult Patients	50
<i>Luis Alberto Gómez-Celaya, Jorge Alonso Débora-Ramírez, Rolando Corral-Barrón, Poleth Rodríguez-Domínguez, Aldair Verdugo-Valerio, Carlos Eduardo Cañedo-Figueroa, Abimael Guzmán Pando, and Hugo Varela-Rodríguez</i>	

Computation of Human-Sperm Local Flagellar Instantaneous Velocity	59
<i>Dan Sidney Díaz-Guerrero, Fernando Montoya, Haydee Olínca Hernández, Paul Hernández-Herrera, Alberto Darszon, and Gabriel Corkidi</i>	
Machine Learning Model to Classify Patients with Complicated and Uncomplicated Type 2 Diabetes Mellitus in the New Civil Hospital of Guadalajara “Juan I. Menchaca”	67
<i>Víctor Manuel Medina-Pérez, Isaac Zúñiga-Mondragón, José Alfonso Cruz-Ramos, Kevin Javier Arellano-Arteaga, Iryna Rusanova, Gerardo García-Gil, and Gabriela del Carmen López-Armas</i>	
Simulator for Cardiovascular Procedures	84
<i>Tomás Efraín Sánchez Pérez, Raúl Martínez Ruiz, Jesus Moises Martinez Buendía, Iliana González Pérez, Carlos Alberto Gallegos Hernández, Efraín Muñoz Montero, Luis Antonio Jiménez Salinas, and Arturo Abundes Velasco</i>	
Design and Characterization of a Microfluidic Biological System for Bone Tissue	93
<i>Odín Ramirez-Fernandez, Fabian Equihua-Guillen, Adrian Garcia-Lara, and Esmeralda Zuñiga-Aguilar</i>	
Tumor Tissue Classification in Hyperspectral Histopathology Images Through Individual and Ensemble of Machine Learning Algorithms	100
<i>Daia Natalia Martínez-Arboleyda, Inés Alejandro Cruz-Guerrero, and Daniel Ulises Campos-Delgado</i>	
Classification of COVID-19 Mortality Risk: A Neural Network-Based Approach Using Mexican Healthcare Sector Data	112
<i>Enrique Mena-Camilo, Gerardo Hernández-Nava, Sheila Leyva-López, and Sebastián Salazar-Colores</i>	
Design of a Convolutional Neural Network for Hippocampal Segmentation in Epileptics and Healthy Patients	120
<i>Alina Andrea García Huizar and José Manuel Mejía Muñoz</i>	
Development and Evaluation of a Diagnostic Exam for Undergraduate Biomedical Engineering Students Using GPT Language Model-Based Virtual Agents	128
<i>Alberto Isaac Perez Sanpablo, María del Carmen Arquer Ruiz, Alicia Meneses Peñaloza, Gerardo Rodriguez Reyes, Ivett Quiñones Uriostegui, and Leonardo Eliú Anaya Campos</i>	

Transcriptional Expression of Bioactive Antimicrobial Peptides with Biomedical Potential in Diverse Organs of the Mexican Axolotl	137
<i>Edgar Cera-Domínguez, Gabriel Arenas-Ballesteros, Luis Varela-Rodríguez, Javier Camarillo-Cisneros, Abimael Guzman-Pando, and Hugo Varela-Rodríguez</i>	
Drugs and Therapeutic Targets in Breast Cancer: Physicochemical Analysis by Computational Chemistry	145
<i>Aarón Morales Stockton, Andrés Kaleb Aguiñaga Miramontes, Raquel Portillo Saenzpardo, Roxana Quintana Jurado, Rodrigo Domínguez García, Javier Camarillo Cisneros, and Alexica Celine Marquez Barreto</i>	
Signal Processing and Medical Images, Data Science	
Automatic Cluster Selection in K-Means Lung Segmentation	155
<i>Natanael Hernández-Vázquez, Stewart René Santos-Arce, Ricardo Antonio Salido-Ruiz, Sulema Torres-Ramos, and Israel Román-Godínez</i>	
Muscle Activation Patterns Differentiate Post-stroke and Healthy Population	164
<i>Esteban Eusebio-Martín, Ricardo Antonio Salido-Ruiz, Sulema Torres-Ramos, and Israel Román-Godínez</i>	
Heterogeneous Behavior in the Iowa Gambling Task: A Clustering Approach	174
<i>Manuel Chávez-Sánchez, Sulema Torres-Ramos, Israel Román-Godínez, and Ricardo Antonio Salido-Ruiz</i>	
Vision Algorithm to Compute Pupil Parameters: Towards Non-invasive Estimation of Intracranial Pressure	183
<i>Gerardo Adrián De La Rosa Hernández, Jesús Alberto Morales Gómez, Griselda Quiroz Compean, and Luis Martín Torres Treviño</i>	
Sex-Based Speech Pattern Recognition for Post-traumatic Stress Disorder	192
<i>Andrea Alejandra García-Valdez, Israel Román-Godínez, Ricardo Antonio Salido-Ruiz, and Sulema Torres-Ramos</i>	
Analysis of EEG Signals Recorded from Persons with Neuronal Alterations due to Covid-19 with Respect to the Cases Considered Normal	201
<i>María Fernanda Arámburo-Castell, María Monserrat Morín-Castillo, Olga Guadalupe Felix-Beltrán, and José Jacobo Oliveros-Oliveros</i>	

Automated Segmentation of Breast Skin for Early Cancer Diagnosis: A Multi-otsu Region Growing Approach for Detecting Skin Thickness Variations	211
<i>Luis Antonio Medina-Ramos, Alfonso Toriz-Vázquez, Yazmín Mariela Hernández-Rodríguez, Rafael Bayareh-Mancilla, and Oscar Eduardo Cigarroa-Mayorga</i>	
Towards Fluid Intake Quantification in Older Adults: An Algorithm for Movement Detection Using Accelerometry and Gyroscope Sensors	222
<i>Paulina Martínez, Nelly Gordillo-Castillo, and David Cortés Sáenz</i>	
Optimizing fMRI Techniques for Post-stroke Motor Rehabilitation: Preliminary Protocol Standardization	232
<i>Abigail Miranda-Hernández, José Carlos López-Mata, Gabriel Valadez-Roque, and Luis Jimenez-Angeles</i>	
Recommendations for ICA Denoising of Task-Based Functional MRI Data of Stroke Patients	241
<i>Martín Emiliano Rodríguez-García, Raquel Valdés-Cristerna, and Jessica Cantillo-Negrete</i>	
EOG Signal Classification Based on Blink-to-Speak Language	249
<i>María Cristina Padilla-Becerra, Diana Karen Macias-Castro, Ricardo Antonio Salido-Ruiz, Sulema Torres-Ramos, and Israel Román-Godínez</i>	
Semiautomatic Generation of a Three-Dimensional Human Anatomical Model of Bone for Biomedical Applications: First Approach	258
<i>Nohemi Barranco Alavez, Christian Isaac Hurtado-Esquivel, Citlalli Jessica Trujillo-Romero, and Álvaro Anzueto Ríos</i>	
Nonlinearities of Heart Rate Variability in Junior High School Students Under Mathematics Anxiety	275
<i>Pilar Cruz-López, Marisela Lázaro-Picazo, Eric Alonso Abarca-Castro, Ana Karen Talavera-Peña, Laura Mercedes Santiago-Fuentes, and José Javier Reyes-Lagos</i>	
Respiratory Rate Estimation from Tracheal Sound Analysis Using a Mobile Application for Smartphones	283
<i>Alison Contreras-Rodríguez, Nemecio Olvera-Montes, Carlos Mariaca-Gaspar, Abel Mosco-Vargas, Victor Maya-Venegas, Sonia Charleston-Villalobos, Tomas Aljama-Corrales, and Bersain Reyes</i>	










Analysis of Cardiovascular and Cerebral Interactions in Response to Cognitive Stressors Stimulus	292
<i>Verónica Janeth Martínez-Hernández and Guadalupe Dorantes-Méndez</i>	
IOT in Health, Bioinstrumentation, Biosensors, Bioimpedance and Micro/Nano Technologies, BCI Systems, Wearable Systems	
Automatic Knee Lock-Release System for a Knee-Ankle-Foot Orthosis (KAFO) with IOT Monitoring System in Real Time	303
<i>Octavio Diaz-Hernandez, Alejandra Santos-Borrález, Antonio Gómez Ruiz, Cynthia Fernanda Carrillo Gómez, and Ileri Gisel Rosas Arellanes</i>	
Portable Chest-Worn Device for Simultaneous Detection of Phonocardiogram and Seismocardiogram Using Piezoelectric Films Sensors	313
<i>Pamela Salas Canales, José Mejía Muñoz, and Rafael Gonzalez-Landaeta</i>	
A Novel Intra-Abdominal Pressure Measurement Device: Design and In Vitro Characterization	321
<i>Alejandra Soto, Ivan Luna, José Camacho, Oscar Núñez, Néstor López, and Bersain Reyes</i>	
Analysis of Signals and Energy Storage in HAp-Collagen-Glucose Piezoelectric Samples Containing an Electrolyte	331
<i>Susana Vargas, Rogelio Rodriguez, and Griselda Rivera</i>	
Intelligent Device for the Control of Electrical Outlet Usage for Medical Devices	343
<i>Héctor G. Delgado Rivera, Zaira L. Bojorquez Muñoz, César F. Hernández Ortiz, Sergio A. Corona Sánchez, Natalia Montañez Prieto, and José M. Macías Macías</i>	
Design of a Wrist-Worn Device for Simultaneous Detection of ECG and Cardiac Pulse: A Preliminary Study	357
<i>Rafael Gonzalez-Landaeta, Aldo Rodrigo Mejía Rodríguez, Guadalupe Dorantes Mendez, and Dora-Luz Flores</i>	
Heart Monitoring Vest (MAC) for Cats and Dogs	365
<i>Ana Lorena García Zárate, Anali Elizabeth González Espinosa, Christian Roberto Ibáñez Nangüelú, and Norma Celeste Estrada Calvo</i>	

A Low-Cost, IoT-Connected Force Platform for Fall Risk Assessment in Older Adults	374
<i>Enrique Hernandez-Laredo, Lorena Parra-Rodríguez, Ángel Gabriel Estévez-Pedraza, and Rigoberto Martínez-Méndez</i>	
Multifrequency Bioimpedance Characterization of DNA Concentration: Sensitivity Analysis	386
<i>Nancy Nallely Gómez-González, Catalina Montoya-Solórzano, Leticia Arias-González, Liliana Becerra-Martínez, Jaime Eduardo López-Madrigal, Virginia Sánchez-Monroy, Jacobo Esteban Munguía-Cervantes, Nadia Mabel Pérez-Vielma, Modesto Gómez-López, and César Antonio González-Díaz</i>	
Design and Heat Transfer Simulation of a Thermal Block for the Integration of a DNA Biosensor Based on PCR	392
<i>Gerardo Ames Lastra and César A. González Díaz</i>	
Magnetic Induction Spectroscopy for DNA Quantification: Preliminary Results	403
<i>Jaime Eduardo López-Madrigal, Liliana Becerra-Martínez, Leticia Arias-González, Nancy Nallely Gómez-González, Catalina Montoya-Solórzano, Virginia Sánchez-Monroy, Nadia Mabel Pérez-Vielma, Modesto Gómez-López, and César Antonio González-Díaz</i>	
Author Index	411

**Artificial Intelligence, Modeling
and Simulation of Biological Systems,
Bioinformatics and Computational
Biology**



Comparison of Machine Learning Algorithms for ADHD Detection with Eye Tracking

Karen P. Rodríguez Rivera , Cynthia D. Márquez Pizarro ,
Astrid J. Ríos Dueñas , Jesús J. Martínez Rodríguez ,
Carlos E. Cañedo Figueroa  , Ana P. Leyva Aizpuru ,
Abimael Guzmán Pando , and Natalia Gabriela Sámano Lira 

Universidad Autónoma de Chihuahua, Circuito Universitario 31109, Campus UACH II,
31125 Chihuahua, México
ccanedo@uach.mx

Abstract. ADHD, or attention deficit hyperactivity disorder, is a persistent pattern of inattention that affects both young people and adults, causing interference with their functioning and overall development. The objective of this study is to develop an efficient diagnostic tool based on machine learning algorithms. The proposed tool utilizes eye-tracking technology to collect data on patients' eye movements while engaging in a concentration game. The eye movement patterns are carefully analyzed and categorized into two groups: patients with ADHD and those without. Initially, a manual classification was performed, followed by the training of algorithms, resulting in F1 scores of 100%, 95.55%, and 60.86% for KNN, ANN, and SVM, respectively. The main goal of this project is to provide a comparison comparison between four machine learning techniques and get base for a diagnostic tool that surpasses the accuracy of current diagnostic methods. By achieving this, it aims to enhance the precision and efficiency of ADHD diagnosis, ultimately improving the quality of care and support provided to individuals with this condition.

Keywords: ADHD · Machine Learning · Eye Tracking · Diagnosis

1 Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is characterized by persistent patterns of inattention and/or hyperactivity that interfere with functioning or development. In ADHD, the inattention aspect of the disorder manifests as distractibility, lack of persistence, difficulty concentrating, and confusion. Hyperactivity involves excessive movement in inappropriate situations or engaging in excessive fidgeting, tapping, or excessive talking. Symptoms of this disorder can range from mild to absent [1].

Individuals with ADHD often exhibit a pattern of hyperactivity or inattention that hinders proper development. Common characteristics of patients with ADHD include frequently not following instructions, inability to complete tasks, being easily bothered or entering a state of denial when faced with mentally demanding tasks that require

prolonged effort. Currently, the diagnosis of this disorder is conducted through various medical examinations, such as auditory and visual tests, to rule out other disorders with similar symptoms to ADHD. A crucial stage in the process is completing a checklist to assess ADHD symptoms. It is important to highlight that diagnosis and treatment should be carried out by an interdisciplinary team available in specialized neurology clinics, consisting of different experts in this pathology [1].

Machine learning is a branch of technology to learn various tasks using data analysis and prediction algorithms. This tool in the field of medicine has a wide range of applications that can contribute to diagnosis. The application of this technology is highly beneficial for data mining in medical research and acquiring knowledge to improve health outcomes [2, 3].

Currently, there are already several investigations on different methods for detecting ADHD, such as electroencephalographic signals (EEG), attention and continuous performance test (CPT), as well as patient behavioral activity (BA). These studies employ different metrics than those obtained in this research [4 – 6].

Previously works using eye tracking for detect some disorders, this technique can be using for diagnosis and detection of spectrum autism, neurological disorders or affect for medications. Thus, is because the eye movements are a principal indicator of concentration, distraction and neurological impulses [13, 14].

The objective of this research is to develop a comparison of four popular algorithm of machine learning for detect ADHD that can accurately identify individuals with this disorder. This algorithm relies on a variety of characteristics associated with ADHD, such as inattention and hyperactivity. By analyzing different variables and features of the obtained samples, the algorithm can be a valuable tool in establishing a reliable diagnosis. It is important to emphasize that this diagnosis cannot replace existing evaluations conducted by a trained mental health professional. However, our intention is for it to serve as a supportive tool, providing additional information for clinical diagnosis.

2 Methodology

2.1 Signal Acquisition

For this work, we developed a graphical user interface (GUI) using the Python language. This GUI displayed a central point, and random images appeared around this point at a frequency of 2 Hz. Eye movement signals were acquired using a camera. To do this, participants were required to hold their heads in a static position and look directly at the central point. See Fig. 1.

This investigation was conducted with the participation of sixteen clinically diagnosed ADHD students and sixteen students without indications of this mental disorder, who were asked for their consent and were shown that their data would be anonymous in the research. All participants were between the ages of 18 and 27 and are students of biomedical engineering at the Autonomous University of Chihuahua.

The experiment acquired data for 15 s at a frequency of 12.4 Hz. This work used 80 signals (S_{ADHD}) for the class of ADHD and 80 signals (S_{NOADHD}) for the class of No ADHD. This was for the application of 10 experiments per participant [13]. The

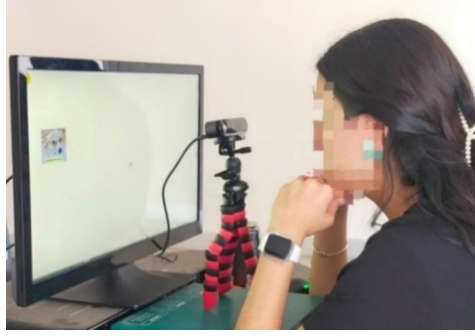


Fig. 1. GUI and a participant for sample collection.

data was divided into two groups: train and test. The train group contained 70% of the signals (tr), while the test group contained the remaining 30% (ts). In order to perform the algorithms, we use windows 10 with 16 GB RAM memory and using MATLAB 2020b.

2.2 Data Preprocessing

Each signal S^D ; $D = \{tr, ts\}$ has two signals corresponding to the movement of the iris in the axis x and the movement on axis y denoted by $pl = \{x_l, y_l\}$; $l = \{1, 2, 3, \dots, L\}$. Using mapminmax algorithm Eq. (1) where c correspond at class $c = \{ADHD, NOADHD\}$, r represents the number of experiment $r = \{1, 2, 3, \dots, R\}$ and i is the sample $i = \{1, 2, 3, \dots, I\}$ on the signal.

$$Norm_{c,r,i}^{D,pl} = \text{mapminmax}\left(S_{c,r,i}^{D,pl}\right) \quad (1)$$

This process was necessary for the variation on the position of the camera during data acquisition.

2.3 Feature Extraction

The feature extraction was necessary for develop the description of each signal. The features selected show a variation with eye movement, for this we used the feature of entropy whit the Eq. (2).

$$H_{c,r}^{D,pl} = \text{Entropy}(Norm_{c,r,i}^{D,pl}) \quad (2)$$

Other feature was the energy calculated with the Eq. (3)

$$E_{c,r}^{D,pl} = \text{Energy}(Norm_{c,r,i}^{D,pl}) \quad (3)$$

The finish feature was the standard deviation, for this we used the Eqs. (4), where $\overline{Norm_{c,r}^{D,pl}}$ represent the average of the signal $Norm_{c,r}^{D,pl}$ and I represents the number of samples in this signal.

$$\sigma_{c,r}^{D,pl} = \sqrt{\frac{\sum_{i=1}^I (Norm_{c,r,i}^{D,pl} - \overline{Norm_{c,r}^{D,pl}})^2}{I}} \quad (4)$$

Each feature was put into a vector how show in the Eq. (5)

$$V_{c,r}^D = [H_{c,r}^{D,x}, H_{c,r}^{D,y}, \sigma_{c,r}^{D,x}, \sigma_{c,r}^{D,y}, E_{c,r}^{D,x}, E_{c,r}^{D,y}] \quad (5)$$

2.4 Support Vector Machine

The algorithm of support vector machine (SVM) using a linear kernel for classification data [7, 8]. In this case, we used a SVM trained into MATLAB with *fitcsvm* function and V^{tr} vectors. The results are shown in the Eq. (6). Where V^{ts} represents the vector for predicting and the subindices are the position agree Eq. (5) and W represent the value of each component calculated from the SVM model.

$$R_{svm} = W_1 * V_1^{ts} + W_2 * V_2^{ts} + W_3 * V_3^{ts} + W_4 * V_4^{ts} + W_5 * V_5^{ts} + W_6 * V_6^{ts} + W_b \quad (6)$$

2.5 Artificial Neural Network

An artificial neural network, or simply a neural network, is a mathematical model based on biological brain networks [9, 10]. The ANN designed contains 6 inputs, 1 hidden layer with 5 neurons and 2 outputs (see Fig. 2). The network's training schedule was carried out using 15 epochs, a validation check of 6, a learning factor of 1×10^{-7} and a minimum error 1×10^{-29} , using the Levenberg-Marquardt backpropagation technique. These numbers were acquired experimentally using vectors of trained group.

2.6 K-Nearest Neighbors

The KNN algorithm, known as K-nearest neighbors, is a machine learning technique used for classification and regression. Its approach is based on identifying the nearest distances between a new sample and a set of training samples, allowing it to make predictions [11].

For this algorithm, is necessary to obtain the training data, which consists of a series of instances with their own features and a class label. This training vector allows training the classification model with the nearest neighbors using the *fitcknn* function in MATLAB, this makes a model calculates the Euclidean distance between the sample data point and the training data. This allows finding the K nearest neighbors to the sample point, where K is a predefined parameter representing the number of closest neighbors to consider for prediction. In this case, the selected value for K was 7.

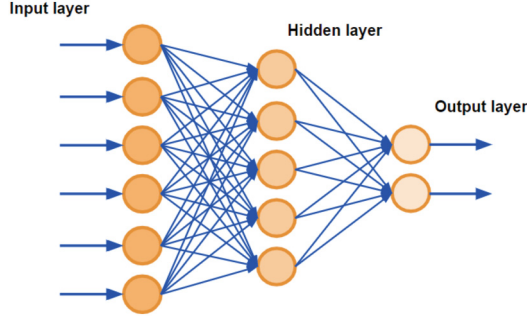


Fig. 2. ANN design with 6 inputs, one hidden layer of 5 neurons with sigmoidal tangential activation function and two outputs.

Finally, the classes to which these K nearest neighbors belong are tallied: ADHD (K_{ADHD}) and non-ADHD (K_{NOADHD}), and the number of identified classes is compared. See Eq. (6)

$$Class = \begin{cases} ADHD & \text{if } K_{ADHD} \geq K_{NOADHD} \\ NOADHD & \text{if otherwise} \end{cases} \quad (7)$$

2.7 Naïve Bayes

The Naive Bayes (NB) classifier is a generative learning method that assumes each feature, where the features are independent and do not interact with each other. It is a computationally efficient tool that can handle large datasets with multiple dimensions [12].

The probability of a specific class was calculated using Eq. (7). This equation refers to the random probability of a vector belonging to a patient with ADHD or a patient without ADHD, respectively. The number of vectors for training that agree at class c is denoted by N . Since this number is equal for both classes, the probability is 50%.

$$P_c = \frac{N_{V_c^{tr}}}{N_{V_{ADHD}^{tr}} + N_{V_{NOADHD}^{tr}}} \quad (8)$$

Afterward, statistical values of the mean ($(\bar{X}_{c,j})$) and variance ($(\sigma_{c,j}^2)$) were determined with the help of Eqs. (8) and (9), where j represents each characteristic, c corresponds to each of the classes, $c = \{ADHD, NOADHD\}$, r represents an experiment and $N_{V_{c,j,r}^{tr}}$ corresponds to the total of data for each characteristic j of the class c .

$$\bar{X}_{c,j} = \frac{\sum (V_{c,j,r}^{tr})}{N_{V_{c,j,r}^{tr}}} \quad (9)$$

$$\sigma_{c,j}^2 = \frac{\sum (V_{c,j,r}^{tr} - \bar{X}_{c,j})^2}{N_{V_{c,j,r}^{tr}} - 1} \quad (10)$$

Subsequently, using Eq. (10), the probability of a new sample belonging to a class was calculated as a function of each characteristic.

$$Pr_c = \left(\prod P(c|j) \right) (P_c) \quad (11)$$

To obtain the priority (Pr_c), the product of the probabilities obtained in Eqs. (10) and Eq. (7) was used with the help of Eq. (11).

$$Pr_c = \left(\prod^P (c|j) \right) (P_c) \quad (12)$$

Then, the evidence (Ev) was calculated using Eq. (12), which was used in Eq. (13) to calculate the posteriori (Ps_c) and determine the probability of each class.

$$Ev = \sum^P r_c \quad (13)$$

$$Ps_c = \frac{Pr_c}{Ev} \quad (14)$$

Finally, to determine the class to which the sample belongs, the rule presented in Eq. (14) must be followed.

$$Class = \begin{cases} ADHD & \text{if } Ps_{ADHD} > Ps_{NOADHD} \\ NOADHD & \text{if } \textit{Otherwise} \end{cases} \quad (15)$$

2.8 Metric Evaluation

With the remaining 30% of the data, we evaluated the accuracy of each algorithm to assess their performance and reliability. To evaluate these metrics, Eqs. (15), (16), (17) and (18) were used to determine the precision, accuracy, recall and F1 score, as can be seen in the equations the data of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN), which were being obtained with the development of the algorithm, are used. Precision focuses on the proportion of correct positive results, recall focuses on the ability to find all positive cases, accuracy focuses on the proportion of correct predictions overall, and the F1 score combines precision and recall to provide a balanced measure of model performance.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (16)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (17)$$

$$Recall = \frac{TP}{TP + FN} \quad (18)$$

$$F1 = 2 \frac{PrecisionRecall}{Precision + Recall} \quad (19)$$

The development of the eye-tracking algorithm for ADHD diagnosis is a complex procedure that must be carefully elaborated, as it takes into account multiple factors as mentioned before, including data selection and preparation. (See Fig. 3).

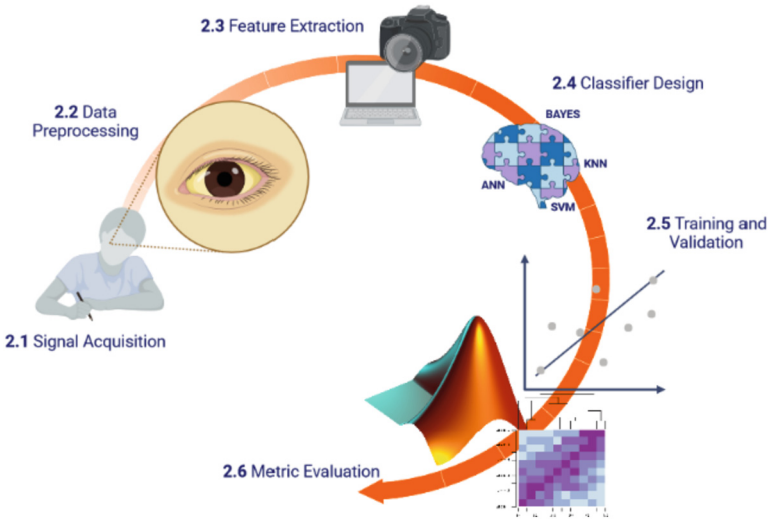


Fig. 3. Diagram illustrating research methodology

3 Results

The designed algorithms of SVM, ANN, Naïve Bayes, and KNN was evaluated using V^{IS} data to obtain our results. Where our SVM is show in Fig. 4. This was our lowest performing algorithm, with an accuracy of 46.66%, precision of 43.75%, recall of 87.50%, and an F1 score of 60.86%.

Next, the results obtained from our artificial neural network can be observed in Fig. 5. The metric results show an accuracy of 95.83%, precision of 96.66%, recall of 93.33%, and an F1 score of 95.55%.

SVM			
Real	ADHD	21	24
	No ADHD	3	0
		No ADHD	ADHD
		Prediction	

Fig. 4. SVM algorithm confusion matrix

ANN			
Real	ADHD	23	1
	No ADHD	1	23
		No ADHD	ADHD
		Prediction	

Fig. 5. Confusion matrix of the artificial neural network

For the Bayesian algorithm, the results are presented in a confusion matrix shown in Fig. 6. The metric results with an accuracy of 93.75%, precision of 93.75%, recall of 87.5%, and an F1 score of 93.33%.

NB			
Real	ADHD	21	0
	No ADHD	3	24
		No ADHD	ADHD
		Prediction	

Fig. 6. Confusion matrix of the Naïve Bayes algorithm

Lastly, our best-performing algorithm, KNN, is presented in Fig. 7. This show 100% accuracy, precision, recall, and F1 score.

		KNN	
		ADHD	No ADHD
Real	ADHD	24	0
	No ADHD	0	24
		No ADHD	ADHD
		Prediction	

Fig. 7. Confusion matrix of the KNN algorithm

4 Discussion

The development of an eye tracking algorithm for the diagnosis of ADHD is a promising area of research in the integration of biomedical engineering and psychology. The eye tracking method can provide objective information about attention patterns and gaze control in patients with ADHD, which could improve traditional diagnostic methods based on clinical observations and subjective evaluations. It is important to consider that an eye tracking algorithm for ADHD diagnosis may present significant challenges.

The main advantage of this type of systems is the reduction of time for detection ADHD in comparison with method of recompilation of data from family, clinical history and psychological test. If the tool mentioned on this paper is only a probe, in a future this can be used how a first intervention in case suspicious for a deep analysis agree the results and congruence with a specialist. As mentioned earlier, several studies have been conducted using different methods for ADHD diagnosis. The inclusion of the biomedical engineer in this field offers several important advantages.

Firstly, the eye tracking method provides a continuous and non-invasive measure of visual behavior, allowing for the capture of subtle patterns and important features of eye movement. For example, features such as fixation frequency, duration, movement speed, and deviation from the fixation point demonstrate significant differences between patients with ADHD and individuals without indications of the disorder. Furthermore, it offers advantages in terms of objectivity and standardization, reducing the inherent bias in clinical assessments and improving diagnostic consistency. By automating the process of analyzing eye tracking data, variability is reduced, and the reliability of diagnostic results is increased. Table 1 shows a comparison of metrics obtained from different previously investigated diagnostic methods, highlighting the significant performance difference of the mentioned algorithms.

This approach also has the potential to reduce costs and time in the diagnosis of ADHD. By using eye tracking algorithms and machine learning, a supportive model has been developed to detect individuals with ADHD. This automation allows for a more efficient and rapid evaluation, facilitating early detection and timely intervention.

While further development and validation are required, this project offers an innovative and effective way to objectively assess ADHD and improve the quality of life for affected individuals. The inclusion of biomedical engineering in psychology, through the

development of tools such as the eye tracking algorithm, enables interdisciplinary collaboration that drives significant advancements in the diagnosis and treatment of mental disorders, thereby enhancing patient care and well-being.

Table 1. Comparison of accuracy represented in percentage for various methods used in the algorithms.

Algorithm	EEG (%) [4]	CPT (%) [6]	BA (%) [5]	ET (Ours) (%)
SVM	96.4	-	71.9	46.66
ANN	96	89	72.1	95.83
KNN	81.2	-	84	100
NB	-	-	69.8	93.75

The present work is designed to focus on the diagnosis of children, as it is a difficult field to detect ADHD, but it would be introducing another graphical interface that could be suitable for preschool infants. Some of the main changes that would be made in the GUI would be the way in which the images are presented to children, since compared to adults in children it is easier for them to be distracted whether they are patients with ADHD or not, so it would be ideal to present another series of images to be able to evaluate them with eye tracking






References

1. Gatell Carbó, A.: Trastorno específico del aprendizaje **26**(1) (2022)
2. Sidey-Gibbons, J.A.M., Sidey-Gibbons, C.J.: Machine learning in medicine: a practical introduction. *BMC Med. Res. Methodol.* **19**(1), 1–18 (2019). <https://doi.org/10.1186/s12874-019-0681-4>
3. Chávez, H.G., Figueroa, C.C.: Vista de Diseño de algoritmo compuesto por machine learning y un modelo probabilístico para la detección de diabetes. In: *Memorias del Congreso Nacional de Ingeniería Biomédica*, pp. 57–60 (2021). <https://memoriascnib.mx/index.php/memorias/article/view/828/488>
4. Maniruzzaman, M., Shin, J., Hasan, M.A.M., Yasumura, A.: Efficient feature selection and machine learning based ADHD detection using EEG signal. *Comput. Mater. Contin.* **72**(3), 5179–5195 (2022). <https://doi.org/10.32604/cmc.2022.028339>
5. A. M. L. Analysis. applied sciences predicting children with ADHD using behavioral activity (2022)
6. Slobodin, O., Yahav, I., Berger, I.: A machine based prediction model of ADHD using CPT data. *Front. Hum. Neurosci.* **14**, 560021 (2020)
7. Bledsoe, J.C., et al.: Diagnostic classification of ADHD versus control: support vector machine classification using brief neuropsychological assessment. *J. Atten. Disord.* **24**(11), 1547–1556 (2020). <https://doi.org/10.1177/1087054716649666>
8. Paul, Y., Goyal, V., Jaswal, R.A.: Comparative analysis between SVM & KNN classifier for EMG signal classification on elementary time domain features. In: *2017 4th International Conference on Signal Processing, Computing and Control (ISPCC)*, pp. 169–175. IEEE (2017). <https://doi.org/10.1109/ISPCC.2017.8269670>

9. Zhang, Z.: *Multivariate Time Series Analysis in Climate and Environmental Research* (2017)
10. Valdez Hernández, K., et al.: Design and comparison of artificial intelligent algorithms for breast cancer classification. In: *XLV Mexican Conference on Biomedical Engineering: Proceedings of CNIB 2022, 6–8 October, Puerto Vallarta, México*, pp. 46–54 (2022)
11. Vázquez, S.R., Vidal, A., Borges, M., Valentín, J., Ginori, L.: Clasificación de células cervicales mediante el algoritmo KNN usando rasgos del núcleo. *Rev. Cuba. Cienc. Informáticas* **10**(1), 95–109 (2016)
12. Kaur, A., Kahlon, K.S.: Accurate identification of ADHD among adults using real time activity data. *Brain Sci.* **12**(7), 831 (2022). <https://doi.org/10.3390/brainsci12070831>
13. Ahmed, I.A., et al.: Eye tracking based diagnosis and early detection of autism spectrum disorder using machine learning and deep learning techniques. *Electronics* **11**, 530 (2022). <https://doi.org/10.3390/electronics11040530>
14. Zuñiga, M.I., López, E.E., Rodríguez, F.J., Soto, A.T.: Eye tracking for detection of ADHD patterns in children between 6 to 8 years old. In: *2022 International Conference on Inclusive Technologies and Education (CONTIE)*, Cartago, Costa Rica, pp. 1–7 (2022). <https://doi.org/10.1109/CONTIE56301.2022.10004422>



SkinSight: A Melanoma Detection App Based on Deep Learning Models with On-Device Inference

Adrian Chavez-Ramirez¹ , Andrea Romero-Ramos¹ ,
Monica Aguirre-Ortega¹ , Samantha Aguilar-Gameros¹ ,
and Graciela Ramirez-Alonso² 

¹ Facultad de Medicina y Ciencias Biomédicas, Universidad Autónoma de Chihuahua, 31125 Chihuahua, Chihuahua, Mexico

² Facultad de Ingeniería, Universidad Autónoma de Chihuahua, 31125 Chihuahua, Chihuahua, Mexico
galonso@uach.mx

Abstract. A timely and accurate skin cancer diagnosis is a key factor in reducing mortality rates, especially with melanoma which often resembles in its early stages with moles. Convolutional neural networks (CNNs) are models commonly used to classify dermoscopy images into benign or malignant. CNNs are frequently implemented on Graphical Processing Units (GPUs), which are not always available in rural areas. This paper compares three CNNs to classify benign and malignant melanoma images. We select the most appropriate neural architecture by comparing accuracy results and model lightness to load it on a mobile device. With this strategy, the training of the CNN is performed on the GPU and the inference in portable devices that can be used in rural areas. The developed app is named SkinSight. This app was evaluated with images of two different datasets achieving competitive results compared to state-of-the-art models. Considering that most people have a mobile device, this app could be used in areas where it is difficult to have specialized GPUs and highly trained personnel in cancer detection.

Keywords: melanoma detection · convolutional neural network · mobile application

1 Introduction

According to the American Cancer Society, melanoma is the most dangerous type of skin cancer; its early diagnosis is essential for successful treatment and patient survival [1]. In a study published in the Skin Cancer Foundation [3], late diagnosis of melanoma is a significant problem in many parts of the world, including Latin America, where lack of access to health services and awareness about skin cancer contributes to a late diagnosis. In 2020, according to data from the International Agency for Research in Cancer of the World Health Organization, through the

GLOBOCAN project, the incidence of Melanoma in Mexico was 2,051 cases with 773 deaths [4].

The diagnosis of melanomas is mainly made by visually inspecting skin lesions by highly trained dermatologists. Asymmetry, border, color, diameter, and lesion enlargement are the standard features that specialists consider. Another common way to diagnose cancer is by performing a biopsy, a pathological examination that takes much time and resources to provide the results. The Sierra Tarahumara is a mountain range part of the Sierra Madre Occidental located in Chihuahua. This rural and remote area lacks sufficient pathologists and medical resources to diagnose and treat skin cancer. The lack of information and awareness about this type of cancer in these communities can lead to delayed seeking care and late diagnosis of the disease. This problem can seriously affect the population's health and lead to higher mortality and morbidity rates in the region. A comprehensive approach is needed to address the lack of access to pathology services, including actions to increase skin cancer awareness, improve local doctors' training, and provide resources and technology for diagnostic testing and treatment.

Deep learning techniques, especially convolutional neural networks (CNNs), have been widely used in different image recognition tasks to automatically classify specific patterns on images [11]. Particularly for classifying skin cancer, different CNN models have been proposed achieving very accurate classification results [7, 16, 22].

Unfortunately, these systems have not yet been incorporated into daily clinical practice because most CNN models need the usage of Graphical Processing Units (GPUs), a hardware not very common in most hospitals. As an alternative to using expensive hardware equipment, TensorFlow (an open-source machine learning ML framework) has launched a lightweight version named TensorFlow Lite (TFLite) [5]. TFLite is optimized for deploying deep learning models on mobile and embedded devices with limited computational resources. Then, CNNs can be implemented in low-cost, low-power, portable, easy-to-use devices for classification and detection tasks. The training is performed on the GPU, but the inference can be executed on mobile devices, also known as on-device inference.

This work presents a comparison of state-of-the-art CNN models to classify images into benign or malignant melanoma lesions automatically. These models are trained and tested on two skin cancer datasets, demonstrating their robustness in different scenarios. The inference of the selected CNN model can be performed on a mobile device, known as on-device inference. The TFLite framework, in combination with Android Studio, allows us to convert the CNN model to a light version capable of working on low-cost, low-power devices. In this way, this CNN can be easily used by medical specialists with access to dermoscopy images and have the opportunity to diagnose suspicious cases in an early manner. Even when this methodology has already been implemented in recent research, most of them only evaluate their proposal in one dataset with few samples, achieve low-performance results, or perform the inference of the model in a server computer. In our proposal, we could maintain a balance

between accurate performance results considering two different datasets demonstrating the robustness of our proposal. We named our application SkinSight, which can be loaded on Android devices. Considering that most people have a smartphone, this tool could be used where it is difficult to have highly specialized GPUs and/or trained personnel in cancer detection. It is worth mentioning that this paper aims to identify the best CNN configuration that achieves a comparable performance with state-of-the-art models trained and tested on GPUs and with those developed to be used in portable devices.

2 Literature Review

The International Skin Imaging Collaboration (ISIC) is a global organization with an online repository of dermoscopic and clinical images of skin lesions [2]. The objective is that researchers from all over the world can work in the development of computer-aided systems to detect and diagnose melanoma and other skin cancers. With the advancement in computer vision algorithms based on deep learning models, different researchers have reported accurate results in classifying benign and malignant skin lesions. Cassidy et al. performed a benchmark study in [9] with images of the ISIC dataset and 19 state-of-the-art deep learning architectures. The VGG19, DenseNet121, and EfficientNetB2 architectures achieved the best area under the Receiver Operating Characteristic Curve (AUC) results. Benyahia, Meftah, and Lezoray [8] also investigate the efficiency of 17 deep learning architectures and 24 machine learning classifiers using the ISIC dataset. They concluded that the DenseNet201 neural architecture combined with the Cubic SVM algorithm produces the best classification results.

Rehman et al. [25] use a modified pre-trained DenseNet201 by staking three convolutional layers at the end of the model, followed by a global average pooling, a batch normalization, and two dense layers. The authors used a contrast stretching enhanced technique to improve the quality of the images reporting an average accuracy of 95.5%. In [21], was adapted a ResNet101 architecture to classify benign and malignant skin cancer images. Two convolutional layers were included at the end of the model, followed by pooling and two fully connected layers. The authors reported an average accuracy of 90.67%.

All these previous research papers perform their training and testing in a specialized GPU, achieving state-of-the-art performance in skin lesions classification tasks. After deeply analyzing their results, we select the ResNet101, DenseNet201, and a CNN of the EfficientNet family in our experiments. The accurate reported results and reduced number of parameters in these neural architectures make them ideal candidates for our research.

Figure 1 shows a block diagram of the process we follow in developing our SkinSight app. First, it is necessary to train the different deep learning models on TensorFlow with the appropriate datasets and compare their performance to select the most appropriate model. Then, convert the selected CNN to TensorFlow Lite. Next, set up Android Studio for Android App development with the appropriate Android SDK and NDK components installed, add TensorFlow Lite

dependencies, and copy the TF-Lite model into the project. The TFLite interpreter is necessary to load the model in the project. A user interface is designed to create the views and controls to interact with the model and display the prediction results appropriately. Then, connect an Android device to the computer and build the app with Android Studio. Finally, test SkinSight with images to confirm that the CNN model works as required.

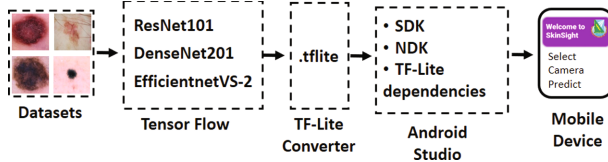


Fig. 1. Block diagram of deploying a CNN in a mobile device using TF-Lite and Android Studio.

The general methodology of performing the training of the CNN in the GPU and the inference in a mobile device (to be used by the medical sector) has already been proposed in different research papers. In [19] is presented a mobile app to classify skin diseases considering their severity based on the MobileNetV2 architecture. A dataset of 1,220 images is processed, achieving an accuracy of 94.32% in the classification task. In [14], a dataset of 2,358 images was classified as melanoma or benign using the InceptionV3 neural architecture. The accuracy reported by the authors is 81%. Dai et al. [10] presented an on-device inference app using 10,015 images. The accuracy achieved by the model was 75.2%. In [15] is presented an augmented reality app that classifies skin lesions for identifying melanoma. The app continuously tracks the lesion, implementing different image pre-processing algorithms to remove hair and segment the lesion before analyzing the image in the CNN model. Their method achieved an accuracy of 78.8%. Kousis et al. [20] load a light version of a DenseNet169 network on a mobile Android device to classify benign or malignant images. The DenseNet169 model achieved an accuracy of 91.10%, considering a dataset of 10,015. The authors mentioned that when testing their app in a real environment, it was necessary to transfer the image to a server for better performance. In [12], the MobileNetV2 architecture classifies skin lesion images considering three datasets. The overall accuracy performance reported when testing their proposal in a new dataset with the mobile app was 91.33%. Arani et al. [6] presented the Melanlysis app for detecting skin cancer based on the EfficientNetLite-0 architecture. The authors use only the dataset's dermoscopy images, achieving an accuracy of 94%. In [13] is presented a lesion segmentation and classification method based on a DenseNet201 model loaded on a mobile device. The classification task considers the identification of seven skin lesion classes achieving an accuracy of 89%.

3 Methodology

3.1 Deep Learning Models

The ResNet (Residual Neural Network) architecture introduces the concept of residual or skip connection to address the vanishing gradient problem present in deep neural networks [17]. The Residual Blocks of the ResNet model have convolutional and batch normalization layers and ReLu activation functions. The number of residual blocks defines the variant of the ResNet architecture. We select the ResNet101 in our experiments considering the results reported in [21].

DenseNet, or Dense Connected Convolutional Network, uses the concept of dense blocks to connect the output of every other layer within each of its blocks [18]. That is, the output of each layer is concatenated before passing it to the input of the subsequent layer within each dense block. To reduce the spatial dimensions between dense blocks and the number of channels, DenseNet defines Transition Layers. Similar to ResNet, DenseNet defines different variants, and in our experiments, the DenseNet201 is selected according to the results in [25].

EfficientNet is a family of deep neural network architectures that use a neural architecture search method to uniformly scale the network’s depth, width, and input image size. EfficientNetV2 [24] aims to optimize the training speed and parameter efficiency. Regularization techniques are adaptively adjusted during training, considering different input image sizes. The authors define this particularity as *Progressive Learning with Adaptive Regularization*. In TensorFlow are implemented seven versions of EfficientNetV2. In our experiments, we select the EfficientNetV2-S variant because it has almost the same number of parameters as DenseNet201.

In order to adapt these three different CNN architectures to the skin cancer datasets, we consider two options. The first one only includes a global average pooling in the last convolutional layer of these architectures, followed by a fully connected layer. Inspired in [25], a second option considers including three convolutional layers, a global average pooling, and a batch normalization, followed by fully connected with dropout layers. A transfer learning strategy was used to train these neural architectures where initially, only the extra layers were trained by ten epochs (freezing the layers of the CNN architectures). Then, a fine tune strategy unfreezes 20% of the CNN architecture, and a new training is performed with a reduced learning rate.

3.2 TensorFlow Lite (TFLite)

TensorFlow Lite (TFLite) [5] is a lightweight deep learning framework specifically designed for deploying CNN models to mobile and embedded devices created by Google. TFLite optimizes the size and speed of the models without neglecting their performance. TFLite uses quantization methods to compress the deep learning model by using fewer bits to represent model parameters [23].

Once the model is converted to a TFLite format, the integrated development environment (IDE) of Android Studio for Android App is used to load the CNN model into the mobile device. The TFLite interpreter is in charge of running the inference of the model and producing the predictions. Then, deploying deep learning models on mobile devices is possible by combining Tensor Flow, TFLite, and Android Studio.

4 Experimental Settings and Results

In our experiments, we use two datasets presented on Kaggle that consider images of the ISIC challenges. Dataset one (DS1)¹ has 3,297 dermoscopic images. 1,800 images are classified as benign and 1,497 as malignant, respectively. Kaggle provides a data partition where 80% of the data is separated to train and 20% to test. In our experiments, the training data was re-partitioned into train and validation with a final distribution of 60% to train, 20% to validate, and 20% to test. The second dataset (DS2)² has 10,605 images. Kaggle defines 9,605 for training and 1,000 for testing. Same as the previous dataset, the training data was re-partitioned to provide a validation set. The final data split corresponds to 80% to train, 10% to validate, and 10% to test.

The training of the CNN models used in this work is performed on Google Colaboratory, a cloud-based platform with pre-installed libraries and dependencies. In our case, we use the TensorFlow library to train the CNN models. Table 1 shows the accuracy classification results of the different CNN architectures. The second column specifies if the CNN considers the three convolutional extra layers, global average pooling, and batch normalization, followed by fully connected and dropout layers. The third column indicates the number of parameters of each CNN. The fourth and fifth columns indicate the accuracy percentage achieved by each CNN.

Table 1. Accuracy results of the different neural architectures.

CNN model	Extra layers	Parameters	DS1	DS2
ResNet101	No	42,660,225	87.87%	92.10%
ResNet101	Yes	43,575,129	86.97%	92.40%
DenseNet201	No	18,832,905	86.06%	91.9%
DenseNet201	Yes	19,189,785	86.96%	90.7%
EfficientnetV2-S	No	20,332,641	85.61%	91.40%
EfficientnetV2-S	Yes	20,953,401	87.57%	91.90%

¹ <https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign>.

² <https://www.kaggle.com/datasets/hasnainjaved/melanoma-skin-cancer-dataset-of-10000-images>.

The accuracy results of the models are very similar. The best accuracy and the model with fewer parameters are highlighted in bold. ResNet101 obtains the best classification results but is the CNN with the largest number of parameters. EfficientnetV2-S and DenseNet201 obtain comparable performance, but in our implementation, it is very important to have a reduced number of parameters because our objective is to deploy the CNN model in an Android application running on a mobile device. For this reason, we select the DenseNet201 model. Figure 2 shows the confusion matrix results obtained with the DenseNet201 model considering the two datasets.

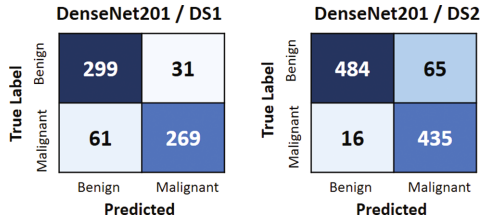


Fig. 2. Confusion matrix results

By visually inspecting the images of the datasets, we realize that some of them are very difficult to classify as benign or malignant. Figure 3 shows some of them where, despite being difficult samples, the DenseNet201 model correctly classifies them.

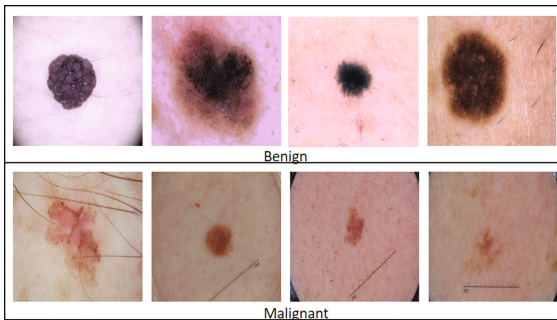


Fig. 3. Examples of difficult samples of the datasets

Once the model was trained, it was converted to a light version with TF-Lite and loaded into the mobile device using Android Studio for Android App. Figure 4 shows the final user interface designed for SkinSight with prediction results. SkinSight can load images from the smartphone gallery. With this option, we could select the testing images of DS1 and DS2 and confirm that the accuracy

performance of the model is maintained on the light version obtaining the same results reported on the confusion matrix of Fig. 2. By comparing these results with those models reported in Sect. 2, our accuracy performance is superior to most mobile apps. Only two of them achieved better results. The first only considers one dataset of few samples (1,220 images), and the second eliminates images not obtained with a dermoscopy (the ISIC dataset has images obtained with simple cameras and are commonly incorrectly classified).

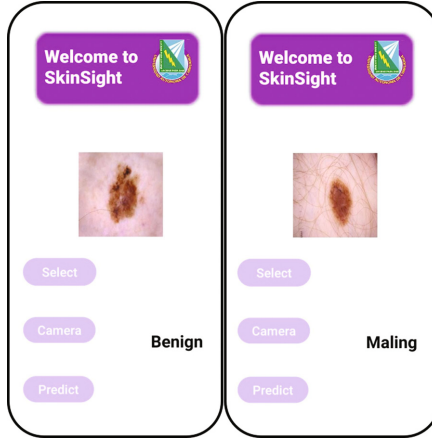


Fig. 4. Prediction results of the SkinSight app.

5 Conclusions

This paper presents the process we follow to design an Android app named SkinSight to detect melanoma automatically. First, we compare the performance of state-of-the-art CNN models trained and tested with images of two datasets of the ISIC challenge. The accuracy results obtained with EfficientnetV2-S, ResNet101, and DenseNet201 are very similar. However, considering that our objective is to develop a mobile app that medical personnel can use to diagnose suspicious cases early, we select the CNN model with the fewest parameters. The combination of using TensorFlow, TensorFlow Lite, and Android Studio offers a powerful solution for deploying deep learning models on mobile devices.

Recent models that surpass the results reported in this paper implement highly cost pre-processing techniques to remove noise and artifacts from the images. Also, some of these publications stack more than five machine learning algorithms, but the improvement is only 3% compared to our implementations. Considering that our SkinSight app is designed to be used by the medical sector with limited resources, we bear in mind a balance between accurate classification results and a few parameters of the model. In this paper, we only perform the

testing of SkinSight with images already analyzed by specialists. Because we want to bring this tool closer to rural areas of our location, our next step is to work with local medical doctors and patients already diagnosed with this disease and test the app in a real environment to identify how to handle different skin tonalities and factors not considered on the ISIC dataset.







References

1. Estadísticas importantes sobre el cáncer de piel tipo melanoma. <https://www.cancer.org/es/cancer/tipos/cancer-de-piel-tipo-melanoma/acerca/estadisticas-clave.html>. Accessed 13 Jun 2023
2. ISIC archive. <https://www.isic-archive.com/#!/topWithHeader/wideContentTop/main>
3. Skin cancer foundation. <https://www.skincancer.org/skin-cancer-information/melanoma/>. Accessed 13 Jun 2023
4. Skin cancer foundation. <https://gco.iarc.fr/today/fact-sheets-populations>. Accessed 20 Jun 2023
5. TensorFlow Lite. <https://www.tensorflow.org/lite?hl=es-419>. Accessed 01 Jun 2023
6. Arani, S., Zhang, Y., Rahman, M., Yang, H.: Melanlysis: a mobile deep learning approach for early detection of skin cancer. In: 2022 IEEE 28th International Conference on Parallel and Distributed Systems (ICPADS), Los Alamitos, CA, USA, January 2022, pp. 89–97. IEEE Computer Society (2022). <https://doi.org/10.1109/ICPADS56603.2022.00020>
7. Bansal, P., Garg, R., Soni, P.: Detection of melanoma in dermoscopic images by integrating features extracted using handcrafted and deep learning models. *Comput. Ind. Eng.* **168**, 108060 (2022). <https://doi.org/10.1016/j.cie.2022.108060>. <https://www.sciencedirect.com/science/article/pii/S0360835222001309>
8. Benyahia, S., Meftah, B., Lezoray, O.: Multi-features extraction based on deep learning for skin lesion classification. *Tissue Cell* **74**, 101701 (2022). <https://doi.org/10.1016/j.tice.2021.101701>
9. Cassidy, B., Kendrick, C., Brodzicki, A., Jaworek-Korjakowska, J., Yap, M.H.: Analysis of the ISIC image datasets: usage, benchmarks and recommendations. *Med. Image Anal.* **75**, 102305 (2022). <https://doi.org/10.1016/j.media.2021.102305>
10. Dai, X., Spasić, I., Meyer, B., Chapman, S., Andres, F.: Machine learning on mobile: an on-device inference app for skin cancer detection. In: 2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC), pp. 301–305 (2019). <https://doi.org/10.1109/FMEC.2019.8795362>
11. Deng, T.: A survey of convolutional neural networks for image classification: models and datasets. In: 2022 International Conference on Big Data, Information and Computer Network (BDICN), pp. 746–749 (2022). <https://doi.org/10.1109/BDICN55575.2022.00145>
12. Ech-Cherif, A., Misbhauddin, M., Ech-Cherif, M.: Deep neural network based mobile dermoscopy application for triaging skin cancer detection. In: 2019 2nd International Conference on Computer Applications and Information Security (ICCAIS), pp. 1–6 (2019). <https://doi.org/10.1109/CAIS.2019.8769517>
13. Emam Ananna, M., Nayeem, J., Jahangir Alam, M., Islam, S.: Skin cancer detection using machine learning framework with mobile application. In: 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI), pp. 1073–1080 (2023). <https://doi.org/10.1109/ICOEI56765.2023.10125640>

14. Emuoyibofarhe, J., Ajisafe, D.: Early skin cancer detection using deep convolutional neural networks on mobile smartphone. *Int. J. Inf. Eng. Electron. Bus.* **12**, 21–27 (2020). <https://doi.org/10.5815/ijieeb.2020.02.04>
15. Francese, R., Frasca, M., Risi, M., Tortora, G.: A mobile augmented reality application for supporting real-time skin lesion analysis based on deep learning. *J. Real-Time Image Proc.* **18**(4), 1247–1259 (2021). <https://doi.org/10.1007/s11554-021-01109-8>
16. Gajera, H.K., Nayak, D.R., Zaveri, M.A.: A comprehensive analysis of dermoscopy images for melanoma detection via deep CNN features. *Biomed. Sig. Process. Control* **79**, 104186 (2023). <https://doi.org/10.1016/j.bspc.2022.104186>
17. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778 (2016). <https://doi.org/10.1109/CVPR.2016.90>
18. Huang, G., Liu, Z., Weinberger, K.Q.: Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2261–2269 (2017)
19. Jaikishore, C., Udutalapally, V., Das, D.: AI driven edge device for screening skin lesion and its severity in peripheral communities. In: 2021 IEEE 18th India Council International Conference (INDICON), pp. 1–6 (2021). <https://doi.org/10.1109/INDICON52576.2021.9691666>
20. Kousis, I., Perikos, I., Hatzilygeroudis, I., Virvou, M.: Deep learning methods for accurate skin cancer recognition and mobile application. *Electronics* **11**(9) (2022). <https://doi.org/10.3390/electronics11091294>
21. Polat, Ö., Kartal, M.S.: Detection of benign and malignant skin cancer from dermoscopic images using modified deep residual learning model. *Artif. Intell. Theor. Appl.* **2**, 10–18 (2022)
22. Pereira, P.M., et al.: Melanoma classification using light-fields with morlet scattering transform and CNN: surface depth as a valuable tool to increase detection rate. *Med. Image Anal.* **75**, 102254 (2022). <https://doi.org/10.1016/j.media.2021.102254>
23. Shi, Y., Yang, K., Yang, Z., Zhou, Y.: Model compression for on-device inference, chap. 5. In: Shi, Y., Yang, K., Yang, Z., Zhou, Y. (eds.) *Mobile Edge Artificial Intelligence*, pp. 71–82. Academic Press (2022)
24. Tan, M., Le, Q.V.: EfficientNetV2: smaller models and faster training. In: Meila, M., Zhang, T. (eds.) *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18–24 July 2021, Virtual Event*, vol. 139, pp. 10096–10106. *Proceedings of Machine Learning Research (PMLR)* (2021)
25. Zia Ur Rehman, M., Ahmed, F., Alsuhibany, S.A., Jamal, S.S., Zulfıqar Ali, M., Ahmad, J.: Classification of skin cancer lesions using explainable deep learning. *Sensors* **22**(18) (2022). <https://doi.org/10.3390/s22186915>



Classification of Acid-Base Disorders Through Machine Learning

Rodrigo Ruiz de la Peña Martínez¹ , Eutzhel André Del Val Saucedo¹ ,
Paola Aidee de la Cruz Gallardo¹ , Carlos Eduardo Cañedo Figueroa¹ ,
Abimael Guzmán Pando² , and Natalia Gabriela Sámano Lira¹ 

¹ Facultad de Medicina y Ciencias Biomédicas, Universidad Autónoma de Chihuahua,
31125 Chihuahua, Mexico
ccanedo@uach.mx

² Laboratorio de Física Química Computacional, Facultad de Medicina y Ciencias Biomédicas,
Universidad Autónoma de Chihuahua, 31125 Chihuahua, Mexico

Abstract. The present research aims to develop a tool in the form of an algorithm that can provide an accurate diagnosis of a patient's acid-base balance without the need for manual calculations by the attending physician. During the research, three different algorithms (Bayesian, KNN and a neural network) were used, tested and compared in order to achieve a reliable result, speeding up the diagnostic process for the patient and reducing the human error that can arise from manual calculations. The results show that the Bayesian algorithm had the lowest performance achieved with an MF1 of 0.7850, followed by the KNN algorithm with an MF1 of 0.8553, the next was the neural network which obtained an MF1 of 0.9711. Finally, an algorithm assembled by the three mentioned above was tested generating an MF1 of 1, which was tested on 70 data samples. This suggests that the design can be used for the classification of acid-base problems.

Keywords: ANN · Acid-base diagnostic · Classification

1 Introduction

One of the inherent attributes to blood is its characteristic pH level, which is indicated in any solution by the pH scale. This scale ranges from 0 for strong acids to 14 for strong bases, with neutral value in the middle of both [1].

Normal pH values found in blood range from 7.35 to 7.45, typically staying around 7.40, which is considered slightly basic [1]. The measurement of these pH values is performed by the physician to assess the acid-base balance of a patient, and the values of carbon dioxide levels (CO₂) as well as blood bicarbonate levels (HCO₃⁻) are used [2].

The acid-base balance mentioned earlier is constantly maintained by the precise action of both the renal system and the respiratory system, preserving the harmony between CO₂ and HCO₃⁻. Any deviation from the normal values would also imply an alteration in the functions of various organs in the body, leading to physiological disturbances and even death [1, 2]. The body also has buffering systems that can combine

with acids or bases, depending on the situation at hand, helping to prevent rapid changes in pH levels [2].

To date, the Henderson-Hasselbalch equation, in conjunction with various other equations, is used to express the patient's pH, which utilizes the values of carbon dioxide partial pressure (pCO_2) and HCO_3^- , determining the relationship between acids and bases. This equation aids in classifying different acid-base disorders that may occur in the body and helps define whether these problems are respiratory or metabolic in nature [2]. The Henderson-Hasselbalch equation is written as Eq. (1)

$$pH = 6.1 + \log \left[\frac{(HCO_3^-)}{(0.03xpCO_2)} \right] \quad (1)$$

Currently, there are also acid-base maps available, especially the Du Bose basic map, which help visually identify the specific disorder [2]. The acid-base map is only helpful from a graphical perspective, but it's not practical for daily patient care use.

The problem that arises when using the Henderson-Hasselbalch equation and identification maps is that, when it comes to making a precise diagnosis, the physician is forced to manually perform these calculations or measurements once they have the patient's laboratory data. This process is unnecessarily time-consuming and can be tedious, leading to erroneous results. Therefore, there is now a search for the automation of this process, completely eliminating the possibility of human calculation errors. This automation is seen as a tool for healthcare personnel.

2 Methodology

2.1 Database

The information used in each of the developed algorithms was gathered from a synthetic database (SDB), which is artificially generated data rather than data based on real-world events [3]. This database was created by the collaborators of this work based on the acid-base map, which serves as a graphical tool for doctors to diagnose acid-base disorders. This enables them to visualize a graphical representation showing potential chemical states between blood concentrations of HCO_3^- , pH, and CO_2 .

The following acronyms were used to facilitate the later mention of the characteristic disorders, which correspond to the areas that compose the DuBose acid-base map. These classes are: normal state, *SN*, metabolic acidosis state, *SACM*, acute respiratory acidosis, *SACRA*, chronic respiratory acidosis, *SACRC*, metabolic alkalosis, *SAM*, acute respiratory alkalosis, *SARA*, and chronic respiratory alkalosis, *SARC*. Random points of the DuBose acid-base map distributed across these seven different regions were used to create the data for each of the possible patient conditions, resulting in a total of 268 data points $p(x,y)$, being x the HCO_3^- and y the pCO_2 concentration respectively. From these points we randomly utilized 198 for training and 70 for testing considering the seven state classes. Table 1 shows the partitioning of the database data [2].

Table 1. Composition of the data set

Full database	Data partitioning	Data consider by class	acronym
268 randomly selected points of the DuBose acid-base map	198 for training	17 -Normal state	SN
		33 -Metabolic acidosis	SACM
		33 -Acute respiratory acidosis	SACRA
		33 -Chronic respiratory acidosis	SACRC
		33 -Metabolic alkalosis	SAM
		33 -Acute respiratory alkalosis	SARA
		16 -Chronic respiratory alkalosis	SARC
	70 for testing	10 for each class	

2.2 Neural Network

The ANN models have the specific architecture format, which is inspired by a biological nervous system. ANN models are made up of neurons in a complex, nonlinear manner, just like the human brain. Weighted links are used to connect the neurons. [4]. In biological neural networks, learning is primarily driven by two forms of synaptic plasticity: long-term potentiation (LTP) and long-term depression (LTD). LTP strengthens the connections between neurons when they consistently exhibit correlated activity, while LTD weakens the connections when the activity is uncorrelated or weakly correlated. These changes in synaptic strength enable neurons to form new connections and modify existing ones, which is essential for learning and memory formation [5].

ANNs on the other hand are made up of an input node layer and an output node layer coupled by one or more hidden node layers. By activating functions, input layer nodes transmit information to hidden layer nodes, which then either activate or do nothing in response to the evidence. When the value of a certain node or collection of nodes in the hidden layer hits a certain threshold, a value is transmitted to one or more nodes in the output layer. The hidden layers apply weighting functions to the evidence. A lot of examples (data) must be used to train ANNs. [6].

For the realization of this project, a feedforward ANN with a single hidden layer consisting of 5 neurons was used, with the following hyperparameters: a learning rate of 0.1, 1000 epochs, and an error tolerance of $1e-29$ these values were selected using a grid search technique, which shows the most optimal hyperparameters that will increase the efficiency of the algorithm. The value of the bias remains 1, but the values of their weights are being constantly updated. Figure 1 illustrate the ANN proposed model.

The output provided by the neural network determines the patient's acid-base balance situation as Eq. (2), Where: $Output_1$ represents a normal state, $Output_2$, chronic respiratory acidosis, $Output_3$, metabolic alkalosis, $Output_4$, acute respiratory alkalosis, $Output_5$, chronic respiratory alkalosis, $Output_4$, metabolic acidosis, and $Output_6$, acute

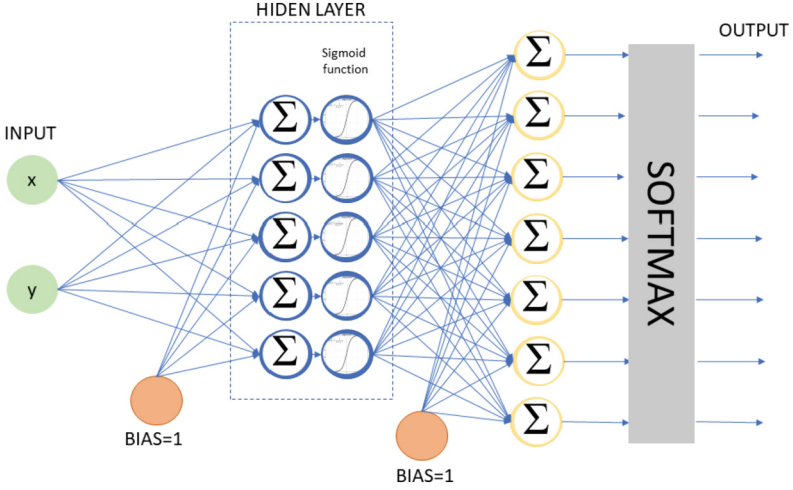


Fig. 1. Neural network

respiratory acidosis. Also, z_i stands for the final softmax output vector obtained after evaluating each i -th output, being N the total number of ANN outputs evaluated. Finally, arg max function is applied to obtain the winner class.

$$Z_i = \frac{e^{Output_i}}{\sum_{n \in N} e^{Output_n}}$$

$$Class = \arg \max(\{Z_i\}) \quad (2)$$

2.3 Bayesian Algorithm

The Bayesian algorithm (AB) functions as a probabilistic classifier that takes into account a particular feature and assigns a label to an element. AB operates using probability and employs empirical reasoning based on training data [7]. Its functionality lies in the correlation between a sample and its membership in a specific class, based on a set of features that are associated with different classes, each having its representative characteristics [8].

The AB training was performed using the set of classes $C = \{SN, SACM, SACRA, SACRC, SAM, SARA, SARC\}$ shown in Table 1. The probability of a training class P_C was obtained using Eq. (3), where T represents the total number of data points $p(x,y)$ consider for a c -th training class.

$$P_c = \frac{T_c}{\sum_{c \in C} T_c} \quad (3)$$

Also, BA involves Eqs. (4) and (5) that correspond to the mean and variance respectively.

$$\overline{p_c^l} = \frac{\sum_{j=1}^{T_c^l} p_{j,c}^l}{T_c^l}; l = \{x, y\} \quad (4)$$

$$(\sigma_c^l)^2 = \frac{\sum_{j=1}^{T_c^l} (p_{c,j}^l - \overline{p_c^l})^2}{T_c^l} \quad (5)$$

where j stands for the j -th point p evaluated on the characteristic l . Once the results of Eqs. (4) and (5) were obtained, Eq. (6) was used, where $P(c|p^l)$ correspond to the probability of a class c given a point p with features l [8, 9].

$$P(c|p^l) = \frac{1}{\sqrt{2\pi(\sigma_c^l)^2}} e^{-\frac{(p^l - \overline{p_c^l})^2}{2(\sigma_c^l)^2}} \quad (6)$$

Subsequently, Eq. (7) was implemented, which determines the relationship between all the probabilities obtained from Eqs. (3) and (6) for each of the seven classes.

$$Pr_c = \left(\prod P(c|p^l) \right) \cdot (P_c) \quad (7)$$

Equation (8) was used to obtain the evidence, which is the sum of all the different results from Eq. (7). This allows for the calculation of the likelihood that a given sample belongs to a particular class, based on Eq. (9) [8].

$$Ev = \sum_{c \in C} Pr_c \quad (8)$$

$$Ps_c = \frac{Pr_c}{Ev} \quad (9)$$

Finally, the AB determines the class to which it belongs by considering the Eq. (10)

$$Class = \operatorname{argmax}(Ps_c) \quad (10)$$

where $Class$ can have values from 1 to 7 according to C .

2.4 KNN

The KNN algorithm, also known as the k-nearest neighbors' algorithm, is a machine learning algorithm. Similar to the previous two algorithms, it is supervised, as the data provided to it already has labeled output values, allowing for the creation of a model to predict new input data [10].

The prediction of new input data is done by calculating the distance between the sample data and the data for each of the seven classes used in the training set [8]. In this case 5 neighbors (K) were used to classify the test dataset. This is done using

Euclidian distance between two samples p^l and q^l depicted on Eq. (11), being D the total components of the sample p^l , in this case $D = 2$ for the two components $l = \{x, y\}$.

$$d_{p,q} = \sqrt{\sum_{l=1}^D |p^l - q^l|} \quad (11)$$

3 Results

The results of the evaluation on each previously trained algorithm are presented. The metrics show the precision, recall, and F1 scores (calculated using Eq. 12) for each class. The MF1 score (calculated using Eq. 13) is also shown for the overall results of the analysis algorithm. For this evaluation, we used 70 test data (10 per class) shown on Table 1.

$$F1_c = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (12)$$

$$MF1 = \frac{\sum_{c=1}^7 F1_c}{7} \quad (13)$$

Table 2 shows the metrics of the ANN classifier, with an MF1 score of 0.97114 (97.114%). This value is affected by the precision and recall scores for class one and class six. Where the precision score for class one is 0.8 (80%) and the recall for class six is 0.83 (83%).

Table 2. Effectiveness metrics for ANN

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
Precision	0.8	1	1	1	1	1	1
Recall	1	1	1	1	1	0.833	1
F1	0.8888	1	1	1	1	0.9090	1
MF1	0.97114						

Table 3 shows the metrics for the KNN algorithm. The data shows a variation for each class, with the best-evaluated classes being class one and class six, both with a F1 score of 0.9523 (95.23%). However, this algorithm has an MF1 score that is 11.584% lower than the ANN algorithm.

Table 4 shows that the classes two and three were perfectly classified, with an F1 score of 1.0 (100%). However, the class six was not classified at all, with an F1 score of 0.0 (0%). This data affected the MF1 score, which was 0.785036 (78.5036%).

Table 3. Effectiveness metrics for KNN

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
Precision	1	1	0.8	0.7	0.8	1	0.7
Recall	0.9090	0.7142	1	0.7777	0.8	0.9090	1
F1	0.9523	0.8333	0.8888	0.73684	0.8	0.9523	0.82352
MF1	0.8553						

The lower MF1 score is due to the fact that the class six was not classified at all. This means that the algorithm was not able to identify any of the instances of class six in the test dataset.

Table 4. Effectiveness metrics for Bayes

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7
Precision	1	1	1	1	1	0	1
Recall	0.83333	1	1	1	0.76923	1	0.625
F1	0.90909	1	1	0.94736	0.86956	0	0.7692
MF1	0.785036						

The neural network algorithm was the most efficient of the three individual algorithms, with an MF1 score of 97.11%. This suggests that the neural network can be used to classify the seven classes. However, there is an approximate 3% error rate in the classification of a data point. This problem could be fixed by retraining each algorithm with a larger dataset and comparing the future results with those obtained in this work.

4 Conclusion



At the end of the project, we managed to create a classification algorithm that, unlike the conventional method based on the Henderson-Hasselbach equation, uses only two parameters for classification. Additionally, our classification algorithm provided us with a 97% effectiveness score with its most reliable method, this being the neural network algorithm, which can serve as a prototype and later as a replacement for highly fallible and unpredictable human interference. This given that machine learning techniques provide numerous advantages over traditional diagnostic methods in the classification of acid-base disorders. These benefits encompass increased efficiency, enhanced accuracy, objectivity, scalability, decision support, early detection, and seamless integration with existing healthcare systems. The utilization of machine learning models enables the rapid processing of substantial amounts of data, facilitating prompt diagnosis in time-sensitive scenarios.

References

1. Lewis, J.L.: III Introducción al equilibrio ácido-básico. Manual MSD Versión Para Público General (2023). <https://www.msmanuals.com/es-mx/hogar/trastornos-hormonales-y-metab%C3%B3licos/equilibrio-%C3%A1cido-b%C3%A1sico/introducci%C3%B3n-al-equilibrio-%C3%A1cido-b%C3%A1sico>
2. Del Pilar Triana Reyes, M., Pérez, V.H.E., Durán, D.P.J.: Chapter two - análisis de gases en la sangre. In: Estupiñan, V.H. (Ed.) Bases Para la Interpretación y Análisis de Gases Arteriovenosos. Cali, Colombia: Editorial Universidad Santiago de Cali; 2020, pp. 29–100 (2020). <https://doi.org/10.35985/9789585583801>
3. Durugkar, S.R., Raja, R., Nagwanshi, K.K., Kumar, S.: Introduction to data mining. Data Min. Mach. Learn. Appl. 1–19. <https://doi.org/10.1002/9781119792529.ch1>
4. Malekian, A., Chitsaz, N.: Concepts, procedures, and applications of artificial neural network models in streamflow forecasting. In: Advances in Streamflow Forecasting, pp. 115–147. Elsevier (2021). <https://doi.org/10.1016/b978-0-12-820673-7.00003-2>
5. Shehab, M., et al.: Chapter eight - artificial neural networks for engineering applications: a review. In: Elsheikh, A.H., Elaziz, M.E.A. (Eds.) Artificial Neural Networks for Renewable Energy Systems and Real World Applications, pp. 189–206. Academic Press (2022). ISBN 9780128207932. <https://doi.org/10.1016/B978-0-12-820793-2.00003-3>
6. Sadiq, R., Rodriguez, M.J., Mian, H.R.: Empirical Models to Predict Disinfection by-Products (DBPs) in Drinking Water: an Updated Review, pp. 324–338. Elsevier (2019). <https://doi.org/10.1016/b978-0-12-409548-9.11193-5>
7. Marimuthu, R., Shivappriya, S.N., Saroja, M.N.: Chapter 14 - a study of machine learning algorithms used for detecting cognitive disorders associated with dyslexia. In: Jude, H.D. (ed.) Handbook of Decision Support Systems for Neurological Disorders, pp. 245–262. Academic Press (2021). ISBN 9780128222713. <https://doi.org/10.1016/B978-0-12-822271-3.00008-6>
8. Figueroa, C.C., Chávez, H.G.: Diseño de algoritmo compuesto por machine learning y un modelo probabilístico para la detección de diabetes. In: Memorias del Congreso Nacional de Ingeniería Biomédica, vol. 8, No. 1, pp. 57–60 (2021). <http://memoriascnib.mx/index.php/memorias/article/view/828/488>
9. Chaitanya, B.K., Yadav, A., Pazoki, M., Abdelaziz, A.Y.: Chapter 8 - a comprehensive review of islanding detection methods. In: Zobia, A.F., Abdel Aleem, S.H.E. (Eds.) Uncertainties in Modern Power Systems, pp. 211–256. Academic Press (2021). ISBN 9780128204917. <https://doi.org/10.1016/B978-0-12-820491-7.00008-6>
10. Team, D.: Descubra el algoritmo KNN : un algoritmo de aprendizaje supervisado. Formation Data Science | DataScientest.com (2022). <https://datascientest.com/es/que-es-el-algoritmo-knn>



Identification of Pneumonia with X-ray Images Using Deep Transfer Learning

Zarah Campos-Lopez¹ , Jose Diaz-Roman¹ , Boris Mederos-Madrado² ,
Nelly Gordillo-Castillo¹ , Juan Cota-Ruiz¹ , and Jose Mejia-Muñoz¹ 

¹ Departamento de Ingeniería Eléctrica y Computación, Instituto de Ingeniería y Tecnología, Universidad Autónoma de Ciudad Juárez, Ciudad Juárez Ch. 32310, Mexico

david.roman@uacj.mx

² Departamento de Física y Matemáticas, Instituto de Ingeniería y Tecnología, Universidad Autónoma de Ciudad Juárez, Ciudad Juárez Ch. 32310, Mexico

Abstract. Radiology plays an essential role in the identification of pathologies; however, image interpretation and the guarantee of accurate diagnoses continue to represent a challenge that involves expert radiologists. This study proposes a model to identify pneumonia in chest X-rays using the deep transfer learning technique, where five pre-trained network architectures and a classifier are tested. The images used in this work are categorized into bacterial pneumonia, viral pneumonia, and normal cases. The pre-trained models used include DenseNet201, MobileNet, VGG16, VGG19 and ResNet50. A support vector machine is used as a classifier. Results show that the ResNet50 model performs best in binary classification (pneumonia vs. non-pneumonia) with 98.1% accuracy and 98.7 F1-score. For multiclassification, VGG19 performs best with an accuracy of 84.7% and an average F1-score of 81.1%. The methodology employed proved to be competent and outstanding when compared to other studies in the state of the art.

Keywords: Pneumonia · X-ray images · Deep Transfer Learning · Support Vector Machine

1 Introduction

The probability of lung disease is notably elevated, particularly in lower-middle-income nations undergoing development, where countless individuals are exposed to poverty and air pollution. Based on estimations by the World Health Organization (WHO), over 4 million premature deaths happen every year due to household air pollution, which can lead to the emergence of health issues like asthma and pneumonia. Pneumonia is an infectious condition impacting either one or both lungs, resulting in the air sacs, known as alveoli, becoming filled with fluid or pus. Pneumonia can be caused by bacteria, viruses, or fungi. This illness poses a significant public health concern and stands as a leading contributor to illness and death in both Mexico and worldwide. It is estimated that pneumonia caused 921,000 deaths in children under 5 years of age in 2015 globally, which represents 15% of all deaths in children under 5 years of age worldwide, a considerable figure related to this disease [1].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024

J. de J. A. Flores Cuautle et al. (Eds.): CNIB 2023, IFMBE Proceedings 96, pp. 32–40, 2024.

https://doi.org/10.1007/978-3-031-46933-6_4

To diagnose pneumonia, a doctor reviews your medical history, performs a physical exam, and orders diagnostic tests, such as a chest X-ray. Diagnosing pulmonary diseases like pneumonia from chest X-rays or computed tomography scans can be challenging, often needing experienced physicians or radiology specialists to interpret the signs due to similar-looking diseases in the images. This sometimes requires additional time or studies for an accurate diagnosis. Moreover, issues like low resolution or varying characteristics of the images can further complicate identification. Therefore, developing diagnostic systems to assist in decision-making for lung disease diagnosis is valuable [2].

In the last decade, deep learning methods, specifically convolutional neural networks (CNNs), have been chosen as they are able to automatically learn multiple invariant features in signals or images for a given task. Because of their feature extraction ability, CNNs have proven to perform well in many applications showing strong robustness against geometric distortions, skew, scale, etc. Furthermore, CNNs trained with large amounts of data (images) in a large scenario of a demanding task can be used to extract image features from another particular context and perform efficiently; this technique is known as deep transfer learning (DTL). The objective of this work is to develop an algorithm capable of identifying pneumonia in chest X-ray images based on the DTL technique, using pre-trained convolutional network architectures as feature extractors and a machine learning classification model. A recent review reported that, regarding the use of DTL in medical image classification tasks, the VGG16 and DenseNet networks have been used more frequently in lung X-ray studies [3]. In this way, the authors of this study decided to explore the performance that the different networks would have for classification and define which of them had the best performance. Thus, five networks were tested that present differences in terms of complexity, number of parameters, depth, size, etc. It is important to highlight that while there are numerous studies that utilize DTL, a majority of them are centered around fine-tuning pre-trained networks. Notably, our proposal surpasses the algorithms found in the current state of the art [4–8].

2 Methodology

This section outlines the steps taken to create and evaluate the proposed classification model. It starts by introducing the database employed, followed by clarifying the method for extracting features via pre-trained convolutional neural networks. The training of the classifier is then elaborated upon, and lastly, the conducted performance tests on the generated models are described.

The algorithms were developed using the Python programming language. The Keras library was used to handle the convolutional neural networks [9], and the Sklearn library was used to develop the support vector machine-based classifier model [10].

2.1 Database

This project utilized images sourced from the Kaggle website, a repository that offers a wide range of beneficial databases for data science projects. The specific database

used, named “Chest X-ray Images”, comprises X-ray images of patients with bacterial pneumonia, viral pneumonia, and patients without the disease [11].

The database is organized into training and test sets, and contains images labeled as normal, bacterial pneumonia and viral pneumonia. The images were stored in grayscale 8-bit depth.jpeg format, and their sizes ranged from 494×151 pixels to 2024×2036 pixels (width \times height). Table 1 below provides a summary of the number of images for each class in each set of the database.

Table 1. Database summary.

Set	Normal	Bacterial pneumonia	Viral pneumonia
Training	1068	1931	1104
Test	270	555	209

2.2 Feature Extraction by Deep Transfer Learning

Feature extraction by DTL refers to a technique in deep learning where pre-trained neural network models are used as a starting point to extract relevant features from different datasets. Transfer learning leverages the knowledge and representations learned by a model on a large dataset to improve performance on a smaller or different dataset.

In DTL, the initial layers of a pre-trained model, typically trained on a large-scale dataset (such as ImageNet), are used as a feature extractor. These initial layers are responsible for learning low-level features such as edges, textures, and shapes, which are generally applicable across various visual tasks. By freezing the parameters of these layers and removing the final classification layers, the pre-trained model can be transformed into a feature extraction network [3]. When constructing a classifier model using the technique of deep transfer learning with feature extraction, it is necessary to employ a machine learning model that is trained taking the extracted features from the pre-trained networks as input and subsequently performs inference on the test observations. Commonly used machine learning models include support vector machines, random forests, k-nearest neighbors, among others. More details about this technique can be found at [3, 12].

In this study, the pre-trained networks VGG16, VGG19, ResNet50, DenseNet201, and MobileNet were utilized for feature extraction from images. These networks were pre-trained on the ImageNet dataset, a renowned dataset for object recognition [13].

The networks were fed resized database images of 224×224 pixels, with intensity values scaled between 0 and 1. Each network produced an output feature vector with sizes of 512 (VGG16 and VGG19), 1024 (DenseNet and MobileNet), and 2048 (ResNet50) for each image.

2.3 Classifier Training

A Support Vector Machine (SVM) with a Gaussian radial basis kernel function was chosen as the classification model. Five models were trained, each corresponding to a

network used for feature extraction. A matrix of feature vectors from the training set images was used as input for SVM training. The regularization hyperparameter ‘C’ of the SVM loss function was determined through a grid search with possible values, utilizing 5-fold cross-validation.

2.4 Experimentation and Evaluation

Two experiments were conducted: the first focused on creating multiclassification models to distinguish normal, bacterial pneumonia, and viral pneumonia classes within the database. The second experiment involved binary classification models to identify normal images from those with pneumonia, without specifying the type of pneumonia.

For the second experiment, bacterial and viral pneumonia image files were pooled, and the task consisted of discriminating patients with pneumonia vs. normal. Since in this experiment, the number of total pneumonia images (3875) almost tripled the number of normal case images (1341) in the training set, it was decided to artificially create a larger number of normal case images with the intention of balancing the number of images from both classes. To achieve this, the data augmentation technique was applied. In this process, a subset of normal class images ($N = 2540$) underwent random rotations (within the range of 0 to $\pm 15^\circ$), zooming (within the range of 0 to $\pm 15\%$), horizontal and vertical shifts (within the range of 0 to $\pm 10\%$), and shearing (within the range of 0 to $\pm 15\%$). These augmented images were then incorporated into the training set alongside the pre-existing ones.

To assess the effectiveness of the constructed classifier models, evaluation metrics including classification accuracy (Acc), recall (Rec), precision (Pre), and F1-score (F1) were computed solely using images from the test set. Equations 1 to 4 were employed to calculate these metrics.

$$Acc = \frac{tp + tn}{tp + tn + fp + fn} \quad (1)$$

$$Rec = \frac{tp}{tp + fn} \quad (2)$$

$$Pre = \frac{tp}{tp + fp} \quad (3)$$

$$F1 = \frac{2 \cdot tp}{2 \cdot tp + fp + fn} = 2 \frac{Rec \cdot Pre}{Rec + Pre} \quad (4)$$

where tp and tn mean true positives and true negatives respectively, and fp and fn mean false positives and false negatives respectively. To calculate the metrics in the multilabel classification task, the strategy of one vs. all was used.

3 Results

This section presents and describes the results of this study. Various deep learning architectures were utilized for feature extraction and subsequent training and testing of classification models. The performance of these models in relation to the two proposed experiments is presented, employing the performance metrics discussed earlier.

Table 2 shows the performance obtained by the multiclassification models using the different pre-trained neural network architectures.

Table 2. Performance of multiclassification models for the three classes of images.

Network	Class	Acc (%)	Pre (%)	Rec (%)	F1 (%)
DenseNet201	Bacterial pneumonia	79.5	80.5	87.8	84.0
	Viral pneumonia		96.1	34.9	51.2
	Normal		74.2	97.0	84.1
MobileNet	Bacterial pneumonia	81.0	79.5	92.3	85.4
	Viral pneumonia		94.3	39.2	55.4
	Normal		80.2	90.0	84.8
VGG16	Bacterial pneumonia	84.0	85.9	87.8	86.8
	Viral pneumonia		70.5	56.0	62.4
	Normal		87.7	97.8	92.5
VGG19	Bacterial pneumonia	84.7	85.7	89.6	87.6
	Viral pneumonia		71.7	54.5	62.0
	Normal		89.8	98.2	93.8
ResNet50	Bacterial pneumonia	84.4	85.9	87.8	86.8
	Viral pneumonia		66.7	57.4	61.7
	Normal		92.7	98.5	95.5

As can be seen in Table 2, the model trained with the features extracted using the VGG19 network presented the highest accuracy in the classification of the 3 classes of images. While this model demonstrates high sensitivity (recall) in detecting patterns of bacterial pneumonia and normal cases with 89.6% and 98.2% respectively, its sensitivity in identifying images of viral pneumonia is notably low at 54.5%. In fact, all models exhibited strong sensitivity for normal cases and bacterial pneumonia, yet their sensitivity for identifying viral pneumonia was consistently lower. Notably, the model utilizing ResNet50 exhibited the highest sensitivity for this class, albeit reaching only 57.4%. On the other hand, the DenseNet201 and MobileNet models demonstrated high precision in detecting viral pneumonia at 96.1% and 94.3% respectively. However, their low sensitivity indicates they may be misclassifying many viral pneumonia cases as normal or bacterial pneumonia. This is further supported by their low F1-score values.

Figure 1 shows the confusion matrix for evaluating the model using the VGG19 network. It reveals that the model committed 5 errors in identifying normal patient images, where one was wrongly labeled as viral pneumonia and 4 as bacterial pneumonia. For bacterial pneumonia identification, the model inaccurately classified 14 instances as normal and 44 as viral pneumonia. The main confusion within the model is evident between the two pneumonia categories.

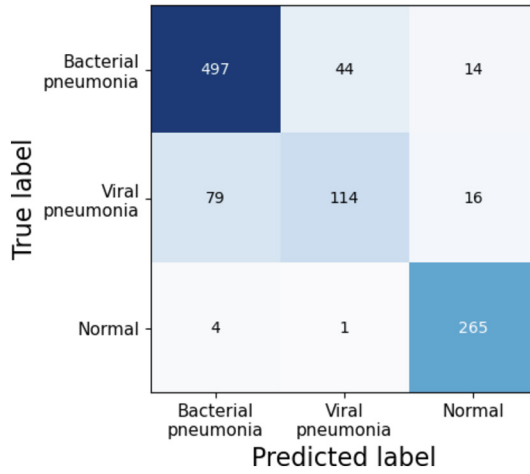


Fig. 1. Confusion matrix of model evaluation using the VGG19 network

A summary of the evaluation of the models in relation to the binary classification task (normal vs. pneumonia) is presented in Table 3.

Table 3. Performance of models for binary classification (normal vs. pneumonia).

Network	Acc (%)	Pre (%)	Rec (%)	F1 (%)
DenseNet201	95.2	97.1	96.3	96.7
MobileNet	95.9	97.6	96.9	97.2
VGG16	96.9	99.7	96.1	97.9
VGG19	96.9	99.1	96.7	97.9
ResNet50	98.1	99.3	98.0	98.7

As indicated in Table 3, the ResNet50 model exhibited the highest performance, boasting a precision of over 99.3% in identifying positive pneumonia cases (regardless of the pneumonia type) and an F1-score of 98.7%. Notably, all models demonstrated outstanding performance, achieving sensitivities exceeding 96% and F1 scores surpassing 96%.

Figure 2 shows the confusion matrix of the model test using the ResNet50 network. Notably, model misclassification was observed in only 20 out of a total of 1,034 images, resulting in a classification error rate of 1.9%. It is worth highlighting that the model also exhibits substantial specificity, achieving a remarkable value of 98.1%.

Prior to training the models, data augmentation was applied to balance the numbers of normal and pneumonia cases using the methodology described in Sect. 2.4. Initial results without data augmentation indicated that the models never exceeded 91% accuracy (with

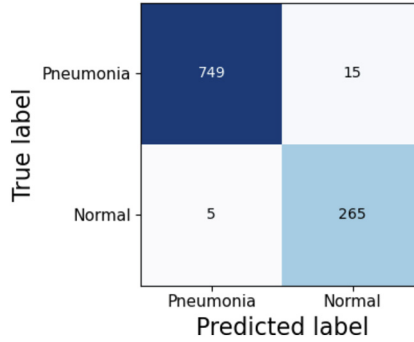


Fig. 2. Confusion matrix of the ResNet50 model for binary classification.

specificity values below 80%). Consequently, the decision was made to employ the data augmentation technique.

Table 4 presents a summary of the best results obtained in previous studies that utilize the same database as our work to perform the binary classification task (normal vs. pneumonia).

Table 4. Comparison with previous works

Paper	Method / model	Acc (%)	Pre (%)	Rec (%)	F1 (%)
Laing et al. [4]	Fine-tuning / VGG19	90.5	89.1	96.7	92.7
Jain et al. [5]	Training a Custom CNN	92.31	--	98	94
Patel et al. [6]	Fine-tuning / VGG19	91.19	95	92	88
Sharma and Guleria [14]	Feature extraction / VGG16 + SVM	92.15	94.28	93.08	93.7
Dey et al. [8]	Feature extraction + handcrafted features / VGG19 + Random-Forest	97.94	95.02	97.56	96.27
Hashmi et al. [7]	Fine-tuning / Weighted combination of five networks	98.43	98.26	99	98.26
Our work	Feature extraction / ResNet50 + SVM	98.1	99.3	98.0	98.7

Table 4 displays that our model, based on ResNet50 + SVM, achieved the highest precision and F1-score (99.3% and 98.7%) among reviewed studies. It was second only to Hossain et al. in accuracy and sensitivity. Notably, our model's advantage lies in simplicity, using features from one network, unlike Hossain's model which combines outputs from five networks: ResNet18, Xception, InceptionV3, DenseNet121, and MobileNetV3.

4 Conclusions

Identifying lung diseases like pneumonia in X-ray images through automated computer vision algorithms is challenging due to the resemblance of patterns and characteristics to other diseases in these images.

This project developed a model to identify pneumonia in X-ray images, utilizing deep transfer learning for feature extraction and an SVM classifier. Two tasks were executed: a multi-classification task for identifying bacterial, viral pneumonia, and normal images, and a binary classification task for pneumonia versus normal cases. The VGG19 model showed the best performance in the three-class classification task with an average precision of 82.4%, sensitivity of 80.8%, and F1-score of 81.1%. Furthermore, the ResNet50-based model outperformed other models in both this work and state-of-the-art models for binary classification, achieving a precision of 99.3% and an F1 score of 98.7%.

The results of this project demonstrate the reliability of the use of models designed using the deep transfer learning technique for the task of detecting pneumonia in X-ray images. Bearing in mind that timely identification of the disease is essential for adequate treatment and thus achieve the recovery of the patient, the development of this type of models would open the possibility of clinical applications in which they can serve as a support tool in decision-making for the diagnosis of this type of pathology.


References

1. McAllister, D., et al.: Global, regional, and national estimates of pneumonia morbidity and mortality in children younger than 5 years between 2000 and 2015: a systematic analysis. *Lancet Glob. Heal.* **7**(1), e47–e57 (2019)
2. Bharati, S., Podder, P., Mondal, M.: Hybrid deep learning for detecting lung diseases from X-ray images. *Informatics Med. Unlocked* **20**, 100391 (2020)
3. Morid, M., Borjali, A., Del Fiol, G.: A scoping review of transfer learning research on medical image analysis using ImageNet. *Comput. Biol. Med.* **128**, 104115 (2021)
4. Liang, G., Zheng, L.: A transfer learning method with deep residual network for pediatric pneumonia diagnosis. *Comput. Methods Programs Biomed.* **187**, 104964 (2020)
5. Jain, R., Nagrath, P., Kataria, G., Kaushik, V., Hemanth, D.: Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning. *Measurement* **165**, 108046 (2020)
6. Patel, M., Sojitra, A., Patel, Z., Bohara, M.: Pneumonia Detection Using Transfer Learning. *Int. J. Eng. Res. Technol.* **10**(10), 252–261 (2021)
7. Hashmi, M., Katiyar, S., Keskar, A., Bokde, N., Geem, Z.: Efficient pneumonia detection in chest xray images using deep transfer learning. *Diagnostics* **10**(6), 417 (2020)
8. Dey, N., Zhang, Y., Rajinikanth, V., Pugalenthi, R., Raja, N.: Customized VGG19 architecture for pneumonia detection in chest X-rays. *Pattern Recognit. Lett.* **143**, 67–74 (2021)
9. Chollet F., et al.: Keras, 2015. <https://keras.io>, Accessed: 15 May 2023
10. Pedregosa, F., et al.: Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011)
11. Mooney, P.: Chest x-ray images (pneumonia), kaggle (2018). <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>, Accessed 15 Mar 2023
12. Tan C., et al.: A survey on deep transfer learning, In: International Conference on Artificial Neural Networks, pp. 270–279 (2018)

13. Russakovsky, O., et al.: ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* **115**(3), 211–252 (2015)
14. Sharma, S., Guleria, K.: A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks. *Procedia Comput. Sci.* **218**, 357–366 (2023)



In Silico Coaxial Antenna Design Applicator Optimization for Microwave Ablation Therapy in Medium Adipose Tissue Density Breast with Ductal Carcinoma *In-Situ*

Eduardo Pinto García¹ (✉) , Mario Francisco Jesús Cepeda Rubio¹, Geshel David Guerrero López², Kristian Segura Felix², José Irving Hernández Jaquez¹, Sergio Israel Velázquez Chávez¹, and Francisco Flores García¹

¹ Instituto Tecnológico de la Laguna, 27000 Torreón, COAH, Mexico
d.epintog@correo.itlalaguna.edu.mx

² Universidad Tecnológica de Torreón, 27400 Torreón, COAH, Mexico

Abstract. Since 2020, breast cancer has been the one with greater incidence across the world. It is important to evaluate the possibility of the application of this minimally invasive procedure to treat the different types of premalignant lesions present in this type of cancer. Since microwave ablation therapy has been proven effective against liver and bone cancer, this research aims to assess the feasibility of an optimization method for the design of an antenna applicator that takes into consideration the effective wavelength of the breast tissue. Therefore, the computational finite element method is used to evaluate the therapy in an *in-silico* environment, which considers the thermodynamic, dielectric properties of the materials, modelling the response of the heat transfer in biological tissues due to microwave and the matching of the antenna. Achieving a $S_{11} = -24.79$ dB, SAR = 25 dB, and maximum temperature of 110 °C for a medium fat tissue density breast model. Inserting a 14 mm of diameter sphere with the characteristics of Ductal Carcinoma In-Situ, resulted in a $S_{11} = -11.47$ dB. Thus, not only maintaining the desired values for the coupling parameters, but also reaching greater ablation zones and temperatures with the lesion.

Keywords: Microwave Ablation · Standing Wave Ratio · Power Reflection

1 Introduction

Microwave ablation (MWA) therapy has been proven a reliable and effective treatment for bone and liver cancer [1, 2]. During MWA therapy, the antenna applicator is guided to the target tissue percutaneously with the guidance of computerized tomography, ultrasound, or magnetic resonance imaging. High-frequency electromagnetic waves are radiated into the tissue in which the antenna is inserted, which causes tissue death due to the coagulation and protein denaturation caused by the heating effect of microwaves to the

polar molecules present in the tissue [3]. Many optimization methods have been used to obtain large and round ablation zones in this type of cancers [4–6], but few have been applied to breast cancer MWA therapy [7]. Since 2020 breast cancer has been the type of cancer with greater incidence and mortality across the world [8], it is important to evaluate the possibility of the application of this minimally invasive procedure to these types of lesions. However, since the dielectric properties in the breast tissue varies with the fat density, it is important to evaluate the feasibility of MWA in different types of fat tissue densities. Hence, this paper aims to assess the possibility of an optimization method that takes into consideration the effective wavelength of breast tissue with medium adipose density tissue.

2 Materials and Methods

2.1 Bioheat Transfer Due to Electromagnetic Waves

MWA is achieved thanks to the interaction of the electromagnetic waves with the polar molecules of the biological tissue [9], this interaction produces movement, therefore friction and heat are generated [10]. This heat is transferred across the tissue and is described by the *Pennes* Bioheat Equation, shown in Eq. (1). It takes into consideration the thermodynamic characteristics of blood in the perfused tissue. Where, ρ_{bl} is the blood density, C_{bl} is the blood-specific heat capacity, ω_{bl} is the blood perfusion rate, T_{bl} is the blood temperature, ρ is the tissue density, C is the tissue-specific heat capacity, k is the tissue thermal conductivity, T is the final temperature.

$$\frac{\rho C \partial T}{\partial t} = \nabla \cdot (k \nabla T) \rho_{bl} C_{bl} \omega_{bl} (T_{bl} - T) + Q_{met} + Q_{ext} \quad (1)$$

The term Q_{met} is the metabolic activity, which is minimal during the therapy, therefore its neglected from the computational analysis. The external heat produced by the antenna applicator is described by the term Q_{ext} , which is proportional to the conductivity of the tissue and the electric field, described as follows:

$$Q_{ext} = \sigma E^2 \quad (2)$$

To evaluate the efficiency of the power delivered to the tissue with the antenna, several parameters need to be considered. Such as the frequency dependent coefficient (S_{11}). Which requires the input power and reflected power [11], hence:

$$S_{11} = 20 \log_{10} \Gamma = 10 \log_{10} \left(\frac{P_r}{P_{in}} \right) (\text{dB}) \quad (3)$$

where the Γ term enables the calculation of the *Standing Wave Ratio* (SWR). This parameter allows us to determine the mismatch of the antenna, thus:

$$\text{SWR} = \frac{1 + |\Gamma|}{1 - |\Gamma|} \quad (4)$$

Γ is useful to calculate the percentage of reflected power to the source of the microwaves, obtained with the following equation:

$$\% P_r = 100 |\Gamma|^2 (\%) \quad (5)$$

The *specific absorption rate* (SAR) represents the amount of power deposited per unit mass, which depends on the tissue conductivity σ , the electric field E , and the density of the tissue ρ :

$$\text{SAR} = \frac{\sigma}{2\rho} |\vec{E}|^2 (\text{W} \cdot \text{Kg}^{-1}), \quad (6)$$

This parameter defines the absorption of microwave energy, which causes temperature to rise, however, it does not determine the temperature distribution in the tissue. At last, the geometric parameters for the antenna applicator design are given by the effective wavelength of the breast tissue, given by:

$$\lambda_{eff} = \frac{c}{f \sqrt{\epsilon_r \mu_r}} \quad (7)$$

where c is the speed of light, f is the operating frequency, ϵ_r is the relative permittivity of the breast tissue, and μ_r is the relative permeability of the medium.

2.2 Model Definition

The physics behind the *Pennes* bioheat equation and electromagnetic waves is described by systems of partial differential equations that can be modeled and solved by numerical methods and computational means. One such numerical method is the *Finite Element Method* (FEM), that aims to solve the physics involved in the given dimensional model (which can be 1D, 2D, 2D-Axi-Simetric or 3D) by dividing the geometry into simple elements. Thus, by applying this discretization solving the system of equations at the nodes of edges of each element. For this study, we use *COMSOL Multiphysics 5.5* software to apply a parametric sweep, frequency domain, and time dependent studies to our breast model with medium fat tissue density. The model solution took 34 min, a 6 core i7-8750H Intel CPU 32 GB of DDR4 RAM Laptop was used for the simulations.

The breast model tissue proposed is a semi-sphere of a 72 mm radius, as seen on Fig. 1. For the study simulations we need to consider the dielectric and thermodynamic characteristics of breast tissue, that's why Table 1 shows these parameters for a breast with a medium adipose tissue density. Specifically, we feed the tissue Cole-Cole parameters to a modified Debye expression to get the relative permittivity and conductivity of a breast with medium fat density [12].

Equation (7) was the foundation for the maximum element size for the model mesh, taking into consideration the effective wavelength of the intermediate fat density tissue, which resulted in 879013 domain elements, 128478 boundary elements, and 7709 edge elements. For electromagnetic wave physics, we chose the external domains of the breast and antenna as the scattering boundary condition. The boundary condition for the bioheat transfer is the thermal insulation in the external boundaries of the geometries.

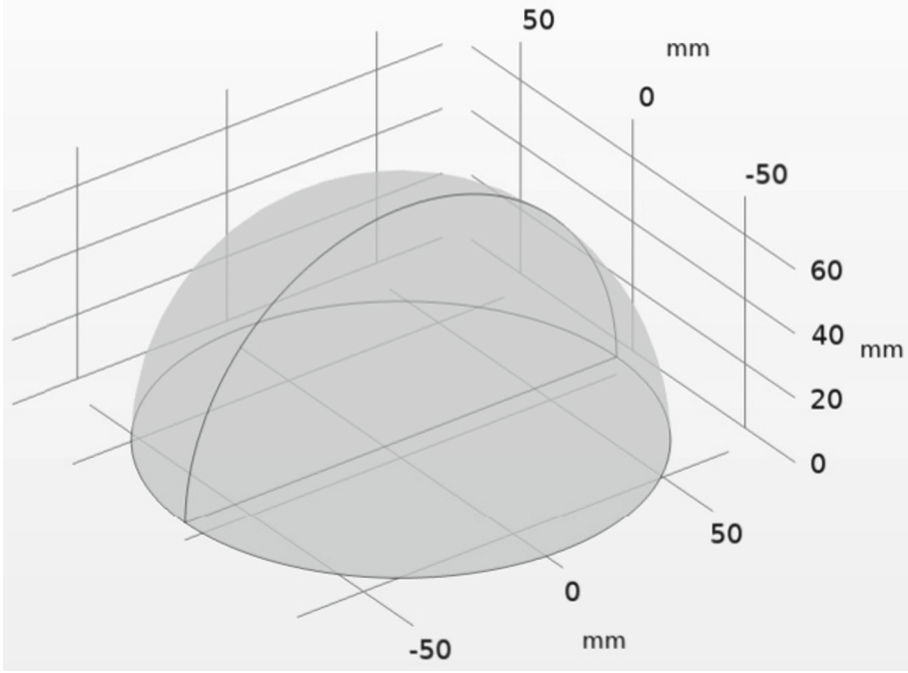


Fig. 1. Breast Model

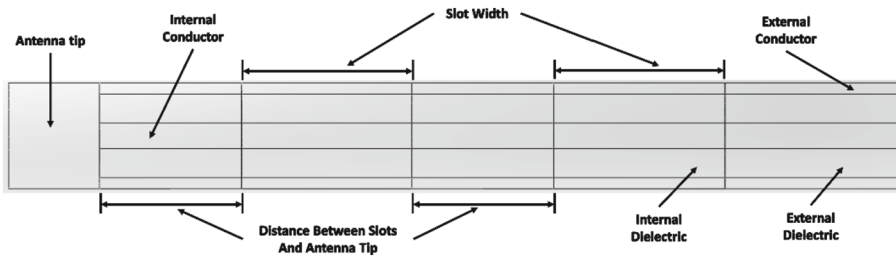
Table 1. Thermodynamic and Dielectric Parameters of the Proposed Breast Tissue.

Parameter	Magnitude	Unit	Reference
Density ρ	932	$\text{kg} \cdot \text{m}^{-3}$	[12]
Thermal Conductivity K	0.171	$\text{W} \cdot \text{m}^{-1} \cdot \text{K}^{-1}$	
Heat Capacity at Constant Pressure C	2,200	$\text{J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$	
Relative Permeability μ_r	1	–	[13]
Relative Permittivity ϵ_r	23.5959	–	
Electrical Conductivity σ_s	0.5	$\text{S} \cdot \text{m}^{-1}$	
Blood Density ρ_{bl}	1040	$\text{kg} \cdot \text{m}^{-3}$	[14]
Blood Specific Heat C_{bl}	3,639	$\text{J} \cdot \text{Kg}^{-1} \cdot \text{K}^{-1}$	
Blood Perfusion Rate ω_{bl}	0.0036	s^{-1}	

For the antenna applicator we propose an original 3D model of a double slot coaxial antenna that is comprised of an outer and inner conductors made of copper and an internal dielectric made of PTFE (As seen in Fig. 2), where the UT-47 coaxial cable standard is used. The geometric and dielectric parameters are shown in Table 2.

Table 2. UT-047 antenna parameters.

Parameter	Magnitude	Unit	Reference
Internal Conductor Diameter	0.287	<i>mm</i>	–
External Conductor Diameter	1.19		
Internal Dielectric Diameter	0.94		
Relative Permittivity of Copper	1	–	[13]
Relative Permittivity of PTFE	2.03		
Electrical Conductivity of Copper	5.998×10^7	$S \cdot m^{-1}$	
Electrical Conductivity of PTFE	5.1×10^{-17}		

**Fig. 2.** UT-047 coaxial double slot antenna model.

2.3 Effective Wavelength (λ_{eff}) Driven Optimization

The proposed optimization method is based on the effective wavelength of the breast tissue, as seen in Eq. (7), which relies on the relative permittivity of the medium density of adipose tissue (MD) ϵ_r , the operating frequency $f = 2.45$ GHz, with an arbitrary power of 8 W, and the speed of light $c = 3 \times 10^8$ m/s. Given the relative permittivity of the MD breast, the calculated $\lambda_{eff} \approx 25$ mm, therefore, we propose rational multiples of λ_{eff} , as seen in Table 3, for the Slot Width (SW) and Distance Between Slots (DBS) parameters of the antenna.

Table 3. Slot Width and Distance Between Slots Parameter definition based on λ_{eff}

Parameter	λ_{eff} rational multiples			
Slot Width	$\frac{\lambda_{eff}}{4}$	$\frac{3 \cdot \lambda_{eff}}{16}$	$\frac{\lambda_{eff}}{8}$	$\frac{\lambda_{eff}}{16}$
	6.2500 mm	4.6875 mm	3.1250 mm	1.5625 mm
Distance Between Slots	$\frac{\lambda_{eff}}{10}$	$\frac{3 \cdot \lambda_{eff}}{40}$	$\frac{\lambda_{eff}}{20}$	$\frac{\lambda_{eff}}{40}$
	2.50 mm	1.875 mm	1.250 mm	0.625 mm

3 Results

The optimization of the SW and the DBS was done by applying a parametric sweep through all the rational multiples of the calculated effective wavelength given in Table 3. The goal is to find the SW and DBS that achieve the desired values of $S_{11} \leq -10 \text{ dB}$, $\%P_r \leq 10\%$, and $SWR \leq 2$ [15, 16]. Therefore, Table 4 shows the combination of values for the SW and the DBS, where with a $SW = 1.875 \text{ mm}$ and a $DBS = 1.5652 \text{ mm}$ the better parameters are obtained, achieving values for $S_{11} = -24.79 \text{ dB}$, $\%P_r = 0.36\%$, and $SWR = 1.1286$.

Table 4. Combination of SW and DBS values with the desired values for s_{11} , SWR, and $\% P_r$.

SW (mm)	DBS (mm)	SWR (Eq. (4))	S_{11} dB (Eq. (3))	$\%P_r$ (Eq. (5))
0.625	3.125	1.3986	-15.589	2.76%
1.25	1.5652	1.1799	-21.666	0.68%
1.25	3.125	1.7807	-11.033	7.88%
1.875	1.5652	1.1286	-24.379	0.36%
2.5	1.5652	1.439	-14.895	3.24%

With the parameters that yielded the best results, we evaluated thermal performance by modelling a 520 s MWA therapy in the proposed MD breast tissue, as seen in Fig. 3. Notice that the greater concentration of power given by the normalized SAR distribution, in Fig. 3a, corresponds with the heat distribution seen in Fig. 3b.

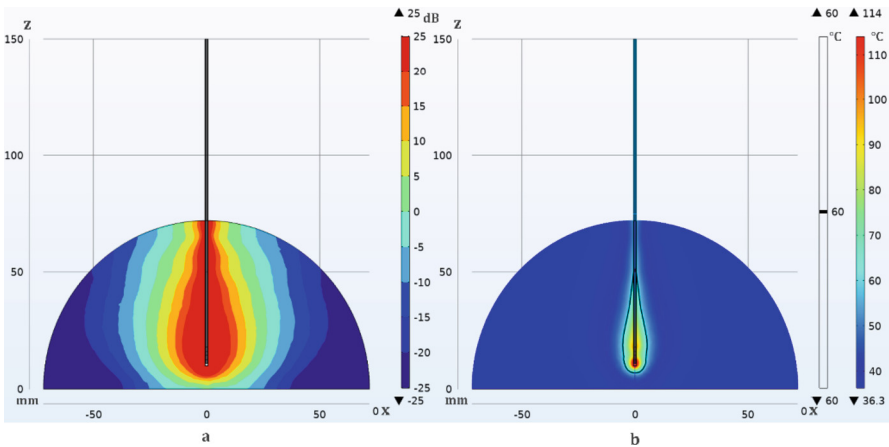


Fig. 3. Normalized SAR and temperature distribution in MD breast tissue.

With these results, the MWA therapy was also applied to a sphere that stands as a *Ductal Carcinoma In-Situ* (DCIS) mass of 14 mm of diameter, to know the thermal characteristics and the coupling of the antenna with a malignant tissue. Where the thermal and dielectric parameters of the DCIS lesion are taken from [17], thus, resulting in a $S_{11} = -11.47\text{dB}$, $SWR = 1.7198$, and a $\%P_r = 7.004\%$. Where it is clearly seen in Fig. 4, that the energy is more concentrated within the DCIS boundaries, thus achieving more heat in the lesion.

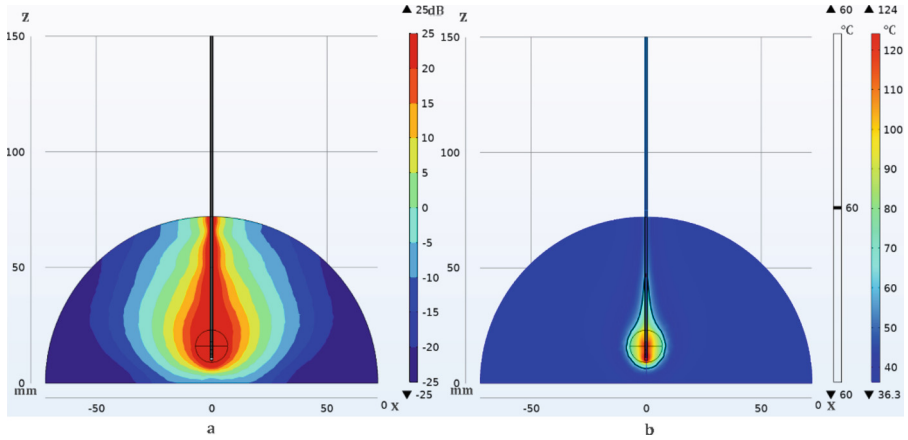


Fig. 4. Normalized SAR and temperature distribution in MD breast tissue with DCIS.

4 Discussion

This computational optimization has achieved the main goal of an antenna design capable of having the desired coupling parameters for matching the applicator with a breast tissue with medium fat tissue density, making this study relevant, as previous works [17–19] only dealt with a breast with high adipose tissue density, which does not represent the entirety of patients who could experience these conditions. The tissue irreversible damage starts at $45\text{ }^{\circ}\text{C}$, but it is best to achieve temperatures between $50\text{ }^{\circ}\text{C}$ and $100\text{ }^{\circ}\text{C}$ [20]. Therefore, the desired ablation temperatures are reached in the $520\text{ s} \approx 8.66\text{ min}$ therapy time, shown in Fig. 5, which also shows that the antenna worked within its operational $150\text{ }^{\circ}\text{C}$ temperature.

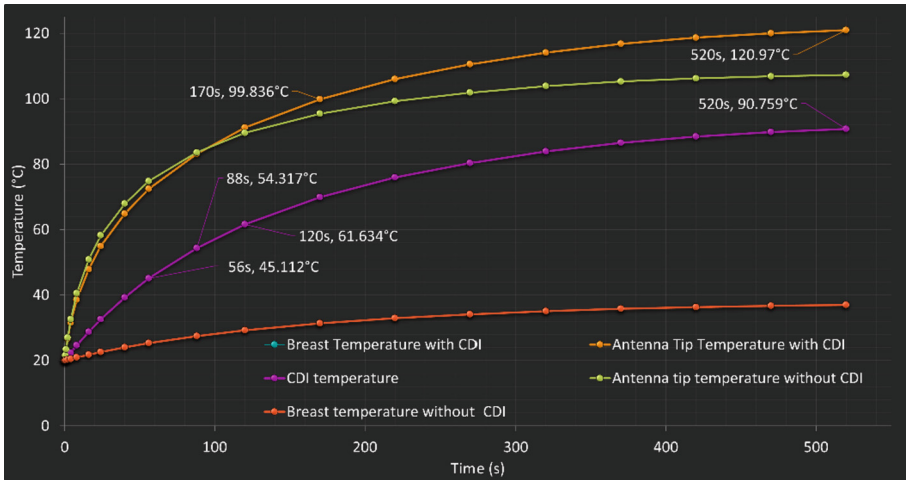


Fig. 5. Average temperatures reached over a 520 s MWA therapy time within the breast tissue models, antenna tip and DCIS.

5 Conclusions

The presented computational optimization successfully achieved the desired coupling parameters, $S_{11} = -11.47\text{dB}$ which represents a $\%P_r = 7.004\%$, for effective microwave ablation therapy in breast tissue. The proposed antenna design demonstrated efficient energy delivery and temperature control, resulting in targeted and controlled thermal treatment, because the ablation temperature of 55°C is reached on the tumor domain before 100 s and can be sustained for at least 20 s before reaching temperatures that might compromise the proper functioning of the antenna, as seen in Fig. 5. Since this can be obtained with 8 watts of power, it is not necessary to simulate other power conditions. Moreover, since blood perfusion is not considered in this model, it is expected that the temperature distribution due to this factor will be more efficient, allowing for a longer duration of the ablation treatment. Comparing this study with those previously conducted, it can be observed that the breast's adipose tissue density is crucial for optimization and proper treatment planning. This approach holds promise for potential application in breast cancer treatment, warranting further exploration and validation through experimental studies.









References

1. Zheng, K., et al.: Clinical guideline for microwave ablation of bone tumors in extremities. *Orthop. Surg.* **12**, 1036–1044 (2020). <https://doi.org/10.1111/os.12749>
2. Radjenović, B., Sabo, M., Šoltes, L., Prnova, M., Čičak, P., Radmilović-Radjenović, M.: On efficacy of microwave ablation in the thermal treatment of an early-stage hepatocellular carcinoma. *Cancers* **13**, 5784 (2021). <https://doi.org/10.3390/cancers13225784>
3. Brace, C.L.: Microwave tissue ablation: biophysics, technology, and applications. *Crit. Rev. Biomed. Eng.* **38**, 65–78 (2010). <https://doi.org/10.1615/CritRevBiomedEng.v38.i1.60>

4. Etoz, S., Brace, C.L.: Analysis of microwave ablation antenna optimization techniques. *Int. J. RF Microw. Comput. Aided Eng.* **28**, e21224 (2018). <https://doi.org/10.1002/mmce.21224>
5. Gas, P., Szymanik, B.: Shape optimization of the multi-slot coaxial antenna for local hepatic heating during microwave ablation. In: 2018 International Interdisciplinary PhD Workshop (IIPhDW), pp. 319–322. IEEE, Swinoujście (2018)
6. Xu, Y., Moser, M.A.J., Zhang, E., Zhang, W., Zhang, B.: Large and round ablation zones with microwave ablation: a preliminary study of an optimal aperiodic tri-slot coaxial antenna with the π -matching network section. *Int. J. Therm. Sci.* **140**, 539–548 (2019). <https://doi.org/10.1016/j.ijthermalsci.2019.03.022>
7. Zhao, Z., Wu, F.: Minimally-invasive thermal ablation of early-stage breast cancer: a systemic review. *Eur. J. Surg. Oncol. (EJSO)* **36**, 1149–1155 (2010). <https://doi.org/10.1016/j.ejso.2010.09.012>
8. Sung, H., et al.: Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CAA Cancer J. Clin.* **71**, 209–249 (2021). <https://doi.org/10.3322/caac.21660>
9. Irving, J., Mario, C., Francisco, V., Geshel, G.: Microwave ablation: state-of-the-art review. *OTT* 1627 (2015). <https://doi.org/10.2147/OTT.S81734>
10. Krenacs, T., et al.: Modulated electro-hyperthermia-induced tumor damage mechanisms revealed in cancer models. *IJMS.* **21**, 6270 (2020). <https://doi.org/10.3390/ijms21176270>
11. Gas, P.: Optimization of multi-slot coaxial antennas for microwave thermotherapy based on the S11-parameter analysis. *Biocybern. Biomed. Eng.* **37**, 78–93 (2017). <https://doi.org/10.1016/j.bbe.2016.10.001>
12. Martellosio, A., et al.: Dielectric properties characterization from 0.5 to 50 GHz of breast cancer tissues. *IEEE Trans. Microwave Theory Tech.* **65**, 998–1011 (2017). <https://doi.org/10.1109/TMTT.2016.2631162>
13. Said Camilleri, J., et al.: Review of thermal and physiological properties of human breast tissue. *Sensors* **22**, 3894 (2022). <https://doi.org/10.3390/s22103894>
14. Segura Félix, K., et al.: Computational FEM model and phantom validation of microwave ablation for segmental microcalcifications in breasts using a coaxial double-slot antenna. *Biomed. Res. Int.* **2021**, 1–10 (2021). <https://doi.org/10.1155/2021/8858822>
15. Balanis, C.A. (ed.): *Modern Antenna Handbook*. Wiley, Hoboken, NJ (2008)
16. Huang, Y., Boyle, K.: *Antennas: From Theory to Practice*. Wiley (2008)
17. Guerrero López, G.D., et al.: Computational FEM model, *phantom* and *ex vivo* swine breast validation of an optimized double-slot microcoaxial antenna designed for minimally invasive breast tumor ablation: theoretical and experimental comparison of temperature, size of lesion, and SWR, preliminary data. *Comput. Math. Methods Med.* **2017**, 1–11 (2017). <https://doi.org/10.1155/2017/1562869>
18. Jesús Cepeda Rubio, M.F.: Coaxial slot antenna design for microwave hyperthermia using finite-difference time-domain and finite element method. *TONMJ* **3**, 2–9 (2011). <https://doi.org/10.2174/1875933501103010002>
19. Ortega-Palacios, R., et al.: Feasibility of using a novel 2.45 GHz double short distance slot coaxial antenna for minimally invasive cancer breast microwave ablation therapy: computational model, phantom, and *in vivo* swine experimentation. *J. Healthc. Eng.* **2018**, 1–10 (2018). <https://doi.org/10.1155/2018/5806753>
20. Rhim, H., et al.: Essential techniques for successful radio-frequency thermal ablation of malignant hepatic tumors. *Radiographics* **21**, S17–S35 (2001). https://doi.org/10.1148/radiographics.21.suppl_1.g01oc11s17



Convolutional Neural Network for Classifying Thoracic Diseases in Adult Patients

Luis Alberto Gómez-Celaya , Jorge Alonso Débora-Ramírez ,
Rolando Corral-Barrón , Poleth Rodríguez-Domínguez ,
Aldair Verdugo-Valerio , Carlos Eduardo Cañedo-Figueroa ,
Abimael Guzmán Pando , and Hugo Varela-Rodríguez 

Universidad Autónoma de Chihuahua, Circuito Universitario 31109, Campus UACH II,
31125 Chihuahua, Mexico
ccanedo@uach.mx

Abstract. This paper proposes an automatic classifier, based on a convolutional neural network, capable of identifying different pathologies and diseases seen in anterior chest radiographs. The dataset was obtained from the National Institutes of Health (NIH) and Kaggle to be used as training for the classifier. For example, healthy, cardiomegaly, infiltration, effusion, mass, pneumothorax, emphysema, fibrosis, oedema, nodules and others. The average results for the evaluated metrics recall, accuracy and F1 score were 92%, 82% and 87% respectively, showing the flexibility of the proposed model to handle different tasks. This paper proposes an automatic classifier based on a convolutional neural network capable of identifying different pathologies and diseases shown in anterior chest X-ray studies. The data set was obtained from the National Institutes of Health (NIH) and Kaggle to be used as training for the classifier. E.g., Healthy, Cardiomegaly, Infiltration, Effusion, Mass, Pnemo-thorax, Emphysema, Fibrosis Edema, Nodule and others.

Keywords: Deep Learning · Thoracic Diseases · Automatic Classifier

1 Introduction

Imaging the chest cavity is crucial for diagnosing heart and lung disease. These interconnected systems can affect each other. Narrowed coronary arteries due to ischemic heart disease can cause breathing problems. Similarly, lung diseases such as COPD (chronic obstructive pulmonary disease) can increase the likelihood of heart failure [1].

Cardiovascular disease is a major cause of death and disability worldwide. The most common types of cardiovascular disease include ischemic heart disease, cerebrovascular disease and peripheral vascular disease [2].

On the other hand, common lung diseases include asthma, chronic obstructive pulmonary disease (COPD), pulmonary fibrosis and pneumonia.

Machine learning, a branch of Artificial Intelligence (AI), allows machines to learn and improve from data without explicit programming [4]. It uses algorithms to recognize patterns and make decisions from inputs, including supervised and unsupervised learning.

Deep learning, a subset of machine learning, uses deep neural networks to extract features and learn through complex layers to capture intricate data features [4]. Some articles use convolutional neural networks (CNN) for image classification.

CNN is a mathematical model that use as input data each pixel of an image, this information is processed for extracting patterns and weight adjustments, and developing recognition systems. A limitation is that it is necessary to specify the resolution and size on the input image. Therefore, many times it is necessary to develop a conditioned algorithm for normalizing and match the systems requirements [10]. In medical field, the AI systems can be used for detection and classification of diseases by using images studies such as magnetic resonance, ultrasound, tomographic data, X-ray and other similar.

In this way, we developed a convolutional neural network algorithm, which uses x-ray studies, but considering that a single x-ray may reveal one or more diseases, the neural model was trained with disease categories.

Similar works in literature includes; A convolutional neural network algorithm that detects lesions in chest x-rays, mentioned in [10], and a Deep Convolutional Neural Network for Computer-Aided Detection and Classification of Anomalies in Frontal Chest Radiographs [11]. However, these works can only detect between three to four diseases.

Here, we are developing an algorithm to classify images using two different data sets solving 15 disease tasks through X-ray images. For this purpose, we took into a count doctor's analyses of anatomical or physiological abnormalities, images, and the patient's medical history. The neural network is trained to find relations, and to identify conditions in the study images. These conditions are extracted from true positive databases. Our aim is not to replace the doctor diagnosis, but to make his job easier.

2 Methodology

2.1 Database

For this work, we used two databases. The first database was generated from information collected by the NIH (National Institutes of Health), with images of 1024×1024 pixels and uniform size. These images came from the NIH Clinical Centre itself, representing 60% of all frontal chest studies.

For the second database, chest studies were collected from institutions such as the Cohen Medical Center and Actmed, among others. Although image resolution and size varied, we corrected this variation by adjusting the dimensions of the images to ensure that both databases had the same characteristics. We collected a total of 39,296 chest radiology images and studies from patients with cardiomegaly, infiltration, effusion, mass, pneumothorax, emphysema, fibrosis, oedema, nodule, atelectasis, consolidation, pneumonia, COVID-19, pleural thickening and healthy. Each condition in the patients can be described as follows:

- Cardiomegaly: An enlarged heart.
- Infiltration: A hazy opacity in the lungs.
- Effusion: A collection of fluid in the space between the lungs and the chest wall.
- Mass: A solid tumor or other abnormality in the lung.

- Pneumothorax: A collapsed lung.
- Emphysema: A condition in which the air sacs in the lungs become damaged and enlarged.
- Fibrosis: The formation of scar tissue in the lungs.
- Oedema: An accumulation of fluid in the tissues.
- Nodule: A small, round mass in the lungs.
- Atelectasis: The collapse of a lung.
- Consolidation: The filling of a lung with fluid or other material.
- Pneumonia: An infection of the lungs.
- COVID-19: A respiratory illness caused by the SARS-CoV-2 virus.
- Pleural thickening: Thickening of the lining of the lungs.
- Healthy patients: Patients who have none of the above.

The diseases described above are the most commonly diagnosed medical problems in the chest area by x-ray, so it is important to have an idea of the condition of each disease in order to properly build the database and achieve the goal of the neural network.

The training process for our network involved randomly using 80% of the images for training and the remaining 20% for testing. The distribution of the data is shown in Table 1 [5].

2.2 Convolutional Neural Network

Our first database, as mentioned earlier, shared uniformity in resolution (1024×1024), size, and font. This is contradicted by the second images database. Therefore, we performed image size adjustments considering the computational capability available, setting the size of 500×500 pixels. Images were grouped into batches of 64 units for training due to the maximum processing capacity of the computers.

Dense convolutional networks (DenseNets) offer a beneficial strategy of referencing prior feature maps from the network's inception. Each feature map layer is combined with the input of subsequent layers within a dense block. This facilitates subsequent layers to directly utilize preceding layer features, promoting feature reuse [3].

The model was trained and structured using Adam optimization and the Binary Cross Entropy loss function. The DenseNET architecture was adopted for multi-label classification [3]. For this work, we used DenseNet201 architecture, as shown in Fig. 1, where the input is an image in RGB format.

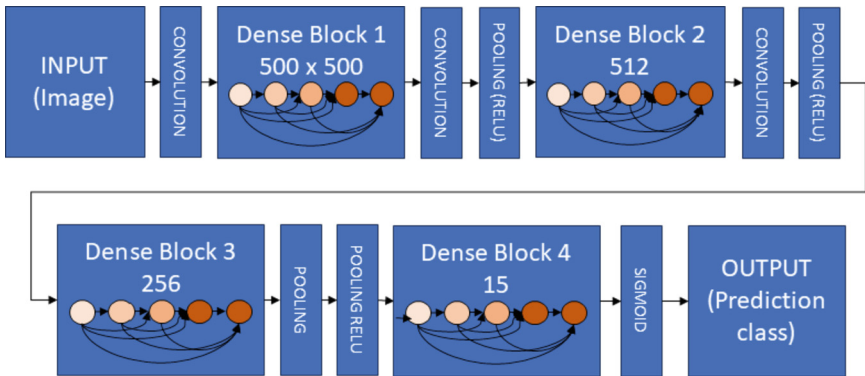
We first created the base for the model using the DenseNet201 architecture with ImageNet weights and an input of $500 \times 500 \times 3$ (L,W,Ch).

We then froze these parameters, created a sequential model and added the base model as the first layer. We then added our own 512 and 256 layers that would change their weights according to the image feed. These layers have a ReLU activation and a kernel of L2 with a value of 0.1. Finally, our last layer has 15 neurons and a sigmoid activation, which is recommended for multi-label work.

To implement this CNN, we used Eq. (1), where $CNN()$ represents a function that propagates an x_j image for the test to obtain a z_i output vector, being i the i -th output class category. The winning class is chosen according to the higher value of z_i . See Table 2.

Table 1. Categories included in the study for training and test Convolutional Neural Network.

Categories	Data per class	Data for training	Data for test
Healthy	5353	4282	1071
Cardiomegaly	2776	2221	555
Infiltration	2894	2315	579
Effusion	2317	1854	463
Mass	2782	2226	556
Pneumothorax	2302	1842	460
Emphysema	2516	2013	503
Fibrosis	1686	1349	337
Edema	2303	1842	461
Nodule	2331	1865	466
Atelectasis	2559	2047	512
Consolidation	2667	2134	533
Pneumonia	1431	1145	286
Covid-19	2994	2395	599
Pleural Thickening	2385	1908	477

**Fig. 1.** Diagram block of DenseNET.

$$z_i = CNN(x_j)$$

$$Class = \operatorname{argmax}(z_i) \quad (1)$$

Table 2. Outputs from the CNN

Class	Value of i-th output
Healthy	1
Cardiomegaly	2
Infiltration	3
Effusion	4
Mass	5
Pneumothorax	6
Emphysema	7
Fibrosis	8
Edema	9
Nodule	10
Atelectasis	11
Consolidation	12
Pneumonia	13
Covid-19	14
Pleural Thickening	15

3 Results

To evaluate the algorithm, we used the data presented in Table 1, column 4, together with Eqs. (2), (3) and (4) to calculate precision, recall and accuracy, respectively. In these equations, TP denotes true positives, FP denotes false positives and FN denotes false negatives [7].

It is essential to know some concepts related to neural networks and their characteristics to have a better perception of the information, these are:

Accuracy (%): Represents the allowable variation between measured and actual values when a known value is used as a reference. Suitable for balanced classes of target variables, as in project datasets.

Accuracy: Measures correct predictions across all types of predictions. Indicates the proportion of positive diagnoses that are truly positive.

Recall: Informs about the performance of the classifier with respect to false negatives and complements the insight of precision with respect to false positives.

F1 Score: Combines Recall and Accuracy into a single score, often calculated as their arithmetic mean.

Average: is the average of all metrics obtained in this work.

The aim is to accurately categorize them as either ‘sick’ or ‘not sick’ and to identify the specific disease within these categories. In addition, the evaluation extends to the accurate diagnosis of images with multiple diseases, confirming whether they have been correctly classified or treated as false positives.

$$precision = \frac{TP}{(TP + FP)} \quad (2)$$

$$recall = \frac{TP}{(FN + TP)} \quad (3)$$

$$accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (4)$$

Using the results of each equation, we calculated the F1 score for a given class. We agreed on Eq. (5) [7].

$$F1 = 2 \left(\frac{precision \cdot recall}{precision + recall} \right) \quad (5)$$

The model is a multi-label convolutional neural network with the ability to classify the diseases selected for analysis. It is necessary to mention that the metrics of the different diseases are different, as the same model was used for all the categories mentioned in Table 2. These results are shown in Table 3.

These are promising results, but the data for the test were from the same sources (obviously not used for training, but still sharing conditions from the other images in training, so we would like to try with different x-ray images from different hospitals to confirm these results).

The metrics were compared with those obtained in some articles with similar diseases. The progress observed in this field is mainly concentrated on diseases that have increased in recent years, such as the different pathologies analysed in this article. Comparative Table 4 shows the different results obtained in previously published articles.

Despite our best efforts, as any other system, there are some limitations. Mainly related to the absence of information and the high costs access associated with the articles outlined in our theoretical framework. Consequently, full data collection for Table 4 was elaborated with these limitations. In particular, the article ‘Lesion-aware convolutional neural network for chest radiograph classification’ [10] remains excluded as it lacks explicit model metrics. Instead, it only provides an average 95% confidence index for 4 conditions - atelectasis, mass and nodule - while remaining silent on the remaining 11 conditions.

Table 4 compares the metrics within three categories: Health, Pneumonia and COVID-19, as these are the only metrics reported in the articles. Notably, our developed network identifies 15 indicators, which is a wider range compared to other studies that typically focus on two or three indicators. Previous studies used different database splits, including 80–20 [1], 70–30 [2, 3], and none shared the same database.

Unlike previous research [7–9], the implementation of our model is distinguished by our base architecture. Previous models used ResNet [8], LeNet [7] and VGGNet-19 [7], while we used DenseNet. Furthermore, previous studies considered a maximum of three diseases, whereas our model effectively encompasses and recognizes 15 diseases without sacrificing accuracy.

Table 3. Analysis categories with their respective percentages of precision, recall, and F1 Score.

Name	Accuracy (%)	Recall (%)	F1-Score (%)	Precision (%)
Healthy	96	80	87	96
Cardiomegaly	91	80	85	91
Infiltration	98	90	94	98
Effusion	88	81	84	88
Mass	88	80	84	88
Pneumothorax	93	82	87	93
Emphysema	94	83	88	94
Fibrosis	92	77	84	92
Edema	89	83	86	89
Nodule	93	79	85	93
Atelectasis	89	78	83	89
Consolidation	91	85	88	91
Pneumonia	90	76	82	90
Covid-19	93	88	90	93
Pleural Thickening	88	84	86	88
Average	92	82	87	92

Table 4. Comparison results respect state-of-the-art algorithms.

	<i>Accuracy (%)</i>					<i>Recall (%)</i>				<i>F1 score (%)</i>			
	Ours	[7]	[8]	[9]	[11]	Ours	[8]	[7]	[9]	Ours	[8]	[7]	[9]
Healthy	96	83	91	-	96	80	89.2	99.6	-	87	90.6	91.0	-
Pneumonia	90	-	92.1	74	-	76	98	-	74	82	96.2	-	82
Covid-19	93	99	94	-	-	88	98	80.8	-	90	96.2	89.2	-
Pleural Thickening	88	-	-	-	96	78	-	-	-	83	-	-	-
Edema	89	-	-	-	86	83	-	-	-	86	-	-	-
Consolidation	91	-	-	-	85	85	-	-	-	88	-	-	-
Cardiomegaly	91	-	-	-	87	80	-	-	-	85	-	-	-
Pneumothorax	93	-	-	-	86	82	-	-	-	87	-	-	-

4 Conclusion

The DenseNet-based model developed in this study provides accurate results for a wide range of lung diseases. The results obtained by our model encourage the developing of deep learning efforts into different areas, introducing additional data such as blood counts as inputs to the neural network, thus improving the detection capabilities.

In particular, the model developed provides a viable solution in a resource-constrained environment, providing accurate detections that support accurate disease diagnoses. This comprehensive strategy allows for the creation of an internal database, eliminating the need for additional infrastructure expenditures and streamlining future network expansion.

For future efforts, an alternative approach could be to build disease-specific models and ensemble them. This ensemble approach would include 15 neural networks (1 for each disease), improving the accuracy of individual models and resulting in superior overall system performance. The use of category-specific models and ensemble techniques has effectively increased disease classification accuracy by capitalizing on the strengths of individual models.

References

1. Veloza, L., Jiménez, C., Quiñones, D., Polanía, F., Pachón-Valero, L.C., Rodríguez-Triviño, C.Y.: Variabilidad de la frecuencia cardiaca como factor predictor de las enfermedades cardiovasculares. *Rev. Colomb. Cardiol.* **26**(4), 205–210 (2019)
2. Hall, J.E.: *Guyton and Hall Textbook of Medical Physiology*, 14th ed. Elsevier (2021)
3. Artola Moreno, Á.: *Clasificación de imágenes usando redes neuronales convolucionales en Python*, Trabajo Fin de Grado Inédito. Universidad de Sevilla, Sevilla (2019)
4. Shinde, P.P., Shah, S.: A review of machine learning and deep learning applications. In: 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, pp. 1–6 (2018). <https://doi.org/10.1109/ICCUBEA.2018.8697857>
5. Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., Summers, R.M.: ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3462–3471 (2017)
6. Lin, C.-H., et al.: Posteroanterior chest X-ray image classification with a multilayer 1D convolutional neural network-based classifier for cardiomegaly level screening. *Electronics* **11**(9), 1364. <https://doi.org/10.3390/electronics11091364>
7. Fuentes, K., Vera, S.: *DESARROLLO DE UN SISTEMA WEB UTILIZANDO PROCESAMIENTO DE IMÁGENES DE RAYOS X DE TÓRAX PARA LA DETECCIÓN DE COVID-19 MEDIANTE DEEP LEARNING* [Trabajo de integración curricular]. Universidad Estatal de Milagro (2021)
8. Gómez, M., et al.: *Detección de COVID-19 en radiografías de tórax mediante aprendizaje profundo*. Tecnología e Innovación en Educación Superior (2023)
9. Naar Pérez, A., Barreto Martínez, F.: *Modelo de red neuronal convolucional para el diagnóstico de neumonía en imágenes radiológicas*. Universidad del SINÚ (2019, junio). <http://repositorio.unisinucartagena.edu.co:8080/jspui/bitstream/123456789/50/1/MODELO~1.PDF>

10. Li, F., et al.: Lesion-aware convolutional neural network for chest radiograph classification. *Clin. Radiol.* **76**(2), 155.e1–155.e14 (2021). <https://doi.org/10.1016/j.crad.2020.08.027>
11. Cicero, M., et al.: Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs. *Invest. Radiol.* **52**(5), 281–287 (2017). <https://doi.org/10.1097/RLI.000000000000034>



Computation of Human-Sperm Local Flagellar Instantaneous Velocity

Dan Sidney Díaz-Guerrero¹, Fernando Montoya¹, Haydee Olínca Hernández², Paul Hernández-Herrera³, Alberto Darszon⁴, and Gabriel Corkidi¹(✉)

¹ Laboratorio de Imágenes y Visión por Computadora, Departamento de Ingeniería Celular y Biotatálisis, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

gabriel.corkidi@ibt.unam.mx

² Posgrado en Ciencia e Ingeniería de la Computación, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

³ Laboratorio Nacional de Microscopía Avanzada, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

⁴ Departamento de Genética del Desarrollo y Fisiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

Abstract. The instantaneous velocity generated by human-sperm flagellum is key for the fundamental description and understanding of flagellar movement and sperm motility. Furthermore, since the flagellar movement is intrinsically three-dimensional, only with 3D dynamic data of the flagellar trajectory it is feasible to compute the instantaneous velocity of flagellar point elements. It was only recently that real flagellar 3D+t motility data were available, which is indispensable for the calculation of instantaneous velocity. In this work we present the first results of flagellar instantaneous velocity obtained from 3D+t data of human sperm by means of a Fourier fitting method. One descriptor and its distribution over the sperm sample are presented to compare the experimental conditions (induced to capacitate, and not induced to capacitate as control) related to the fertilization process. While the descriptor distributions are evidently distinct across both experimental conditions, a statistical analysis confirmed this difference supporting the visual perception and importance of the finding.

Keywords: Instantaneous velocity · Fourier transform · Fourier series · Vector resultant · Human sperm · 3D microscopy

1 Introduction

The study of sperm motility has been mainly based on the bidimensional (2D) head trajectory over time [2, 14–19, 21, 23, 26]. This approach has made possible important advances in basic research (see [9] and references therein) and in the health sciences [20]. Nonetheless, this framework is clearly incomplete as sperm

swim in three dimensions (3D) in certain regions of the female reproductive tract. A comprehensive understanding of the sperm dynamics needs to incorporate the third spatial dimension and the flagellum. This approach was technologically impaired until about 20 years ago [7]. Recently there have been numerous and important developments (see [5, 6, 8, 10, 11, 25]) that include the flagellar beating in the sperm motility analysis of sperm flagellar beating [4, 6, 13].

More recently [13], an experimental setup and a methodological framework were developed to obtain dynamical 3D human whole-sperm data (3D+t). This opened a myriad of possible research lines regarding flagellar motility. However, this significant achievement came with big challenges in the form of enormous quantities of images and numerical data.

In this work, the 3D + t data set consists of 60744 data files (from 147 sperm), each containing the x , y , z coordinates of the segmented flagellum [12]. Each cell belongs to one of two conditions: not induced to capacitate (NoCap, 60 cells) and induced to capacitate (Cap, 87 cells). Using these data, a Fourier series was constructed and used to compute the instantaneous velocity for the mid-flagellar point for each one of the 147 sperm cells in the dataset. The Fourier series of the component functions allowed us to compute instantaneous velocities, and opening other scientific venues.

2 Methodology

The x , y , z coordinates of a sperm at a given time will be called a flagellar curve, thus the entire trajectory of a given sperm will consist of T flagellar curves. Due to the experimental setup there are 90 flagellar curves per second [13]. For instance, if there are 270 flagellar curves for a given sperm, then the time recorded is three seconds.

To compute the instantaneous velocity of each point of the flagellum, first let's use the position vector notation

$$\{\mathbf{r}_{i,t}\}_{i=0}^{n-1} = \{\mathbf{r}_{0,t}, \mathbf{r}_{1,t}, \dots, \mathbf{r}_{n-1,t}\} \quad (1)$$

where $\mathbf{r}_{i,t} = (x_i(t), y_i(t), z_i(t))$ is the i -th flagellar point at time t . The flagellar point with index $i = 0$ represents the head-flagellum junction and the last flagellar point index $i = n - 1$ represents the tail of the flagellum. For instance, take a given point of the flagellum, let's say the ninth point at time five, then it is denoted by $\mathbf{r}_{9,5} = (x_9(5), y_9(5), z_9(5))$. So, the instantaneous velocity at the point $\mathbf{r}_{i,t}$ is its time derivative, this is denoted by $\dot{\mathbf{r}}_{i,t}$. Although there are several options to compute discrete derivatives, in general they all have numerical instabilities. Considering this, the ideal option is to have the mathematical expression of the function describing the experimental points over time, a difficult solution to achieve. However, there are feasible options: series expansions of mathematical functions. Therefore, it is necessary to obtain an accurate curve fitting of $\mathbf{r}_{i,t}$ over time. Let's call this curve a time-point curve, that is the set $\{\mathbf{r}_{i,t}\}$ for $t \in \{1, 2, \dots, T\}$ (abbreviated as $\{\mathbf{r}_{i,t}\}_{t=1}^T$) at a fixed flagellar position i over the flagellum, see Fig. 1.

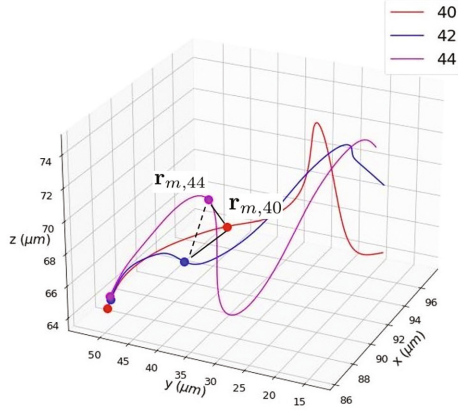


Fig. 1. Flagellar curves, corresponding to times 0.44, 0.47, and 0.49 s (or indices 40, 42, 44). The points in the lower left corner indicate the head of the flagellum, meanwhile the points in the middle are precisely their middle points. Therefore, fixing the flagellar point to value m and considering all the registered times one gets the time-point curve indicated by joining the middle points with lines.

The Taylor series are frequently used to expand a mathematical function. In the Taylor series each term involves a derivative of the objective function, so the mathematical expression of the function must be known *a priori*. Nevertheless, there is another well known series more suited to the case of time-point curves that is the Fourier series. One of the advantages of these series is that they do not use the derivative of the objective function. The Fourier series require the Fourier transform and their associated frequencies, as shown next in Eq. 2.

$$g(t) = \frac{1}{n} \sum_{k=0}^{n-1} \phi(k) \exp(2\pi j \varepsilon(k)t) \quad (2)$$

where j is the imaginary unit ($j^2 = -1$), $\exp(t)$ is the exponential function, and $\phi(k)$ is the k -th point of the Fourier transform, and $\varepsilon(k)$ is its frequency. Given the Fourier series, its time derivative is straight-forward

$$\frac{d}{dt}g(t) = \frac{2\pi j}{n} \sum_{k=0}^{n-1} \phi(k) \varepsilon(k) \exp(2\pi j \varepsilon(k)t). \quad (3)$$

Hence, by numerically computing the Fourier transform and its frequencies a time-point curve is modeled by its Fourier series of the form (2) with the corresponding Fourier transform and frequencies, computed numerically with *python-numpy-fft* package.

3 Results

The Fourier-models of the function components $(x_i(t), y_i(t), z_i(t))$ over time for the time-point curve $\{\mathbf{r}_{i,t}\}_{t=1}^T$ were obtained and used to compute the instantaneous velocity. Thus, $\mathbf{r}_{i,t} = (g_x(t), g_y(t), g_z(t))$ for each $t \in \{1, 2, \dots, T\}$. Here $g_x(t), g_y(t), g_z(t)$ are obtained by applying Eq. 2 to $x_i(t), y_i(t), z_i(t)$, respectively. In Fig. 2 a comparison of the model with the experimental $3D + t$ points is illustrated.

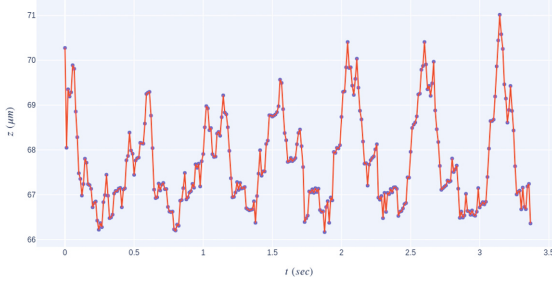


Fig. 2. Fourier model. The experimental data points of the $z_0(t)$ function component are plotted in blue, their corresponding Fourier series is plotted in red with lines for visual clarity.

Given the excellent fitting of the experimental points by the Fourier model, we used the velocity (Eq. 3) on all three $x_i(t), y_i(t), z_i(t)$ function components and for all flagellar points. The latter process produces the set of velocities $\{\dot{\mathbf{r}}_{i,t}\}_{i=0}^{n-1}$ for $t \in \{1, 2, \dots, T\}$. As this set consists of all the velocities for every point for each sperm in the dataset, we choose the middle flagellar point to apply a statistical analysis of the velocities. Even with the restriction to one point, there are 147 sperm and several times for each one of them. Therefore, we considered one descriptor based on the norm of the velocity vectors: the norm-resultant velocity.

The norm-resultant velocity (Eq. 4) is an integrative measure, as it is the norm of the vector sum of all the velocities. In Fig. 3 the violin plots of the norm-resultant velocity for each sperm conditions are shown. The resultant velocity vector is the vector sum of the velocities of the middle flagellar point over all available times.

$$\|\dot{\mathbf{r}}_{res}\| = \left\| \sum_{t=1}^T \dot{\mathbf{r}}_{m,t} \right\|. \quad (4)$$

The norm-resultant velocity distributions for both conditions are visually distinguishable. In the NoCap condition, the distribution is more uniformly dispersed over its range. Meanwhile, in the Cap condition its distribution is localized over its boxplot.

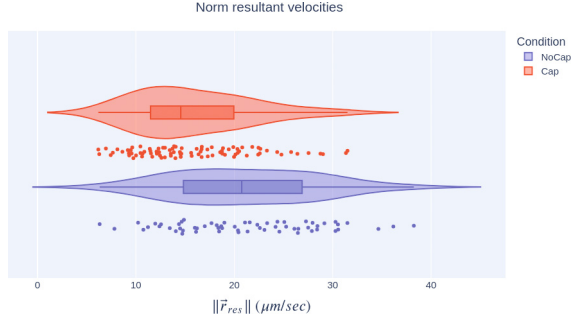


Fig. 3. Norm-resultant velocity distributions. This is a single descriptor for each sperm cell defined as the euclidean norm of the resultant velocity vector for the middle flagellar point. In this figure the norm-resultant velocity is used to compare both experimental conditions: spermatozoa not induced to capacitate (NoCap), and spermatozoa induced to capacitate (Cap). It is visually clear that those cells induced to capacitate have a more localized norm-resultant velocity distribution.

In the norm-resultant distribution it seems that there are more sperm, for the Cap condition, with values less than $10 \mu\text{m}/\text{s}$ (see Fig. 3). This suggests that induction to capacitate changes the distribution of the velocity vectors of the middle flagellar point. This change is such that more sperms have norm-resultant values lower than most of the sperms for the NoCap condition.

A statistical analysis of significance was applied to the Norm Resultant Velocities (NRVs). First, a Shapiro-Wilk normality test was conducted, yielding a p-value of 0.002. This value implies that the data does not follow a normal distribution. Second, a Wilcoxon signed-rank test was performed, resulting in a p-value of 0.02. This p-value allows us to reject the null hypothesis that the NRVs for both the NoCap and Cap conditions come from the same distribution.

In summary, based on the obtained results, it can be concluded that the Norm Resultant Velocities (NRVs) do not follow a normal distribution and that NoCap and Cap conditions have significant different distributions.

4 Discussion and Conclusions

The knowledge of the instant velocity for every single flagellar point is of great value due to the possible venues of research that it opens, such as the analysis of the local forces and their relation with hydrodynamics. Furthermore, with the results presented in this work there are differences in the distributions of the considered descriptor for conditions NoCap and Cap. The existence of statistical differences is of crucial importance in this context. This significance stems from the fact that, according to the literature (e.g., [3]), only a relatively small percentage, ranging from 10 to 20%, of cells under the Cap condition attain hyperactivation. Moreover, the diversity in local dynamical properties could obscure hyperactivation features. However, among those cells with a norm

resultant velocity lower than $10\ \mu\text{m/s}$ it is possible that local hyperactivation characteristics may be exhibited.

A key element in the presented analysis is the local character of the instantaneous velocity. In fact, this locality is the reason why this descriptor was selected to give global information about each cell. But this doesn't imply that a local analysis is not possible or desirable. On the contrary, it indicates that local analysis for both experimental conditions could be very relevant as there are features in the Cap condition absent in the NoCap sample.

However, a local analysis is as interesting as is challenging due to the size of the set of velocity vectors, and its diversity (magnitudes and directions). In this sense, previous findings [4,21] showed that a hyperactivated sperm could display motility features belonging to both NoCap and Cap conditions. Thus, a local analysis demands not only to know if a sperm was or wasn't induced to capacitate, but the identification of the flagellar curves (times) that have distinctive features too. In this line of thought it would be ideal to have a labeled set of hyperactivated sperms (those sperm cells induced to capacitate that truly reach capacitation), this would enable a target-oriented analysis with the main aim of locally characterizing the hyperactivated motility features.

Recently other features of the flagellar beating have been studied: the second harmonics of human flagellar sperm beat [24], the flagellar energetics of beating patterns in mouse sperm [22], and the change of flagellar beating forces before and after capacitation in human sperm [1]. These will be complemented and may be improved by our results because, although in this work only the flagellar mid-point was considered, the methodology described above can be directly applied to any point in the dataset. In contrast the methodology described in [1] is limited to the tip-point of the sperm tail.

In conclusion, we implemented a Fourier fitting method and used it to compute the instantaneous velocity of a mid-flagellar point. Based on this set of instantaneous velocity vectors we proposed a descriptor which showed differences between the two experimental conditions under consideration. Moreover, the obtained set of instantaneous velocity has a wide potential of describing the local details of flagellar dynamics. This local level of accuracy in the flagellar movement is only limited by the experimental setup from which the dataset was obtained.

Meanwhile, the immediate research question that can, and will be, explored are those concerning with the characterization of the distinctive flagellar forms that are empirically identified as signatures of hyperactivated sperm cells. Also, the flagellar movement in different viscosity conditions will be analyzed under the hypothesis that due to high viscosity the average curvature over the flagellum should decrease.

Notes and Acknowledgments. This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was done by the Bioethics Committee from Instituto de Biotecnología, UNAM. Signed written informed consent forms were signed by all healthy donors. This work was supported in part by Dirección General de Asuntos del Personal Académico, Universidad Nacional

Autónoma de México under the Grants IN204922 and IN105222. The work of Paul Hernández-Herrera was supported by the Chan Zuckerberg Initiative DAF under the Grant 2020-225643. The work of D. S. Díaz-Guerrero and H. O. Hernández was supported by a Posdoctoral fellowship and Ph.D. scholarship respectively, from CONACYT-Mexico.









References

1. Battistella, A., Andolfi, L., Stebel, M., Ciubotaru, C., Lazzarino, M.: Investigation on the change of spermatozoa flagellar beating forces before and after capacitation. *Biomater. Adv.* **145**, 213242 (2023)
2. Brokaw, C.J.: Microtubule sliding in swimming sperm flagella: direct and indirect measurements on sea urchin and tunicate spermatozoa. *J. Cell Biol.* **114**(6), 1201–1215 (1991)
3. Cohen-Dayag, A., Tur-Kaspa, I., Dor, J., Maschiach, S., Eisenbach, M.: Sperm capacitation in humans is transient and correlates with chemotactic responsiveness to follicular factors. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11039–11043 (1995)
4. Corkidi, G., Hernández-Herrera, P., Montoya, F., Gadêlha, H., Darszon, A.: Long-term segmentation-free assessment of head-flagellum movement and intracellular calcium in swimming human sperm. *J. Cell Sci.* **134**(3), jcs250654 (2021)
5. Corkidi, G., et al.: Are there intracellular Ca²⁺ oscillations correlated with flagellar beating in human sperm? A three vs. two-dimensional analysis. *Mol. Hum. Reprod.* **23**(9), 583–593 (2017)
6. Corkidi, G., Taboada, B., Wood, C.D., Guerrero, A., Darszon, A.: Tracking sperm in three-dimensions. *Biochem. Biophys. Res. Commun.* **373**(1), 125–129 (2008)
7. Crenshaw, H.C.: Analysis of three-dimensional trajectories of organisms: estimates of velocity, curvature and torsion from positional information. *J. Exp. Biol.* **203**, 961–982 (2000)
8. Gadêlha, H., Hernandez-Herrera, P., Montoya, F., Darszon, A., Corkidi, G.: Human sperm uses asymmetric and anisotropic flagellar controls to regulate swimming symmetry and cell steering. *Sci. Adv.* **6**(31) (2020)
9. Gaffney, E.A., Ishimoto, K., Walker, B.J.: Modelling motility: the mathematics of spermatozoa. *Front. Cell Dev. Biol.* **9**, 71082 (2021)
10. Guerrero, A., Carneiro, J., Pimentel, A., Wood, C.D., Corkidi, G., Darszon, A.: Strategies for locating the female gamete: the importance of measuring sperm trajectories in three spatial dimensions. *Mol. Hum. Reprod.* **17**(8), 511–523 (2011)
11. Hernandez-Herrera, P., Montoya, F., Rendón, J.M., Darszon, A., Corkidi, G.: Sperm Flagellum Center-Line Tracing in Fluorescence 3D+t Low SNR Stacks Using an Iterative Minimal Path Method. In: Karray, F., Campilho, A., Cheriet, F. (eds.) *ICIAR 2017*. LNCS, vol. 10317, pp. 437–445. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59876-5_48
12. Hernández-Aviña, H.O., et al.: Feature-based 3D+t descriptors of hyperactivated human sperm beat patterns. *bioRxiv*, 4 (2023)
13. Hernández-Herrera, P., Montoya, F., Rendón-Mancha, J.M., Darszon, A., Corkidi, G.: 3-D+t human sperm flagellum tracing in low snr fluorescence images. *IEEE Trans. Med. Imaging* **37**, 2236–2247 (2018)
14. Katz, D.F., Cherr, G.N., Lambert, H.: The evolution of hamster sperm motility during capacitation and interaction with the ovum vestments in vitro. *Gamete Res.* **14**, 333–346 (1986)

15. Katz, D.F., Davis, R.O.: Automatic analysis of human sperm motion. *J. Androl.* **8**, 170–181 (1987)
16. Katz, D.F., Diel, L., Overstreet, J.W.: Differences in the movement of morphologically normal and abnormal human seminal spermatozoa. *Biol. Reprod.* **26**(4), 566–570 (1982)
17. Mortimer, D., Kossakowski, J., Mortimer, S.T., Fussell, S.: Prediction of fertilizing ability by sperm kinematics. *J. Assist. Reprod. Genet.* **14**(5), 525 (1997)
18. Mortimer, D., Pandya, I.J., Sawers, R.S.: Relationship between human sperm motility characteristics and sperm penetration into human cervical mucus in vitro. *J. Reprod. Fertil.* **78**(1), 93–102 (1986)
19. Mortimer, D., Mortimer, S.T.: Influence of system parameter settings on human sperm motility analysis using cellsoft. *Hum. Reprod.* **3**(5), 621–625 (1988)
20. Mortimer, S.T.: Casa-practical aspects. *J. Androl.* **21**, 515–524 (2000)
21. Mortimer, S.T., Swan, M.A.: Variable kinematics of capacitating human spermatozoa. *Hum. Reprod.* **10**(12), 3178–3182 (1995)
22. Nandagiri, A., et al.: Flagellar energetics from high-resolution imaging of beating patterns in tethered mouse sperm. *eLife* **10** (2021)
23. Owen, D.H., Katz, D.F.: Sampling factors influencing accuracy of sperm kinematic analysis. *J. Androl.* **14**(3), 210–221 (1993)
24. Saggiorato, G., Alvarez, L., Jikeli, J.F., Kaupp, U.B., Gompper, G., Elgeti, J.: Human sperm steer with second harmonics of the flagellar beat. *Nat. Commun.* **8**(1), 1415 (2017)
25. Sánchez-Cárdenas, C., et al.: Intracellular Ca²⁺ threshold reversibly switches flagellar beat off and on. *Biol. Reprod.* **99**(5), 1010–1021 (2018)
26. Suarez, S.S., Katz, D.F., Overstreet, J.W.: Movement characteristics and Overstreet. and acrosomal status of rabbit spermatozoa recovered at site and time of fertilization. *Biol. Reprod.* **29**(5), 1277–1287 (1983)



Machine Learning Model to Classify Patients with Complicated and Uncomplicated Type 2 Diabetes Mellitus in the New Civil Hospital of Guadalajara “Juan I. Menchaca”

Víctor Manuel Medina-Pérez¹ , Isaac Zúñiga-Mondragón¹ ,
José Alfonso Cruz-Ramos^{2,3} , Kevin Javier Arellano-Arteaga^{2,4} ,
Iryna Rusanova⁵ , Gerardo García-Gil¹ ,
and Gabriela del Carmen López-Armas¹  

¹ Centro de Enseñanza Técnica Industrial, Subdirección de Investigación y Extensión, Laboratorio de Biomédica/Mecatrónica, C. Nueva Escocia 1885, Guadalajara, Jal, Mexico
glopez@ceti.mx

² Departamento de Clínicas Médicas, Universidad de Guadalajara, C. Sierra Mojada 950, Guadalajara, Jal, Mexico

³ Instituto Jalisciense de Cancerología, Coronel Calderón 715, Guadalajara, Jal, Mexico

⁴ Servicio de Medicina Interna, Nuevo Hospital Civil de Guadalajara “Dr. Juan I. Menchaca” Salvador Quevedo y Zubieta 750, Independencia Oriente, 44340 Guadalajara, Jal, Mexico

⁵ Departamento de Bioquímica y Biología Molecular, Facultad de Ciencias, Universidad de Granada, 18019 Granada, Spain

Abstract. Diabetes Mellitus (DM) is a chronic disease worldwide. By 2030 are estimated to be 643 million people with DM, and by 2045 is projected to be 783 million, according to International Diabetes Federation. Machine Learning (ML) can be used as a smart preventive medicine tool for clinical records in hospitals, clinical laboratories, medical personnel, and patients. Implementing ML in current healthcare systems could translate into early diagnosis of DM. This work aimed to implement a classification algorithm for complicated Type 2 DM(T2DM), uncomplicated T2DM, and healthy Mexican participants. For this work, we enrolled 82 subjects from New Hospital Civil Juan I. Menchaca of Guadalajara, divided into 26 complicated T2DM, 26 uncomplicated T2DM, and 30 healthy subjects. ML algorithms used were decision tree (DT), Random Forest, AdaBoost, Bagging Classifier, and Support Vector Machine (SVM). The models use a dataset of 24 different clinical, biological, and molecular variables to discriminate between the 3 groups. The average accuracy was 78% from the C4.5 DT classifier, and we performed an AUC-ROC curve with value = 0.88. ML models can serve for early diagnosis of T2DM in healthcare systems, implementing this in preventive medicine clinics, developing an APP for smart mobile for personal care, and improving pharmaceutical approaches for treating T2DM.

Keywords: Diabetes Mellitus · Machine Learning · smart preventive medicine · decision tree

1 Introduction

1.1 Diabetes Mellitus

Diabetes Mellitus (DM) is an old chronic disease that remains a significant challenge for the public health system. The main feature of diabetes is lifting blood glucose levels over 126mg/dl, caused by defective insulin secretion accompanied by several biological adjustments [1]. The most affected systems are well described and can lead to retinopathy, neuropathy, cardiovascular diseases, and several stages of kidney damage [2]. Until today DM is considered in three categories: type 1 (T1DM), type 2 (T2DM), and gestational diabetes; nevertheless, DM occurs in young adults, in people between the fourth and sixth years of life, and minus degree in the elderly and children; mainly T2DM is often associated with the development of metabolic syndrome, micro and macrovascular complications of DM besides sedentarism and improper eating habits [3]. The external biological factor contributing to this disease is living in a country with low income and middle-income, where 80% of people with T2DM; in this sense, México is considered a middle-income country [4]. This dilemma is increasing without a strategy effective until today. Therefore, exploring T2DM early diagnosis by Machine Learning (ML) tools is nowadays a reality implemented in medicine, primarily due to access to hospital database repositories in developed countries that contain biochemical and anthropometric variables that diagnose DM, such as glycosylated hemoglobin (HbA1c) fasting plasma glucose (FPG), oral glucose tolerance tests (OGTT), random blood glucose levels, Body Mass Index (BMI), waist circumference, insulin among others [5]. The early diagnosis of T2DM makes it easier to control the natural course of the disease and delay micro and macrovascular complications; this way, ML can help healthcare personnel make a preliminary diagnosis of T2DM based on their monthly physical examination data as well as improve their pharmacological management [6, 7].

1.2 Machine Learning in Diabetes Mellitus

Artificial intelligence has emerged as an area of help in medical diagnoses as well as in the most appropriate pharmacological approach for different pathologies; in this case, ML is a technique that has gained confidence in many professional guilds and especially in the medical area and is based on this there are several works with ML and diabetes mellitus this because it is a health emergency. It is necessary to implement more effective technological resources for early diagnosis and extend the micro and macrovascular complications as much as possible or alert people predisposed to developing DM to more significant health care [8]. In this sense, ML tools are affordable to be used even in a mobile application or the creation of a computational aid system (CAS) in more vulnerable populations. Today, in Berlin, exists a CAS since 1974, which is powered with data from 55,000 patients with T2DM and gives them prevalence, incidence, pharmacological treatments, and duration of DM; this is achieved by recording blood and urinary glucose for metabolic criteria [9].

Algorithms classifiers like multilayer perceptron (MLP) and long short-term memory (LSTM) for DM were used by Butt et al. [10], obtained an accuracy of 86.08% and 87.26%, respectively, using PIMA Indian Diabetes database consisting in 8 variables,

and I was considered like target variable, is important to mention that they only compared the state-of-the-art with previously work reported in the same database. On the other hand, Phongying et al. [11] worked on the dataset from the Department of medical services in Bangkok from 2019-2021 with 20,227 samples with 10 attributes related to developing DM; they applied ML tools like DT, random forests, SVM, and K-nearest neighbors and obtained after tuning the hyperparameters, and the interaction terms that the random forest tree had the best performance classification, with 97.5% accuracy, 97.4% precision, 96.6% recall, and a 97% F1-score. Also, Habibollah et al. [12] worked with DT and random forest tools, and they explored the (MASHAD) Study program from Iran; they obtained 9258 records and 18 attributes, considering 17 like predictors and 1 like target; they only identified T2DM or no event of T2DM. In conclusion, they observed the best results with the random forest model and reported 71.1% accuracy, 71.3% sensitivity, and 69.9% specificity, and the AUC of the ROC curve measured 77.3% correspondingly. It should be noted that they used several attributes biochemical and anthropometric, like our work, but not cellular or molecular attributes. In this sense, there is no report of ML with this kind of attributes for diabetes mellitus disease; mainly, the rest of the reported manuscript focus on only classifier to predict DM. Recently, Schallmoser et al. [13] worked with ML models, logistic regression (LR), and gradient-boosted decision trees (GBDTs) for the prediction of the risk of developing micro or macrovascular complications in 13,904 participants with 105 predictors in prediabetes or DM participants from dataset EHRs (Israeli health provider); they included 3 microvascular complications (retinopathy, nephropathy, and neuropathy), and 3 macrovascular complications: peripheral vascular, cerebrovascular and cardiovascular diseases. Their results for prediabetic individuals were plotted in a AUC-ROC of the LR, and GBDTs correspondingly, 0.657 and 0.681 for retinopathy, 0.807 and 0.815 for nephropathy, 0.727 and 0.706 for neuropathy, 0.730 and 0.727 for peripheral vascular disease, 0.687 and 0.693 for cerebral vascular disease, 0.707 and 0.705 for the cardiovascular condition; on the other hand DM participants obtained AUC-ROC in the same ML models LR and GBDTs, respectively, 0.673 and 0.726 for retinopathy, 0.763 and 0.775 for nephropathy, 0.745 and 0.771 for neuropathy, 0.698 and 0.715 for peripheral vascular disease, 0.651 and 0.646 for cerebral vascular disease, and 0.686 and 0.680 for the cardiovascular condition. They conclude that ML models can predict prediabetes and DM patients to develop micro and macrovascular complications in this population. Hence, ML methods are widely used in diabetes prediction and obtain favorable results; in this work, a DT model was implemented to classify T2DM patients into uncomplicated, complicated patients and healthy participants.

2 Methods

2.1 Database

For this study, we enrolled 82 Mexican participants from New Hospital Civil of Guadalajara obtained through 2020–2021 previously reported [14]. The study was conducted following the Helsinki Declaration and the General Health Law on research; the hospital's Ethical Committee accepted this study with the number approbation 17 CI 14 039

116 COFEPRIS, and by the no: 940/CEIH/2019 of the Ethics Committee of the University of Granada, Spain. Before biological samples were collected, all participants were informed about the study and signed informed consent forms. All participants were adults from 40–65 years of age and divided into three groups of study: 1. Healthy individuals without a family history of first-degree diabetes, besides non-disease of kind inflammatory, endocrine, vascular, or autoimmune, were not taking any pharmaceutical treatment. For patients with T2DM who had at least 5 years with a diagnosis of DM according to ADA criterion [5] (fasting plasma glucose ≥ 126 mg/dL, oral glucose tolerance test ≥ 200 mg/dL, or glycosylated hemoglobin (HbA1c) $\geq 6.5\%$) and were divided as T2DM uncomplicated and T2DM complicated. The uncomplicated patients were defined with reasonable glucose control and free of micro (diabetic neuropathy, diabetic nephropathy, and diabetic retinopathy) or macrovascular complications of DM (atherosclerosis, myocardial infarction, stroke, and cerebrovascular disease) [15]. Patients with DM complications were those with any micro or macro complications of DM before mentioned. We obtained a complete clinical history of each patient and measured anthropometric parameters, Body Mass Index (BMI) and waist circumference, and biochemical and molecular variables.

2.2 Data Preprocessing

The initial dataset comprised 56 clinical, biological, and molecular variables from 82 subjects. Aiming to create a more concise dataset, variables with greater significance were selected to be kept in the final dataset. Also, variables representing T2DM complications in patients were removed (retinopathy, nephropathy, neuropathy, stroke, pharmacological treatment). Variables with a small percentage of missing data were imputed with the mean value of the corresponding group [16]. For variables such as sex, a value of 1 for females and 2 for males was given, as well as for the variables of alcohol consumption, smoking, and physical activity, 1 represents an affirmation and 0 a negation. The final dataset comprises the following 24 variables (see Table 1).

The 24 variables listed in the table above correspond to the more significant attributes that helped the DT to discriminate between healthy subjects, patients with uncomplicated T2DM, and patients with complicated T2DM. The primary reason for analyzing the groups in a split manner corresponds precisely to differentiating clinical profiles; dividing patients into groups according to the presence or absence of complications helps to understand the unique characteristics of each group better and to identify specific risk factors associated with complications; similarly by analyzing and comparing these three groups, unique patterns, biomarkers, and associations can be discovered that might not be evident if all diabetes patients were considered as a homogeneous group. This may lead to new insights into the underlying mechanisms of complications and diabetes in general.

2.3 Implementation and Training of Classifier Models

Different classifier models were trained and validated with the corresponding datasets, and all classifier models were implemented using the scikit-learn (version 1.0.2) and Python library existing models (version 3.9.13). The dataset was divided into train and

validation subsets using the scikit-learn train/test splitter; 70% ($n = 57$) of the records were used for the model training, while the remaining 30% ($n = 25$) were used for the model validation and we performed a five-fold cross-validation procedure. The split was performed using stratification; train and validation subsets were tested and had a significant representation of the 3 groups; the training subset was composed of 21 healthy subjects, 18 uncomplicated, and 18 complicated T2DM patients, and the validation subset with 9 healthy subjects, 8 uncomplicated and 8 complicated T2DM patients, are shown in Fig. 1.

The dataset split train-validation process, training, and model validation were executed 200 times for the classifier models. The optimization of hyperparameters was carried out exhaustively through the modification (tuning) in its maximum depth, the proportion or number of cases of the training and test set, the minimum number of subjects per sheet, and all pertinent biological markers were selected according to the physiology of Type 2 Diabetes mellitus, that is, the characteristics were designated in such a way that those variables that were not in the definition of complicated diabetes were included. The number of iterations for optimizing and adjusting the model in a total of 200 executions was also modified and adjusted. For this work, it was the performance of those classifier models.

2.3.1 Support Vector Machine

This algorithm was built with the default parameters provided by the Scikit library), is a supervised learning algorithm used for classification and regression. It searches for a hyperplane that best separates data classes in a multidimensional space. SVM maximizes the margin between classes, resulting in good generalization to new data. It can handle both linearly separable and non-linearly separable data using kernel functions.

2.3.2 Random Forest

This algorithm was built using 350 estimators, while the remaining parameters were set to the default option; this ML tool creates multiple decision trees and combines their predictions to obtain a result. Each tree is trained on a random subsample of the data and uses bootstrapping to build different data sets. The predictions of the individual trees are then averaged or voted to make a final decision.

2.3.3 ADABOOST

This classifier was built with 100 estimators and a 0.01 learning rate; all other parameters were set as the default option; this ensemble algorithm combines several weak learning models (e.g., weak decision trees) to form a stronger model. Adaboost gives more weight to instances misclassified in the previous iteration at each iteration, allowing the model to focus on hard-to-classify cases. The weak models are then weighted according to accuracy and combined to make joint decisions.

Table 1. Summary of clinical, biochemical, and molecular characteristics of T2DM subjects. ‡Relative expression was calculated using the $2^{-\Delta\Delta\text{CT}}$ method (ΔCt target miR- ΔCt control gene)

N°	Variable	Type	Value
0	Age (Years)	Numeric	[1–100]
1	Sex	Nominal	[1–2]
2	Waist circumference (cm)	Numeric	[70–150]
3	BMI index (kg/m ²)	Numeric	[17–47]
4	Systolic Blood pressure (mm/Hg)	Numeric	[80–205]
5	Diastolic Blood pressure mm/Hg)	Numeric	[60–115]
6	HOMA IR (fasting glucosa mg/dL * fasting insuline IU/L/405)	Numeric	[0.25–20]
7	HbA1c (%)	Numeric	[4–13]
8	TG (mg/dL)	Numeric	[40–360]
9	Cholesterol(mg/dL)	Numeric	[100–320]
10	Urea (mg/dL)	Numeric	[13–135]
11	Creatinine (mg/dL)	Numeric	[0.1–2.6]
12	miR_21 (relative expression) ‡	Numeric	[30–350]
13	miR_126 (relative expression) ‡	Numeric	[10–195]
14	IL_6 (pg/mL)	Numeric	[0.01–5.50]
15	IL_10 (pg/mL)	Numeric	[0.20–13.5]
16	IL_18 (pg/mL)	Numeric	[10–426]
17	TNF_alfa (pg/mL)	Numeric	[6–127]
18	Diabetic Family	Nominal	[0–6]
19	Alcohol intake	Nominal	[0–1]
20	Smoke	Nominal	[0–1]
21	Physical activity	Nominal	[0–1]
22	Glucose (mg/dL)	Numeric	[70–380]
23	Years of diabetic (years)	Numeric	[0–35]

2.3.4 Bagging Classifier

This model was built using 200 estimators, with the remaining parameters set to the default value; this algorithm is an ensemble technique used to improve the accuracy and stability of ML models. Using bootstrapping, the Bagging Classifier builds multiple base models (e.g., decision trees) on random subsets of the training data. Then, the predictions of the base models are combined by voting to determine the final class of an instance, and finally, we implemented the C4.5 DT (the class to measure the quality of a split was set to “gini,” splitter to “best,” no max depth, and the rest of the parameters as the default option). Since this last ML model was the one that gave us the best overall

performance, we will focus on describing it in detail. Finally, the precision, recall, F1-score, and accuracy results of every execution were saved and averaged to review the results further.

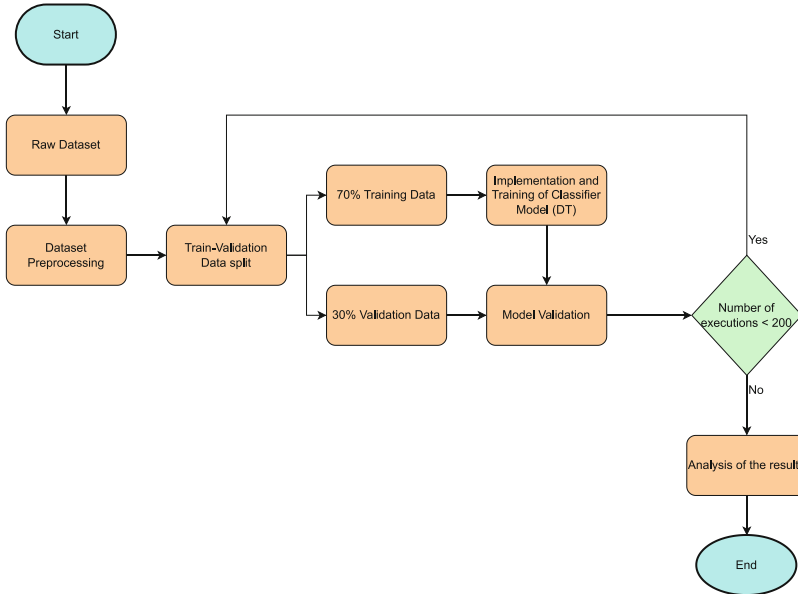


Fig. 1. Flow diagram for data processing and functioning sequences of the proposed diabetes classification model.

2.4 Model Decision Tree

The DT methodology was used, C4.5, as a classifier. It was implemented in Python, a high-level programming language whose function emphasizes the readability of its code [17]. Due to its free software characteristics, it is used to develop academic, scientific, and technological applications. It is a programming language that allows the coexistence of different paradigms, such as oriented object programming, structured programming, imperative programming, reactive programming, and functional programming searching to improve the production of the development of the projects. The DT model is a primary and regression method with a tree structure describing classifying instances based on features [18]. It is considered a set of “if-then” rules, which also function as conditional probability distributions defined in feature and class space.

The DT utilizes a tree structure, starting with a single node representing the training samples [19]. If the samples are located inside the same class, the node is transformed into a leaf and labeled with the same class. Otherwise, the algorithm chooses the discriminative attributes as the actual node of the DT. According to the value of the current attribute of the decision node, the training samples are divided into different subsets, each in different forms, values, and branches. The anterior steps are repeated for each

subset or for each branch obtained, recursively forming a DT on each partitioned sample [20]; see the pseudocode in Table 2. Typical DT algorithms are ID3, C4.5, and CART, among others.

This study used the DT C4.5 and some parameters like entropy and GINI index that work in DT with the information that was input from the dataset used to classify each subset test. The dataset generated has nominal attributes for the classification tasks with non-missing values. In general, if a probability distribution $P = (p_1, p_2, p_3, \dots, p_n)$ and an example S is given, then the information carried by this distribution is known as the entropy P , and is calculated by: Eq. (1):

$$\text{Entropy } P = - \sum_{i=1}^n p_i \cdot \log(p_i) \quad (1)$$

In this work, when the entropy level is 0, it will be the level of order, 1. On the other hand, the information gain allows us to measure the degree of impurity of the classes for all the examples and, therefore, any position of the tree under construction. It must have a new function that allows one to select the attribute that should label the current node. This defines the gain Eq. (2) for a test T and a position p :

$$\text{Gain}(p, T) = \text{Entropy } P - \sum_{j=1}^n (p_j \cdot \text{Entropy})(p_j) \quad (2)$$

Where the values (p_j) is the set of all possible values for attribute T . This measure can then be used to determine which attribute is best and build the DT where at each node is the attribute with the highest information gain of all the attributes not yet considered in the path from the root node [21].

Information gain of attribute A Eq. (3):

$$\text{Gain}(A) = \text{Info}(D) - \text{Info}_A(D) \quad (3)$$

Pre-segmentation of information entropy Eq. (4):

$$\text{Info}(D) = \text{Entropy}(D) = - \sum p(j|D) \log_2 p(j|D) \quad (4)$$

Entropy of distributed information Eq. (5):

$$\text{Info}_A(D) = \sum (n_i/n) \text{Info}(D_i) \quad (5)$$

The impurity GINI is one of the possible measures for generating the DT. It provides more information about the data distribution per node than the classification used in the accuracy reporting. The impurity of the model nodes is calculated using each count of each objective category divided by every record obtained by one node. The total GINI impurity is calculated as a sum of squares of the count of proportion divided by all the objective categories per node from which one is subtracted. The result is multiplied by the quantity of records Eq. (6). For example, when splitting a tree node, the algorithm searches for a field with the highest improvement in total impurity, calculated as the total

impurity among all potential child nodes subtracted from the total impurity of the parent node. The goal is for the gini index to be as small as possible.

$$\text{Gini}(Xq) = 1 - \sum_{k=1}^k (p_{k,q})^2 \quad (6)$$

2.5 Algorithm C4.5

C4.5 is an algorithm used to generate a DT developed by Ross Quinlan [18], and this is an extension of the ID3 algorithm. DT generated by C4.5 can be used for classification; therefore, C4.5 is almost always referred to as a statistical classifier. C4.5 builds DT from a training data set as ID3 does, using the concept of information entropy. In a DT, the training data refers to a collection of examples used to train the tree. Each example is denoted as S_i and represents a data point with associated information. This collection of examples forms the dataset the DT algorithm uses to learn and build the tree structure. Each example S_i is represented as a vector of attributes or features, such as X_1, X_2 , etc. These features describe the characteristics or properties of the data point S_i .

The training data are augmented with a vector $C = C_1, C_2$, where C_1, C_2 , represent the class to which each sample belongs. Each example S_i is associated with a specific class or category, and a class label represents this class. The vector C contains these class labels for each example S_i . These labels indicate the predefined categories that each data point belongs to.

In every tree node, the DT C4.5 selects the variable that better divides the sample into different data subsets with a predominant class; the criteria are normalized for information gain resulting in the selection of the variable that better divides the data. The decision parameter will be the variable that achieves the best-normalized information gain [17]. See pseudocode Table 2.

This algorithm has 3 main features; firstly, the DT C4.5 creates a leaf node when all samples are classified in the same class. In the second instance, when none of the variables show a significant information gain, the DT C 4.5 creates a decision node to continue with the tree; finally, when an unseen class is found, the DT also creates a decision node to continue with the tree C4.5.

2.6 Metrics to Evaluate the Performance of Classifiers on Unbalanced Datasets

It is essential to mention that the performance of a classification method on data sets with two classes is to use a confusion matrix. This contains information about the predictions made by the classifier using the number of true positives, true negatives, false positives, and false negatives. Table 3 shows an example of the confusion matrix.

The confusion matrix should only be used for classification accuracy to determine the performance of a classifier; it is not suitable when dealing with unbalanced data sets. This is because high classification accuracy values can be obtained in these types of applications by simply ignoring the minority class; however, the latter is essential in cases such as web mining detection, direct marketing, and medical diagnostics [22].

Table 2. Algorithm C4.5 for the construction of DT

C4.5 Gnal Pseudocode
<pre> FUNCTION CreateDecisionTree(data, attributes) IF all data belongs to a single class THEN RETURN a leaf node with that class IF no attributes left THEN RETURN a leaf node with the majority class in the data ELSE chosenAttribute = SelectWinningAttri- bute(data, attributes) NEW NODE decisionNode WITH chosenAttribute FOR each value in chosenAttribute.VALUES DO subsetData = FilterData(data, chosenA- ttribute, value) IF subsetData is empty THEN RETURN a leaf node with the majo- rity class in the data ELSE NEW NODE child WITH CreateDeci- sionTree(subsetData, attributes - {chosenAttri- bute}) ADD child TO decisionNode WITH la- bel value END IF END FOR RETURN decisionNode END IF END FUNCTION FUNCTION SelectWinningAttribute(data, attributes) bestAttribute = NULL bestGain = 0 FOR each attribute IN attributes DO gain = CalculateGain(data, attribute) IF gain > bestGain THEN bestGain = gain bestAttribute = attribute END IF END FOR RETURN bestAttribute END FUNCTION </pre>

(continued)

Table 2. (continued)

```

FUNCTION CalculateGain(data, attribute)
  dataEntropy = CalculateEntropy(data)
  conditionalEntropy = 0

  FOR each value IN attribute.VALUES DO
    subsetData = FilterData(data, attribute,
value)
    weight = size of subsetData / size of data
    conditionalEntropy = conditionalEntropy +
(weight * CalculateEntropy(subsetData))
  END FOR

  gain = dataEntropy - conditionalEntropy
  RETURN gain
END FUNCTION

FUNCTION CalculateEntropy(data)
  countClass(class) counts the number of exam-
ples in data with the given class label

  entropy = 0
  FOR each class IN classes DO
    probability = countClass(class) / size of
data
    IF probability > 0 THEN
      entropy = entropy - (probability *
log2(probability))
    END IF
  END FOR

  RETURN entropy
END FUNCTION

FUNCTION FilterData(data, attribute, value)
  NEW SET subset
  FOR each example IN data DO
    IF example.attribute = value THEN
      ADD example TO subset
    END IF
  END FOR
  RETURN subset
END FUNCTION

```

According to the Confusion Matrix, True Positive (TP) is a Positive Value that is Correctly Classified as Positive, then False Positive (FP) is a Negative Value that is Incorrectly Classified as Positive; also, We Have False Negative (FN) that is a Positive Value Classified as Negative and Finally We Have True Negative (TN) Which Indicates a

Table 3. Confusion matrix

		Actual class	
		Positive	Negative
Predicted class	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

Negative Value Predicted Classified Negative (See Table 3). In This Work, the Confusion Matrix Was Calculated in a Python Program and Obtained the Subsequent Metrics: $Accuracy = [(TP + TN) / (TP + TN + FP + FN)] \times 100$, Sensitivity or $Recall = [(TP / (TP + FN))]$, $Specificity = [TN / (TN + FP)]$, $Precision = [TP / (TP + FP)]$, and $F1 - score = (2 \times Precision \times Recall) / (Precision + Recall)$ See Fig. 3. Besides, We Performed in the Phytion Program an Area Under the Curve (AUC) with a Cutoff of 0.5, and Values Were Plotted in the AUC- Receiver Operation Curve (ROC) of DT's Best and Worst Performance. Fig. 4.

3 Results and Discussion

The design objective of this DT C4.5 was programmed to identify 3 groups of patients belonging to the following categories: 1. Healthy subjects, 2. Patients with uncomplicated T2DM, and 3. Patients with T2DM with complications; the results are shown in Fig. 2.

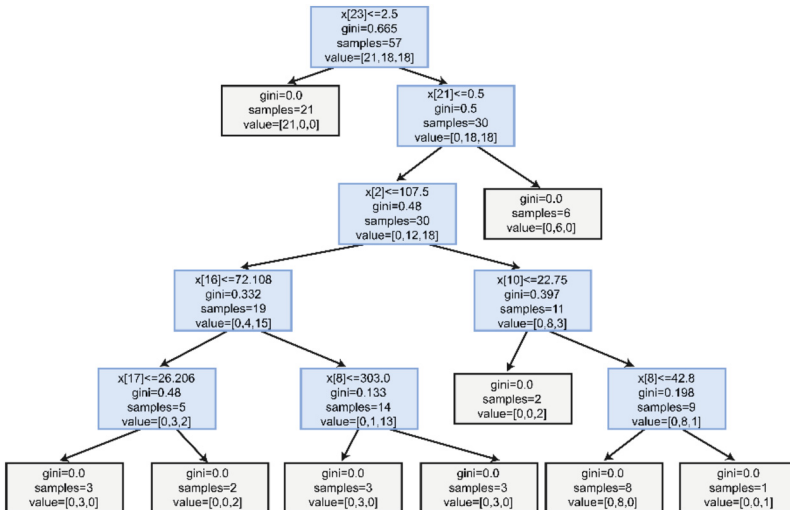


Fig. 2. DT C4.5: this tree shows in the leaves (white boxes) where “x” represents the variable taken into consideration to make a classification; “gini” expresses the degree of impurity; the “samples” are subjects. Finally, the “value” represents the vector where the first index belongs to healthy, the second is T2DM uncomplicated, and the last is T2DM complicated subjects.

The results of the ML classifiers are shown in Table 4, precision, recall, F1 score, and classification accuracy corresponding to the mean \pm SD of the dataset obtained from the New Hospital Civil Juan I. Menchaca, considering 24 attributes of 82 patients (Table 1), subdivided in healthy, uncomplicated T2DM and complicated T2DM patients. We obtain from DT C 4.5 the best results of all classifiers, showing for precision of 0.77 ± 0.04 , recall of 0.76 ± 0.03 , F1 score of 0.76 ± 0.02 and accuracy of $78\% \pm 0.02$. The rest of the classifier are described in the Table 4.

Table 4. Performance metrics of ML model

Classifier	Precision	Recall F1 Score	Accuracy
C4.5 (DT)	0.77 ± 0.04	0.76 ± 0.03 0.76 ± 0.02	$78\% \pm 0.02$
Adaboost	0.75 ± 0.0	0.75 ± 0.0 0.76 ± 0.0	$76\% \pm 0.0$
Bagging	0.73 ± 0.04	0.72 ± 0.03 0.70 ± 0.04	$73\% \pm 0.03$
RF	0.74 ± 0.04	0.72 ± 0.03 0.69 ± 0.04	$73\% \pm 0.03$
SVM	0.66 ± 0.0	0.66 ± 0.0 0.64 ± 0.0	$68\% \pm 0.0$

We also performed a confusion matrix multiclass to determine the classification problem of 3 groups: healthy subjects, 2TDM uncomplicated, and 2TDM complicated the results are the average of 200 executions of the DT C4.5.

Finally, we plotted the results of DT C4.5 in T2DM complicated and was analyzed by ROC curve showing accuracy in the best performance with an area under the curve (AUC) = 0.88 and the worst performance with AUC = 0.76. Fig. 4.

Several ML works have been developed to manage databases of patients with diabetes mellitus to establish a prediction in the course of the disease; some focus on complicated T2DM, such as the work of Ljubic et al. [19], who worked with a dataset of more than 1 million people diagnosed with DM over 9 years, they focused on predicting micro and macrovascular complications of DM. They considered 10 of these and the number of hospitalizations on average 1 to 4 and employing the Recurrent neural network (RNN) with a gated recurrent unit (GRU), they obtained an accuracy of 73% (myocardial infarction) and 83% (chronic ischemic heart disease). On the other hand, Agliata et al. [23] worked with 3 different datasets to diagnose DM2 and included characteristics such as insulin sensitivity, age, sex, BMI, and glycosylated hemoglobin, among other parameters; to train their data; they chose a binary classifier trained in scratch which gave them an area under the receiver operating characteristic curve (AUROC) of 0.934; however, these authors mention the importance of the characteristics to be used in a database; since there are too many variables that may not be considered, they decided to work with the simplest to obtain from almost any medical history. Pan et al. [24].

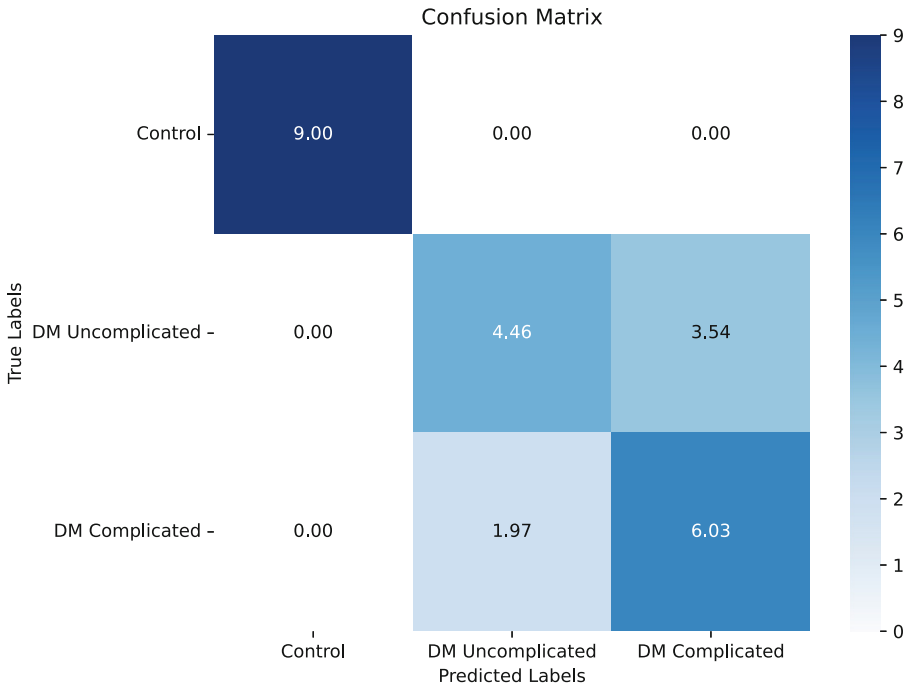


Fig. 3. Confusion matrix for multiclass DT classifier.

They focused on building a multivariate logistic regression model to predict the risk of diabetic retinopathy in a Chinese population of 2385 patients already diagnosed with DM, where they built model I with classical predictors such as glycosylated hemoglobin, systolic blood pressure, ongoing disease, postprandial blood glucose, and the urinary albumin/creatinine ratio, obtaining an AUROC of 0.703 with an accuracy of 0.79. In this sense, our database is of Mexican patients, and with variables beyond only clinical, anthropometric, and biochemical data, but also included cellular (interleukins) and molecular (microRNAs) variables that, to date, there are no similar reports in Mexican patients where DM2 is a severe health problem, our ML model consisted of a DT C4.5 which gave us an accuracy of 78%, which for the number of patients in a small sample shows that it is very efficient at the time of classifying healthy patients with uncomplicated T2DM and with T2DM with complications.

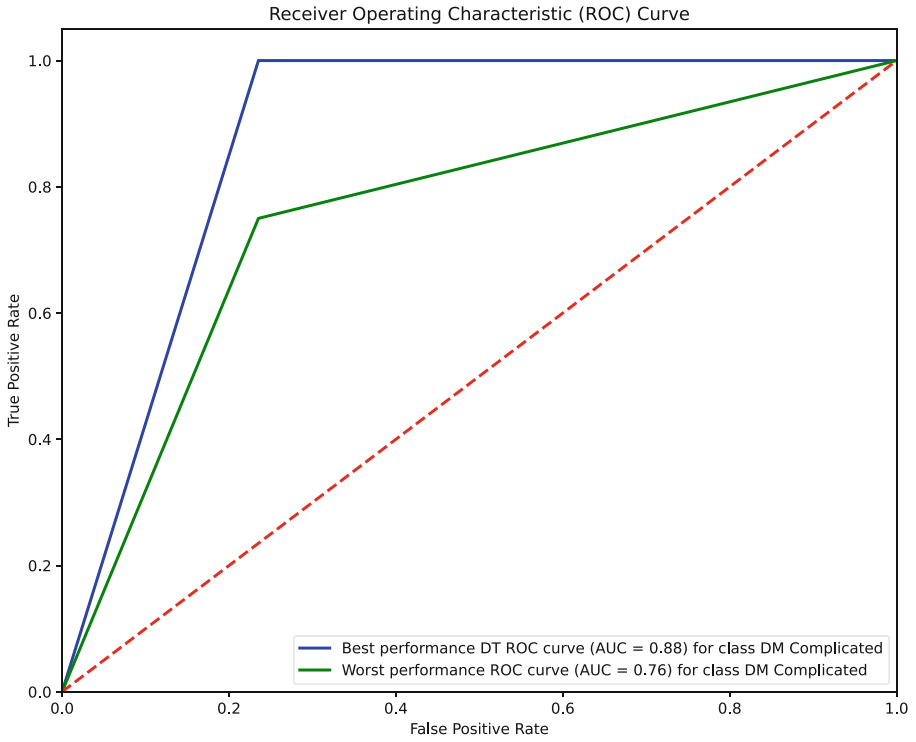


Fig. 4. ROC curve analysis for DT C4.5 classifier during the use of the best performance with a value of $AUC = 0.88$ and worst performance with a value of $AUC = 0.76$ for T2DM complicated class.

4 Conclusion

This work was carried out in Mexican patients with T2DM and is the first to our knowledge in the classification by ML of the complications of the disease itself, as well as the inclusion of novel cellular and molecular parameters already studied T2DM, such as proinflammatory and anti-inflammatory interleukins (IL-6, TNF- α , IL-10, and IL-18) and microRNAs (Mirs-21 and 126) that can regulate gene expression and thus, may become a therapeutic target in the future; our primary intention is to encourage the existence of more robust databases with markers that can contribute to a better early diagnosis in order to avoid the disabling complications of DM2 since it is a worldwide public health problem. Our future work will be focus on making this database bigger, since this was only the implementation of the machine learning algorithms as an exploratory classification method as well as developing deep learning to discover possible best results with a larger database.








References

1. Ludwig, D.S.: The glycemic index: Physiological mechanisms relating to obesity, diabetes, and cardiovascular disease. *JAMA* **287**(18), 2414 (2002). <https://doi.org/10.1001/jama.287.18.2414>
2. Chawla, A., Chawla, R., Jaggi, S.: Microvascular and macrovascular complications in diabetes mellitus: Distinct or continuum? *Indian J. Endocr. Metab.* **20**(4), 546 (2016). <https://doi.org/10.4103/2230-8210.183480>
3. Ceriello, A., Ihnat, M.A., Thorpe, J.E.: The “metabolic memory”: Is more than just tight glucose control necessary to prevent diabetic complications? *J. Clin. Endocrinol. Metab.* **94**(2), 410–415 (2009). <https://doi.org/10.1210/jc.2008-1824>
4. Chan, J.C.N., et al.: The Lancet commission on diabetes: Using data to transform diabetes care and patient lives. *Lancet* **396**(10267), 2019–2082 (2020). [https://doi.org/10.1016/S0140-6736\(20\)32374-6](https://doi.org/10.1016/S0140-6736(20)32374-6)
5. ElSayed, N.A., et al.: 10. Cardiovascular disease and risk management: Standards of care in Diabetes—2023. *Diabetes Care* **46**(Supplement_1), S158–S190 (2023). <https://doi.org/10.2337/dc23-S010>
6. Fujihara, K., Sone, H.: Machine learning approach to drug treatment strategy for diabetes care. *Diabetes Metab. J.* **47**(3), 325–332 (2023). <https://doi.org/10.4093/dmj.2022.0349>
7. Lee, B.J., Kim, J.Y.: Identification of type 2 diabetes risk factors using phenotypes consisting of anthropometry and triglycerides based on machine learning. *IEEE J. Biomed. Health Inform.* **20**(1), 39–46 (2016). <https://doi.org/10.1109/JBHI.2015.2396520>
8. Chaki, J., Thillai Ganesh, S., Cidham, S.K., Ananda Theertan, S.: Machine learning and artificial intelligence based Diabetes Mellitus detection and self-management: A systematic review. *J. King Saud Univ. - Comput. Inform. Sci.* **34**(6), 3204–3225 (2022). <https://doi.org/10.1016/j.jksuci.2020.06.013>
9. Thoeleke, H., Meusel, K., Ratzmann, K.-P.: Computer-aided system for diabetes care in Berlin, G.D.R. *Comput. Methods Programs Biomed.* **32**(3–4), 339–343 (1990). [https://doi.org/10.1016/0169-2607\(90\)90118-S](https://doi.org/10.1016/0169-2607(90)90118-S)
10. Butt, U.M., Letchmunan, S., Ali, M., Hassan, F.H., Baqir, A., Sherazi, H.H.R.: Machine learning based diabetes classification and prediction for healthcare applications. *J. Healthcare Eng.* **2021**, 1–17 (2021). <https://doi.org/10.1155/2021/9930985>
11. Phongying, M., Hiriote, S.: Diabetes classification using machine learning techniques. *Computation* **11**(5), 96 (2023). <https://doi.org/10.3390/computation11050096>
12. Esmaily, H., Tayefi, M., Doosti, H., Ghayour-Mobarhan, M., Nezami, H., Amirabadizadeh, A.: A comparison between decision tree and random forest in determining the risk factors associated with type 2 diabetes. *J. Res. Health Sci.* **18**(2), 1–7 (2018)
13. Schallmoser, S., Zueger, T., Kraus, M., Saar-Tsechansky, M., Stettler, C., Feuerriegel, S.: Machine learning for predicting micro- and macrovascular complications in individuals with prediabetes or diabetes: Retrospective cohort study. *J. Med. Int. Res.* **25**, e42181 (2023). <https://doi.org/10.2196/42181>
14. López-Armas, G.C., et al.: Role of c-miR-21, c-miR-126, redox status, and inflammatory conditions as potential predictors of vascular damage in T2DM patients. *Antioxidants* **11**(9), 1675 (2022). <https://doi.org/10.3390/antiox11091675>
15. Bin Rakhis, S.A., AlDuwayhis, N.M., Aleid, N., AlBarrak, A.N., Aloraini, A.A.: Glycemic control for type 2 diabetes mellitus patients: A systematic review. *Cureus* **14**, e26180 (2022). <https://doi.org/10.7759/cureus.26180>
16. Zhang, Z.: Missing data imputation: Focusing on single imputation. *Ann. Transl. Med.* **4**(1), 9 (2016). <https://doi.org/10.3978/j.issn.2305-5839.2015.12.38>

17. Quinlan, J.R.: Improved use of continuous attributes in C4.5. arXiv, 29 de febrero de (1996). Accedido: 30 de junio de 2023. [En línea]. Disponible en: <http://arxiv.org/abs/cs/9603103>
18. Quinlan, J.R.: Induction of decision trees. *Mach. Learn.* **1**(1), 81–106 (1986). <https://doi.org/10.1007/BF00116251>
19. Friedl, M.A., Brodley, C.E.: Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* **61**(3), 399–409 (1997). [https://doi.org/10.1016/S0034-4257\(97\)00049-7](https://doi.org/10.1016/S0034-4257(97)00049-7)
20. Habibi, S., Ahmadi, M., Alizadeh, S.: Type 2 diabetes mellitus screening and risk factors using decision tree: Results of data mining. *Glob. J. Health Sci.* **7**(5), p304 (2015). <https://doi.org/10.5539/gjhs.v7n5p304>
21. Hssina, B., Merbouha, A., Ezzikouri, H., Erritali, M.: A comparative study of decision tree ID3 and C4.5. *Int. J. Adv. Comput. Sci. Appl.* **4**(2), 13–19 (2014). <https://doi.org/10.14569/SpecialIssue.2014.040203>
22. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: SMOTE: Synthetic minority over-sampling technique. *Jair* **16**, 321–357 (2002). <https://doi.org/10.1613/jair.953>
23. Agliata, A., Giordano, D., Bardozzo, F., Bottiglieri, S., Facchiano, A., Tagliaferri, R.: Machine learning as a support for the diagnosis of type 2 diabetes. *IJMS* **24**(7), 6775 (2023). <https://doi.org/10.3390/ijms24076775>
24. Pan, H., et al.: A risk prediction model for type 2 diabetes mellitus complicated with retinopathy based on machine learning and its application in health management. *Front. Med.* **10**, 1136653 (2023). <https://doi.org/10.3389/fmed.2023.1136653>



Simulator for Cardiovascular Procedures

Tomás Efraín Sánchez Pérez , Raúl Martínez Ruiz ,
Jesus Moises Martinez Buendia , Iliana González Pérez ,
Carlos Alberto Gallegos Hernández , Efraín Muñoz Montero ,
Luis Antonio Jiménez Salinas , and Arturo Abundes Velasco

Instituto Nacional de Cardiología, Juan Badiano 1, Belisario Domínguez Secc 16,
Tlalpan, 14080 Ciudad de México, Mexico
setth2023@gmail.com

Abstract. This project integrates an interface and a prototype that allows the simulation of vascular procedures, both previously performed at the “Instituto Nacional de Cardiología Ignacio Chávez”, with the objective that these can work together and achieve a simulator with greater competitiveness, performance, and functionality. This simulator incorporates a mechanical, hydraulic, digital, and imaging system like the one used in fluoroscopy, as well as interchangeable and personalized anatomical models. The user can operate it manually by using analog controls or more precise instructions through the interface control panel. By using this simulator, medical specialists, and residents will be able to practice the placement of various cardiovascular devices and have an approach to specific clinical cases.

Keywords: simulator · anatomic models · cardiovascular devices

1 Introduction

Cardiovascular procedures pose a significant challenge in the medical field, particularly within the educational context. This is due to the inherent limitations of the traditional training approach, which relies on expert-guided supervision. The constraints of this approach become apparent when considering factors such as the student-to-teacher ratio, available working hours, the risk stemming from technical complexity, and the legal and ethical implications [1].

In response to these limitations, alternatives to the traditional teaching approach have emerged. These options involve the utilization of animals and cadavers as instructional resources. However, both approaches come with considerable disadvantages. In the case of using animals for experimental procedures, ethical dilemmas and discrepancies arise, as the animals used are often healthy specimens, leading to substantial differences in physiological responses, pharmacological interactions, and disease manifestations compared to patients and their cardiac conditions. On the other hand, the use of cadavers raises legal implications, and specifically, in the manipulation of a guide through the radial artery, it becomes hindered by blood coagulation [2].

Another emerging training method is the use of virtual reality or augmented reality, which is conceptually appealing but doesn't provide a reliable physical feedback medium, potentially lacking realism due to current technological limitations [3].

Currently, various research fields have been working with a range of physically accurate anatomical models, which are based on image capture (Magnetic Resonance Imaging, Computed Tomography, and Ultrasound) and employ 3D printing technology with a variety of materials, each with distinct properties and finishes [4, 5]. By applying these methodologies, the capability to generate models in both common and catastrophic clinical scenarios is attainable. When needed, it is possible to make suitable modifications to replicate situations like serious medical complexities, like arterial blockage or a congenital anomaly.

The focus of this study was directed towards the development of a simulator for transcatheter cardiovascular procedures, which include guided placement of stents. This simulator was constructed within a controlled environment, resembling real procedure situations, utilizing 3D printing technology. Furthermore, it was meticulously designed to comprehensively incorporate critical tools for the procedure, including radial puncture and motion control for fluoroscopy image acquisition, all without exposure to ionizing radiation.

2 Materials and Methods

2.1 Segmentation

Segmentation was executed using the Mimics software (Materialise. <https://www.materialise.com/en/healthcare/mimics-innovation-suite/mimics>) and with the support of additive manufacturing, the generation of multiple resin-printed models was achieved (Clear Resin by Formlabs). The segmented files are processed into the area of interest and adapted to the inputs and outputs of the hydraulic system (Fig. 1). The integration of these models with a hydraulic system is essential and will be discussed further, as such incorporation enables the reproduction of both radial and coronary flows. These models are interchangeable, allowing for the selection of the most appropriate one according to the specific scenario of interest.

2.2 Hydraulic System

The hydraulic system is based on circulating water through the segmented model to simulate flow conditions. To achieve this, two tools were employed: a peristaltic pump and a set of hoses with outer and inner diameters like physiological values (Average reported diameter 2.68 ± 0.15 mm [6]). This approach aims to achieve a more realistic pulsation.

Peristaltic Pump. The choice was made to use a 12V peristaltic pump with a 2-roller design. The operating voltage was determined within a range of 4 to 7 V, allowing the generation of simulated pulsations at a frequency comparable to the 60–100 bpm range. This setting can be modulated to meet specific requirements according to the scenario's needs (Fig. 2).



Fig. 1. Aortic artery adapted for integration into the hydraulic system.

Radial Puncture. The puncture process is performed through a simulated artery surrounded by a portion of radial bone, where the hose is routed. This point has been designated specifically for palpation and insertion of the introducer sheath, followed by guide wire insertion and device navigation (Fig. 3).

2.3 Motion Axes

The operation of a fluoroscope involves a C-arm that performs movements known as cephalocaudal and oblique motions. Additionally, there is a linear motion on the bed. The spectrum of movements covers a range of 90° in the cranial direction, 90° in the caudal direction (Fig. 4), 90° in the left oblique direction, and 90° in the right oblique direction (Fig. 5). For this device, a linear bed displacement of 30 cm is controlled. This measurement is determined as the optimal range to allow comprehensive visualization of the printed model.

The system features three axes of movement: cephalocaudal, oblique, and linear. The first two axes can be controlled using a joystick, while the third is operated using two buttons. It is important to emphasize that all three systems are in constant feedback, always enabling precise determination of their position.

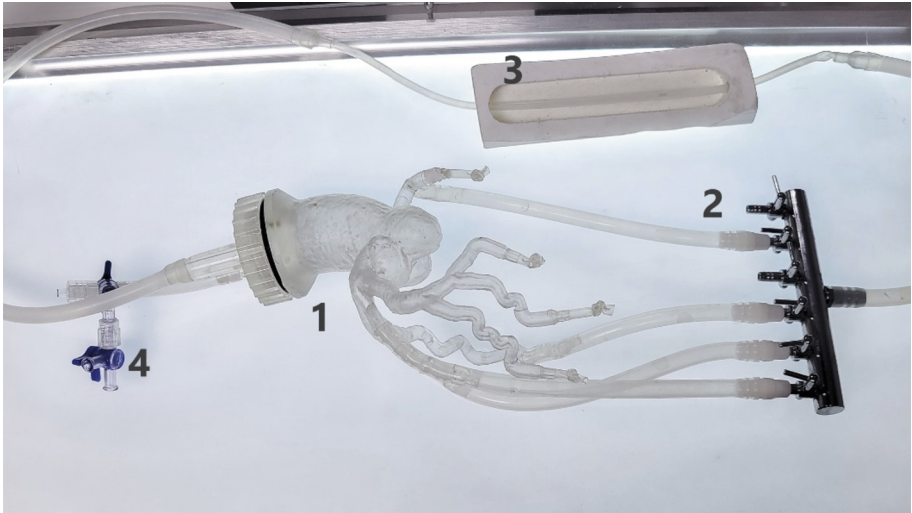


Fig. 2. Detailed representation of the comprehensive hydraulic system. The segmented anatomical model of the aortic artery, accurately segmented beforehand. 4-way return configuration for fluid flow. Internal view of the subsystem dedicated to radial puncture simulation. Includes arterial pulse simulation and a portion of radial bone for reference. 2-way system arrangement, centered on air removal within the model.

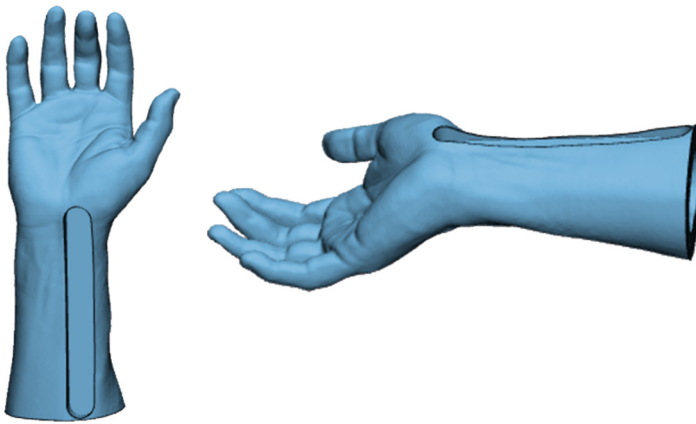


Fig. 3. System for performing radial puncture.

The cephalocaudal movement is based on a mechanical transmission and a Nema 23 motor, adapted to move an arc where the camera is situated in a range of 0° to 90° (Fig. 6). A mechanical transmission was incorporated, which also serves to maintain the position of the arc even when the system is not powered by a power source.

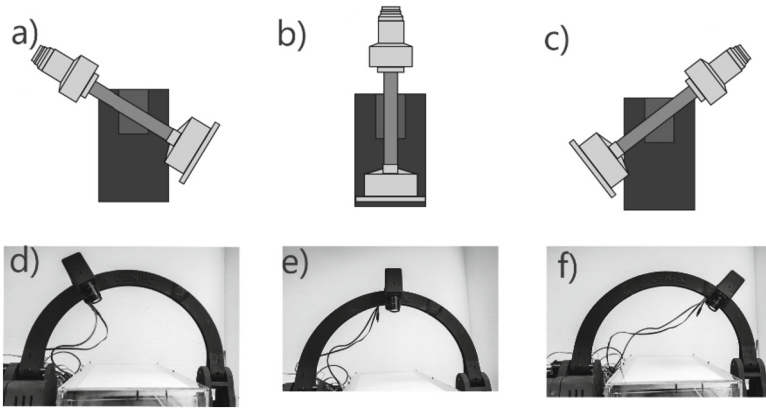


Fig. 4. The top section depicts the oblique axis of the fluoroscope, while the bottom section shows the oblique axis of the simulator. In references a) and d), the positive side is observed; in references b) and e), the neutral position is shown; and finally, in references c) and f), the negative side is observed.

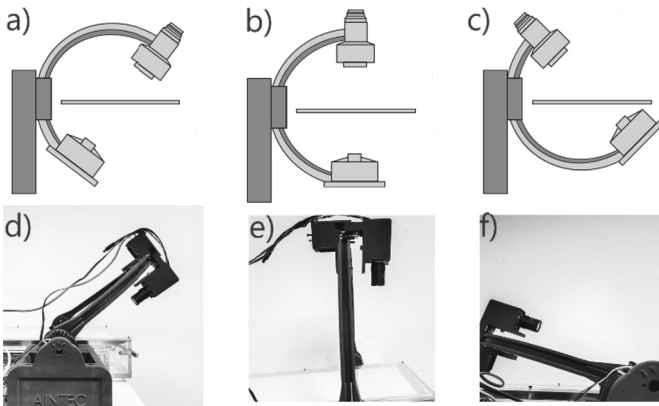


Fig. 5. The top section depicts the cephalocaudal axis of the fluoroscope, while the bottom section shows the cephalocaudal axis of the simulator. In references a) and d), the caudal side is observed; in references b) and e), the neutral position is shown; and finally, in references c) and f), the cranial direction is observed.

The oblique movement comprises a Nema 17 motor and a camera, anchored to a housing. With the assistance of the motor and a toothed belt, it's possible to move the camera along the arc, providing left and right oblique positions (Fig. 7).

The bed's movement is designed using a 12V linear actuator at a speed of 10 mm/s. It's mounted on an acrylic base, and above it, a box holds the lamp and features a rail system to allow smoother movement.

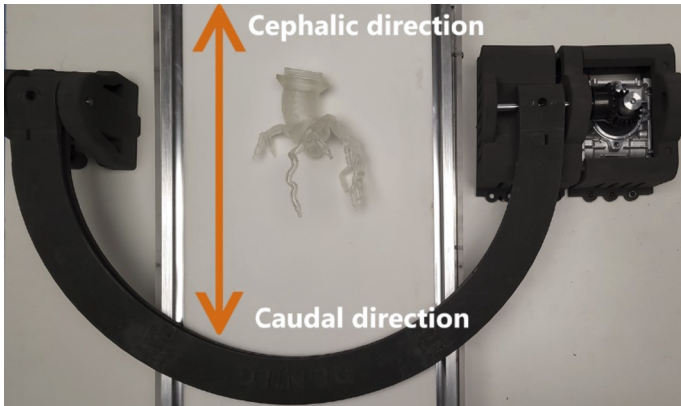


Fig. 6. Cephalocaudal Axis. On the right side, you can observe the transmission mechanism responsible for supplying the required torque for mobility and stabilization of the arc.

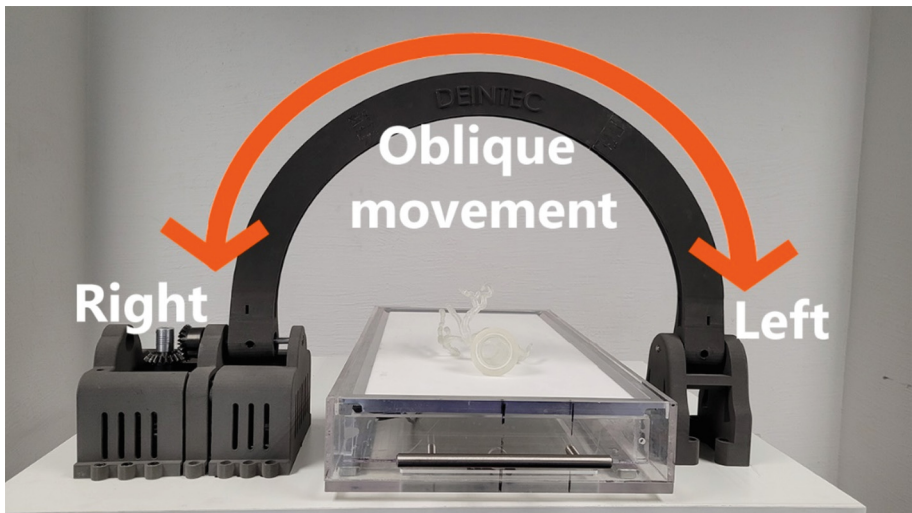


Fig. 7. Oblique Axis. The camera will move along the perimeter of the arc to acquire different viewing angles. The arc's configuration is designed to maintain a constant focal point.

2.4 Mechanical and Digital System

Graphical Interface. The interface was designed using Visual Studio Community (Microsoft Corporation. <https://visualstudio.microsoft.com/es/vs/community/>) and incorporates image acquisition and motion control features (Fig. 8).

The interface provides predefined positions, including right cranial oblique (-30° , 30°), right caudal oblique (-30° , -30°), right neutral oblique (-30° , 0°), left cranial oblique (45° , 30°), left caudal oblique (45° , -30°), left neutral oblique (45° , 0°), among others. However, in case, at some point during the simulation, the specialist or resident

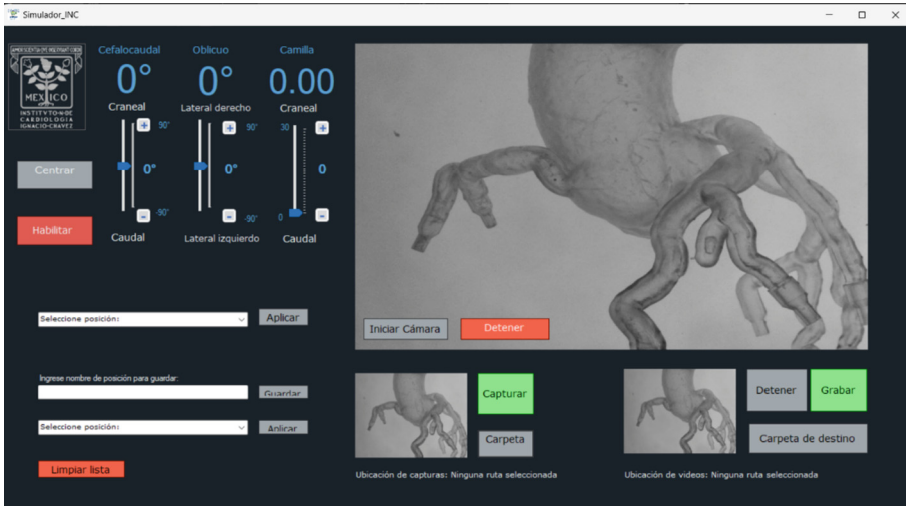


Fig. 8. Perspective of the interface in operation. The three axes of movement are controllable, where the “Center” button resets the three axes to their corresponding neutral coordinates. The “Enable” option authorizes system control directly from the interface. The first line manages predefined positions, while the second line enables the preservation of the current position with its respective identifier. The third line selects a previously established position. On the right-hand side, the camera display is present, with a button to activate and stop the display. Below this section, controls for image and video capture and storage in a predefined folder are situated.

identifies a specific useful position, there is an option to save it for accessing it when needed.

Image Acquisition. With the aim of obtaining images that emulate those acquired through fluoroscopy, the image acquisition technique in grayscale with backlight support, known as “Backlight,” was implemented. This strategy involves placing the object of interest between the camera and a light source, allowing the visualization of its silhouette through the generation of contrasts. The resulting signal is presented on the graphical interface in the form of a grayscale image (Fig. 9).

3 Results

The environment has been meticulously designed to replicate the conditions of a trans-catheter procedure. During the simulation, the medical professional can perform a radial puncture guided by the cardiac pulse. They can then insert an introducer, guide a catheter through a real anatomical model assisted by the fluoroscopy-mimicking system, and position the area of interest for the release of trans-catheter medical devices. This process is achieved without exposure to ionizing radiation, is repeatable in a brief period, and is adaptable to any clinical case obtained through imaging techniques such as CT, MRI, or echocardiography.

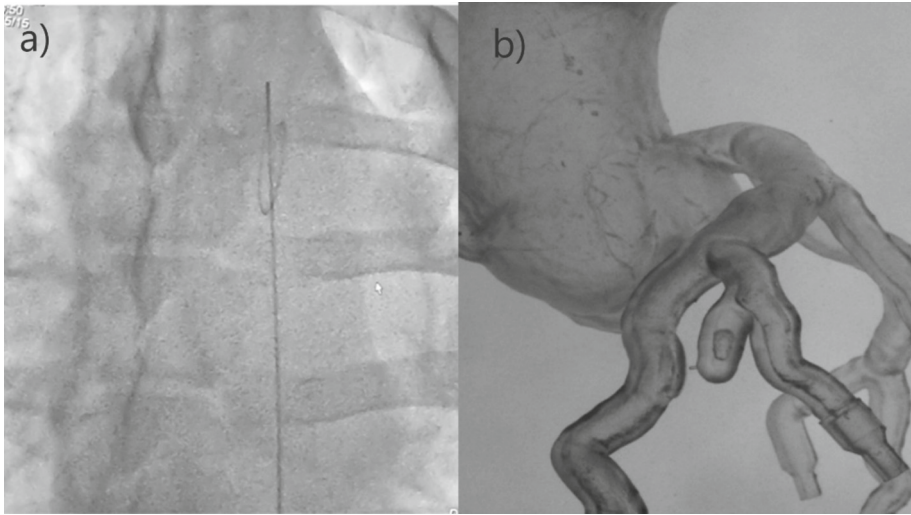


Fig. 9. Comparison between fluoroscopy (a) and its equivalent in the simulator (b). The representation on the left displays the insertion of a guide in a porcine model.

4 Discussion

The cardiovascular procedure simulation system introduces innovative approaches, such as precise segmentation of anatomical models and the implementation of 3D printing techniques. Additionally, it integrates a fluoroscopy system along with radial puncture technique. It is worth noting that being a design developed by the INC, this system stands out for its reduced cost compared to commercially available simulators with similar features. Furthermore, it presents unique attributes, such as the ability to adapt to any clinical case and the potential to adjust based on current needs. These qualities make the simulator a highly suitable tool for training physicians in the field of radial puncture and trans-catheter procedures.

5 Study Limitations

Achieving a complete replication of an environment identical to that of a cardiovascular procedure is a task of great complexity. The proposed system represents an approximation to the procedure's conditions; however, there are opportunities for further improvements. These enhancements could include the use of a blood-analog fluid, precise measurement of pressure and flow velocity, the establishment of a contrast medium injection system, utilization of more transparent and flexible materials for better visualization, among other potential areas of improvement.

6 Conclusions

This work highlights the advantages of incorporating physical simulators for training and educating professionals. The flexibility to encompass a wide variety of clinical cases and the ability to adapt the system to changing needs in the medical field make it a highly valuable tool for training in trans-catheter procedures and radial puncture.


This system marks a step forward in medical education by allowing medical professionals to practice procedures in a controlled and realistic environment. While there is still a path to perfection in cardiovascular simulation, current and future advancements in this field will undoubtedly contribute to the improvement of training and healthcare in the realm of cardiovascular procedures.

References

1. Waran, V., et al.: Injecting realism in surgical training – initial simulation experience with custom 3D models. *J. Surg. Educ.* **71**(2), 193–197 (2014). <https://doi.org/10.1016/J.JSURG.2013.08.010>
2. Zimmermann, J.M., et al.: Novel augmented physical simulator for the training of transcatheter cardiovascular interventions. *Catheter. Cardiovasc. Interv.* **95**(6), 1202–1209 (2020). <https://doi.org/10.1002/CCD.28493>
3. Ghaednia, H., et al.: Augmented and virtual reality in spine surgery, current applications and future potentials. *Spine J.* **21**(10), 1617–1625 (2021). <https://doi.org/10.1016/j.spinee.2021.03.018>
4. Wang, D.D., et al.: 3D printing, computational modeling, and artificial intelligence for structural heart disease. *JACC* **14**(1), 41–60 (2021). <https://doi.org/10.1016/J.JCMG.2019.12.022>
5. Aimar, A., Palermo, A., Innocenti, B.: The role of 3D printing in medical applications: a state of the art. *J. Healthc. Eng.* **21**, 5340616 (2019). <https://doi.org/10.1155/2019/5340616>
6. Wahood, W., Ghozy, S., Al-Abdulghani, A., Kallmes, D.F.: Radial artery diameter: a comprehensive systematic review of anatomy. *J. Neurointerv. Surg.* **14**(12), 1274–1278 (2022). <https://doi.org/10.1136/NEURINTSURG-2021-018534>



Design and Characterization of a Microfluidic Biological System for Bone Tissue

Odin Ramirez-Fernandez^{1,2}(✉) , Fabian Equihua-Guillen³, Adrian Garcia-Lara³, and Esmeralda Zuñiga-Aguilar⁴

¹ Departamento de Ingeniería Biomédica, Universidad Politécnica de Cuautitlán Izcalli, 54720 Cuautitlán Izcalli, Mexico

odinramirezfernandez@gmail.com

² Departamento de Investigación, Universidad Tecnológica de Mexico-UNITEC Mexico-Campus en Línea, Col. Anáhuac, 11320 Ciudad de México, Mexico

³ Facultad de Ingeniería Mecánica y Eléctrica, Universidad Autónoma de Coahuila, FIME, 25720 Monclova, Mexico

⁴ Departamento de Ingeniería Biomédica, Universidad Autónoma de Ciudad Juárez Chihuahua, 32315 Ciudad Juárez, Mexico

Abstract. Tissue engineering has managed to revolutionize the transplantation and regenerative medicine, it is responsible for developing biomaterials to generate a promising approach for studying complex physiological processes in vitro. The design and application of biological systems evaluate a microfluidic platform based on organs physiology, modeling the nutrient distribution, and testing potential implants.

The present investigation aims to design and characterize a biological system to simulate the microfluidic environment of the bone tissue. The system must provide the conditions for adequate cell regulation, to achieve this, adhesion, migration, proliferation, and differentiation are used, as well as the adequate delivery of bioactive factors such as growth and adhesion.

A novel hydraulic circuit with radial flow and their components were designed in SolidWorks, then the pieces were calculated with the Navier-Stokes equations using ANSYS and COMSOL software, to have a laminar environment and their behavior using microfluidics and cells mimicking the bone structure. The components were generated by 3D printing and the additive stereolithography resin technique.

According to the CFD simulations, it was found that the system had a media flow of 18.56 nl/s with the smallest pressure of 146.32 mPa our chamber is the optimum model for the bone cells.

Our microfluidic system's design shows the flow change can be used to follow the bone anatomy, which indicates an appropriate irrigation of nutrients for the cells. The proposed biological system has confirmed to be an adequate model for the bone tissue with a continuous irrigation of the media.

Keywords: biomaterials · bone tissue · tissue engineering · biological system · microfluidic simulation

1 Introduction

Bone tissue engineering seeks to develop innovative approaches to regenerate and repair bone defects. The biological system technology provides a platform to mimic the complex microenvironment and microfluidic of bone tissue and study its cellular interactions and physiological processes in function of the structural and functional aspects of the bone tissue. The key components such as microchannels, cell culture chambers, and perfusion systems are discussed, emphasizing their role in creating dynamic and biomimetic bone tissue models [1].

The aging of the population leads to an increased incidence of bone fractures due to osteoporosis. Understanding the mechanisms of bone physiology is essential to design advanced treatments. Bone is a dynamic tissue in continuous formation and resorption through coordinated communication between osteoblasts (Ob) and osteoclasts (Oc). Mechanical loads modulate bone architecture and cell physiology and play an important role in bone tissue homeostasis. Although many *in vitro* models of Ob/Oc cultures exist for material analysis, little is known about cell communication under mechanical stimulation in this system [2].

Around 115 million animals were used in experiments around the world last year, with most tests being carried out without anesthesia or painkillers, and the toxic substances cause long-term pain, according to the report. European Coalition to End Animal Experiments (ECEAE). Mexico spends 10 million on animals for experimentation. Worldwide, more than 500,000 animals suffer and die each year because of laboratory tests [3].

The development and characteristics of biological models, explores the integration of relevant cell types, such as Ob, Oc, and mesenchymal stem cells, within the designed platforms, to avoid the animal experimentation in every experiment. The platform includes the incorporation of mechanical properties, such as fluid flow, shear stress and strain, and the importance of extracellular matrix components in creating physiologically relevant bone tissue microenvironments [4].

The strategies employed in design biological models' considerate aspects of the bone tissue microenvironment. This includes the integration of vascular networks, the introduction of biochemical gradients, and the simulation of bone remodeling processes. It discusses the ability of the systems to mimic bone-related pathologies, such as osteoporosis, bone metastasis, and bone infections. The use of this platforms for high-throughput drug screening, assessment of drug efficacy, and personalized medicine approaches is also addressed [5].

The challenges associated with bone tissue models, including the need for improved complexity and scalability, the incorporation of other tissue interfaces like muscle-bone interfaces, and the integration of patient-specific cells. It also discusses potential future directions, such as combining biological system models with advanced imaging techniques, multi-organ systems, and the use of bioprinting for 3D printing of bone tissues [6].

For the design of the Bioreactor, the COMSOL Multiphysics software was used based on the Solidworks design and the ANSYS mesh, The combination of those tools allowed us to solve the physics problems necessary to carry out the fluid simulation and control. In the simulation process it is considered the geometry of the Bioreactor, the fluid

physics with their respective boundary conditions and domains, define the simulation parameters and define and simulate the control strategy of the device [7].

This research presents a design of bone tissue biological system. The models were designed using SolidWorks, and hydrodynamic simulations were done in ANSYS and COMSOL afterward. The flow behavior in the irrigation system's channels were studied to determine which configuration of pressure and stress were the optimum values to imitate the anatomy and irrigation of bone microenvironment.

2 Methodology

2.1 Biological System Design

The blood flow in bone tissue, circulates through a closed cavity where the pressure must stay constant in the vessels and above the veins. This circulation allows the interchange of nutrients and minerals among the blood and bone tissue. In other hand, the capillaries within bone structure have the same anatomic assembly as those located elsewhere in the body [7].

We design a platform based on bone circulation with two tubes of 2 mm of diameter for the input and output media, the base is an octagon of 30 mm side and 3 mm wide, the circular base is 25 mm long and 3 mm wide. The base of the test system is circular because it refers to an internal part of a bone, the two external parts are for introducing the cells and in the same way they can be introduced by lifting the lid. The in vitro test system was designed in an octagon shape with 1 tube at each end using SolidWorks CAD software to observe cell flow. It has 1 removable lid that prevents the cells from coming out as shown in Fig. 1), inside it has a tube with holes that allow the cells to flow.

2.2 System Flow Simulation

The objectives of the flow simulation were to see in which zones of zero flow within the in vitro test system, the assembled system was exported to the ANSYS software to obtain the optimum mesh for the model and then it was used in COMSOL Multiphysics to resolve the Navier-Stokes equations to determine the flow fields, speed, and tension alongside bone tissue.

The simulation carried out allowed us to study the behavior of the fluid, such as the pressure profiles and velocity in the chamber, with these values it was possible to define a qualitative range of stresses present in the chamber for different flow values. With the variable control of flow and stress, it was possible to establish a range in which the simulation converges. For the validation of the Bioreactor simulation, it was proposed to design an object with specific characteristics and like that of a human bone tissue.

Once imported the model we can generate their mesh, a volume to represent the fluid in the channels had to be created. This volumetric body meshed with hexagonal hexahedral volumetric elements. The elements were refined at the channels. Working pressure at 1 atm, and a viscous-laminar model to describe the fluid as Newtonian, therefore, viscosity is considered constant. The working fluid was set with a characteristic

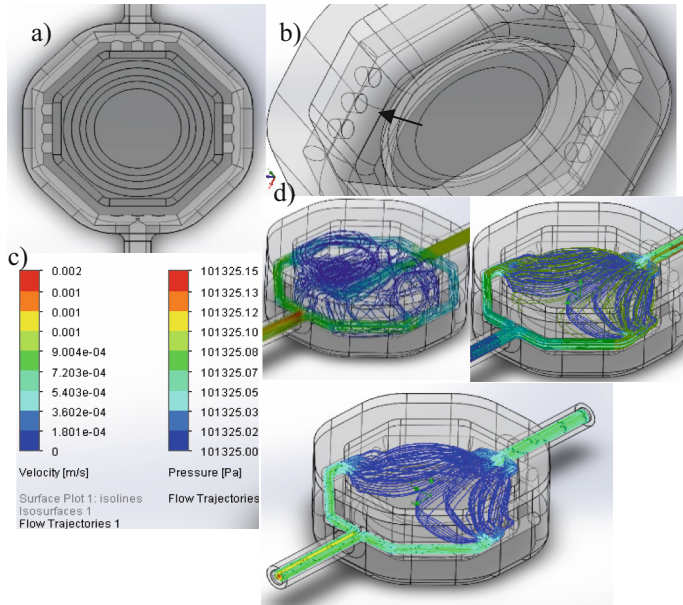


Fig. 1. Components of the microfluidic biological system. a) Top view of the system. b) The pipes and holes of the in vitro test system. c) The top view of the pipes of the in vitro test system. d) preliminary Flow simulation with inlet velocity of 8.3×10^{-5} m/s, value referred to a peristaltic pump; and outlet condition at 1 atm.

$Pr = 4.34$. The boundary conditions were set as: inlet velocity 8.3×10^{-5} m/s. Solver algorithm was set to *Simple*, with a *Green-Gauss Cell* gradient, and residuals were set to 1×10^{-06} . Set was started at the inlet, and 500 iterations as shown in Fig. 2.

2.3 Platform 3D Print

The parts of the biological system were 3D printed on a photolithography additive printer (Crealty LR-002) using a photosensitive resin to generate the complex figures in one piece to fabricate the final test. To complete this, Creality® Chitubox V2.2® slicing software was used to print the resin parts with the following parameters used for resin printing:

Layer Height 0.4 mm Bottom/Top, Thickness 0.2 mm Print Speed (mm/s) 6 s, Print Temp 27 °C, Fan Enable, Cut Bottom of Object 0.0 mm, Overlap 0.15 mm, supports along the structure 5 mm.

3 Results

The results of the flow simulation are shown in Fig. 3. a) at the entrance and exit is constant 9×10^{-4} m/s. b) the simulation shows that the flow decreases as it passes through the pipes. c) top view of the simulation and it is observed how the flow changes

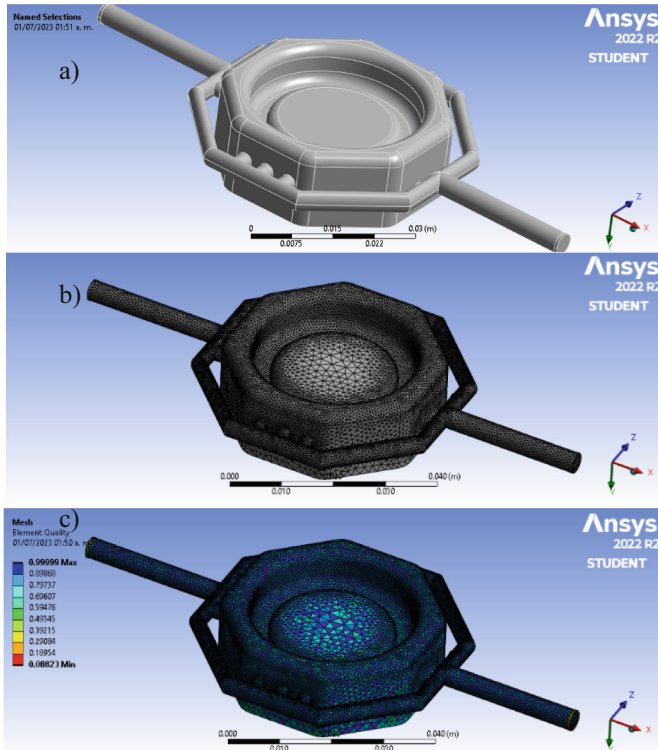


Fig. 2. Components of the microfluidic biological system meshed, Fluid volume: $1.009169 \times 10^{-08} \text{ m}^3$, Solid volume: $2.716106 \times 10^{-06} \text{ m}^3$. a) Volume of the system. b) Generic mesh. c) Element quality and functional cells.

color. The spiral lines are continuous colors, without turbulent flow, it is observed as is, ensuring that the cells will not have shear stress. The simulation values were Fluid volume: $1.009169 \times 10^{-06} \text{ m}^3$, Solid volume: $2.716106 \times 10^{-06} \text{ m}^3$, Total cells: 11,913, fluid cells: 11,913, Fluid cells in contact with solids: 7,132.

3.1 Printed Model

Finally, the images of the printed model are shown in Fig. 4. This prototype could be probed with a colored liquid to demonstrate the computational results. The system with the bone cell culture will be on a CO_2 incubator to control the gases temperature and measure the density of living cells and biomass, it allows to monitor critical parameters continuously and automatically to the changing conditions of the process.

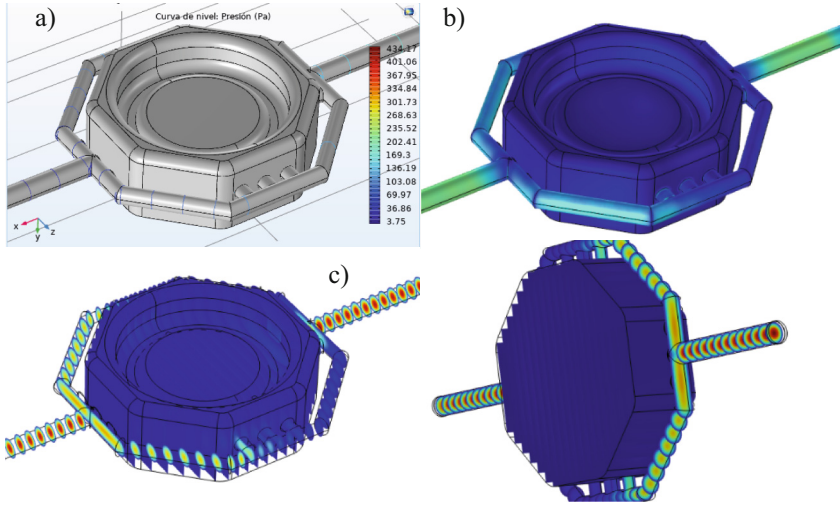


Fig. 3. Navier-Stokes COMSOL solver of the microfluidic biological system. Flow simulation components. a) System Volume with conditions. b) Chamber flow fields and their distribution. c) Vectors fields through the support in the form of the chamber.

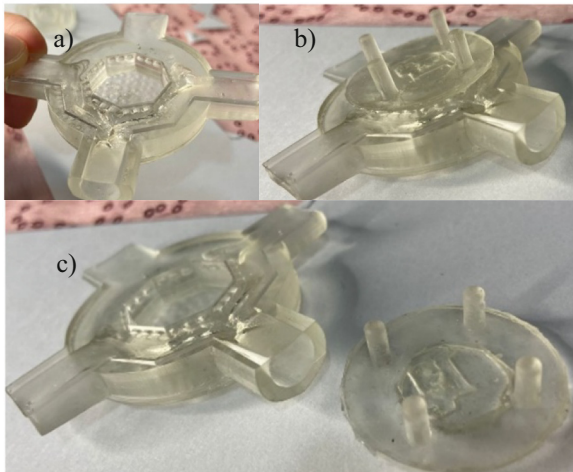


Fig. 4. Final prototype printed model of the in vitro test system. The clear resin was selected to facilitate the visual inspection of the flow and culture.

4 Conclusion

The biological system was designed and manufactured, it will provide a constant laminar flow, since in the flow lines it can be observed that there were no regions with zero flow, since a good fluid exchange was carried out, and the cells will not be affected.

The microfluidic technology holds immense potential for advancing bone tissue engineering and regenerative medicine. The ability to create biomimetic bone tissue models in a controlled and dynamic microenvironment allows for better understanding of bone biology, disease mechanisms, and drug responses. With continued advancements and interdisciplinary collaborations, bone tissue models have the potential to revolutionize personalized medicine, accelerate drug discovery, and improve patient outcomes in the field of bone health.

Acknowledgement. The authors would like to express their deepest gratitude to the CONHACYT and the SNII program, for the technical and financial support for this study and to the Project PAPIIT IT100222.

References

1. Botting, J.H., Morrison, A.R.: Animal experimentation, essential for medicine. *Res. Sci.* 69–72 (1997)
2. Badylak, S.F., Taylor, D., Uygun, K.: Whole-organ tissue engineering: decellularization and recellularization of three-dimensional matrix scaffolds. *Annu. Rev. Biomed. Eng.* **13**, 27–53 (2011)
3. Rosa, V.: What and where are the stem cells for Dentistry? *Singapore Dent. J.* **34**(1), 13–18 (2013)
4. Silva, M.: From fiction to reality: organs-on-chips in service of science and medicine. *Mexican Dent. Mag.* **20**(2), P74–P76 (2016)
5. Busch, M., Bredeck, G., Kämpfer, A.A.M., Schins, R.P.F.: Investigations of acute effects of polystyrene and polyvinyl chloride micro- and nanoplastics in an advanced in vitro triple culture model of the healthy and inflamed intestine. *Environ. Res.* **193**, 110536 (2021). <https://doi.org/10.1016/j.envres.2020.110536>
6. Probst, C., Schneider, S., Loskill, P.: High-throughput organ-on-a-chip systems: current status and remaining challenges'. *Curr. Opin. Biomed. Eng.* **1**(6), 33–41 (2018)
7. Laroche, M.: Intraosseous circulation from physiology to disease. *Joint Bone Spine* **69**(3), 262–269 (2002)



Tumor Tissue Classification in Hyperspectral Histopathology Images Through Individual and Ensemble of Machine Learning Algorithms

Daia Natalia Martínez-Arboleyda¹(✉), Inés Alejandro Cruz-Guerrero¹,
and Daniel Ulises Campos-Delgado^{1,2}

¹ Facultad de Ciencias, Universidad Autónoma de San Luis Potosí, 78295 San Luis Potosí, San Luis Potosí, Mexico

a282863@alumnos.uaslp.mx

² Instituto de Investigación en Comunicación Óptica, Universidad Autónoma de San Luis Potosí, 78290 San Luis Potosí, Mexico

Abstract. Hyperspectral imaging (HSI) is a versatile modality that can provide a noninvasive tool for cancer detection, image-guided surgery, tissue classification, among other applications. In this work, we demonstrate the integration of hyperspectral imaging, spectral unmixing, and machine learning models to accurately classify tumor and non-tumor tissue from histopathology samples. The studied database contains 494 images from 13 different patients diagnosed with glioblastoma. Our approach is based on identifying characteristic spectral signatures for each hyperspectral image by spectral unmixing, and using them as an input feature vector for machine learning models: support vector machine, random forest, and a voting ensemble. The resulting average accuracy in our evaluation of four folds was 89.4% that improves the reference value of 85.5%, which was the best performance in the state-of-the-art. In this way, our proposed methodology shows promising results to assist in pathological analysis and provide support to healthcare professionals.

Keywords: Hyperspectral imaging · Machine Learning · Binary Classification · Glioblastoma · Spectral Unmixing

1 Introduction

In histopathology, digital imaging is a powerful tool for achieving a systematic evaluation of the studied samples [2]. Nonetheless, different factors affect the evaluations, such as sample preparation and pathologist expertise. In this sense, machine learning has recently been suggested in medical applications to avoid bias in diagnosis [4], and in particular for digital histopathology samples [5]. Unlike traditional images, which only work with three color channels (Red, Green, and Blue - RGB), hyperspectral (HS) images provide much more detailed spectral information about the object and its composition. Hyperspectral imaging (HSI) is an emerging imaging modality that provides a noninvasive tool in medical applications: cancer detection, image-guided surgery, tissue

classifications, etc. HSI is a hybrid modality that combines imaging and spectroscopy, by collecting spectral information at each pixel of a two-dimensional detector array, generates a three-dimensional dataset of spatial and spectral information, known as hypercube. In this way, each pixel in the HS image has a unique spectral signature, which refers to the pattern or distribution of light intensity across different wavelengths. Through analyzing these spectral signatures, these images allow the detection and identification of various pathological conditions [1]. HSI can be incorporated to improve diagnosis, as the intensity of the light transmitted through the sample could be analyzed beyond the visible spectrum. Some examples of this trend include the application of deep learning for the diagnosis of breast cancer cells [6], the use of machine learning for the automatic recognition of colon and esophagogastric cancer [7], and the implementation of an ensemble of deep neural networks to identify cancer samples in brain tissue [13].

In this work, we focus on utilizing spectral unmixing and machine learning models (support vector machine with polynomial and radial basis function kernels, random forest, and a voting ensemble) to classify HS images of histopathology samples of brain tissue. A supervised classification approach was employed, where a pathologist previously labeled the samples in the HSI database as non-tumor and tumor. The database was organized into four folds to address the variance in the dataset during the learning stage due to the imbalance in the classes. The input feature vectors for the machine learning models were obtained by the concatenation of the estimated endmembers from the spectral unmixing stage. The results obtained were compared with those produced in [13], where the same database was considered, along with an ensemble of deep neural networks for classification. Our results show an improvement in the average accuracy with respect to the state-of-the-art method (89.4% vs. 85.5%).

2 Materials and Methods

The proposed methodology in this work consists of six steps: (i) HSI database, (ii) pre-processing, (iii) spectral unmixing, (iv) features extraction, (v) training and testing, and (vi) results, as illustrated in Fig. 1. In this section, we describe each step highlighting their collective contribution to the overall research objective.

2.1 HSI Database

This research work used a database consisting of human biopsies obtained during brain tumor resection procedures. The study protocol and consent procedures were approved by the “Comité Ético de Investigación Clínica - Comité de Ética en la Investigación” (CEIC/CEI) of “Hospital Universitario de Gran Canaria Doctor Negrin” [8]. The HS camera was configured with a magnification.

20 \times and captured images with a size 1004 \times 800 pixels. Further details about the dataset can be found in [8]. After the resection, the samples were diagnosed as glioblastoma (GB) based on the World Health Organization (WHO) classification of nervous system tumors by a pathologist. Regions of interest (ROI) were selected for subsequent analysis, using the mentioned instrumentation [8], resulting on 494 HS images from 13 different patients (P1, P2, ..., P13) with GB. An example of tumor and non-tumor tissue

samples from the database are illustrated in Fig. 1.(i). In fact, not all patients presented simultaneously tumor and non-tumor tissue samples. In Fig. 2, the distribution of tumor and non-tumor HS images is presented. In this way, patients P9 to P13 only contain tumor tissue samples.

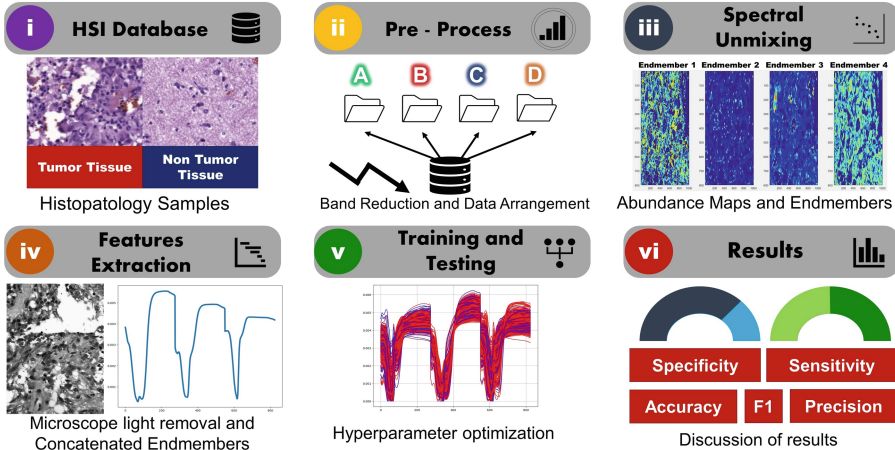


Fig. 1. Diagram of the proposed methodology to classify hyperspectral histopathology images.

2.2 Pre-processing Stage

A pre-processing of the database was performed to adjust the spectral responses of the HS images, and reduce noise and redundant information. The HS images covered a spectral range of 400 to 1000 nm with a resolution of 2.8 nm. Thus, the original HS images had a spatial size of 1004×800 pixels with 826 spectral channels. In [8], a reduction of the spectral bands was proposed to remove redundant information by their high correlation, which led to a decrease from 826 to 275 channels.

As mentioned earlier, not all patients in the database presented both types of classes (non-tumor and tumor), nor they have the same number of HS images for each class (as explained in detail in [8]). This imbalance in class representation, along with the limited number of patients (just 13 different), presented a challenge for the classification scheme. To address these issues and mitigate the potential bias and overfitting during the learning stage, the database was organized into four distinct folds (A, B, C, and D) for cross-validation, with subsets for training, validation, and test. Folds, A, B, and C included nine patients for training, one for validation, and three for testing. Meanwhile, Fold D contained eight patients for training, one for validation, and four for testing. These folds are illustrated in Fig. 3, with different colors (red, green, and blue) indicating each subset. All the folds were randomized to ensure that each sample was presented once in the testing subset, and that patients with both types of labeled data (non-tumor and tumor) were selected for the validation subset. With this four folds approach, the database is presented in four different ways to the ML algorithms for cross-validation.

This perspective follows previous works in [8] and [9], which served as guides for this study.

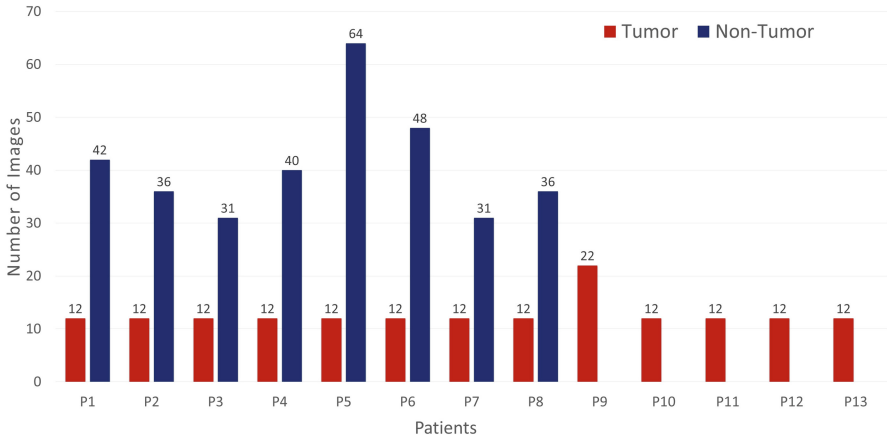


Fig. 2. Distributions of tumor and non-tumor HS images acquired from 13 patients (P1, P2,...,P13).

2.3 Spectral Unmixing

In the field of HS imaging, spectral unmixing (SU) plays a crucial role in analyzing and extracting information about the fundamental materials present in a scene or image. By identifying specific spectral signatures known as endmembers, which represent the fundamental materials, the unmixing technique uses them as reference points. This process enables the computation of concentrations or abundances of each component at pixel level. The resulting abundance maps, like the ones depicted in Fig. 1. (iii), provide valuable insights into the spatial distribution of materials within the image. This technique is a valuable tool to understand the composition and characteristics of the scene or image captured through HSI [10]. Hence, the identification of the fundamental signatures in the HS image by spectral unmixing allows the implementation of machine learning models to perform classification tasks. In addition, the computational cost of the classification task can be reduced through the application of spectral unmixing.

In this study, the SU stage was carried out by using the EBEAE algorithm [10]. Other unmixing algorithms such as sparse non-negative matrix factorization, non-smooth non-negative matrix factorization, or their variants could also be applied [11], but EBEAE was chosen due to its robustness to different types and levels of noise and its low complexity. EBEAE has the ability to estimate the abundances of individual endmembers by minimizing the approximation error and maximizing their entropy. In addition, EBEAE normalizes the cost function to mitigate the dependence on the size of the database. The EBEAE hyperparameters were determined according to the recommendations provided in [10], and they were set to, $\rho = 1$ (*similarity weight*), $\mu = 0$ (*entropy weight*), error threshold $\epsilon = 1 \cdot 10^{-6}$, and a maximum number of 10 iterations (*maxiter*). The linear unmixing process assumed four characteristic components, that is, $N = 4$, as in [13].

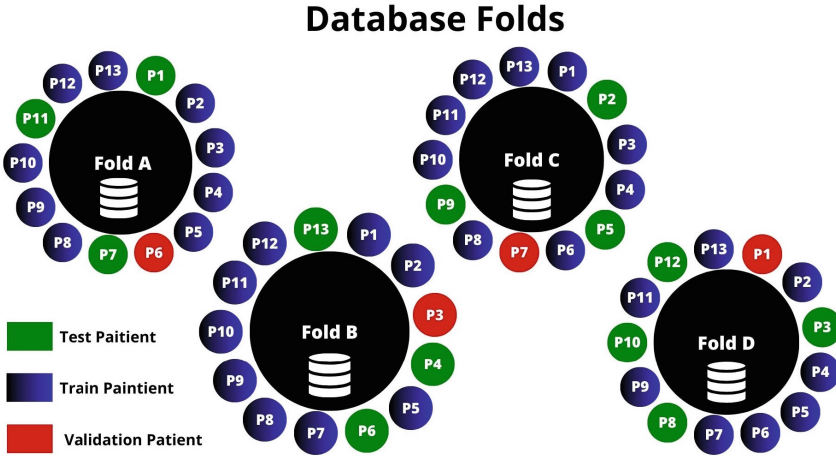


Fig. 3. Database folds for the 13 patients (P1, P2,...,P13) with subsets for training, validation, and testing.

2.4 Features Extraction

In the histopathology samples, there are areas without brain tissue where the microscope light can transmit directly. These areas without tissue do not contain information and need to be eliminated to avoid bias in the classification process. To address this issue, the first step is to identify the endmember with a flat spectral signature that corresponds to no-tissue regions, where the microscope light is transmitted directly. Next, this endmember must be eliminated since it does not provide characteristic information of the sample. This phenomenon can be easily identified due to the sum-to-one restriction applied by EBEAE [10], which results in a spectral signature with uniform information across most spectral bands. In Fig. 4.(i), the green-colored endmember, denoted as endmember 3, corresponds to the microscope light (roughly flat spectral signature). Another challenge is that in EBEAE, the order of the resulting endmembers could vary for each HS image. Therefore, a mathematical method needs to be devised to identify this flat endmember. For this goal, the Euclidean distance between each resulting endmember and a flat spectral signature is computed by recalling that there are 275 spectral bands:

$$E_i = \left\| p_i - \frac{1}{275} \mathbf{1} \right\| \quad \forall i \in \{1, 2, 3, 4\} \quad (1)$$

where p_i represents the i -th estimated endmembers by EBEAE, $\mathbf{1}$ is a vector with just unitary entries, and $\|\cdot\|$ denotes the Euclidean norm. Consequently, for each HS image, we eliminate the endmember with the smallest value E_i .

Next, the remaining three endmembers were organized in a specific order to define an overall feature vector for each HS image. To achieve this, we selected a HS image as a reference in the database and utilized the order of endmembers in this image as a guide to arrange the remaining ones. In this stage, we used again the Euclidean distance as a similarity index among endmembers. At the end of this stage, the arranged and

concatenated endmembers (as visualized in Fig. 4.(ii)) were used as a single feature vector per HS image. As a final step, the feature vectors were randomly mixed in each set of the different folds to ensure that the processes were not biased towards any specific order.

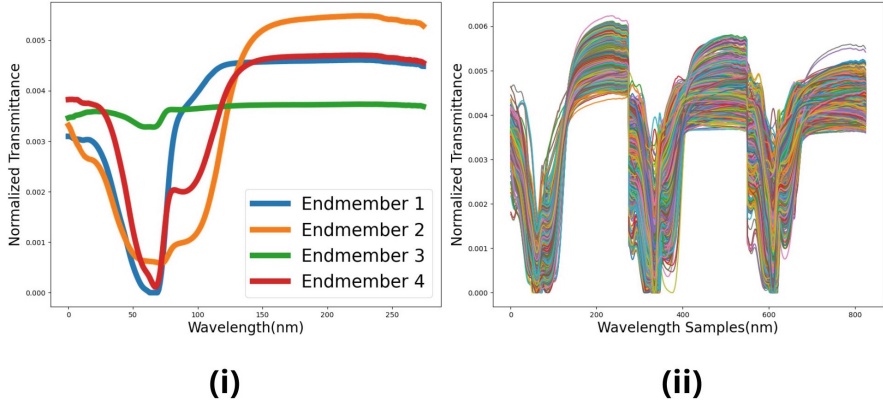


Fig. 4. Features extraction step: (i) Example of the estimated endmembers; and (ii) features vectors by concatenating the endmembers with the same order.

2.5 Training and Testing

In this study, a supervised approach is employed for the classification task, since each HS in the database is already labeled as tumor or non-tumor (binary problem) by the pathologist [8]. Machine learning models have varying strengths and weaknesses, and their suitability depends on factors, such as the source of the data, the complexity of the problem, and the specific requirements of the task [12]. For instance, certain classifiers may excel in handling linearly separable data, while others may be more effective for non-linear relationships or high-dimensional data. Now, by looking at the endmembers and input feature vectors in Fig. 4, there is some similarity in the input data for classification and the vectors are high-dimensional, so a complex nonlinear problem could be expected. As a result, we focus on two well-known strategies to deal with this complex scenario: Support Vector Machine (SVM) and Random Forest (RF). In the following, we present a short description of these classification models:

- SVM is a powerful, versatile, and popular machine learning model. SVMs are particularly well suited for the classification of complex small or medium-sized datasets [12]. For nonlinear classification problems, the selection of the kernel can help to adjust the decision boundaries, for example, polynomial and radial basis functions. Each of these kernels has hyperparameters that can be adjusted to fit the training dataset.
- RF is an ensemble of decision trees, generally trained with the bagging method and based on the wisdom of the crowds paradigm [12]. Therefore, the RF classifier is also

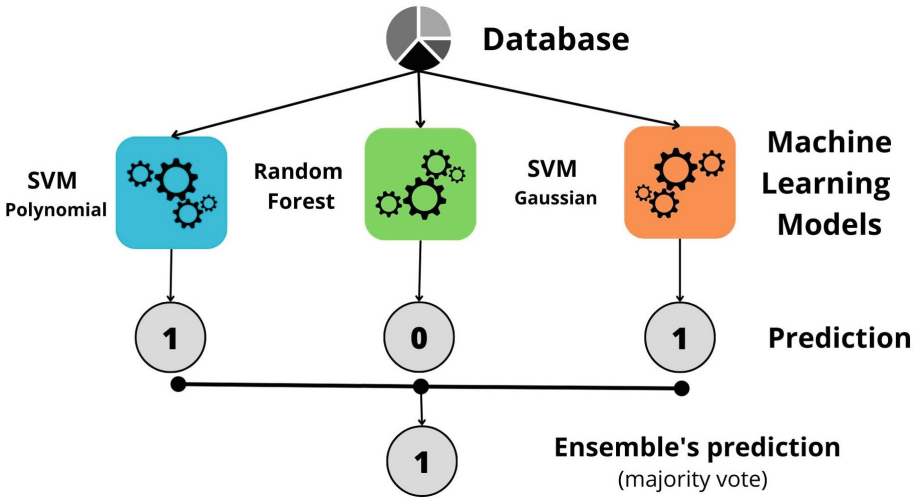


Fig. 5. Structure of the voting ensemble.

a popular machine learning model that can capture complex relationships between features and targets.

Three classification models were initially considered: SVM with a polynomial kernel (SVM-Polynomial), SVM with a Gaussian radial basis function kernel (SVM-Gaussian), and RF. Furthermore, to utilize the specific properties of each classifier and enhance the decision's robustness, an ensemble of them was also explored, resulting in the consideration of a voting methodology. In this methodology, the prediction is based on the class that receives most votes [12]. Typically, the voting ensemble outperforms the best classifier within the ensemble in terms of accuracy. The structure of the voting ensemble is depicted in Fig. 5. To have the best hyperparameters, we used a Bayesian Optimizer in this tuning process, and the final values are listed in Table 1. Two principal hyperparameters are associated with SVM-Polynomial: the degree of the polynomial kernel function (*PolynomialDegree*), and the regularization weight C . The degree of the kernel establishes a compromise between model complexity and computational time, and the regularization weight C shows an inverse relationship with the margin size of the decision regions and must be strictly positive. In the case of SVM-Gaussian, there are two hyperparameters: scaling coefficient of the radial basis function (*Gamma*), and once more, the regularization weight C . In this case, *Gamma* is related to the influence radius of the support vectors. Finally, for RF, there are three hyperparameters: the number of estimators (*Estimators*), the minimum samples split (*MinSamplesSplit*), and maximum depth (*MaxDepth*). The number of estimators corresponds to the number of decision trees in the forest, while *MinSamplesSplit* denotes the minimum number of samples required to split an internal node. Lastly, *MaxDepth* indicates the maximum depth of each decision tree [12].

3 Results

In this section, we discuss the results obtained from the supervised classification process of the HS images in the database. Every fold proposed in Sect. 2.2 was trained with different hyperparameters to get the best results in each one. The pre-processing stages were carried out in MATLAB. Meanwhile, the machine learning classification was performed and evaluated in Python using Scikit-learn [14] to take advantage of the libraries and tools for this goal.

Table 1. Optimized hyperparameters of the machine learning models.

Classifier	Partitions				
	Parameter	Fold A	Fold B	Fold C	Fold D
SVM-Polynomial	<i>PolynomialDegree</i>	3	2	3	3
	<i>C</i>	1	0.126	120	0.21
SVM-Gaussian	<i>Gamma</i>	110	101	800	500
	<i>C</i>	8000	9	8	5
Random Forest	<i>Estimators</i>	100	140	88	100
	<i>MinSamplesSplit</i>	0.015	2	0.001	1
	<i>MaxDepth</i>	10	18	3	20

Table 2. Accuracy scores of the machine learning models in the test subset.

Classifier	Folds				Average Accuracy
	A	B	C	D	
SVM-Polynomial	95.22	79.27	94.53	95.22	91.06
SVM-Gaussian	93.85	79.67	94.46	92.83	90.20
Random Forest	90.00	77.25	88.92	85.66	85.46
Voting Ensemble	96.51	79.67	94.31	93.09	90.89
Average	93.89	78.96	93.05	91.70	89.40
Reference in [13]	96	79	77	90	85.5

Table 2 presents the accuracy score (%) obtained in each fold with every classifier in the test subset. Hence, Fold A reached the best performance with the Voting Ensemble at 96.51%, and the average of all machine learning models was 89.40%. However, the best machine learning model by averaging four folds was SVM-Polynomial, followed by the Voting Ensemble. Meanwhile, Fold B had the worst performance out of the four folds. To validate the results, we compared our work with [13], which used the same database and applied a similar partition methodology, but the classification model is based on an ensemble of deep neural networks. In the following, the metrics to evaluate the

performance of the machine learning models in the test subset were: accuracy, precision, F1-score, sensitivity, and specificity.

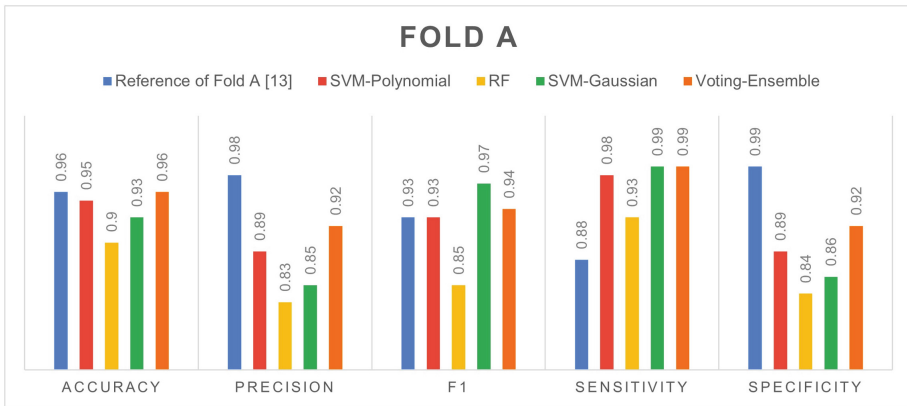


Fig. 6. Performance metrics for Fold A with the proposed machine learning models compared to [13] in the test subset.

In Fig. 6, Fold A shows good results and high similarity with the reference scores in [13]. Among the proposed classifiers, RF exhibited the poorest performance in all metrics. The reference in [13] showed a clear improvement in precision and specificity, but in accuracy, all the proposed machine learning models were close to each other. Meanwhile, in F1-score, SVM-Gaussian achieved top-performance, and in sensitivity, all the proposed models improved the reference. Hence, the resulting high sensitivity indicates the ability to accurately identify true positives.

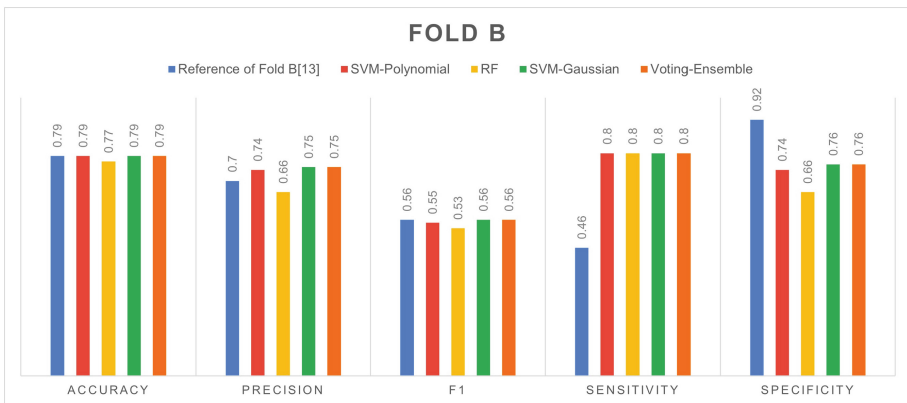


Fig. 7. Performance metrics for Fold B with the proposed machine learning models compared to [13] in the test subset.

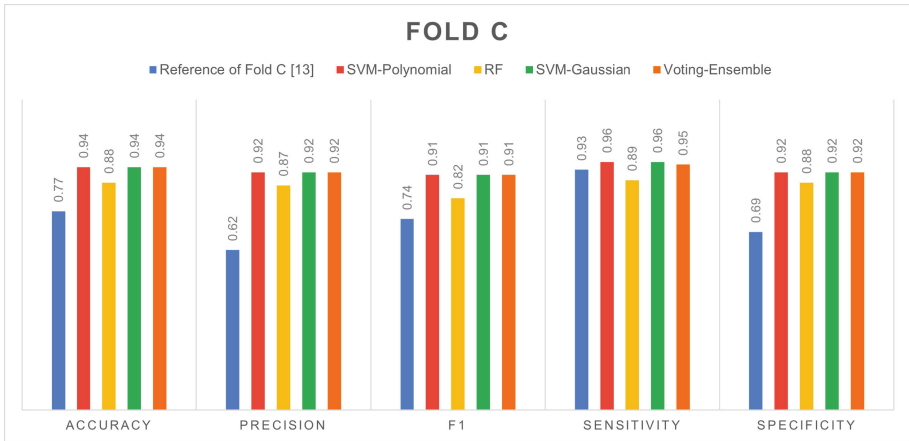


Fig. 8. Performance metrics for Fold C with the proposed machine learning models compared to [13] in the test subset.

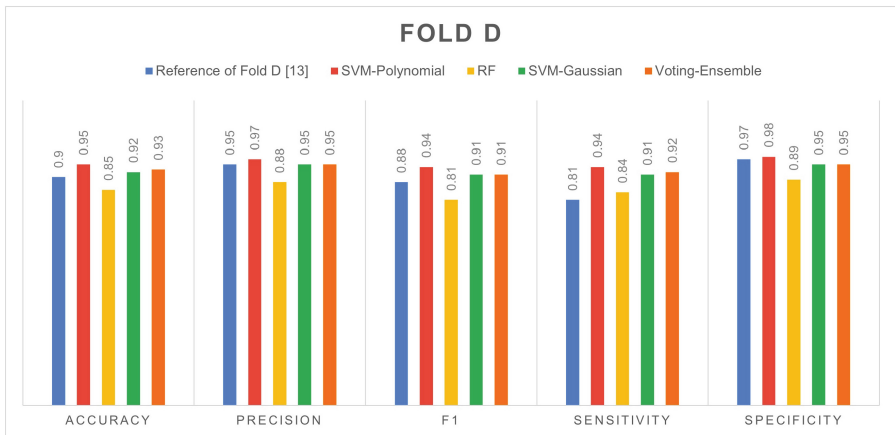


Fig. 9. Performance metrics for Fold D with the proposed machine learning models compared to [13] in the test subset.

As mentioned earlier, the lowest performance among the four partitions was observed in Fold B, where the results are illustrated in Fig. 7. All four machine learning models and the reference in [13] achieved roughly the same performance in accuracy, precision, and F1-score. But the proposed models improved sensitivity, and the reference in [13] reached the best specificity. In overall, the sensitivity rates were relatively higher than accuracy, precision, F1-score, and specificity suggesting that the proposed models were more successful in identifying true positives. This observation could be attributed to the limited number of samples and the imbalanced nature of the database.

Finally, Folds C and D demonstrated the superior performance of our machine learning models compared to the reference in [13]. Figure 8 shows the results for Fold C,

where all four proposed machine learning models outperformed the reference. For Fold D, the performance metrics are illustrated in Fig. 9, where SVM-Polynomial achieved the highest metrics among all, and improved both the reference and other classifiers in every metric. It can be inferred that the more robust results observed in Fold D might be attributed to the difference in the number of samples of each class in every fold, which can be visualized in Figs. 2 and 3.

4 Conclusions

In this work, the effectiveness of machine learning models in conjunction with SU techniques has been demonstrated for the classification of HS histopathological images for tumor tissue diagnosis. The performance metrics obtained from our proposals have surpassed in most cases the reference performance in [13]. In our study, we obtained an average accuracy of 89.40%, which improves the previous result of 85.5% in [13]. Moreover, the proposed machine learning models are based on shallow classifiers (SVM and RF) compared to the scheme in [13]. On average, the best machine learning models were SVM-Polynomial and the Voting Ensemble (see Table 2). These results open up exciting possibilities for clinical applications, where our methodology could be utilized to assist in pathological analysis and support healthcare professionals to improve decision-making processes and patient diagnosis.

Nevertheless, a limitation of this proposal is that specific hyperparameters and independent training stages were required for each fold. As future work, we will explore common hyperparameters in the four folds that could maintain or exceed our performance metrics. Besides that we will investigate other pathologies that can be analyzed using HSI and this classification methodology.

References

1. Lu, G., Fei, B.: Medical hyperspectral imaging: a review. 10.1117/1. JBO.19.1.010901
2. Madabhushi, A.: Digital pathology image analysis: opportunities and challenges. *Imaging Med.* **1**(1), 7 (2009)
3. Halicek, M., Fabelo, H., Ortega, S., Callico, G.M., Fei, B.: In-vivo and ex-vivo tissue analysis through hyperspectral imaging techniques: revealing the invisible features of cancer. *Cancers* **11**(6), 756 (2020). <https://doi.org/10.3390/cancers11060756>
4. Batta, M.: Machine learning algorithms-a review. *Int. J. Sci. Res. (IJSR)* **9**, 381–386 (2020)
5. Magoulas, G.D., Prentza, A.: Machine learning in medical applications. In: Paliouras, G., Karkaletsis, V., Spyropoulos, C.D. (eds.) *ACAI 1999*. LNCS, vol. 2049, pp. 300–307. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-44673-7_19
6. Sortega, S., et al.: Hyperspectral imaging and deep learning for the detection of breast cancer cells in digitized histological images. *Proc. SPIE* **11320** (2020). <https://doi.org/10.1117/12.2548609>
7. Collins, T., et al.: Automatic recognition of colon and esophagogastric cancer with machine learning and hyperspectral imaging. *Diagnostics* **11**(10), 1810 (2021). <https://doi.org/10.3390/diagnostics11101810>
8. Ortega, S., et al.: Hyperspectral imaging for the detection of glioblastoma tumor cells in H&E slides using convolutional neural networks. *Sensors* **20**(7), 1911 (2020). <https://doi.org/10.3390/s20071911>

9. Ortega, S., et al.: Hyperspectral superpixel-wise glioblastoma tumor detection in histological samples. *Appl. Sci.* **10**(13), 4448 (2020). <https://doi.org/10.3390/app10134448>
10. Campos-Delgado, D.U., et al.: Extended blind end-member and abundance extraction for biomedical imaging applications. *IEEE Access* **7**, 178539–178552 (2019)
11. Gan, J., Liu, T., Li, L., Zhang, J.: Non-negative matrix factorization: a survey. *Comput. J.* **64**(7), 1080–1092 (2021)
12. Geron, A.: *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc., Sebastopol (2022)
13. Cruz-Guerrero, I.A., et al.: Hybrid brain tumor classification scheme of histopathology hyperspectral images using linear unmixing and deep learning. <https://doi.org/10.2139/ssrn.4292742>
14. Raschka, S., Liu, Y.H., Mirjalili, V., Dzhulgakov, D.: *Machine Learning with PyTorch and Scikit-Learn: Develop Machine Learning and Deep Learning Models with Python*. Packt Publishing Ltd., Birmingham (2022)



Classification of COVID-19 Mortality Risk: A Neural Network-Based Approach Using Mexican Healthcare Sector Data

Enrique Mena-Camilo¹(✉) , Gerardo Hernández-Nava¹ ,
Sheila Leyva-López¹ , and Sebastián Salazar-Colores² 

¹ Facultad de Ingeniería, Universidad Autónoma de Querétaro,
Santiago de Querétaro, 76010 Querétaro, Mexico
enriquemece97@gmail.com

² Centro de Investigaciones en Óptica, 37150 León, Guanajuato, Mexico

Abstract. This study focuses on the use of neural networks and clinical data collected by the Mexican Ministry of Health to classify the risk of death from COVID-19. A multi-layer perceptron neural network model was designed, achieving remarkable results with an accuracy of 96.28%, sensitivity of 99.23%, and an F1 score of 0.9773. The model was optimized through meticulous exploration of various network configurations and performance enhancement techniques. The results showcase the efficacy of neural networks in predicting the risk of death, allowing healthcare professionals to prioritize treatment and allocate resources more efficiently. The value of artificial intelligence in the fight against the pandemic is emphasized, along with its potential application in diverse geographical and healthcare contexts. This work contributes to the advancement of predictive models and encourages further research in the fields of epidemiology and artificial intelligence to combat COVID-19.

Keywords: COVID-19 · Artificial Neural Network · Multi-Layer Perceptron · Mortality Risk

1 Introduction

The COVID-19 pandemic, has posed an unprecedented global health challenge. The World Health Organization (WHO) reports that as of June 2023, the virus has caused the death of over 6.9 million people worldwide, with the actual number possibly exceeding 20 million [13, 15]. Despite the declaration of the national emergency as concluded, the number of infections and fatalities continues to rise. In just June 2023, WHO reports 1 million new cases and 5,700 deaths [14], while the Mexican Government during the same period reports 14,820 new cases and 143 deaths [4].

The pandemic has highlighted the need for more sophisticated data processing techniques to handle the vast amount of generated information. Artificial

Intelligence (AI), specifically Artificial Neural Networks (ANNs), holds great potential in these circumstances. ANNs are computational models that mimic the pattern recognition capabilities of the human brain, enabling the identification of complex patterns within extensive datasets [9, 10].

In this context, various AI-based models predict patient mortality, aiming for early at-risk identification. Notable examples include using LASSO regression and XGBoost to predict mortality from clinical and inflammatory features [5]; machine learning on lab markers like C-reactive protein and D-dimer for mortality forecasting [3]; combining demographic, clinical, and lab data in logistic regression achieving high effectiveness [6]; using advanced algorithms like XGBoost and CatBoost for ventilation and mortality anticipation [16]; a model combining neural networks and random forests for hospital admission mortality prediction from blood samples [7]; adopting federated learning across hospitals to enhance COVID-19 patient mortality prediction [12]; and a risk scoring system based on an ANN showing high accuracy in COVID-19 patient mortality prediction at admission [1].

In contrast to these studies, it is essential to emphasize that several of them are based on specific hospital datasets or particular regions, potentially limiting the extrapolation of their findings to broader populations. Furthermore, while some articles rely on existing laboratory data to construct their models, there have also been cases where the availability of such data is restricted within the scope of the proposed research. The lack of laboratory information could hinder the replication and applicability of the models in diverse clinical settings. These considerations underscore the importance of addressing data heterogeneity and the need for accessibility to comprehensive clinical information to achieve robust and generalizable predictions.

Hence, this research aims to design an ANN capable of capturing patterns from a dataset encompassing an entire nation, thereby incorporating diversity and enriching the data. Additionally, it includes data from patients of other nationalities who were treated in Mexico.

2 Materials and Methods

2.1 Dataset Description and Preprocessing

The data used during this project comes from the open data portal of the Mexican Ministry of Health [11]. It consists of a CSV file containing a total of 7,294,030 instances and 40 attributes.

Out of the available 40 attributes, the *date of death* attribute was selected as the target variable through a labeling process, where, if the date of death was present, the label *death* was assigned, and in the absence, the label *no death* was assigned. Out of the remaining 39 attributes, 24 were excluded because they did not provide information related to clinical indicators, such as: *registration ID*, *birth entity*, *residence entity*, *treatment entity*, *indigenous*, among others. This resulted in a final set of 15 clinical indicators, of which 14 are categorical

(*immunosuppression, hypertension, cardiovascular disease, obesity, gender, intubation, pneumonia, pregnancy, diabetes, chronic obstructive pulmonary disease, chronic kidney disease, smoking, asthma, and nationality*) and 1 is numerical (*age*).

Given the objective of this research to develop a classifier for COVID-19 mortality risk, a class balance analysis was conducted. It was revealed that 99% of the utilized dataset contained instances labeled as *no death* (coded as 0), while only the remaining 1% represented instances labeled as *death* (coded as 1). Due to this class imbalance, a data reduction technique was employed by randomly selecting 46,729 instances from the majority class (category 0), equalizing the number of available samples in the minority class (category 1). This data reduction approach resulted in an equal number of instances for each class, totaling 93,458 instances.

After selecting the instances, a process of data encoding and imputation was carried out. Since the attributes to be encoded were categorical in nature, ordinal encoding was applied to them. For the imputation process, mode imputation was applied to categorical attributes, while for the *age* attribute, mean imputation was performed with the decimal part truncated to the nearest integer.

2.2 Search and Selection of ANN Architectures

Inspired by a Multi-Layer Perceptron (MLP) architecture, a heuristic search was conducted to identify optimal network configurations to achieve the objective of classifying mortality risk. The search involved modifying various hyperparameters, such as the number of hidden layers, the number of neurons per layer, dropout regularization, and batch normalization. However, certain hyperparameters, including training epochs, batch size, optimizer, learning rate, and loss function, remained fixed. The range of values tested for each variable hyperparameter is illustrated in Table 1, while Table 2 provides the values of the fixed hyperparameters used during the training process.

Table 1. Set of hyperparameters and values modified during the search for MLP architectures.

Hyperparameter	Tested values
Number of hidden layers	2, 3, 4
Number of neurons per layer	16, 32, 64, 128, 256, 512, 1024
Regularization by dropout	Enabled or disabled
Batch normalization	Enabled or disabled

As shown in Table 1, the architecture search process encompassed evaluating various MLPs ranging from 2 to 4 hidden layers. The number of neurons in each hidden layer spanned powers of 2 from 4 to 10. The ReLU activation

Table 2. Set of hyperparameters and fixed values during the search for MLP architectures.

Hyperparameter	Value
Training epochs	10
Batch size	1024
Optimizer	Adam
Learning rate	0.001
Loss function	Binary crossentropy

function was consistently applied to all hidden layers. Furthermore, performance-enhancing techniques like dropout regularization and batch normalization were flexible options that could be independently enabled or disabled for each hidden layer within a given architecture. As a result, certain layers could have these enhancements while others did not, allowing for greater flexibility in optimizing network performance.

Throughout this search process, a total of 100 unique hyperparameter combinations were tested using a holdout evaluation method. From this comprehensive evaluation, the five architectures that exhibited the highest accuracy were selected. To ensure robustness, these architectures underwent further assessment via cross-validation in later stages of the analysis.

2.3 Architectures Training and Validation

Performance metrics are crucial tools for evaluating the effectiveness of a classification model in machine learning. They help us understand how well the model performs and where improvements can be made.

Five widely used metrics were employed, as commonly found in the literature: $accuracy = \frac{TP+TN}{TP+FP+FN+TN}$, $sensitivity = \frac{TP}{TP+FN}$, $specificity = \frac{TN}{TN+FP}$, $F1 - score = \frac{2TP}{2TP+FP+FN}$, and $precision = \frac{TP}{TP+FP}$ [2]. Where: TP, true positive; TN, true negatives; FP, false positives; FN, false negatives [2].

To obtain reliable performance values, a cross-validation methodology was employed. The dataset is divided into subsets for training and testing. In k-fold cross-validation, the dataset is divided into k subsets, and each subset is used as a test set while the others are used for training [8]. This process is repeated k times, and at the end, the metrics obtained in each iteration are averaged to derive the model's performance. Common values for k are 5 and 10, and in this project, a 10-fold cross-validation was utilized, resulting in a distribution of 90% training and 10% testing in each fold.

3 Results and Discussions

Related to the class balancing and data imputation process, a description of the distribution for each of the utilized attributes can be observed in Table 3

and Table 4. In these tables, it can be observed that the performed sampling and imputation process ensures that the data used for analysis constitutes a representative sample of the original data. In other words, the data preprocessing performed does not affect the distribution of the data.

Table 3. Distribution of features in original data set.

Feature	Original dataset	
	Non dead	Dead
Sex	Female (57.0), Male (43.0)	Female (40.9), Male (59.1)
Intubed	97 (96.2), 2 (3.7), 1 (0.1)	97 (0.3), 2 (84.2), 1 (14.9), 99 (0.6)
Pneumonia	2 (97.6), 1 (1.4), 99 (1.0)	2 (44.4), 1 (55.5), 99 (0.1)
Age	36 (25–49)	69 (56–80)
Pregnancy	2 (55.0), 97 (43.0), 1 (1.4), 99 (0.6)	2 (40.8), 97 (59.1), 1 (0.1), 99 (0.1)
Diabetes	2 (93.7), 1 (6.0), 98 (0.4)	2 (63.3), 1 (36.3), 98 (0.4)
COPD	2 (99.1), 1 (0.5), 98 (0.4)	2 (99.2), 1 (7.4), 98 (0.4)
Asthma	2 (97.7), 1 (1.9), 98 (0.4)	2 (98.0), 1 (1.7), 98 (0.3)
INMUSUPR	2 (99.1), 1 (0.5), 98 (0.4)	2 (95.0), 1 (4.6), 98 (0.4)
Hypertension	2 (91.2), 1 (8.5), 98 (0.4)	2 (55.8), 1 (43.8), 98 (0.4)
Cardiov Disease	2 (98.8), 1 (0.8), 98 (0.4)	2 (91.2), 1 (8.4), 98 (0.4)
Obesity	2 (93.6), 1 (6.0), 98 (0.3)	2 (88.7), 1 (11.1), 98 (0.2)
CKD	2 (98.9), 1 (0.7), 98 (0.4)	2 (87.6), 1 (12.0), 98 (0.3)
Smoking	2 (95.5), 1 (4.2), 98 (0.4)	2 (91.2), 1 (8.4), 98 (0.4)
Nationality	MX (99.3), USA (0.3), Others (0.4)	MX (99.7), USA (0.1), Others (0.2)

Categorical features are expressed as ‘label (%)’. Feature ‘age’ is expressed as ‘mean (IQR)’.

Abbreviations: IQR, Interquartile range; Cardiov disease, Cardiovascular Disease; INMUSUPR, Immunosuppression; CKD, Chronic kidney disease; COPD, Chronic obstructive pulmonary disease.

Labels: 1, Yes; 2, No; 97, Not Apply; 98, Unknown; 99, Unknown.

The resulting metrics from the 10-fold cross-validation can be observed in Table 5, presenting the mean values and their corresponding standard deviations (std).

Based on the findings presented in Table 5, it is noticeable that the results obtained for each of the models exhibit slight variations. Therefore, an ANOVA test was applied to determine whether there is a significant difference among the various metrics. This statistical test yielded a p -value of 1.00, indicating that there is no significant difference in the performance of each model.

These conditions of non-significant difference led to the selection of the model with lower complexity as the best model. Consequently, the third model was chosen as the best. The architecture proposed by the third model is summarized in Table 6, where we can observe the number of layers used, as well as their respective sizes.

Table 4. Distribution of features in a sampled data set.

Feature	Sampled dataset	
	Non dead	Dead
Sex	Female (57.1), Male (42.9)	Female (40.9), Male (59.1)
Intubed	97 (96.2), 2 (3.7), 1 (0.1)	97 (0.8), 2 (84.2), 1 (14.9)
Pneumonia	2 (98.6), 1 (1.4)	2 (44.5), 1 (55.5)
Age	36 (25–49)	69 (56–80)
Pregnancy	2 (98.6), 1 (1.4)	2 (99.9), 1 (0.1)
Diabetes	2 (94.1), 1 (5.9)	2 (63.7), 1 (36.3)
COPD	2 (99.4), 1 (0.6)	2 (92.6), 1 (7.4)
Asthma	2 (98.1), 1 (1.9)	2 (98.3), 1 (1.7)
INMUSUPR	2 (99.5), 1 (0.5)	2 (95.4), 1 (4.6)
Hypertension	2 (91.4), 1 (8.6)	2 (56.2), 1 (43.8)
Cardiov disease	2 (99.1), 1 (0.9)	2 (91.6), 1 (8.4)
Obesity	2 (93.8), 1 (6.2)	2 (88.9), 1 (11.1)
CKD	2 (99.2), 1 (0.8)	2 (88.0), 1 (12.0)
Smoking	2 (96.1), 1 (3.9)	2 (91.6), 1 (8.4)
Nationality	MX (99.2), USA (0.3), Others (0.5)	MX (99.7), USA (0.1), Others (0.2)

Categorical features are expressed as ‘label (%)’. Feature ‘age’ is expressed as ‘mean (IQR)’. Abbreviations: IQR, Interquartile range; Cardiov disease, Cardiovascular Disease; INMUSUPR, Immunosuppression; CKD, Chronic kidney disease; COPD, Chronic obstructive pulmonary disease.

Labels: 1, Yes; 2, No.

Table 5. Comparison of the performance metrics obtained in the cross-validation.

Model	Accuracy [%]		Sensitivity [%]		Specificity [%]		Precision [%]		F1-score	
	Mean	Std ($\times 10^{-5}$)	Mean	Std ($\times 10^{-5}$)	Mean	Std ($\times 10^{-5}$)	Mean	Std ($\times 10^{-5}$)	Mean	Std ($\times 10^{-5}$)
Model 1	97.68	9.2	99.16	29.9	96.31	16.0	96.31	16.0	0.9771	9.4
Model 2	97.69	4.2	99.22	5.5	96.28	5.2	96.28	5.2	0.9773	4.2
Model 3	97.70	8.1	99.23	24.1	96.28	16.4	96.28	16.4	0.9773	8.1
Model 4	97.68	9.0	99.20	22.8	96.28	15.4	96.28	15.4	0.9772	9.0
Model 5	97.69	33.4	99.22	68.2	96.27	12.9	96.27	12.9	0.9773	33.6

Table 6. Description of the architecture with the best performance obtained.

Layer	Units	Activation	Dropout	Batch normalization
Input	15	Linear	False	True
Hidden 1	256	ReLU	False	False
Hidden 2	128	ReLU	False	False
Output	1	Linear	False	False

Lastly, taking the best model as the reference (model number 3), a comparison was conducted with the reported metrics from other algorithms implemented for similar purposes as in the present study. This comparative analysis of metrics is presented in Table 7.

As shown in Table 7, the architecture proposed in model 3 outperforms the metrics reported by other state-of-the-art algorithms. However, it’s important to highlight that this comparison of metrics may be subject to biases, as each

Table 7. Comparison of the best model obtained against the best ones found in the state of the art.

Article	Data	Algorithm	Accuracy	Sensitivity	Specificity	Precision	F1-score	AUC
	Provider		[%]	[%]	[%]	[%]		[%]
Our best model	Secretaria de Salud México*	MLP	97.70	99.23	96.28	96.28	0.97	-
Booth et al. [3]	University of Texas Medical Branch	SVM	-	91.00	91.00	-	-	93.00
Guan et al. [5]	Wuhan China Hospitals	XGBoost	-	85.00	-	90.00	0.90	-
Hu et al. [6]	Tongji Hospital, Wuhan	RF	-	91.40	76.00	-	-	92.20
Yu et al. [16]	Beaumont Health	Catboost	80.3	-	-	79.00	-	85.00
Ko et al. [7]	Hospitals from China and Korea*	DNN+RF	93.00	92.00	93.00	-	-	-
Vaid et al. [12]	Mount Sinai Health System	MLP	78.00	80.50	70.20	-	0.32	83.60
Abdulaal et al. [1]	West London Teaching Hospital	ANN	86.25	87.50	85.90	60.87	-	90.12

* *Open access dataset*

The symbol - indicates that the information is not provided or is not calculated.

Abbreviations: RF, Random Forest; XGBoost, Extreme Gradient Boosting; DNN, Dense Neural Network; SVM, Support Vector Machine.

research study employs different datasets, which in turn vary in the number of instances and attributes used.

4 Conclusions

This study demonstrates the potential and effectiveness of ANNs in accurately predicting the risk of COVID-19 mortality. By leveraging tabular data collected from the healthcare sector in Mexico, we successfully developed and trained a model that excels in performance evaluation metrics, exhibiting high precision and accuracy.

The detailed and meticulous methodology, which encompassed an extensive exploration of MLPs and advanced performance enhancement techniques, played a pivotal role in the success of this project. The evaluation process, utilizing the holdout method followed by cross-validation on the most promising architectures, facilitated robust model optimization and validation.

With an impressive accuracy of 97.70%, notable sensitivity of 99.23%, an F1 score of 0.9773, precision of 96.28%, and specificity of 96.28%, the final model showcased outstanding performance across all metrics. These exceptional outcomes underscore the potential of neural networks and artificial intelligence in the realm of healthcare and epidemiology.

In conclusion, this work highlights the significant contribution of artificial intelligence in the battle against the COVID-19 pandemic. While tailored to the Mexican context, this study holds immense potential for application in other geographical and healthcare settings. This approach has the potential to become a valuable tool for healthcare professionals worldwide, as it aims to utilize readily accessible clinical parameters from electronic health records, assisting in guiding treatment decisions and enhancing patient outcomes.

References

1. Abdulaal, A., Patel, A., Charani, E., Denny, S., Mughal, N., Moore, L.: Prognostic modeling of COVID-19 using artificial intelligence in the United Kingdom: model

- development and validation. *J. Med. Internet Res.* **22**(8), e20259 (2020). <https://doi.org/10.2196/20259>
2. Aceves Fernández, M.A.: *Inteligencia artificial para programadores con prisa*. Universo de Letras (2021). ISBN 9788418854613
 3. Booth, A.L., Abels, E., McCaffrey, P.: Development of a prognostic model for mortality in COVID-19 infection using machine learning. *Mod. Pathol.* **34**(3), 522–531 (2021). <https://doi.org/10.1038/s41379-020-00700-x>
 4. Gobierno de México: COVID-19 México. <https://datos.covid-19.conacyt.mx/>. Accessed 1 July 2023
 5. Guan, X., et al.: Clinical and inflammatory features based machine learning model for fatal risk prediction of hospitalized COVID-19 patients: results from a retrospective cohort study. *Ann. Med.* **53**(1), 257–266 (2021). <https://doi.org/10.1080/07853890.2020.1868564>
 6. Hu, C., et al.: Early prediction of mortality risk among patients with severe COVID-19, using machine learning. *Int. J. Epidemiol.* **49**(6), 1918–1929 (2020). <https://doi.org/10.1093/ije/dyaa171>
 7. Ko, H., et al.: An artificial intelligence model to predict the mortality of COVID-19 patients at hospital admission time using routine blood samples: development and validation of an ensemble model. *J. Med. Internet Res.* **22**(12), e25442 (2020). <https://doi.org/10.2196/25442>
 8. Laura-Ochoa, L.: Evaluation of classification algorithms using cross validation. In: *Industry, Innovation, and Infrastructure for Sustainable Cities and Communities*, pp. 24–26 (2019). <https://doi.org/10.18687/LACCEI2019.1.1.471>
 9. Leyva-López, S., Salazar-Colores, S., Hernández-Nava, G., Pedraza-Ortega, J.C.: Aprendizaje automático para la detección del daño pulmonar a través de parámetros clínicos, chap. 19. In: *Diseño y Planeación Mecatrónica*, pp. 262–271. Asociación Mexicana de Mecatrónica A.C (2022). ISBN 978-607-9394-25-7
 10. Patterson, J., Gibson, A.: *Deep Learning*. O'Reilly Media, Inc. (2017). ISBN 9781491914250
 11. Secretaría de Salud: Información referente a casos COVID-19 en México. <https://datos.gob.mx/busca/dataset/informacion-referente-a-casos-covid-19-en-mexico>. Accessed 2 July 2023
 12. Vaid, A., et al.: Federated learning of electronic health records to improve mortality prediction in hospitalized patients with COVID-19: machine learning approach. *JMIR Med. Inform.* **9**(1), e24207 (2021). <https://doi.org/10.2196/24207>
 13. World Health Organization: Coronavirus disease (COVID-19) Weekly Epidemiological Updates and Monthly Operational Updates. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>. Accessed 1 July 2023
 14. World Health Organization: Weekly epidemiological update on COVID-19, 29 June 2023. <https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---29-june-2023>. Accessed 1 July 2023
 15. World Health Organization: WHO Director-General's opening remarks at the media briefing, 5 May 2023. <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing---5-may-2023>. Accessed 1 July 2023
 16. Yu, L., et al.: Machine learning methods to predict mechanical ventilation and mortality in patients with COVID-19. *PLoS ONE* **16**(4), 1–18 (2021). <https://doi.org/10.1371/journal.pone.0249285>



Design of a Convolutional Neural Network for Hippocampal Segmentation in Epileptics and Healthy Patients

Alina Andrea García Huizar^(✉)  and José Manuel Mejía Muñoz 

Instituto de Ingeniería y Tecnología de la Universidad Autónoma de Ciudad Juárez,
Ave. del Charro 450, 32310 Ciudad Juárez, Mexico
alina.garhuizar@gmail.com

Abstract. Manual segmentation is still the gold standard in hippocampal segmentation due to its precision. However, it requires a considerable amount of time. Convolutional neural networks offer a less resource-intensive alternative. In this study, we propose a parallel convolutional neural network architecture for segmenting the hippocampus in patients with epilepsy and healthy patients based on magnetic resonance images. Our network design resembles a wavelet filter bank but utilizes adaptive filters generated by convolutional layers. By employing a limited number of convolutional layers, our approach achieves improved computational efficiency compared to existing network models in the literature. The performance evaluation was conducted using the Jaccard index, Dice coefficient, sensitivity, and precision, and compared to the widely used U-Net network segmentation. The results, based on similarity metrics, indicate that the parallel network demonstrates higher predictive similarity to the test dataset, achieving a precision score of 0.80 outperforming the U-Net network.

Keywords: Convolutional Neural Network · Hippocampus Segmentation · Wavelet

1 Introduction

Around 50 million people in the world experience epilepsy, making it one of the most common neurological disorders, and organizations such as the World Health Organization recognize it as a major health problem. One of the most common form of human epilepsy is temporal lobe epilepsy. Nowadays, manual segmentation software allows professionals to perform automatic hippocampus segmentation in the temporal lobe in magnetic resonance imaging for epilepsy detection. These methods are costly and time-consuming, so neural network-based solutions have been used instead. The U-Net architecture is often used as the base network in several models for image segmentation tasks in the biomedical field [1–3].

In 2021, in [2] a network employing the U-Net architecture was introduced. In their work several alterations were implemented to the original U-Net structure: residual connections were integrated between the convolutional block's input and output, to each

convolutional layer, batch normalization was used, and instead of utilizing a single input, 2D adjacent patches were employed. The devised approach underwent training and validation using the HCUncamp database as outlined in their study. Likewise, the method presented by [3] in 2019 utilized a 3D U-Net with a fusion of the anatomical planes' outputs, which incorporates different segmentation and error correction steps employing replacing and refining networks. Finally, the Quick Nat software from [4] uses various 2D U-Net approximation for each image orientation and masks are manually refined. Free Surfer software is used for auxiliary data augmentation.

Prior research has showcased the utilization of intricate and resource-intensive networks for segmentation tasks, such as U-Net. These networks often pose challenges in terms of implementation, particularly when integrating algorithms into embedded systems or when dealing with volumetric data that necessitates a specialized server for processing. The complexity and computational demands of these networks impede their practicality in real-world scenarios, where efficiency and resource utilization are crucial. Additionally, the need for specialized hardware or dedicated servers for processing volumetric data further adds to the complexity and cost of implementing such segmentation algorithms. Addressing these limitations, our study aims to explore alternative approaches that prioritize simplicity and computational efficiency without compromising segmentation accuracy [5–7].

Hereby, we introduce a 2D Parallel convolutional neural network (CNN) for automating hippocampal segmentation for MRI images of the temporal lobe. Our proposed architecture exhibits several advantages over existing models, such as the well-known U-Net. Notably, our network comprises fewer components, resulting in improved computational efficiency. Moreover, it surpasses the time-consuming manual segmentation process in terms of speed. One key feature of our architecture is the utilization of multiple channels to process the input image. Each channel applies filters of varying sizes, allowing us to capture diverse levels of detail comparable to a wavelet filter bank. However, unlike fixed wavelet filters, our adaptive filters are dynamically determined through the convolutional layers. This approach enables our network to adapt and extract relevant features from the input image, optimizing the accuracy of the segmentation process. By incorporating these adaptive filters within our architecture, we enhance the network's ability to capture intricate hippocampal structures and boundaries more effectively. By proposing this 2D Parallel CNN, we aim to provide a more computationally efficient and accurate solution for hippocampal segmentation in MRI images.

2 Methods

In this section, we provide a comprehensive description of the Parallel and U-Net architectures, along with the preprocessing steps, training process, and evaluation techniques utilized in our study.

2.1 Database

Several databases were searched on the Kaggle site. The Hippocampus Segmentation in MRI Image database was selected, which contains 50 MRI images in ANALYZE format

with T1W weighting, from 40 epileptic and 10 non-epileptic patients. A magnetic field of 1.5 T was applied to 20 epileptic and 10 non-epileptic subjects, with the remaining subjects a 3 T magnetic field was used: in both cases with impaired recovered gradient echo sequence (SPGR).

In Fig. 1 is depicted the anatomical planes in a sample from the set of images from the database. Note that the following is included for segmentation, subiculum, head, body, and tail of the hippocampus. In contrast, the alveus and fimbria were not included: neither the amygdala nor the temporal horn of the lateral ventricle [8].

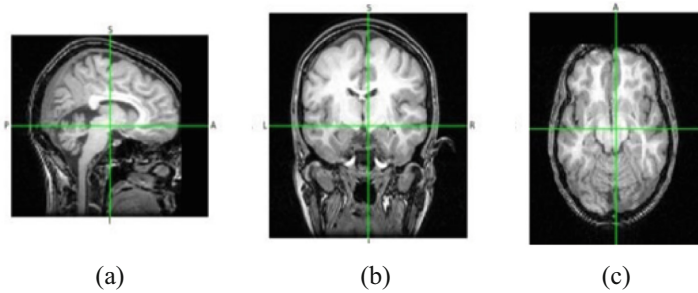


Fig. 1. MRI from the database. (a) Sagittal, (b) Coronal and (c) Transverse slices.

2.2 Preprocessing

Originally the database contained images size 516. However, the images were modified with a resize function to 128×128 , to speed up the training and evaluation of the networks. Only the coronal slices were chosen, in a range of 40 to 70 slices, where the hippocampal segmentation is best appreciated. Figure 2 shows both images already modified. Additionally, the label's set images were binarized with a threshold of 0.5. Finally, data partitioning was performed. The `train_test_split` function was used on a total of 755 images, to obtain a test set of 90% and a training set of 10%.



Fig. 2. Images obtained after preprocessing. (a) Original coronal slice (b) Hippocampal mask resized segmentation.

2.3 Architecture for Segmentation

To accomplish the hippocampus segmentation, we proposed a modified and adapted deep learning network architecture based on the framework presented in reference, [9, 10]. Figure 3 shows a schematic of the network. Our architecture consists of three channels, each incorporating a total of nine convolutional layers. The purpose of utilizing multiple channels is to process the input image using filters of varying sizes, enabling us to capture different levels of detail akin to a wavelet filter bank. However, unlike fixed wavelet filters, our adaptive filters are obtained through the convolutional layers.

Specifically, the three channels operate at different resolutions to extract distinct representations from the input image. The first channel operates at a coarse resolution using a 9×9 filter, enabling the acquisition of broader contextual information. The second channel operates at a medium resolution using a 4×3 filter size, facilitating the extraction of intermediate-level features. Lastly, the third channel employs a fine resolution with a 2×2 filter size, allowing for the capture of fine-grained details.

The information from each channel is then concatenated and fed into a series of convolutional networks. By incorporating multiple channels, we aim to obtain diverse representations of the input signal, which in turn aids the final layers of the network in accurately segmenting the different regions of the hippocampus.

The sigmoid activation function is employed in the network. The rationale behind this is worth noting. It facilitates the tendency of the output image to become binary, which aligns perfectly with the requisite for our specific two-class segmentation.

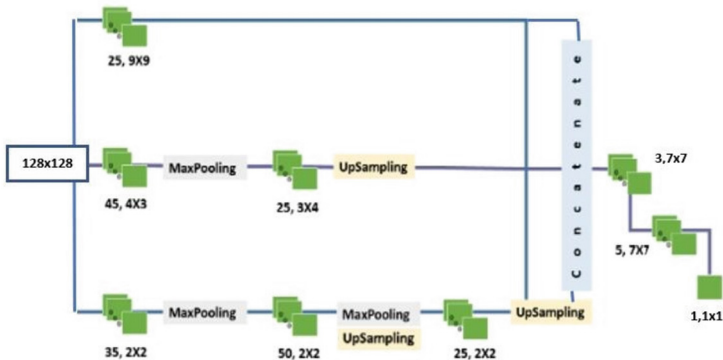


Fig. 3. Parallel architecture.

To compare the proposed network, we used a U-Net, with the following attributes. For the contracting path, a Conv2D layer of 32 kernels of size 3×3 with ReLU activation was used, followed by a MaxPooling layer of size 2×2 . Another three layers were added for the downstream phase, with 64, 128, and 256 filters of the same size as the input layer, with their respective reduction layers. The intermediate layer also consisted of Conv2D using 512 kernels of size 3×3 with ReLU activation. For the expansion pathway, three layers with 64, 128, and 256 filters of the same size as the input layer were used, with their respective up sampling layer, which was performed with the Conv2DTranspose

function, with 25 filters of size 2×2 . Finally, the output layer consisted of a convolution with a filter of size 1×1 with sigmoid activation.

Both networks employed as loss function the binary crossentropy, and as optimizer the Adam algorithm, and accuracy metric. In the training phase of the U-Net network, it underwent 70 epochs, a batch size of 50 and with a 0.001 learning rate. On the other hand, the second network underwent a longer training period of 400 epochs, a batch size of 40 and with a 0.0001 learning rate. The choice of binary crossentropy loss function is because the models were trained to perform binary classification tasks in this case the pixels were classified into two classes: hippocampus or not. The Adam optimizer, known for its efficiency and ability to handle large-scale optimization problems, was employed to update the weights of the network in the training.

Dice coefficient, Jaccard index (IoU), sensitivity and precision were obtained to assess the performance of both networks.

3 Results and Discussion

This section presents the results from the networks with 697 samples from the training set, of which 20% was used as the validation set. The U-Net was trained with 70 epochs and the parallel network with 400 epochs.

3.1 Test Set Result

We compare the coronal slices image, and the results of the hippocampal segmentation predictions of the images with higher accuracy, with the masks of the U-Net test set in Fig. 4. Both show the segmentation of the left and right lobes of the hippocampus. However, in the test set the edges are well defined while in the predictions it is seen to take neighboring pixels belonging to the amygdala as part of the hippocampus.

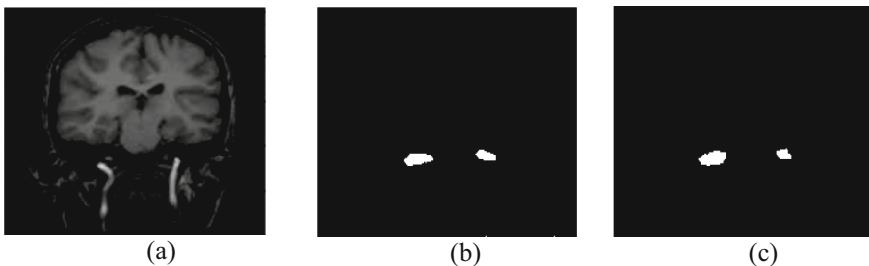


Fig. 4. Comparison of the A) coronal slices, B) the hippocampal segmentation from the test set and C) the predicted segmentation mask of the U-Net network.

Analogous, Fig. 5 put side by side the coronal slices image, its corresponding segmentation from the test set and the prediction made by the parallel network. The segmentation of both hippocampal lobes can be distinguished in yellow, where the contours are more similar to the test set compared to the first network, however, it can be seen that the network segmented pixels corresponding to the amygdala.

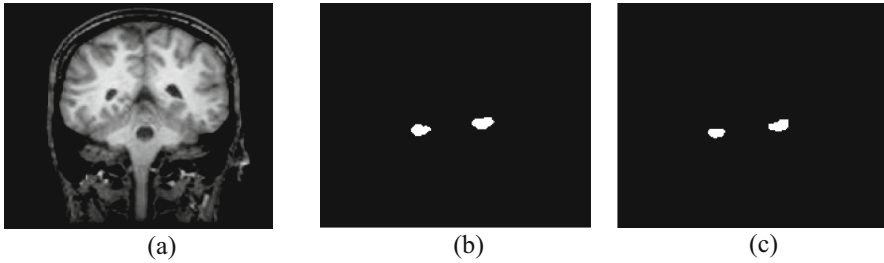


Fig. 5. Comparison of the (a) coronal slices, (b) the hippocampal segmentation from the test set and (c) the predicted segmentation mask of the parallel network.

3.2 Evaluation Metrics

Table 1 reports the average values of the evaluation metrics obtained by the U-Net and parallel network. In contrast, Table 2 shows the average values obtained by using two automatic segmentation algorithms to predict segmentation: Classifier Fusion software and Labelling (CFL) and Brain Parser using the same database. The U-Net and parallel convolutional neural networks evaluation metrics have similar values for the IoU segmentation metrics and the Dice coefficient to those resulting from using the two segmentation algorithms CFL and Brain Parser. In the case of the parallel neural network, the average in the segmentation coefficients is higher, indicating a more accurate segmentation. In the case of the sensitivity metric, a value of 0.92 was reported when using Brain Parser, higher than that achieved in the two proposed neural networks. On the other hand, when comparing the precision of the two designed neural networks, it is observed that higher values were achieved than those obtained by the algorithms.

Table 1. Evaluation Metrics of U-Net and Parallel networks on the test dataset

Architecture		IoU	Dice Coefficient	Sensitivity	Precision
U-Net	Mean	0.48	0.60	0.61	0.68
	SD	0.21	0.22	0.25	0.19
Parallel	Mean	0.63	0.76	0.75	0.80
	SD	0.12	0.10	0.12	0.13

In addition, the work in [2] obtained a Dice of 0.71 using Quick Nat network [4], a Dice of 0.76 for Ext2D, and a Dice of 0.74 for Hippodeep [12]. However, these networks were using a different database they were not included in Table 2.

The value obtained in the metrics is relatively low due to the smaller area to be segmented in each image concerning the entire image; consequently, the loss function obtains a good value even when the hippocampal pixels have not been segmented since the rest of the pixels (background) are the majority. An essential factor for segmentation is image quality. In this case, the current resolution of the images, $0.78 \times 0.78 \times 2.00 \text{ mm}^3$, does not allow adequately differentiating the edges so that the hippocampus blends with

Table 2. Evaluation Metrics of CFL and Brain Parser algorithms on the test dataset [8].

Algorithm		IoU	Dice Coefficient	Sensitivity	Precision
Brain Parser	Mean	0.47	0.64	0.92	0.50
	SD	0.06	0.06	0.06	0.08
CFL	Mean	0.60	0.75	0.74	0.77
	SD	0.07	0.07	0.10	0.07

the amygdala. This causes a lower Dice coefficient due to false positive cases. In addition, the presence of a high number of anisotropic voxels makes segmentation difficult. Other factors are intrinsic or extrinsic lesions such as hippocampal atrophy, which affect the hippocampal texture area [10, 11].

To achieve the best contrast and higher resolution in structural MRI in patients with epilepsy is recommended an MRI basic protocol that includes 3D-T1 sequences, coronal T2 and FLAIR slices, and axial FLAIR and T2 slices.

4 Conclusion

The gold standard for hippocampal segmentation has long been manual segmentation, primarily due to its high precision. However, this method is time-consuming, requiring a significant investment of time and effort. To address this issue, CNNs have emerged as a promising alternative that significantly reduces the time required for segmentation. In this study, two specific network architectures were compared: 1) the U-Net architecture, and 2) a Parallel architecture. Both networks aimed to perform hippocampal segmentation using magnetic resonance imaging data obtained from both epileptic and healthy patients. To compare the performance of the two models, evaluation metrics such as the IoU and Dice coefficient were utilized.

The U-Net network yielded the following metrics: a mean IoU value of 0.48, an average Dice coefficient of 0.60, an average sensitivity of 0.61, and a precision of 0.68. On the other hand, the parallel network demonstrated superior performance, with a mean IoU value of 0.63, a mean Dice coefficient of 0.75, an average sensitivity of 0.75, and a precision of 0.8. These results clearly indicate that the parallel network outperforms the U-Net network in terms of producing more precise segmentations of the hippocampus.

However, despite these encouraging findings, there is still room for improvement to enhance the overall reliability of the developed networks. For instance, it is recommended to augment the training dataset by including a larger number of data samples specifically obtained from the hippocampus of epileptic patients. Additionally, adopting an imaging protocol that provides higher resolution images could potentially improve the accuracy of the segmentation. Moreover, incorporating sagittal slices into the training process could lead to more precise delineation of the hippocampal boundaries, particularly in capturing the leading edges.

By addressing these suggested improvements, the performance and reliability of the networks could be significantly enhanced, making them even more valuable tools in the


field of hippocampal segmentation. Continued research and development in this area are vital to refine these neural network models further and unlock their full potential in medical image analysis applications.

References

1. WHO Homepage. <https://www.who.int/news-room/fact-sheets/detail/epilepsy>. Accessed 11 June 2023
2. Carmo, D., Silva, B., Yasuda, C., Rittner, L., Lotufo, R.: Alzheimer's disease neuroimaging initiative: Hippocampus segmentation on epilepsy and Alzheimer's disease studies with multiple convolutional neural networks. *Heliyon* **7**(2), e06226 (2021). <https://doi.org/10.1016/j.heliyon.2021.e06226>.
3. Ataloglou, D., Dimou, A., Zarpalas, D., Daras, P.: Fast and precise hippocampus segmentation through deep convolutional neural network ensembles and transfer learning. *Neuroinformatics* **17**(4), 1–20 (2019)
4. Abhijit G, R. et al.: QuickNAT: a fully convolutional network for quick and accurate segmentation of neuroanatomy. *NeuroImage* **186**, 713–727 (2019)
5. Kumar, A., Komaragiri, R., Kumar, M.: Time–frequency localization using three-tap biorthogonal wavelet filter bank for electrocardiogram compressions. *Biomed. Eng. Lett.* **9**(3), 407–411 (2019)
6. Saad, O.M., Shalaby, A., Sayed, M.S.: Automatic discrimination of earthquakes and quarry blasts using wavelet filter bank and support vector machine. *J. Seismol.* **23**(2), 357–371 (2018)
7. Samantaray Aswini, K., Rahulkar Amol, D.: New design of adaptive Gabor wavelet filter bank for medical image retrieval. *IET Image Proc.* **14**(4), 679–687 (2020)
8. Jafari-Khouzani, K., Elisevich, K., Patel, S., Soltanian-Zadeh, H.: Database of magnetic resonance images of nonepileptic subjects and temporal lobe epilepsy patients for validation of hippocampalsegmentationtechniques. *Neuroinformatics* **9**, 335–346 (2011)
9. Cordero Lares, A.K.: Diseño de una arquitectura de red para la segmentación de regiones tumorales en pulmón por medio de tomografía computarizada, aplicando redes convolucionales. Dissertation, University of Juarez, pp. 19–20 (2019)
10. Medina Bonilla, E.: Diseño de un sistema termográfico para el análisis de tablillas electrónicas fabricadas en la industria automotriz. Dissertation, University of Juarez (2019)
11. Alvarez-Linera, J.: Structural magnetic resonance imaging in epilepsy. *Radiologia* **54**(1), 9–20 (2012). <https://doi.org/10.1016/j.rx.2011.07.007>
12. Thyreau, B., Sato, K., Fukuda, H., Taki, Y.: Segmentation of the hippocampus by transferring algorithmic knowledge for large cohort processing. *Med. Image Anal.* **43**, 214–228 (2018)



Development and Evaluation of a Diagnostic Exam for Undergraduate Biomedical Engineering Students Using GPT Language Model-Based Virtual Agents

Alberto Isaac Perez Sanpablo^{1,2}(✉) , María del Carmen Arquer Ruiz¹, Alicia Meneses Peñaloza³, Gerardo Rodríguez Reyes⁴, Ivett Quiñones Uriostegui², and Leonardo Eliú Anaya Campos²

¹ Biomedical Engineering Department, La Salle University, Cuauhtemoc, 06140 Mexico City, Mexico

² Laboratory of Motion Analysis and Rehabilitation Engineering, National Institute of Rehabilitation, Tlalpan, 14389 Mexico City, Mexico
albperez@inr.gob.mx

³ Pediatric Rehabilitation Department, National Institute of Rehabilitation, Tlalpan, 14389 Mexico City, Mexico

⁴ Technological Research Department, National Institute of Rehabilitation, Tlalpan, 14389 Mexico City, Mexico

Abstract. The creation of a diagnostic exam for biomedical engineering undergraduate students using virtual agents based on GPT language models is analyzed in this work. Thirty-nine eighth-semester students answered a 20-question exam generated by ChatGPT-3 covering the topics of acquisition, amplification, processing, and visualization of biomedical signals encompassing different levels of thinking according to the taxonomy of Bloom, including the application, analysis, and evaluation levels. Three academic experts assessed the quality of the questions based on clarity, relevance, level of thinking, and difficulty. Also, difficulty and discrimination indexes and Rasch analysis were calculated. Students obtained an average grade of 5.91, with a standard deviation of 1.39 points. Subject reliability was 0.599, and the p-value for the fit of the model of Rasch was 0.017. High correlations between some questions were observed. Based on their difficulty, a few questions could be considered irrelevant. The Wright competency map showed a good distribution on the ability scale with some redundancies and gaps. In conclusion, virtual agents have great potential to create diagnosis exams in biomedical engineering. However, it is necessary to consider their limitations and conduct a rigorous evaluation of the quality and reliability of the questions generated.

Keywords: ChatGPT-3 · education · biomedical engineering · teaching · artificial intelligence

1 Introduction

Artificial intelligence and virtual agents based on GPT language models have recently gained relevance in academia [1]. Texts on specific academic topics can be generated by these tools facilitating the teaching and learning process. Their use allows academics to focus on essential tasks through time savings [2]. These tools can simplify tasks for scholars and improve their efficiency by making summaries and translating information [3]; Virtual agents can help students learn by acting as tutors and answering their questions. For example, in microbiology, ChatGPT-3, a virtual agent based on GPT language models, has proven to be effective for automatically answering questions, providing students with accurate and relevant information [4].

However, scholars must know the limitations and potential biases of virtual agents. Although the information generated by these tools may seem authentic to a not fully trained reader, such as a student-in-training, it is essential to recognize that ChatGPT needs help interpreting and understanding the content profoundly, which can probably cause it to generate incorrect information [5]. Thus, there is concern about the reliability and possible biases of virtual agents since these tools are trained on large amounts of data and may reflect the tendencies present in that data. Therefore, it is essential to verify the facts of all virtual agent statements and be aware of the possibility of incorrect content. Another reported aspect is the tendency of virtual agents to generate information not based on actual events, known as “hallucinating” [6]. Therefore, it is necessary to exercise critical judgment when evaluating the text generated by virtual agents and use reliable sources to contrast the information.

Despite these considerations, the use of virtual agents has the potential to increase scholarly output. These tools can help in the creation of questionnaires. These agents have been used in medicine to organize material and generate and correct texts [7]. Thus, virtual agents have potential benefits for the academy, such as the generation of exams; however, their limited interpretation capacity, biases, and errors must be considered, so it is essential to verify the information and use a critical approach to evaluate its content. Therefore, this work aims to evaluate the use of virtual agents based on GPT language models to create diagnostic tests for biomedical engineering courses. This will be done by analyzing the quality of the questions generated.

There are multiple reports exploring the use of ChatGPT to answer quizzes. Although some authors propose the creation of questions, only one have tested this hypothesis by creating multiple-choice questions. The study compared various virtual agents to generate reasoning-based multiple-choice questions in medical physiology [8]. The authors found that virtual agents require improvement and that each tool had limitations, as Bing developed less valid and ChatGPT less complex questions. This work contributes to the field of biomedical engineering education by exploring the potential of virtual agents in creating personalized diagnostic tests at different cognitive levels. The work validates the quality of the exam and considers practical issues. The findings of the article could improve assessment methods in biomedical engineering education to improve the quality of education for students in this field.

2 Methodology

The methodology used in this study was based on the ChatGPT-3 virtual agent due to its more equitable and unrestricted access compared to more recent models such as ChatGPT-4. The study aimed to create a 20-question multiple-choice diagnostic exam for biomedical engineering students in their final year before the biomedical measurements course. The biomedical measurements course focuses on designing instruments for electrophysiological signal measurements to support the clinical diagnosis of various pathologies. The course is part of the curriculum line of applied engineering. The questions evaluated topics related to the acquisition, amplification, processing, and visualization of biomedical signals, which consider different levels of thinking according to the taxonomy of Bloom, including the application, analysis, and evaluation levels. An example of a prompt for ChatGPT-3 to generate questions of analysis level is shown in Fig. 1.

Create a diagnostic exam for biomedical engineering undergraduate students with ten multiple-choice questions to evaluate Bloom's Cognitive Dimension of Analyze. Each question has three options; you must indicate the correct answer. Cover the topics of acquisition, amplification, processing, display of biosignals

Fig. 1. Example of a prompt for ChatGPT-3 to generate questions of analysis level.

The exam creation process began with a researcher reviewing the content and selecting twenty out of thirty questions generated by the virtual agent based on their clarity, relevance, level of thinking, and difficulty. Subsequently, the selected questions were captured in a question bank using the Moodle learning platform. With this bank of questions, the exam was created, where each question and its response options were presented individually and sequentially, with a maximum resolution time of 20 min.

The exam was applied to all students in the last year of a degree program in biomedical engineering, certified according to CACEI. The quality of each question was evaluated by an external sample of academic experts in the subject who had training in biomedical engineering and at least five years of teaching experience related to the topics to be evaluated. The consistency of the evaluators was calculated using the Intraclass correlation coefficient (ICC).

Academics were asked to rate the clarity, validity and relevance, level of thinking, and difficulty of each question and answer through a survey using a standardized form in Word. The document contained a supporting explanation for each domain to be assessed. Clarity and relevance were assessed using a dichotomous variable (yes/no), the thinking level was classified according to the levels of Bloom (application, analysis, or evaluation), and the difficulty was classified as low, medium, or high.

General descriptive statistics were performed for test subjects and scores. The difficulty and discrimination indexes were calculated for each question. The Rasch analysis was carried out in the free software Jamovi [9], where the reliability of the subjects, the p-value for the model fit, the correlation matrix of all the exam questions, and the skills map of Wright for the test were calculated. The difficulty index measures the difficulty of the question calculated by dividing the number of students who answered the question correctly by the total number of students. The discrimination index measures the ability

of questions to distinguish between high and low-achieving students. Subject reliability is a measure of the consistency of students in their responses. The question correlation matrix shows the correlation between all questions and may indicate that a pair of questions measure the same thing. The Wright Skill Map is a graphical representation of the relationship between student skills and the difficulty of test questions. The ability and error in estimating item ability measures how well the Rasch model estimates the skills of students. The fit of the answer to each question (infit) and the general fit of all the answers (outfit) to the model indicates how well the question fits the model.

3 Results

The 20-question diagnostic exam was applied to 39 eighth-semester biomedical engineering Spanish-speaking students (see Fig. 2).

Question ID	Description
A01	Question about types of biosignal acquisition systems.
A02	Inquiry about the definition of gain in the context of an amplifier.
A03	Question about digital signal processing filters.
A04	Question about the purpose of signal averaging in biosignal processing.
A05	Question about display methods for electrocardiogram signals.
A06	Inquiry about biosignal amplifiers.
A07	Question about the Nyquist frequency.
A08	Exploration of methods of biosignal processing.
A09	Question about the characteristics of biosignal acquisition systems.
A10	Inquiry about types of biosignal processing techniques.
Ap01	Question about the component responsible for converting analog signals to digital in a biosignal acquisition system.
Ap02	Question about a method to reduce noise in biosignals.
Ap03	Exploration of the purpose of amplification in biosignal processing.
Ap04	Question about the primary function of a biosignal processing system.
Ap05	Inquiry about a type of biosignal that requires a high sampling rate for accurate acquisition.
E01	Question about the primary purpose of signal amplification in biomedical signal processing.
E02	Exploration of disadvantages of using digital filters for signal processing.
E03	Question about factors considered when selecting the appropriate sampling rate for a biosignal.
E04	Inquiry about the type of electrode commonly used for recording electroencephalography (EEG) signals.
E05	Question about the most critical parameter to consider when selecting a filter for biosignal processing.

Fig. 2. Description of the 20 questions obtained by ChatGPT for the diagnostic exam. Analysis, application, and evaluation questions are coded as A, Ap, and E.

The average grade obtained on the exam was 5.91, with a standard deviation of 1.39, in a total score range of 3.00 to 9.00 points. The exam was designed to assign half a point to each correct answer with a maximum total score of 10 points. The exam was written in English. Some students needed help with an English written exam. Students were helped by paraphrasing and using synonyms to clarify questions. All questions were classified as relevant by the three experts. Nine questions were related to the amplification topic, five to acquisition and processing, and one to display. Of the 20 questions, ten selected questions corresponded to the analysis level of thinking (A01 to A10) and five to the

application and evaluation levels (Ap01 to Ap05 and E01 to E05, see Fig. 3). There was no consistency among evaluators regarding the clarity and level of thinking of questions ($ICC < 0.01$, $p\text{-value} = 0.48$). However, in most of the questions, at least two experts agreed on the level of thinking ($N = 15$) and clarity ($N = 14$). On the other hand, a moderate agreement ($ICC = 0.55$, $p\text{-value} < 0.01$) was reached between experts for the difficulty of the questions.

Question ID/ Experts	Clarity			Level of thinking			Difficulty		
	1	2	3	1	2	3	1	2	3
A01	Yes	Yes	Yes	Applying	Evaluating	Applying	low	low	low
A02	Yes	Yes	Yes	Analyzing	Evaluating	Analyzing	low	low	medium
A03	Yes	Yes	Yes	Applying	Evaluating	Applying	low	low	low
A04	Yes	Yes	Yes	Analyzing	Evaluating	Analyzing	medium	medium	medium
A05	Yes	No	Yes	Applying	Analyzing	Analyzing	low	high	medium
A06	Yes	No	Yes	Applying	Analyzing	Applying	low	low	low
A07	Yes	No	Yes	Applying	Evaluating	Applying	medium	medium	low
A08	Yes	Yes	Yes	Analyzing	Evaluating	Applying	low	medium	low
A09	Yes	Yes	Yes	Applying	Evaluating	Analyzing	low	medium	medium
A10	Yes	No	Yes	Applying	Analyzing	Analyzing	low	high	medium
Ap01	Yes	Yes	Yes	Applying	Analyzing	Applying	low	medium	low
Ap02	Yes	No	Yes	Applying	Evaluating	Analyzing	low	low	medium
Ap03	Yes	Yes	Yes	Analyzing	Evaluating	Applying	medium	medium	low
Ap04	Yes	Yes	Yes	Analyzing	Analyzing	Analyzing	medium	high	medium
Ap05	Yes	Yes	Yes	Analyzing	Analyzing	Analyzing	high	high	high
E01	Yes	Yes	Yes	Applying	Analyzing	Applying	low	high	low
E02	Yes	Yes	Yes	Analyzing	Analyzing	Analyzing	high	high	high
E03	Yes	No	Yes	Analyzing	Analyzing	Analyzing	high	high	high
E04	Yes	Yes	Yes	Applying	Evaluating	Applying	low	low	low
E05	Yes	Yes	Yes	Analyzing	Analyzing	Analyzing	medium	high	high

Fig. 3. Analysis by experts of the validity: clarity, level of thinking, and difficulty of each question. Analysis, application, and evaluation questions are coded as A, Ap, and E.

All students answered one knowledge application question correctly (Ap02) and was therefore omitted from the Rasch analysis. Rasch analysis revealed a subject reliability of 0.599 and a p-value for the model fit of 0.017, indicating acceptable consistency in student responses. The correlation matrix showed a few questions with a strong correlation between them (see Fig. 4). The highest correlation was found between knowledge application questions (A01 and A10), while another knowledge application question (A08) presented several high correlations (>0.3) with three questions of the same thinking level (A04, A05, and A06).

The diagnostic exam questions presented an average difficulty index of 0.57, with a standard deviation of 0.27 (see Fig. 5). However, according to this index, most of the questions (12 in total) could be considered invalid or irrelevant because they were straightforward (>0.7) or very difficult (<0.3). On the other hand, the average ability of the students was -0.47 , with a standard deviation of 1.66. The average infit and outfit values were close to 1, with standard deviations of 0.09 and 0.31, respectively. However, one analysis question (A10) was identified that presented a significant outfit deviation (2.040), suggesting the presence of unexpected answers, such as a low-ability student

	A01	E01	E02	E03	E04	E05	A02	A03	A04	A05	A06	A07	A08	A09	A10	Ap01	Ap03	Ap04	Ap05	
A01	—																			
E01	0.27	—																		
E02	-0.38	0.17	—																	
E03	0.04	-0.03	0.05	—																
E04	0.26	0.13	-0.08	0.04	—															
E05	-0.21	-0.16	0.14	0.10	-0.14	—														
A02	0.09	-0.01	-0.26	0.01	-0.19	0.21	—													
A03	-0.11	0.01	-0.10	0.09	0.03	0.21	0.01	—												
A04	0.14	0.07	-0.05	0.06	-0.20	0.17	0.35	-0.21	—											
A05	-0.21	-0.04	0.22	0.18	-0.17	0.09	-0.46	0.17	-0.17	—										
A06	0.24	0.04	-0.09	-0.14	0.35	0.03	0.09	-0.04	0.08	-0.29	—									
A07	0.12	-0.13	0.11	-0.23	-0.01	-0.45	-0.06	-0.23	-0.24	-0.06	0.25	—								
A08	-0.25	-0.21	0.25	0.06	0.05	0.33	-0.07	0.24	-0.37	0.32	-0.36	0.02	—							
A09	-0.08	-0.14	0.14	0.11	-0.11	0.07	-0.14	0.00	-0.15	0.21	-0.17	-0.34	-0.08	—						
A10	-0.56	-0.34	0.05	0.03	-0.26	0.01	0.19	-0.19	-0.14	0.05	-0.27	-0.06	0.14	0.13	—					
Ap01	-0.10	0.32	-0.29	0.02	0.31	0.17	-0.21	0.54	-0.13	0.15	0.27	-0.13	0.10	-0.21	-0.41	—				
Ap03	0.30	0.54	-0.23	-0.08	0.00	-0.11	0.25	-0.19	0.29	-0.14	0.25	0.08	-0.30	-0.13	-0.12	0.02	—			
Ap04	0.10	0.08	-0.19	0.07	-0.09	-0.32	-0.06	-0.12	-0.12	-0.17	0.05	0.26	-0.34	0.09	-0.06	0.09	0.16	—		
Ap05	-0.22	-0.04	0.03	-0.15	-0.43	-0.06	-0.28	-0.09	0.10	-0.06	-0.44	-0.25	-0.16	0.10	0.06	-0.14	-0.25	-0.03	—	

Fig. 4. Correlation Matrix of Rasch analysis.

obtaining a high score on a difficult question. To evaluate the discrimination index, the sample of students was divided into a group with the worst performance (36%, $N = 14$) and a group with the best performance (33%, $N = 13$) based on the total scores obtained. The average discrimination index was 0.19, with a standard deviation of 0.14.

The competency map of Wright showed a good distribution throughout the competency scale, with more significant clustering around the mean (see Fig. 6). Redundancy was also observed in four questions of the same dimension (E02 and E03, A04 and A05, E01, and E04, A03 and A09), as well as in two pairs of questions of different dimensions (A02 and Ap05, A08 and Ap04). In addition, gaps in the scale were identified around extreme difficulty levels (± 2), which suggests that the questionnaire might not have sufficient capacity to discriminate students with extreme performance. No groupings of students were observed at the extremes that could be attributed to the ceiling or floor effects of the questionnaire.

4 Discussion

The questions generated by ChatGPT followed a normal difficulty index distribution with an average close to the ideal value of 0.5. Thirty percent of the questions had a discrimination index above the excellent value of 0.3. Subject confidence was slightly less than the ideal value 0.8. The p-value for model fit was lower than the excellent value of 0.05. Some correlations between the questions were higher than the ideal value of 0.3. The person-question map of Wright followed the perfect smooth curve with no outliers. Only one question had an outfit value above the ideal range of 1.3. A previous investigation showed that ChatGPT generated easy questions for medical physiology based on the assessment of experts using a 3-point scale [8]. This result is lower than the median rate of difficulty assigned by our experts using a similar scale. Our results showed no consistency among evaluators regarding the clarity and level of thinking of questions. Training experts and the supporting explanation included in the document for evaluation could improve those aspects. For future work, index values for Moodle could also be integrated for analysis.

Item	Difficulty Index	Ability	Ability S.E.	Infit	Outfit	Discrimination Index
A01	0.97	-3.86	1.02	1.02	0.95	0.00
E01	0.77	-1.33	0.40	0.92	0.84	0.26
E02	0.21	1.49	0.41	0.97	1.10	0.21
E03	0.21	1.49	0.41	0.89	0.81	0.31
E04	0.74	-1.18	0.38	1.05	1.05	0.05
E05	0.62	-0.53	0.35	0.90	0.86	0.41
A02	0.33	0.77	0.36	0.98	1.02	0.31
A03	0.87	-2.09	0.49	0.95	0.82	0.15
A04	0.54	-0.18	0.34	0.98	0.98	0.36
A05	0.56	-0.29	0.34	0.99	0.97	0.26
A06	0.72	-1.04	0.37	0.98	0.94	0.26
A07	0.44	0.28	0.34	1.12	1.16	0.10
A08	0.41	0.40	0.34	0.99	1.00	0.15
A09	0.87	-2.09	0.49	1.04	0.99	0.10
A10	0.08	2.69	0.61	1.06	2.04	-0.05
Ap01	0.95	-3.13	0.74	0.91	0.50	0.05
Ap03	0.79	-1.49	0.41	0.89	0.75	0.36
Ap04	0.41	0.40	0.34	1.02	1.01	0.31
Ap05	0.33	0.77	0.36	1.28	1.41	-0.05

Fig. 5. Analysis of the difficulty index, ability measurements, standard errors of ability estimation, infit and outfit values, and the discrimination index for each question in the diagnostic exam. The questions are categorized by their respective code, including ‘A’ for analysis, ‘E’ for evaluation, and ‘Ap’ for application questions.

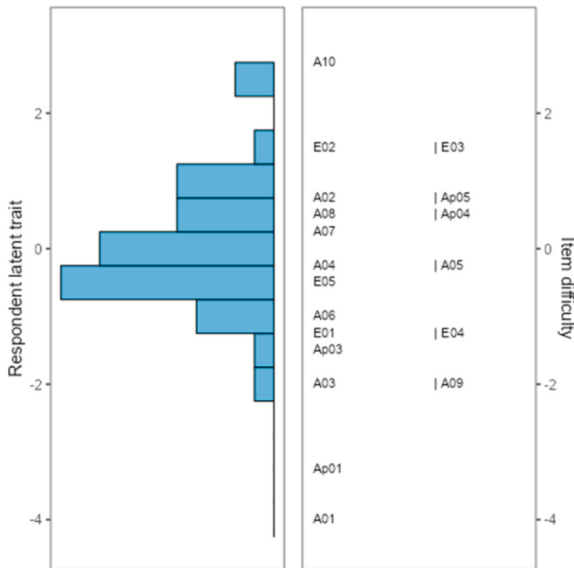


Fig. 6. Person-question map of Wright shows the relative position of questions and students. The Scale increases vertically from less capable student less capable/more straight-forward question to more skilled/more difficult question.

The use of virtual agents in academia raises validity issues. To ensure the responsible use of these tools, it is necessary to reach a consensus on regulating their use to prevent possible abuse and guarantee academic integrity. Virtual agents should not be used for cheating but to enhance personalized learning and provide new opportunities for

scholars and students. Teachers can integrate these tools into their teaching by creating new exercises that encourage critical thinking and problem-based learning, using AI to support and enrich the educational process [3]. Potential advantages of virtual agents for creating exams are their availability at any time, scalability to respond to significant work demands in less time and effortlessly, and customization to individual needs with high variability.

Our results showed that it is necessary to conduct tests and studies to determine the quality of the information these tools provide. Academics and students should always verify the content validity of the questions with these tools. Scholars and students who use virtual agents must be transparent about their use and take responsibility for the veracity of the information [10, 11]. As we can see, more work is still needed to improve the performance of virtual agents for academic use, [4] especially in specific areas such as biomedical engineering. Finally, we consider it essential to educate teachers and students about the limitations of virtual agents and how to use them effectively [3]. Giving scholars the authority and independence to use them is also essential [12]. We also consider it necessary to guarantee that everyone has the same access to these technologies [13].

5 Conclusion

In conclusion, using artificial intelligence and virtual agents based on GPT language models in academia raises several practical and ethical considerations. Implementing clear guidelines, a code of ethics, and consensus in regulating its use are required. Academics must be transparent in their use, assuming responsibility for the information generated. Literacy in this area, constant evaluation, and continuous improvement are essential to maximize the benefits and minimize the risks associated with these technologies.

References

1. D'Amico, R.S., White, T.G., Shah, H.A., et al.: I asked a ChatGPT to write an editorial about how we can incorporate chatbots into neurosurgical research and patient care. *Neurosurgery* **92**, 663–664 (2023)
2. Sallam, M.: ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare* **11**, 887 (2023)
3. Yang, H.: How I use ChatGPT responsibly in my teaching. *Nature* (2023). <https://doi.org/10.1038/d41586-023-01026-9>
4. Das, D., Kumar, N., Longjam, L.A., et al.: Assessing the capability of ChatGPT in answering first- and second-order knowledge questions on microbiology as per competency-based medical education curriculum. *Cureus* **15**, 3–11 (2023)
5. Ariyaratne, S., Iyengar, K.P., Nischal, N., et al.: A comparison of ChatGPT-generated articles with human-written articles. *Skelet. Radiol.* **52**, 1755–1758 (2023)
6. Eysenbach, G.: The role of ChatGPT, generative language models, and artificial intelligence in medical education: a conversation with ChatGPT and a call for papers. *JMIR Med. Educ.* **9**, e46885 (2023)

7. Sedaghat, S.: Early applications of ChatGPT in medical practice, education and research. *Clin. Med. (Northfield Il)* **23**, 278–279 (2023)
8. Agarwal, M., Sharma, P., Goswami, A.: Analysing the applicability of ChatGPT, bard, and bing to generate reasoning-based multiple-choice questions in medical physiology. *Cureus* (2023). <https://doi.org/10.7759/cureus.40977>
9. The Jamovi Project. Jamovi (2023). <https://www.jamovi.org>
10. Gurha, P., Ishaq, N., Marian, A.J.: ChatGPT and other artificial intelligence chatbots and biomedical writing. *J. Cardiovasc. Aging* 3–5 (2023)
11. Hosseini, M., Horbach, S.P.J.M.: Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other large language models in scholarly peer review. *Res. Integr.* 1–14 (2023)
12. Anders, B.A.: Is using ChatGPT cheating, plagiarism, both, neither, or forward thinking? *Patterns* **4**, 100694 (2023)
13. Salvagno, M., Taccone, F.S., Gerli, A.G.: Can artificial intelligence help for scientific writing? *Crit. Care* **27**, 75 (2023)



Transcriptional Expression of Bioactive Antimicrobial Peptides with Biomedical Potential in Diverse Organs of the Mexican Axolotl

Edgar Cera-Domínguez , Gabriel Arenas-Ballesteros , Luis Varela-Rodríguez ,
Javier Camarillo-Cisneros , Abimael Guzman-Pando ,
and Hugo Varela-Rodríguez  

Universidad Autónoma de Chihuahua, Circuito Universitario Campus II, Chihuahua,
31125 Chihuahua, México
hvrodriquez@uach.mx

Abstract. A notable challenge in modern biomedicine is the development of innovative therapies that mitigate resistance to conventional antibiotics. Antimicrobial peptides (AMPs), represent a promising source of new antibiotics due to their broad-spectrum activity and low incidence of resistance development, even against new infectious diseases. In nature, anuran amphibians (such as frogs and toads) are the largest source of recorded AMPs. However, salamanders and newts (such as the Mexican axolotl) have received less attention despite the fact that some promising bioactive peptides have been detected in their skin secretions. The axolotl is a critically endangered caudate amphibian endemic from Xochimilco's Lake in Mexico, which is considered a biological model par excellence for genomic studies related to salamanders. In the present study, the wide availability of genomic resources of the axolotl was used to investigate the presence of possible new AMPs with biomedical potential, whose findings could be also extensive to other salamander species unexplored. In the axolotl, seventeen different coding transcripts for presumptive antimicrobial peptides, such as Leap2, Cathelicidins, β -Defensin, Hepcidins, Transferrins, and Cystatin-C were identified, whose transcriptional expression is mainly concentrated in the liver and spleen. This could potentially foster the development of new pharmacological treatments in the future.

Keywords: Axolotl · Transcriptomics · Antimicrobial Peptides

1 Introduction

Antimicrobial peptides (AMPs) are organic polymers of amino acids that are produced endogenously by numerous multicellular metazoans in order to protect the host from pathogenic microorganisms [1]. Therefore, AMPs are also known as defense peptides due to their involvement in innate immunity [1]. Among the various characteristics described for AMPs, there is a relatively short length (\sim 12 to 100 residues), a net positive charge ranging from +2 to +9, as well as a certain amphipathicity that enables them

to dissolve in aqueous environments [1]. By exhibiting antimicrobial and immunomodulatory activities, AMPs are humoral factors able to activate immune response mechanisms such as phagocytosis, prostaglandin release, neutralization of lipopolysaccharide effects, immune cells recruitment at inflammatory sites, wound healing, and promotion of angiogenesis, between others [1]. The largest source of AMPs found in literature records are anuran amphibians, which account for ~30% (mainly frogs followed by toads), being some typical classes described of amphibian AMPs the β -Defensins, Cathelicidins, Hecpидins, Cystatins, Bombinins, Magainins, Dermaseptins, Esculentins, among many others [1, 2]. In fact, some classes of AMPs are characterized by containing disulfide bridges in their structure, being commonly expressed in epithelial cells and leukocytes. In addition, many AMPs also give to a prepropeptide consisting of a secretion signal peptide, a linker proregion or protease inhibitor domain, as well as a highly variable mature peptide segment that is responsible for the antimicrobial biological function [2]. Particularly, most of scientific attention for the identification and characterization of novel bioactive peptides have been directed towards the anuran families of Pipidae and Bombinatoridae, given their global distribution and degree of diversification [3]. Meanwhile, genera from the orders Caudata and Apoda, which include tailed and limbless amphibians respectively, have been largely overlooked in this regard [3].

The axolotl (*Ambystoma mexicanum*) is a representative specie of salamander (caudate amphibian) that has garnered significant attention as an emergent biological model organism [4]. Native from Mexico, the axolotl exhibits interesting characteristics, such as its regenerative ability, neoteny, scar-free healing, and robust immune system, which has led to the availability of extensive genomic resources, make it an invaluable biological model [4]. These features impact multiple disciplines, including regenerative biomedicine, developmental biology, wound healing, immunology, biomedical engineering, and genomic sciences [4]. Indeed, the axolotl genome has been sequenced, being one of the few complete sequenced genomes of salamanders to date, providing comprehensive resources to study the genes and regulatory elements involved in regeneration, and immunity related to AMPs [5]. While several countries regulate and protect certain amphibian species, such as the axolotl (by NOM-059-SEMARNAT-2010 in Mexico), it is now possible to obtain antimicrobial peptides (AMPs) through genomic information implementing heterologous expression or similar biotechnological systems [6]. This allows sustainability and bioavailability to be guaranteed without endangering amphibian biodiversity, which is already threatened. Thus, the rich genomic dataset offers an unprecedented opportunity to discover new representative antimicrobial peptides in salamanders and potentially translate them into therapeutic strategies for human health to future.

Given the evident increase in resistance to conventional drugs by pathogens, the scientific community has been forced to develop and improve new anti-infective agents. Since AMPs are one of the best alternatives to address this problem, the search and characterization of new AMPs has become imperative, with amphibians being a promising source of these peptides, as previously mentioned [7]. The role of antimicrobial peptides in the fight against antibiotic-resistant bacterial infections seems set to be a turning point in the coming decades [7]. Therefore, the present study aims to identify new antimicrobial peptides present in the Mexican axolotl, through *in silico* screenings

with bioinformatics tools and taking advantage of available genomic resources, which could represent a great alternative with biomedical potential in the pharmacological field, currently threatened due to increasing resistance to conventional drugs available.

2 Materials and Methods

2.1 Screening of Coding Transcripts for Bioactive Peptides in the Axolotl Transcriptome

We used online data from the UniProt repository to create a non-redundant reference database of bioactive antimicrobial peptides from amphibians. This database helped us to detect AMP-encoding axolotl transcripts via BLAST v2.12 with the options “task blastp-short -evalue 1e-3 -max_target_seqs 1” (bitscore \sim 28). The axolotl transcriptomes used in this study were obtained from the NCBI GenBank database with the following TSA (transcriptome shotgun assembly) accessions: GFBM010000000, and GFZP010000000. All axolotl transcripts were merged and redundancy reduced with CD-HIT-EST v4.8.1 under the parameters “-c 0.95 -g 1 -b 50 -n 10 -A 0.95” before being employed for AMPs screening. We then used the EMBOSS v6.6 package to identify and translate all possible six-frame open reading frames (ORFs) in the axolotl transcripts, by running getorf with the parameters “-table 1 -minsize 30 -find 3”. After, we applied transeq to convert the ORFs into protein sequences for BLAST searches, first discarding any sequences with non-standard undefined amino acids.

2.2 Annotation and Domain Architecture Analysis of Candidate Sequences with Potential Antimicrobial Activity

The annotation of the selected axolotl transcripts was done with Trinotate v4.0. The annotated amphibian protein sequences were classified according to their gene family with InterProScan v5, SMART v9.0, and ScanProsite v20.0 to determine the composition and structural organization of the characteristic protein domains. In addition, Cyscon was used for the inference of disulfide bridges and PeptideCutter for cleavage sites. The results of domain architecture identification for the determined AMP classes were plotted with IBS v2.0. The prediction of antimicrobial properties was done with AMPLify v1.1 (which is an attentive deep learning model), and AmpGram v1.1 (based on n-gram encoding and random forests), using the default parameters.

2.3 Transcript Expression Analysis with Axolotl Transcriptomic Data

Raw mRNA-seq data publicly available at NCBI was acquired to study the gene expression patterns of previously identified axolotl AMPs in different organs and excised segments of structures. The corresponding accessions of the SRA studies used were SRP093628 and SRP065567. Quality profiling was done with FastQC v0.12.1. For adapter trimming and read filtering, the BBTools v37.62 BBDuck was used with “ktrim=r k=23 mink=11 hdist=1 tpe=t tbo=t qtrim=r1 trimq=20 minlen=15 forcetrimleft=14”, and BBMap with “maxindel=1 minid=0.95”. RiboDetector v0.2.7 was employed to discard

rRNA reads from data set. Then, sequencing reads were mapped to the axolotl transcriptome with Bowtie2 v2.5 using the parameters “--sensitive --dpad 0 --gbar 99999999 --mp 1,1 --score-min L,0,-0.1 --no-mixed --no-discordant -k 200 -I 1 -X 1000”. RSEM v1.3.1 was after assessed for estimating gene expression levels by expected counts with EM algorithm to mapped reads. Subsequently, abundance analysis was performed with EdgeR obtaining the counts per million (CPM) for axolotl AMP transcripts, and calculating the Fold Change based on housekeeping transcript abundance of *amx-odc-1*. Results were graphed in a bar plot and a heatmap with the libraries gplots v3.1.3 and RColorBrewer v1.1 of R v4.1.2.

3 Results and Discussion

In the Mexican axolotl transcriptome, seventeen transcripts were identified encoding for at least 3 different classes of Transferrins (one Melanotransferrin, one Saxiphilin and one Serotransferrin), four Hepcidins, two Leap2 (Liver-Enriched Antimicrobial Peptide 2), six Cathelicidins, one Cystatin-C, and one β -Defensin. Through a domain architecture analysis (see Fig. 1), it was possible to verify the presence of characteristic functional peptide domains, as well as the presence of signal peptides. In this sense, it was observed that the vast majority of the axolotl bioactive peptides analyzed have a clear signal peptide of the Sec/SPI type, a standard secretory signal through the Sec pathway found in the endoplasmic reticulum and which directs the insertion of membrane proteins; destined in turn to the vesicle sorting pathway and which are cleaved by a signal peptidase I [8].

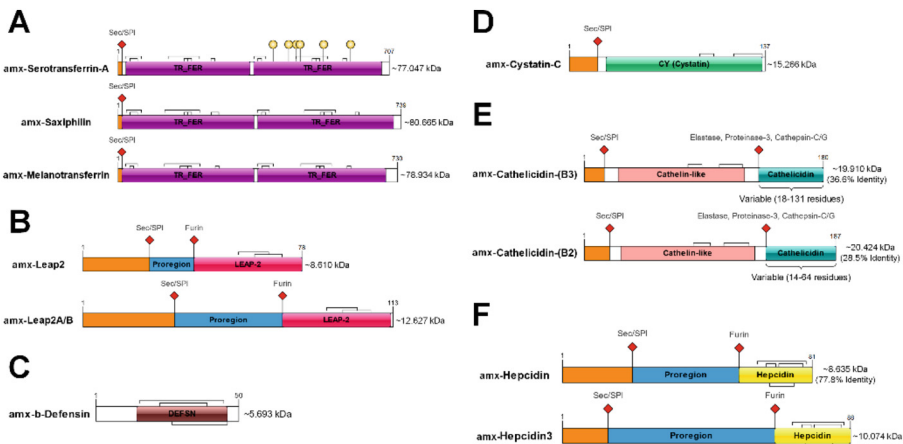


Fig. 1. Domains organization schemes for sequences translated of axolotl transcripts identified encoding antimicrobial peptides; where domains are distinguished with different colors and names; predicted secretory signal in orange; red diamonds represent cleavage sites; yellow spheres correspond to binding sites. (A) Transferrins. (B) Liver-Enriched Antimicrobial Peptide 2. (C) β -Defensin. (D) Cystatin-C. (E) Cathelicidins. (F) Hepcidins.

Furthermore, the size range of the predicted axolotl antimicrobial proteins was between approximately 5 and 20 kDa, with exception of Transferrin proteins of around

~80 kDa, all as expected [7]. Among the most outstanding findings, the presence of multiple axolotl paralogs for the Leap2 and Hecpudin classes can be mentioned. These are proteins intimately related to iron homeostasis (together with Transferrins), which display a classical tripartite organization of a signal peptide, with a linker variable pro-region, and the functional peptide cleaved by a Furin during prepropeptide maturation [9, 10]. However, unlike what happened in mammals, there are at least two different paralogs for each class of Leap2 and Hecpudin (see Fig. 1, B and F panels), more similar to what is observed in fish and reptile species [9, 10]. In the case of Transferrins, these axolotl proteins showed two transferrin domains that are known to form two homologous lobes called N-lobe and C-lobe (see Fig. 1, panel A), the latter domain being the most important due to its high affinity for Fe^{3+} binding, as can be seen in Serotransferrin [11]. On the other hand, the axolotl β -Defensin and Cystatin-C identified (see Fig. 1, C and D panels) had three and two predicted disulfide bridges respectively, as reported in other studies [12, 13]. For axolotl Cathelicidins (see Fig. 1, panel E), these proteins presented a classical architecture with a secretory signal, followed by a Cathelin-like domain related to Cathepsin-L-like inhibition, commonly cleaved by Elastase or Proteinase-3 proteases to give rise a mature peptide highly variable in length [14].

Table 1. Antimicrobial activity prediction for protein sequences identified in axolotl.

Sequence ID	Probability score (AMPlify)	Probability score (AmpGram)
amx-b-Defensin	0.4764318	0.987373016
amx-Cathelicidin-B2-1	0.6586954	0.403123016
amx-Cathelicidin-B2-3	0.7404948	0.872134127
amx-Cathelicidin-B3-1	0.37388664	0.944224603
amx-Cathelicidin-B3-2	0.9541861	0.976533622
amx-Cathelicidin-B3-3	0.34033796	0.950809524
amx-Cathelicidin-B3-4	0.9080199	0.909112698
amx-Cystatin-C	0.23207442	0.513125397
amx-Hecpudin3-A	0.98758775	0.99975
amx-Hecpudin-A	0.8232733	1
amx-Hecpudin-B	0.8644066	1
amx-Hecpudin-C	0.84842086	0.9995
amx-Leap2	0.987711	0.21107619
amx-Leap2AB	0.99623394	0.989477778

Sequences with a cut-off value equal to or greater than 0.5 for both cases were highlighted with bold letters. The probability scores given by the prediction algorithms range from 0 to 1 only. Due to limitations of the AMPlify algorithm, Transferrins were omitted in the table.

When carrying out the prediction of antimicrobial activity with AMPLify and AmpGram [15, 16] for the annotated protein sequences of the axolotl, a high probability score was denoted for the mature forms of Hecpидins, as well as for some Cathelicidins and Leap2 (see Table 1). In other studies, some *in vitro* antibacterial activity has also

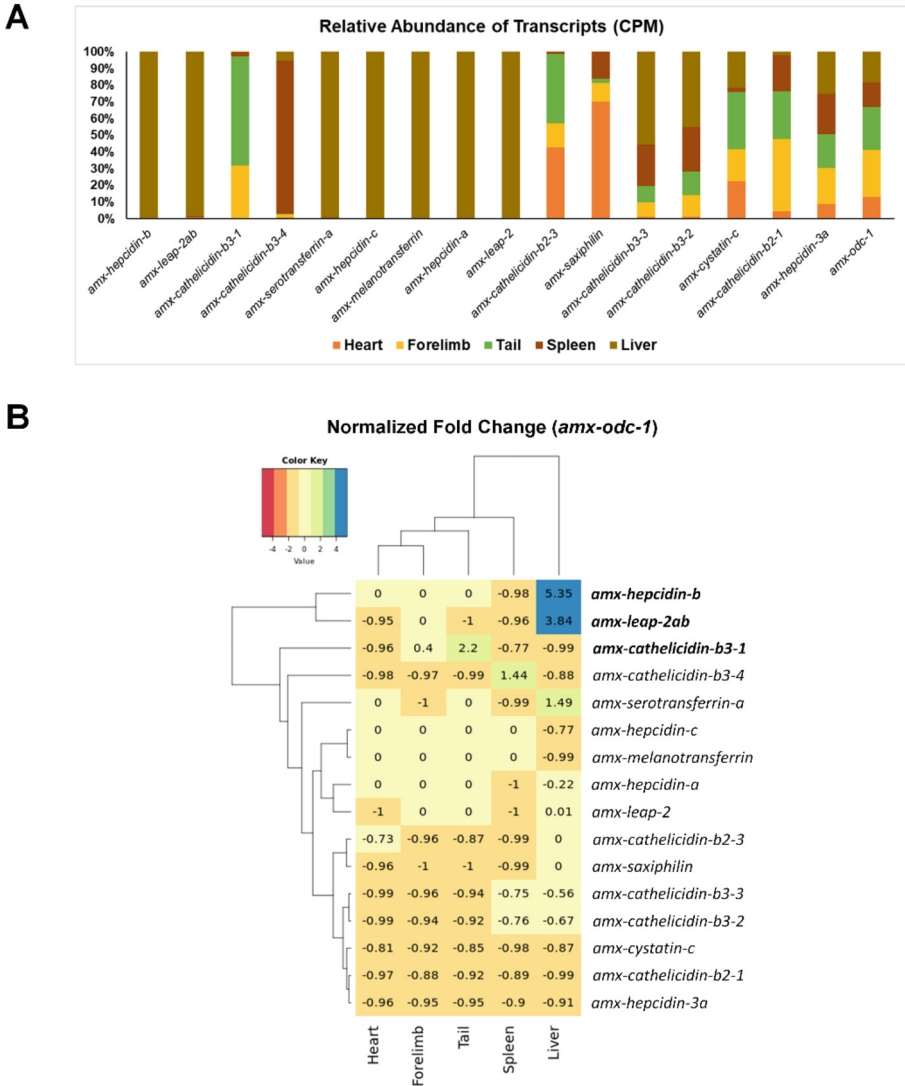


Fig. 2. Transcript expression profiling of axolotl sequences identified as presumptive AMPs. (A) Relative abundance in counts per million (CPM) for transcripts expressed in different organs, including biological structure segments. (B) Fold Change normalized with housekeeping *amx-odc-1* abundance of axolotl transcripts and evidencing their similarities to each other. Fold changes (above 2) that were significant were highlighted in bold.

been reported for similar amphibian Hepcidins, Cathelicidins, and Leap2, which is also consistent with our predictions shown [17].

The axolotl organs and structure portions that revealed a high abundance of transcribed AMPs were the liver and spleen, followed by the tail (see Fig. 2, A and B panels). It has been documented in previous works that the vast majority of the regulatory factors of iron homeostasis in mammals are mainly produced by the liver, as was also evidenced by the abundance of messenger RNAs coding for axolotl *leap-2*, and several of *hepcidins* and *transferrins* [9–11]. Likewise, an organ that drew particular attention is the spleen. This organ, being histologically made up of red and white pulp, the latter abundant in leukocytes, was expected to show a high abundance of transcripts for several *cathelicidins* in the axolotl (see Fig. 2, panel A), since *cathelicidin* transcripts are usually expressed in epithelium and white cells of the immune system [14]. This transcriptional expression profiling, referring to the two previously contrasted conditions, allows us to note that the same trend is evolutionarily conserved even up to amphibians. Particularly, a significant change in abundance of the transcripts encoding *hepcidin-b* and *leap-2ab* is observed (see Fig. 2, panel B). More additional analyzes would be needed to unravel functional differences to the canonical ones described for the different paralogs identified in the present study.

4 Conclusion

Through a conservative comparative approach, using only public genomic resources, the identification of seventeen potential antimicrobial bioactive peptides in the axolotl was achieved. Furthermore, the structural organization of domains for the vast majority of the proteins encoded in the identified axolotl transcripts was observed to be highly conserved with respect to their counterparts reported in other species. Although predictions of antimicrobial activity are promising for many of the axolotl peptide sequences found, further exploratory analyzes are required to confirm this finding. Finally, the abundance of the axolotl transcripts coding AMPs analyzed in the transcriptional expression profiling, allowed the liver and spleen to be highlighted as important biological sources, information that could contribute to the isolation of such sequences or their products for future assays.









References

1. Jenssen, H., Hamill, P., Hancock, R.E.: Peptide antimicrobial agents. *Clin. Microbiol. Rev.* **19**(3), 491–511 (2006)
2. Olascoaga Del Angel, K., et al.: Antimicrobial peptides, a promising alternative for the treatment of infectious diseases. *Gac. Med. Mex.* **154**, 8 (2018)
3. Raaymakers, C., et al.: A new family of diverse skin peptides from the microhylid frog genus *phrynomantis*. *Molecules* **25**(4), 912 (2020)
4. Bölük, A., Yavuz, M., Demircan, T.: Axolotl: a resourceful vertebrate model for regeneration and beyond. *Dev. Dyn.* **251**(12), 1914–1933 (2022)
5. Varela-Rodríguez, H., et al.: Functional characterization of the *lin28/let-7* circuit during forelimb regeneration in *ambystoma mexicanum* and its influence on metabolic reprogramming. *Front. Cell Dev. Biol.* **8**, 562940 (2020)

6. Deng, T., et al.: The heterologous expression strategies of antimicrobial peptides in microbial systems. *Protein Expr. Purif.* **140**, 52–59 (2017)
7. Xiao, Y., Liu, C., Lai, R.: Antimicrobial peptides from amphibians. *Biomol. Concepts* **2**(1-2), 27–38 (2011)
8. Natale, P., Brüser, T., Driessen, A.J.: Sec and tat-mediated protein secretion across the bacterial cytoplasmic membrane distinct translocases and mechanisms. *Biochem. Biophys. Acta.* **1778**(9), 1735–1756 (2008)
9. Kim, C.H., Kim, E.J., Nam, Y.K.: Subfunctionalization and evolution of liver ex-pressed antimicrobial peptide 2 (LEAP2) isoform genes in Siberian sturgeon (*Acipenser baerii*), a primitive chondrosteian fish species. *Fish Shellfish Immunol.* **93**, 161–173 (2019)
10. Barroso, C., Carvalho, P., Nunes, M., Gonçalves, J.F.M., Rodrigues, P.N.S., Neves, J.V.: The Era of antimicrobial peptides: use of hepcidins to prevent or treat bacterial infections and iron disorders. *Front. Immunol.* **12**, 754437 (2021)
11. Mohd-Padil, H., Mohd-Adnan, A., Gabaldón, T.: Phylogenetic analyses uncover a novel clade of transferrin in nonmammalian vertebrates. *Mol. Biol. Evol.* **30**(4), 894–905 (2013)
12. Jurczak, P., Groves, P., Szymanska, A., Rodziewicz-Motowidlo, S.: Human cystatin C monomer, dimer, oligomer, and amyloid structures are related to health and disease. *FEBS Lett.* **590**(23), 4192–4201 (2016)
13. Correa, P.G., Oguiura, N.: Phylogenetic analysis of β -defensin like genes of bothrops crotalus lachesis snakes. *Toxicon* **69**, 65–74 (2013)
14. Kościuczuk, E.M., et al.: Cathelicidins: family of antimicrobial peptides a review. *Mol. Biol. Rep.* **39**(12), 10957–10970 (2012)
15. Li, C., Warren, R.L., Birol, I.: Models and data of amplify: a deep learning tool for antimicrobial peptide prediction. *BMC. Res. Notes* **16**(1), 11 (2023)
16. Burdukiewicz, M., et al.: Proteomic screening for prediction and design of antimicrobial peptides with amprgram. *Int. J. Mol. Sci.* **21**(12), 4310 (2020)
17. Vasconcelos, I.A., et al.: Salamanders and caecilians, neglected from the chemical point of view. *Toxin Rev.* **41**(4), 1304–1332 (2022)



Drugs and Therapeutic Targets in Breast Cancer: Physicochemical Analysis by Computational Chemistry

Aarón Morales Stockton¹ , Andrés Kaleb Aguiñaga Miramontes¹ ,
Raquel Portillo Saenzpardo¹ , Roxana Quintana Jurado¹ ,
Rodrigo Domínguez García² , Javier Camarillo Cisneros¹ ,
and Alexica Celine Marquez Barreto¹  

¹ Laboratorio de Física Química Computacional, Facultad de Medicina y Ciencias Biomédicas, Universidad Autónoma de Chihuahua, 31125 Chihuahua, Mexico

abarreto@uach.mx

² Centro de Investigación en Materiales Avanzados, S.C., Tecnología de la Información, 31136 Chihuahua, Mexico

Abstract. Breast cancer is the most common neoplasm and the leading cause of cancer death in women worldwide, followed by cervical cancer. The process of developing anticancer drugs focuses on synthesis and biological testing. However, it is a process that requires a large investment of resources. In the present work, the visualization of their most stable structure, that is, the one that requires less energy for its conformation, and the visualization of their molecular orbitals have been carried out in the analysis of some drugs commonly used in breast cancer patients. To generate useful information for the design and study of potential analogs to treat breast cancer. To evaluate the feasibility and properties of drug candidates, DFT allows the theoretical prediction of pharmacological properties. The HOMO and LUMO are particularly important in the study of drugs because they are related to properties such as light absorption and emission, electron-donating or accepting capacity, and the formation of chemical bonds with other molecules. Pharmacological activity and the design of new compounds depend on these properties.

Keywords: Computational chemistry · Breast cancer · Tamoxifen · Anastrozole · Lapatinib · Tucatinib · Capecitabine · Ifosfamide

1 Introduction

Breast cancer is a disease where cells in the breast tissue grow uncontrollably [1]. In 2020, the World Health Organization detected around 2 million new breast cancer cases, adding to an existing 7.8 million cases. This data confirms breast cancer is the most prevalent and common cancer worldwide [2, 3].

The breast is composed of three main parts: lobes that produce milk, ducts that carry milk to the nipple, and connective tissue that supports and surrounds all parts of the breast.

Breast cancer has several subtypes, including invasive ductal carcinoma, invasive lobular carcinoma, and inflammatory breast carcinoma. Other subtypes are diagnosed based on molecular characteristics and specific markers found in cancer cells. The first subtype is HR-positive, which indicates the expression of hormone receptors, estrogen, and/or progesterone in the cancer cells. The second subtype is HER2-positive, where the cancer cells overexpress the HER2 receptor. Finally, triple-negative cancers are those that do not express hormone receptors (ER-/PR-) or overexpress HER2. Cancer treatment options depend on the tumor type and may include drugs, hormonal and biological therapies, surgery, and radiation therapy.

Biological testing and synthesis are the heart and ground of the cancer drug development process; however, it is a long process that requires over 15 years and more than 800 million dollars without guaranteeing the test's success [4]. Therefore, integrating new technologies and reducing development times and costs would be necessary. An alternative option is computer-assisted drug design and testing which has proven to be a favorable implement in development cost and time [5]. Computational chemistry comprises the application of computation in the areas of chemistry, biology, and physics for the study of atoms, molecules, and macromolecules through modeling and representations in three dimensions [6, 7].

Specifically in breast cancer, computational chemistry has been used for the design of drug analogs to those already demonstrated to be effective. In a recent study, drug analogs were generated by modifying the functional groups of tamoxifen to improve the activity and selectivity of the drugs [8]. Using computer modeling techniques such as molecular dynamics, docking, and density functional theory (DFT), it has been possible to simulate molecular interactions and predict how a drug will interact with a specific protein or enzyme in the body. A review was published highlighting and demonstrating the efficacy of tucatinib and trastuzumab for HER2-positive breast cancer based on molecular docking results [11]. Also, another study in which a computational analysis of different candidate compounds for breast cancer was performed, obtained promising results for these compounds, designating them as good candidates for preclinical studies [10]. This saves time and resources by directly reducing the number of compounds that need to be synthesized and tested in laboratory experiments. In addition, computational science is used to predict the toxicity of compounds, helping to identify the potential side effects and risks associated with a drug before it is tested in animals or humans.

Although computational chemistry has been widely used in the study of drugs and their interactions with their therapeutic targets, including studies for breast cancer treatment, there is still a lack of physicochemical information on drugs that are effective against breast cancer and the chemical properties of their interactions with therapeutic targets. In the present work, the analysis of some drugs commonly used in breast cancer patients was performed, obtaining the visualization of their most stable structure and their molecular orbitals (HOMO-LUMO), to generate helpful information for the design and study of potential analogs drugs for the treatment of breast cancer using a computational chemistry methodology, reduce cost and investment time than the original drug design and study. To accomplish the objective of this study, a physicochemical characterization was performed, that is, to obtain chemical measurements to determine the properties of the drugs that are currently effective for breast cancer.

2 Methodology

The molecules used in this study were tamoxifen, anastrozole, lapatinib, tucatinib, capecitabine and ifosfamide. These are the most used molecules in the clinical treatment of the three types of cancer mentioned: HR, HER2, and triple negative. The full set of molecules shown in Fig. 1 was searched and downloaded from the open-access database PubChem, grouped according to the type of cancer being treated.

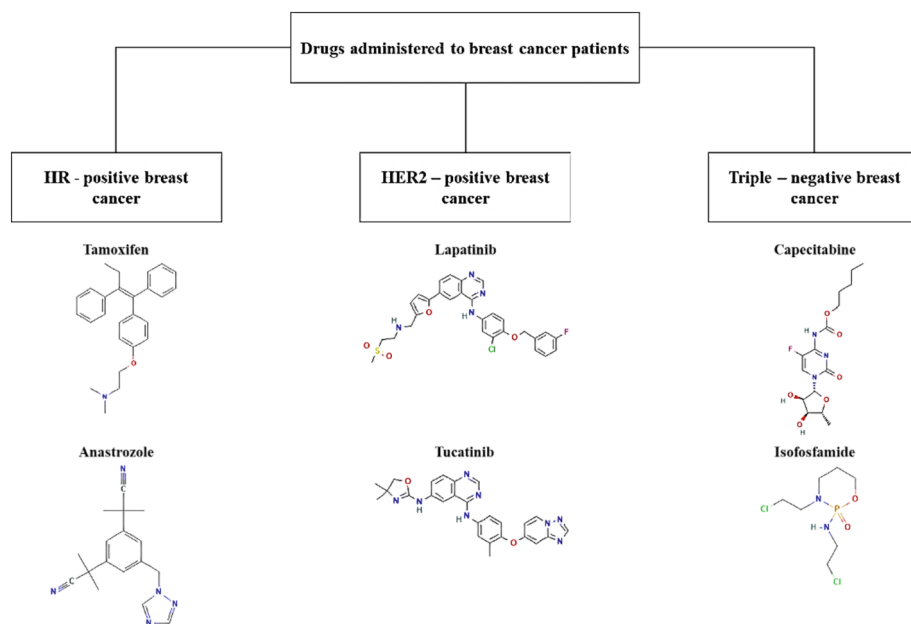


Fig. 1. Most common drugs administered to breast cancer patients.

The first set of calculations consisted of a geometric optimization of all the structures. The aim is to obtain the most energy-stable molecular structure. We first obtained the structure by molecular dynamics using Tinker software with optimization parameters of structural convergence of energy 2×10^{-5} kcal/mol, force 0.5 kcal/mol/Å, and maximum displacement 1×10^{-5} Å with a maximum of 500 iterations. After the atoms are positioned at a given coordinate, Molecular Dynamics calculates each of these quantities. It then performs the same energy and force calculation by changing the position and torsion angle of each atom. These two measurements show that for some pairs of atoms, there is an increase/decrease in force and energy. The software then looks for combinations that will decrease both the force between the pairs of atoms and the total energy of the system.

The second set of calculations is performed with the most stable molecule obtained by molecular dynamics. This next level of theory is the Density Functional Theory (DFT) method. The DFT method in quantum chemistry uses electron density to calculate the properties and energies of compounds. It provides information on electronic states and

energy transitions relevant to pharmacological activity. The Highest Occupied Molecular Orbital (HOMO) and the Lowest Unoccupied Molecular Orbital (LUMO) orbitals are of special interest at this level of theory as they indicate where a molecule can lose or give up electrons, and where it can accept electrons.

All calculations were performed on a desktop computer with an i7-processor with 32 GB of RAM and a dedicated video card Nvidia GTX-1080 Ti with 3584 CUDA cores.

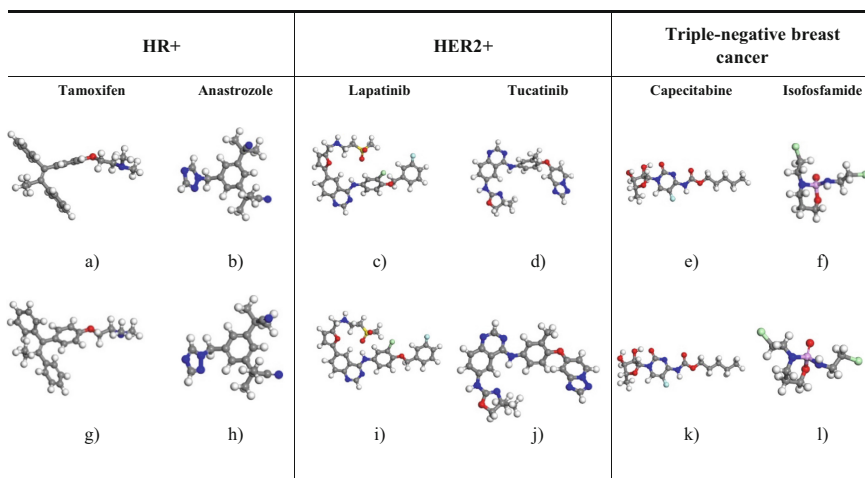
3 Results and Discussion

The set of six molecules is shown in Table 1. The first row (molecules a) to f)) corresponds to the molecules assembled from the Lewis structure and optimized at the molecular dynamics level. The energetically optimized structures by DFT are shown in the second row (molecules g) to l)). In terms of distances, the atom-atom variations were less than 5 Å in absolute terms, while the energy decreased by more than 2 eV.

Since our molecules are in an isolated state, it is easy to see that they are different from the state they are in reality. This is because the drugs in clinical use are presented in a medium that would modify our ideal structure due to the different temperatures, pH, enzymes, proteins, and H₂O that surround them and are not taken into account in the calculations. However, the molecules of the present work, after the mentioned energetic optimization, can be used to know each of their affinity and chemical selectivity towards therapeutic targets in a reliable way. The B3LYP functional we are considering has proven to be very powerful in the search for minimum energies in isolated molecules, so these structures can be used at this point to obtain their electronic structures and reliably compare them with each other. However, electronic quantities of interest, such as HOMO-LUMO, suffer from slight modifications in this isolation simulation compared to what would be expected in a real medium. It is the knowledge of the volumes of the HOMO-LUMO orbitals that allows us to move to the next point of analysis and use it for drug correlation.

Table 2 shows the electronic structures of the 6 molecules energetically obtained at the level of DFT theory. The interactions between the drugs and the therapeutic target would be expected to occur in the volumes framed by the yellow and blue clouds in the images. The chemical properties listed would also depend on two factors, the position of the HOMO and the LUMO. On one hand, the HOMO of a molecule is associated with the most energetic electrons and thus represents the orbital with the greatest capacity to donate electrons. On the other hand, the LUMO is associated with the lower energy electrons and can act as an electron acceptor or acceptor orbital. These properties are relevant to the ability of the molecule to participate in chemical bonding and electron transfer reactions.

For tamoxifen, the chemical interactions would be at the ends of the benzene or phenyl rings (six carbon atoms). For anastrozole, the phenyl and triazole rings (two N and two C atoms) would be the sites of chemical interactions. For the remaining four drugs, the same trend is found: the HOMO-LUMO orbitals are found around the rings, which in these cases create more complex structures between more than one ring. For the six drugs, the volumes are not found in the other functional groups, such as methyl, amino, and cyan, so it would be expected that they do not intervene in the chemical

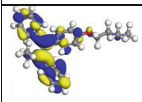
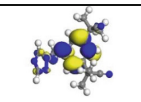
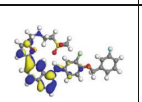
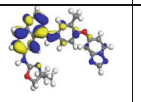
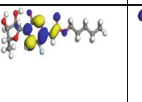
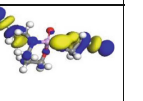
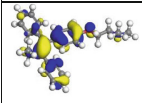
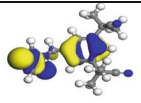
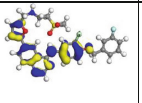
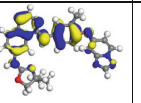
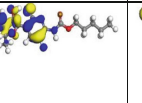
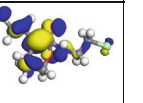
Table 1. Molecular dynamic optimization (a through f)) and DFT energetic optimization (g through l)) of drugs commonly used in breast cancer patients.

bonds of the drugs. Thus, it is only in the part of the molecule formed by the rings that the anchoring points with enzymes or proteins present in the cancer process will occur. Non-covalent interactions such as hydrogen bonding, electrostatic interactions, and pi-pi stacking would be some of the crucial mechanisms for molecular recognition and therapeutic activity in chemically reactive zones defined by HOMO-LUMO volumes.

The HOMO orbital can only donate electrons to orbitals with a lower energy modulus (negative values closer to zero) that are present in the therapeutic targets. The tendency of the molecule to donate electrons increases as the energy level of the HOMO decreases. For example, tamoxifen, lapatinib, and tucatinib, which have energies closer to zero, can donate electrons to more molecules than anastrozole. LUMO can be thought of as the orbital that can accept electrons or participate in electron transfer reactions. It is an indicator of a molecule's ability to accept electrons or act as an oxidant. The tendency of a molecule to accept electrons increases as the energy level of the LUMO decreases. Of our group of drugs, capecitabine has the highest oxidizing property and isofosfamide has the lowest. The difference in energy between the HOMO and LUMO, also known as the electronic gap, is also important. A smaller gap can indicate a higher reactivity and interaction capacity of a molecule.

As mentioned above, there are several computational studies of drugs used in cancer, however, the objectives of these studies are oriented to modify the properties to measure their efficacy, analyze interactions with other drugs or analyze the therapeutic activity of combined drugs, but there is no information available on the physicochemical characteristics of isolated drugs. An area of interest for this type of result is the development of targeted therapies since the drug and the therapeutic target must be known in detail to make the treatment more effective, however, a limitation to the present work would be that future cross-validations and comparisons with experimental data are important to support and verify the theoretical results. In addition, once the properties of the drugs are

Table 2. Molecular orbitals (HOMO-LUMO) of drugs commonly used in breast cancer patients.

Treatment for HR-positive breast cancer		Treatment for HER2 – positive breast cancer		Treatment for triple-negative breast cancer	
Tamoxifen	Anastrozole	Lapatinib	Tucatinib	Capecitabine	Isofosfamide
					
LUMO	LUMO	LUMO	LUMO	LUMO	LUMO
-0.0775 eV	-0.0717 eV	-0.0904 eV	-0.0923 eV	-0.1074 eV	-0.0208 eV
0.0976 eV	0.1697 eV	0.092 eV	0.0959 eV	0.1168 eV	0.2066 eV
HOMO	HOMO	HOMO	HOMO	HOMO	HOMO
-0.1751 eV	-0.2414 eV	-0.1824 eV	-0.1882 eV	-0.2242 eV	-0.2274 eV
					

known, it is possible to perform molecular docking studies to determine their interactions with the therapeutic targets.

4 Conclusions

Cancer has impacted the scientific community since its beginnings, not only because of the complexity of its treatment, which does not always guarantee a complete cure but also because of all the social effects that the fight against this disease entails. One of the types of cancer that is currently most present in humanity is breast cancer, which impacts many women of all ages and backgrounds every year, and even if all risk factors are treated, only 30% of the probability of suffering from this cancer is reduced; even if it's not present in the background of a lineage that precedes all patients. This motivates waves of fear and uncertainty in society, which at some point in their lives may be part of the statistics. This is why in modern times we find ourselves in the necessity of using all the tools and technology available to solve health problems as relevant as this one, to offer new effective and less invasive alternatives for breast cancer.

To accelerate drug discovery and development, improve efficiency, and reduce the costs associated with drug discovery and testing, computational science has provided powerful tools and methods. It is important to note, however, that computational science does not completely replace traditional methods. Rather, it is used in conjunction with them to enhance and complement the drug development process. It is important to note that DFT is used in conjunction with other experimental techniques and approaches, although it is a valuable tool in drug design. To ensure the efficacy and safety of the compounds being developed, drug design and development is a complex and multi-disciplinary process that requires the integration of information and data from multiple

sources. As mentioned before, knowing the drugs and properties that are currently effective against cancer provides relevant information for the search for analogous drugs or the design of new drugs with these same characteristics. HOMO and LUMO measurements in studies of molecules designed to treat breast cancer provide information about the reactivity, molecular interactions, and optical properties of the molecules that may be relevant to their therapeutic activity.






References

1. Organización Mundial de la Salud. Organización Mundial de la Salud. (2021). <https://www.who.int/es/news-room/fact-sheets/detail/breast-cancer>
2. Breast Cancer. Breast Cancer (2023). <https://www.breastcancer.org/es/datos-estadisticas>
3. Pashayan, N., Antoniou, A., Ivanus, U., Esserman, L., Easton, D., French, D., et al.: Personalized early detection and prevention of breast cancer: ENVISION consensus statement. *Nat. Rev. Clin. Oncol.* **17**(11), 687–705 (2020)
4. Saldivar Gonzalez, F., Prieto Martinez, F., Medina, F.J.: Descubrimiento y desarrollo de farmacos: enfoque computacional. *Educ. Quim.* **28**(1), 51–58 (2017)
5. Macalino, S.J., Gosu, V., Hong, S., Choi, S.: Role of computer-aided drug design in modern drug discovery. *Arch. Pharm. Res.* **38**(9), 1686–1701 (2015)
6. Garofalo, M., Grazioso, G., Cavalli, A., Sgrignani, J.: How computational chemistry and drug delivery techniques can support the development of new anticancer drugs. *Molecules* **25**(7), 1756 (2020)
7. Universidad Nacional Autonoma de Mexico. *Naturalis: Quimica computacional*. Boletín. Mexico: Universidad Nacional Autonoma de Mexico, Division de Ciencias Basicas
8. Landeros Martinez, L.: Modelado molecular de la funcionalizacion de nanodiamantes con Tamoxifeno dirigido a receptores del cancer mamario. Tesis doctoral. Chihuahua: Centro de Investigacion en Materiales Avanzados, S. C
9. Sirhan, Z., Thyagarajan, A., Sahu, R.: The efficacy of tucatinib-based therapeutic approaches for HER2-positive breast cancer. *Military Med. Res.* **9**(1), 39 (2022)
10. Abdulrahman, H., Uzairu, A., Uba, S.: Computer modeling of some anti-breast cancer compounds. *Struct. Chem.* **32**, 679–687 (2021)
11. Centro para el Control y la Prevencion de Enfermedades. ¿Que es el cancer de mama? (2022). https://www.cdc.gov/spanish/cancer/breast/basic_info/what-is-breast-cancer.htm. Accessed 06 July 2023
12. Palermo Picazo, J., Lassad Rosenthal, J., Juarez Aguilar, L., Medina Nuñez, C.A.: Cancer de mama: una vision general. *Acta medica Grupo Angeles.* **19**(3), 356–360 (2021)
13. Abdulrahman, H., Uzairu, A., Uba, S.: QSAR, Ligand Based Design and Pharmacokinetic Studies of Parviflorons Derivates as Anti-Breast Cancer Drug Compound Against MCF-7 Cell Line. *Chemistry Africa* **4**(1), 175 (2021)
14. Landeros, L., Glossman, D., Flores, N.: Interaction of tamoxifen analog with the pocket of some hormone receptors. A molecular docking and density functional theory study. *Front. Chem* **6**, 293 (2018)

Signal Processing and Medical Images, Data Science



Automatic Cluster Selection in K-Means Lung Segmentation

Natanael Hernández-Vázquez^(✉) , Stewart René Santos-Arce ,
Ricardo Antonio Salido-Ruiz , Sulema Torres-Ramos ,
and Israel Román-Godínez 

Division of Cyber-Human Integration Technologies, University of Guadalajara,
44430 Guadalajara, Jalisco, Mexico
natanael.hernandez4729@alumnos.udg.mx

Abstract. Lung segmentation is a critical step in machine-learning-based radiomics using thoracic computed images. It involves isolating a specific region of interest, but variations in lung intensity values caused by diseased lung tissue can difficult correct segmentation. Although K-means is commonly used, it requires manual intervention to select each cluster related to the region of interest, leading to an efficiency decrease in terms of the specialist's time and effort, especially for large image volumes. To address these limitations, an automatic cluster selection methodology is proposed. It involves a training process to determine a threshold for discriminate clusters; then, morphological transformations and image processing techniques enhance segmentation. Evaluation using DICOM images from the Interstitial Lung Diseases Database yielded a Jaccard Similarity Index of 0.9056 and a Dice Similarity Coefficient of 0.9475, demonstrating the effectiveness and accuracy of the proposed approach.

Keywords: Lung segmentation · K-means · Computational Vision · Image Processing

1 Introduction

Segmentation is the process of delineating specific areas within an image, achieved by assigning labels to each pixel based on regions of interest (ROIs). Various methodologies exist for performing semi-automatic or automatic segmentation, intelligent image processing analysis in the scenario [1, 2]. Thresholding is an easiest algorithm used for image segmentation which groups pixels based on intensity differences [2].

In lung images, segmentation serves as a preprocessing step to separate the lungs, enabling various machine learning tools to concentrate their actions exclusively on this tissue, thus improving performance. Thresholding algorithms work well when lung pathologies are absent, as there is a noticeable contrast between the lungs and the surrounding tissue in both computed tomography (CT) images and plain X-rays. However, complications arise when the lung tissue density increases due to diseases such as pulmonary fibrosis [3], interstitial lung disease [4], and cancer [5], among others, resulting in X-ray beams interacting similarly with both lung and surrounding tissues [6, 7].

To overcome this challenge, state-of-the-art (SOA) approaches often employ K-means due to their simplicity and interpretability. For instance, Gupta et al. (2022) used K-means, fuzzy C-means to separate anatomical structures within the CT images, incorporating wavelet techniques to enhance the obtained mask. Their method achieved an accuracy, Dice Similarity Coefficient (DSC), and Jaccard Similarity Index (JSI) of 0.9928, 0.9872, and 0.9787, respectively [1]. Similarly, Hu et al. (2020) employed Convolutional Neural Networks for lung region mapping and utilized Bayes, Support Vector Machines, and K-means as kernels, achieving an accuracy of 0.97 [8]. Liu et al. (2023) utilized K-means in conjunction with Hough transform to remove cavities, obtaining a DSC and JSI of 0.9786 and 0.9512, respectively [9].

While these studies effectively tackle lung segmentation (LS), they rely on the manual identification of lung clusters per image to generate masks. This can be time-consuming for specialists dealing with sizable image volumes during diagnosis and treatment, where accuracy is vital. Hence this work proposes an approach to automate the selection of lung and non-lung clusters. The goal is to have a methodology for the segmentation of diseased lungs that achieves competitive performances regarding SOA methods.

2 Materials and Methods

K-means is an unsupervised clustering algorithm designed to identify K groups within a dataset. In the context of image segmentation, the dataset consists of elements represented by a dimension $M \times N$ image matrix denoted as \mathbf{X} . The algorithm groups the pixels in \mathbf{X} into K based on their similarity. The user determines the number of clusters, K , based on prior \mathbf{X} data analysis. The resulting segmentation, \mathbf{Y} , allows for drawing conclusions and characterizing the K groups [10].

2.1 Image Segmentation Process using K-means

The image segmentation process is showed in Fig. 1 [11]. It involves treating the image as a dimension $M \times N$ matrix, denoted as \mathbf{X} , where each element X_{mn} represents the intensity information. In CT images, this intensity is expressed in Hounsfield Units [12]. To facilitate processing, the matrix is transformed into a vector using lexicographical ordering $\cdot\mathcal{L}\{\cdot\}$, this consists of the column-by-column to left-to-right stacking of the matrix \mathbf{X} [13]. This vector is then subjected to the K-means algorithm, resulting in a vector where each pixel is assigned to one of the K clusters. Then K-means clustering model output is reorganized through lexicographical reordering $\cdot\mathcal{L}^{-1}\{\cdot\}$, resulting in a \mathbf{Y} matrix of dimension $M \times N$. With this, the user can identify and select the clusters of interest. In the context of LS, the user designates all the pixels within the lung clusters with an intensity value of 1, while non-lung pixels are assigned an intensity value of 0. This process is referred to as binarization, given its binary decision nature in this study.

One drawback of the image segmentation process using K-means is the manual assignment of labels (such as lung or not-lung) to each cluster in every image, which can compromise its effectiveness.

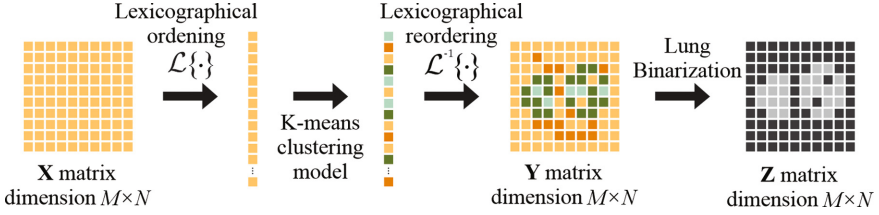


Fig. 1. Representation of image segmentation process using K-means.

2.2 Automatic Cluster Selection

To address the issue of manual selection and labeling of clusters in K-means LS, an automatic methodology is proposed (see Fig. 2). The main idea of this methodology is to determine an upper threshold α and a lower threshold β . Once these thresholds are computed, the ratio of pixels within each cluster that falls within the limits of α and β is calculated.

First, the data set A is divided into two sets: 70% of the studies for training (A_1) and 30% for test (A_2). The training set, A_1 , contains \mathbf{B}_i thorax CT images, while the test set, A_2 , consists of \mathbf{C}_i thorax CT images. Additionally, the \mathbf{D}_i mask is required as a counterpart for \mathbf{C}_i and \mathbf{B}_i ; every image matrix of $M \times N$ dimension. The \mathbf{D}_i mask contains labeled pixels indicating whether they belong to lung and non-lung regions, with lung pixels assigned an intensity value of 1 and non-lung pixels assigned an intensity value of 0. This resulting mask serves as gold-standard for each \mathbf{B}_i and \mathbf{C}_i images.

The training process involves the following steps: for each \mathbf{B}_i image in A_1 , the \mathbf{D}_i mask is applied to compute the element-wise product (\times) between \mathbf{B}_i and \mathbf{D}_i resulting in the extraction of lung intensities. This procedure is repeated for all A_1 images. Using the lung intensities from all A_1 images, the global mean \bar{x} and global standard deviation σ are calculated. These values are then used to determine the threshold using the following equations:

$$\alpha = \bar{x} + \sigma, \quad (1)$$

$$\beta = \bar{x} - \sigma. \quad (2)$$

During the experiment, all \mathbf{C}_i images from A_2 are collected and concatenated into a matrix \mathbf{C}' with dimensions $M \times N_i$. Subsequently, the K-means image segmentation process, as described in Sect. 2.1, is applied to \mathbf{C}' .

To determine the number of clusters, various experiments were conducted with values of k ranging from 2 to 10. Optimal results were achieved with 4 clusters. Through the K-means segmentation process every pixel in the matrix \mathbf{C}' is assigned a label corresponding to one of the K clusters. To determine whether a cluster represents the lungs, a ratio is calculated based on the pixels within the threshold defined by α and β . This ratio is computed using the following equation:

$$\gamma = \frac{p}{P} \quad (3)$$

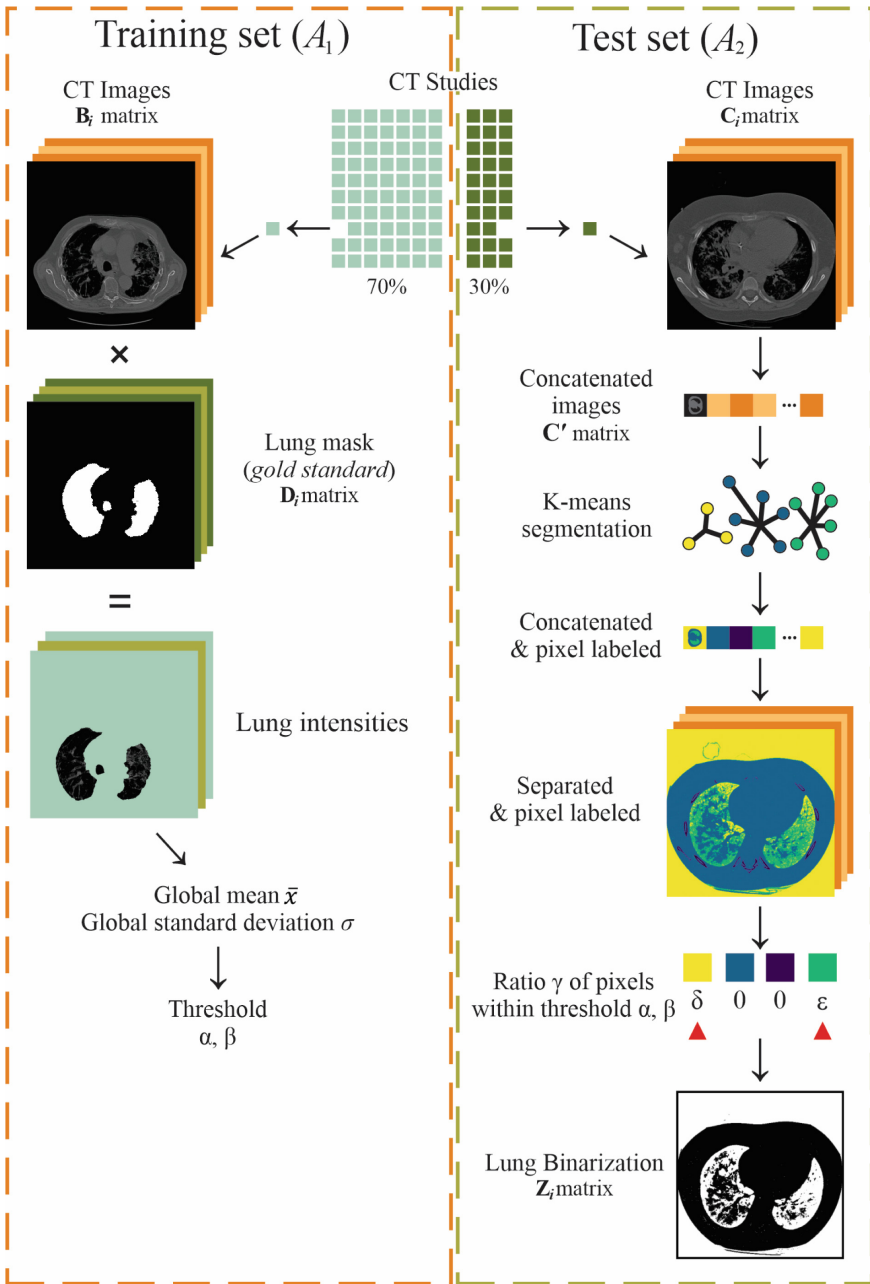


Fig. 2. Schematic of the proposed methodology. δ and ϵ are values greater than 0.001.

where p is the number of pixels within the threshold for a given cluster, while P represents the total number of pixels in that cluster. By calculating γ , clusters with values greater than 0.001 are selected. Subsequently, all pixels belonging to the selected clusters are assigned a value of 1, while the remaining pixels are assigned 0. This generates the $M \times N_i$ matrix binary mask C_i , that needs to be divided into i separate matrices. As a result, we obtain Z_i matrices with $M \times N$ dimensions.

2.3 Post-processing Mask Strategy

The resulting Z_i is obtained for each C_i and undergoes post-processing, which involves applying morphological transformation and image-processing techniques. This process aims to refine the mask and ensure that it covers most lung pixels. Further details of this post-processing procedure are shown in Fig. 3.

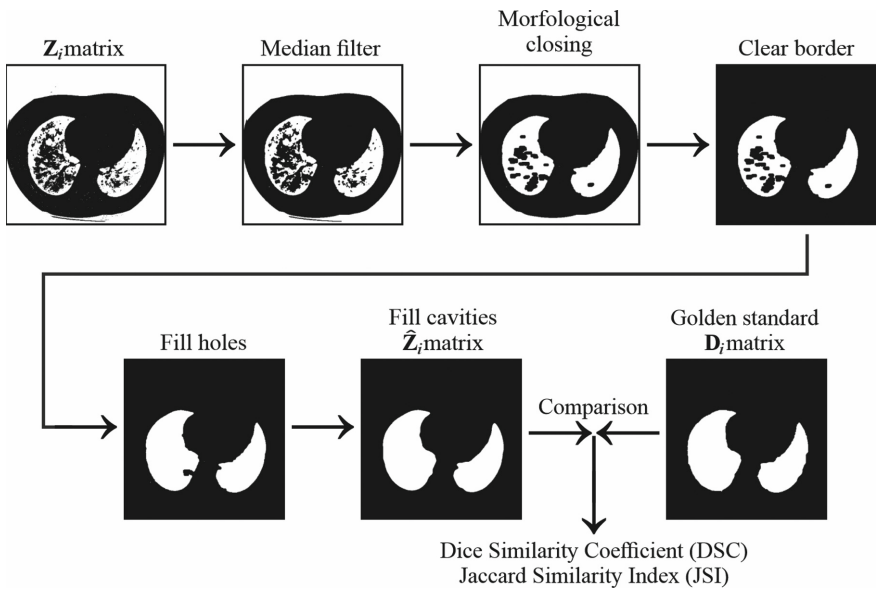


Fig. 3. Schematic post-processing including the gold-standard comparison.

The initial step involves applying a median filter as suggested by Tukey [13], with the defined kernel size of three pixels. This process eliminates scattered pixels in the image and smooth out the relevant ones. Next, a morphological closing operation, initially proposed by Matheron and Serra [14] is performed. The image undergoes dilation followed by an erosion to enhance the solidity of the mask. A circular structural element with a radius of eight pixels is utilized for this operation.

The presence of air outside the body and within the lungs leads to similar intensity values, making it necessary to remove it from the mask. To accomplish this, a morphological transformation is applied, taking advantage of the proximity of air pixels along the image edge. By intersecting the input image with its edge, a marker image is created,

containing seeds for each connected pixel or particle at the edge. Through reconstruction, an image consisting of these particles is obtained and subsequently erased [15]. To fill any remaining unconnected holes, an erosion-based reconstruction is performed using the mask and a marker image with consistent lung values [15].

Depending on the disease, cavities can appear at the edge of the region of interest (ROI). To address this, the circular shape of the cavities can be leveraged using an algorithm described by Liu et al. [9]. The algorithm utilizes the Hough transform to detect circles to be filled. Selection criteria were defined as follows: circles with less than 1/2 of lung pixels, and circles with more than 2/3 of lung pixels on their perimeter. This process results in the post-processing matrix \hat{Z}_i with $M \times N$ dimensions.

2.4 Evaluation Metrics

To assess the performance of the proposed methodology, the post-processed matrix \hat{Z}_i was compared to the corresponding gold-standard mask D_i using DSC [16] and JSI [17]. These metrics provide a similarity value ranging from 0 to 1, where 0 indicates no spatial overlap and 1 represents complete spatial agreement.

Following extensive comparisons using DSC and JSI metrics, the global mean and standard deviation for each metric were calculated and then compared to previous works in the SOA.

We use DSC and JSI metrics in this study based on their common usage in segmentation methodologies and their ability to evaluate performance. However, it is important to note that JSI offers advantages over DSC. Unlike DSC, JSI satisfies all the properties of a metric, including the crucial triangular inequality property [18]. The relaxed triangular inequality in DSC can affect efficiency and approximation ratios, rendering it not fully considered as a metric [19].

2.5 Computational Tools

This methodology and experimentation were developed using Python 3.9.16. Pydicom 2.3.1 for reading DICOM files and their metadata. OpenCV 4.7 for image morphological transformations, NumPy 1.21.5 and Pandas 1.5.3 for matrix analysis, operations, data transformation, and structures, and Matplotlib 3.6.3 for image and results displaying. All this work was carried out using the hardware CPU AMD Ryzen 7 5800 H 3.20 GHz, GPU Nvidia GeForce RTX 3060 6 GB VRAM and 16 GB RAM.

3 Results and Discussion

The evaluation of the proposed methodology use the Multimedia Database of Interstitial Lung Diseases created by Depeursinge et al., and its use for research purposes permitted by the ethics committee of the University Hospitals of Geneva [20]. This database comprises 3076 CT images in DICOM format, obtained from 113 patients diagnosed with various lung diseases within ILDs. Each image has dimensions of 512×512 pixels and is of uint16 value type.

Additionally, a gold-standard lung mask, manually annotated by a medical specialist, is available for each image in the database. For the quantitative assessment, the metrics DSC and JSI were employed. Figure 4 presents box plots illustrating the outcomes obtained by applying the proposed methodology to all 3076 images split randomly into a training set (70% - 2,153 images) and test set (30% - 923 images). Furthermore, the results reported by Gupta et al. [1] and Liu et al. [9] are also depicted in Fig. 4.

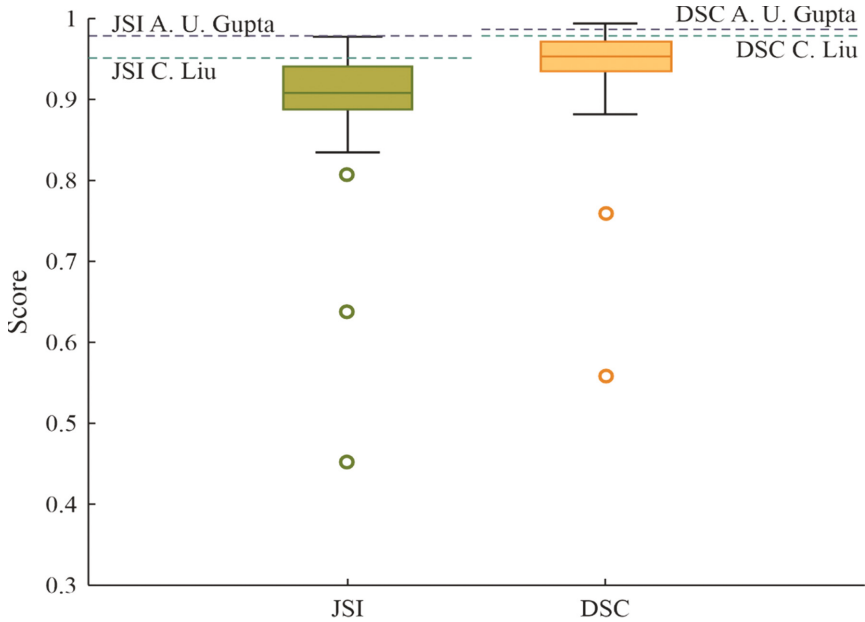


Fig. 4. Performance of the proposed methodology showed in JSI and DCS.

The performance of the proposed automatic cluster selection using K-means can be observed. The box plot reveals a low dispersion of 0.0660 and 0.0520 for JSI and DSC respectively. Notably, Quartile 1 for JSI is 0.887 and for DSC is 0.9351, while Quartile 3 for JSI is 0.9403, and for DSC is 0.9723. These results indicate that approximately half of the automatic segmentation indexes fall within these ranges, suggesting stability in the proposal. However, there are three outliers for JSI and two outliers for DSC, implying that certain images pose challenges for segmentation using this approach. Hence, it is expected to achieve performances close to the median values of 0.9092 for JSI and 0.9539 for DSC. In the comparison with the SOA, Fig. 4 demonstrates that the segmentation results obtained by this proposal are comparable to Gupta et al. [1], and in certain cases, they even surpass the results reported by both authors.

Additionally, Table 1 presents the results obtained using the same metrics and dataset, with K-means as the clustering method. First, Gupta's approach [1] incorporated fuzzy C-means and wavelets into their methodology, while Liu et al.'s approach [9] incorporated the Hough transform. Both authors noted that manual intervention is required for

selecting lung or not-lung clusters. Gupta et al. also reported the need for manual selection of clusters during image decomposition and reconstruction using wavelets, which further increases the level of user intervention required for creating the mask.

Table 1, demonstrates that with the proposed automatic cluster selection approach, it is possible to achieve performances above 0.90 for both JSI and DSC metrics, with a standard deviation of 0.066 or lower. In contrast, Gupta et al. [1] and Liu et al. [9] did not provide information on error rate results or standard deviation resulting from manual interaction. Consequently, the potential impact of this manual selection on reproducibility within their methodologies cannot be determined.

Table 1. Comparison of the proposal with the SOA.

Methodology	JSI		DSC	
	<i>Mean</i>	<i>Standard Deviation</i>	<i>Mean</i>	<i>Standard Deviation</i>
Gupta et al. [1]	0.9787	-	0.9872	-
Liu et al. [6]	0.9512	-	0.9786	-
Proposed methodology	0.9056	0.0660	0.9475	0.0520

Furthermore, results above 0.90 in DSC and JSI (see Table 1) were accomplished, which are near to the values reported by each author, but without the manual intervention of experts that spend important time and effort.

4 Conclusions

This work introduces a methodology that enables automatic cluster segmentation eliminating the need for manual selection of clusters by experts. This approach facilitates fully automated LS, which can be valuable for various applications such as disease classification and delineating ROIs within the lungs for radiological analysis allowing to decrease specialist's work and time spent in the analysis of large image volumes.

In future work, this methodology will be enhanced by incorporating additional image characteristics such as textures and pixel relationships. It is also crucial to evaluate the performance of the proposed methodology on different datasets to assess its robustness. By testing the methodology in diverse scenarios, its applicability and generalization capabilities can be examined.



References

1. Gupta, A.U., Singh-Bhadauria, S.: Multi level approach for segmentation of interstitial lung disease (ILD) patterns classification based on superpixel processing and fusion of k-means clusters: SPFKMC. *Comput. Intell. Neurosci.* **2022**, 22 (2022)

2. Pham, D.L., Xu, C., Prince, J.L.: Current methods in medical image segmentation. *Annu. Rev. Biomed. Eng.* **2**(1), 315–337 (2000)
3. Christie, A., Peters, A.A., Drakopoulos, D., Heverhagen, J.T., Geiser, T., Stathopoulou, T., et al.: Computer aided diagnosis of pulmonary fibrosis using deep learning and CT images. *Invest. Radiol.* **54**(10), 627–632 (2019)
4. Pawar, S.P., Talbar, S.N.: Two stage hybrid approach of deep learning networks for interstitial lung disease classification. *Biomed. Res. Int.* **2022**, 1–10 (2022)
5. Liu, X., Li, K.W., Yang, R., Geng, L.-S.: Review of deep learning based automatic segmentation for lung cancer radiotherapy. *Front. Oncol.* **11**, 717039 (2021)
6. Mansoor, A., Bagci, U., et al.: Segmentation and image analysis of abnormal lungs at CT: current approaches, challenges, and future trends. *Radiographics* **35**(4), 1056–1076 (2015)
7. Yoo, S.J., et al.: Automated lung segmentation on chest computed tomography images with extensive lung parenchymal abnormalities using a deep neural network. *Korean J. Radiol.* **22**(3), 476 (2021)
8. Hu, Q., et al.: An effective approach for CT lung segmentation using mask region based convolutional neural networks. *Artif. Intell. Med.* **103**, 101792 (2020)
9. Liu, C., Xie, W., Zhao, R., Pang, M.: Segmenting lung parenchyma from CT images with gray correlation based clustering. *IET Image Proc.* **17**(6), 1658–1667 (2023)
10. Han, J., Kamber, M.: *Data Mining: Concepts and Techniques*, 3rd edn. Elsevier, Burlington MA (2012)
11. OpenCV: K-Means Clustering in OpenCV. https://docs.opencv.org/3.4/d1/d5c/tutorial_py_kmeans_opencv.html (2023). Accessed 06 May 2023
12. Dougherty, G.: *Digital Image Processing for Medical Applications*. Cambridge University Press, Cambridge UK; New York (2009)
13. Tukey, J.: Nonlinear nonsuperposable methods for smoothing data. *Cong Rec EASCON* **74**, 673 (1974)
14. Matheron, G., Serra, J.: The birth of mathematical morphology. In: *International Symposium on Mathematical Morphology*, pp. 1–16 (2001)
15. Soille, P.: *Morphological Image Analysis: Principles and Applications*. Springer, Berlin Heidelberg (1999)
16. Dice, L.R.: Measures of the amount of ecologic association between species. *Ecology* **26**(3), 297–302 (1945)
17. Jaccard, P.: The distribution of the flora in the alpine zone. *New Phytol.* **11**(2), 37–50 (1912)
18. Kosub, S.: A note on the triangle inequality for the jaccard distance. *Pattern Recogn. Lett.* **120**, 36–38 (2016)
19. Gragera, A., Suppakitpaisarn, V.: Relaxed triangle inequality ratio of the sørensen dice and tversky indexes. *Theoret. Comput. Sci.* **718**, 37–45 (2018)
20. Depeursinge, A., Vargas, A., Platon, A., Geissbuhler, A., Poletti, P.A., Müller, H.: Building a reference multimedia database for interstitial lung diseases. *Comput. Med. Imaging Graph.. Med. Imaging Graph.* **36**(3), 227–238 (2012)



Muscle Activation Patterns Differentiate Post-stroke and Healthy Population

Esteban Eusebio-Martín^(✉) , Ricardo Antonio Salido-Ruiz ,
Sulema Torres-Ramos , and Israel Román-Godínez 

División de Tecnologías para la Integración Ciber-Humana, CUCEI,
Universidad de Guadalajara, 44430 Guadalajara, Jalisco, Mexico
esteban.eusebio4727@alumnos.udg.mx

Abstract. Upper limb impairments can give rise to alterations or functional limitations that may result from strokes. These limitations have conducted to the development of novel tools to aid in various stages of the rehabilitation process. One such tool is the assessment of the patient's level of impairment. Traditionally, the assessment necessitates the continuous involvement of a specialized therapist, which leads to substantial consumption of time and medical resources. To address this, some works have proposed the automatic evaluation of the patient's affection level by using electromyography, by extracting parameters like the muscle activation level or the co-contraction index, others use the muscle activity patterns to look at movement compensation. Nonetheless, none of those works attempt to use muscle activation patterns and the corresponding muscle compensation to automatically assess the impairments. Hence, this work proposes a method for automated detection of muscle activations in electromyography and utilizes them in statistical analysis to compare healthy and post-stroke subjects. It highlights similarities for both groups in muscles with higher activation during tasks, particularly in muscles such as the deltoid and extensor carpi, as well as differences in overall muscle activation percentages. Additionally, significant differences between activation patterns were observed.

Keywords: EMG Muscle Activation Patterns · Stroke · Impairment Assessment · Upper Limb · Daily Living Activities

1 Introduction

Upper limb impairments can arise from various circumstances, resulting in limitations in the performance of activities of daily living. These limitations can arise from various causes, including Acquired Brain Injuries (ABI) such as strokes or Traumatic Brain Injuries (TBI) [1] which according to the National System of Statistical and Geographical Information (INEGI, by its acronym in Spanish) in Mexico in 2021 there were 118 cases per 100,000 inhabitants [2]. They can also be a result of muscle overexertion or repetitive use of the limb, such as in sports activities [3], and can be associated with age-related limitations in older individuals [4]. In order to restore lost functions in the upper limb and enhance the quality of life for individuals affected by such impairments, various rehabilitation strategies are being implemented [5].

The rehabilitation process can be conceptualized as a cyclical approach, including the assessment of the patient, the establishment of goals, intervention, and subsequent assessment to evaluate progress according to the goals [6]. Various scales are employed for the assessment process, including the Functional Ability Scale (FAS) [7], the Wolf Motor Function Test (WMFT) [8], and the Fugl-Meyer Assessment Upper Extremity (FMA-UE) [9], among others.

These scales are typically administered by a therapist and present certain drawbacks, such as significant time and medical resource requirements. Additionally, it can be affected by the subjectivity of the individual conducting the assessment [10]. To mitigate these challenges, alternative approaches have been developed to automate the assessment process by utilizing biosignals during the execution of specific movements.

One example of such biosignals is electromyography (EMG). Hu et al. [11] employed a rehabilitation robotic hand equipped with EMG sensors to detect movement intentions and replicate them. Additionally, the sensors recorded EMG signals to assess whether there was improved muscle coordination following the interventions. Also, McCrea et al. [12] utilized EMG signals to examine muscle saturations (i.e., muscle is maximally activated) in four specific muscles and investigate potential differences between healthy and post-stroke subjects.

Despite there are several studies using EMG signals to evaluate post-rehabilitation improvements or measure muscle saturation to distinguish between groups of subjects, to the best of our knowledge, no studies have yet been found that specifically investigate the utilization of muscle activation patterns for assessing impairments, which is an important source of information due to the compensatory mechanism caused by the impairments. If such information can be automatically detected, it can provide new insights for automated assessment purposes. In this work, we employed a database obtained by Averta et al. [13]. These signals were conditioned and processed to calculate the muscle activation patterns and then, a comprehensive statistical analysis was conducted to compare and contrast these patterns between the two groups. This work aims first to identify differences in muscle patterns between the groups to aid in assessing impairments in post-stroke patients and second in further work to use this information for automated assessment purposes.

2 Materials and Methods

The methodology consists of four main activities: Data Exploration, Preprocessing, Muscle Activation Detection, and Statistical Analysis.

2.1 Data Exploration

This study utilizes a database from Hannover Medical School [13] comprising physiological and kinematic data of 20 healthy (8 female, age 46.77 ± 15.25 years) and 20 post-stroke subjects (6 female, age 49.88 ± 16.92 years), from now on Healthy Group (HG) and Stroke Group (SG), respectively. The electromyographic data was recorded using a wireless sEMG system (Trigno Delsys, Inc., Natick, MA, USA) 12 bipolar

electrodes in the impaired upper limb at 2000 Hz while the subjects performed three repetitions of 30 Daily Living Activities (DLA) (see Fig. 1), each execution time depends on the subject.

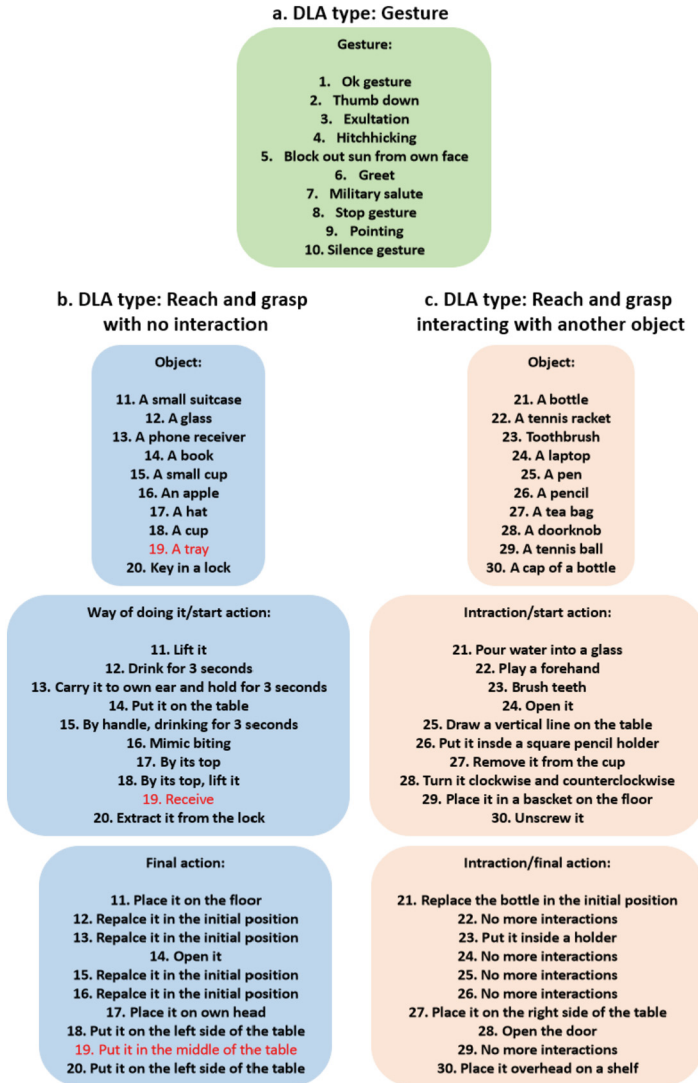


Fig. 1. List of the 30 DLAs. Gestures are shown in a), Reach and grasp with no interaction (in red DLA #19, did not start with a reach and grasp as the rest of the DLAs in this block) are shown in b), Reach and grasp interacting with another object are shown in c), some of the tasks from “c” group do not have a 3rd execution stage and is labeled as “No more interaction”.

2.2 Preprocessing

The EMG data underwent several filtering steps to enhance the quality of the signal. Firstly, a second-order IIR band-stop filter at 50 Hz was applied to eliminate power line noise. Then, a third-order IIR high-pass filter at 5 Hz was used to remove DC offset. Finally, a third-order IIR low-pass filter at 600 Hz was applied to eliminate high-frequency noise. The EMG signals were used without a normalization to take advantage of the magnitudes of each recording.

2.3 Muscle Activation Detection

To automatically detect the muscle activations in the EMG signals a two-step procedure was employed. This procedure involved calculating the threshold and subsequently binarizing the signal.

Two thresholds were calculated for HG based on mean and standard deviation (std) of rectified and windowed EMG signals. A 50-sample window with a 98% overlap was utilized and mean and std were calculated for each window. The maximum value from each measure was recorded, and each threshold was defined as 30% of the maximum values obtained.

For the SG the thresholds from the HG were applied, resulting in only two general thresholds: one based on the mean and the other based on the standard deviation. This was accomplished by taking all the individual thresholds from the HG and calculating their mean value.

Then, for both HG and SG, two binary vectors were generated. When the EMG signal reached or exceeded the thresholds, the corresponding values were mapped to 1; otherwise, they were mapped to 0 (see Fig. 2b).

Some abnormal peaks and holes were observed in the new vectors, therefore a refinement of the binarization by morphological hole filling, which consists of an erosion (see Fig. 2c) and dilation (see Fig. 2d) procedure, was conducted to address this issue [14].

Both, erosion, and dilation operations, were calculated using a 60-sample window, chosen empirically with the aim of having a good resolution and representing around 1% of the average length of the recordings, with a 98% overlap.

A final vector was then calculated by multiplying the two resulting vectors and finally dilated one last time (see Fig. 2e).

This process was performed for the 12 EMG signals of each repetition of the DLA, to every subject.

2.4 Statistical Analysis

Using the previously obtained vectors, the percentage of muscle activation during each of the DLAs was computed by dividing the active state values by the total signal duration. This calculation was performed for each muscle in each DLA and was recorded. Subsequently, the mean percentage across all DLAs was calculated for each muscle. This process was performed for both HG and SG. Figure 3 depicts all the muscles along with their respective percentage of activation for Task #9 and #10. Similar plots were generated for each DLA (not shown here for space restrictions).

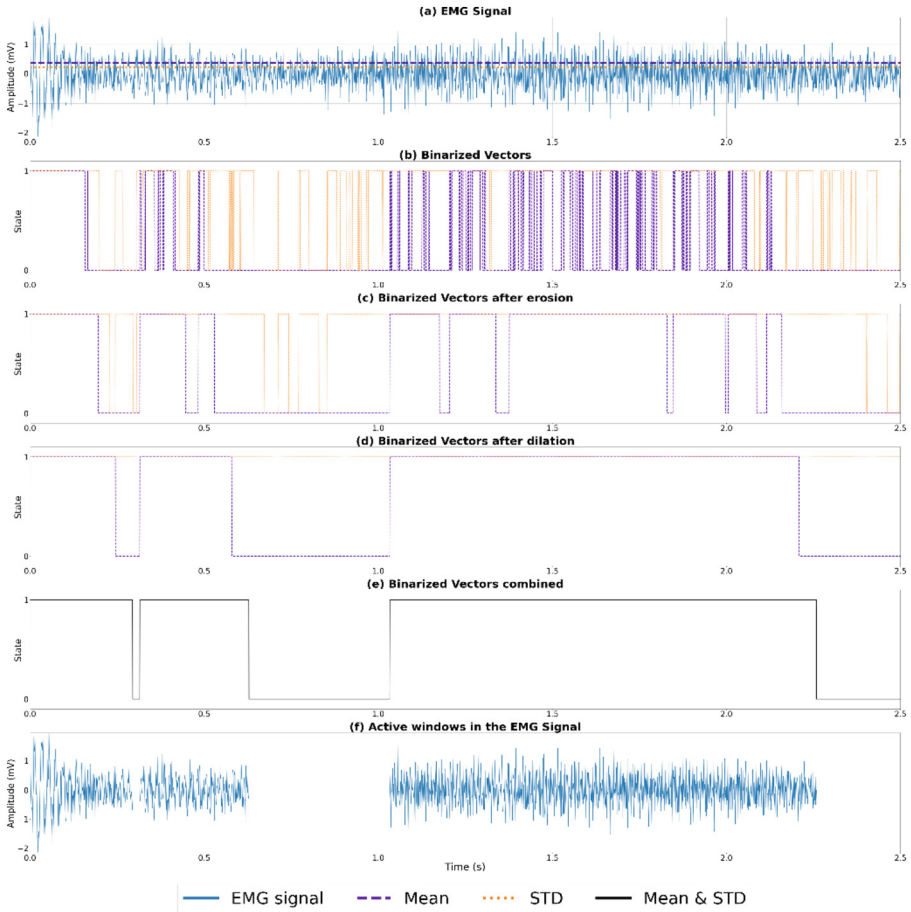


Fig. 2. Detection of muscle activation: (a) EMG Signal (b) Binarized vectors (c) Vectors eroded (d) Vectors dilated (e) Final refined vector (f) Windows with activity in the EMG signal.

After obtaining the percentage values for both HG and SG, a statistical analysis was performed. First, the normality of the data was assessed. If the data followed a normal distribution, a T-test was employed. However, if the data did not meet the assumptions of normality, the Mann-Whitney Test (U-test), which is the non-parametric equivalent of the T-test, was utilized. Both tests allow to verify the probability that 2 samples (HG & SG) come from the same populations and have the same mean. And finally, and Holm-Bonferroni (H-B) correction was made due to the large number of comparisons.

3 Results and Discussion

The results obtained can be divided into two different parts: the muscle activation patterns, and the statistical analysis.

3.1 Muscle Activation Patterns

Muscle activation patterns in terms of activation percentages are shown in Fig. 3, for example, in DLA #9 and DLA #10 the muscle activation pattern can be observed as the changes in muscle activation in function of the DLA, in each of the muscles involved in the task. For instance, it is possible to see that most of the muscles have an activation grouped in range of around 20 percentage units for each of the 30 DLAs, 40%–60% in the HG (blue lines) and 20%–40% in the SG (red lines), but there are few muscles below or above this range.

The difference in the range between HG and SG can be seen as a lack of activation for most of the muscles during the realization of the tasks from the SG.

3.2 Statistical Analysis

Regarding the statistical analysis, box plots were created for each DLA. An example for DLA# 10 can be seen in Fig. 4. These box plots were based on the calculated percentages, allowing the identification of muscles with significantly different activation patterns between the groups. Regarding the median values, the HG demonstrates a range of approximately 40% to 60% for muscle activation whereas the SG exhibits lower values around 20% to 40%. Thus, reinforcing the idea that for the SG the activation of most of the muscles is lower than the activation from HG.

Table 1 presents the p-values resulting from the statistical analysis conducted on all repetitions of each DLA for each subject. The values highlighted in the table indicate significant differences ($p < 0.05$) after H-B correction. Notably, only seven out of the 30 DLAs did not show any significant difference, with five of them belonging to the “Reach and grasp with no interaction” group (see Fig. 1). This could be attributed to the specific movements required for these tasks. In all five cases, the objects to be held were described as “big” in size and/or “light” in weight, and intricate finger movements were not necessary for their manipulation.

Among the other groups of DLAs, the “Gesture” type exhibited the lowest p-values. This observation could be attributed to the specific phase of the movement where the upper limb is required to maintain a gesture position for an extended period of time. This task appears to be more demanding for the SG group compared to the HG group.

The boxplots also provided insights into the presence of outliers, which helped identify muscles with higher or lower percentages of activation during certain DLAs. For instance, in Fig. 4, it is evident that for the HG, the M. Deltoideus pars clavicularis (DPC) exhibit higher activity, while the M. Brachioradialis demonstrated lower activity.

Another important point was that in certain DLAs the outliers from HG matched with those from the SG. This may be due to the nature of the task, where a particular muscle is expected to be consistently active, or nearly inactive throughout the activity. This observation highlights another notable distinction between groups and subjects. For instance, if the task requires consistent activation of the DPC, and a subject from the SG exhibits a lack of activation in this muscle, it could indicate an impairment in that particular muscle for that subject.

To gain a clearer understanding of these effects, the number of times each muscle was identified as an outlier was recorded. This data was used to create a histogram displaying

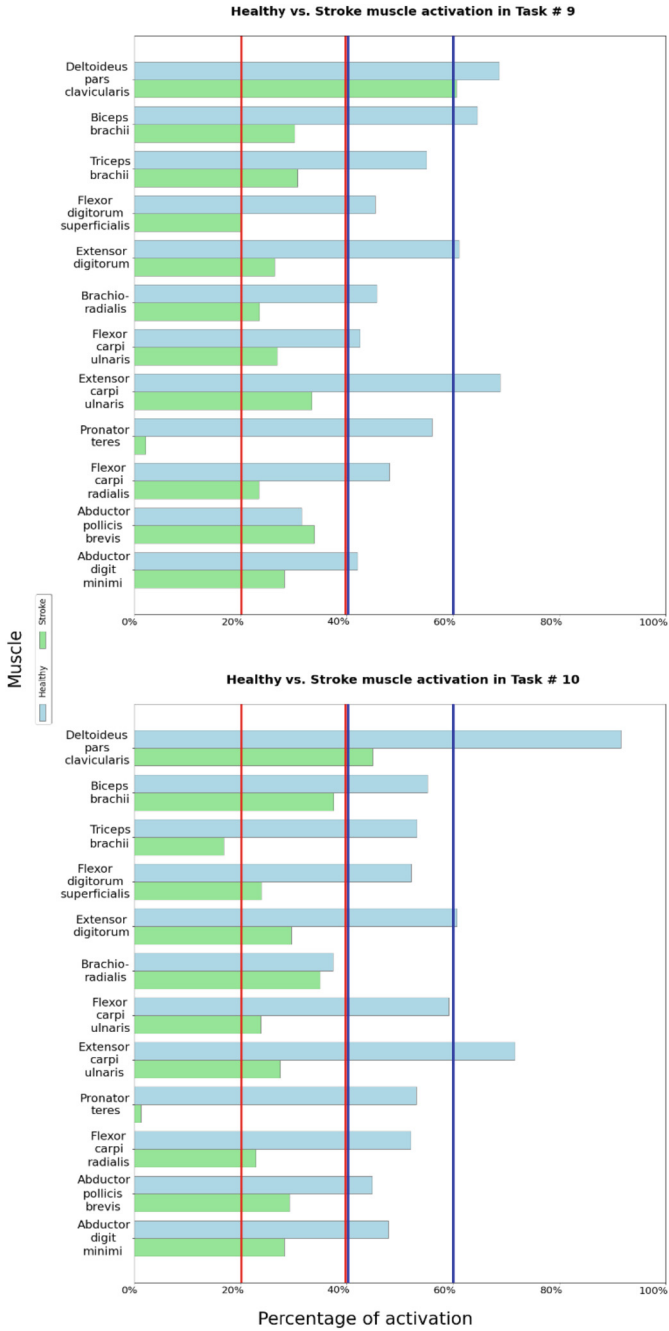


Fig. 3. Percentage of muscle activation during the DLA #9 and #10 of the HG & SG.

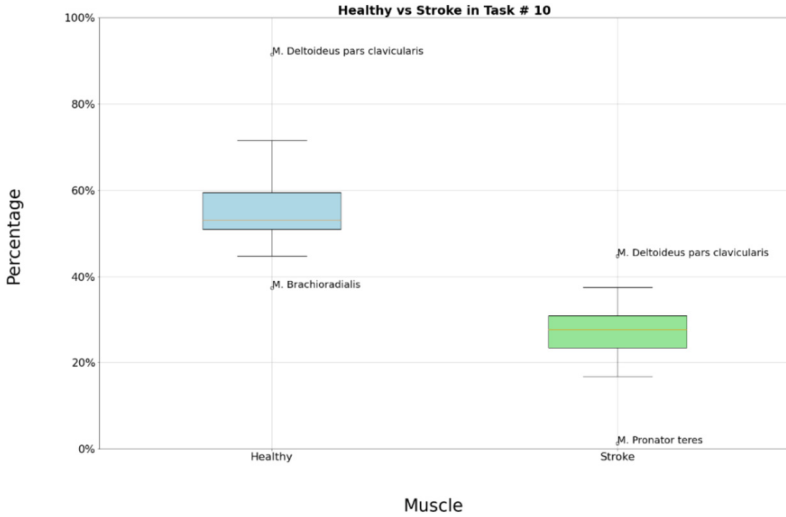


Fig. 4. Boxplot of the DLA #10 for both groups where the muscle names are the muscles showing an activation outside IQR.

Table 1. P-values table of each DLA

No. DLA	P-value	No. DLA	P-value	No. DLA	P-value
1	2.86 e-04	11	5.96 e-05	21	8.93 e-04
2	9.75 e-04	12	1.31 e-02	22	8.86 e-01
3	5.71 e-03	13	2.14 e-01	23	3.02 e-02
4	4.69 e-05	14	5.65 e-02	24	5.06 e-04
5	9.01 e-04	15	7.25 e-03	25	7.80 e-03
6	2.01 e-03	16	1.01 e-02	26	6.01 e-04
7	9.66 e-07	17	2.82 e-01	27	6.09 e-03
8	3.84 e-04	18	5.05 e-04	28	1.93 e-02
9	3.84 e-04	19	2.47 e-03	29	2.94 e-03
10	6.79 e-06	20	3.63 e-03	30	1.01 e-02

Gesture
 Reach and grasp with no interaction
 Reach and grasp interacting with another object

the muscles with either higher or lower activity in both groups (see Fig. 5). In the HG group, only three distinct muscles were identified, with the DPC muscle being the most frequently observed for high activity.

In the case of the SG group, a higher number of muscles (five) exhibited variations in activity. This is evident in Fig. 5, where each green bar represents a muscle with either increased or decreased activity in the SG. These findings align with the observations reported in [12], which indicated that stroke subjects tend to rely more on certain muscles

compared to healthy subjects. Specifically, the M. Pronator teres was the muscle with the highest frequency of appearances for lower activity. It is worth noting that the second highest frequency of appearances was the same as in the HG group, suggesting that this muscle may consistently be active during specific DLAs, as is the case with the M. Extensor carpi ulnaris.

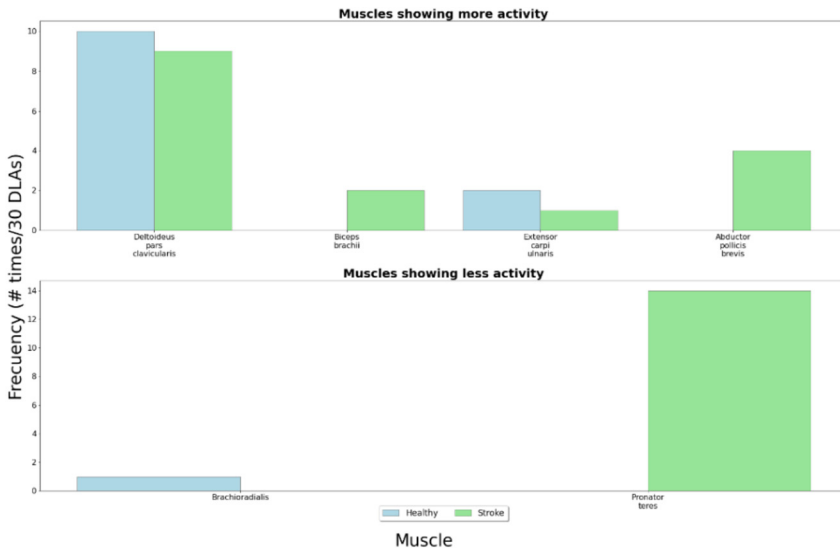


Fig. 5. Frequency of muscle activity variations across DLAs.

4 Conclusions

In this study, the muscle activation percentages were computed for various DLAs in two groups: HG and SG. The objective was to compare these percentages and acquire additional information, specifically the muscle activation patterns, that can distinguish between the two groups. The findings confirmed that there are notable differences in muscle activations between the two groups. Specifically, it was observed that the differences are more pronounced in the first type of DLAs, while the disparity becomes less prominent in the second type of DLAs due to the nature of the movements involved. Furthermore, it was observed that certain muscles exhibit similar activation patterns, characterized by increased activity; however, the percentages of activation differ between the two groups. Notably, SG demonstrated a greater number of individual muscles with heightened activity. This phenomenon can be attributed to the compensatory mechanisms employed when a particular muscle is unable to execute a movement independently.

As part of the limitations of this work some factors that could affect the signals used were not considered, such as the muscle fatigue due to the execution of 90 tasks, or the used of the impaired limb even if it was the non-dominant arm that could make the execution slower.





Finally, for further works the information obtained here, specifically the muscle activation patterns will be useful along with other EMG characteristics for implementing new assessment tools based in machine learning techniques, using a normalization of the EMG signals could be useful as well to look at similar shapes between the groups, the ratio of agonist to antagonist muscles could be a good information to use in future works to confirm the muscle compensation as well.

References

1. Centers for Disease Control and Prevention: Prevalence and most common causes of disability among adults- -United States, 2005. *MMWR: Morbidity Mortality Weekly Rep.* **58**(16), 421–426 (2009)
2. Secretaría de Salud. [https://www.gob.mx/salud/prensa/531-en2021-ictus-o-enfermedad-vascular-cerebral-ocasiono-mas-de-37-mil-decesos-en-mexico#:~:text=Arauz%20G%C3%B3ngora%20inform%C3%B3%20que%2C%20en,Estad%C3%ADstica%20y%20Geograf%C3%ADa%20\(Inegi\)](https://www.gob.mx/salud/prensa/531-en2021-ictus-o-enfermedad-vascular-cerebral-ocasiono-mas-de-37-mil-decesos-en-mexico#:~:text=Arauz%20G%C3%B3ngora%20inform%C3%B3%20que%2C%20en,Estad%C3%ADstica%20y%20Geograf%C3%ADa%20(Inegi)), Accessed 8 Aug 2023
3. Breceda, E.: Motor rehabilitation in stroke and traumatic brain injury: stimulating and intense. *Curr. Opin. Neurol.* **26**(6), 595–601 (2013)
4. de Oliveira, L.: Health-related quality of life of institutionalized older adults: influence of physical, nutritional and self-perceived health status. *Arch. Gerontol. Geriatr.* **92**, 104278 (2021)
5. Adans-Dester, C.: Enabling precision rehabilitation interventions using wearable sensors and machine learning to track motor recovery. *NPJ digital Med.* **3**(1), 1–10 (2020)
6. Langhorne, P.: Stroke rehabilitation. *The Lancet* **377**(9778), 1693–1702 (2011)
7. Wolf, S.: Assessing wolf motor function test as outcome measure for research in patients after stroke. *Stroke* **32**(7), 1635–1639 (2001)
8. Patel, S.: A novel approach to monitor rehabilitation outcomes in stroke survivors using wearable technology. *Proc. IEEE* **98**(3), 450–461 (2010)
9. Del Din, S.: Estimating Fugl-Meyer clinical scores in stroke survivors using wearable sensors. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 5839–5842. IEEE (2011)
10. Lohse, K.: On the reporting of experimental and control therapies in stroke rehabilitation trials: a systematic review. *Arch. Phys. Med. Rehabil.* **99**(7), 1424–1432 (2018)
11. Hu, X.: The effects of post-stroke upper-limb training with an electromyography (EMG)-driven hand robot. *J. Electromyogr. Kinesiol.* **23**(5), 1065–1074 (2013)
12. McCrea, P.: Saturated muscle activation contributes to compensatory reaching strategies after stroke. *J. Neurophysiol.* **94**(5), 2999–3008 (2005)
13. Averta, G.: U-Limb: a multi-modal, multi-center database on arm motion control in healthy and post-stroke conditions. *GigaScience* **10**(6), 043 (2021)
14. Kang, K.: Event detection of muscle activation using an electromyogram. *Appl. Sci.* **10**(16), 5593 (2020)



Heterogeneous Behavior in the Iowa Gambling Task: A Clustering Approach

Manuel Chávez-Sánchez^(✉) , Sulema Torres-Ramos , Israel Román-Godínez ,
and Ricardo Antonio Salido-Ruiz 

División de Tecnologías para la Integración Ciber-Humana, Universidad de Guadalajara,
Marcelino García Barragán 1421, 44430 Guadalajara, Jalisco, México
manuel.chavez2913@alumnos.udg.mx

Abstract. Iowa Gambling Task is a widely used tool that involves the evaluation of decision-making ability while playing a strategic gambling game. The assessment of the game results has been analyzed using tools such as inferential statistics, knowledge-based systems, and machine-learning classification techniques. Hence, the purpose of this research is to examine a publicly available subset of data from over 300 non-clinical participants who have performed the Iowa Gambling Task (IGT), using K-means to identify performance-related subgroups during IGT. The findings suggest that while the samples were initially reported as homogeneous, there are significant differences observed based on their database origin. Furthermore, the application of the K-means algorithm revealed numerous distinct subgroups of behavior in relation to the credits earned during the test. Thus, the analysis of this feature may suggest previously unconsidered information that could help us to obtain a more detailed view of the behavior in decision making, in both non-clinical and patients with decision making disorders.

Keywords: Iowa Gambling Task · Decision-Making · K-means · Heterogeneity Behavior

1 Introduction

Data science applied in biomedical engineering provides information from different types of evaluation: biological, physiological, or through data obtained from neuropsychological tests, which have become a widely employed tool for evaluating the behavioral characteristics of a specific population. Among them, for evaluation of decision-making ability, Iowa Gambling Task (IGT) [1] is noticeable. It is conducted through a gambling game in an arrangement of four decks (A, B, C or D), in which the participants must create a strategy that allows them to earn more credits according to the choices they made. The analysis of behaviors resulted in the neuropsychological test has shown that different dispositions occur according to the experience generated; thus, depending on the temporal moment being evaluated, the participant will have different knowledge and tendencies. IGT is typically divided into five blocks of 20 decisions each, until all 100 decisions are completed. Different versions have been developed, where 95 [2], 120 [3]

or up to 200 elections [4] are applied. Thereby, the processing and analysis of decisions made in the IGT, allow recognition of the strategic predispositions generated by the participants and evaluation of their performance.

For analysis purposes, ANOVA is the most widely used tool comparing IGT scores. [4–8] Although its explainability is easy to understand, this approach only takes in consideration the final credits of the participants or its division by blocks, so that the generated strategies visible in the credits of the participants during the test, result in loss of information, which is important in diagnostic accuracy to identify differences between groups. This differentiation can be performed at each trial of the test, this is the statistical analysis of every decision; however, this procedure may not be useful or interpretable. Given this problematic, the purpose of this work is to use machine learning tools that are able to identify specific differentiation features in IGT data.

In this perspective, several techniques have been used to identify patterns in IGT; for example, knowledge-based models have been designed [9] or classification tools have been also used for pattern identification in IGT [10]. However, to the best of our knowledge, no clustering techniques have been used to identify subgroups related to the performance of participants during IGT, based on the number of credits obtained; this could be useful to study heterogeneity in the decisions made. Therefore, the aim of this research is to propose a different approach to identify performance-related subgroups during the IGT, using one of the most popular and easily understood clustering techniques, the K-means algorithm. This approach will provide useful and clearly interpretable information in the analysis of the performance of participants in the IGT, additionally it could change the way in which data from this neuropsychological test is processed, improving the quality of interpretation and the explicability of the results.

2 Materials and Methods

The proposed methodology consists of the following phases: 1) Database selection, 2) Feature extraction and datasets creation, 3) Experimental setup, 4) Clustering process and 5) Clustering evaluation (see Fig. 1).

2.1 Database Selection

This study uses a publicly available database compilation, provided by Steingroever et al. [11], which collects results from 10 independent studies evaluating the performance of neurotypical participants on the IGT. Institutional Review Board approval was obtained for each data collection. The data set comprises results from 617 participants; however, there are variations in the protocols used, ranging from 95 to 150 decisions. For this research, only the results of subjects who applied the 100-trial version were used. The dataset used in this study comprises 315 participants obtained from Horstmann's research (data originally published by Anette Horstmann in Steingroever et al. [12]) ($n = 162$) and Wood et al. [13] ($n = 153$). Each study specifies three types of variables: the choice made by the participant in each trial (A, B, C or D), the positive feedback (credits ranging from 50 to 100) and negative feedback (credits ranging from -50 to -1250 ; 0, in case there was none).

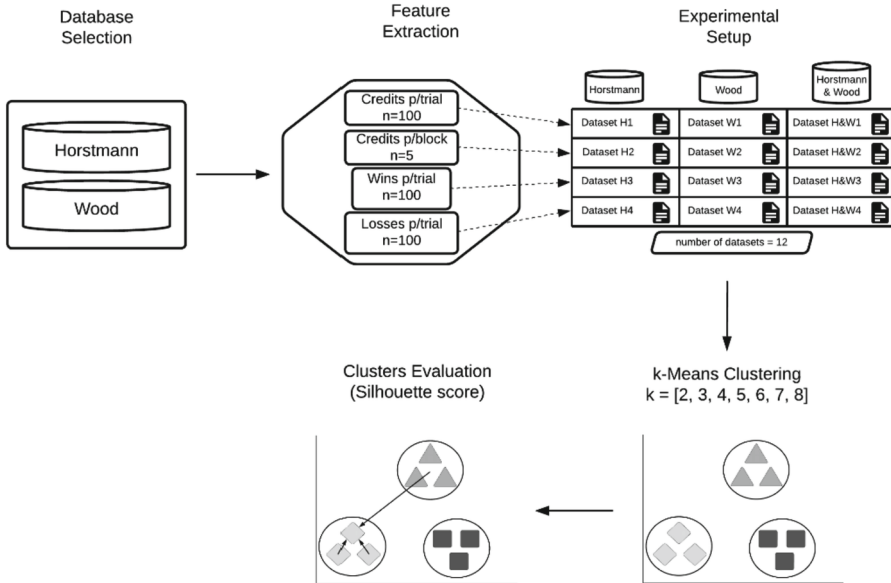


Fig. 1. Methodological diagram.

2.2 Feature Extraction

To analyze different behavioral patterns according to the credits obtained by the participants, four features per module were analyzed: 1) Credits per trial (datasets H1, W1, H&W1), obtained by the addition of the feedback on each trial, resulting in 100 features; 2) Credits per block (datasets H2, W2, H&W2), analyzing only the scores obtained at the end of each block (decisions 20, 40, 60, 80 and 100), resulting in five features; 3) Wins per trial (datasets H3, W3, H&W3), where the earnings of the participants in each trial are evaluated, resulting in 100 features; and finally 4) Losses per trial (datasets H4, W4, H&W4), which evaluates the defeats at each trial, resulting in 100 features. Each characteristic was chosen according to two conditions: 1) the information provided by the database, 2) the general test arrangements, which make it possible to recognize where differentiating patterns can be found.

2.3 Experimental Setup

The experiments were structured into three modules: first, using only data extracted by Horstmann et al.; second, for data extracted by Wood et al. and last, for the data combined (Horstmann et al. & Wood et al.), this was performed with the aim of identifying differences between experiments, according to their origin. For each dataset, clustering experiments, as described in Sect. 2.4, were conducted.

2.4 Clustering

To perform the clustering process, one of the most employed and comprehensive clustering algorithms, namely K-means, was selected. K-means is an unsupervised machine-learning algorithm designed to identify elements that are closest in the solution space, given a Euclidean-type geometry [14]. Thus, given a series of K-centroids, the distance of each instance with respect to each centroid is calculated and clustered according to the smallest distance.

$$E = \sum_{i=1}^k \sum_{p \in C_i} \text{dist}(p, c_i)^2 \quad (1)$$

where E is the sum of the squared distance for all objects in the data set, p is the point in space representing a given object (instance), and c_i is the centroid of cluster C_i . The K-means clustering method, implemented using the Orange3® data mining tool [15], was employed to cluster each instance of the experiments based on the extracted features. By the nature of the algorithm, which requests the number of centroids k to be used for clustering (see Eq. 1), clustering was performed using a brute force method, applying the algorithm for k values ranging from two to eight. Therefore, for each dataset, seven different experiments were performed, one for each k value, and all of them were evaluated through silhouette coefficient metric.

2.5 Cluster Evaluation

The silhouette coefficient is a measure used to evaluate the quality of clustering. Its value ranges from -1 to 1. A higher silhouette coefficient refers to a model with more cohesive clusters [14]. For each object o , coefficient incorporates two spacing types: $a(o)$, representing the compactness of the cluster to which o belongs (see Eq. 2) and $b(o)$ corresponding to the degree of separation of o from the other clusters (see Eq. 3).

$$a(o) = \frac{\sum_{o' \in C_i, o' \neq o} \text{dist}(o, o')}{|C_i| - 1} \quad (2)$$

$$b(o) = \min_{C_j: 1 \leq j \leq k, j \neq i} \left\{ \frac{\sum_{o' \in C_j} \text{dist}(o, o')}{|C_j|} \right\} \quad (3)$$

where C_i are each cluster evaluated, C_j are the other each cluster; each object o belongs to C_i ; whereas each o' belongs to C_j cluster and, k is the clusters which the dataset is partitioned. Finally, the silhouette coefficient of o , is then defined as

$$s(o) = \frac{b(o) - a(o)}{\max\{a(o), b(o)\}} \quad (4)$$

Thus, $\max\{\}$ is a function that selects the maximum value from the indicated functions: $a(o)$ and $b(o)$. As can be observed, the silhouette coefficient is computed for each individual sample, and the resulting values can be averaged to represent the entire cluster. In this study, the averaged silhouette coefficient was employed to represent a clustering experiment. This allows us to identify those experiments in which the cohesion between samples has been achieved in a better quality, and in this way, to identify the characteristic that best indicates this differentiation between samples.

3 Results and Discussion

To determine the general tendency of the studied population, credits of the participants were evaluated. Both studies reported the same initial conditions in the application of the IGT (see Fig. 2 at trial zero), starting with 2000 credits. Also, observe that at trial twenty, the most significant differences between databases are observed. In addition, negative feedback seems more punitive in the Horstmann dataset, whereas the Wood dataset shows larger positive feedback. On the other hand, between trials 50 and 60, minimal differences are observed across all databases. However, after this point, differences start to increase again. The tendency shifts, with the Horstmann dataset displaying a shift towards more weighted positive feedback, and the Wood dataset showing a shift towards more punitive negative feedback. The combined Horstmann and Wood dataset represents an average tendency observed in each dataset separately.

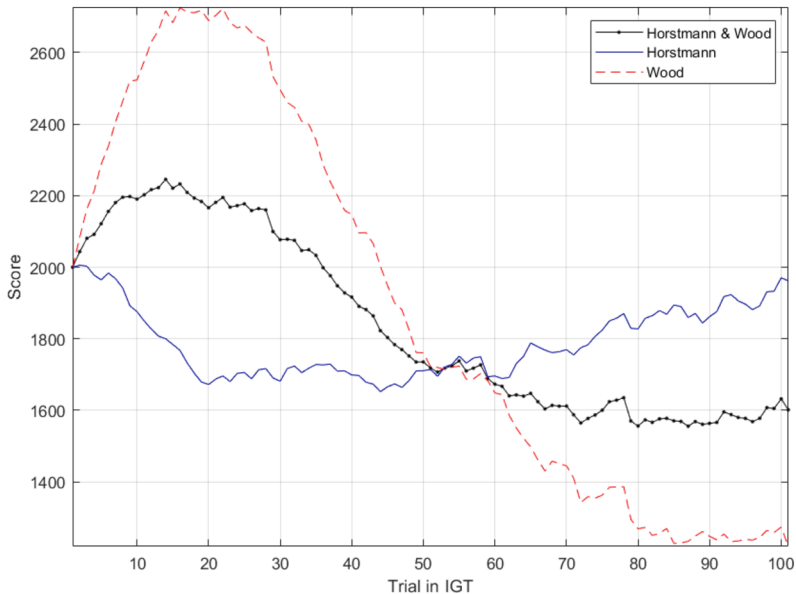


Fig. 2. The average scores for Horstmann, Wood and combined databases are shown. Note that, for all cases, the number of credits at the end of the test does not exceed the initial credits and that different behaviors are described in relation to the credits obtained during the test, among participants of both studies.

Quantitative clustering evaluation indicates that the most differentiating features are the *Scores per trial* and the *Scores per block* (see Table 1). To do so, the silhouette results were used to rank the features and find the highest silhouette coefficient. These features exhibit differences of approximately one order of magnitude compared to profits and losses. For Horstmann & Wood, the highest silhouette coefficient was observed in clusters two and four for *Scores per trial* and *Scores per block*, respectively. In the case

of Horstmann, it occurred in two clusters for both features, whereas for Wood, it was observed in clusters three and four.

Table 1. Silhouette coefficients for each feature at each experimental module.

Horstmann & Wood				
Clusters	Scores per trial	Scores per block	Profits	Losses
2	.205	.257	.072	.080
3	.202	.258	.032	.036
4	.200	.269	.015	.058
5	.166	.234	.007	.022
6	.166	.224	.018	.016
7	.158	.222	.003	.015
8	.145	.228	.006	.013
Horstmann				
2	.186	.242	.066	.020
3	.143	.214	.043	.051
4	.154	.237	.041	.037
5	.156	.240	.038	.014
6	.131	.228	.032	.030
7	.119	.224	.022	.010
8	.115	.222	.035	.005
Wood				
2	.212	.281	.072	.017
3	.218	.294	.021	.052
4	.194	.312	.042	.087
5	.186	.266	.027	.031
6	.130	.268	.024	.025
7	.119	.267	.026	.024
8	.125	.254	.019	.049

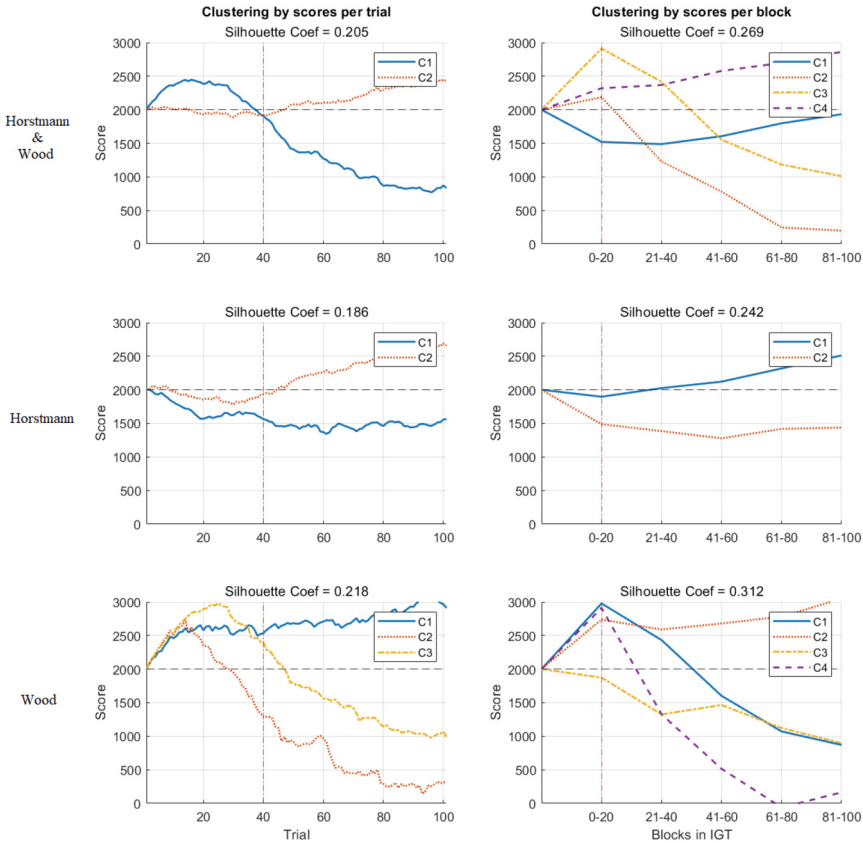


Fig. 3. Scored averaged according to the clustering analysis. Each column corresponds to the differentiating features, whereas each row indicates the experimental module: Horstman, Wood and Horstmann & Wood. The center dotted line indicates the credits at IGT startup. Two fundamental performances can be observed: good and bad performance.

This k distribution for each feature could indicate that, as previously reported, the behavior of neurotypical subject is heterogeneously distributed [8, 12]. This heterogeneity may be due to several factors: gender [3], nationality [8] or medical history [16]. However, the findings suggest that there may be more than just two behavioral subgroups, with some cases potentially exhibiting three or even four subgroups. Considering each clustering experiment with the highest silhouette coefficient, the credits obtained during IGT were evaluated. Each curve represents the averaged behavior of k -means clusters founded. These observations, as an initial remark, may indicate distinct phases of learning within the IGT. For instance, when clustering by scores per trial (see Fig. 3, left column) a notable shift in direction can be observed around trial 40, for the three datasets (H&W1, H1 and W1). This change potentially means a shift in the strategies employed by the participants. On the other hand, when clustering by scores per block (see Fig. 3, right column), it is observed the influence of the dataset on the tendencies

among the blocks. However, for the three datasets (H&W2, H2 and W2) a notable shift in direction can be observed between blocks 0-20 and 21-40.

For all experiments, two main trends are observed: participants who performed better and those who performed worse. This pattern of performance behavior has been previously reviewed in several research studies [8], nevertheless, their classification has not been previously reviewed using unsupervised artificial intelligence tools, such as *k*-means. The possible implications that could be associated with changes in *k* value, due to the clustering process, might be reviewed. Moreover, although the different phases in IGT have been reviewed [6, 12], there are internal trends that could be studied in detail, given the clustering performed with machine-learning techniques. The implications of a better analysis of participants' behavior within the IGT could lead to a better interpretation of the results for the identification of related decision-making deficits.

4 Conclusions

The methodical review of the evaluated parameters through the experimentation performed, in which data were extracted from participants, and features were extracted and evaluated using K-means as a clustering machine-learning tool, has provided several suggestions: the processing of results obtained from IGT or other neuropsychological assessments tools is not only reliable, but also complementary in behavioral interpretation. Also, the subgroups differentiation among a set of participants, such as the neurotypical population, in this case, is an interpretation that deserves to be reviewed from different points of view: previous work already indicates the existence of such non-homogeneous behavior; however, the causes of this differentiations and its relation to the health status of the individual are still unknown. Therefore, future work would consider the identification of such differentiation factors, relating them to the performance of participants during the IGT, and the comparison between different clustering tools.

References

1. Bechara, A., Damasio, A.R., Damasio, H., Anderson, S.W.: Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* **50**(1–3), 7–15 (1994)
2. Fridberg, D.J., et al.: Cognitive mechanisms underlying risky decision-making in chronic cannabis users. *J. Math. Psychol.* **54**(1), 28–38 (2010)
3. van den Bos, R., Jolles, J., van der Knaap, L., Baars, A., de Visser, L.: Male and female Wistar rats differ in decision-making performance in a rodent version of the Iowa Gambling Task. *Behav. Brain Res.* **234**(2), 375–379 (2012)
4. Giustiniani, J., Gabriel, D., Nicolier, M., Monnin, J., Haffen, E.: Neural correlates of successful and unsuccessful strategical mechanisms involved in uncertain decision-making. *PLoS ONE* **10**(6), e0130871 (2015)
5. Serrano, J.I., Iglesias, Á., Woods, S.P., del Castillo, M.D.: A computational cognitive model of the Iowa Gambling Task for finely characterizing decision making in methamphetamine users. *Expert Syst. Appl.* **205**, 117795 (2022)
6. Bechara, A., Tranel, D., Damasio, H., Damasio, A.R.: Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cereb. Cortex* **6**(2), 215–225 (1996)

7. Chiu, Y.C., Lin, C.H., Huang, J.T., Lin, S., Lee, P.L., Hsieh, J.C.: Immediate gain is long-term loss: are there foresighted decision makers in the Iowa Gambling Task? *Behav. Brain Funct.* **4**(1), 1–10 (2008)
8. Lee, W.K., Lin, C.J., Liu, L.H., Lin, C.H., Chiu, Y.C.: Recollecting cross-cultural evidence: are decision makers really foresighted in Iowa gambling task? *Front. Psychol.* **11**, 537219 (2020)
9. Iglesias, A., del Castillo, M.D., Serrano, J.I., Oliva, J.: A computational knowledge-based model for emulating human performance in the Iowa Gambling Task. *Neural Netw.* **33**, 168–180 (2012)
10. Kornev, D., et al.: Gaming behavior and brain activation using functional near-infrared spectroscopy, Iowa gambling task, and machine learning techniques. *Brain Behav.* **12**(4), e2536 (2022)
11. Steingroever, H., et al.: Data from 617 healthy participants performing the Iowa gambling task: a “many labs” collaboration. *J. Open Psychol. Data* **3**(1), 340–353 (2015)
12. Steingroever, H., Wetzels, R., Horstmann, A., Neumann, J., Wagenmakers, E.J.: Performance of healthy participants on the Iowa Gambling Task. *Psychol. Assess.* **25**(1), 180 (2013)
13. Wood, S., Busemeyer, J., Kolling, A., Cox, C.R., Davis, H.: Older adults as adaptive decision makers: evidence from the Iowa Gambling Task. *Psychol. Aging* **20**(2), 220 (2005)
14. Han, J., Pei, J., Tong, H.: *Data Mining: Concepts and Techniques*. Morgan kaufmann (2022)
15. Demšar, J., et al.: Orange: data mining toolbox in python. *J. Mach. Learn. Res.* **14**(Aug), 2349–2353 (2013)
16. Gorlyn, M., Keilp, J.G., Oquendo, M.A., Burke, A.K., Mann, J.J.: Iowa Gambling Task performance in currently depressed suicide attempters. *Psychiatry Res.* **207**(3), 150–157 (2013)



Vision Algorithm to Compute Pupil Parameters: Towards Non-invasive Estimation of Intracranial Pressure

Gerardo Adrián De La Rosa Hernández^{1(✉)}, Jesús Alberto Morales Gómez², Griselda Quiroz Compean¹, and Luis Martín Torres Treviño¹

¹ Facultad de Ingeniería Mecánica y Eléctrica, Universidad Autónoma de Nuevo León, Ave. Universidad S/N, Cd. Universitaria, 66455 San Nicolás de los Garza, Nuevo León, Mexico

gerardo.delarosahe@uanl.edu.mx

² Departamento de Neurocirugía del Hospital Universitario “Dr. José Eleuterio González”, Universidad Autónoma de Nuevo León, 64460 Monterrey, Nuevo León, Mexico

Abstract. Conventional intracranial pressure monitoring consists of an invasive neurosurgical intervention in which a hole is made in the cranium, at the frontal level, through which a sensor is inserted and lodged in the epidural, intracerebral or intraventricular space. The sensor is connected to a system that continuously measures intracranial pressure and is maintained for the necessary time, depending on the pathology. For this reason, this paper proposes the development of a vision algorithm to compute pupil parameters in order to collect data from the pupil that could be considered to estimate an approximate value of intracranial pressure, which can be used by neurosurgeons to assess the possible disease of a patient without the need to perform an invasive neurosurgical procedure.

Keywords: ICP monitoring · ICP vision system · Pupillometer

1 Introduction

The central nervous system is a structure made up of four main components which are: the brain, spinal cord, cerebrospinal fluid (CSF) and blood [1]. An increment in the volume of any of the constituents will lead to a reduction in volume in one or two of the remaining constituents. This change in volume has the consequence of elevating intracranial pressure (ICP) [2,3]. The normal ICP range for adults is approximately 10 to 15 mmHg, for younger children is 3 to 7 mmHg, for newborns is 1.5 to 6 mm Hg, and for pediatrics is not clearly established [1].

Among adults, the dimensions of the pupil typically span from 2 to 4 mm under well-lit circumstances, but can dilate to a range of 4 to 8 mm in conditions of dim lighting. Both pupils are usually equal in size. The pupil enlarges,

or dilates, when exposed to darkness. The pupil constricts when focusing on a nearby object. It responds to direct light and consensual illumination. Abnormalities occur when the pupil fails to dilate in darkness or constrict in response to light or accommodation [4]. Pupil constriction lasts about one second, with diameter changes ranging from 0.25 mm to 2.2 mm [5, 6]. ICP is not the only condition that affects the constriction or dilatation of the pupil, other conditions that can affect the pupil include: medications, traumatic brain injuries or neurological disorders [7–9].

ICP is a condition characterized by elevated pressure within the cranium, exceeding 20 mmHg. Various factors can contribute to ICP, including brain volume changes, injuries, tumors, abscesses, blood clots, cerebrospinal fluid buildup, increased cerebral blood flow, and cranial deformities [2]. Controlling ICP is crucial to prevent brain damage or death. High ICP can lead to cerebral ischemia, as well as bradycardia and respiratory irregularities caused by Cushing’s response, a sympathetic-renal mechanism [10].

Pupils’ responses to visible light provide valuable information for determining ICP levels. Signals controlling pupil contraction and dilation travel from retina nerves to midbrain and hypothalamus, affecting the response if damaged [5, 11]. The parasympathetic and sympathetic nervous systems control pupil size. The retina relays light data to the pretectal olive nucleus, which signals the midbrain’s Edinger-Westphal nucleus. This nucleus, via cranial nerve III, prompts pupil constriction [7].

Numerous approaches have been suggested for the non-invasive assessment of elevated ICP. These encompass: pupillometry, optic nerve sheath measurements, transcranial doppler, magnetic resonance imaging, computed tomography, and electroencephalography analysis [12–16]. Non-surgical techniques employed for measurement have the benefit of avoiding procedures that might result in brain injury. Nonetheless, it is acknowledged that these methods still exhibit limited accuracy when compared with invasive approaches, also, they are unable to facilitate uninterrupted monitoring. This paper proposes the development of a vision algorithm that can measure the different parameters from the human eye that could be used to estimate an approximate value of intracranial pressure.

2 Methodology

The computational vision algorithm proposed here consists in the development of a tool based on Python and OpenCV to detect pupil diameter changes. A disturbance, such as light, can alter the pupil’s size over time, allowing us to calculate dilation or constriction speed based on its position and time. Pupil constriction velocity less than 0.8 mm/sec is considered abnormal [17]. To have a better knowledge of the vision algorithm, the diagram of Fig. 1 is proposed.

2.1 Video Acquisition

Here, an offline system is proposed, then it is first necessary to consider recording a video through a camera and then turning that video into information that can

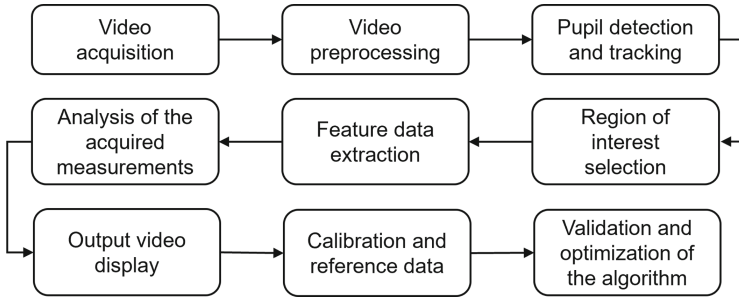


Fig. 1. Vision algorithm flow chart.

determine if a person has elevated ICP. To achieve the above, an algorithm is developed to continuously examine and analyzes the video. The video analyzed in this study comprises a frame of 640×480 pixels, and operates at a velocity of 10 fps.

2.2 Video Processing

Haar cascade classifiers, introduced by Viola and Jones [18], are effective for pupil detection. They use positive and negative images (Positive images show the desired object to detect, while negative images do not) in a cascading machine learning approach, enabling object detection in various images.

The video starts with color-to-gray conversion. Then, DetectMultiScale from CascadeClassifier identifies objects of various sizes, including eye states (open, closed, or blinking). The detected objects are returned as a list of rectangles [19]. The *DetectMultiScale* function has four return values. These four values correspond to the x and y axis of the initial point and height and width of the position, so considering this the rectangle around the eye can be made using the *rectangle* function.

Once there is a rectangle of the area of interest, the second stage of processing is to analyze the space obtained. For that, some *ImageFiltering* functions are used, which are *GaussianBlur* and *erode*. *GaussianBlur* applies convolution to the source image using a designated Gaussian kernel and *erode* generates a structuring element with the defined size and shape for use in morphological procedures. In the thresholding process, the video is performed in gray-scale since the same threshold value is applied to each pixel. The pixel value is set to 0 if it falls below the threshold, if not it is set to a maximum value.

After that, the video must be analyzed using *HoughCircles* function, which is used to find circles in an image or video. Hough Circle Transform works based on the mathematical circle theory. In mathematical terms, a circle can be expressed as,

$$(x - x_c)^2 + (y - y_c)^2 = r^2 \tag{1}$$

where r is the radius of the circle, and (x_c, y_c) are the centers. Finally, the eye pupil is rounded by a circle. To do that, *Circle* from *Drawing Functions* is used.

Taking advantage that the pupil's circular shape and prior processing, ideally, a single black circle remains, simplifying detection. Table 1 lists the relevant functions, while Algorithm 1 provides the video processing pseudocode.

Table 1. Image and video processing functions.

Function	Python code	Description
CvtColor	<code>cv2.cvtColor(src,code,dst,dstCn)</code>	Converts image color
CascadeClassifier	<code>cv2.CascadeClassifier(filename)</code>	Object detection
DetectMultiScale	<code>x.detectMultiScale(src,sF,minN)</code>	List of rectangles
GaussianBlur	<code>cv2.GaussianBlur(src,ksize,dst)</code>	Blurs and image
Erode	<code>cv2.erode(src,kernel,iterations)</code>	Erodes an image
Threshold	<code>cv2.threshold(src, thresh,maxval,type)</code>	Segmentation process
HoughCircles	<code>cv2.HoughCircles(src,c,m,dp,md,p)</code>	Circles in an image
Circle	<code>cv2.circle(src,cent,rad,col,th)</code>	Draws a circle
Append	<code>x.append(elmnt)</code>	Makes a data list

Algorithm 1. ICP monitoring vision system pseudocode

```

1: Initialize Video and Cascade Classifier
2: while 1:
3:   ret, img = cap.read()
4:   gray = cv2.cvtColor()
5:   eyes = eyecascade.detectMultiScale()
6:   if (len(eyes) > 0) then "Eye Open"
7:     for = (ex,ey,ew,eh) in eyes:
8:       roigray = gray[ey : ey + eh, ex : ex + ew]
9:       blur = cv2.GaussianBlur(roigray)
10:      erosion = cv2.erode(blur)
11:      ret3, th3 = cv2.threshold(erosion)
12:      circles = cv2.HoughCircles(erosion)
13:      circles = np.uint16(np.around(circles))
14:      for = i in circles[0, :]:
15:        if = i[2] > 0 and i[2] < 55:
16:          cv2.circle(roigray)
17:            diameter calculation
18:            velocity calculation
19:        end
20:      end
21:    end
22:  else: "Eye Close"
23:    cv2.imshow()
24:  end
25: end

```

2.3 Analysis of the Acquired Measurements

From the prior processing, we detect the pupil diameter. So, now we focus on two key data points: pupil size and dilation/constriction speed.

The *HoughCircles* function returns as output $i[0]$ corresponding to the x-axis, $i[1]$ corresponding to the y-axis and $i[2]$ corresponding to the radius of the circle. To store the diameter and speed, lists were created with the *Append* function while the program collects the information until the end of the video time. So, considering that the radius of the circle is available according to the *HoughCircles* function and that a list can also be made using the *Append* function, the diameter data is calculated when executing the algorithm, once the video ends, two variables are generated where the list of pupil diameters and velocities is stored. Position or diameter of the pupil is obtained from the video processing tools of OpenCV. Velocity can be calculated as follow,

$$v = \frac{\Delta d}{\Delta f} = \frac{d_n - d_{n-1}}{f_n - f_{n-1}} \quad (2)$$

For $n = 1, 2, 3 \dots$ last calculated diameter and where d_{n-1} is the first calculated diameter in the first frame f_{n-1} and d_n is the second calculated diameter in the second frame and continues consecutively.

2.4 Medical Research Protocol

Elevated ICP often arises after a brain injury, causing skull pressure and affecting the midbrain and brainstem, leading to changes in pupil reactivity. Pupil reactivity alterations observed with a vision system are associated with elevations in ICP [20]. The Neurologic Pupil Index (NPi) is a developed algorithm aimed at measuring pupillary reactivity and eliminating subjective elements from this evaluation [8]. The NPi algorithm integrates several pupil-related factors, including latency, dimension, rate of dilation, and rate of constriction, which are all included as parameters in Table 2. Each aspect obtained from a pupil measurement is contrasted with a standard distribution derived from individuals without health issues, serving as a reference. The difference between the measurements and the mean is calculated and standardized using the standard deviation. A composite score ranges from 0 to 5, with a score of 3 or higher indicating normal pupil behavior. A score below 3 indicates abnormal pupillary light reflex, a slow response compared to the normative model. A score of 1 is considered more abnormal than a score of 3.

Table 2. Pupil size variables.

Variable	Units
Pupil Min/Max Size	mm
Constriction	%
Latency	seconds
Constriction Velocity	mm/sec
Max. Constriction Velocity	mm/sec
Dilation Velocity	mm/sec
NPi	Scalar value

In individuals with elevated ICP, an inverse correlation between pupil reactivity and ICP is presented. Pupil metrics summarized in Table 2 are: (a) minimum size (pupil's size when it reaches the maximum level of constriction), (b) constriction percentage (difference between maximum and minimum size divided by maximum size), (c) time difference (duration between light stimulation and pupillary constriction onset), (d) average velocity (extent of constriction divided by constriction duration), with the highest velocity occurring during constriction, (f) recovery rate (extent of pupil size recovery divided by recovery duration), and (g) an algorithm comparing these variables to a conventional model to generate an overall score for pupillary reaction.

As mentioned, pupil reactivity shows an inverse correlation with ICP in individuals with elevated ICP, particularly in those with impaired NPi. More research is required to investigate the temporal relationship between NPi and ICP and evaluate its potential as a non-invasive biomarker [20]. To conduct offline ICP monitoring, a suitable test environment must be established. This involves ensuring consistent video acquisition by using a single camera model with the same resolution, position, and orientation for all participants. Ambient lighting should be kept constant. Additionally, the light source used to stimulate pupil dilation or contraction should be standardized across the study.

3 Results

Having considered the previous work, the outcomes of the algorithm creation for the ICP vision monitoring system are showcased. As mentioned in video acquisition section, proposed video consists of a frame of 640×480 pixels, and a speed of 10 fps, a screenshot of the original model could be seen in Fig. 2. Figure 3 show the functions mentioned in Table 1.

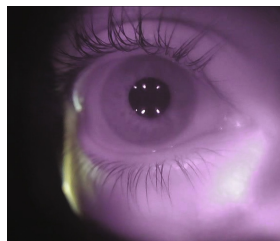


Fig. 2. Screenshot of a frame from the proposed original video.

As mentioned in Subsect. 2.3, the code calculates the pupil diameter and velocity and stores them in lists, Fig. 4 shows when the eye is closed, open and the pupil diameter in millimeters in addition to the graphs obtained of diameter and velocity once the video concludes.

Normal pupil diameter size is from 2 to 4 mm when exposed to light and 4 to 8 mm while is out of light [4,21]. According to Fig. 5(a), average pupil diameter is approximately 4mm, so measurement presents a normal size. In case of the velocity, rate change can be positive or negative depending on a dilatation or constriction of the pupil but observing the Fig. 5(b), an average velocity could be 0.6 to 0.8 mm/s [17]. Pupil velocity less than 0.8 mm/s is considered abnormal,

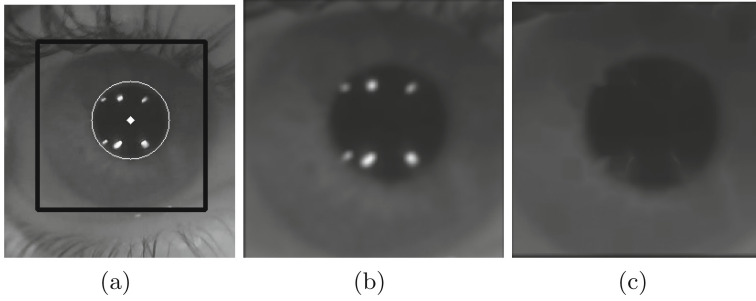


Fig. 3. (a) Rectangle and Circle functions, (b) Blur function and (c) Erosion function.

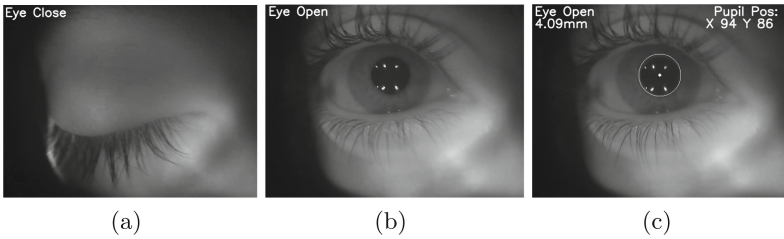


Fig. 4. Response of the proposed system to the following states: (a) Eye is closed, (b) Eye is open and (c) Eye is open and detect the pupil.

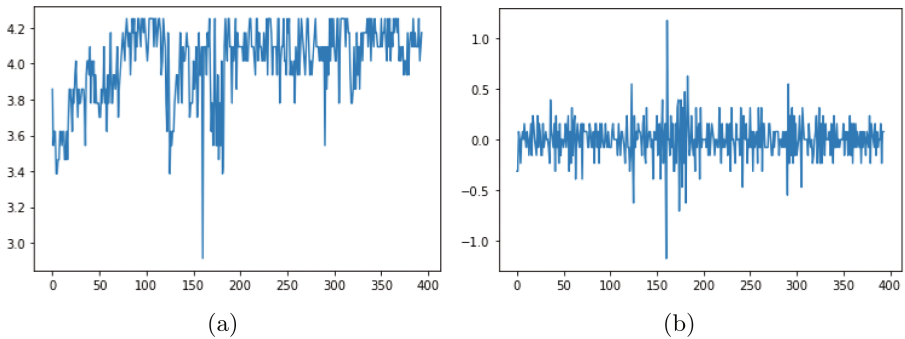


Fig. 5. (a) Pupil diameter graph and (b) Pupil rate of change graph.

so velocity results must be revised to check that velocity coding based on the diameter and frames data is correct.

4 Discussion

The proposed methodology shows that the development of the visual algorithm accomplished pupil diameter and velocity calculation. Pupil diameter results are in adequate ranges according to background investigations, but physical implementation is needed to proceed with the calibration, validation, and optimization of the algorithm because velocity rate presents an abnormal functionality. The development of the vision algorithm to compute pupil parameters in order to collect data from the human eye could be considered to develop a non-invasive ICP estimation method. As mentioned in Subsect. 2.4, a medical research protocol needs to be made to have several study subjects that help us to improve the algorithm calibration. So, future research needs to be continuing to obtain an approximate value of ICP related to parameters of the NPi.

5 Conclusions

Vision systems based on video processing could be used to measure human morphological changes. In the case of measurement and tracking of the pupil diameter, the vision system proposed here worked in a regular way. Haar feature-based cascade classifiers was introduced about 2000s and is just like a kernel in a Convolutional Neural Network (CNN), nevertheless, in CNN, the kernel values are determined through training, whereas in the case of a Haar Feature, manual selection is employed. This is important because the parameter adjustment of the different functions spends a lot of time when cascade classifiers are applied to videos.

Results show that the pupil diameter could be measure in a regular way because of the imprecision used in the function parameters don't allow a perfect pupil detection. Velocity codification needs to be checked to be sure that the diameter data or the frame data are correct. A medical research protocol must be carried out when the algorithm works properly, making reliable measurements of pupil diameter and velocity, at the same time the way of recording video of the patients who could participate in the protocol must be standardized.





References

1. Rangel, L.: Management of intracranial hypertension. *Neurol. Clin.* **26**, 521–541 (2008). <https://doi.org/10.1016/j.ncl.2008.02.003>
2. Pinto, V.: Increased Intracranial Pressure. *StatPearls* (2022). <https://www.ncbi.nlm.nih.gov/books/NBK482119/>
3. Doczi, T.: Volume regulation of the brain tissue - a survey. *Neurochir* **121**, 1–8 (1993). <https://doi.org/10.1007/BF01405174>

4. Spector, R.: The pupils. clinical methods: the history, physical, and laboratory examinations (1990). <https://www.ncbi.nlm.nih.gov/books/NBK381/>
5. Ciuffreda, K.: Understanding the effects of mild traumatic brain injury on the pupillary light reflex. *Concussion* **2**, 08 (2017). <https://doi.org/10.2217/cnc-2016-0029>
6. Ellis, C.: The pupillary light reflex in normal subjects. *Br. J. Ophthalmol.* **65**, 754–759 (1981). <https://doi.org/10.1136/bjo.65.11.754>
7. Adoni, A.: The pupillary response in traumatic brain injury: a guide for trauma nurses. *J. Trauma Nurs.* **14**, 191–196 (2007). <https://doi.org/10.1097/01.jtn.0000318921.90627.fe>
8. Chen, J.: Pupillary reactivity as an early indicator of increased intracranial pressure: the introduction of the neurological pupil index. *Surg. Neurol. Int.* (2011). <https://doi.org/10.4103/2152-7806.82248>
9. Dunn, L.: Raised intracranial pressure. *J. Neurol. Neurosurg. Psychiatr.* **73**, 23–27 (2002). https://doi.org/10.1136/jnnp.73.suppl_1.i23
10. Schmidt, E.: Intracranial pressure is a determinant of sympathetic activity. *Front. Physiol.* **9**, 02 (2018). <https://doi.org/10.3389/fphys.2018.00011>
11. Manley, G.: Infrared pupillometry during uncal herniation. *J. Neurosurg. Anesthesiol.* (2002). <https://doi.org/10.1097/00008506-200207000-00009>
12. Rosenberg, J.: Non-invasive methods of estimating intracranial pressure. *Neurocrit. Care* **15**, 599–608 (2011). <https://doi.org/10.1007/s12028-011-9545-4>
13. Price, D.: Review of non-invasive intracranial pressure measurement techniques for ophthalmology applications. *Br. J. Ophthalmol.* **104**, 887–892 (2020). <https://doi.org/10.1136/bjophthalmol-2019-314704>
14. Andersen, M.: A new novel method for assessing intracranial pressure using non-invasive fundus images: a pilot study. *Sci. Rep.* **10**, 8 (2020). <https://doi.org/10.1038/s41598-020-70084-0>
15. Stevens, A.: Optical pupillometry in traumatic brain injury: neurological pupil index and its relationship with intracranial pressure through significant event analysis. *Brain Inj.* **33**, 1032–1038 (2019). <https://doi.org/10.1080/02699052.2019.1605621>
16. Kristiansson, H.: Measuring elevated intracranial pressure through noninvasive methods: a review of the literature. *J. Neurosurg. Anesthesiol.* **25**, 372–385 (2013). <https://doi.org/10.1097/ANA.0b013e31829795ce>
17. Shoyombo, I.: Understanding the relationship between the neurologic pupil index and constriction velocity values. *Sci. Rep.* **8**, 04 (2018). <https://doi.org/10.1038/s41598-018-25477-7>
18. Viola, P.: Rapid object detection using a boosted cascade of simple features. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1 (2001). <https://doi.org/10.1109/CVPR.2001.990517>
19. Open Source Computer Vision. Cascadeclassifier class reference (2022). [https://docs.opencv.org/](https://docs.opencv.org/docs.opencv.org/)
20. Hutch, M.: Quantitative pupillometry and intracranial pressure in neuro icu patients. *Neurology* **94** (2020). ISSN 0028–3878
21. Mathôt, S.: Pupillometry: psychology, physiology, and function. *J. Cogn.* **21**, 02 (2018). <https://doi.org/10.5334/joc.18>



Sex-Based Speech Pattern Recognition for Post-traumatic Stress Disorder

Andrea Alejandra García-Valdez^(✉) , Israel Román-Godínez ,
Ricardo Antonio Salido-Ruiz , and Sulema Torres-Ramos 

Division of Cyber-Human Integration Technologies, University of Guadalajara,
44430 Guadalajara, Jalisco, Mexico
andrea.garcia4730@alumnos.udg.mx

Abstract. Post-Traumatic Stress Disorder (PTSD) is a mental health disorder diagnosed by physical and psychological evaluations, scales of assessment, and diagnosis criteria. PTSD diagnosis can be affected by subjective evaluations and limited accessibility. Hence, alternative methodologies for diagnosing PTSD have emerged, like speech analysis, it is recognized as a source of predictors to psychiatric disorders, can be evaluated remotely and suggest sex-based differences. Few studies have been focused on changes in the speech of PTSD population, and, as far as we can ascertain do not explored features of sex-based speech associated with PTSD. For this reason, the present study aims to automatically detect sex-based speech patterns linked to indicators of PTSD through the characterization of speech signals. This research employed decision trees to recognize speech predictors of PTSD. The resulting models achieved favorable performance, being the women's model the most accurate. Moreover, when not considering sex-based differences, the spectral domain features highlighted and confirmed that PTSD population is more likely to have uniform speech. Additionally, the results suggest sex-based differences. Such as, the men's model considers the dispersion of speech and if a voice is tense, and the women's model focuses on the frequency bands of the speech.

Keywords: PTSD · Speech · Sex-based Patterns · Decision Tree

1 Introduction

Post-traumatic stress disorder (PTSD) is a psychiatric disorder stemming from direct or indirect exposure to shocking, scary, or dangerous event. It can impact the psychological, physiological, interpersonal, and spiritual aspects [1], affecting individuals of diverse ethnicities and nationalities without age-related restrictions.

According to World Health Organization reported that approximately 3.6% of the global population experienced PTSD in 2013 [2]. Additionally, the American Psychiatric Association estimated that one in eleven people would be diagnosed with PTSD [1]. Among women, the occurrence of PTSD throughout their lifetime is recorded at a rate of 9.7%, while men have a lower rate of 3.6% [3].

To diagnose PTSD, therapists use several assessment instruments, such as perform a physical exam, a psychological evaluation, scales of interview-based assessment, which usually estimate symptom presence and severity, and diagnosis criteria described in manuals like the Diagnostic and Statistical Manual of Mental Disorders [4], often known as the “DSM-V”.

However, PTSD diagnosis can be affected by accessibility barriers, including economic factors, social stigma, the experience of clinicians, patient acceptability, or the episodic nature of psychiatric conditions [5].

Hence alternative methodologies for PTSD diagnosis have emerged, managing to improve the evaluations by using biosignals [6], neuroimaging [7], clinical lab markers [8], facial expressions [9, 10], text data [9, 11] and speech [12–15], among others. Speech is recognized as a source of predictors to psychiatric disorders and presents advantages like expressing emotion and thoughts directly through verbal content, which can generalize in different languages.

In this regard, most studies have used automatic extraction of speech features to classify the presence or absence of PTSD [13, 14]. Some others focused on changes in the speech PTSD population. For example, Marmar et al. [12] found features that contribute to high performance in detecting PTSD, these features indicated slower speech and less change in tonality. Similarly, Scherer et al. [15] discovered reduced tonality in the vowel space and exhibited that the participants with PTSD and depression show more tense voice features.

Although promising, these studies did not explore the sex-based variability of speech features associated with PTSD. In this respect, previous findings suggest differences in women and men evaluations of psychiatric disorders. Stratou et al. [10] demonstrate that employed a sex-specific approach notably enhances the identification of nonverbal signs associated with PTSD. However, this study does not consider speech features as potential biomarkers of PTSD.

At present, few publications have described and explained indicators that could be related to PTSD. To a lesser extent, studies do not consider sex-based differences in searching for such indicators; which leads us to the following research question: Are there differences in speech patterns between women and men in the PTSD population that can be identified by using signal characterization and machine learning techniques? For this reason, the present study aims to automatically detect sex-based patterns linked to indicators of PTSD through the characterization of speech signals.

2 Materials and Methods

The proposed approach is divided into three main phases: data pre-processing, experimental dataset, and machine learning modeling. A methodology overview is shown in Fig. 1.

2.1 Dataset

The Distress Analysis Interview Corpus - Wizard of Oz (DAIC-WOZ) is a dataset created to aid in diagnosing mental health disorders, such as PTSD, developed in 2014 by Gratch et al. [16] at the University of California.

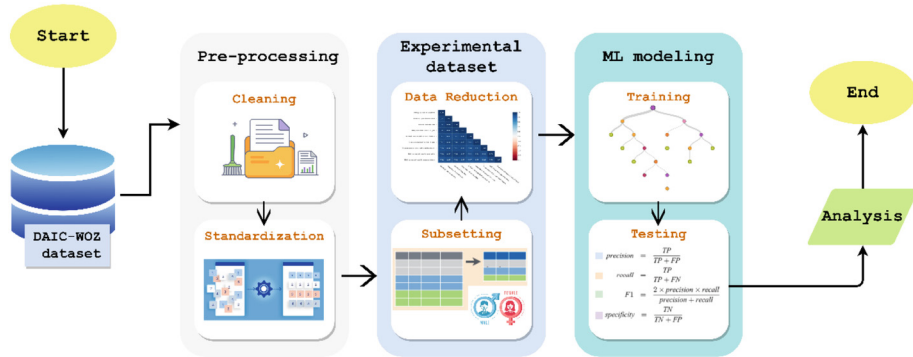


Fig. 1. Proposed methodology to detect sex-based patterns linked to speech indicators of PTSD.

This resource contains audio and transcriptions of clinical interviews by Ellie, an animated virtual interviewer (a person who asked a set of questions in a semi-structured way operated Ellie).

The database involves 189 participants aged from 18 to 70 years, who were fluent English speakers for which speech features were extracted through COVAREP [16]. Such features were: fundamental frequency (F0), maxima dispersion quotient (MDQ), the first two harmonics of the differentiated glottal source spectrum (H1, H2), voicing (VUV), normalized amplitude quotient (NAQ), quasi-open quotient (QQQ), Mel Frequency Cepstral Coefficients (MCEP 0-12), Mel cepstral coefficients (MCEP 13-24), parabolic spectral parameter (PSP), spectral slope of wavelet responses (peak/slope), shape parameter of the Liljencrants-Fant model of the glottal pulse dynamic (Rd), Rd conf, harmonic model and phase distortion mean values (HMPDM 0-24) and deviations (HMPDD 0-12), for a total of 73 features.

Also, the database provides the scores from a standardized self-report questionnaire called the PTSD Checklist-Civilian version (PCL-C) [17] based on the DSM-IV criteria for PTSD recognition. Scores range from 17 to 85, representing the magnitude of PTSD severity. The first range (17–29) represents little to no severity, the second range (30–44) indicates moderate to moderately high severity, and the last range (45–85) signifies high severity and a positive diagnosis for PTSD. The distribution of the DAIC-WOZ participants is noted in Table 1.

Table 1. Distribution of DAIC-WOZ database participants

Severity	Score	Women	Men	Total
Little to no severity	17–29	43	54	97
Moderate severity	30–44	27	10	37
High severity	45–85	17	38	55
Total		87	102	189

2.2 Data Pre-Processing

The first phase of the methodology consists of data cleaning and standardization.

Initially, features with zero value across all instances were dropped. Then, interviewer parts and spaces without speech were removed to analyze only the voice of the participants.

Furthermore, feature vectors were standardized using z-score [18], a statistical measure to evaluate data from populations with different standard deviations and means, and placed them on a common scale.

2.3 Experimental Dataset

The second phase involves grouping participants by sex and PCL-C scores. In addition, a dimensionality reduction was generated for speech features.

As can be seen in Table 1, each participant has a PTSD severity score. The 97 participants with no severity are considered as the control group, and the 55 participants with high severity conform to the group with active PTSD. In this study, the moderate group is not considered because it could bias the indicators related to PTSD.

As previous studies have noted, there are differences in the automatic evaluations of mental disorders of women and men [10]. During subsetting, women and men were considered as different groups to recognize sex-based patterns linked to PTSD.

Therefore, this work is divided in three experiments. The general experiment is not sex-aggregated and has a total of **152** participants. The following experiments involve a women group (**60** participants), and a men group (**92** participants).

Subsequently, a dimensionality reduction approach was used to remove features that exhibit the same trends and to reduce the computation time. In this sense, Spearman correlation [18] was calculated, which is a non-parametric method for determining the correlation between two tests of measurements taken from the same individuals and is robust to outliers. According to Dancey and Reidy [18], the threshold for considering a high correlation ratio is 0.7. Finally, the general, women, and men experiments contain 53, 55 and 58 features, respectively.

2.4 Machine Learning Modeling

The last phase of the methodology consists of generating and evaluating the machine learning models, corresponding to the three experiments: general, women, and men.

Initially, supervised classification with a decision tree algorithm [19] was used to train and classify participants as either PTSD or non-PTSD (controls). This method, employing graphical decision rules, was chosen for its interpretability and feature recognition capabilities.

The parameters set were a binary tree approach, a minimum of two instances in leaves, no split subsets smaller than 5, a maximal tree depth of 5, and an unemployment rate when the majority reaches 95%. The decision trees were pruned to six levels for graphic representation.

Finally, the testing step evaluates the generated models to measure the performance. A ten-fold cross-validation was utilized, calculating performance measures including

F1 score, Precision, Recall and Specificity (see Eqs. 1 to 4). In these equations, true positive (TP) refers to correctly identified positives, while false positive (FP) represents incorrectly identified positives. By contrast, true negative (TN) accurately identified negative, while false negative (FN) signifies incorrectly identified negatives.

$$F_1 = \frac{\textit{Precision} \cdot \textit{Recall}}{\textit{Precision} + \textit{Recall}} \quad (1)$$

$$\textit{Precision} = \frac{\textit{TP}}{\textit{TP} + \textit{FP}} \quad (2)$$

$$\textit{Recall} = \frac{\textit{TP}}{\textit{TP} + \textit{FN}} \quad (3)$$

$$\textit{Specificity} = \frac{\textit{TN}}{\textit{TN} + \textit{FP}} \quad (4)$$

Data were processed using Orange, a robust data analysis platform [20].

3 Results and Discussion

The results are discussed into main two parts: classification performance of the models and the analysis of the features discovered by the models.

Table 2 presents the performance measures for classifying PTSD vs control group for general, women and men. The general model obtained an F1 score of 0.7, precision of 0.702, recall of 0.706 and specificity of 0.649. The women's model achieved an F1 score of 0.791, precision of 0.790, recall of 0.792, and specificity 0.756. The men's model attained F1 score of 0.733, precision of 0.743, recall of 0.742, and specificity score of 0.681. The results demonstrate improved performance when considering sex separation, with the women's model exhibiting the highest metrics.

Table 2. Performance measures results for each experiment

Experiment	Performance			
	F1 Score	Precision	Recall	Specificity
General	0.7	0.702	0.706	0.649
Women	0.791	0.790	0.792	0.756
Men	0.733	0.743	0.742	0.681

On the other hand, Figs. 2, 3 and 4 show the decision trees generated and blue shaded features representing the decision rules to classify PTSD. As can be seen in the three models, the 11th value Harmonic Model Phase Distortion-Mean (HMPDM-11) is the root node of all decision trees and the major indicators of PTSD in this analysis, unbiased by sex. This feature is included in a spectral group of parameters based on the harmonic

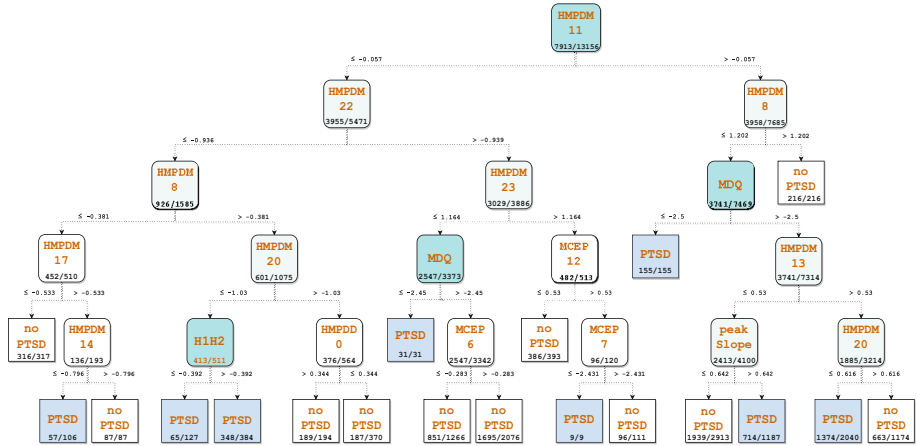


Fig. 2. Schematic of a decision tree generated from general experiment; it was pruned to six levels.

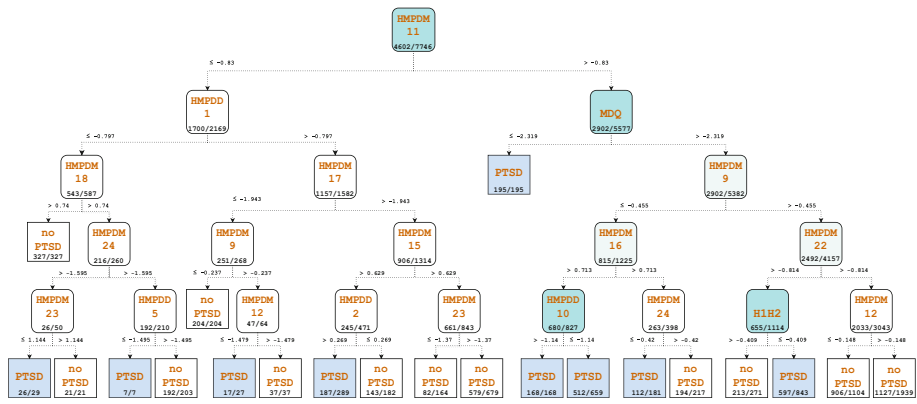


Fig. 3. Schematic of a decision tree generated from a men experiment; it was pruned to six levels.

model characterized by the phase distortion through mean. Particularly, these parameters are distinguished to reflect the uniformity of a signal.

Previous work reported similar patterns, Marmar et al. [12] selected voice markers that indicated more monotonous speech with less activation. Findings from this proposal might confirm that a person diagnosed with PTSD is likelier to have uniform speech.

Regarding the above, the research results have also demonstrated sex-based differences. The general model (see Fig. 2) found two groups of features, spectral and glottal source underlining the Maxima Dispersion Quotient (MDQ) due to generating a leaf node, which means that the classification process considers only three features. The MDQ is a glottal source feature linked to the sharpness of the glottal excitation used to quantify the dispersion. In addition, multiple features related to Harmonic Model Phase Distortion-Mean (HMPDM), and the first two harmonics of the differentiated glottal source spectrum (H1H2) were found.

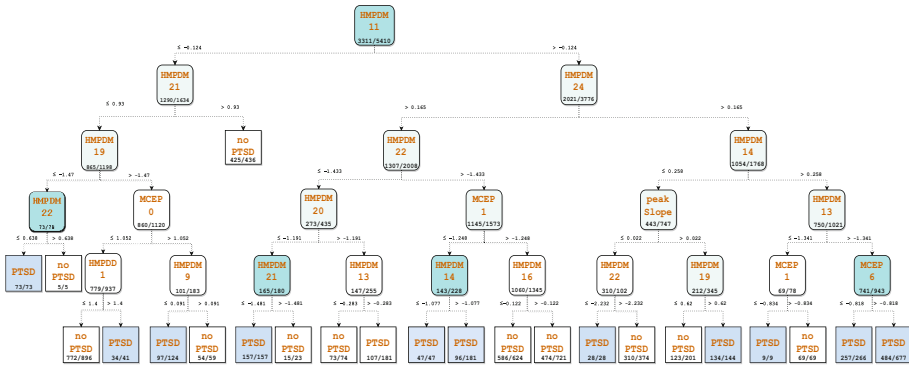


Fig. 4. Schematic of a decision tree generated from a women experiment; it was pruned to six levels.

The men’s model (see Fig. 3) also reflects features of the HMPDM group, and the MDQ feature increases its importance as a predictor since, to classify PTSD, the decision tree evaluates only MDQ and the 11th value HMPDM. The model also shows the 10th deviation HMPD and the H1-H2 feature, used to search changes in phonation and related to opening of the glottis phase, when a second harmonic is stronger than the first harmonic (negative value) indicates a pressed phonation and causes a tense and restricted voice.

The women’s model (see Fig. 4) finds multiple features of the HMPDM group, particularly, the 14th, 21st and 22nd values HMPDM. Similarly, the 6th Mel Frequency Cepstral coefficient is considered, these represent the short-term power spectrum in different frequency bands. This approach suggests that the voice of control or PTSD woman depends on the frequency band in which her speech is.

The previous mentioned findings provide new possibilities and implications in detecting PTSD patients, offering a quantitative diagnostic tool, and potentially enhancing the understanding of PTSD. As future trends, the proposal is to engage in the characterization of alternative data sources and use others interpretable automated algorithms.

4 Conclusion

This paper proposed a speech pattern recognition method to analyze sex-based indicators linked to PTSD through a decision tree algorithm.

Overall, it was proved that patterns could be found in the speech of patients with PTSD and promising performance of machine learning models, with subtle improvements when considering participant sex.

The results found indicators linked to PTSD based on speech, such as uniformity in a voice, recognized by all experiments. In addition, the findings suggest that indicators may vary by sex differences. The men’s model evaluates the dispersion of speech and if a voice is tense. In contrast, the women’s model focuses on the frequency bands of the speech.

Among the limitations of this study, even though promising performance metrics were obtained, it would be desirable to evaluate the models in a different database to test their stability.

As future work, it would be interesting to evaluate other possible sources of PTSD indicators, such as text data, facial movements, and among others; due to only DAIC-WOZ dataset features being considered in this analysis, there is a high potential to find other indicators with more defined patterns of PTSD.






References

1. American Psychiatric Association. <https://www.psychiatry.org/patients-families/ptsd>. Accessed 12 May 2023
2. World Health Organization. <https://www.who.int/news/item/06-08-2013-who-releases-guidance-on-mental-health-care-after-trauma>. Accessed 12 May 2023
3. Mitchell, K.S., Mazzeo, S.E., Schlesinger, M.R., Brewerton, T.D., Smith, B.N.: Comorbidity of partial and subthreshold PTSD among men and women with eating disorders in the national comorbidity survey-replication study. *Int. J. Eat. Disord.* **45**(3), 307–315 (2012)
4. Blake, D.D., Weathers, F.W., Nagy, L.M., Kaloupek, D.G., Charney, D.S., Keane, T.M.: Clinician-Administered PTSD Scale for DSM-IV. National center for posttraumatic stress disorder, Boston (1998).
5. Muhorakeye, O., Biracyaza, E.: Exploring barriers to mental health services utilization at kabutare district hospital of Rwanda: perspectives from patients. *Front. Psychol.* **12**, 638377 (2021)
6. Laxminarayan, S., Wang, C., Oyama, T., Cashmere, J.D., Germain, A., Reifman, J.: Identification of veterans with PTSD based on EEG features collected during sleep. *Front. Psych.* **11**, 532623 (2020)
7. Harnett, N.G., Goodman, A.M., Knight, D.C.: PTSD-related neuroimaging abnormalities in brain function, structure, and biochemistry. *Exp. Neurol.* **330**, 113331 (2020)
8. Schultebrucks, K., Qian, M., Abu-Amara, D., Dean, K., Laska, E., Marmar, C.R., et al.: Pre-deployment risk factors for PTSD in active-duty personnel deployed to Afghanistan: a machine-learning approach for analyzing multivariate predictors. *Mol. Psychiatry* **26**(9), 5011–5022 (2021)
9. Schultebrucks, K., Yadav, V., Shalev, A.Y., Bonanno, G.A., Galatzer-Levy, I.R.: Deep learning-based classification of posttraumatic stress disorder and depression following trauma utilizing visual and auditory markers of arousal and mood. *Psychol. Med.* **52**(5), 957–967 (2022)
10. Stratou, G., Scherer, S., Gratch, J., Morency, L.P.: Automatic nonverbal behavior indicators of depression and PTSD: exploring gender differences. In: 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, pp. 147–152. IEEE, Geneva (2013)
11. Sawalha, J., Yousefnezhad, M., Shah, Z., Brown, M.R., Greenshaw, A.J., Greiner, R.: Detecting presence of PTSD using sentiment analysis from text data. *Front. Psych.* **12**, 811392 (2022)
12. Marmar, C.R., et al.: Speech-based markers for posttraumatic stress disorder in US veterans. *Depress. Anxiety* **36**(7), 607–616 (2019)
13. Scherer, S., Stratou, G., Gratch, J., Morency, L.P.: Investigating Voice Quality as a Speaker-Independent Indicator of Depression and PTSD. *Interspeech*, pp. 847–851 (2013)

14. Gupta, S., Goel, L., Singh, A., Agarwal, A.K., Singh, R.K.: TOXGB: teamwork optimization Based XGBoost model for early identification of post-traumatic stress disorder. *Cogn. Neurodyn.* **16**(4), 833–846 (2022)
15. Scherer, S., Lucas, G.M., Gratch, J., Rizzo, A.S., Morency, L.P.: Self-reported symptoms of depression and PTSD are associated with reduced vowel space in screening interviews. *IEEE Trans. Affect. Comput.* **7**(1), 59–73 (2015)
16. Gratch, J., et al.: *The Distress Analysis Interview Corpus of Human and Computer Interviews*. University of Southern California Los Angeles (2014)
17. Weathers, F.W., Litz, B.T., Herman, D.S., Huska, J.A., Keane, T.M.: The PTSD Checklist (PCL): Reliability, validity, and diagnostic utility. In: *Annual Convention of the International Society for Traumatic Stress Studies*, vol. 462. San Antonio, TX (1993)
18. Dancey, C.P., Reidy, J.: *Statistics Without Maths for Psychology*, 8th edn. Pearson, London (2017)
19. Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., et al.: Top 10 algorithms in data mining. *Knowl. Inf. Syst.* **14**, 1–37 (2008)
20. Demšar, J., Curk, T., Erjavec, A., Gorup, Č, Hočevar, T., Milutinovič, M., et al.: Orange: data mining toolbox in Python. *J. Mach. Learn. Res.* **14**(1), 2349–2353 (2013)



Analysis of EEG Signals Recorded from Persons with Neuronal Alterations due to Covid-19 with Respect to the Cases Considered Normal

María Fernanda Arámburo-Castell¹ , María Monserrat Morín-Castillo¹  ,
Olga Guadalupe Felix-Beltrán¹ , and José Jacobo Oliveros-Oliveros² 

¹ Facultad de Ciencias de la Electrónica, BUAP, Puebla, Puebla 72570, México
maria.morin@correo.buap.mx

² Facultad de Ciencias Físico Matemáticas, BUAP, Puebla, Puebla 72570, México

Abstract. COVID-19 disease has caused a significant impact on the global health sector, due to its high rate of contagiousness, along with the repercussions that this disease has on the people who have suffered from it. Some types of sequels are often presented, among which stand out those affecting the nervous system. These have been detected due to the mental fog that some people present after their recovery, or when they present persistent symptoms, these can usually be analyzed by an electroencephalogram. In this work, we compare the electroencephalographic record of four persons affected by COVID-19 concerning records of this same type of people, considered with a normal signal. To be able to do this, it was necessary to elaborate an algorithm to recover the quantitative data of persons with alterations due to COVID-19, after which the analysis of this signal was performed together with normal cases. This analysis shows that the COVID-19 signal presents smaller values than variance, amplitude, zero crossings, and dominant frequency compared to normal cases.

Keywords: COVID-19 · EEG · Neurological sequels

1 COVID-19

1.1 Introduction

COVID-19 disease is the result of the SARS-CoV-2 virus, a member of the coronavirus family. Within this family, two other strains have greatly influenced the well-being of various geographic regions. One of these is the SARS-CoV virus, which emerged in 2002 in the Chinese province of Guangdong, infecting more than 8,000 people in 32 countries. Despite this, it did not reach pandemic status [1]. Similarly, in 2012, Middle East Respiratory Syndrome (MERS), caused by MERS-CoV, occurred in Jeddah, Saudi Arabia, with 2,499 cases reported [1]. These numbers, although significant, are small compared to the millions reached by SARS-CoV-2. They have contributed to improving the understanding and management of diseases caused by this virus.

On the 11th of March 2020, the World Health Organization (WHO) designated COVID-19 as a global pandemic owing to its substantial incidence rates [1]. This pandemic has received its conclusion designation; nevertheless, its consequences endure unabated. The effects of this disease have impacted the ordinary life of inhabitants worldwide, exerting a profound influence on sectors such as economics, education, and, particularly, public health. One reason is that even when people recover from the disease, in many cases, they present some sequels, often affecting various organs and systems, one of which is the nervous system [2].

1.2 EEG

Electroencephalography is a noninvasive technique used to study and diagnose the state of the nervous system and measure, on electrodes located on the scalp, the brain's electrical activity produced by clusters of neurons working simultaneously. These electrodes are placed following the international 10/20 system [3]. This system names the anatomical areas of the brain Fp (frontopolar), F (frontal), T (temporal), O (occipital), C (central), and P (parietal), and the various electrodes are numbered using even numbers for the left side and odd numbers for the right, with the lower numbers reflecting locations closer to the midline [3].

An integral aspect to remember when measuring this kind is establishing the precise configuration that will be utilized. These include monopolar montage, i.e., one electrode concerning another reference electrode, or bipolar, in which the measurement is performed concerning two electrodes that are not reference points [3].

The resulting measurement is a voltage versus time signal composed of four brain-wave groups (signals with different frequencies) and artifacts. Therefore, the signal is filtered before analysis to remove noise caused by external and internal signals. Once this has been done, the resulting signal is analyzed in the frequency domain to visualize the brainwaves. Those are alpha (8–12 Hz), beta (13–30 Hz), theta (4–7 Hz), and delta (<4 Hz), each of which appears in a normal case in specific circumstances and areas of the brain [3]. If the brainwaves appear in abnormal circumstances, it implies some neurological alteration.

1.3 Neurological Sequels

COVID-19 presents many sequels affecting several body parts, some of the most common being muscle pain, headache, loss of sense of taste or smell, among others. According to [4, 5], these sequels can be correlated with alterations in their EEG, which are usually epileptiform discharges, epileptic discharges, and slow waves in the frontal and temporal areas [5].

According to Furlannis [4], patients with some alteration in their EEG signal exhibit cognitive disturbance and occasionally some persistent symptoms, described in Table 1, which shows the persistent symptoms and the frequency of their occurrence, considering the patients with EEG alterations caused by COVID-19. It is observed that the most frequent symptoms are headaches, dizziness, and hyposmia. On the other side, it is also shown that the absence of these symptoms does not imply a normal EEG, especially considering that 53.85% did not present persistent symptoms.

Table 1. Neurological symptoms present concurrently with EEG abnormalities [4].

Neurological symptoms	Relative frequency (%)
Hyposmia	23.08
Tinnitus	15.38
Myalgia	7.69
Paresthesia	7.69
Headache	23.08
Dizziness	23.08
No symptoms	53.85

Also, in [4], different EEG abnormalities are mentioned and shown in Table 2 with their relative frequencies. The slow wave is the most frequent in this study. The abnormalities found in this study were also found in patients with COVID-19 [5–7].

Table 2. Electroencephalographic abnormalities [4].

Neurological symptoms	Relative frequency (%)
Slow wave	69.23
Epileptiform discharges	23.07
Epileptic discharges	7.69

From this, it is found that EEG abnormalities exist due to COVID-19 disease and the brain areas where they occur. According to data in [4, 5], frontal and temporal areas are the most frequent.

Everything described above corresponds to qualitative data, which allows knowing the characteristics of EEG impairments caused by COVID-19 documented by medical research. Although it is essential to analyze electroencephalographic signals qualitatively, it is not feasible due to the absence of available data [7, 8]. Thus, as an alternative way, using images were found that exemplify EEG affectations caused by COVID-19, from which it is possible to obtain quantitative data using image processing algorithms [8, 9].

2 Signal Analysis

2.1 Signal Recovery

As mentioned above, EEG images of (anonymous) patients affected by COVID-19 are available in different research articles. In this work, we use an EEG image with the described features. Signal data was extracted from the available images using image

processing functions accessible in Python. We consider some relevant characteristics needs, like having around 200 pixels per second and a scale, which is essential to obtain a voltage/pixel ratio that allows a proper conversion of the values. In some cases, EEG recordings have overlapping channels, which would interfere with the conversion, so it is necessary to separate each channel into a new image, where the lowest value of the signal is in the last column of the image, considering it like a matrix. Another important consideration is to create an image corresponding to the scale measurements, which will allow obtaining the contribution of each pixel in the voltage and check that it matches the resolution mentioned above.

So, an algorithm was implemented using the scale and channel images, considering these as a $n \times m$ matrix [10], and generating a vector ranging from 1 to n and multiplying this by the *pixel/voltage* correlation factor. Fig. 1 shows the flowchart of the proposed algorithm. The algorithm was implemented in Python using the OpenCV, NumPy, and Pandas modules.

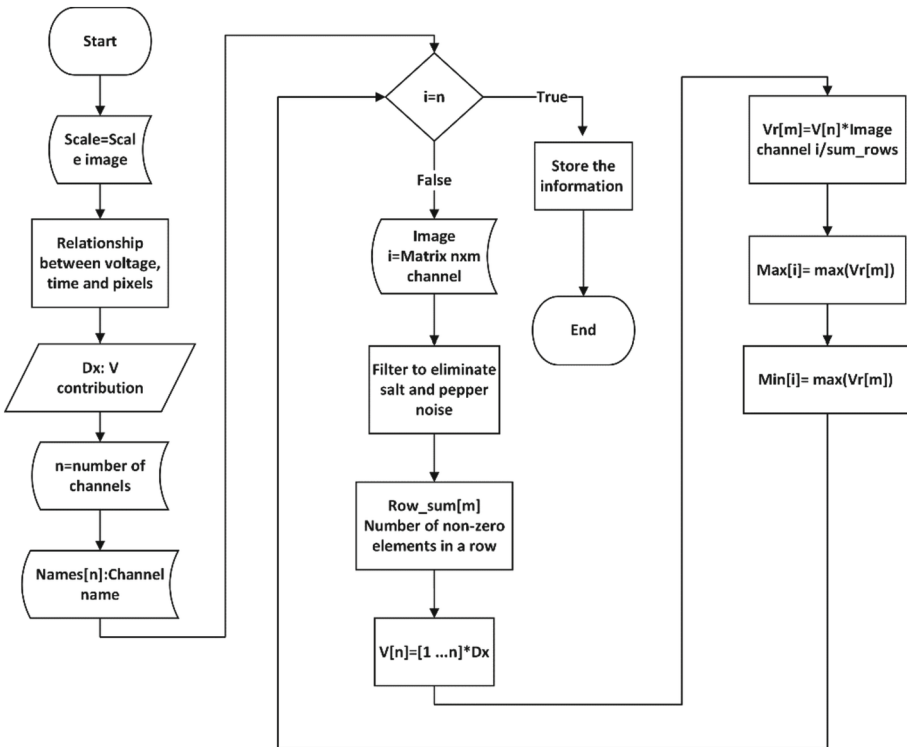


Fig. 1. Signal recovery algorithm flowchart.

Example. Using a signal from the NMT database, available in [10], plotted with Python to obtain an image with the previously mentioned characteristics. Once the image and the signal scale have been obtained, it is possible to use signal restoration from the

image. Fig. 2 shows the restored signal in blue and the original signal in red. We can see overlapped signals.

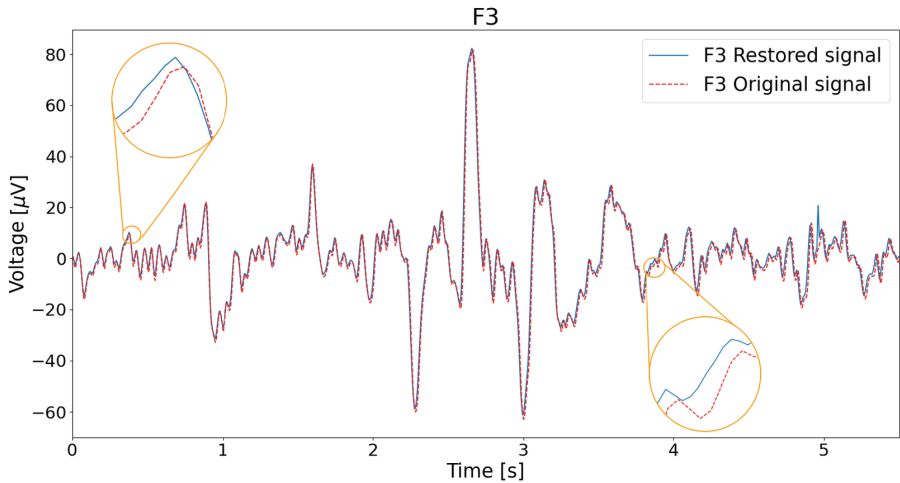


Fig. 2. Original and recovered signals.

Table 3 shows the main statistical values of both signals, where there are some differences. In the cases of the maximum and minimum, the error is $1.593 \mu\text{V}$ and $0.4744 \mu\text{V}$ respectively, which implies an error of 2.53% and 0.57% in each case.

Table 3. Electroencephalographic averages values.

Parameters	Original signal	Restored signal
Samples	1401	1817
Average	$-0.2054 \mu\text{V}$	$0.662 \mu\text{V}$
Standard deviation	15.4474	15.231
Minimum	$-62.751 \mu\text{V}$	$-61.158 \mu\text{V}$
Maximum	$81.8656 \mu\text{V}$	$82.340 \mu\text{V}$

2.2 Statistics Analysis for Cases of Study

Once the signal data are available, extracting them quantitatively for each measurement is possible. In this sense, each EEG channel for the studied cases is analyzed in terms of its parameters in the time domain, such as standard deviation, median, maximum, minimum, variance, skewness, kurtosis, amplitude, and zero crossings. These parameters are used in

other EEG analyses [11, 12]. Moreover, in this type of signal, the dominant frequency is fundamental since it allows for finding neurological abnormalities [3]. These parameters allow for highlighting the qualitative characteristics of the signal quantitatively. In this research, it is observed the parameters for the cases of healthy and affected by COVID-19 persons. The cases of healthy persons will be analyzed NMT database [13], and in the cases of persons with COVID-19, the signals retrieved from the article [6].

For the statistical analysis, it is imperative to consider that the NMT database employs a monopolar montage, whereas the COVID-19 cases utilize a bipolar configuration. To make a feasible compare is necessary to retake the database NMT and get the difference between the channels Fp2- C4, C4-O2, T4-O2, Fp1-C3, C3-O1, Fp1-T3, and T3-O1 to achieve a similar type of measurement that was performed in the patient's EEG with COVID-19, such that all the recordings have signals coming from the corresponding location.

Four healthy cases will be contrasted with four cases affected by COVID-19. To achieve this, database files exhibiting analogous attributes to those of the COVID-19-affected cases, matched explicitly in gender and approximate age, will be employed [6, 11]. We use the parameters mentioned for each measurement on each record, especially amplitude and dominant frequency, which are usually signal alterations caused by COVID-19 [6].

3 Results

As mentioned above, abnormalities attributed to COVID-19 typically manifest in the temporal and frontal brain regions. Consequently, it is important to investigate the channel signals within EEG records. Fig. 3 shows the first seconds in the channels: Fp2-C4, C4-O2, T4-O2, Fp1-C3, C3-O1, Fp1-T3, and T3-O1 of a healthy patient, and one affected by COVID-19. In both cases, they are signals from women aged 83 and 84 years, respectively. It can be seen how the signal affected by COVID-19 has a smaller amplitude and a lower frequency than the healthy case, which translates into fewer zero crossings. Similar results were found when comparing the other cases.

Statistical parameters were retrieved and averaged across all channels, encompassing four healthy cases and four COVID-19-affected cases, as illustrated in Table 4. The data shows that the amplitude, variance, dominant frequency, and zero crossings exhibit notably higher values in healthy cases compared to those affected by COVID-19.

Figure 4 depicts the average dominant frequency observed in the channels studied. In particular, the dominant frequency shows variations, especially in the temporal and occipital regions, with disparities of approximately 10 Hz more pronounced in individuals without COVID-19 than those affected by the disease.

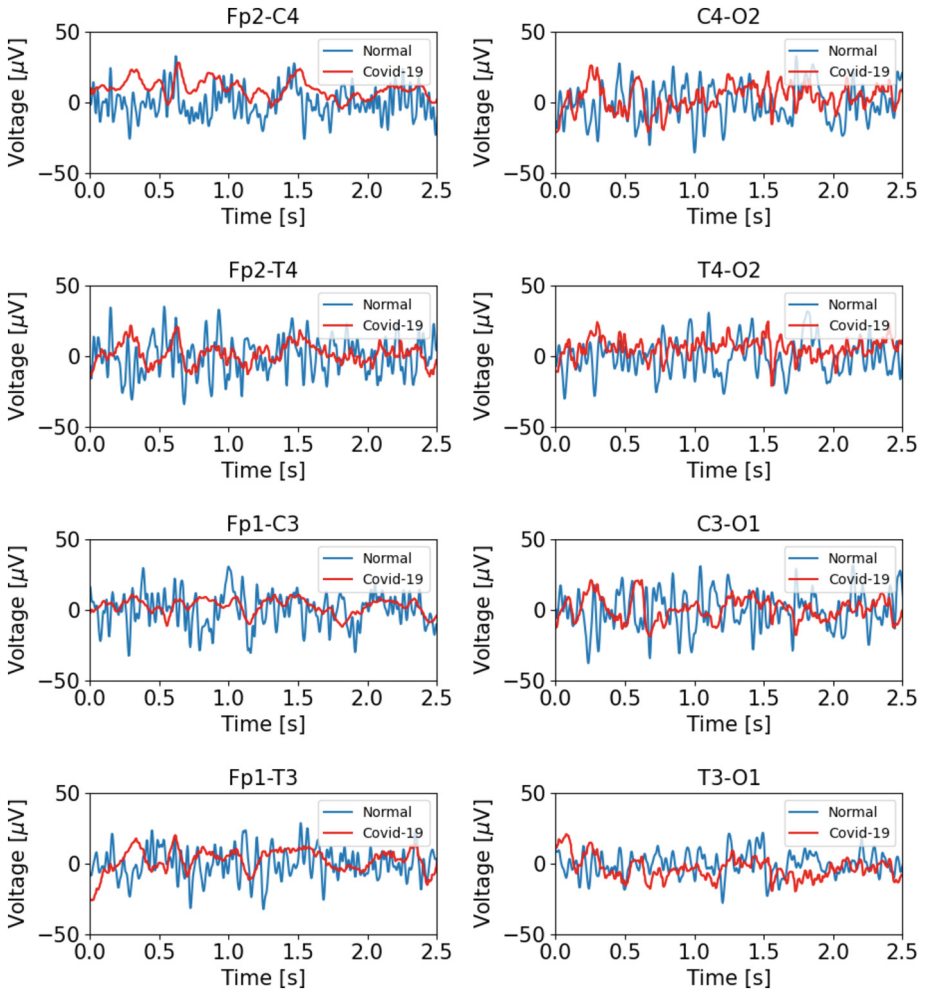


Fig. 3. EEG signals of a person considered normal, and a person affected by COVID-19.

Drawing from [4, 5], the regions exhibiting more pronounced alterations because of COVID-19 are primarily the temporal and frontal areas. Hence, the statistical parameters of the Fp1-T3 and Fp2-T4 channels are highlighted to establish a contrast with the values of the reference cases. Table 5 presents the values pertinent to the right hemisphere, while Table 6 encompasses the signals from the left hemisphere; both tables show lower values in terms of amplitude, variance, dominant frequency, and zero crossings in the instances of COVID-19-affected cases, in alignment with the data presented in Table 4.

Table 4. Average statistical parameters for the cases: a) normal and b) COVID-19 persons.

Parameters	Normal	COVID-19
Average	0.0414 μ V	3.0949 μ V
Standard deviation	-1.5627	3.9456
Median	0.7996 μ V	2.9757 μ V
Maximum	344.0706 μ V	119.04328 μ V
Minimum	-535.2224 μ V	-110.7787 μ V
Variance	1862.8078	244.4142
Skewness	-3.0685	0.1019
Kurtosis	33.5497	6.3346
Amplitude	879.2930 μ V	229.8219 μ V
Zero crossings	1260	496.5
Dominant frequency	10.6428 Hz	4.84093 Hz

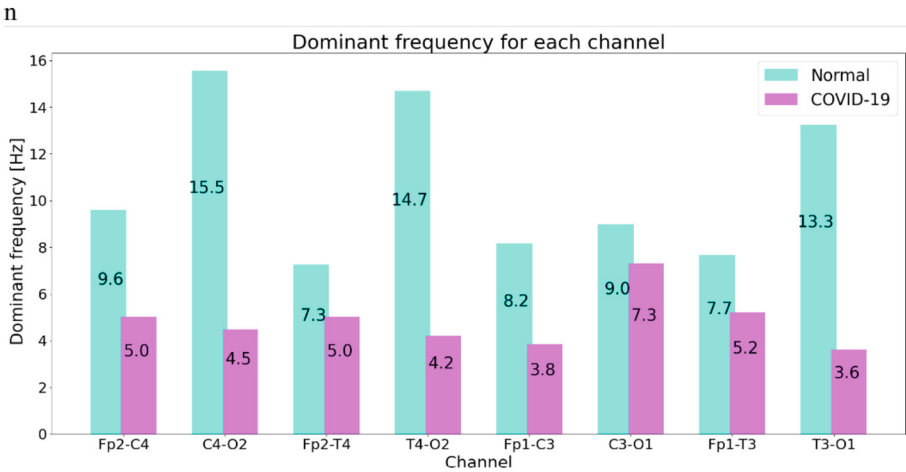


Fig. 4. Dominant frequency for each channel for the cases: a) normal and b) COVID-19 persons.

Table 5. Statistical parameters for the cases: a) normal and b) COVID-19 persons in channel Fp2-T4.

Parameters	Normal	COVID-19
Average	-0.3952 μV	7.0974 μV
Standard deviation	-0.6926	9.1264
Median	1.2796 μV	6.8930 μV
Maximum	474.4208 μV	118.9079 μV
Minimum	-865.4444 μV	-112.0145 μV
Variance	3,159.8254	250.4935
Skewness	-6.2287	0.0779
Kurtosis	62.6974	4.9386
Amplitude	1,339.8652 μV	230.9224 μV
Zero crossings	1211	579
Dominant frequency	7.2571 Hz	5.0111 Hz

Table 6. Statistical parameters for the cases: a) normal and b) COVID-19 persons in channel Fp1-T3.

Parameters	Normal	COVID-19
Average	-0.0133 μV	8.8855 μV
Standard deviation	-3.4262	5.9624
Median	0.2555 μV	9.3062 μV
Maximum	449.4642 μV	119.9983 μV
Minimum	-834.8713 μV	-114.5507 μV
Variance	2,671.6986	261.2239
Skewness	-7.1214	-0.5069
Kurtosis	72.7827	6.5072
Amplitude	1,284.3355 μV	234.5490 μV
Zero crossings	1304	634
Dominant frequency	7.6571 Hz	5.2180 Hz

4 Discussion and Conclusions

From the above results, it can be affirmed that the main alteration of EEG signals in patients with Covid-19, for the cases studied, is observed in their dominant frequency. This observation aligns with the diagnoses depicted in the study [6]. As a result, we get analogous conclusions regarding the recovered signal to those drawn from the original signal.

We also found that the abnormalities in the signal caused by the disease also affect their amplitude, as mentioned in [6], being much lower than the normal cases used in this work. Although other studies [14] mention alterations in the frontal and temporal areas, it is interesting to note that in the cases reviewed, the frequency alterations were mainly in the occipital and temporal areas of the right side of the brain, which is different from what is shown in the Tables of the articles [3, 5].







Therefore, it is important to highlight that no conclusions can be drawn due to the small number of cases studied. However, this work could contribute to a more in-depth study of these signals. The absence of these databases is an important limitation. Although it is possible to reconstruct the EEG signal from images, they do not substitute the use of a sturdy database with greater detail and a more significant number of records.

References

1. Alsafi, R.: Lessons from SARS-CoV, MERS-CoV, and SARS-CoV-2 infections: what we know so far. *Can. J. Infect. Diseases Med. Microbiol.* (2022)
2. Mao, L., et al.: Neurologic manifestations of hospitalized patients with coronavirus disease 2019 in Wuhan, China. *JAMA Neurol.* **77**(6), 683–690 (2020)
3. William, O.: *Handbook of EEG Interpretation*, 3rd. Demos Medical, Florida (2022)
4. Furlanis, G., et al.: Cognitive deficit in post-acute COVID-19: an opportunity for EEG evaluation? *Neurol. Sci. Offic. J. Italian Neurol. Soc. Italian Soc. Clin. Neurophysiol.* **44**(5), 1491–1498 (2023)
5. Hwang, S., et al.: EEG Abnormalities and their radiographic correlates in a COVID-19 inpatient cohort. *Neurol. Clin. Pract.* **12**(1), 52–59 (2022)
6. Sáez-Landete, I., et al.: Retrospective analysis of EEG in patients with COVID-19: EEG recording in acute and follow-up phases. *Clin. EEG Neurosci.* **53**(3), 215–228 (2022)
7. Petrescu, A., Taussig, D., Bouillere, V.: Electroencephalogram (EEG) in COVID-19: a systematic retrospective study. *Clin. Neurophysiol.* **50**(3), 155–165 (2020)
8. Antony, A., Haneef, Z.: Systematic review of EEG findings in 617 patients diagnosed with COVID-19. *Seizure* **83**, 234–241 (2020)
9. Perrottelli, A., et al.: Cognitive impairment after post-acute COVID-19 infection: a systematic review of the literature. *J. Personal. Med.* **12**(12) (2022)
10. Rosebrock, A.: *Practical python and OpenCV: an introductory, example driven guide to image processing and computer vision*. 3rd ed. Pyimagesearch (2016)
11. Shirley, G., Jerritta, S.: Time domain analysis of electroencephalogram (EEG) signals for word level comprehension in deaf graduates with congenital and acquired hearing loss. In: *IOP Conference Series: Materials Science and Engineering*, vol. 1070 (2021)
12. Iqbal, S., Muhammed Shanir, P., Khan, Y.U., Farooq, O.: Time domain analysis of EEG to classify imagined speech. In: Satapathy, S., Raju, K., Mandal, J., Bhateja, V. (eds.) *Proceedings of the Second International Conference on Computer and Communication Technologies. AISC*, vol. 380. Springer, New Delhi (2016). https://doi.org/10.1007/978-81-322-2523-2_77
13. Khan, H., et al.: The NMT scalp EEG dataset: an open-source annotated dataset of healthy and pathological EEG recordings for predictive modeling. *Front. Neurosci.* (2022)
14. Appelt, P., et al.: Changes in electrical brain activity and cognitive functions following mild to moderate COVID-19: a one-year prospective study after acute infection. *Clin. EEG Neurosci.* **53**(6), 543–557 (2022)



Automated Segmentation of Breast Skin for Early Cancer Diagnosis: A Multi-otsu Region Growing Approach for Detecting Skin Thickness Variations

Luis Antonio Medina-Ramos¹ , Alfonso Toriz-Vázquez² ,
Yazmín Mariela Hernández-Rodríguez³ , Rafael Bayareh-Mancilla³ ,
and Oscar Eduardo Cigarroa-Mayorga³  

¹ CICATA-National Polytechnic Institute, Legaria 694, 11500, Mexico, Mexico

² Academic Unit of the Institute of Applied Mathematics and Systems Research, UNAM, 97357 Ucu, 97302 Merida, Yucatan, Mexico

³ Department Advanced Technologies, UPIITA-Instituto Politécnico Nacional, Av. IPN No. 2580, 07340 Mexico, Mexico
ocigarroam@ipn.mx

Abstract. Accurate segmentation of breast skin in mammograms is of great interest for computer-aided diagnosis systems aiming for early cancer detection, as it can highlight contralateral asymmetries linked to the development of different types of carcinomas. We present an automated skin segmentation algorithm based on Otsu multi-thresholding. It was evaluated on a dataset of 102 pairs of mammograms from females of ages ranging from 30 to 74 years. We then proposed two novel characteristics for asymmetry assessment: mean skin thickness, which resulted in 0.88 pixels long, with a standard deviation of 1.16 pixels; and skin area, with an average of 285.79 pixels and a standard deviation of 399.35 pixels. Under these considerations, the algorithm identified 26 out of 102 cases (25.49%) as exhibiting asymmetry in skin thickness, and 24 out of 102 cases (23.52%) as displaying variations in skin area. These results demonstrate the effectiveness of the proposed algorithm in accurately segmenting breast skin and detecting potential asymmetries. Our advanced skin segmentation method enhances breast imaging for cancer detection. Analyzing BI-RADS correlation with skin thickness reveals significant asymmetries, aiding early diagnosis. Our innovative approach outperforms previous techniques, although sample size and subjectivity warrant consideration.

Keywords: Automatic segmentation · Breast cancer · Multi-otsu segmentation · Skin thickness

1 Introduction

Breast cancer is a pressing global health issue. Each year, more than 2 million new cases of breast cancer are diagnosed worldwide [1]. Studies have consistently shown that identifying breast cancer at an early stage significantly increases the chances of

treatment, with a rate of 90% of successful survival [2, 3]. One common characteristic observed in breast cancer is the occurrence of contralateral asymmetry. Notably, such asymmetry often serves as an early indication of the disease. Some breast cancer related to asymmetry involves Invasive Carcinoma, and Inflammatory Breast tissues, each with distinct clinical manifestations and prognoses [4, 5]. There is a 12.8% likelihood that the individual may develop breast cancer. Moreover, if the degree of asymmetry exceeds 20%, it can be regarded as a potential indication of breast cancer with a 99% confidence level [6–9]. Skin pattern thickness has emerged as a potential indicator for detecting early signs of breast cancer. Variations in skin pattern thickness have been shown to correlate with underlying pathological changes. By analyzing and quantifying these subtle differences in skin pattern thickness, our study aims to contribute to the development of non-invasive and accurate diagnostic tools. This approach holds promise for enhancing the precision and reliability of breast cancer detection, ultimately leading to improved patient outcomes and more effective healthcare interventions [10]. Table 1 presents an overview of the current state-of-the-art in terms of algorithmic advancements in skin segmentation techniques.

As depicted in Table 1, algorithms applied to mammograms attempt to identify the skin-line rather than the complete skin tissue, except in the case of MRI images where it is often excluded in favor of internal tissues [14]. Moreover, neither of these techniques attempts to correlate skin characteristics with a diagnosis, as is presented in our work. The Breast Imaging Reporting and Data System (BI-RADS) classification could potentially hold a significant correlation with contralateral asymmetries. Therefore, the exploration of algorithms for skin segmentation in mammograms and the quantification

Table 1. Table summarizing the state of the art related to automatic segmentation algorithms for breast skin.

Description	Algorithm	ROI	Reference
Edge detection algorithms, scale space concepts, and a growing active contour method. Applied to mammograms	Segmentation process with a seed point Contour growth is based on the gradient factor	Skin-line	Martí R. <i>et al.</i> [11]
Segmentation of the skin and dense tissue in MRI to produce a biomechanical model of the breast	C-means (4 classes)	Skin	Solves Llorens J. A., <i>et al.</i> [12]
Fractal and Hausdorff Distance analysis applied to breast skin line segmentation in mammograms	Fractal-based segmentation	Skin-line	Don. S., Min. D [13]

(continued)

Table 1. (continued)

Description	Algorithm	ROI	Reference
An automated skin segmentation algorithm was developed to remove the skin in MRI before 3D reconstruction	Active contour model	Skin	Lee C. <i>et al.</i> [14]
Automatic Breast Segmentation algorithm for low contrast, indistinct boundaries, noise, and artifacts. Used on mammograms	Iterative movement of the original seed along the diagonal reference line	Skin-line	Zhang Z., Lu J., Yip Y. [15]

of differences in skin thickness across contralateral breasts within a BI-RADS classification dataset provide novel and valuable insights into early breast cancer detection and diagnosis.

2 Methods

2.1 BI-RADS Database

The BI-RADS dataset used in this study consists of Medial Lateral Oblique (MLO) and Craniocaudal (CC) projections, including left and right views for each patient. The dataset involves a total of 102 patients, all women with ages ranging from 30 to 74 years, and categorized by oncologists from B1 to B5. The data used in this study was sourced from the University of Pittsburgh Text Information Extraction System (TIES) [16].

2.2 Multi-otsu Region Segmentation

Otsu segmentation is a widely used image segmentation technique that aims to automatically determine optimal thresholds for separating regions of interest according to their brightness. Otsu's algorithm calculates the optimal thresholds by maximizing the between-class variance, which represents the separability between the two classes [17].

In the initial step, the image is extracted from the DICOM file, which is subsequently converted to a grayscale format. Then, a tissue contrast enhancement technique, known as The Look-Up Table Region of Interest (VOILUT), was applied. VOILUT technique is an image processing tool utilized in mammography to enhance the visualization of breast lesions, mapping different shades of gray to adjust the image intensity and accentuate specific features such as the skin [18, 19]. To focus solely on the breast region, a mask is created to isolate it from the surrounding areas. To ease computational procedures and maintain consistency, every image is resized to a spatial resolution of 512×512 pixels. Figure 1 depicts the differences between the original mammography image, the VOI LUT process, and a mask enclosing the breast.

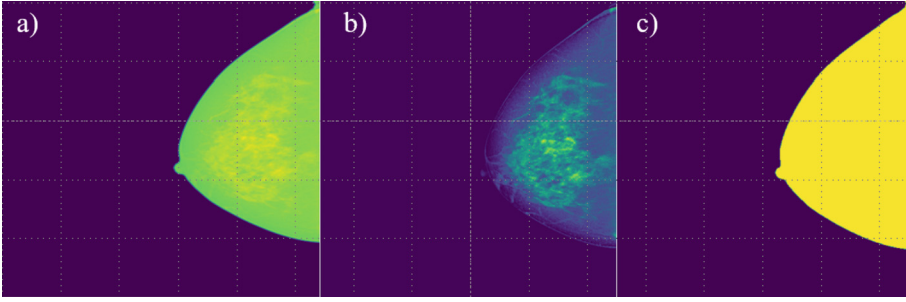


Fig. 1. Preprocessing steps applied to a mammography image, a) the mammogram is extracted from the DICOM file, exhibiting low tissue contrast, b) VOI LUT result, c) breast mask. While applying VOI LUT is not essential for obtaining the mask, it becomes necessary for differentiating the intensity of the skin from other tissues.

For the second step, a 13×13 kernel was used to perform an erosion operation. Next, by applying element-by-element XOR operation between the mask and the eroded image, only the pixels closer to the edges of the breast are retained. A histogram is computed using only these pixels. Subsequently, the Otsu method is employed to determine an optimal thresholds. The optimum thresholds computed that maximize the variance is given by Eq. (1). In Fig. 2, we can observe the variation in variance as the threshold is adjusted across the intensity values.

$$\sigma_G^2 = \sum_{i=0}^{L-1} (i - m_G)^2 P_i \quad (1)$$

where σ_G^2 is the global intensity variance of all the pixels in the image. For the particular case of having 2 classes, σ_B^2 is also given by Eq. (2)

$$\sigma_B^2 = P_1 P_2 (m_1 - m_2)^2 = \frac{(m_G P_1 - m)^2}{P_1 (1 - P_1)} \quad (2)$$

where m_G is the global intensity average, and m_1 and m_2 are the averages of both classes, P_1 is the probability calculated with the summation from 0 to the hypothetical threshold over the normalized histogram, and P_2 would be from the threshold up to 255, which is the maximum value. Next, the variances are calculated for each hypothetical threshold value, around all possible values divided into four different thresholds. The threshold that yields the maximum variance is determined, indicating the optimal point where the image is segmented. Algorithm I shows the pseudocode that presents the implementation described in this section.

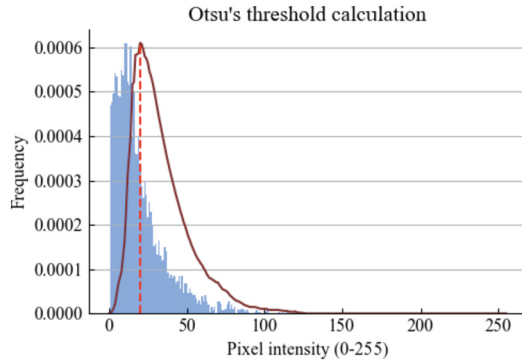


Fig. 2. This graphical representation provides insights into the selection of an optimal threshold, as it allows for the identification of the threshold value that maximizes the variance. The dashed red line indicates the location where the maximum value occurs for the pixel collection.

Algorithm I: skin segmentation by Otsu multi -thresholding

Inputs:

- Breast image I
- Erosion kernel size N

Outputs:

- Skin image mask S

- 1 Create binary matrix Z from the image turning any pixel value greater than 0 into a 1
 - 2 Apply element-by-element XOR operation between E and Z and save as X.
 - 3 Apply X as a mask to I, creating a border matrix B
Create new list L where non-black pixel values will be collected
 - 4 **for** every pixel P in B:
 - 5 **if** P > 0
 - 6 **Append** P to L
 - 7 Ts = threshold_multiotsu(L, 5); we use 5 as argument to produce 4 thresholds (for 5 classes) and store them to the T list
 - 8 Make the final threshold Tf equal to the second element of Ts.
 - 9 Generate mask F by evaluating the cases where B > Tf
 - 10 Create a kernel K2 of ones having size 5x5
 - 11 Erode the Z image using K2 in a new R matrix.
 - 12 Combine mask F and R with an element-by element OR operation to include the most outer pixels, and store as S
-

2.3 Analyzing BI-RADS Cases and Breast Skin Asymmetry

After acquiring the skin segmentations, a study was conducted to assess the correlation between BI-RADS cases exhibiting skin asymmetry. The evaluation revolved around

computing the average difference between skin pixels and the area of the left and right breast. The underlying premise was that in the presence of asymmetry, both breasts would exhibit high differences. Figure 3 presents a comparison between a case displaying asymmetry and another case demonstrating normality.

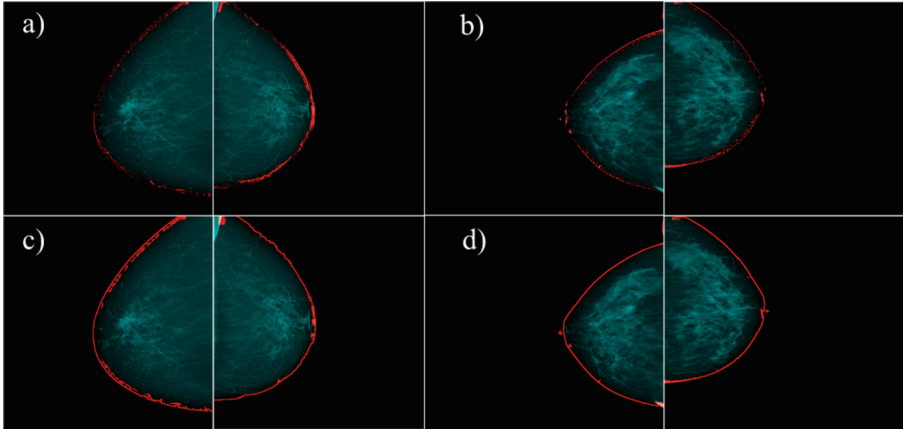


Fig. 3. Segmentation cases, a) skin symmetry of patient with BI-RADS1 with single threshold, b) skin asymmetry of a patient with BI-RADS2 with single threshold, c) and d) cases of skin symmetry and asymmetry with class division respectively.

3 Results

Preliminary results obtained from segmentation techniques based solely on a single class can be observed in Fig. 4. Some skin pixels that are slightly distant from the main border fail to be included in the initial boundary delineation.

Therefore, the pixels were divided into five classes based on these thresholds, depicted in Fig. 6a). Pixels closer to the skin are red, while those closer to blue are not part of the skin. To refine the segmentation, a new mask is generated by including only the pixels within the top three classes of highest intensity. Additionally, to ensure that the boundary pixels are included in the segmentation, they are combined with a small edge obtained through the erosion of three pixels from the original breast region. Figure 5 presents a histogram in which the threshold values are computed for separate five classes where: 17, 40, 90, and 179.

Figure 6 showcases a result of the skin segmentation process applied to the original mammography. It is important to note that skin segmentation does not yield a uniformly defined border. The resulting segmentation highlights the intricate and non-uniform nature of the skin region.

In terms of validation concerning the labeled BIRADS cases, our paper presents the results regarding the level of variation in skin thickness between the contralateral sides of each patient. These findings shed light on the extent of differences observed in skin

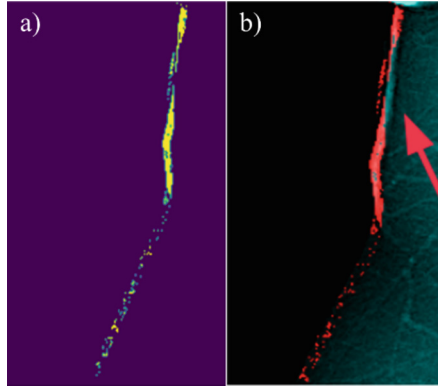


Fig. 4. Results of skin segmentation with a single class, a) shows the pixels that were considered part of the skin, while b) shows the segmentation on the original mammogram. The red arrow denotes skin regions that were omitted.

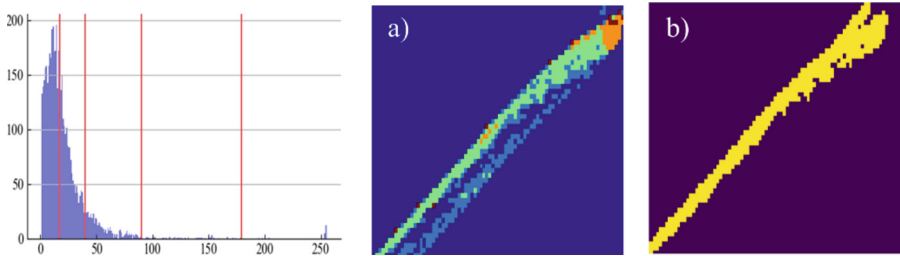


Fig. 5. Segmentation results with the distinction of five different classes, a) presents the segmented subregions in which 5 classes can be observed within different colors and b) mask segmentation.

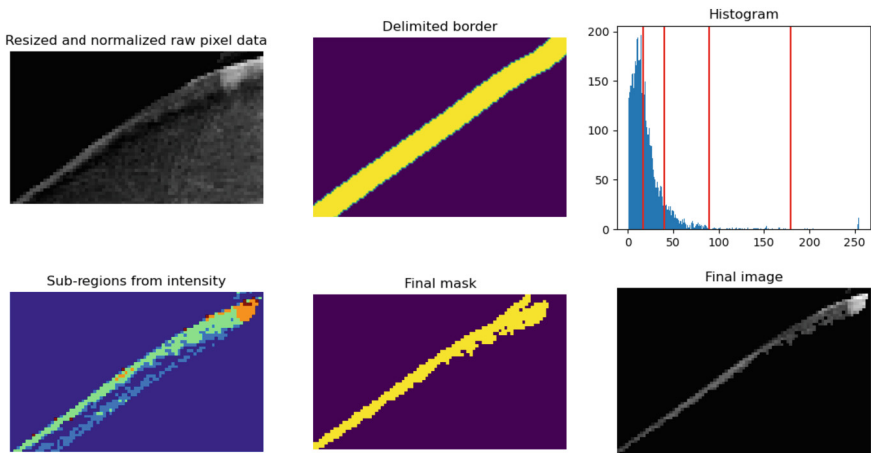


Fig. 6. Processing results for segmenting breast skin based on Multi-Otsu Region Growing Approach.

thickness between the corresponding sides of the breasts in Table 2, presented in Fig. 7 and Table 3 presents the cases found concerning each BI-RADS type.

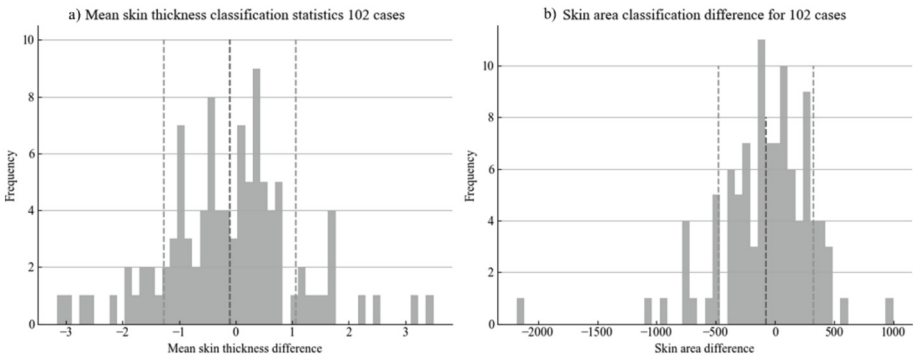


Fig. 7. Results of the analysis, a) average skin thickness differences and b) difference according to skin area.

Table 2. Results of the analysis for BI-RADS cases. The cases considered asymmetric are those that present anomalies in the dataset whose value exceeds the σ threshold.

Statistics	Skin thickness	Skin area
Pairs analyzed	102	102
μ	0.88	285.79
σ	1.16	399.35
Asymmetry cases	26/102	24/102
Asymmetry %	25.49	23.52

Table 3. Results divided by BI-RADS cases.

BI-RADS	Number of cases	Number of asymmetries (Mean difference)	Number of asymmetries (Area difference)
1	19	4	2
2	7	2	2
3	33	10	10
4	31	5	7
5	11	5	2

4 Discussion

The results of our study underscore the significance of skin segmentation within breast imaging, for potential application in early cancer detection. Our methodology addressed challenges posed by tissue overlap and indistinct boundaries. We assessed the association between BI-RADS classification and variations in skin thickness, revealing substantial dissimilarities across contralateral breasts. These disparities serve as markers for asymmetry and abnormalities. However, it is essential to acknowledge the limitations inherent in our study. The relatively limited sample size employed might not comprehensively represent the broader population, prompting the necessity for further investigation using larger, more diverse datasets.

Comparing our findings, it is evident that our approach represents a distinct advancement. The incorporation of advanced techniques, including multiclass segmentation, allows us to achieve a more refined delineation of skin regions. This is a substantial departure from previous works that predominantly focused on identifying the skin-line. Our exploration of skin thickness and area as indicators of breast health further sets our study apart, enabling a comprehensive assessment of asymmetries.

5 Conclusion

Based on our findings, an asymmetry between contralateral sides in terms of skin thickness was found. Notably, 25.49% of cases exhibited variations in skin thickness, while 23.52% displayed differences in skin area. The cases considered asymmetric are those that present anomalies in the dataset whose value exceeds the σ threshold. Furthermore, our study highlights the potential of using a multiclass approach to enhance skin segmentation. By assigning different classes to distinct regions based on the initial segmentation, we can achieve more accurate and refined results. This multiclass segmentation approach enables a finer delineation of boundaries and aids in the identification of specific areas of interest within the breast images. These findings underscore the value of implementing advanced techniques, such as the systematic testing of threshold combinations, to improve skin segmentation in breast imaging analysis. By incorporating these techniques into computer-aided diagnosis systems, we can enhance the detection and diagnosis of breast abnormalities, ultimately leading to improved patient care and outcomes.

Acknowledgements. The authors acknowledge the Mexican *Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT)* for the economic support to this research through the project FORDEC-YT-PRONACES/6005/20. The authors also thank the *Secretaría de Investigación y Posgrado* of the *Instituto Politécnico Nacional (IPN)* for the partial economic support to this research. The research was carried out in the *Laboratorio de Sistemas para Diagnóstico y Tratamiento de Cáncer* at *UPIITA-IPN*.

References

1. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A.: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424 (2018). <https://doi.org/10.3322/CAAC.21492>
2. Chávarri-Guerra, Y., et al.: Breast cancer in Mexico: a growing challenge to health and the health system. *Lancet Oncol.* **13** (2012). [https://doi.org/10.1016/S1470-2045\(12\)70246-2](https://doi.org/10.1016/S1470-2045(12)70246-2)
3. Marmot, M.G., Altman, D.G., Cameron, D.A., Dewar, J.A., Thompson, S.G., Wilcox, M.: The benefits and harms of breast cancer screening: an independent review: a report jointly commissioned by Cancer Research UK and the Department of Health (England) October 2012. *Br. J. Cancer* **108**, 2205 (2013). <https://doi.org/10.1038/BJC.2013.177>
4. Kim, M.Y., Choi, N.: Mammographic and ultrasonographic features of triple-negative breast cancer: a comparison with other breast cancer subtypes. *Acta radiol.* **54** (2013). <https://doi.org/10.1177/0284185113488580>
5. Pediconi, F., et al.: Breast imaging and cancer diagnosis during the COVID-19 pandemic: recommendations from the Italian College of Breast Radiologists by SIRM. *Radiologia Medica.* **125** (2020). <https://doi.org/10.1007/s11547-020-01254-3>
6. Hanafy, M.M., Ahmed, A.A.H., Ali, E.A.: Mammographically detected asymmetries in the era of artificial intelligence. *Egypt. J. Radiol. Nuclear Med.* **54** (2023). <https://doi.org/10.1186/s43055-023-00979-1>
7. Skaane, P.: Malignant mammographic asymmetric densities without US correlate remain challenging. *Radiology* **302**, 533–534 (2022). <https://doi.org/10.1148/RADIOL.212628>
8. Scutt, D., Manning, J.T., Whitehouse, G.H., Leinster, S.J., Massey, C.P.: The relationship between breast asymmetry, breast size and the occurrence of breast cancer. **70**, 1017–1021 (2014). <https://doi.org/10.1259/BJR.70.838.9404205>
9. Price, E.R., Joe, B.N., Sickles, E.A.: The developing asymmetry: revisiting a perceptual and diagnostic challenge. *Radiology* **274**, 642–651 (2015). <https://doi.org/10.1148/RADIOL.14132759>
10. Wong, S., Kaur, A., Back, M., Lee, K.M., Baggarley, S., Lu, J.J.: An ultrasonographic evaluation of skin thickness in breast cancer patients after postmastectomy radiation therapy. *Radiat. Oncol.* **6**, 1 (2011). <https://doi.org/10.1186/1748-717X-6-9>
11. Martí, R., Oliver, A., Raba, D., Freixenet, J.: Breast Skin-Line Segmentation Using Contour Growing
12. Solves Llorens, J.A., Rupérez, M.J., Monserrat, C., Feliu, E., García, M., Lloret, M.: Segmentation of the breast skin and its influence in the simulation of the breast compression during an X-ray mammography. *Sci. World J.* **2012** (2012). <https://doi.org/10.1100/2012/876489>
13. Don, S., Min, D.: Breast skin line segmentation on digital mammogram using fractal approach. *Indian J. Sci. Technol.* **9** (2016). <https://doi.org/10.17485/ijst/2016/v9i31/85420>
14. Lee, C.Y., Chang, T.F., Chang, N.Y., Chang, Y.C.: An automated skin segmentation of Breasts in Dynamic Contrast-Enhanced Magnetic Resonance Imaging. *Sci. Reports* 2018 8:1. **8**, 1–9 (2018). <https://doi.org/10.1038/s41598-018-22941-2>
15. Zhang, Z., Lu, J., Yip, Y.J.: Automatic segmentation for breast skin-line. In: Proceedings - 10th IEEE International Conference on Computer and Information Technology, CIT-2010, 7th IEEE International Conference on Embedded Software and Systems, ICES-2010, ScalCom-2010, pp. 1599–1604 (2010). <https://doi.org/10.1109/CIT.2010.283>
16. The TIES Cancer Research Network (TCRN) Homepage
17. Khairnar, S., Thepade, S.D., Gite, S.: Effect of image binarization thresholds on breast cancer identification in mammography images using OTSU, Niblack, Burns, Thepade's SBTC. *Intelligent Systems with Applications*, pp. 10–11, (2021). <https://doi.org/10.1016/j.iswa.2021.200046>

18. Albiol, A., Corbi, A., Albiol, F.: Automatic intensity windowing of mammographic images based on a perceptual metric. *Med. Phys.* **44**, 1369–1378 (2017). <https://doi.org/10.1002/MP.12144>
19. Iswardani, A., Hidayat, W.: *Mammographic Image Enhancement using Digital Image Processing Technique* (2018)



Towards Fluid Intake Quantification in Older Adults: An Algorithm for Movement Detection Using Accelerometry and Gyroscope Sensors

Paulina Martínez¹ , Nelly Gordillo-Castillo¹  , and David Cortés Sáenz² 

¹ Department of Electrical and Computer Engineering,
University of Ciudad Juarez, Ciudad Juarez, Mexico
nelly.gordillo@uacj.mx

² Department of Design, University of Ciudad Juarez, Ciudad Juarez, Mexico

Abstract. Dehydration in older adults leads to numerous adverse outcomes and is associated with cognitive diseases and medical treatments affecting thirst sensitivity. Technological solutions, including wearable sensors in devices like smartwatches, play a crucial role in monitoring physical activity and health. The integration of wearable technology, data processing, and advanced analysis offers a non-intrusive, accurate method for quantifying fluid intake in older adults by analyzing wrist and arm movements. This paper presents an algorithm for detecting liquid intake movements in older adults using signal samples collected from a wearable sensor equipped with accelerometry and gyroscope technology. The algorithm focuses on preprocessing the signals and extracting relevant features related to fluid intake in two different experimental conditions: standing and sitting. The proposed algorithm accurately identifies fluid intake events, including instances with uncontrolled movements. Various validation metrics were used to evaluate the algorithm's performance, including precision, accuracy, sensitivity, and F1-Score. The recorded metrics consistently showed high percentages ranging from 92% to 100%. These results indicate the algorithm's effectiveness in accurately detecting and classifying fluid intake events in older adults.

Keywords: fluid intake · older adults · accelerometry · gyroscope

1 Introduction

Dehydration in older adults is a significant contributing factor to accidents, falls, delirium, confusion, reduced appetite, depression, dementia, and impaired body mobility, among other conditions. These factors account for a substantial percentage of hospital admissions and readmissions [1–4]. These consequences are directly associated with age-related cognitive diseases and treatments involving medication, which can decrease the patient's sensitivity to thirst and disrupt water balance in the body. Additionally, various pathologies such as diabetes, Alzheimer's disease, renal failure, colon cancer, fragility syndrome, infections, constipation, and other factors can limit daily water intake due to concerns about urinary or fecal incontinence [1].

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024

J. de J. A. Flores Cuautle et al. (Eds.): CNIB 2023, IFMBE Proceedings 96, pp. 222–231, 2024.

https://doi.org/10.1007/978-3-031-46933-6_24

Several technological solutions have been proposed to quantify fluid intake, both in the general population and in specific groups like the elderly. These solutions aim to improve hydration monitoring and promote healthy habits. Some of the technologies that have been used include mobile applications designed to track and record fluid intake [5]. Users can input the quantity and type of beverage they consume, and the app calculates their daily intake, some apps also provide reminders to drink water and offer hydration tips; smart water bottles equipped with ultrasonic, RGB color, temperature, and accelerometer sensors as well as a computational framework for machine-learning-based fluid intake type classification, volume estimation, and bottle-state-recognition [6]; smart cups that use weight sensors to measure how much liquid is consumed [7]. Some models sync with mobile apps to provide real-time information about fluid intake. Connected health tracking systems like smart scales or body composition monitors, can indirectly measure hydration by considering factors such as variations in body weight and the percentage of body water [8].

Advancements in biomedicine and related technologies have led to the development of wearable sensors, which utilize accelerometry and gyroscope techniques to monitor three-dimensional movement patterns of the human body. These wearables consist of reference points placed on different body parts to collect raw data. The accelerometer captures linear acceleration, enabling the detection of changes in body movements. Additionally, the gyroscope measures alignment and detects rotational movements, providing valuable data for analysis [9].

The combination of data collection and analysis techniques is commonly used in human activity recognition (HAR) and applications involving artificial intelligence (AI), machine learning (ML), and deep learning (DL). Through these techniques, algorithms can be trained to classify and identify specific issues by processing the collected data using signal or image processing techniques. This integration of wearables, data processing, and advanced analysis methods offers promising opportunities for various fields of research and practical applications.

Wearable sensors, such as those found in smartwatches and smartphones, play a vital role in monitoring both physical activity and health of individuals. These devices can acquire electrocardiographic signals to provide valuable insights into human well-being [10]. Additionally, they utilize reference points on various body parts to identify and track specific human activities, including walking, running, sitting, standing, and object manipulation, employing the principles of HAR [9]. In addition to the daily activities, the application of sensors combined with HAR and AI has led to the development of smartwatch-like devices capable of identifying various other behaviors. For instance, these devices can accurately determine the timing of food consumption [11]. For fluid intake detection, these sensors can provide real-time data on the frequency and duration of wrist movements and calculate the amount of water consumed when lifting the hand to drink [12]. The integration of wearable technology, data processing, and advanced analysis methods enables the quantification of fluid intake in a non-intrusive and accurate manner.

The objective of this study is to develop an algorithm for detecting fluid intake in older adults using accelerometry and gyroscope techniques. By collecting data from wearable sensors, a comprehensive database will be generated, specifically focusing on older

adults' water consumption. The algorithm will then analyze this data to automatically determine the amount of fluid intake within a specified time frame. This research aims to contribute to a better understanding of fluid intake patterns in older adults and provide valuable insights for healthcare professionals and caregivers.

2 Methods

The WitMotion Bluetooth 2.0 wearable sensor, specifically the BWT901CL model, was carefully chosen for data acquisition in this study. The sensor was calibrated using the MinilMU.exe program, which can be installed either on a PC or a mobile device. The calibration process involved adjusting the accelerometer, magnetic field, gyroscope, and z-axis of the sensor.

During calibration, the configuration algorithm was set to accommodate 9 axes with a range of ± 16 G. The horizontal direction of the sensor was specified, and the settings included a bandwidth (BW) of 256 Hz, a sampling frequency (fs) of 50 Hz, and a baud rate of 9600 bits per second for data transmission. These calibration parameters were carefully selected to ensure accurate and reliable data acquisition from the wearable sensor for further analysis and interpretation in the study.

2.1 Experimental Design

The experimental design and protocol for this project adhered to the guidelines outlined by the Research Ethics Committee (CEI) of the UACJ for approval. The design involved recruiting individuals aged 60 years or older, regardless of their gender, who appeared to be in good health. Participants were provided with a letter of informed consent, which they signed, and were given the opportunity to choose a location for the sample collection.

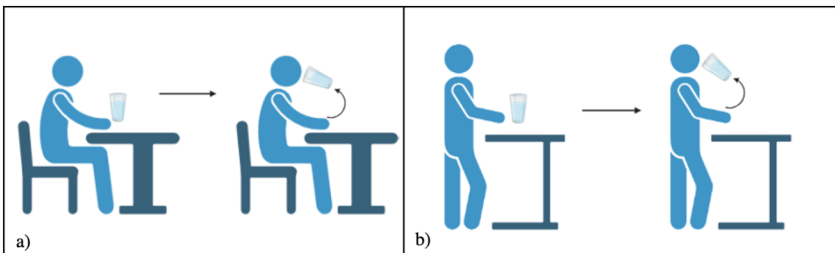


Fig. 1. Activity scheme of experiment A (a) and experiment B (b).

During the signal recording stage, three experiments were conducted with the wearable sensor placed on the participant's wrist in the form of a watch. Each experiment had a total duration of approximately 10 min. Experiment A (Fig. 1a) involved participants seated while drinking liquids, while experiment B (Fig. 2b) involved participants standing while drinking liquids. In experiment C, participants were asked to include hand movement activities other than fluid intake. Signals of similar movements such as

touching the hair or adjusting the glasses were captured to denote the differences between both signals and thus facilitate the characterization of the signal of interest. The timing of fluid intake for each experiment was recorded on a data recording sheet. The protocol was designed to ensure ethical considerations and the collection of relevant data for the study's objectives. In the signal safeguarding stage, personal data from the participants were collected, including age, dominant hand, and sex. To ensure confidentiality and anonymity, these personal identifiers were assigned a numerical code instead of using the participant's actual identification data. This approach was implemented to protect the privacy and identity of the participants throughout the study.

2.2 Database Collection for Training and Test Sets

For the database collection of training and test sets, a total of 120 signal samples were gathered from each of the 10 participants, 70% of the samples were randomly selected for the training set and 30% for the test set. The samples were divided into three types: Type A, which involved participants seated during fluid intake (three samples per participant), Type B, which involved participants standing while consuming fluids (three samples per participant), and Type C activities involving hand movements other than fluid intake (six samples per participant including similar movements such as touching the hair or adjusting the glasses). The signal data was automatically synchronized from the Wit Motion sensor to the MiniIMU.exe program, and each sample was saved in a.txt file for further analysis. Table 1 provides an overview of the participants' general data, including their age and sex, all participants were right-handed.

Table 1. General information about the participants.

ID	Age (years)	Sex
1	63	Male
2	84	Male
3	79	Female
4	75	Male
5	73	Female
6	60	Female
7	61	Male
8	63	Female
9	65	Male
10	60	Female

2.3 Data Analysis and Pre-processing

The original signals obtained from the accelerometer and gyroscope of the motion sensor were analyzed based on the type of liquid intake (A and B) in all three axes (x, y, and z).

Among the axes, the Y-axis of both the accelerometer and gyroscope were considered the most relevant due to distinct patterns observed across the samples in both types of intakes.

In the accelerometer’s Y-axis (Fig. 2a), a consistent pattern was observed initially, indicating the hand resting on the table. This was followed by a gradual increase in signal amplitude as the glass was lifted and brought to the mouth for drinking. The amplitude formed an almost perfect curve that lasted a few seconds, representing the duration of liquid intake. Subsequently, the amplitude gradually decreased as the glass was lowered towards the table until it returned to the initial point of minimum amplitude, indicating the hand and glass back on the table.

Regarding the gyroscope’s Y-axis (Fig. 2b), the pattern displayed a constant minimum amplitude over 0°/s at the beginning, representing the hand on the table. This was followed by a series of peaks that converged into a curve with a slight decrease, followed by an increase in amplitude after a few seconds, indicating rotational movements and alignment during the action of bringing the glass to the mouth. The signal then partially stabilized with slight movements in alignment, oscillating at 0°/s for several seconds during liquid ingestion. Finally, positive and negative peaks of varying amplitudes were observed, culminating in a return to a minimum amplitude of 0°/s, signifying the glass being returned to the table.

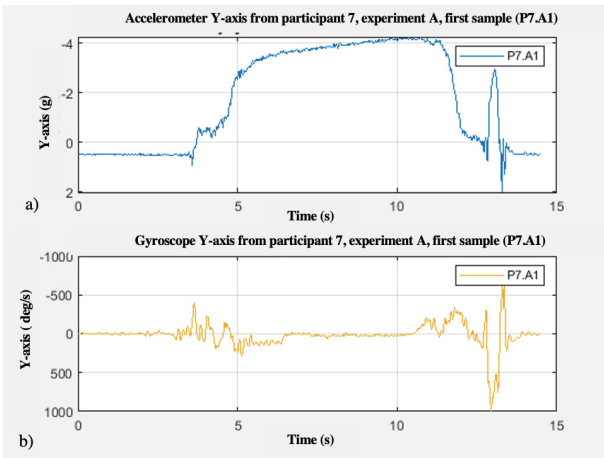


Fig. 2. Y-axis of accelerometer (a) and gyroscope (b).

To process the signals, a low-pass Butterworth filter of order 5 with a cutoff frequency (f_c) of 3 Hz, which represents the maximum frequency of interest, was applied. The signals were sampled at a frequency (f_s) of 50 Hz, complying with the Nyquist theorem and considering the frequency range of 15 Hz to account for arm movements and the targeted signal of interest, i.e., liquid intake. After applying the filter, the first derivative was calculated for the gyroscope’s Y-axis. This was done to attenuate the representative peaks of the signal pattern mentioned earlier and to smoothen the slightly stable line corresponding to liquid intake in this axis.

To determine the start time and end time of the liquid intake, the onset of the action was defined 1 s prior to the detected start time, and the completion of the action was defined 1 s after the detected end time. These time points were crucial for accurately capturing the duration of liquid intake during the experiments conducted in this project.

Algorithm 1. Fluid intake detection
<p>Input: a pre-processed Y-axes signals from the accelerometer and gyroscope. Output: start and end points of fluid intake.</p> <ol style="list-style-type: none"> 1. Define a sliding window of 10 seconds with a step or overlap of 3 seconds along the signal axes. 2. Define initial thresholds: <ol style="list-style-type: none"> a. $t_1=7$, the characteristic peaks of the signal in older adults are equal to or greater than this value. b. t_2 corresponds to the peak with the maximum amplitude in the first derivative of the y-axis of the gyroscope. 3. if $t_2 > t_1$. <ol style="list-style-type: none"> a. Locate the peaks of interest in the signal considering conditions such as the minimum peak height (t_1) and the minimum distance between peaks (3 seconds, reflecting the minimum duration of liquid intake in elderly individuals). b. Store the locations of the peaks that meet the condition in a variable (P_1) based on their timing (peaks with lower heights and a minimum distance of at least 3 seconds between them). 4. t_3 is the 80% of the maximum amplitude of the accelerometer signal, indicating the characteristic curve of the signal on the y-axis. 5. Create a matrix of zeros with a length equal to the number of peaks in P_1 and the length of the accelerometer signal. 6. for $i=1$ to the length of the peak locations <ol style="list-style-type: none"> a. Set a window at 45% within the length of the signal of interest. b. Determine the length of the peaks' positions within the zone of interest, spanning one second before and one second after. c. Calculate the average of the data between the possible peaks of interest based on the length of the previous positions (p). d. if $p > t_3$. <ol style="list-style-type: none"> (1) Calculate the difference between the peaks in P_1 and store it in the variable $DifT$. e. if $DifT > 3 \ \&\& \ DifT < 34$. <ol style="list-style-type: none"> (1) Display the detection of fluid intake on a graph, including the signals from both Y-axes (gyroscope and accelerometer), as well as the start and end points of fluid intake by the participant. 7. end

Fig. 3. Algorithm for fluid intake detection.

2.4 Algorithm Development

The development of the algorithm was conducted in the MATLAB environment. Figure 3 outlines the developed algorithm for detecting fluid intake in signal samples from individuals aged over 60 years, single or repetitive sip gestures were not considered.

This algorithm aims to identify and visualize instances of fluid intake based on the processed signal samples. It utilizes various thresholds, windowing techniques, and signal analysis to detect and represent the occurrence of fluid intake by older adults.

3 Results

The acquired signals representing fluid intake had a minimum duration of 3.86 s, a maximum of 33.92 s, and an average of 9.8 s across the different participants. To compare the signals based on their type (A and B), Fig. 4 displays the graphical representations. It can be observed that the signal of interest for type B, which corresponds to fluid intake while the participant is standing, exhibits a higher amplitude compared to type A. However, the general patterns in both types of signals remain similar.

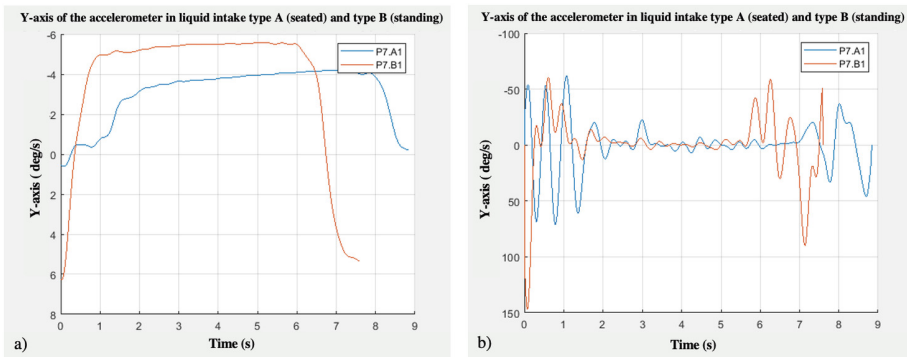


Fig. 4. Comparison of both types of fluid intake on the accelerometer (a) and gyroscope (b) Y-axes.

To visually represent the detected fluid intake in a signal, two dotted lines are utilized to mark the beginning and end of the intake. Figure 5 displays the outcome of intake detection in a healthy 84-year-old individual. In this specific case, the accelerometer signal (shown in blue) exhibits a sustained peak at its maximum point, while the gyroscope peaks (indicated by the red signal) are well-defined. As a result, the algorithm successfully detects the fluid intake event, and the demarcation lines accurately correspond to its duration. This visualization provides clear evidence of the algorithm’s capability to identify and delimit fluid intake periods in the signal.

By implementing a 10-s window with a 3-s step in the algorithm, the detection of multiple fluid intakes within a single signal, even in the presence of uncontrolled movements, was achieved. Figure 6 illustrates an example where two fluid intake events were correctly detected, as indicated by the dotted lines marking the beginning and end of each intake. Importantly, the algorithm successfully distinguished these fluid intake events from other participant movements, ensuring accurate identification specifically for liquid ingestions. This capability underscores the algorithm’s effectiveness in discerning and isolating the targeted fluid intake events while disregarding other unrelated movements performed by the participant.

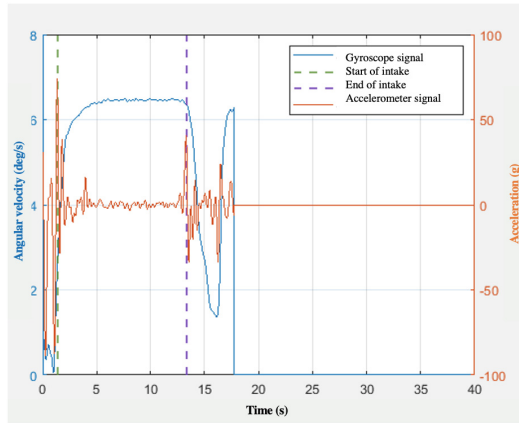


Fig. 5. Example of a fluid intake result in an 84-year-old male participant.

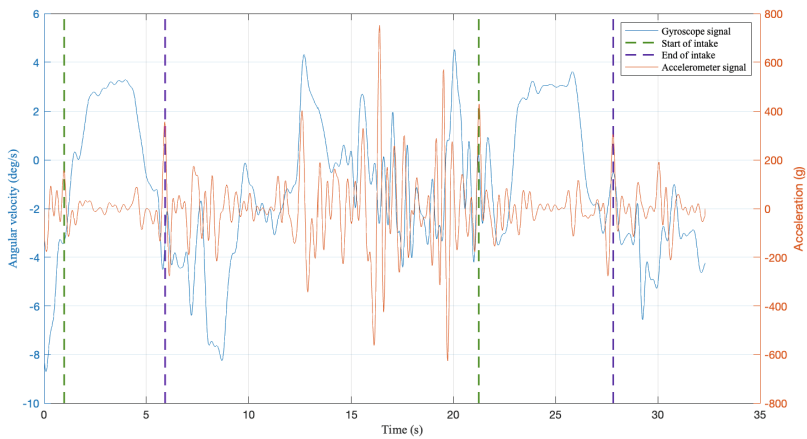


Fig. 6. Example of a signal with fluid intake and uncontrolled movements. Two ingestions delimited by the dotted lines are detected.

		Training set				Test set	
		Predicted classes				Predicted classes	
		Intake	Other			Intake	Other
Actual classes	Intake	78	0	Actual classes	Intake	36	0
	Other	6	84		Other	0	36
		a)				b)	

Fig. 7. Confusion matrix in training set (a) and test set (b) of the liquid intake detection algorithm.

The results of the algorithm validation for all samples in both the training and test datasets are presented in Fig. 7.

The corresponding data is summarized in Table 2. The algorithm demonstrated excellent performance in detecting fluid intake in elderly individuals, as evidenced by the high accuracy rates. In the training set, the algorithm achieved an accuracy of 92.8%, while in the test set, it achieved a perfect accuracy of 100%. Additionally, the algorithm displayed a high sensitivity of 100% in both the training and test sets, indicating its ability to correctly identify all instances of fluid intake. The F1-Score, which combines precision and recall, was 96.29% in the training set and 100% in the test set, further indicating the algorithm's robust performance in accurately detecting fluid intake events. In accordance with another proposition documented in the literature [12], the attained accuracy exhibited a notable level of proficiency, as evidenced by a comparison to the 94% accuracy achieved in [12], which exclusively focused on signal analysis from younger participants.

Table 2. Training and test set evaluation metrics.

Metric	Training	Test
Precision	0.9285	1
Accuracy	0.9642	1
Sensibility	1	1
F1-score	0.9629	1

4 Conclusions

In conclusion, the utilization of human activity recognition technologies and the development of an algorithm in MATLAB have demonstrated promising results in detecting liquid intake events in signals obtained from individuals and over the age of 60. The application of this technology, in conjunction with a thorough analysis for the quantification of the amount of liquids ingested and reminders for water consumption based on a predetermined schedule, holds great potential to benefit individuals in the future.

The validation of the algorithm showcased strong performance, including instances with uncontrolled movements. The validation metrics, which were derived from testing a total of 120 samples (60 liquid intake samples and 60 samples of different activities), consistently exhibited high accuracy rates ranging from 92% to 100%. These results indicate the algorithm's effectiveness in accurately detecting and distinguishing fluid intake events, even amidst challenging conditions.






The successful implementation of this algorithm and its validation highlight its potential for practical use in monitoring and promoting adequate fluid intake in various populations, particularly among older adults. Further development and integration with wearable sensor technologies hold promise for enhancing personalized hydration reminders and improving overall health and well-being.

References

1. Edmonds, C.J., Foglia, E., Booth, P., Fu, C.H.Y., Gardner, M.: Dehydration in older people: a systematic review of the effects of dehydration on health outcomes, healthcare costs and cognitive performance. *Arch. Gerontol. Geriatr.* **95** (2021). <https://doi.org/10.1016/j.archger.2021.104380>
2. Tobar, E., Alvarez, E.: Delirium in elderly hospitalized patients. *Revista Medica Clinica Las Condes* **31**, 28–35 (2020). <https://doi.org/10.1016/j.rmcl.2019.11.008>
3. Hamrick, I., Norton, D., Birstler, J., Chen, G., Cruz, L., Hanrahan, L.: Association between dehydration and falls. *Mayo Clin. Proc. Innov. Qual. Outcomes* **4**, 259–265 (2020). <https://doi.org/10.1016/j.mayocpiqo.2020.01.003>
4. Engelheart, S., Forslund, H.B., Brummer, R.J., Ljungqvist, O.: Dehydration and loss of appetite: key nutrition features in older people receiving home health care. *Nutrition* 91–92 (2021). <https://doi.org/10.1016/j.nut.2021.111385>
5. Rocco, M.V., Rigaud, M., Ertel, C., Russell, G., Zemdegs, J., Vecchio, M.: Fluid intake management in maintenance hemodialysis using a smartphone-based application: a pilot study. *Kidney Med.* 100703 (2023). <https://doi.org/10.1016/j.xkme.2023.100703>
6. Pedram, M., et al.: LIDS: mobile system to monitor type and volume of liquid intake. *IEEE Sens. J.* **21**, 20750–20763 (2021). <https://doi.org/10.1109/JSEN.2021.3081012>
7. Liu, K.-C., Hsieh, C.-Y., Huang, H.-Y., Chiu, L.-T., Hsu, S.J.-P., Chan, C.-T.: Drinking event detection and episode identification using 3D-printed smart cup. *IEEE Sens. J.* **20**, 13743–13751 (2020). <https://doi.org/10.1109/JSEN.2020.3004051>
8. Manimegala, P., Anitha Mary, X., Biji, N., Selshia Shiny, V.: Dehydration measurement using sweat sensor patch and pulse sensor. In: *Journal of Physics: Conference Series*. IOP Publishing Ltd (2021). <https://doi.org/10.1088/1742-6596/1937/1/012013>
9. Ehatisham-Ul-Haq, M., Malik, M.N., Azam, M.A., Naeem, U., Khalid, A., Ghazanfar, M.A.: Identifying users with wearable sensors based on activity patterns. In: *Procedia Computer Science*, pp. 8–15. Elsevier B.V. (2020). <https://doi.org/10.1016/j.procs.2020.10.005>
10. Lin, J., et al.: Wearable sensors and devices for real-time cardiovascular disease monitoring (2021). <https://doi.org/10.1016/j.xcrp.2021.100541>
11. He, S., et al.: A comprehensive review of the use of sensors for food intake detection (2020). <https://doi.org/10.1016/j.sna.2020.112318>
12. Huang, H.Y., Hsieh, C.Y., Liu, K.C., Hsu, S.J.P., Chan, C.T.: Fluid intake monitoring system using a wearable inertial sensor for fluid intake management. *Sensors (Switzerland)* **20**, 1–17 (2020). <https://doi.org/10.3390/s20226682>



Optimizing fMRI Techniques for Post-stroke Motor Rehabilitation: Preliminary Protocol Standardization

Abigail Miranda-Hernández¹ , José Carlos López-Mata² ,
Gabriel Valadez-Roque³ , and Luis Jimenez-Angeles¹  

¹ Departamento de Ingeniería en Sistemas Biomédicos, DIMEI-Facultad de Ingeniería UNAM, Ciudad de México, Mexico

luis.jimenez@ieee.org

² Laboratorio de Neuroimagenología, Departamento de Ingeniería Eléctrica, Universidad Autónoma Metropolitana, Ciudad de México, Mexico

³ División de Rehabilitación Integral, Instituto Nacional de Rehabilitación “Luis Guillermo Ibarra Ibarra”, Ciudad de México, Mexico

Abstract. This paper presents preliminary results of the standardization for the acquisition of fMRI scans for post-stroke motor rehabilitation in the MRI service of the Mexican National Institute of Rehabilitation. fMRI scans of 9 healthy volunteers ($30.22 \pm \text{SD } 3.41$ years) with no prior history of neurological or psychiatric disorders were acquired while they performed a motor paradigm that included tasks of action observation plus motor execution (AOME) and action observation plus motor imagery (AOMI) for fist and finger tapping. Our findings revealed that the primary cortical surfaces (precentral gyrus, postcentral gyrus, supplementary motor area, and premotor cortex) had a functional gradient with higher values for AOME and lower values for AOMI, which is in line with existing literature, indicating that this standardized approach can be beneficial for routinely carrying out post-stroke motor rehabilitation research projects.

Keywords: fMRI · standardization · post-stroke · motor imagery

1 Introduction

Biomedical Engineers (BME) study the physics of magnetic resonance imaging (MRI), the hardware, software and understand the clinical relevance of MRI. In a multidisciplinary field, BME are responsible for the development of new types of imaging [1], improve the image quality [2], reduce imaging artifacts and acquisition times [3], all with the goal of maximizing the clinical utility of MRI techniques. In this work BME students worked alongside psychologists, radiologists, and technicians on a global project aimed at refining a protocol for motor rehabilitation using functional MRI (fMRI).

Numerous studies have highlighted the potential benefits of Motor Imagery (MI, dynamic processes whereby actions are mentally generated and unfold over time without physical movement execution) and Action Observation (AO, dynamic mental processes that represent action content and are triggered when observing other's actions); as strategies for aiding post-stroke motor rehabilitation [4–8]. However, the combination of Action Observation and Motor Imagery (AOMI) produces increased activity in motor-related brain areas compared with AO or MI alone. AOMI involve greater neural activity in the caudal supplementary motor area (SMA), left precentral gyrus, basal ganglia, and cerebellum compared with AO, and the bilateral cerebellum and precuneus compared with MI [9, 10]. These findings indicate that AOMI may be more effective in enhancing motor function during post-stroke rehabilitation. fMRI has proven to be useful for monitoring post-stroke rehabilitation because it has valuable characteristics for both clinical and research applications. However, in MRI facilities with initial experience in hardware and software, there is a continuous need for methodological improvements to ensure the validity and reliability of studies, as errors, variations, or inconsistencies are often observed, which have an impact on obtaining reliable results, and the design of the paradigm is a key factor in optimizing these parameters [11]. Therefore, the contribution of BME expertise and skills are crucial in optimizing the time spent on selection and interviews with patients, reducing the acquisition time of fMRI, programming and validating the paradigm task, and achieving functional results in line with those reported in the literature.

This paper reports on the initial validation of the paradigm task as a means of obtaining functional outcomes comparable to those reported in the literature for healthy volunteers, with the aim of using fMRI studies as a monitoring tool for post-stroke motor rehabilitation in the MRI service of the Mexican National Institute of Rehabilitation.

2 Methods

2.1 Participants

Nine healthy participants (5 women) with an average age of $30.22 \pm \text{SD } 3.41$ years (range: 26–36 years) were selected. The participants had no history of neurological or psychiatric disorders and provided written consent to participate in the study. The participants were enrolled through direct invitation. Interested subjects voluntarily attended a brief interview at the MRI service to perform a series of psychological tests, verify the inclusion and exclusion criteria, sign a letter of informed consent if they were met and participate in a training session. The inclusion criteria were: ages between 25 and 40 years; dexterous (Edinburgh Inventory ≥ 10); between 9 and 18 years of schooling; Kinaesthetic and Visual Imagination Questionnaire R (KVIQ) test ≥ 3 and Perceived Stress Control Questionnaire (PCE) ≤ 13 . The exclusion criteria were: structural brain lesions

(insults, dysplasias or tumors); systemic or chronic disease; claustrophobia; contraindication to MRI procedure (prostheses, pacemakers, valves, etc.). The elimination criteria were: voluntary abandonment; excessive motion that produces artifacts during MRI acquisition and loss or deterioration of subject information.

2.2 Training Session

Subjects who met the selection criteria were scheduled for a training session where they performed the following tasks: 1) opening and closing the fist (fist AOME), 2) joining and separating the tips of the thumb and index finger (finger tapping AOME), 3) imagined opening and closing the fist (fist AOMI), and 4) imagined joining and separating the tips of the thumb and index finger (finger tapping AOMI). Subjects had to synchronize their performance with a short video (30 s) of the expected task with 1 Hz frequency presented at the beginning of each trial. The video also included a 1 Hz beat to facilitate synchronization. The approximate duration of the training session was 12 min, in addition to the time used for explanation.

2.3 fMRI Paradigm

On the day following the training session, MRI images were acquired while participants performed the trained tasks. A block design paradigm was employed (5 blocks for Rest and 5 blocks for Task), each lasting 30 s (10 volumes, TR = 3 s) while the subjects were shown the indications and videos of the tasks to be performed as well as the beginning and end of the activation periods. The basic structure of the paradigm is illustrated in Fig. 1. Each paradigm was presented separately by randomizing the order of exposure, and the tasks were randomized to avoid carryover and interaction effects.

2.4 fMRI Acquisition and Preprocessing

The MRI equipment was synchronized to a laptop computer via an event synchronizer device designed specifically for this research. The objective of this device was to synchronize the electrical pulses from the MRI equipment with the stimulation paradigm controller software (PsychoPy v3.0) to start the sequences of each trial. Anatomical and functional images were acquired using a Philips Magneto Ingenia 3 Tesla with a standard 16-channel head antenna. For functional MRI, a whole-brain echo planar imaging (EPI) sequence with the following parameters was used: Repetition Time (RT) = 3000 ms, Echo Time (ET) = 36 ms, acquisition matrix = 80×95 voxels, Field of View (FOV) = 190×230 mm, Flip Angle (FA) = 90° , slices = 40, slice thickness = 4 mm, inter-slice gap = 0 and 4000 volumes over the experimental scan time of 5 min. Functional image preprocessing and analysis were conducted using the FMRIB Software Library (FSL, FLIRT version 6.0). The Brain Extraction Tool (BET) was applied to each structural image from the command line before preprocessing and for functional

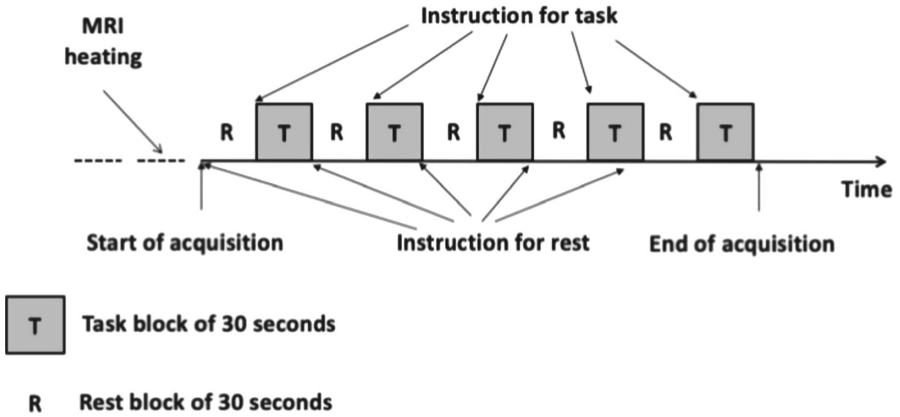


Fig. 1. The basic structure of the paradigm execution during fMRI scanning. The Task and Rest blocks were composed of five blocks, each lasting 30 s (10 volumes, TR = 3 s). The entire paradigm was carried out for each condition (fist AOME, fist AOMI, finger tapping AOME, and finger tapping AOMI).

data with the BET option within the Prestats module of FEAT. The rest of the preprocessing steps (motion correction, slice-timing correction, smoothing, registration, and normalization) were performed using the FEAT GUI with the default options. To correct for motion, each volume from the BOLD images was first rigidly registered to the middle volume using a normalized correlation cost function and linear interpolation (MCFLIRT12 tool). After spatial smoothing with a 5 mm FWHM, registration and normalization were performed using the default linear registration method with full search and 12 degrees of freedom.

2.5 Functional Analysis

The first-level analysis was performed using a general linear model (GLM), and the contrast images were computed for the following conditions: fist AOME, fist AOMI, finger tapping AOME and finger tapping AOMI. On the second level analysis, we used separated models to compute the main effect between conditions by conducting one-sample t-tests with the respective first-level contrast images (Fist AOME > AOMI, Finger tapping AOME > AOMI, Fist AOME > Finger tapping AOME, and Fist AOMI > Finger tapping AOMI) on whole brain level concerning all clusters considering a p-value below 0.05 as significant. As results the cluster sizes higher than 20 voxels, the coordinates of the cluster peaks as well as further relevant local maxima within these clusters are reported.

3 Results and Discussions

3.1 Differences and Similarities Between Fist AOME and AOMI

Figure 2 shows the functional analysis of the Fist AOME, Fist AOMI, and Fist AOME > Fist AOMI in the cluster coordinates $[-44 -18 54]$. This cluster caused a considerable rise in the left-sided BOLD signal of the precentral and postcentral gyrus, as well as the premotor cortex and supplementary motor area (PM+SMA). The voxel volume in AOME was significantly larger than AOMI in the precentral gyrus (3357 vs. 82), postcentral gyrus (3190 vs. 81), and PM+SMA (1739 vs. 17). The contrast to Fist AOME > AOMI shows a significant decrease in voxel volume across all structures while preserving the proportions as seen in AOME. These findings indicate that Fist AOME and AOMI involve the same cortical areas but differ in their volume sizes, and that Fist AOMI can be beneficial for post-stroke motor rehabilitation. Our results are consistent with those reported by Chepurova et al. [12] which verified the activation of primary motor cortex (M1, situated in the precentral gyrus) during both active and passive hand movements. They monitored the brain activity of the motor cortex while the participants performed independent (active) and assisted (passive) fist opening and closing tasks. During active movement, the contralateral M1 showed slightly more activation and there were also some overlapping

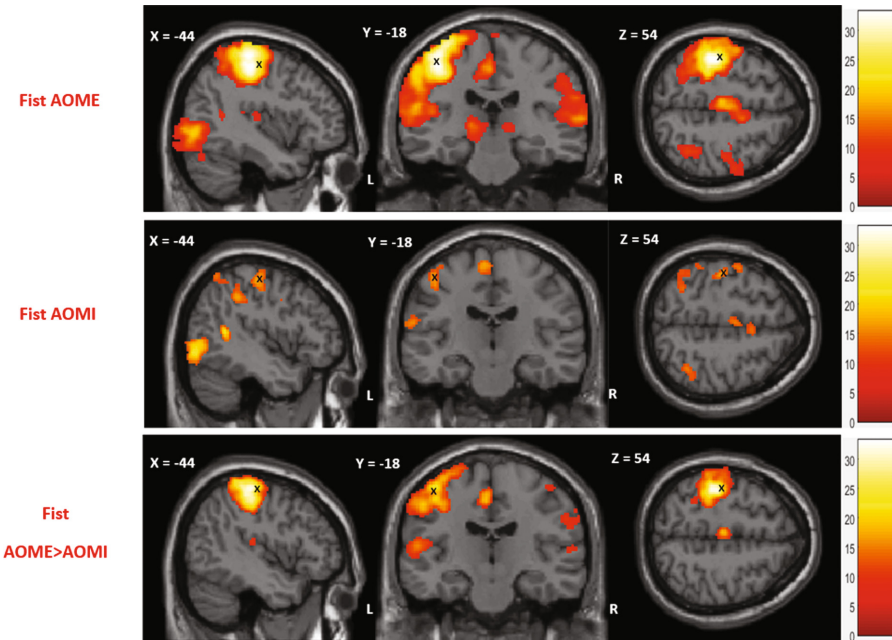


Fig. 2. T-maps showing group brain activation related to the fist condition: AOME, AOMI and AOME>AOMI.

activity in the contralateral primary somatosensory cortex (S1, located in the postcentral gyrus). The authors conclude that the fist opening and closing task is successful in stimulating M1 and S1, even with the help of assisted movement, making it a viable option for patients with paresis.

3.2 Differences and Similarities Between Movement and Imagined of Finger Tapping

Figure 3 shows the functional analysis of the Finger tapping AOME, Finger tapping AOMI, and Finger tapping AOME > Finger tapping AOMI in the cluster coordinates $[-44 -18 54]$. As in Fist analysis, Finger tapping caused a significant increase in the left-sided BOLD signal of the precentral and postcentral gyrus, as well as in PM+SMA. The voxel volume in AOME was significantly bigger than AOMI in the precentral (1171 vs. 201), postcentral gyrus (1574 vs. 50), and PM+SMA (634 vs. 89). The contrast to Finger tapping AOME > AOMI shows a significant decrease in voxel volume of precentral gyrus (337) and PM+SMA (24) but a lesser change in postcentral gyrus (673).

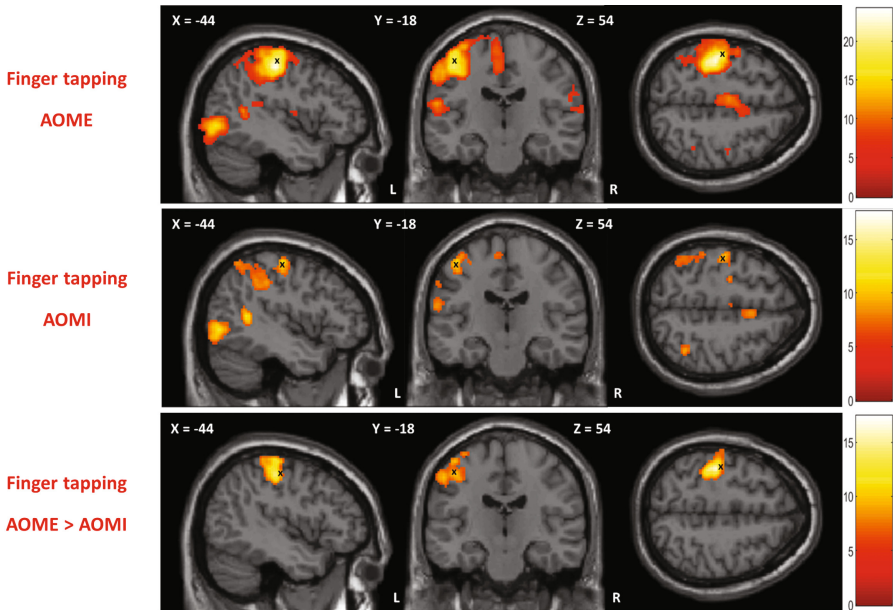


Fig. 3. T-maps showing group brain activation related to the finger tapping condition: AOME, AOMI and AOME>AOMI

These findings are in line with those of Rao et al. [13], who showed that basic free-frequency finger extension and flexion movements activate M1, and that complex finger tapping movements with the same fingers activate the contralateral (sometimes ipsilateral) M1, as well as both SMA and PM bilaterally.

They also showed that when the frequency of the movement is set to 2 Hz for complex movements, there is less intense activation (due to the lower frequency compared to free frequency movements) in the (M1), (SMA), (PM). Lastly, the authors discovered that during a first-person complex movement imagination task, the SMA is more active than the PM.

3.3 Differences and Similarities Between Fist and Finger Tapping in AOME and AOMI

Figure 4 shows the functional analysis of the Fist AOME > Finger tapping AOME and Fist AOMI > Finger tapping AOMI in the cluster coordinates $[-44 -28 54]$. Analyzing the same structures in AOME contrast (first row) the voxel volume for precentral gyrus, postcentral gyrus and PM+SMA were 128, 411 and 9 voxels respectively. While non-significant differences were found for AOMI contrast (second row).

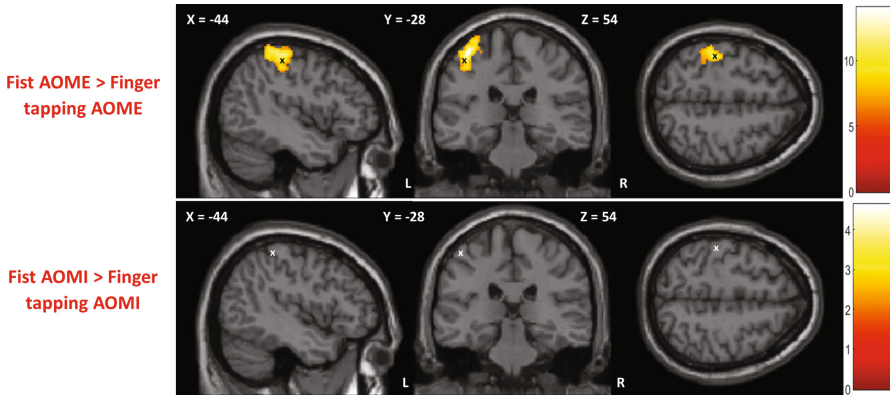


Fig. 4. T-maps showing group brain activation between Fist and Finger Tapping in AOME and AOMI.

Hanakawa et al. [14] reported similar results identifying common cortical brain areas that were activated during the Finger tapping AOMI and Finger tapping AOME tasks. The M1, dorsolateral premotor area in its caudal portion (PMdc) and supplementary motor area in its caudal portion (SMAc) were more active during movement execution than during imagination. Nevertheless, activation was equal in the rostral part of the dorsolateral premotor area (PMdr), ventral premotor area (PMv) and rostral part of the supplementary motor area (SMAr). Their findings suggest a functional gradient from more “executive” to more “imaginative” areas, which further corroborates Rao et al. [13], adding specificity to the neuroanatomy of equivalent activation.

4 Conclusions

The aim of this work was to provide an overview of the standardization process for the acquisition of fMRI studies as a monitoring tool for post-stroke motor rehabilitation in the MRI service of the Mexican National Institute of Rehabilitation. To this point, we have established a standard set of psychological tests and a pre-training protocol that is indicated to the subjects one day before the MRI images are taken. The connection between the MRI machine and the computer that displays the results of the motor paradigm has been established and standardized, making it possible for the service to routinely conduct this type of research. The motor paradigm designed with AOME and AOMI motor tasks of fist and finger tapping showed that, in a control group of healthy volunteers, the primary cortical surfaces (precentral gyrus, postcentral gyrus, SMA, and PM) had a functional gradient from executive to imaginative surfaces (AOME > AOMI), which is in line with the literature. In future work, we will assess the efficiency (time/effect) of the training processes and performance of motor tasks, and a linguistic processing paradigm will be included to evaluate the accuracy (percentage) of motor and linguistic tasks during image acquisition, and analyze the overall brain BOLD signal intensity during motor and linguistic processing tasks.

Acknowledgments. The authors are grateful to Dr. Rubén Conde Espinosa, Dra. Ana Lilia Dotor Llerena and medical students from the Mexican National Institute of Rehabilitation, for their invaluable assistance in granting access to the MRI infrastructure and offering clinical and psychological advice for this project. This project was partially funded by the program UNAM-DGAPA-PAPIME “PE103323: Herramientas de enseñanza para el desarrollo de dispositivos médicos basados en inteligencia artificial”.




References

1. Ruetten, P.P.R., Gillard, J.H., Graves, M.J.: Introduction to quantitative susceptibility mapping and susceptibility weighted imaging. *Brit. J. Radiol.* **92**(1101), 20181016 (2019). <https://doi.org/10.1259/bjr.20181016>
2. Haldar, J.P., Liu, Y., Liao, C., Fan, Q., Setsompop, K.: Fast submillimeter diffusion MRI using gSlider-SMS and SNR-enhancing joint reconstruction. *Magn. Reson. Med.* **84**(2), 762–776 (2020). <https://doi.org/10.1002/mrm.28172>
3. Munoz, C., Fotaki, A., Botnar, R.M., Prieto, C.: Latest advances in image acceleration: all dimensions are fair game. *J. Magn. Reson. Imag.* **57**(2), 387–402 (2023). <https://doi.org/10.1002/jmri.28462>. <https://onlinelibrary.wiley.com/doi/abs/10.1002/jmri.28462>
4. Emerson, J., Binks, J., Scott, M., Kenny, R.P.W., Eaves, D.L.: Combined action observation and motor imagery therapy: a novel method for post-stroke motor rehabilitation. *AIMS Neurosci.* **5**, 236–252 (2018). <https://doi.org/10.3934/neuroscience.2018.4.236>
5. Choi, J.B., Yang, S.W., Ma, S.R.: The effect of action observation combined with motor imagery training on upper extremity function and corticospinal excitability in stroke patients: a randomized controlled trial. *Int. J. Environ. Res. Public Health* **19**, 12048 (2022). <https://doi.org/10.3390/ijerph191912048>

6. Ma, J., Yang, B., Qiu, W., Zhang, J., Yan, L., Wang, W.: Recognizable rehabilitation movements of multiple unilateral upper limb: an fMRI study of motor execution and motor imagery. *J. Neurosci. Methods* **392**, 109861 (2023). <https://doi.org/10.1016/j.jneumeth.2023.109861>
7. Binks, J., Emerson, J., Scott, M., Wilson, C.J., Schaik, P.V., Eaves, D.: Enhancing upper-limb neurorehabilitation in chronic stroke survivors using combined action observation and motor imagery therapy. *Front. Neurol.* **14**, 1097422 (2023). <https://doi.org/10.3389/fneur.2023.1097422>
8. Scott, M., et al.: Combined action observation and motor imagery: an intervention to combat the neural and behavioural deficits associated with developmental coordination disorder. *Neurosci. Biobehav. Rev.* **127**, 638–646 (2021). <https://doi.org/10.1016/j.neubiorev.2021.05.015>
9. Macuga, K.L., Frey, S.H.: Neural representations involved in observed, imagined, and imitated actions are dissociable and hierarchically organized. *NeuroImage* **59**, 2798–2807 (2012). <https://doi.org/10.1016/j.neuroimage.2011.09.083>
10. Taube, W., et al.: Brain activity during observation and motor imagery of different balance tasks: an fMRI study. *Cortex* **64**, 102–114 (2015). <https://doi.org/10.1016/j.cortex.2014.09.022>
11. Engström, M., Ragnehed, M., Lundberg, P., Söderfeldt, B.: Paradigm design of sensory-motor and language tests in clinical fMRI. *Neurophysiologie Clinique-Clin. Neurophysiol.* **34**, 267–277 (2004). <https://doi.org/10.1016/j.neucli.2004.09.006>
12. Chepurova, A., Hramov, A., Kurkin, S.: Motor imagery: how to assess, improve its performance, and apply it for psychosis diagnostics. *Diagn.* **12**(4) (2022). <https://www.mdpi.com/2075-4418/12/4/949>
13. Rao, S.M., et al.: Functional magnetic resonance imaging of complex human movements. *Neurology* **43**, 2311 (1993). <https://doi.org/10.1212/wnl.43.11.2311>
14. Hanakawa, T., Dimyan, M.A., Hallett, M.: Motor planning, imagery, and execution in the distributed motor network: a time-course study with functional MRI. *Cerebral Cortex* **18**, 2775–2788 (2008). <https://doi.org/10.1093/cercor/bhm036>



Recommendations for ICA Denoising of Task-Based Functional MRI Data of Stroke Patients

Martín Emiliano Rodríguez-García¹ (✉) , Raquel Valdés-Cristerna¹ ,
and Jessica Cantillo-Negrete² 

¹ Departamento de Ingeniería Eléctrica, Universidad Autónoma Metropolitana, Iztapalapa, 09340 Ciudad de México, Mexico

emilianorodriguez@xanum.uam.mx

² División de Investigación en Neurociencias Clínica, Instituto Nacional de Rehabilitación “Luis Guillermo Ibarra Ibarra”, 14389 Ciudad de México, Mexico

Abstract. Independent component analysis (ICA) denoising represents a useful tool in functional magnetic resonance imaging (fMRI) preprocessing pipelines. The most used methods for ICA denoising involve automatic artifact component selection for resting-state fMRI studies on healthy populations. However, these automated methods have important limitations that are magnified if they are to be used in clinical populations with commonly limited sample sizes, like stroke, and in task-based fMRI (tb-fMRI), which requires the execution of a given task. Nevertheless, this imaging modality can offer helpful information about the neural activation patterns post-stroke produced by the execution of a motor task, which could be relevant in clinical assessments. However, stroke populations show higher artifact presence in tb-fMRI studies and increased variability in neural activation patterns, which can complicate the identification of signal components. In this work, recommendations to manually identify features of artifact and signal components in ICA denoising of tb-fMRI stroke data are provided. These suggestions could allow a more reliable analysis and interpretation of tb-fMRI studies in stroke populations.

Keywords: BOLD signals · ICA denoising · stroke

1 Introduction

Independent component analysis (ICA) has been shown as a useful denoising tool in functional magnetic resonance imaging (fMRI) preprocessing pipelines [1]. An ICA cleanup can remove different types of artifacts (noise) from the signal of interest, commonly related to neural activity (i.e., blood oxygen level dependent (BOLD) signals), after decomposing the 4D fMRI data into independent components (ICs) [2]. The main step towards an ICA data denoising is to identify which ICs reflect artifacts and which ones reflect BOLD signal. To this end, automated ICA cleanup methods have been proposed [3, 4]. However, their use is focused on resting-state fMRI (rs-fMRI) studies of

healthy populations, limiting their impact in other modalities, such as task-based fMRI (tb-fMRI), and in clinical populations, like stroke, where denoising can have a beneficial impact in data interpretation.

In stroke populations, tb-fMRI has proven to be a helpful tool in the assessment of upper-extremity (UE) motor neural activity [5, 6]. This imaging modality can allow a better understanding of the modified motor control patterns produced post-stroke [5–7]. During tb-fMRI, patients are commonly instructed to perform a motor task following a block design, which allows the analysis of the BOLD signal changes during the task execution [8]. However, tb-fMRI stroke data can contain increased artifact presence due to the patients' motor limitations and discomfort when attempting to move their affected UE [9]. Nevertheless, ICA cleanups in this type of data are scarce [10], which can be attributed, in part, to the automated algorithms' limitations, including the requirement of a training set or focusing on the removal of one type of artifact [3, 4]. However, these limitations can be avoided if manual identification of signal and artifact ICs is performed, as it still remains the gold-standard in ICA denoising [11].

Manually identifying the ICs' information after an ICA decomposition of tb-fMRI stroke data can be difficult. Stroke patients present greater head motion in tb-fMRI, increasing artifact presence in the acquired data [9]. Additionally, the identification of neural activation patterns in stroke patients produced by motor tasks can be challenging due to vascular alterations or neurological changes [7, 12]. Nevertheless, correct identification of the ICs' information is crucial in artifact removal and in preserving the BOLD signal. To this end, guidelines have been proposed to manually identify the ICs' information [11]. However, these suggestions focus on healthy subjects' rs-fMRI data. For ICA denoising purposes, the main difference between the tb-fMRI and rs-fMRI modalities is what is expected to be BOLD signal, whereas artifacts can be of similar origin but with increased presence in the task-based modality, especially for stroke populations. Additionally, accounting for altered BOLD-related activation patterns in stroke patients is required [7]. Thus, particular indications for tb-fMRI data of stroke populations are needed, especially to handle increased artifact presence and to identify the BOLD signal components associated to the performed motor task. This, in turn, could allow a more reliable assessment of the neural activation patterns of stroke populations. In this work, recommendations for manual identification of artifact and BOLD signal ICs during ICA denoising of tb-fMRI data of stroke patients are presented. The ensuing sections include an introduction to ICA, suggestions to identify noise and signal ICs, and, lastly, an ICA cleanup example in tb-fMRI stroke data.

2 Independent Component Analysis in fMRI Data

The spatial probabilistic ICA (PICA) model allows a non-square data mixing process while assuming that additive Gaussian noise ($\eta(t)$) is present, making it useful for fMRI data [1]. This type of data can be described by Eq. 1:

$$\mathbf{x}_i = \mathbf{A}\mathbf{s}_i + \boldsymbol{\mu} + \boldsymbol{\eta}_i \quad (1)$$

where \mathbf{x}_i represents the individual measurements at the i -th voxel, \mathbf{A} is the mixing matrix, \mathbf{s}_i denotes the non-Gaussian sources, and $\boldsymbol{\mu}$ is the mean value of all \mathbf{x}_i [1]. PICA looks to estimate an unmixing matrix $\mathbf{W} = \mathbf{A}^{-1}$ that allows a good approximation $\hat{\mathbf{s}}$ to the real

sources \mathbf{s} following Eq. 2:

$$\hat{\mathbf{s}} = \mathbf{W}\mathbf{x} \quad (2)$$

In spatial PICA, the estimated sources (the ICs) are comprised by a 3D spatial map and its corresponding time series, which can be used to identify noise or artifact features in each component. In turn, this allows the removal of noise ICs from the data, i.e., data denoising, by substituting the corresponding columns of \mathbf{A} with zeros and, subsequently, regressing the components' information out of the data according to Eq. 3:

$$\mathbf{x}' = \mathbf{A}'\hat{\mathbf{s}} \quad (3)$$

where \mathbf{x}' denotes the filtered data and \mathbf{A}' the modified mixing matrix. However, it is necessary to correctly identify the ICs' information to adequately denoise the data.

3 Identifying Artifact and Signal Components in tb-fMRI Data of Stroke Patients

Classifying the ICs as artifact or BOLD signal is the most important step in an ICA cleanup. In this section, suggestions to identify the type of information that a component contains during ICA denoising of tb-fMRI stroke data are presented. Here, the following examples were obtained from a private database of tb-fMRI data of stroke patients acquired in a 3T Philips Ingenia scanner (TR/TE = 3000/35 ms, flip angle = 90°, 120 volumes, 36 slices/volume, voxel size = 2.4 × 2.4 × 4 mm³). During each sequence, patients performed (attempted to perform) continuous finger flexion and extension of their affected hand interleaved with rest periods following a block design. In this work, the image processing and analysis was performed via the FSL (v6.0.4) package [13]. The preprocessing stage included traditional steps like realignment, outlier detection, smoothing (5 mm full width at half maximum), grand-mean intensity normalization, and high-pass filtering (0.011 Hz). Lastly, the ICA decomposition was performed after preprocessing via the PICA algorithm implemented in the FSL's MELODIC tool using default parameters [1].

For tb-fMRI data of stroke patients, it is necessary to account for a high artifact presence. This could result in many components containing noise information, which exhibit specific characteristics depending on the type of artifact it reflects. Several examples of noise ICs' patterns obtained from multiple tb-fMRI sequences of stroke patients are shown in Fig. 1, where the presented axial slices of the spatial distributions highlight each artifact source. To identify a noise component, the following features can be looked out for:

1. The spatial distribution has:
 - a. Small and sparse clusters (Fig. 1a–d).
 - b. Cluster peaks around blood vessels, corticospinal fluid (CSF), or white matter (Fig. 1a–d).
 - c. A ring-like pattern, associated to motion artifacts (Fig. 1e).
 - d. Cluster peaks around air-tissue areas, causing magnetic resonance imaging (MRI) susceptibility distortion (Fig. 1f).

- 2. The time series shows:
 - e. Saw-like morphology, reflecting pulsation artifacts (Fig. 1a, 1d).
 - f. Abrupt changes, associated to motion artifacts (Fig. 1e–f).
- 3. The time series' power spectrum is predominantly of high frequency (Fig. 1a–f) [11].

Additionally, a single component can contain multiple artifact features (Fig. 1a–b) or multiple components can show the same noise feature due to an increased number of artifact sources in tb-fMRI stroke data. Thus, if similar patterns are recognized in a components' spatial distribution, time series and/or power spectrum, it can then be identified as a noise component and, subsequently, regressed out of the data.

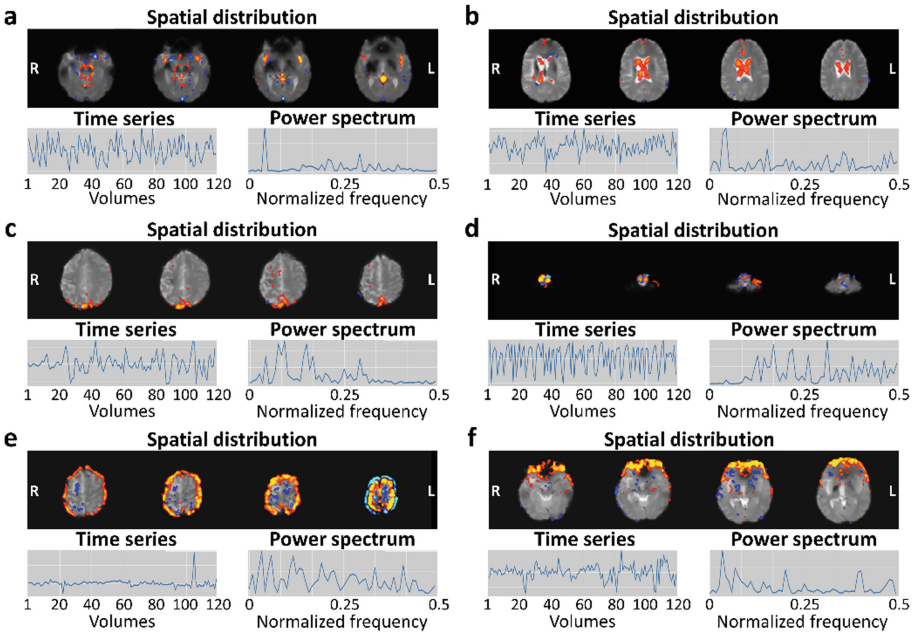


Fig. 1. Selected axial slices of spatial distributions (on top of the corresponding slices of fMRI images), time series, and power spectra of artifact ICs. (a) Artery-, vein-, and CSF-related artifacts component. (b) CSF artifact component. (c) Vein-related artifact component. (d) Pulsation artifact component. (e) Motion artifact component. (f) MRI susceptibility artifact component. R: right, L: left.

Identifying the BOLD-related components in a tb-fMRI ICA decomposition requires the definition of an expected signal according to the selected design. For example, the BOLD signal in a block design paradigm is expected to follow an “off-on” pattern, similar to the one shown in Fig. 2a [8]. Here, the estimated hemodynamic response function (HRF) in each block is shown in red. In this design, patients are instructed to rest (light blocks) and to perform a task (dark blocks) during the “off” and “on” periods, respectively [6, 10]. Both the selected task and design can provide an expected BOLD signal location (for example, around the contralateral primary motor cortex (M1) when executing a UE motor task) and morphology (“off-on” pattern), which can be reflected

in a component's own spatial distribution and time series, respectively. Additionally, BOLD-related components have features that can be used to identify them, including:

1. The spatial distribution has:
 - a. Cluster peaks around brain regions associated with the performed task during the tb-fMRI sequence.
 - b. Large and few clusters.
 - c. Low number of activations around blood vessels, CSF, and white matter.
2. The time series shows regular oscillatory patterns similar to the selected design for the tb-fMRI sequence.
3. The time series' power spectrum is mainly of low frequency [11].

These features can be seen in Fig. 2b, where a component that contains BOLD patterns produced when a stroke patient performs a motor task with its affected hand is presented [12]. Here, the axial slices of the component's spatial distribution around the ipsilesional M1 are shown. Albeit establishing the BOLD signal hypothesis is important to identify the associated ICs, it is not as straightforward in stroke patients due to vascular alterations, lesion site, or neurological changes in activation patterns [7, 12]. This can be observed in some components where either the spatial distribution, time series, or power spectrum reflect, in part, the expected BOLD signal while the other features do not match it completely. In Fig. 2c–f, multiple ICs with this behavior obtained from a single ICA decomposition are shown, focusing on the axial slices that highlight their distribution patterns. Here, the ICs show spatial distributions surrounding the left (Fig. 2c) and right (Fig. 2d–e) M1, and the cerebellum (Fig. 2f), while including additional non-motor related areas (Fig. 2c–f), with changing time series that have, mainly, low frequency power spectra, suggesting that the BOLD signal was partitioned into multiple ICs. Additionally, the BOLD signal in these components is mixed with artifact information, possibly due to increased variability in the spatial distribution of the signal of interest and higher artifact presence in stroke patients' tb-fMRI studies, making the ICA decomposition a more complex task [7, 9, 12]. Therefore, the following suggestions can also be established:

1. Label the component as BOLD-related if its spatial distribution is mainly located around gray matter.
2. Do not regress the component out of the data if it is a noise and signal mixture.

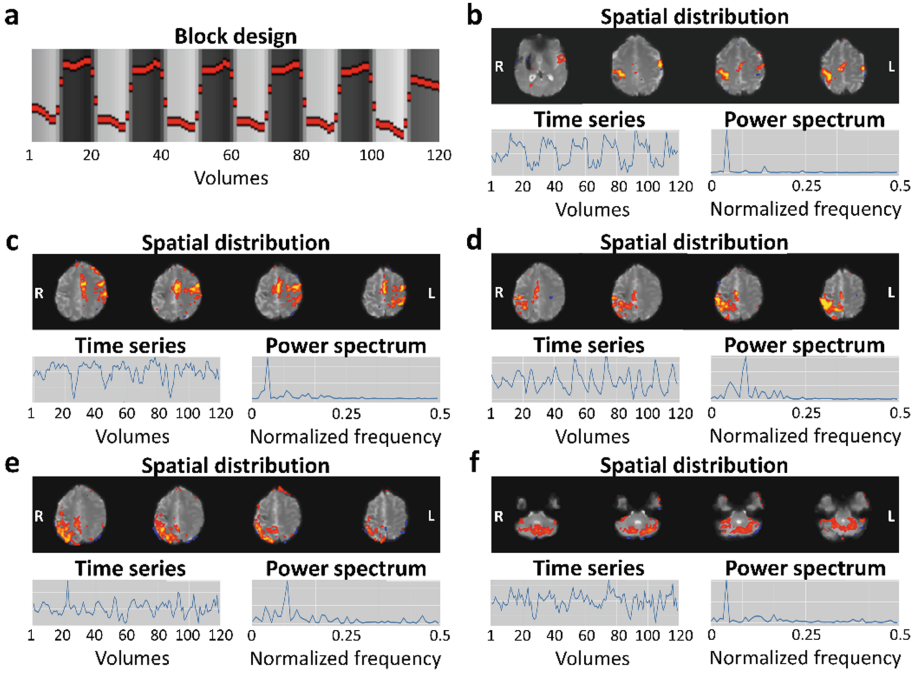


Fig. 2. Block design of a tb-fMRI study involving finger flexion and extension with the affected (left) hand of a stroke patient and ICs containing BOLD signal information on top of selected axial slices of fMRI images. (a) Block design with 6 rest conditions (light) alternating with 6 task conditions (dark) with the expected HRF superimposed (red line). (b) Component that mainly reflects the spatial distribution, time series, and power spectrum of the expected BOLD signal. (c) Component with partitioned BOLD signal (left M1) and artifact information. (d) Component with partitioned BOLD signal (right M1) and artifact information. (e) Component with partitioned BOLD signal (right M1) and artifact information. (f) Component with partitioned BOLD signal (cerebellum) and artifact information.

4 ICA Denoising Example in tb-fMRI Data of a Stroke Patient

Significant activation maps showing the ICA denoising effect in tb-fMRI stroke data are presented in Fig. 3. These maps were obtained after the processing and analysis of a tb-fMRI sequence of a stroke patient before (Fig. 3a) and after (Fig. 3b) including ICA denoising as the last step of the previously described preprocessing stage. Statistical analysis was carried out via a generalized linear model (GLM) including regressors of the expected HRF and its temporal derivative, of 24 realignment parameters [13], and of the detected outliers. Additionally, the regressors were high-pass filtered (0.011 Hz). The resulting Z-maps show the statistically significant activations (cluster threshold of $Z > 3.1$, corrected cluster significance threshold of $p = 0.05$) produced during the tb-fMRI sequence. In Fig. 3a, significant activations can be found around the ipsilesional (right) M1 (marked in red dashed circles), which can be associated to the expected BOLD response given the left-hand motor task [12]. However, additional significant activations,

possibly associated to vein, motion, and CSF artifacts, can also be observed. In Fig. 3b, after ICA denoising, significant activations are preserved around the ipsilesional M1. Moreover, the activations deemed as artifact related are no longer present. This suggests that an ICA cleanup can further denoise tb-fMRI data even after traditional preprocessing pipelines and the inclusion of several motion-related regressors in the GLM analysis. Reliability in the obtained significant activations is of particular importance in quantification methods of tb-fMRI data, such as the often used laterality index [14], which are dependent on the number and location of these activations.

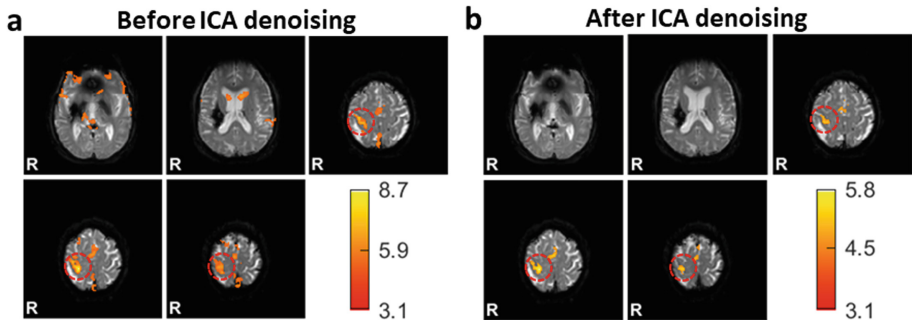


Fig. 3. Significant activation maps of a stroke patient on top of selected axial slices of fMRI images when performing continuous finger flexion and extension with its affected hand (left hand) during a tb-fMRI sequence. Ipsilesional M1 is marked in red dashed circles. (a) Activation maps obtained before ICA denoising. (b) Activation maps obtained after ICA denoising. R: right hemisphere.

5 Conclusion

ICA denoising represents an important tool in tb-fMRI preprocessing that can lead to more reliable results. The recommendations to recognize noise and signal components presented in this work aim to provide a guideline to correctly identify by hand each component obtained after an ICA decomposition of tb-fMRI stroke data. The included example illustrates that ICA denoising can be helpful in removing persistent activations associated with artifacts while preserving BOLD information. In turn, an adequate tb-fMRI artifact cleanup could allow for a more reliable interpretation and quantification analyses of neural activation patterns in stroke patients.

Acknowledgments. MER-G gratefully acknowledges the Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) for Ph.D. scholarship 844633. The authors thank CONAHCYT for grant number SALUD-2018-02-B-S-45803.






References

1. Beckmann, C.F., Smith, S.M.: Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans. Med. Imaging* **23**, 137–152 (2004)

2. Murphy, K., Birn, R.M., Bandettini, P.A.: Resting-state fMRI confounds and cleanup. *Neuroimage* **80**, 349–359 (2013)
3. Salimi-Khorshidi, G., Douaud, G., Beckmann, C.F., Glasser, M.F., Griffanti, L., Smith, S.M.: Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage* **90**, 449–468 (2014)
4. Pruim, R.H.R., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J.K., Beckmann, C.F.: ICA-AROMA: a robust ICA-based strategy for removing motion artifacts from fMRI data. *Neuroimage* **112**, 267–277 (2015)
5. Demers, M., Varghese, R., Winstein, C.: Retrospective analysis of task-specific effects on brain activity after stroke: a pilot study. *Front. Hum. Neurosci.* **16**, 871239 (2022)
6. Yuan, K., Wang, X., Chen, C., Lau, C.C.Y., Chu, W.C.W., Tong, R.K.Y.: Interhemispheric functional reorganization and its structural base after BCI-guided upper-limb training in chronic stroke. *IEEE Trans. Neural Syst. Rehabil. Eng.* **28**, 2525–2536 (2020)
7. Dodd, K.C., Nair, V.A., Prabhakaran, V.: Role of the contralesional vs. ipsilesional hemisphere in stroke recovery. *Front. Hum. Neurosci.* **11**, 469 (2017)
8. Soares, J.M., et al.: A Hitchhiker’s guide to functional magnetic resonance imaging. *Front. Neurosci.* **10**, 515 (2016)
9. Seto, E., et al.: Quantifying head motion associated with motor tasks used in fMRI. *Neuroimage* **14**, 284–297 (2001)
10. Krishnamurthy, V., et al.: A method to mitigate spatio-temporally varying task-correlated motion artifacts from overt-speech fMRI paradigms in aphasia. *Hum. Brain Mapp.* **42**, 1116–1129 (2021)
11. Griffanti, L., et al.: Hand classification of fMRI ICA noise components. *Neuroimage* **154**, 188–205 (2017)
12. Wang, L.E., Fink, G.R., Diekhoff, S., Rehme, A.K., Eickhoff, S.B., Grefkes, C.: Noradrenergic enhancement improves motor network connectivity in stroke patients. *Ann. Neurol.* **69**, 375–388 (2011)
13. Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M.: FSL. *Neuroimage* **62**, 782–790 (2012)
14. Seghier, M.L.: Laterality index in functional MRI: methodological issues. *Magn. Reson. Imaging* **26**, 594–601 (2008)



EOG Signal Classification Based on Blink-to-Speak Language

María Cristina Padilla-Becerra^(✉) , Diana Karen Macias-Castro ,
Ricardo Antonio Salido-Ruiz , Sulema Torres-Ramos ,
and Israel Román-Godínez 

Division of Cyber-Human Integration Technologies, University of Guadalajara,
44430 Guadalajara, Jalisco, Mexico
maria.padilla5065@alumnos.udg.mx

Abstract. Given the need for a universal communication method for people who cannot communicate using oral language (as speech disability, limited or no movement in the upper limbs avoiding the use of sign language, among others), the electrooculogram (EOG) is proposed as a tool for recording eye patterns. Therefore, a total of 10 patterns derived from the Blink-to-speak language, which contains essential needs, were carefully chosen to be recorded by 20 participants. From these recordings, various statistical and signal measurements were extracted as distinctive features. Subsequently, a comprehensive database was constructed using these features in order to train a machine-learning algorithm, specifically a decision tree model. The primary objective of this model was to accurately detect and predict the specific eye-movement patterns that corresponded to a given word or communicative concept, all while considering the underlying intention behind each movement. An 83.4% of accuracy were achieved for this multiclass task, the patterns that achieved 100% correct predictions were “I want to sleep” and “danger”, while the patterns with lower performance were “yes” and “toilet”. Hence, our proposal represent a viable communication alternative for those who cannot communicate due to speech and limb movement limitations.

Keywords: Eye movements · EOG · communication · Blink to Speak · machine learning

1 Introduction

Communication is defined as the process of transmitting information between a sender and a receiver [1], this process is inherent to human beings, as it is a basic mechanism for socialization.

According to the census conducted in 2020 by INEGI, in Mexico there are about 945,000 individuals with speech or communication disabilities [2]. People with speech or communication disabilities often rely on Sign Language as a communication method. However, this method presents limitations; for example, both the sender and receiver need to be familiar with hand movements and facial expressions. Furthermore, sign language

varies depending on the geographical region or language, resulting in approximately 300 different sign languages [3]. Additionally, individuals with speech disabilities and limited upper limb mobility face even more significant communication challenges since sign language is not suitable for their situation.

In 2018, as part of the TBWA/India campaign, a universal language called “Blink to Speak” (BTS) [4] was introduced to tackle this problem. It utilizes eight key alphabets (shut, blink, left, right, up, down, wink, and roll).

While the BTS language provides a solution by eliminating the need for limb movement and being universally applicable, its use implies a high level of concentration from both the sender and the receiver. Any distractions must be avoided, and the process can be complex and time-consuming, as it requires the attention of both parties involved. Consequently, current efforts are focused on the creation of a method that allows the recognition of eye movement (EM) patterns, in order to help people with speech disabilities and motor limitations to communicate with their environment.

In order to facilitate the interaction of individuals with speech disabilities and motor impairments, several systems have been proposed that utilize EMs, particularly eye blinking, as they are the least affected muscle movements [5]. Additionally, the EOG has emerged as a practical and user-friendly method for acquiring EM data, making it convenient and straightforward to implement.

For instance, the use of EOG is proposed in [6] for the development of a computer-assisted communication support system that allows the user to make and receive voice calls, send SMS and emails. While this system is capable of identifying simple EM patterns, its intended use is for individuals with motor disabilities who still retain their speech abilities.

Besides, the use of EMs to control virtual keyboards has also been proposed. In [7] an EOG-based writing system is described, employing eight distinct patterns of EMs, and in [8] is presented a system with 10 patterns representing basic needs achieving an accuracy of 80%, however, calibration is required for each user. Although these systems present an alternative means of communication, as virtual keyboards, they necessitate a monitor and their usage is intricate, since selecting each character involves more than one EM.

On the other hand, machine learning (ML) has been used in the patterns eye recognition for communication. In [9] an EOG-based eye-writing system is proposed as a communication tool that recognizes 10 Arabic numerals. The data of each participant underwent individual testing. A comparison between different models is performed, reporting a mean accuracy of 96.79%, but excluding 2 number patterns that were not classified correctly.

Also, computer vision has been employed to recognize EM patterns with communicative intent. An example of this is [10], where the Blink-to-live system is proposed, which is a language of 4 EMs (left, right, up and down), based on “BTS” which uses a camera through a mobile application and a python module for the processing of the images obtained. An important disadvantage of using images is that variations in ambient brightness levels must be considered [9].

Although the mentioned works aim to provide a communication tool for person, these involve multiple EMs to represent a single command and may require the use of additional devices such as monitors located in front of the person. Also, the use of ML has demonstrated its usefulness in detecting ocular patterns for communicative purposes. However, this approach is constrained by the fact that the models were exclusively trained to each individual, and the reported averages of the best accuracies exclude patterns with lower accuracy.

In this work, a ML model is proposed for the automatic detection of eye patterns that are registered using EOG, as a non-invasive and comfortable method for capturing eye patterns based in BTS language.

2 Methodology

In order to identify eye patterns using a ML algorithm, the methodology described in Fig. 1 were followed.

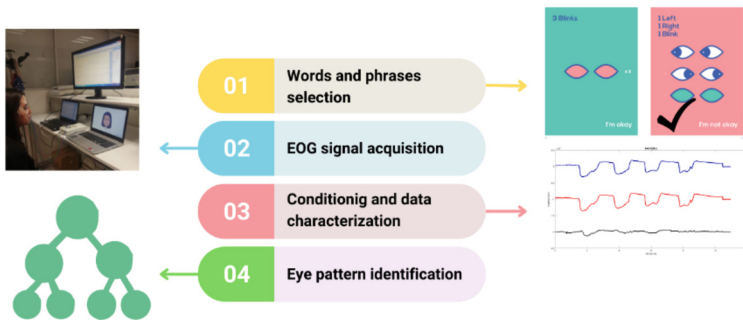


Fig. 1. Diagram of methodology

2.1 Words and Phrases Selection

A total of 10 phrases and words are chosen from the 50 patterns included in the BTS language. The selection process entails evaluating both the communicative intent and the complexity of the movement associated with each pattern. Specifically, the 10 patterns with the highest weights are chosen for further analysis (see Fig. 2).

It is important to mention that among the 50 language patterns, 18 require feedback from the receiver in order to effectively communicate the needs expressed by the sender. Out of the selected patterns due to their communicative intention, three require feedback: “I feel like eating” requires the receiver to indicate or mention food options, “I’m in pain” requires them to point out body parts, and “danger” requires them to indicate potential risky situations, so that the sender can respond with a single blink for “yes” or double blink for “no”.

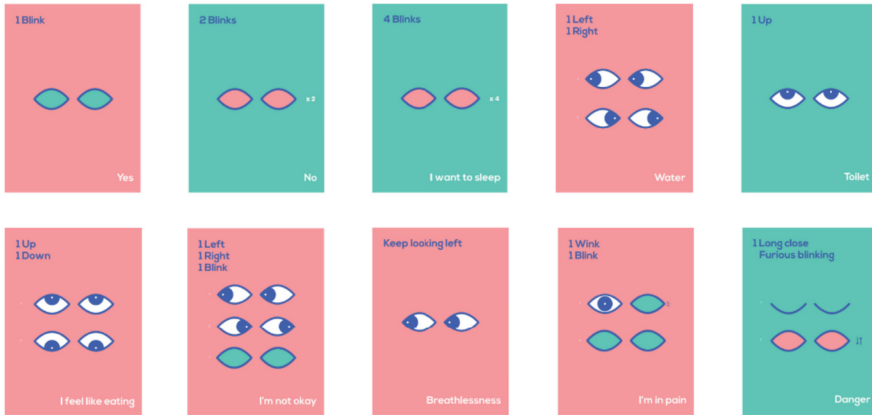


Fig. 2. Phrases and words selected. Images were taken from [4]. In the upper left part the ocular pattern is described, in the center its graphic representation and in the lower right part the phrase or word that represents.

2.2 Data Acquisition

EOG recordings are obtained from a cohort of 20 participants, each replicating the aforementioned set of 10 movement patterns associated with the selected words and phrases. All subjects selected were free from any impediments that could hinder the replication of visual motion patterns, comprising an equal distribution of 10 women and 10 men aged between 18 and 50 years.

The EOG data was captured using the Natus XLTEK EEG 32 system. The recording setup incorporates the following configuration parameters: a sample rate of 256 Hz, a time base of 60 mm/s, 3 channels for EOG measurements, a sensitivity of 100 $\mu\text{V}/\text{cm}$, and a bandpass filter ranging from 0.5 to 35 Hz [11–13].

For the recording of the selected EM patterns, the registration of three channels is required: two vertical corresponding to the right eye and left eye, and a horizontal one, placing the electrodes as shown in Fig. 3. Gold cup surface electrodes with conductive paste are used.

In order to record each pattern, a training video is presented, demonstrating the movements to be performed at a consistent rhythm of one count per second. This process is repeated until the participant can replicate the movements without visual guidance, relying solely on the auditory count.

The videos begin with a four-second period of rest time, accompanied by the sound of a bell indicating the start of the recording. Subsequently, the timer begins, guiding the participant in performing the EMs.

All patterns begin and end with the gaze fixed on a reference point in front of the subject. The participant is instructed to refrain from making any EMs outside of the trained movement pattern.

Each participant individually performs the 10 movement patterns selected, resulting in a total of 200 records for the dataset at the end of the data acquisition process.

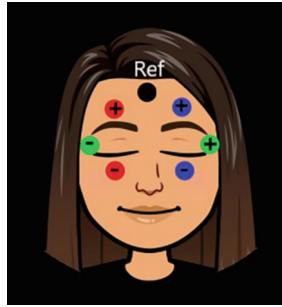


Fig. 3. Distribution of the seven electrodes to record the EMs of the three channels involved: right (red electrodes), left (blue electrodes), and lateral (green electrodes), as well as the reference (black electrode).

2.3 Conditioning and Data Characterization

After obtaining the recorded motion patterns, a zero-phase band-pass filter ranging from 0.5 to 30 Hz is applied to them. The filtering process utilizes a Kaiser window with a 20 dB attenuation [14].

It works with each channel separately, and with the addition and multiplication of these channels, thus having five signals. For each of these, the average, variance, standard deviation, RMS value [15], number of peaks (defining a threshold of the mean plus 1.75 times the standard deviation), and distance between peaks are calculated, as well as the difference between the highest and lowest voltage of each channel [16] and the number of samples of the recording.

In order to reduce the number of features and to minimize the computational cost of the ML algorithm, a correlation matrix was calculated to assess the linear relationship between features and eliminate those that present a high correlation [17]. Initially, a correlation threshold of 0.7 was proposed for feature selection according to [18]. However, testing the model with a slightly constrained threshold of 0.8 resulted in improved performance. This approach aimed to balance between feature reduction and maintaining predictive accuracy. Thus a total of 19 features were selected that had a correlation of less than 0.8.

2.4 Generation and Evaluation of the Classification Model

To train the algorithm, the “Classification Learner” Matlab toolbox [19] was utilized, specifically employing the decision tree (DT) model. The DT is a non-parametric model that employs non-recursive binary partitioning to predict the class to which a case belongs based on its attributes [20]. It was decided to use this model because it’s descriptive, easy to implement, and its computational cost is lower than other algorithms.

The training dataset for the DT model comprises the 19 aforementioned features, extracted from each of the 200 records associated with the 10 selected EM patterns.

The performance evaluation during training was assessed using a 5-fold cross-validation approach to find a model that suffers neither from excessively high bias nor from very high variance [21]. Also, confusion matrices were used to assess performance. The model was trained 10 times, always using the same conditions.

3 Results and Discussion

One of the products derived of this study is a database comprising the characterization of 10 distinct patterns of EMs executed by 20 individuals. This database allows the training of ML algorithms, enabling further advancements in the field.

The ML model generated shows very promising results, with an average of 83.4% correct predictions (see Fig. 4) after having trained 10 times. The patterns that achieved 100% correct predictions were “I want to sleep” and “danger”, while the patterns “I’m not okay” and “I’m in pain” had a regular performance (81% and 82.5%). Unlike the other patterns, “yes” (blink) and “Toilet” (up) require minimal muscle activation (i.e., the signals have more frequently lower SNR compared to others acquired for different words), the model misclassify them, making them the worst performing patterns, with 53% and 65%, respectively.

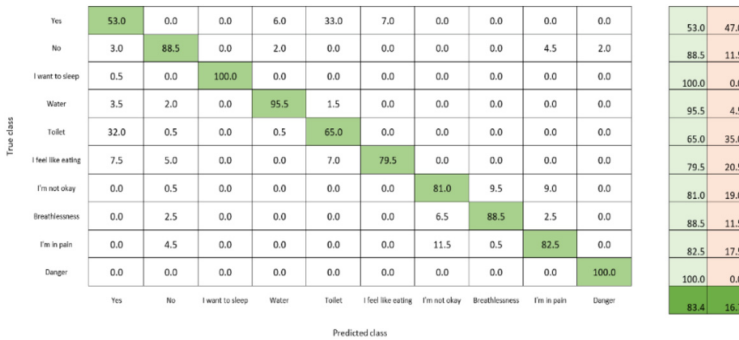


Fig. 4. Confusion matrix of the generated model.

Regarding the features considered by the DT algorithm to classify the patterns (see Fig. 5), the most relevant one is the length, which corresponds to the number of samples in the record.

Other important features include variance, the number of peaks, the mean, and the distance between peaks, especially of the horizontal channel. It is worth noting that all the considered features correspond to the characterization of separate channels, although the database also contained characterization of the sum and multiplication of the three channels. It is important to note that we expected that the sum and multiplication of these signals could provide more differentiable patterns that could help in the classification of these signals; however, the algorithm did not consider these signals in order to classify the words.

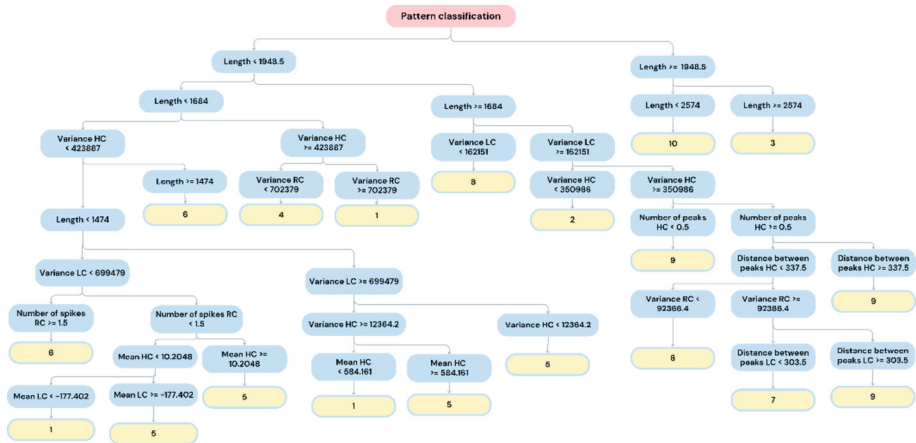


Fig. 5. Decision tree. HC is the horizontal channel, RC is the right channel and LC is the left channel. Numbers meaning: 1) Yes, 2) No, 3) I want to sleep, 4) Water 5) Toilet 6) I feel like eating 7) I'm not OK 8) Breathlessness 9) I'm in pain 10) Danger.

Considering the works mentioned in the state of the art, the present study provides an alternative form of communication based on EOG, similar to [6–9], with the advantage that it can be utilized by individuals with both motor and speech disabilities; it is built upon a universal language with simple patterns, eliminating the need for a monitor in front of the person. Furthermore, by not relying on images [10], it can be employed in any environment regardless of lighting conditions.

The generated model can be used by different individuals since the dataset used for training includes records from various people, unlike the datasets from [8] and [9], which are models trained for each user. Although slightly lower than the average accuracy of [9], no pattern is discriminated against even if it exhibits a lower accuracy. The DT was chosen because, in comparison with the algorithms used in [9], it allows for the observation of the most important attributes in pattern classification.

It should be noted that a direct comparison with the state of the art is not feasible, as each study adopts a distinct approach, which are integrated into this research.

4 Conclusions and Future Work

Based on the analysis of the current state-of-the-art and the results obtained, the proposed method for the acquisition and characterization of records represents a viable alternative. Unlike image processing-based methods [10], EOG does not require additional equipment such as a camera and does not require environmental control such as lighting or distance management. In addition, it offers comfort to the user since the electrodes are superficial and do not cause discomfort.

Thus, the 10 selected EM patterns of the BTS language were successfully recorded with EOG and efficiently classified using a ML algorithm with 83.4% correct predictions.

As part of our future work, we are focusing on enhancing the results and user accessibility. Specifically, we are currently designing a real-time EM acquisition device that aims to serve as a comprehensive interpretation or communication system. Our goal is to ensure the stability of the recorded signals and improve the sensitivity of the device to accurately detect the movement patterns associated with the language. Additionally, we plan to explore alternative algorithms that leverage different signal characteristics, allowing us to expand the classification of words in the language.

References

1. Gómez, F.S.J.: La Comunicación. *Salus* **20**(3), 5–6 (2016)
2. INEGI, Estadísticas a propósito del día internacional de las personas con discapacidad (datos nacionales) (2021)
3. Villa, A.R.: El lenguaje de señas en un mundo globalizado, *Blog Digit. Univ.* **3**(1), 1–12 (2021)
4. Blink to speak. <https://www.blinktospeak.com/>. Accessed 15 May 2023
5. Palma Bustamante, D.A.: Extracción de artefactos oculares de EEG mediante análisis de componentes independientes (2020)
6. Nayak, S.K., Pal, K., Thirugnanam, A.: Development of an EOG based computer aided communication support system. In: 2015 Annual IEEE India Conference (INDICON), pp. 1–6. IEEE (2015)
7. Nathan, D.S., Vinod, A.P., Thomas, K.P.: Electrooculogram based assistive communication system with improved speed and accuracy using multi-directional eye movements. In: 2012 35th International Conference on Telecommunications and Signal Processing (TSP), pp. 554–558. IEEE (2012)
8. Lingegowda, D.R., Amrutesh, K., Ramanujam, S.: Electrooculography based assistive technology for ALS patients. In: 2017 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia), pp. 36–40. IEEE (2017)
9. Du Chang, W., Cha, H.S., Kim, D.Y., Kim, S.H., Im, C.H.: Development of an electrooculogram-based eye-computer interface for communication of individuals with amyotrophic lateral sclerosis. *J. Neuroeng. Rehabil.* **14**(1), 1–13 (2017)
10. Ezzat, M., Maged, M., Gamal, Y., Adel, M., Alrahmawy, M., El-Metwally, S.: Blink-To-Live eye-based communication system for users with speech impairments. *Sci. Rep.* **13**(1), 1–15 (2023)
11. Mahecha, D., Lozano, F., García E.: Control de mouse através de señales EOG y algoritmos de boosting. In: *Memorias del XIII simposio de Tratamiento de Senales, Imágenes y Visión Artificial (STSIVA 2008)*, vol. 2, pp. 57–60 (2008)
12. Barrientos Castaño, A., Herrera Mora, M., García Báez, O., Sierra Hernández, M., Pérez de la Rosa, O.: Valores de referencia del electrooculograma, *Revista Cubana de Investigaciones Biomédicas* **31**(1), 46–52 (2012)
13. López Martínez: Sistema asistencial basado en computador empleando señales electrooculográficas orientado a personas discapacitadas (2012)
14. Oleas Villalba, F.G., Quintana Hermosa, W.H.: Diseño e implementación de filtros FIR con parámetros variables en tiempo real utilizando la ventana de Kaiser (2002)
15. Semmlow, J.: *Circuits, Signals, and Systems for Bioengineers: A Matlab-Based Introduction*. Elsevier Academic Press, Cambridge (2005)
16. Suárez, C.: Análisis del patrón respiratorio y la interacción cardiorrespiratoria en pacientes mayores (2018)

17. Correlation coefficients - MATLAB corrcoef - MathWorks América Latina. <https://la.mathworks.com/help/matlab/ref/corrcoef.html>. Accessed 08 Apr 2023
18. Dancy, P., Reidy, J.: Statistics Without Maths for Psychology, 8th edn. Pearson, London (2018)
19. Entrenar modelos para clasificar datos usando machine learning supervisado - MATLAB - MathWorks América Latina. https://la.mathworks.com/help/stats/classificationlearner-app_es.html. Accessed 18 May 2023
20. Modelos de aprendizaje automático mediante árboles de decisión. <https://www.econstor.eu/handle/10419/238403>. Accessed 19 May 2023
21. James, G., Witten, D., Hastie, T., Tibshirani, R.: An Introduction to Statistical Learning, vol. 112, p. 18. Springer, New York (2013)



Semiautomatic Generation of a Three-Dimensional Human Anatomical Model of Bone for Biomedical Applications: First Approach

Nohemi Barranco Alavez¹, Christian Isaac Hurtado-Esquivel¹,
Citlalli Jessica Trujillo-Romero², and Álvaro Anzueto Ríos¹

¹ Laboratorio de Biomecánica, Departamento de Ingeniería Biónica, Unidad Profesional Interdisciplinaria en Ingeniería y Tecnologías Avanzadas-IPN, Avenida Instituto Politécnico Nacional No. 2580, Barrio la Laguna Ticomán, 07340 Ciudad de México, Mexico

² División de Investigación en Ingeniería Médica, Instituto Nacional de Rehabilitación-LGII, Calz. México-Xochimilco 289, Coapa, Guadalupe Tlalpan, Tlalpan, 14389 Ciudad de México, Mexico

cjtrujillo@inr.gob.mx

Abstract. Third-dimensional human anatomical models have been used to better understand the human body. However, one of their most important uses is the personalization of any type of treatment. 3D human models of bones from the lower extremities are crucial for visualization, treatment planning, and orthotic and prosthetic design. Nonetheless, this 3D models must be compatibles with any modeling software, like those working with the finite element method. Therefore, a methodology to generate 3D anatomical models compatible with CAD systems was proposed. A couple of sets of Computed Tomography (CT) images in DICOM (Digital Imaging and Communications in Medicine) format was used. To reduce the computational cost, different optimization algorithms were implemented. A Gaussian low-pass filter was implemented for bone segmentation, and the thresholding method was applied. Morphological operations to eliminate sparse pixels, gaps and to generate smooth edges were performed. The *Marching Cubes* algorithm was used to generate a three-dimensional mesh. The 3D model scale was adjusted by the metadata in the DICOM files. The 3D model surfaces were smoothed by a Gaussian filter applied to the segmented images and an averaging filter used on the model's vertices. To verify the dimensions of the reconstructed model, they were compared with those obtained from a model reconstructed in InVesalius 3.1. Approximately, it took 1.95 min to generate the exported 3D model in a CAD-compatible STL file.

Keywords: Bone segmentation · CT · 3D models · CAD · 3D anatomical reconstruction

1 Introduction

Recently, targeted personalised therapies have been increased for different types of diseases, such as treatments for bone cancer, [1]. Using computational modeling and numerical methods such as the finite element method (FEM), it is possible to predict the outcome of patient-specific treatments, [2,3]. To make a highly realistic prediction using computational modeling, a 3D reconstruction of the patient must be generated. Therefore, the development of a three-dimensional anatomical model from medical images such as computed tomography (CT) or magnetic resonance imaging (MRI) is a crucial task for the development of patient-specific treatments. Although several studies, [4–6] focus on three-dimensional reconstruction, they do not solve some problems present in DICOM images.

Cardentey et al. [7] offer a general noise reduction approach, excluding CT patient bed noise. Their focus is noise analysis during acquisition, with thresholding segmentation improvement options. Devisivasankari et al. [8] suggest GPUs for parallel computing-based brain structure segmentation. GPU use is constrained to devices with dedicated GPUs, potentially impractical.

Lorca et al. [9] reconstructed a 3D anatomical model from DICOM images. Although focused on skull CT images, they discussed generating a bone model of a lower extremity using semi-automatic threshold-based segmentation. The study evaluated clustering algorithms (K-Means and Fuzzy C-Means) and noted their sensitivity to initial centroid choice, affecting convergence time to a local minimum. Manual centroid initialization restricts algorithm use, and tomographic noise prevents noise-free segmentation.

Generating a segmentation without noise means eliminating any artifact or artifice that is defined as a distortion, addition or error in an image. Moreover, it is not correlated to the anatomical region studied [10]. Therefore, this work presents noise-free, semi-automatic creation of 3D bone models for lower extremities. In-convex areas harmful to treatment planning were eliminated. Additionally, the model is CAD-compatible for FEM computational modeling.

2 Materials and Methods

To generate the three-dimensional models of a bone segment from a lower extremity by using CT images, the work strategy described in Fig. 1 was followed.

2.1 Volume Creation .npy File

To enhance DICOM image portability, two image sets were encoded into a Python-compatible .npy file. These files stores the images in a three-dimensional matrix array, enabling quick access for subsequent steps. The volume, represented by the matrix array $OS = (OS_{i,j,k})$, maintains a constant number of slices (k) while the dimensions of rows (i) and columns (j) vary.

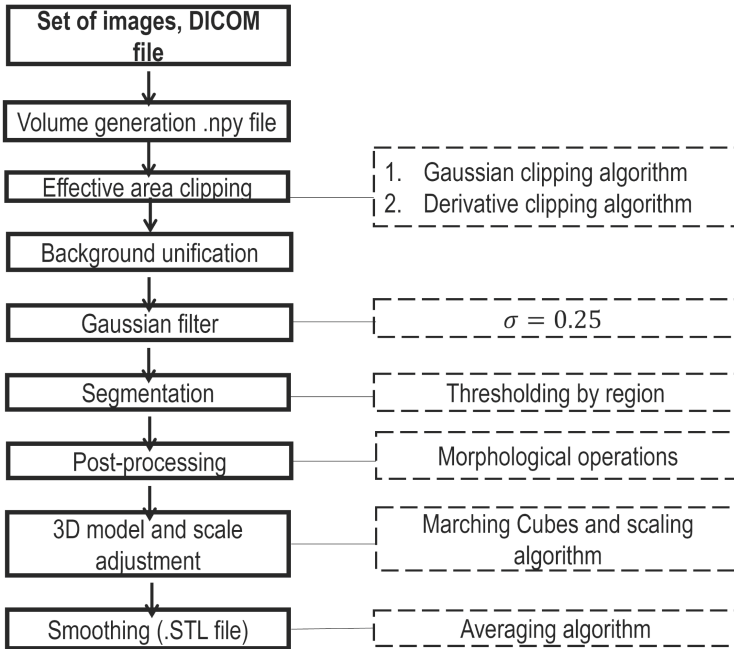


Fig. 1. Flowchart followed to generate a 3D bone segment from lower extremities.

2.2 Effective Area Clipping

To minimize the processing time in subsequent steps, the working area was optimized by focusing just on the regions of interest (RoI) within the images, specifically those areas with tissues. To generate just the 3D model of the RoI, the Derivative and Gaussian Clipping methods were implemented.

Gaussian Clipping Algorithm: In the image sets (OS), relevant tissues were concentrated in the center. Information quality diminishes towards the image edges due to noise from the patient’s bed in the tomography. To exclude undesired regions, the RoI location is represented by a 2D probability distribution function using a Gaussian bell-shaped surface. Adjusting mean and standard deviation parameters aligns the Gaussian shape with the image, enabling selectivity control and width adjustment. This effectively eliminates surrounding regions of the RoI.

To visualize a grayscale image from OS in 3D, a coordinate plane (XY domain) is formed using its row count (M) and column count (N). Each pixel represents a point on this plane. The grayscale intensity level determines the height (Z coordinate) of each point, based on image data. The algorithm utilizes a single image I_m , computed by identifying the maximum grayscale levels using Eq. 1.

$$I_m(i, j) = \max\{OS_{i,j,n}\}, \text{ for } n = 0, \dots, s - 1, \quad (1)$$

where the *max* function identifies the largest shade of gray present in the vector $OS_{i,j,n}$, and n is the number of slices in the original sets with s equal to 481 for both sets.

As a result, an image of $M \times N$, where the value in each pixel is the maximum value found in the volume OS according to each location (i, j) of all the slices, was generated. Since this image has negative values, a normalization was performed. The image I_m is the input to the Gaussian clipping algorithm, and consists of a projection of each pixel onto a Gaussian curve f parameterized according to Eq. 2.

$$f(x, y, \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{1}{2} \left[\frac{(x - \frac{M}{2})^2}{\sigma_x^2} + \frac{(y - \frac{N}{2})^2}{\sigma_y^2} \right]\right). \quad (2)$$

Surface f is characterized by σ_x (standard deviation in columns) and σ_y (standard deviation in rows), controlling curve selectivity for RoI. M and N represent row and column counts. Discrete vectors x and y range from 0 to $M - 1$ and $N - 1$ respectively. Normalizing image I_m scales the largest value to 1. Projection of I_m onto f is achieved through point-to-point multiplication, resulting in the surface f_p .

Data within the 85th percentile were chosen to keep the RoI; this helps to avoid removing the tissue of interest. The result was a new region that ignores the structures farthest from the center of the image, called I_{gm} . This process was applied to an image derived from the OS to identify the RoI boundaries; therefore, it was necessary to select that region on the entire original set. It was implemented by a derivative clipping algorithm that generates a new set of smaller slices.

Derivative Clipping Algorithm: Implemented to determine RoI limits from the OS , a sum of rows and columns in I_{gm} generated epithelial-based silhouettes. Silhouettes underwent normalization and saturation, with values above a 5% threshold indicating RoI boundaries. To locate minimum silhouette values, consecutive numerical derivatives were applied, considering positive values. Peaks caused by noise were mitigated by locating points above 60% of the maximum silhouette value. A correction factor α widened the interval, avoiding tissue loss. Extracting RoI from OS resulted in a new 3D matrix, CS .

2.3 Background Unification

CT images encoded with 12 or 16 bits exhibit negative pixel values due to varying object densities. In this context, tissues denser than water (0 UH) were processed, adhering to the Hounsfield scale (UH) guidelines, [11]. Consequently, an adjustment of image intensities becomes imperative, as demonstrated in the subsequent process.

1. **Translation of each negative value to zero:** The negative values in the background of the images CS are removed, this new set was called CS_T .
2. **Average image acquisition:** From CS_T a single image (I_a) was calculated, by averaging all the slices. The OTSU algorithm was applied to the image. It separates the tissue tones from the remaining background and noise by using a threshold value according to its histogram.
3. **Intensity scaling:** A condition was applied to CS_T , i.e. every value less than the threshold defined by the OTSU algorithm was replaced by this gray intensity, the result was called CS_O .
4. **Translation of each value to zero:** All the intensities in CS_O were traversed so that the minimum absolute value of the volume is 0, resulting in the new set called CS_B .

2.4 Gaussian Filter

CT scans frequently exhibit high-frequency noise, [12]. This noise should be mitigated to prevent distortion in image interpretation. Accordingly, a low-pass filter was employed in the spatial domain, applying a standard deviation of 0.25 to CS_B in accordance with the 2D Gaussian equation (Eq. 3).

$$G(x+i, y+j) = \frac{1}{2\pi\sigma^2} \sum_{i=-1}^1 \sum_{j=-1}^1 \exp\left(-\left[\frac{(x-i)^2 + (y-j)^2}{2\sigma^2}\right]\right) \cdot I(x+i, y+j), \quad (3)$$

where $G(i, j)$ is a filtered image, according to standard deviation σ ; $i, j \in \{-1, 0, 1\}$. The image $I(x+i, y+j)$ is a 3×3 neighborhood as to $x \in [1, M-2]$ and $y \in [1, N-2]$. This process was applied to each image in the CS_B set to generate a new set of filtered images called CS_G . The proposed segmentation process was implemented to each image set by using these filtered images.

2.5 Segmentation

The images set CS_G was divided into three regions to develop the tissue segmentation: 1) proximal zone corresponding to cortical bone, 2) articular area composed by the knee and comprising spongy bone and 3) distal area composed by cortical bone. Equation 4 was implemented to generate a semiautomatic segmentation based on the thresholding technique, [4, 13].

$$H(x, y) = \begin{cases} 1 & \text{if } G(x, y) \geq T_k, k = 1, 2, 3 \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where $H(x, y)$ is a binary image, the result of segmentation according to a threshold defined by zones (T_k), k is the zone number, $G(x, y)$ is an image belonging to CS_G .

Figure 2 shows the methodology and the morphological operations used to segment each region of CS_G corresponding to sets 1 (Fig. 2a) and 2 (Fig. 2),

generating a segmented binary set SS . The morphological operations were part of the post-processing applied to the SS .

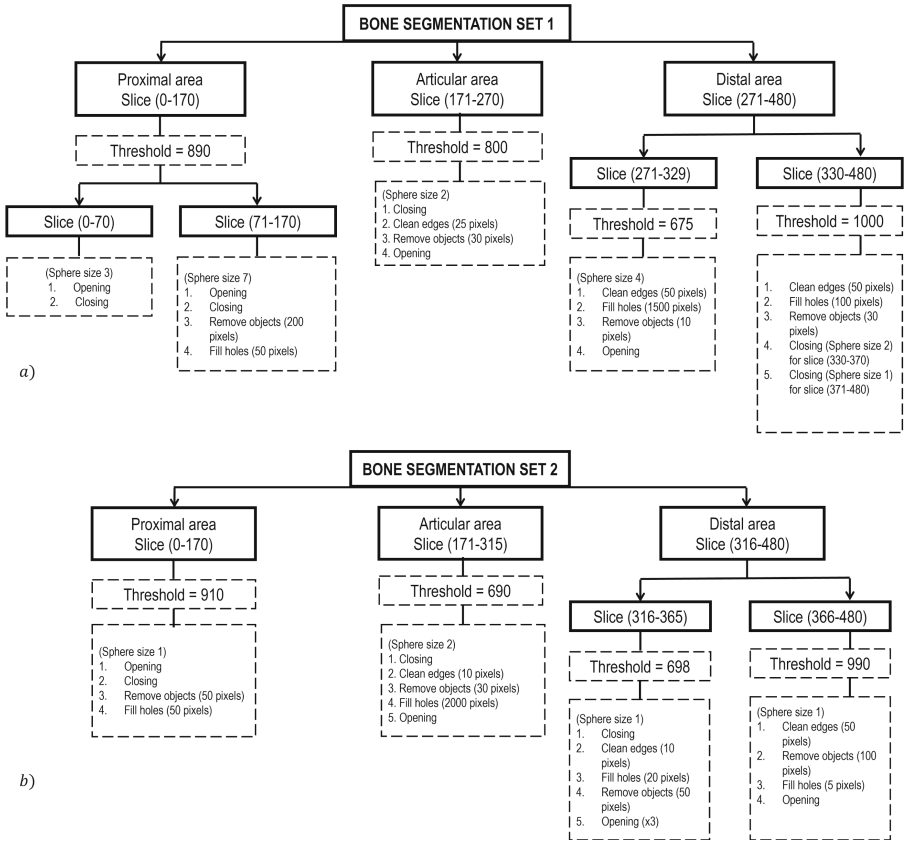


Fig. 2. Flowchart for bone segmentation. Zones divided for segmentation thresholds, followed by tailored morphological operations per tissue characteristics. a) image set 1 and b) image set 2.

2.6 Post-processing

Binary morphological operations eliminated unconnected points from the patient bed and bone texture, [14]. The three-dimensional structuring element was adjusted based on the area, small object size, and holes. These post-processing operations enhanced bone regions, filling gaps, and reducing noise. Additionally, segmentation edges were smoothed.

2.7 3D Model and Scale Adjustment

To generate the three-dimensional model, the *Marching Cubes* algorithm (MC) was used. It is a technique to generate a three-dimensional model from scalar data in a volume, i.e. the scalar values are converted into a polygonal representation (triangle mesh), that approximates the desired surface [15]. The input to the MC algorithm was SS , while the output was the three-dimensional array with all the vertices of the mesh (V), with XYZ positions. This V array was scaled by using three metadata in the DICOM files: *Slice Thickness* (address $[0 \times 0018, 0 \times 0050]$), *Spacing Between Slices* (address $[0 \times 0018, 0 \times 0088]$) and *Pixel Spacing Attribute* (address $[0 \times 0028, 0 \times 0030]$). The first two values were used to correct the image size in the Z direction of V , while the last one has the distance between each pixel in each XY plane.

To adjust the size in the Z -direction called V_z the Eq. 5 was used.

$$L_z = \begin{cases} G \cdot s & , \text{ if } G = D \\ G + D(s - 1) & , \text{ if } G \neq D, \end{cases} \quad (5)$$

where L_z is the original height of the bone, G is the thickness of each slice, D the space between the slices and s the number of slices in the study.

Equations 6 and 7 were used to calculate the scaling factor r_z , it depends on L_z . This factor adjusts the values of the vertices matrix in its Z direction (V_z). As a result, a vertices matrix V_{r_z} adjusted in Z is obtained; its dimensions are in mm.

$$r_z = \frac{L_z}{\max_z \{V_z\} - \min_z \{V_z\}} . \quad (6)$$

and

$$V_{r_z} = V_z \cdot r_z . \quad (7)$$

Model scaling in X and Y directions was achieved using *Pixel Spacing metadata*, represented by the two-element vector $[L_x, L_y]$. L_x indicates the space (mm) between rows, the distance measured between each center of the pixel. L_y is the space between columns. Due to the algorithm automatically sets row and column spacing to the value of unit, it is possible to perform a rescaling by multiplying V_x and V_y volume information. Therefore, Eqs. 8 and 9 were implemented.

$$V_{r_x} = V_x \cdot L_x \quad (8)$$

and

$$V_{r_y} = V_y \cdot L_y, \quad (9)$$

where V_{r_x} and V_{r_y} represent the actual length and width of the bone, respectively, also measured in mm.

2.8 Smoothing

To remove surface imperfections, a Gaussian filter was applied to the binary volume, this helps to generate a smooth mesh in the final exporting 3D solid model. Finally, the averaging algorithm was applied to eliminate either possible disconnections or peaks. This filter is described by Eq. 10.

$$v_i = \frac{v_i + \sum_{n \in N} v_n}{|N| + 1}, \quad (10)$$

where v_i is the current vertex that is smoothing, N is the set of vertices adjacent to the vertex v_i , and $|N|$ is the number of vertices in N .

3 Results

Two DICOM CT image sets were included in the study. Images were originally 16-bit encoded, MONOCHROME2 photometric interpretation. Original size: 512×260 for both sets, totaling 481 slices. Processes in Python 3.9.6 on an 8 GB RAM DELL laptop, AMD Ryzen 5 2500U processor with Radeon Vega Mobile Gfx at 2.00 GHz. Used libraries: *NumPy* 1.25.0, *scikit-learn* 1.3.0, *PyMCubes* 0.1.4, and *os* 3.11.4 for OS interaction.

3.1 Effective Area Clipping

The algorithms described in Sect. 2.2 demonstrated that it is possible to eliminate areas without tissue information by using a Gaussian surface as well as the numerical derivative method.

The application of both algorithms for the RoI selection helps to reduce approximately 50.40% and 65.58% the proportion of rows and columns with respect to the OS for set 1 and 2, respectively. This indicates that the cumulative computational cost in each process was reduced due to focus over the RoI.

Figure 3 depicts the identification of the RoI for set 1. Figure 3a) shows the projection of I_{gm} on Gaussian surface on a yellow-black color scale, in addition to the visualization of the f bell parameterized on a magenta-cyan color scale. The algorithm weighted the central regions more heavily, shifting the noise to levels below 15%. The resultant image was the input to the derivative clipping algorithm. The RoI boundaries were calculated with the derivative clipping algorithm, Fig. 3b) shows the silhouette formed by moving in the direction of the image rows I_{gm} and the boundaries marked with the green line. Likewise, the sum in the column direction is shown in Fig. 3c); the green colored lines spanned a larger interval due to the modulation of the widening factor α .

Figure 4 shows the result of the Gaussian clipping algorithm in grayscale for set 1; Fig. 4a) represent a slice of the original set, Fig. 4b) is the I_m set and the boundaries marked in red in Fig. 4c) correspond to RoI.

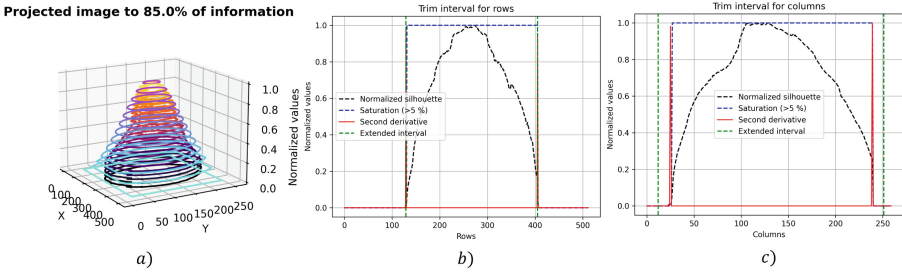


Fig. 3. Identification of the ROI for set 1. a) Projection of I_m on a Gaussian surface, b) the obtained row boundaries, c) the obtained column boundaries.

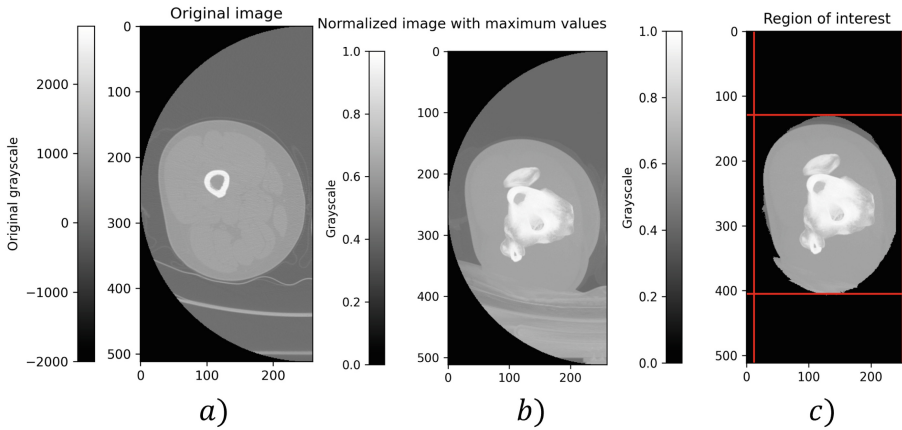


Fig. 4. ROI extraction for set 1. a) Original image, b) Image I_m , c) ROI for the whole set.

3.2 Background Unification and Gaussian Filter

Once the *OS* region of interest was found and extracted, the slices were cleaned. Therefore, the background was unified and a low-pass filter was applied. These process were applied to both image sets. Therefore, Fig. 5 shows the adjustments unification, Fig. 5a) shows the effect of cropped at *OS* for set 1, while Fig. 5b) shows the result of the unification algorithm with the effect of the Gaussian filter. Figure 5c) shows the original image of the set 2, while Fig. 5d) shows the image cropped to the area of interest and the background unification by applying the Gaussian filter with a σ value of 0.25.

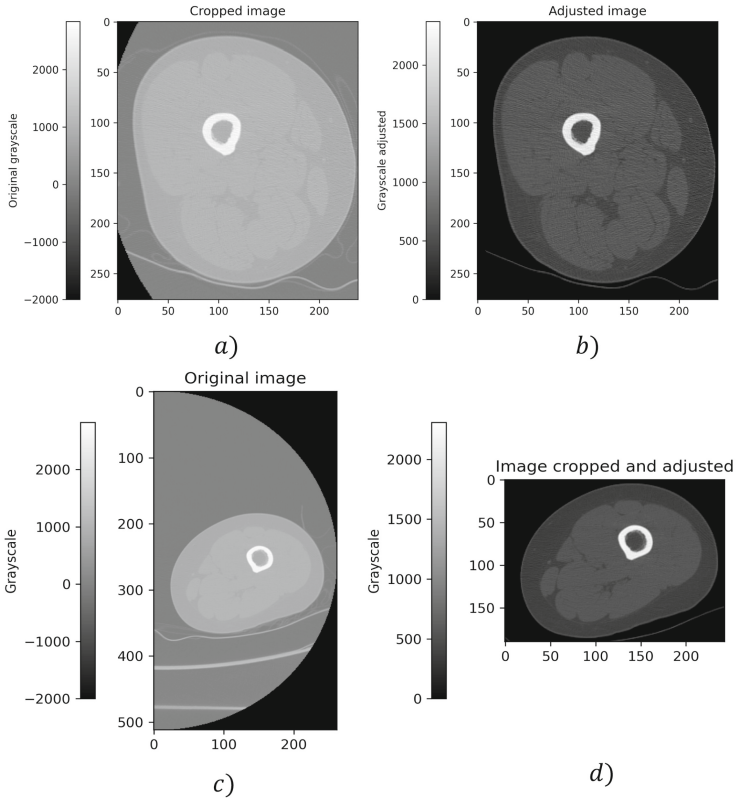


Fig. 5. Adjustment to the set OS for set 1 and 2, a) slice 0 of set 1 clipped and b) slice 0 of set 1 with uniform background and filtered with $\sigma = 0.25$, c) original slice of set 2 and d) slice clipped and background unification.

3.3 Segmentation and Post-Processing

Figure 6 and Fig. 7 depicts three shots corresponding to each of the previously established regions of set 1 and 2, respectively. In Figs. 6a, d and g, slices 100, 230 and 300 are shown cropped, adjusted and low pass filtered, respectively, according to Eq. 3 for set 1. The second column shows the results from the segmentation, according to the selected thresholds, applying a threshold of 890 for slice 100 (Fig. 6b), 800 for slice 230 (Fig. 6e) and 675 for slice 300 (Fig. 6h). These binary images depict isolated points: patient bed, muscle tissue, and bone’s porous structure. The third column (Figs. 6c, f and i) displays outcomes post morphological operations. These operations eliminate unconnected entities, unify regions, and smooth segmentation edges.

Finally, Fig. 7 shows the results of applying the processing steps according to Fig. 2b) to the image set 2. It shows that by modifying some hyper-parameters it is possible to segment bone tissue.

3.4 3D Model and Scale Adjustment

Using these binary images, it was possible to create the three-dimensional reconstruction by using the *Marching Cubes* algorithm present in the *PyMCubes* package. However, as an intermediate process, it was required to implement another Gaussian filter, but this time to the binary images. This operation was performed through the same library. The result was a solid with no sharp edges or cavities, as can be seen in Fig. 8a), b), c) and d) for set 1, and in Fig. 8e), f), g) and h), for set 2.

3.5 Smoothing

Finally, a smoothing process was performed to these models by applying an Averaged algorithm. The results were surface without prominent spikes or incisions, as observed in Fig. 9a), b), c) and d), and Fig. 9e), f), g) and h), for set 1 and 2, respectively. These surfaces are exported as an .STL file compatible with CAD systems.

Table 1 shows a comparison of the reconstructed model dimensions (length, width and height) with those from the 3D models generated by InVesalius 3.1, which was used to manually segment and reconstruct the anatomical models. A maximum discrepancy of 1.24 mm was observed in the length of model 1. Consequently, all the errors can be considered negligible.

Table 2 shows the execution time of the whole 3D reconstruction, from loading the set to the generation of the smoothed 3D model exported in an .STL file for both sets.

Process of segmentation and cleaning

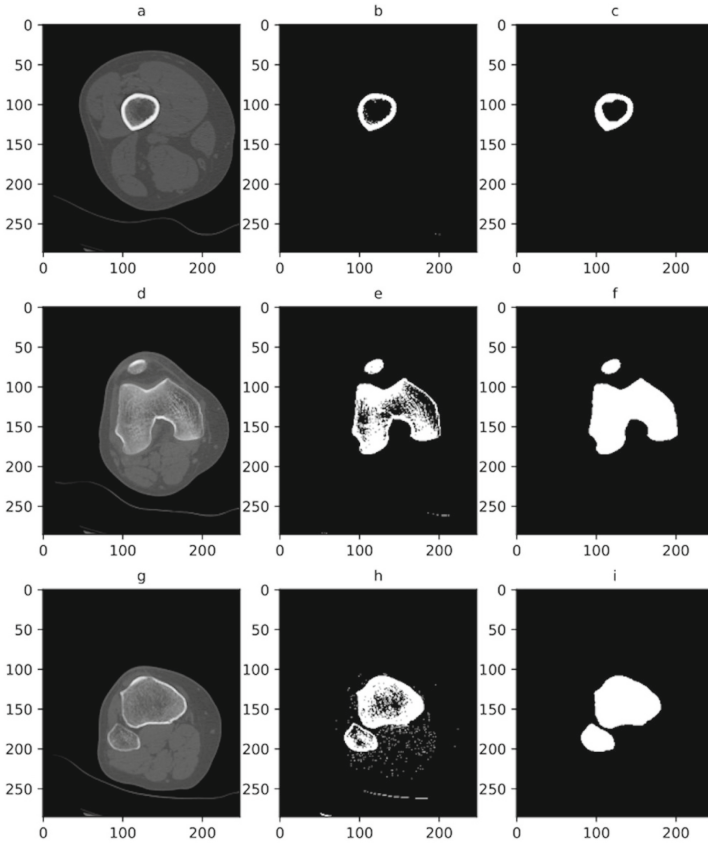


Fig. 6. Image set 1 segmented by zones using threshold-based approach and morphological operations applied per tissue traits. a) Slice 100, d) Slice 230, g) Slice 300, b) Segmentation (threshold: 890), e) Segmentation (threshold: 800), h) Segmentation (threshold: 675), c), f), e) display images after morphological noise reduction.

Table 1. Comparison of dimensions of 3D models

	Model set 1 (mm)	Model set 1 with InVesalius 3.1 (mm)	error model 1 (mm)	Model set 2 (mm)	Model set 2 with InVesalius 3.1 (mm)	error model 2 (mm)
Length	105.52	106.76	1.24	79.36	80.43	1.07
Width	82.22	83.00	0.78	80.81	81.79	0.98
Height	299.08	300.01	0.93	299.13	300.03	0.9

Table 2. Process execution times

Process	Time (sec) Set 1	Time (sec) Set 2
Load set	2.8	2.58
Gaussian clipping	3.5	0.36
Derivative clipping	0.75	0.01
Background unification	1.6	1.45
Zone 1 segmentation	48	3.66
Zone 2 segmentation	2.5	4.52
Zone 3/1 segmentation	3.0	1.16
Zone 3/2 segmentation	2.4	2.01
Save .npy file	0.3	0.58
3D model	14.2	16.76
Smoothing	37.8	16.95
Export .STL file	0.5	0.86
Total	116.85	50.9

Process of segmentation and cleaning: Second set

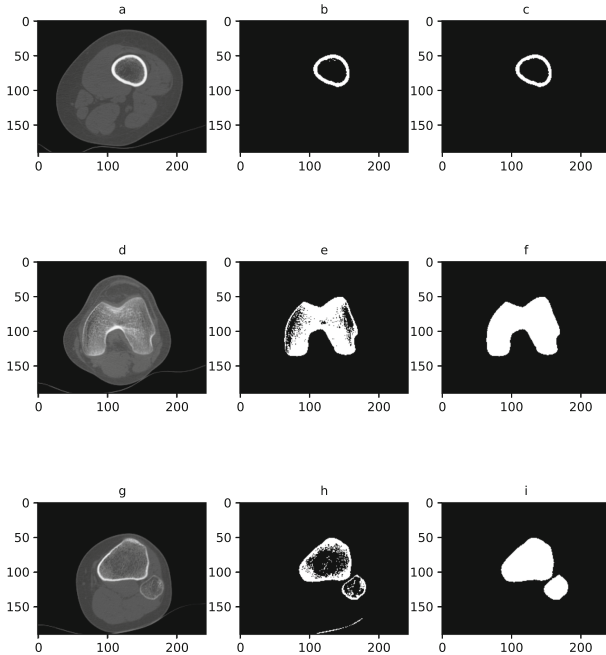


Fig. 7. Image set 2 segmented into zones using threshold-based approach; morphological operations applied per tissue traits. a) Slice 100, d) Slice 230, g) Slice 300, b) Segmentation (threshold: 910), e) Segmentation (threshold: 690), h) Segmentation (threshold: 698), c), f), e) display images after morphological noise reduction.

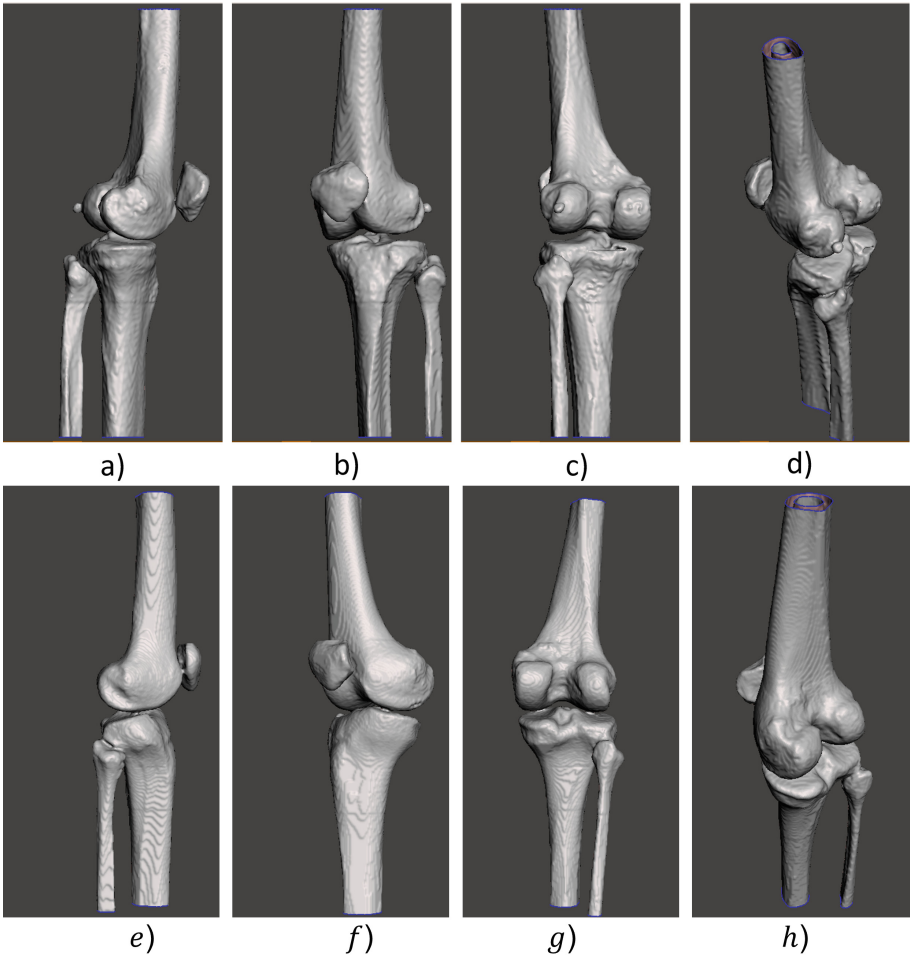


Fig. 8. Views of a three-dimensional reconstruction model using the Marching Cubes algorithm for the set 1 and 2. a), e) side view, b), f) front view, c), g) back view and d), h) isometric view.

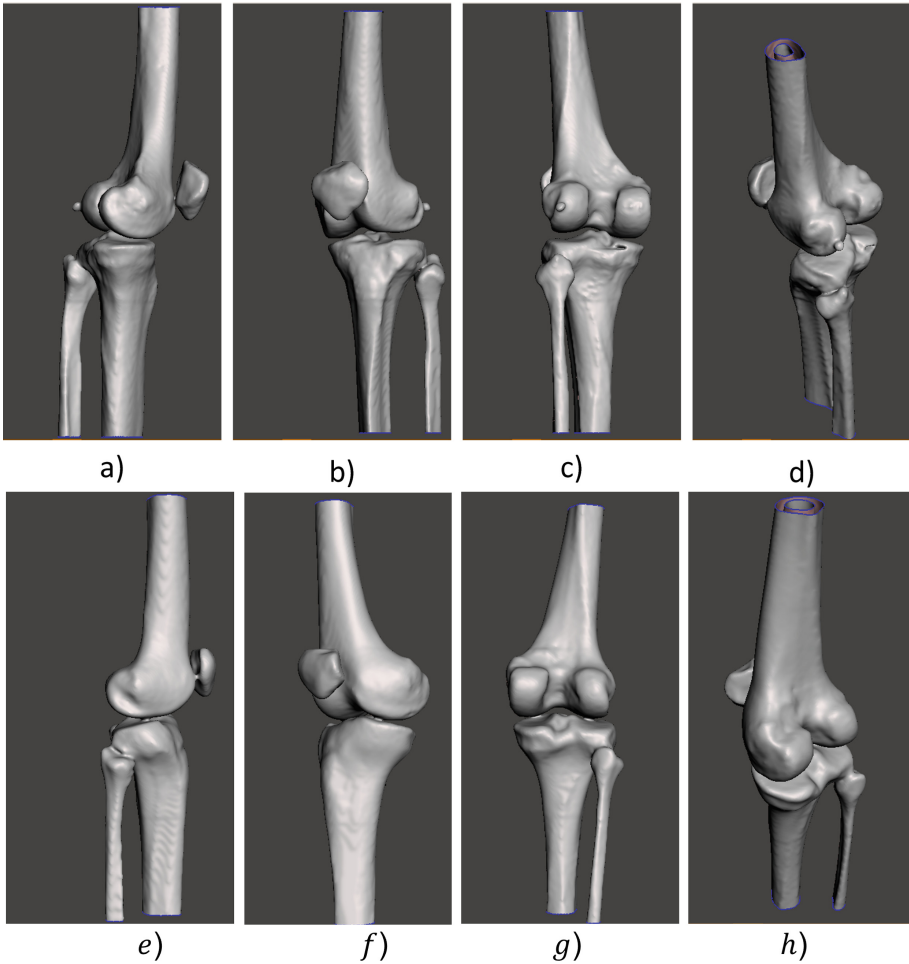


Fig. 9. Views of the smoothed 3D bone model of set 1 and set 2. a), e) side view, b), f) front view, c), g) back view and d), h) isometric view.

4 Conclusions

In this study, a DICOM-based image processing technique for segmenting bone of a lower limb and generates its three-dimensional anatomical models is described. To ensure a realistic 3D reconstruction, the metadata related to the model dimensions were used. Moreover, the dimensions were corroborated by using InVesalius 3.1. Due to the physiology and shape of the bone, a global threshold was ineffective. Therefore, different thresholds were chosen according to the reconstructed region. In addition, the possibility to modify the threshold and post-processing hyper-parameters results in noise-free segmentation. This methodology is replicable for different sets of CT images with the aforementioned characteristics. By

using a Gaussian filter as well as the averaging technique, a three-dimensional model without either prominences or valleys was generated. Consequently, a smoothed mesh was generated.

Manual segmentation is strongly related to the abilities of the operator, it could take from 20 min to 120 min to create a 3D model, [6]. Therefore, one of the main achievements was the reduction of computational cost due to the selection of the RoI. A maximum total process execution time of 1.95 min was achieved; it depends on the images sizes. It is demonstrated that processes optimization completely reduce latency time in all subsequent processes. Moreover, the 3D model reconstructed can be exported as .STL files, to be compatible with CAD systems. These 3D models can be used in multiple biomedical design processes and in the patient-specific treatment planning, such as those required in ablation treatments for bone tumors. The perspectives of this study are focused on the segmentation of the remaining tissues (muscle, fat, and skin) in the CT studies. Therefore, the hyper-parameters of the proposed algorithms must be modified due to the different densities of the tissues.








References

1. Trujillo, C.J., Rico, G., Gutiérrez, J.: Thermal ablation: an alternative to bone cancer. <https://www.medigraphic.com/pdfs/invdiss/ir-2018/ir181e.pdf>
2. Trujillo, C.J., Lejía, L., Vera, A., Rico, G., Gutiérrez, J.: Double slot antenna for microwave thermal ablation to treat bone tumors: modeling and experimental evaluation. *Electronics* **10**(7), 761 (2021)
3. Ramírez, T.J., et al.: Thermal evaluation of a micro-coaxial antenna set to treat bone tumors: design, parametric fem modeling and evaluation in multilayer phantom and ex vivo porcine tissue. *Electronics* **10**(18), 2289 (2021)
4. Pérez-Carrasco, J.-A., Acha, B., Suárez-Mejías, C., López-Guerra, J.-L., Serrano, C.: Joint segmentation of bones and muscles using an intensity and histogram-based energy minimization approach. *Comput. Meth. Program. Biomed.* **156**, 85–95 (2018)
5. Wu, Z., Sullivan, J.M., Jr.: Multiple material marching cubes algorithm. *Int. J. Numer. Meth. Eng.* **58**, 189–207 (2003). <https://doi.org/10.1002/nme.775>
6. Drizdal T., Paulides M.M., van Holthe N., van Rhooon G.C.: Hyperthermia treatment planning guided applicator selection for sub-superficial head and neck tumors heating. *Int. J. Hyperth.* **34**(6), 704–713 (2018). Epub 2017 Oct 16. PMID: 28931333. <https://doi.org/10.1080/02656736.2017.1383517>
7. Ortega, D., Iznaga, A.: Técnicas de Segmentación de Imágenes Médicas [Convención]. 14 CONVENCIÓN CIENTÍFICA DE INGENIERÍA Y ARQUITECTURA, La Habana, Cuba, 2–5 diciembre 2008
8. Devisivasankari, P., Vijayakumar, R.: Parallel watershed method for medical modality image segmentation. In: 2020 International Conference on Emerging Trends in Information Technology and Engineering (ie-ETITE), pp. 1–3. IEEE (2020)
9. Lorca, G., Arzola, J., Pereira, O.: Segmentación de imágenes médicas digitales mediante técnicas de clustering. *Aporte Santiaguino* **3**, 2 (2010)
10. Sartori, P., et al.: Artefactos y artificios frecuentes en tomografía computada y resonancia magnética. *Revista Argentina de Radiología/Argentinian J. Radiol.* **79**(4), 192–204 (2015). <https://doi.org/10.1016/j.rard.2015.04.005>

11. Dougherty, G.: *Digital Image Processing for Medical Applications*. Cambridge University Press, Cambridge, England (2018)
12. Peña-Peñate, A., Silva Rojas, L.G., Alcolea Núñez, R.: Módulo de filtrado y segmentación de imágenes médicas digitales para el proyecto Vismedic. *Rev. Cuba. Cienc. Inform.* **10**(1), 13–27 (2016)
13. Song, F., Tian, Y., Gao, X., Yang, S., Zheng, M.: Research on medical image segmentation method. In: *2021 3rd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*, Taiyuan, China, pp. 577–580 (2021). <https://doi.org/10.1109/MLBDBI54094.2021.00115>
14. Zbarsky, E.S.: *Binary Morphology in Image Processing* (<https://github.com/MathWorks-TeachingResources/Morphology-in-Image-Processing/releases/tag/v1.2.4>). GitHub (2023)
15. Lorensen, W.E., Cline, H.E.: Marching cubes: a high resolution 3D surface construction algorithm. *Comput. Graph. (ACM)* **21**(4), 163–169 (1987)



Nonlinearities of Heart Rate Variability in Junior High School Students Under Mathematics Anxiety

Pilar Cruz-López¹ , Marisela Lázaro-Picazo² , Eric Alonso Abarca-Castro² , Ana Karen Talavera-Peña² , Laura Mercedes Santiago-Fuentes^{1,3} , and José Javier Reyes-Lagos¹  

¹ School of Medicine, Autonomous University of the State of Mexico, 50180 Toluca, Estado de México, Mexico
jjreyesl@uaemex.mx

² Biological and Health Sciences Division, Autonomous Metropolitan University-Lerma (UAM-L), 52005 Lerma, Estado de México, Mexico

³ Department of Health Sciences, Autonomous Metropolitan University-Iztapalapa (UAM-I), 09340 Mexico City, Mexico

Abstract. Heart rate variability (HRV) is commonly assessed using linear methods; however, there is increasing evidence of potential nonlinear aspects in HRV. This study aimed to investigate nonlinear dynamics in Junior High School students with and without math anxiety. We hypothesized that there would be differences in the irregularity and nonlinearity of RR fluctuation time series between students with and without math anxiety. The study encompassed two stages with a sample of 34 first-grade students: an exam stage (EXAM) conducted during the examination period and a regular class day stage (NO EXAM). The Abbreviated Mathematics Anxiety Scale (AMAS) and the multiscale entropy (MSE) of the RR time series were assessed in both stages. Our findings indicated a decrease in the irregularity and nonlinearity of the RR time series in students during the EXAM stage compared to the NO EXAM stage. Additionally, we observed that the AMAS scores were higher in the EXAM group compared to the NO EXAM group. These results suggest that math anxiety may impact specific nonlinear characteristics of the cardiovascular system and may be associated with modifications in cardiac autonomic function. Future investigations should explore alternative entropy-based methods to examine HRV in students experiencing different anxiety-inducing situations.

Keywords: Heart Rate Variability · Nonlinear Dynamics · Math Anxiety

1 Introduction

Regardless of educational level, mathematics provokes negative feelings and symptoms in most students. Among the negative symptoms are anxiety, frustration, and anger [1]. Math anxiety (MA) refers to the negative feelings (fear, tension, or apprehension) that individuals experience when they are faced with situations involving mathematical

tasks [2]. This psychological condition can affect people of all ages and proficiency levels in mathematics. According to the Organization for Economic Co-operation and Development (OECD), 75% of Mexican students have exhibit MA [3].

MA is negatively associated with math achievement; anxiety reduces performance by affecting the process of information acquisition and retrieval [4]. Several studies have suggested that this psychological condition affects working memory and attention capacity [2, 4, 5], leading to a vicious cycle between poor math performance and MA [4]. Some physiological manifestations related to MA have also been reported, including decreased testosterone levels on exam days in young adults [6]; rapid breathing during anxiety episodes [7]; and a decreased heart rate variability (HRV), specifically a low parasympathetic activity under stress conditions in children [8].

Most of the studies have explored how MA is related to math performance and cognitive processes using standardized psychological tests and questionnaires; others have explored the association between MA and linear indexes of HRV to associate the anxiety level with heart rate regulation [8, 9]. Linear indexes to explore autonomic changes in diverse cardiac pathologies are widely used. However, growing evidence of potential nonlinear characteristics in HRV may provide insight into how HRV evolves under different clinical circumstances, adding to the body of predictive knowledge and enhancing conventional time- and frequency-domain analysis [10]. For example, Multiscale entropy is a significant predictor of mortality or heart transplant in congestive heart failure patients [11].

The objective of this study was to longitudinally (pre and post analysis) compare the irregularity and nonlinearity of RR fluctuation time series using an entropy-based method in Junior High School students, both in the presence and absence of math anxiety. We hypothesized that there would be differences in the irregularity and nonlinearity of RR fluctuation time series between Junior High School students with and without math anxiety.

2 Methods

The study obtained approval from the Research Ethics Committee of the School of Medicine at the Autonomous University of the State of Mexico (approval number: 001.2023). The participants comprised 34 first-grade students, aged between 12 and 14 years, from a Junior High School in the State of Mexico. Both boys and girls were included.

The study comprised two longitudinal stages. The first stage (EXAM) occurred during the examination period, where the math test served as a task associated with academic anxiety. The participants completed the Abbreviated Mathematics Anxiety Scale (AMAS). Subsequently, they assumed a supine position for recording two ten-minute electrocardiograms (ECGs), sampled at 200 Hz using the Ganglion Board (Open BCI, New York, United States of America). Following the recording, the participants promptly returned to the classroom to complete the corresponding math test. The second stage (NO EXAM) occurred within the first week after a 14-day spring break period. During this stage, the participants also answered the AMAS scale, and their ECGs were recorded. Once their participation concluded, they returned to the classroom.

The AMAS scale, developed by Hopko et al. (2003), is a widely recognized and psychometrically validated measure of math anxiety [12]. Participants indicated their anxiety levels in response to statements presented in the questionnaire, scored on a Likert-type scale from 1 to 5. The final score ranges from 9 to 45 points, with higher values indicating greater levels of math-related anxiety. The established cut-off point was 16 points, determined based on the questionnaire responses, with scores below 16 points classified as “low math anxiety” according with the AMAS scale.

The Multiscale Entropy (MSE) algorithm was used to calculate the level of nonuniformity or unpredictability of a time series at different scales. Each data point of a coarse-grained time series is estimated using the following Eq. (1) [13].

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-a)\tau+1}^{j\tau} x_i, \quad 1 \leq j \leq N/\tau \quad (1)$$

where τ means the scale factor ($\tau = 1$ to 20), N/τ is the size of each coarse-grained signal for $1 \leq j \leq N/\tau$, x_i is an element in the initial time series and j is an element in the created time series [13].

Using the Iterative Amplitude Adjusted Fourier Transform (iAAFT) method adjusted to 100 iterations, 200 RR time series were generated for each participant in both EXAM and NO EXAM stages. Subsequently, the MSE was calculated for all the resulting signals, with a value of r set to 0.15 and m to 2. The null hypothesis of linear dynamics was assessed for each signal. If the MSE value of the original RR signal was lower than the 5th percentile of the MSE pool generated from 200 randomly created signals, the null hypothesis of linear dynamics was refused, indicating that the initial signal exhibited nonlinearity.

To examine the normality of the data, the D’Agostino & Pearson test was conducted on the results of both stages. If the data met the normality assumption, a paired t-test was utilized to compare the AMAS results between the EXAM and NO EXAM stages. Conversely, a Wilcoxon test was employed if the data did not satisfy the normality assumption. Furthermore, a Wilcoxon test or a paired t-test was employed to compare the MSE results of the original RR time series at the EXAM and NO EXAM stages for each scale. The data were reported as mean \pm SEM to illustrate homogeneous data. Additionally, a chi-square test was performed to compare the percentages of nonlinearity among the 200 generated RR time series.

3 Results

Figure 1 presents the AMAS scores for the EXAM and NO EXAM stages, revealing average scores of 25.7 ± 5.7 and 24.23 ± 5.59 points, respectively. A significant difference ($p < 0.05$), identified by a Wilcoxon test, is observed between two stages with higher scores in the EXAM stage.

In Fig. 2a, significant differences in the RR time series are shown ($p < 0.05$) for scales $\tau = 8$, $\tau = 10$ and $\tau = 12$. Figure 2b presents the percentage of nonlinear surrogate RR signals at different time scales in both stages (EXAM and NO EXAM). A significant difference, identified by a chi-square test, is observed in scale $\tau = 13$. Notably, students

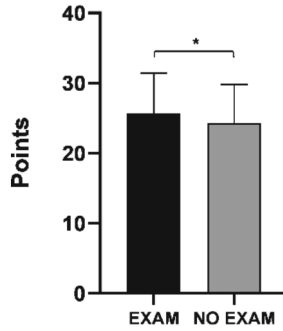


Fig. 1. Abbreviated Mathematics Anxiety Scale (AMAS) scores are represented by the black bar for the EXAM stage and the gray bar for the NO EXAM stage, both assessed on a school day. *($p < 0.05$) differences between EXAM and NO EXAM stages by a Wilcoxon test.

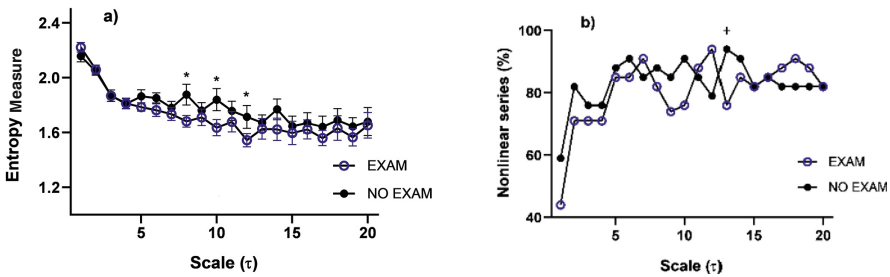


Fig. 2. a) The graph illustrates the Mean \pm SEM of Multiscale Entropy (MSE) for each scale range scale range, on the y-axis shows MSE and on the x-axis scales 1 to 20. The analysis was conducted on RR time series from 34 students, comparing their data before taking the math test (EXAM, represented by blue circles) and on a regular school day (NO EXAM, represented by black circles). b) The graph presents the percentage of nonlinear surrogate RR time series obtained by MSE, utilizing the null hypothesis of linear dynamics for discrimination. A comparison was made between the data collected before the math test (EXAM, represented by blue circles) and on a regular school day (NO EXAM, represented by black circles). Statistical significance was determined as follows: * $p < 0.05$ indicates a significant difference between EXAM and NO EXAM based on a paired t-test, while + $p < 0.05$ indicates a significant difference based on a chi-squared test.

exhibited a lower percentage of nonlinear series before taking the math test (EXAM) compared to measurements on a regular school day (NO EXAM) for these scales.

In Fig. 3, the time scales were divided into short and long-time scales grouped from 1 to 5 beats and 6 to 20 beats respectively. Notably, within the short category, the EXAM stage exhibits a lower percentage of nonlinear series than the NO EXAM stage. This difference is statistically significant, as determined by a chi-square test.

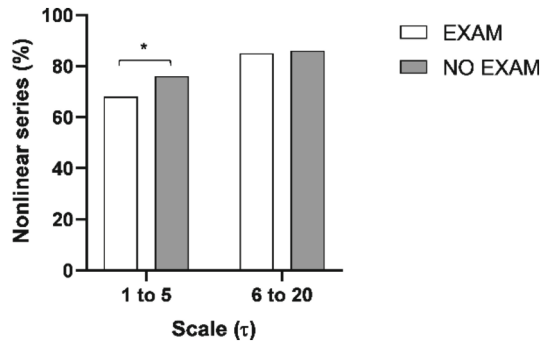


Fig. 3. The bar graph illustrates the percentage of nonlinear dynamics given by MSE, categorized into two scales: short (1 to 5 beats) and long (6 to 20 beats). The bars represent the EXAM stage (open bars) and the NO EXAM stage (gray bars). In particular, the open bars in the short scale are significant (*), indicating differences in the pooled percentage of nonlinear dynamics as determined by a chi-square test ($p < 0.05$).

4 Discussion

We confirmed that junior high school students showed a significant increase in the mean score on the AMAS test in the EXAM stage compared to the NO EXAM stage, as shown in Fig. 1. Some authors suggest that math anxiety may increase when tasks in Mathematics must be solved immediately, as is the case with an exam. Additionally, we observed higher irregularity in the heart rate fluctuations in the NO EXAM stage, as indicated by the Multiscale Entropy (MSE) in Fig. 2a, compared to the EXAM stage. Furthermore, the percentage of nonlinear series was higher in the NO EXAM stage than in the EXAM stage, as shown in Fig. 2b. Figure 3 displays a significant difference in the short time scale between the EXAM and NO EXAM stages, suggesting a linearization of HRV dynamics due to sympathetic hyperactivity as an effect of concomitant vagal withdrawal, while over a longer time period, there is no difference in the modulation of peripheral resistances by the parasympathetic system, which contributes to heart rate nonlinearities in both the EXAM and NO EXAM stage [14]. Our findings suggest that students' math anxiety does not lead to cardiac complications; however, it does exhibit a similar behavior that can be linked to autonomic dysregulation, causing alterations in heart rate dynamics, which can be perceived through their nonlinear characteristics (MSE) [15]. Psychobiological theories propose that emotional dysregulation experienced during anxiety inducing situations can be assessed by measuring resting vagal activity and HRV reductions attributed to the parasympathetic system. Thus, there could be a relationship between vagal responses to anxiety and nonlinear cardiac measures [16]. Other studies on adolescents have also shown that highly anxious individuals exhibit lower entropy of heart rate fluctuations than their less anxious counterparts. Therefore, these complexity measures might help identify adolescents at risk of any disorder of the anxiety field and, with this, take actions to manage anxiety [15, 16].

This study compared the stage before the exam (EXAM) with a regular class day (NO EXAM), and we observed that students exhibited decreased percentage of nonlinearity, as measured by MSE, before taking the exam. This finding is coherent with other studies

that utilize nonlinear methods such as MSE, which suggest that impaired cardiac control affects the distribution of nonlinear HRV dynamics [14].

Currently, conditions of altered cardiac control and multiscale nonlinearity are rarely used, it is often assumed that HRV is nonlinear, however, methods are not often employed to substantiate this claim [14]. It is important to have knowledge of the nonlinear dynamics because it could be associated with the autonomic function and the pathophysiological state (related to the unpredictability and complexity of the signal) [14]. These methods would be more clinically relevant to provide a better interpretation complementing information given by traditional methods that may sometimes not be sufficient to understand the complexity of heart dynamics [17]. Psychological analysis through questionnaires such as the AMAS in conjunction with nonlinear analysis broadens the research landscape to obtain more complete results that lead to a better understanding of the effects of math anxiety.

The justification for the distinctiveness of this study lies in its focus on evaluating math-related anxiety in an academic setting. In contrast to other works [18], this study examines the specific anxiety experienced in relation to mathematics and how it impacts HRV. Additionally, other works [18] calculates both linear and nonlinear indices, but it does not investigate the percentage of nonlinearity in HRV as this study does. Consequently, this study introduces a novel approach by evaluating the level of nonlinearity in HRV, which offers a deeper understanding of the relationship between anxiety and HRV. Furthermore, this study's evaluation of HRV across short and long-time scales is an additional distinctive factor.

Future research will use the percentage of nonlinearity using different entropy-based methods to examine HRV because of potential clinical utility as it provides insights into how math-related anxiety influences HRV across different temporal contexts. The observation of varying HRV patterns over different time scales could lead to refined nuanced conclusions regarding the interplay between anxiety and HRV.

5 Conclusion

In conclusion, this study examined nonlinear dynamics in the heart rate variability (HRV) of Junior High School students with and without math anxiety. The results suggest that math anxiety may influence specific nonlinear features of the cardiovascular system. We found that the RR fluctuation time series exhibits higher nonlinearity and irregular behavior during regular class days. Conversely, before taking the mathematics exam, the reduced nonlinearity of the RR time series is speculated to be influenced by math anxiety, as indicated by the psychological test results. It is suggested that this anxiety may cause subtle cardiac autonomic modifications. Nonlinear analysis of the RR signal can indicate changes in heart rate regulation in anxiety-inducing situations. Further research should explore the potential of nonlinear analysis in assessing heart rate dynamics concerning math anxiety. This type of study is important to recognize the psychophysiological behavior of mathematics anxiety in Junior High School students in order to provide strategies for its management and thus improve their academic performance.

References

1. Martínez Lozano, R.P.: Uso de la técnica de Heart Math para reducir la ansiedad y el estrés ante los exámenes de matemáticas. *Educere* **21**, 275–283 (2017). <https://www.redalyc.org/articulo.oa?id=35655222006>
2. Maloney, E.A., Beilock, S.L.: Math anxiety: who has it, why it develops, and how to guard against it. *Trends Cogn. Sci.* **16**(8), 404–406 (2012). <https://doi.org/10.1016/j.tics.2012.06.008>
3. OECD. Does Math Make You Anxious?, vol. 48. OECD Publishing, Paris (2015). <https://doi.org/10.1787/5js6b2579tnx-en>
4. Carey, E., Hill, F., Devine, A., Szücs, D.: The chicken or the egg? the direction of the relationship between mathematics anxiety and mathematics performance. *Front. Psychol.* **6** (2016). <https://doi.org/10.3389/fpsyg.2015.01987>
5. Passolunghi, M.C., Caviola, S., De Agostini, R., Perin, C., Mammarella, I.C.: Mathematics anxiety, working memory, and mathematics performance in secondary-school children. *Front. Psychol.* **7** (2016). <https://doi.org/10.3389/fpsyg.2016.00042>
6. Jamieson, J.P., Black, A.E., Peláiz, L.E., Reis, H.T.: The impact of mathematics anxiety on stress appraisals, neuroendocrine responses, and academic performance in a community college sample. *J. Educ. Psychol.* **113**(6), 1164–1176 (2021). <https://doi.org/10.1037/edu0000636>
7. Masaoka, Y., Homma, I.: Anxiety and respiratory patterns: their relationship during mental stress and physical load. *Int. J. Psychophysiol.* **27**(2), 153–159 (1997). [https://doi.org/10.1016/S0167-8760\(97\)00052-4](https://doi.org/10.1016/S0167-8760(97)00052-4)
8. Michels, N., et al.: Children’s heart rate variability as stress indicator: association with reported stress and cortisol. *Biol. Psychol.* **94**(2), 433–440 (2013). <https://doi.org/10.1016/j.biopsycho.2013.08.005>
9. Punita, P., Saranya, K., Kumar, S.: Gender difference in heart rate variability in medical students and association with the level of stress. *Natl. J. Physiol. Pharm. Pharmacol.* **6**(5), 431 (2016). <https://doi.org/10.5455/njppp.2016.6.0102325042016>
10. Voss, A., Schulz, S., Schroeder, R., Baumert, M., Caminal, P.: Methods derived from nonlinear dynamics for analysing heart rate variability. *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* **367**(1887), 277–296 (2009). <https://doi.org/10.1098/rsta.2008.0232>
11. Ho, Y.-L., Lin, C., Lin, Y.-H., Lo, M.-T.: The prognostic value of non-linear analysis of heart rate variability in patients with congestive heart failure—a pilot study of multiscale entropy. *PLoS ONE* **6**(4), e18699 (2011). <https://doi.org/10.1371/journal.pone.0018699>
12. Hopko, D.R., Mahadevan, R., Bare, R.L., Hunt, M.K.: The Abbreviated Math Anxiety Scale (AMAS): construction, validity, and reliability. *Assessment* **10**(2), 178–182 (2003). <https://doi.org/10.1177/1073191103010002008>
13. Costa, M., Goldberger, A.L., Peng, C.K.: Multiscale entropy analysis of biological signals. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **71**(2) (2005). <https://doi.org/10.1103/PhysRevE.71.021906>
14. Silva, L.E.V., et al.: Nonlinearities of heart rate variability in animal models of impaired cardiac control: contribution of different time scales. *J. Appl. Physiol.* **123**(2), 344–351 (2017). <https://doi.org/10.1152/japplphysiol.00059.2017>
15. Bornas, X., Llabrés, J., Noguera, M., López, A.M., Gelabert, J.M., Vila, I.: Fear induced complexity loss in the electrocardiogram of flight phobics: a multiscale entropy analysis. *Biol. Psychol.* **73**(3), 272–279 (2006). <https://doi.org/10.1016/j.biopsycho.2006.05.004>
16. Bornas, X., Balle, M., De la Torre-Luque, A., Fiol-Veny, A., Llabrés, J.: Ecological assessment of heart rate complexity: differences between high- and low-anxious adolescents. *Int. J. Psychophysiol.* **98**(1), 112–118 (2015). <https://doi.org/10.1016/j.ijpsycho.2015.07.007>

17. Fernandes de Godoy, M.: Nonlinear analysis of heart rate variability: a comprehensive review. *J. Cardiol. Ther. (Hong Kong)* **3**(3), 528–533 (2016). <https://doi.org/10.17554/j.issn.2309-6861.2016.03.101-4>
18. Chalmers, J.A., Quintana, D.S., Abbott, M.J.-A., Kemp, A.H.: Anxiety disorders are associated with reduced heart rate variability: a meta-analysis. *Front. Psychiatry* **5** (2014). <https://doi.org/10.3389/fpsy.2014.00080>



Respiratory Rate Estimation from Tracheal Sound Analysis Using a Mobile Application for Smartphones

Alison Contreras-Rodríguez¹ ^(✉), Nemesio Olvera-Montes² ,
Carlos Mariaca-Gaspar² , Abel Mosco-Vargas¹ , Victor Maya-Venegas¹ ,
Sonia Charleston-Villalobos³ , Tomas Aljama-Corrales³ , and Bersain Reyes¹ 

¹ School of Sciences, Universidad Autónoma de San Luis Potosí, 78295 SLP, Mexico
A310209@alumnos.uaslp.mx

² Mechatronics, Mechanical and Industrial Engineering Division, Tecnológico de Estudios Superiores de Ecatepec, 55210 Ecatepec de Morelos, Mexico

³ Electrical Engineering Department, Universidad Autónoma Metropolitana Iztapalapa, 09340 Mexico City, Mexico

Abstract. Diagnosing several lung diseases is challenging and usually requires various methods and tests, including a patient's clinical history, auscultation, spirometry, pulmonary function tests, and other methods using more specialized medical devices. For its part, the pulmonary auscultation with the mechanic stethoscope represents an early approach to the disease. However, it is highly subjective. Therefore, acquiring and analyzing respiratory sounds through mobile computerized devices, such as smartphones, has been an attractive alternative for the estimation of physiological parameters, including respiratory rate (RR). This study explored the estimation of RR performed completely on a single smartphone device, from the tracheal sound acquisition, signal conditioning and processing, and results report. To this end, a mobile application was developed for the Android system, and acquisitions were made in ten ($N = 10$) healthy volunteers while breathing at different metronome RR. The results obtained with the app were compared with the ones obtained from a respiratory reference signal. Mean absolute errors of 0.06, 0.18, 0.66 and 0.54 bpm were found for RR of 6, 12, 18 and 24 bpm, respectively. The promising results point out to test the mobile-developed system in breathing maneuvers that include temporal changes in RR.

Keywords: tracheal sounds · respiratory rate · smartphone

A. Contreras-Rodríguez—Data acquisition, data analysis, and first draft of the manuscript.
N. Olvera-Montes and C. Mariaca-Gaspar—Design and implementation of the mobile application.
A. Mosco-Vargas and V. Maya-Venegas—Data acquisition and creation of the database.
S. Charleston-Villalobos, T. Aljama-Corrales and B. Reyes—Conceived the original idea, oversaw the study, and writing process.

1 Introduction

Many years ago, the World Health Organization ranked cardiorespiratory diseases among the top ten causes of death worldwide, independently of the income level of the countries. A vital sign that allows their detection and monitoring is respiratory rate (RR). Among simple methods used to estimate RR are human observation and palpation, as well as auscultation using a mechanical or digital stethoscope [1]. The latter allows the respiratory acoustic signals to be amplified and filtered, which facilitates their hearing that sometimes cannot be achieved through the classical mechanical stethoscope. This study focused on respiratory tracheal sounds (TS) whose vibrations are originated in the trachea and upper lung and transmitted through the surrounding tissue up to the neck surface and its vicinity, where they are heard, analyzed, and related to physiological characteristics that represent health or disease. Therefore, TS and RR are relevant to learn about the respiratory system [2].

Computerized respiratory sound analysis systems (CORSA) have helped overcome traditional auscultation's limitations, making it possible to acquire, store, reproduce, analyze, and display information on various respiratory sounds, including TS. Nowadays, CORSA systems have begun to permeate clinical settings, and commercial medical devices currently exist, e.g., the MASIMO Rainbow equipment, which allows the estimation of RR using an acoustical approach based on TS [3]. Alternatively, efforts have recently been made to use smartphones to develop mobile CORSA systems (mCORSA), given the characteristics of ubiquity, mobility, and cost-effective sensors of these devices. This approach has allowed the implementation of digital signal processing algorithms on smartphones to facilitate the application of CORSA systems in different locations, without traveling to more specialized sites, and even to perform the acquisition and analysis of respiratory sounds directly at the patient's bedside [4].

Our research group has made some efforts regarding RR estimation from TS and the development of mobile applications (apps) for TS analysis. In a first effort, the estimation of RR was performed offline, i.e., the smartphone-acquired TS was analyzed on a regular computer, not in the mobile device itself [5]. In a posterior effort, an mCORSA system allowing the acquisition, processing, and display of results was developed to detect adventitious respiratory sounds [4]. In a recent effort, a mobile app was implemented to compute the Shannon entropy (SE), which quantifies the uncertainty of a stochastic signal of a previously uploaded file, sound or not, to the app [6]. It has been reported that the SE of TS provides a surrogate signal of the normalized respiratory airflow, providing the rationale for using them to estimate RR [7]. Hence, the development of an mCORSA system that allows RR estimation directly from TS acquired by the smartphone device was pending for our research group.

In this study, we addressed the end-to-end estimation of RR on a single smartphone using a tracheal sounds approach which, to the best of our knowledge, has not been addressed yet. We developed a mobile app to govern the acquisition of TS using an acoustical sensor attached to the smartphone. The app computes the RR based on the power spectral density (PSD) of the SE of the acquired TS. The results are displayed right on the mobile device. Tests were made with metronome breathings maneuvers while considering the RR derived from a piezoelectric respiratory band as a reference.

2 Methodology

2.1 Hardware

An acoustical sensor was employed to acquire TS, as described in our previous study [4]. Briefly, the sensor comprises a subminiature electret microphone (BT-2159000, Knowles Electronics, IL, USA) encapsulated in a plastic bell and connected to the 3.5 mm audio input of the smartphone. Regarding the smartphone, the Huawei Y6 2018 (Huawei, Shenzhen, China), which has 2.0 GB of RAM and runs an Android 8.0 operating system, was used. The mCORSA system was governed by the mobile app, whose design and implementation are described below.

2.2 Mobile App

The app was developed in Android Studio Flamingo 2022.2.1 (Google, CA, USA) using Java (Oracle Corp., CA, USA). The app oversaw the setting, starting, and stopping of the TS acquisition stage, as well as performed the digital signal processing required to estimate RR from the acquired sounds. Hence, the design and implementation were divided into two main parts: 1) the graphical user interface (GUI) and 2) the internal digital signal processing of the TS.

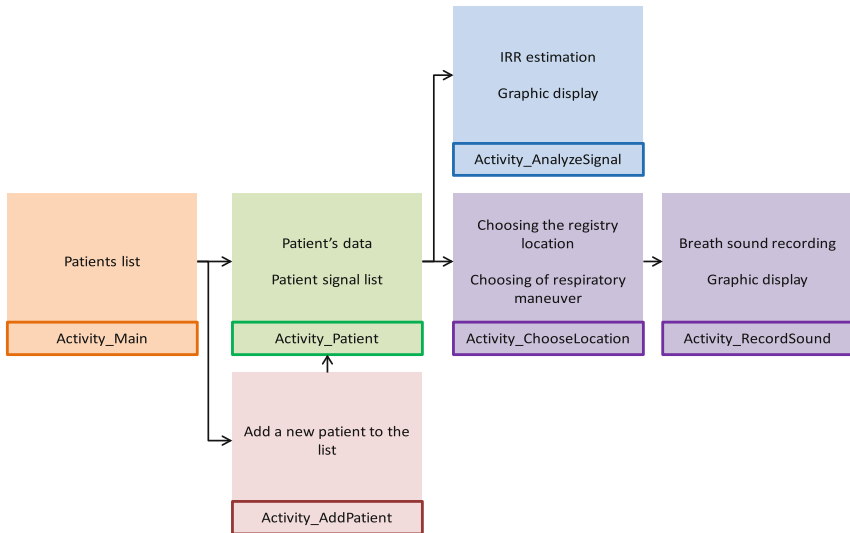


Fig. 1. Main Flow of activities of the developed mobile app.

Several activities were implemented regarding the GUI, as shown in Fig. 1. First, the user can select a patient (volunteer) from an existing list in the main activity. In the add patient activity, the user is prompted to introduce the patient's data, including first and last name, age, and gender. Subsequently, the user can select among different breathing maneuvers and the location of the acquired signal. The neck location was always used

in this study, as only TS was acquired. Finally, the user can start recording the TS by clicking an acquisition signal button. After the acquisition is finalized and saved, the acquired data is displayed if not discarded. If the user selects an existing patient from the list, the app uploads its available data and recordings. Finally, the user can choose the “analyze signal” activity to perform the data analysis, resulting in the visualization of the TS, its normalized SE signal, its PSD, and the corresponding estimated RR.

TS is acquired using the aforementioned acoustic sensor, where the corresponding electrical signal is digitized at a sampling frequency of 10 kHz and 16 bits-per-sample, to comply with Nyquist sampling theorem and international CORSA guideline recommendations. The TS is bandpass filtered using a digital FIR filter with a Hamming window of 500 samples and 100 and 3000 Hz cutoff frequencies. The filtered TS are normalized between 0 and 1, resulting in the signal $x_1[n]$ of the block diagram in Fig. 2. The files in this stage correspond to the raw audio signal and the text file with the amplitudes obtained from the TS. The block diagram shown in Fig. 2 expands the “analyze signal” activity block from Fig. 1.

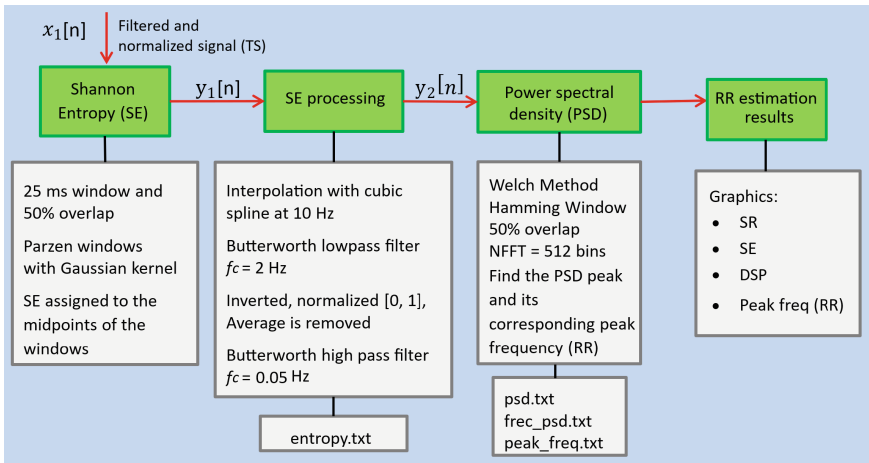


Fig. 2. Block diagram of the digital signal processing performed by the mobile app.

First, the SE signal of the acquired and preprocessed TS, $x_1[n]$, is estimated following the method described in our previous work [4, 5]. Briefly, TS $[n]$ is divided into successive windows of 25 ms with 50% overlap, and the probability density function (pdf) is calculated for each window using the Parzen window method with a Gaussian kernel. The SE value of each window is assigned to its midpoint, resulting in the time series $y_1[n]$ in Fig. 2. To achieve a uniform sampling frequency adequate for RR estimation accordingly to Nyquist sampling theorem, as well as to focus on the RR frequency range, $y_1[n]$ is interpolated using cubic splines at 10 Hz and filtered with an IIR Butterworth lowpass filter with a cutoff frequency of 2 Hz. To facilitate the subsequent signal processing, the SE signal is reversed, normalized between 0 and 1, its mean (DC component) is removed, and filtered again with an IIR Butterworth high pass filter with a cutoff frequency of 0.05 Hz, resulting in the time series $y_2[n]$.

After computing the processed SE signal, the RR is obtained via the frequency corresponding to the peak of the PSD of the SE signal. To this end, the PSD is estimated with the Welch periodogram method, using a Hamming window of 10 s, with 50% overlap between consecutive windows, and $NFFT = 512$ frequency bins.

Once all the calculations are completed, the TS signal, the SE signal, the PSD of the SE, and the corresponding RR are displayed in the mobile app. Finally, the related text files are stored for possible export to external programs like MATLAB or Python. It is worth mentioning that, in this study, the resulting text files with information about the SE signal, its PSD, and its corresponding peak frequency, i.e., estimated RR, were exported to MATLAB R2023a (The MathWorks, MA, USA) only for comparison purposes with the respiratory reference, i.e., no more signal processing was performed outside the developed mobile app.

2.3 Data Acquisition

Data from ten ($N = 10$) respiratory healthy volunteers were acquired, with ages ranging from 19 to 23 years old, seven women and three men. Before the acquisition, the experimental protocol was explained to the volunteers, and their informed consent to participate in the study was obtained according to the Declaration of Helsinki.

To acquire the TS, the acoustic sensor was placed on the lateral surface of the neck, as shown in Fig. 3, using a two-sided adhesive ring. Simultaneously, the respiratory reference and ECG signals were acquired using the Biosignal Plux wireless system (PLUX Wireless Biosignal, Lisboa, Portugal), using a sampling frequency equal to 1000 Hz. It is worth mentioning that the ECG signal was not used in this study, but it was acquired for its future use by our research group when developing and comparing different RR estimation methods. The respiratory reference signal was acquired using a piezoelectric respiratory band, while the ECG was acquired following a configuration that emulated DI by placing three adhesive Ag/AgCl electrodes on the thoracic area.

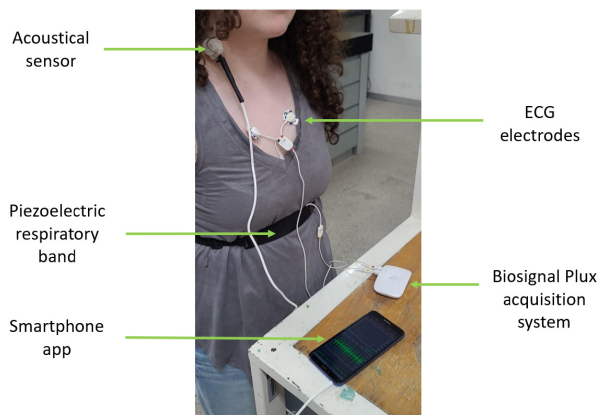


Fig. 3. Experimental setting showing a volunteer during the signal acquisition stage.

For this study, data from fixed RR maneuvers were acquired. To this end, each volunteer was instructed to perform four different metronome breathing maneuvers, each at a fixed RR equal to 0.1, 0.2, 0.3, and 0.4 Hz, equivalent to 6, 12, 18, and 24 breaths-per-minute (bpm). Each recording at a given fixed RR lasted one minute. Visual metronome feedback was provided to the volunteers through a display placed in front of them to facilitate them following the required breathing rhythm.

2.4 Data Analysis

The mCORSAs-based RR estimates were compared to the ones obtained from the reference piezoelectric sensor. The reference RR estimates were automatically computed via the PSD of the respiratory reference signal by searching for the frequency corresponding to its maximum peak. Data analysis included descriptive statistics, boxplots, computing absolute errors (AE), in bpm and normalized percent, as well as Bland-Altman analysis and non-parametrical statistical tests using two-sided Wilcoxon signed-rank tests, considering a 5% significance level.

3 Results and Discussion

Screen captures of the implemented mobile app for the estimation of RR from TS are displayed in Fig. 4, where it can be seen the GUI developed for the “add patient” activity (left panel), for the “sound recording” activity (central panel), and the “results display” activity (right panel). The app lets us quickly introduce the patient’s information and validate the data. The app allows the user to start and stop the recording of TS, as well as to save or discard the acquisition. After performing all the digital signal processing, from the TS to the RR estimation, the app displays several signals, including the TS alone, the TS together with its corresponding SE signal, and the PSD of the SE together with the detected maximum PSD peak and the frequency peak (RR estimate). It is possible to observe that, for the example shown in Fig. 4, the volunteer was asked to breathe at 12 bpm, and the calculated peak frequency using the app was 0.195 Hz, corresponding to a RR equal to 11.70 bpm.

The RR estimation results obtained for all volunteers are summarized in Table 1 for each metronome maneuver, where RR_{ref} denotes the RR reference values based on the piezoelectric respiratory band, and RR_{app} indicates the RR estimated values computed with the developed mCORSAs system. AE represents the absolute error of the estimate, in bpm units and normalized units, concerning the reference value. Each value is presented as mean \pm standard deviation, median, and (minimum, maximum). It was found that the higher AE was 1.2 bpm, corresponding to 6.41% of the reference fixed RR of 18 bpm.

Figure 5 shows the normalized AE for each of the four metronome maneuvers (M1–M4), where it can be noted that data does not follow a normal distribution. No statistically significant differences were found between the median of RR estimates from the mCORSAs system and those from the piezoelectric reference sensor, at the 5% significance level, for all maneuvers. A statistically significant bias of -0.27 bpm was found, and the 95% limits-of-agreement (LoA) were -1.07 and 0.53 bpm.

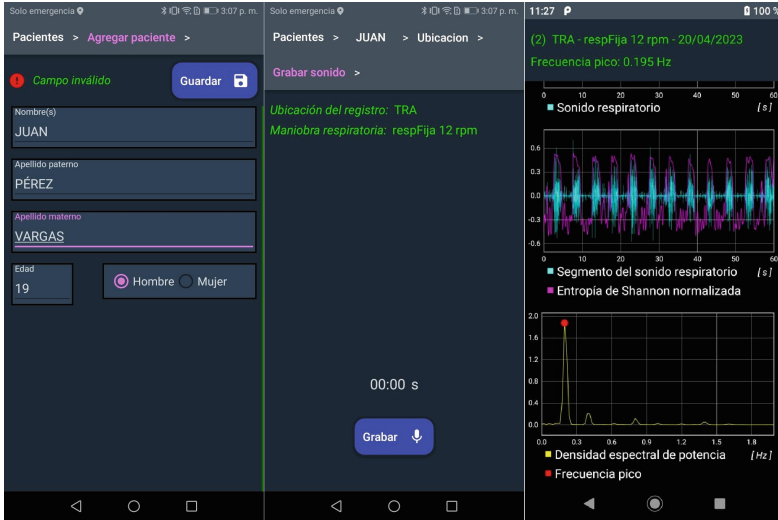


Fig. 4: Screenshots of the developed mobile app for RR estimation using TS. Left: Add patient activity. Center: Sound recording activity. Right: Results display activity.

The obtained results in this study are comparable to previous efforts reported in the literature, e.g., the bias and LoA were found to be 0.11 and -1.41 to 1.63 bpm in a study employing smartphone-acquired TS [5]. In contrast, the median error was less than 1% in a study using smartphone-acquired nasal sounds [6], but our system has the advantage of performing an end-to-end RR estimation in the smartphone.

Table 1: Results obtained for the metronome breathing maneuvers ($N = 10$).

Metronome breathing (bpm)	RRref (bpm)	RRapp (bpm)	AE (bpm)	AE (%)
6	5.76 ± 0.19 5.82 (5.22, 5.82)	$5.82 \pm 8.7 \times 10^{-16}$ 5.82 (5.82, 5.82)	0.06 ± 0.190 0.00 (0.00, 0.60)	1.15 ± 3.63 0.00 (0.00, 11.49)
12	11.88 ± 0.29 11.70 (11.70, 12.30)	$11.70 \pm 1.7 \times 10^{-15}$ 11.70 (11.70, 11.70)	0.18 ± 0.29 0.00 (0.00, 0.60)	1.46 ± 2.39 0.00 (0.00, 4.48)
18	18.18 ± 0.189 18.12 (18.12, 18.72)	$17.52 \pm 3.5 \times 10^{-15}$ 17.52 (17.52, 17.52)	0.66 ± 0.19 0.60 (0.60, 1.20)	3.62 ± 0.98 3.31 (3.31, 6.41)
24	24.06 ± 0.19 24.00 (24.00, 24.60)	23.76 ± 0.58 23.40 (23.40, 24.60)	0.54 ± 0.19 0.60 (0.00, 0.60)	2.25 ± 0.79 2.50 (0.00, 2.50)

Values presented as mean \pm standard deviation, median, (minimum, maximum).

Besides the promising results of this study, some limitations should be addressed in future studies. First, the sensor is sensitive to other sounds coming from the respiratory tract, and for this reason the volunteers were asked not to swallow or talk during the acquisitions. Second, the sample size is small, and we are trying to increase it. Third, the piezoelectric sensor used as a reference is not as good as respiratory bands based on an inductive sensor, but unfortunately the latter is not available in our laboratory. Finally, more breathing maneuvers should be explored, including spontaneous breathing and abrupt changes in RR, where incorporating time-frequency analysis, e.g., the spectrogram, would be helpful to contend with the time-varying nature of the corresponding respiratory rates.

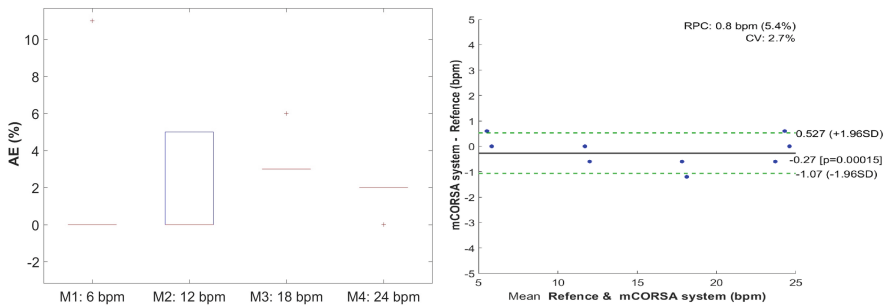


Fig. 5. Results of RR estimation via mCORSA system. Left: Boxplot of the absolute error for each metronome maneuver (M). Right: Bland-Altman graph.

4 Conclusions

The end-to-end estimation of respiratory rate using a smartphone-based system for tracheal sound analysis was explored in this study, obtaining results comparable to the ones reported in the literature for metronome breathing. It is worth mentioning that the recordings of this study were performed in a regular classroom, with the associated acoustical noise from the environment, pointing out the feasibility of extending the RR estimation beyond clinical or research settings.

References

1. Al-Khalidi, F.Q., Saatchi, R., Burke, D., Elphick, H., Tan, S.: Respiration rate monitoring methods: a review. *Pediatr. Pulmonol.* **46**(6), 523–529 (2011). <https://doi.org/10.1002/ppul.21416>
2. Beck, R., Rosenhouse, G., Mahagnah, M., Chow, R.M., Cugell, D.W., Gavriely, N.: Measurements and theory of normal tracheal breath sounds. *Ann. Biomed. Eng.* **33**(10), 1344–1351 (2005). <https://doi.org/10.1007/s10439-005-5564-7>
3. Acoustic Respiration Rate (RRa), MASIMO. <https://www.masimo.com/evidence/pulse-co-oximetry/rra/>. Accessed 28 June 2023

4. Reyes, B.A., Olvera-Montes, N., Charleston-Villalobos, S., González-Camarena, R., Mejía-Ávila, M., Aljama-Corrales, T.: A smartphone-based system for automated bedside detection of crackle sounds in diffuse interstitial pneumonia patients. *Sensors* **18**(11), 3813 (2018). <https://doi.org/10.3390/s18113813>
5. Reyes, B.A., Reljin, N., Chon, K.H.: Tracheal sounds acquisition using smartphones. *Sensors* **14**(8), 13830–13850 (2014). <https://doi.org/10.3390/s140813830>
6. Nam, Y., Reyes, B.A., Chon, K.H.: Estimation of respiratory rates using the built-in microphone of a smartphone or headset. *IEEE J. Biomed. Health Inform.* **20**(6), 1493–1501 (2015). <https://doi.org/10.1109/JBHI.2015.2480838>
7. Yadollahi, A., Moussavi, Z.M.: A robust method for estimating respiratory flow using tracheal sounds entropy. *IEEE Trans. Biomed. Eng.* **53**(4), 662–668 (2006). <https://doi.org/10.1109/TBME.2006.870231>



Analysis of Cardiovascular and Cerebral Interactions in Response to Cognitive Stressors Stimulus

Verónica Janeth Martínez-Hernández^(✉)  and Guadalupe Dorantes-Méndez 

Facultad de Ciencias, Universidad Autónoma de San Luis Potosí, 78295 San Luis Potosí, San Luis Potosí, Mexico
veronicajmh1@gmail.com

Abstract. Stress can cause cardiovascular alterations and maladaptive reactions. Stress management techniques such as controlled breathing could be helpful to decrease the physiological alterations caused by prolonged stress levels. With the use of transfer entropy (TE), we can assess the interactions between the cardiovascular and cerebral systems and assess whether these interactions are affected by the application of controlled breathing. In this study, a test protocol was conducted consisting of the stages of rest, first cognitive task (mental arithmetic + Stroop), controlled breathing, second cognitive task (mental arithmetic + Stroop), and recovery. The goal was to evaluate changes in TE between maneuvers in 17 healthy volunteers. The results showed that most interactions were from brain to heart in both cognitive tasks and that the sympathetic pathway was the most affected. In addition, a higher number of significant interactions from the heart to the brain in the second cognitive task after applying controlled breathing, specifically from the vagal part. This suggests that controlled breathing is indeed influencing task, but further training in the breathing technique is needed to find possible significant differences between the tasks.

Keywords: Cognitive Tasks · Controlled Breathing · Transfer Entropy

1 Introduction

Any stimulus, whether physical or mental, that disrupts in homeostasis and triggers its readiness to confront or escape a threat is considered a stressor [1]. However, when it occurs frequently or for a prolonged period, it can lead to alterations in cardiovascular responses, favoring the development of hypertension, endothelial dysfunction, and atherosclerosis. It can also cause maladaptive reactions including depression, anxiety, and cognitive impairment [1, 2].

Due to the physiological changes associated with stress, various systems are involved in this process. First, the hypothalamus integrates information from sensory and visceral pathways, activating two classical stress pathways known as the hypothalamic-pituitary-adrenal (HPA) axis and the sympathetic-adreno-medullary (SAM) axis [3].

Given the significant consequences of chronic stress, it is important to employ stress management techniques such as controlled breathing, which has been reported to decrease cortisol levels during a sustained attention task after undergoing controlled breathing training [4].

This study includes controlled breathing as a biofeedback signal. Controlled breathing involves maintaining a low and controlled respiratory frequency, which enhances respiratory sinus arrhythmia. By stimulating the baroreflex, deep breathing helps to decrease the elevated heart rate during periods of stress [5].

Therefore, the most common techniques for analyzing cognitive tasks response are time and frequency domain analysis. By applying a mental arithmetic task using electrocardiogram (ECG) and electroencephalogram (EEG) signals, sympathetic activity is only activated in the later stages of the mental arithmetic task [6]. Coherence analysis using EEG has revealed enhanced connectivity within interhemispheric areas in the δ and θ bands [7]. However, the techniques do not reveal the type of interaction between the cardiovascular and cerebral systems. Therefore, an alternative is the use of transfer entropy (TE), as an index that indicates the flow of information and its direction [8].

A study employing TE analyzed a cardiovascular network and a cerebral subnetwork during the application of stressors like mental arithmetic and sustained attention [8]. The findings indicated that mental arithmetic led to an increase in TE in the δ band and bidirectional RR interval. Additionally, an increase in TE from RR intervals to the β band was observed [8]. Conversely, sustained attention showed connections from the α band to RR intervals and from RR intervals to the θ band in a unidirectional manner [9]. Furthermore, when cortical regions were used in a cognitive task factor study, it was found that TE increased from the parietal to frontal region [10].

In this study, the Stroop test and mental arithmetic were used as cognitive task. The Stroop test can induce an increase in sympathetic activity [11]. On other hand, mental arithmetic has been shown to correlate with elevated cortisol levels, which is considered a biological indicator of stress. Therefore, both tasks are considered effective cognitive stressors [12].

This study aims to examine the interactions between the cardiovascular and cerebral systems when responding to cognitive task, and to determine if controlled breathing has an impact on these interactions.

2 Methodology

2.1 Database

The data were obtained from a database with ECG and EEG signals of 17 volunteers, 9 males and 8 females, with an age of 21.82 ± 1.38 years. The data was acquired using a board Cyton of OpenBCI® with software version OpenBCI GUI v3.3.0. The ECG signal was recorded through the CM5 lead, and EEG signals were recorded using unipolar leads F3/F4, C3/C4 and P3/P4, based on International System 10–20. Both signals were sampled to 250 Hz [13].

The signals were acquired during a test protocol, which consisted in 5 stages: rest, mental arithmetic + Stroop (first cognitive task), controlled breathing, mental arithmetic + Stroop (second cognitive task) and recovery. Each stage had a duration of 5 min, except rest and recovery stages which had a duration of 3 min (Fig. 1). All stages were conducted with the participants remaining seated, with their hands resting on knees, and their open eyes [13]. The participants provided their written consent to take part in the study.



Fig. 1. Test Protocol Diagram

The cognitive stimulus involved mental arithmetic tasks through a combination of sums of random numbers alternating with the Stroop test, where the participants identified the hue of the word as opposed to the actual word, each stimulus had a duration of 2 s and were alternated randomly. The stressor stimulus was presented to the volunteers through an interface, which displayed the math operation or word in the center of screen, with the possible answers in each lower corner; the participants provided their answer by closing their left or right hand corresponding to the considered correct answer. For the controlled breathing, a frequency of 6 breaths per minute was used, where the starts and end of inhalation and exhalation was indicated to participants [13].

All volunteers were non-smokers, without obesity, and without a permanent medical prescription. Moreover, they had been requested not ingesting alcohol or coffee 12 h prior to the recording. In addition, for female participants, the recordings were taken during the follicular phase of the menstrual cycle to reduce the variations caused by hormonal changes.

2.2 Data Preprocessing

The EEG and ECG signals were preprocessed using MATLAB® software. The RR intervals of the ECG signal were obtained [14], and subsequently, an adaptive filter [15] was used for artifact correction, which was manually verified for accuracy. For the EEG signal, the independent component analysis (ICA) method [16] was employed for artifact removal, and the signal was filtered using a FIR bandpass filter with a Kaiser window, ranging from 1 to 100 Hz.

From the artifact-free signals, the coupling between the RR interval signal and EEG signal was performed. The RR interval signal was resampled at 4 Hz, and subsequently, the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) technique was applied. This approach offers a precise reconstruction of the original signal and improved signal separation by selecting only the necessary intrinsic mode functions (IMFs) that do not contain Gaussian noise caused by the ensemble empirical mode decomposition (EEMD) technique, which is used within the CEEMDAN method [17]. The IMFs containing their power spectrum within the high-frequency (HF) and low-frequency (LF) ranges, corresponding to the range of 0.15–0.4 Hz and 0.04–0.15 Hz, respectively, are selected. Subsequently, the selected IMFs for each band, LF and HF, are considered to reconstruct the signal.

For the coupling of the EEG signal, the power spectral density (PSD) is calculated for each maneuver in 2-s windows, with a sliding window of 0.25 s. The area under the corresponding bands is obtained for each window: δ (1–4 Hz), θ (4–8 Hz), α (8–12 Hz), β (12–30 Hz), and γ (30–70 Hz), and with this process four values per second are obtained [18].

2.3 Transfer Entropy (TE)

TE determines how much the history of one process, X, influences the transition probabilities of another process, Y, and quantifies this relationship, and it is given by Eq. (1), where $\mathbf{y}_t^{d_y}$ and $\mathbf{x}_t^{d_x}$ are dimensional delay vectors, d_y and d_x are dimensions in y and x, and t is a discrete time index [19].

$$ET(X \rightarrow Y) = \sum_{\mathbf{y}_{t+1}, \mathbf{y}_t^{d_y}, \mathbf{x}_t^{d_x}} p(\mathbf{y}_{t+1}, \mathbf{y}_t^{d_y}, \mathbf{x}_t^{d_x}) \log \left(\frac{p(\mathbf{y}_{t+1}, \mathbf{y}_t^{d_y}, \mathbf{x}_t^{d_x})}{p(\mathbf{y}_{t+1} | \mathbf{y}_t^{d_y})} \right) \quad (1)$$

Before computing TE, it is essential to reconstruct the state space of the raw data. This was done using Takens' embedded delay to map the scalar time series into trajectories in a possibly high-dimensional state space. The mapping uses delay coordinates to form a collection of vectors or points within a space of higher dimensions according to Eq. (2) [19].

$$\mathbf{x}_t^d = (x(t), x(t - \tau), x(t - 2\tau), \dots, x(t - (d - 1)\tau)) \quad (2)$$

The success of this process relies on the characteristics of the parameters associated with it dimension d and delay τ . To estimate these parameters, the false nearest neighbors (FNN) criterion [20] and the average mutual information (AMI) [21] are used, respectively.

After, the joint probability in Eq. (1) is estimated using an adaptive partition with the Darbellay-Vajda (DV) algorithm. The DV method allows for adaptively adjusting the bin size according to the uniform distribution of data within a specific subregion of the data space [22].

Surrogate data generated through random permutation were used to confirm that the coupling determined by TE originates from the signal's dynamics [23]. The null hypothesis is the absence of transfer information, which is rejected if the original value of TE is greater than the 95th percentile of the respective surrogate data [22].

2.4 Statistics Analysis

The Lilliefors test was used to verify if the ET results follows a normal distribution. Comparisons between maneuvers and acquisition channels were performed using a one-way analysis of variance (ANOVA 1) or Kruskal-Wallis test, depending on the data distribution. Comparisons were made for each combination of selected IMF's corresponding to LF and HF with each of the EEG signal rhythms, including δ , θ , α , β and γ . Additionally, a post hoc test based on the Bonferroni test was conducted. Significant differences were considered at a $p < 0.05$.

3 Results and Discussions

Results showed in Fig. 2 to Fig. 3 consider only the interactions that passed the surrogate data test in at least 8 signals for each maneuver, where none of the interactions were rejected, since although some subjects were rejected, all interactions met the condition to be considered.

Figure 2 shows the significant TE results for each channel between rest, first cognitive task, controlled breathing and second cognitive task maneuvers, in the direction of brain rhythms to IMF_{LF} .

As observed in the comparison between rest and the first cognitive task, the predominant interactions are $\beta \rightarrow IMF_{LF}$ and $\delta \rightarrow IMF_{LF}$, decreasing during the first cognitive task. This means that information flux from brain to sympathetic cardiac activation diminished, this could be associated to the increase in β and δ during cognitive tasks [7]. In the second cognitive task there are also a significant decrease in $\delta \rightarrow IMF_{LF}$ in the central and right frontal channels with respect to rest stage. The p -values in the comparison $\delta \rightarrow IMF_{LF}$ vs rest were higher in the first cognitive task in comparison with the second cognitive task, suggesting that there was a greater decoupling in the first cognitive task compared to the second. This change in p -value could be influenced by the controlled breathing stage.

On the other hand, the interactions $\beta \rightarrow IMF_{LF}$, $\delta \rightarrow IMF_{LF}$ and $\gamma \rightarrow IMF_{LF}$ increase during controlled breathing in comparison with the first cognitive task at least in three channels. Meanwhile, in the comparison between controlled breathing and the second cognitive task, the predominant interaction is $\beta \rightarrow IMF_{LF}$ that decrease during the second cognitive task.

The significant differences of TE in the direction of brain rhythms to IMF_{HF} are showed in Fig. 3. As can be observed, there is only one interaction in the comparison between rest and first cognitive task. This could be because during cognitive task is expected a more sympathetic activation and a less vagal participation. In the comparison between first cognitive task and controlled breathing, interactions $\gamma \rightarrow IMF_{HF}$ and $\alpha \rightarrow IMF_{HF}$ predominate, showing an increase during controlled breathing, while the interactions $\alpha \rightarrow IMF_{HF}$ and $\theta \rightarrow IMF_{HF}$ decrease during second cognitive task with respect to controlled breathing. This means that information flux from brain to vagal cardiac activation diminished, this could be associated to the increase in α during controlled breathing stage [24]. However, in comparison with the second cognitive task, it includes interactions with the θ index, which can also be linked to the increase of this index during the controlled breathing stage. Therefore, the comparison with the second cognitive task encompasses more elements that may be associated with controlled breathing, where a higher vagal activity is observed [25].

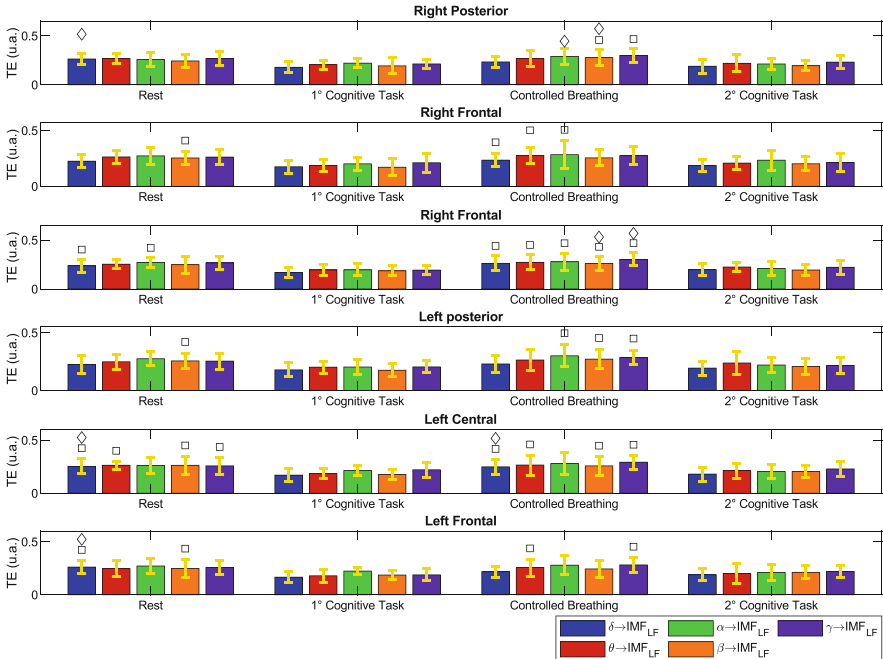


Fig. 2. Significant interactions from brain rhythms to IMF_{LF} , where \square vs-first cognitive task and \diamond vs-second cognitive task

Finally, in the comparison between rest and controlled breathing, it can be observed that all the interactions are directed towards IMF_{HF} , this result could be influenced by the respiratory sinus arrhythmia (RSA) since this phenomenon is highlight with slow respiratory frequency, suggesting an increase in vagal activity.

On the other hand, the number of interactions in the direction of IMF_{LF} to brain rhythms is reduced compared to the interactions from brain rhythms to IMF_{LF} , since the former represent 10.52% of the total significant interactions, and the latter 48.68% of the total significant interactions. In contrast, interactions in the direction of IMF_{HF} to brain rhythms correspond to 11.84% of the total significant interactions, while those directed from IMF_{LF} , represent 10.52%.

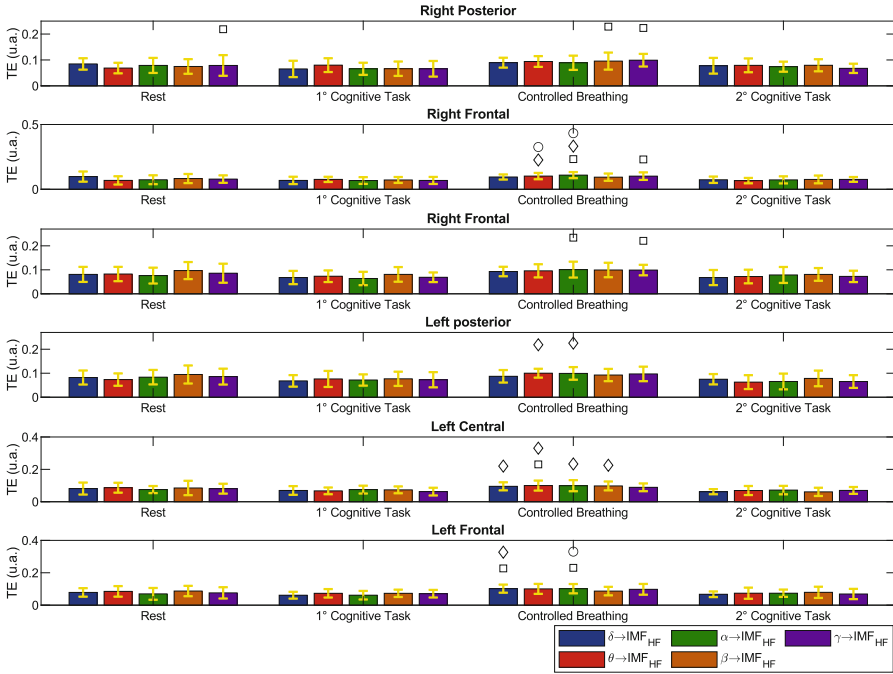


Fig. 3. Significant interactions from brain rhythms to IMF_{LF} , where \square vs-1st cognitive task, \diamond vs-2nd cognitive task and \circ vs-rest

4 Conclusions

In this study, the flow and interactions of the cerebral and cardiovascular systems were analyzed when applying cognitive tasks, and whether these interactions are affected by applying controlled breathing before a cognitive task, using TE. Significant differences were found in the test protocol. However, no significant differences were observed between the tasks. Nonetheless, it was observed that when controlled breathing is applied before a cognitive task, the interactions in the second cognitive task change compared to the first cognitive task. In the second cognitive task, a vagal influence can be observed, which is not present in the first cognitive task. Additionally, there is a greater number of interactions directed from the heart, suggesting that controlled breathing did influence the cognitive task. A controlled breathing training could help to increase the effect of this strategy to manage the physiological responses to stress. As a future perspective, this breathing training could be incorporated in a new experimental protocol. Furthermore, it is important to isolate external stimuli as much as possible to avoid interference with the measurements and potential alteration of the interactions.

References

1. Chu, B., Marwaha, K., Sanvictores, T.: *Physiology, Stress Reaction*. StatPearls Publishing, Tampa (2022)
2. Huang, C.-J.: Cardiovascular reactivity, stress, and physical activity. *Front. Physiol.* **4**, 314 (2013)
3. Burrage, E.: Cerebrovascular dysfunction with stress and depression. *Brain Circ.* **4**, 43–53 (2018)
4. Ma, X.: The effect of diaphragmatic breathing on attention, negative affect and stress in healthy adults. *Front. Psychol.* **8**, 874 (2017)
5. Sutarto, A.P.: Resonant breathing biofeedback training for stress reduction among manufacturing operators. *Int. J. Occup. Saf. Ergon.* **18**, 549–561 (2012)
6. Wang, X.: Cerebral and neural regulation of cardiovascular activity during mental stress. *Biomed. Eng. OnLine* **15**, 160 (2016)
7. Subhani, A.R.: Difference in brain dynamics during arithmetic task performed in stress and control conditions. In *IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES) 2016*, pp. 695–698. IEEE (2016)
8. Lizier, J.T.: Local information transfer as a spatiotemporal filter for complex systems. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **77**, 1–11 (2008)
9. Zanetti, M.: Information dynamics of the brain, cardiovascular and respiratory network during different levels of mental stress. *Entropy* **21**, 275 (2019)
10. Keshmiri, S.: Stress changes the resting-state cortical flow of information from distributed to frontally directed patterns. *Biology* **9**, 236 (2020)
11. Mueller, V.: The stroop competition: a social-evaluative stroop test for acute stress induction. In: *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, pp. 1–4 (2022)
12. Wang, J.: Perfusion functional MRI reveals cerebral blood flow pattern under psychological stress. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 17804–17809 (2005)
13. Rojas Sánchez, D.L.: *Evaluación de la Respuesta Cardíaca y Cerebral ante Estímulos Estresantes Cognitivos y su Regulación con Biofeedback*. Undergraduate thesis, Universidad Autónoma de San Luis Potosí, San Luis Potosí, S.L.P. (2018)
14. Pan, J., Tompkins, W.J.: A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.* **32**, 230–236 (1985)
15. Wessel, N.: Nonlinear analysis of complex phenomena in cardiological data. *Herzschrittmachertherapie Elektrophysiologie* **11**, 159–173 (2000)
16. Gabard-Durnam, L.J.: The Harvard Automated Processing Pipeline for Electroencephalography (HAPPE): standardized processing software for developmental and high-artifact data. *Front. Neurosci.* **12**, 97 (2018)
17. Torres, M.E.: A complete ensemble empirical mode decomposition with adaptive noise. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2011*, pp. 4144–4147. IEEE (2011)
18. Catrambone, V.: Functional brain-heart interplay extends to the multifractal domain. *Philos. Trans. A Math. Phys. Eng. Sci.* **379**, 1–16 (2021)
19. Lindner, M.: TRENTOOL: a Matlab open source toolbox to analyse information flow in time series data with transfer entropy. *BMC Neurosci.* **12**, 119 (2011)
20. Kennel, M.B.: Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. At. Mol. Opt. Phys.* **45**, 3403–3411 (1992)
21. Wallot, S.: Calculation of Average Mutual Information (AMI) and False-Nearest Neighbors (FNN) for the estimation of embedding parameters of multidimensional time series in Matlab. *Front. Psychol.* **9**, 1–10 (2018)

22. Lee, J.: Transfer entropy estimation and directional coupling change detection in biomedical time series. *Biomed. Eng. OnLine* **11**, 19 (2012)
23. Lancaster, G.: Surrogate data for hypothesis testing of physical systems. *Phys. Rep.* **748**, 1–60 (2018)
24. Komori, T.: Extreme prolongation of expiration breathing: effects on electroencephalogram and autonomic nervous function. *Ment. Illn.* **10**, 7881 (2018)
25. Birdee, G., et al.: Slow breathing for reducing stress: the effect of extending exhale, *Complement. Ther. Med.* **73**, 1–21 (2023)

**IOT in Health, Bioinstrumentation,
Biosensors, Bioimpedance
and Micro/Nano Technologies, BCI
Systems, Wearable Systems**



Automatic Knee Lock-Release System for a Knee-Ankle-Foot Orthosis (KAFO) with IOT Monitoring System in Real Time

Octavio Diaz-Hernandez¹ (✉) , Alejandra Santos-Borráz¹ ,
Antonio Gómez Ruiz² , Cynthia Fernanda Carrillo Gómez¹ ,
and Ileri Gisel Rosas Arellanes¹ 

¹ Escuela Nacional de Estudios Superiores Unidad Juriquilla, Universidad Nacional Autónoma de México, 76230 Juriquilla, Querétaro, Mexico
octavio.diaz@unam.mx

² Programa de Ingeniería Biomédica, Facultad de Ingeniería, Universidad Anáhuac, 76246 El Marqués, Querétaro, Mexico

Abstract. Gait disorder have several causes such as neurological problems, aging, or orthopedic problems. A gait disorder affects directly on the lifestyle of a person and one of the options is the use of orthoses when is prescribed, specifically, the Knee-Ankle-Foot Orthosis (KAFO) is commonly used for patients that suffer from quadriceps weakness, poliomyelitis, leg contractions, leg deformity or leg fracture, among others. In our previous work, we proposed a KAFO with an automated knee lock-release system and now, we add a new feature that implements a connection of the KAFO with the cloud to facilitate the remote monitoring. In this report we show the biomechatronic system connected to the cloud and the trials.

Keywords: KAFO · Orthosis · IoT Monitoring · Rehabilitation · gait cycle

1 Introduction

Lower limb disorders and gait disorders affect the person's lifestyle, for example they can suffer falls, joint problems, back problems, and restrict their independence [2, 5]. Therefore, people with gait problems may need orthoses to compensate some physical functions that are lost temporarily or permanently [7]. In this work, we are adding functions to a previous development [1], which is a KAFO prescribed in patients with quadriceps weakness, poliomyelitis, leg deformity or leg fracture. In our previous device there is an automatic knee lock-release system which allows flexion during gait, but we decided to connect our KAFO to the web server with the purpose of remote monitoring and medical assessment in real time, by using Internet of things (IoT) [8, 9].

Related works are few, for example, in 2021, Chanchotisien's and co-workers developed an Ankle Foot Orthoses (AFO) with IoT to keep a track of the patient recovery phase [2]. With this device, the physician can see the patient gait pattern, using sensors to capture the angular velocity and acceleration, other to capture the loads in the foot,

and one more to detect the temperature and humidity [2]. In Ab Wahid's [3] research, they developed an AFO, for rehabilitation within the hospital and the purpose was to change the position of the foot of the patient while the patient was immobilized in the hospital bed to avoid muscle atrophy and muscle shortness [3]. With this technology, the patient can move the foot using a mobile app that they developed, that is linked to the device, to prevent muscle atrophy and shortness in the patient. Finally, in Chetan's [4] approach, the device is a knee brace that helps elder people to sit or to stand up, the person can control the device using a mobile app, and select either assistance to sit down, or to stand up, but the data is not sent to a web server, nor can be considered as a clinic orthosis.

For this reason, we propose to use a Knee-Ankle-Foot Orthosis with connection to the web and demonstrate its advantages. For this device, the advantages would be the possibility to visualize data in graphs in real time (minimum time delay), store data for subsequent clinical assessment or data analytics (e.g., deep learning or machine learning) and notifications for the user and the medical provider, given than compliance to the treatment enhances the improvement of the pathological condition. We believe that these innovative features improve the functionality of the orthosis improving the life quality of the patient.

2 Methodology

2.1 Materials and Methods

For this project, we continue using the KAFO developed in previous work, which followed all the proper steps to be denominated a clinical orthosis [1], and we reprogrammed the system capable of release the knee articulation using an inertial sensor (see Fig. 1) to a Wi-Fi capable board.

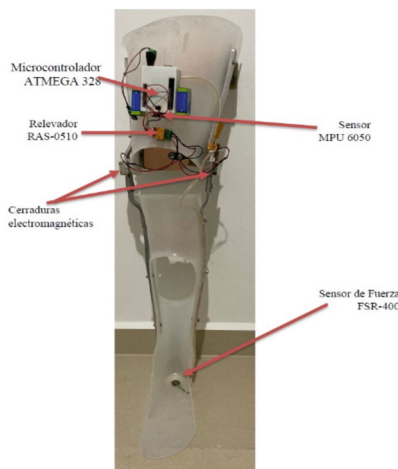


Fig. 1. Automated KAFO orthoses approach [1]

For this system modification, we proposed a ESP8266 board, and it was redesigned the housing of the electronic components (see Fig. 2).

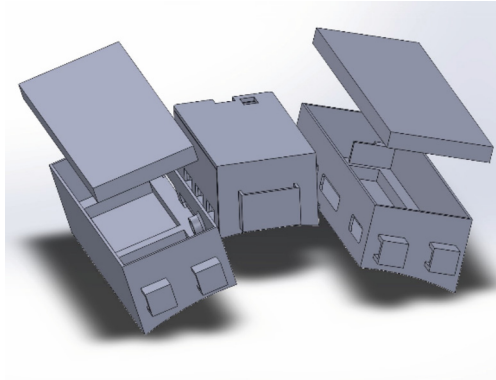


Fig. 2. Case assembly visualization in SolidWorks (2020)

2.2 Circuit Description

The ESP8266 board works with 5 V and required a voltage regulator, the inertial sensor (MPU6050) is connected to I2C pins. For the power control, it was used a transistor (2N222) to send a signal to relay, which controls the electromagnetic lock (release of the knee joint). In the heel of the KAFO it was placed a pressure sensor, so it can detect when the patient steps on the floor. The ESP8266 board needs two signals at the same time so it can send a signal to activate the electromechanical lock. One signal is the angle obtained through the gyroscope, and the second signal comes from the pressure sensor.

Once the electromechanical lock has been activated, the knee joint is unlocked, allowing the patient to bend the leg, and when the lock is off, it gets locked, keeping the leg straight. In the Fig. 3 we can see how our circuit is connected and organized.

2.3 Arduino Program

The board also uses Arduino IDE, first all the sensor values are obtained and stored, for example, the angular position of the gyroscope, and the pressure in the heel, and with both variables, decide whether send a signal to the mechanism or not. The flowchart in Fig. 4 shows the implemented logic, where “Gy” stands for the incoming values from the Gyroscope in the Y-axis, and “PS” stands for the pressure value read from the pressure sensor, which range of values goes from 0 to 100.

The loop in the code is measuring the incoming values from the Gyroscope and from the pressure sensor. Once it gets the data, it enters in the conditional statements, checking whether they are true or not, and depending on that, what part of the code should run next. When the patient is standing, the sensor is at 175° approximately, and when the

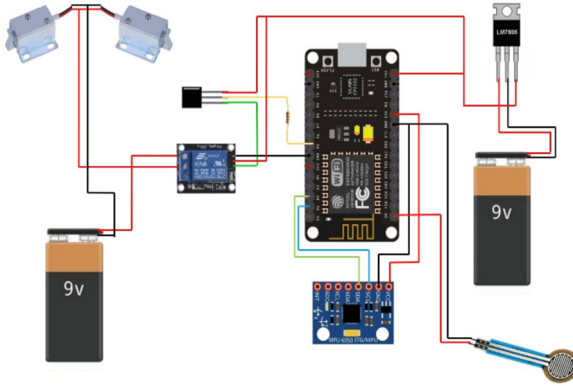


Fig. 3. Circuit design showing all the used components.

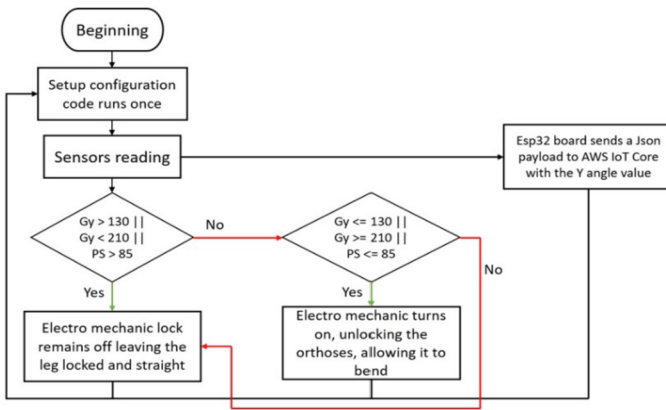


Fig. 4. Flowchart showing the logic implemented in the Arduino code.

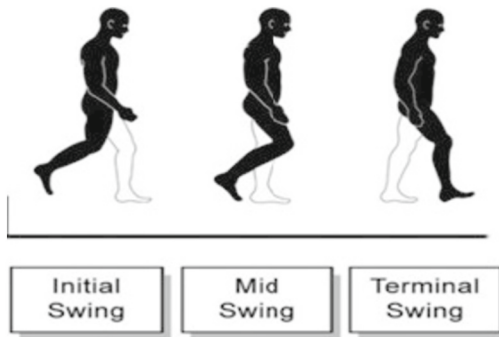


Fig. 5. Swing phase in the gait cycle, referring to the black leg [6].

patient is in the swing phase from the gait cycle, the sensor will get values higher than 210° in the Y axis (in the Fig. 5 we can see the swing phase in the gait cycle).

In the Fig. 6 below, shows when the patient is standing, the gyroscope value in the Y-axis is of 175° . The interval where the leg needs to be straight is 175° – 210° in the extension, and 130° – 175° in the flexion.

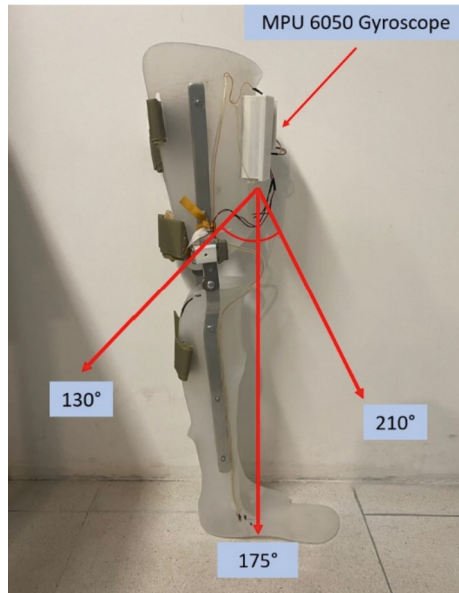


Fig. 6. Y-axis range value where the electro mechanic lock must remain off.

This means that in this stage, the patient needs to bend its leg so it can keep walking naturally. The board will now check on the other variable, the pressure sensor, and see what values it is receiving. Since the leg of the patient is in the swing phase at this point, the pressure sensor will not receive a signal, sending a zero value to the board. The board will be waiting for these two conditions to occur, so it can send a signal from the D5 pin so it can unlock the electromechanical lock.

While this is happening, the code also connected the board to Amazon Web Services (AWS) IoT Core via the MQTT protocol, in the set-up part of the code. The code stores the value from the Y axis obtained from the gyroscope and attached it to a Json payload with the time variable, measured using the Millis function. Then the Json payload is sent to the IoT Core so it can be stored in the database and graphed.

2.4 IoT Monitoring System

The Fig. 7 shows the different stages in the IoT monitoring system and the tools are used from the Amazon Web Services (AWS) IoT Core.

As shown, the process starts when the ESP8266 board sends the data to the AWS IoT Core through a Json payload which has the Y value and the time stored in it, using the

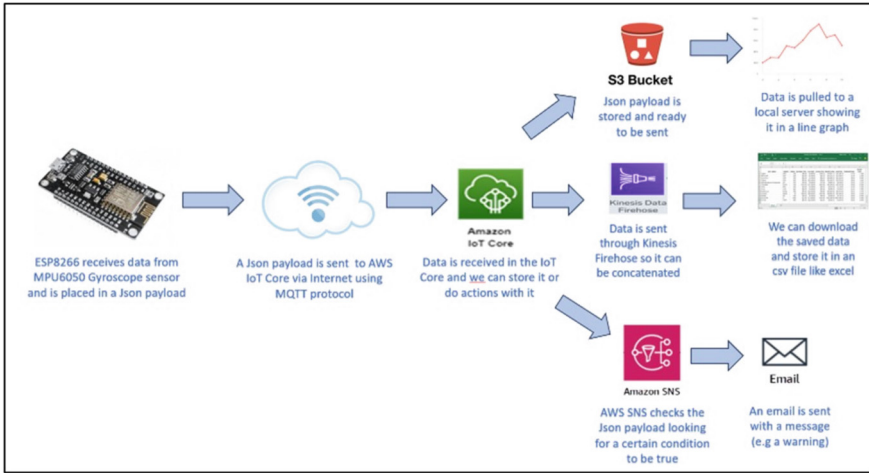


Fig. 7. Diagram showing the stages in the IoT monitoring system.

MQTT communication protocol. The board sends 100 Json payloads for every second. In the Arduino program we used the libraries “ESP8266WIFI”, “WifiClientSecure”, “PubSubClient”, and “ArduinoJson” to be able to connect to the Amazon Web Services (AWS) IoT core and send the data.

Once the payload is received in the IoT Core, there will be three actions waiting to occur. On one action, the data will be sent to the S3 bucket, and it will be stored in a key file. The disadvantage of this service is that the data will be overwritten, and we will only see the last data received. Although this is useful for our static server which will be pulling the data constantly and refreshing the graph plotting the new data.

On the second action, the data will be sent through the Kinesis Firehose action. With this service, the data will be able to concatenate in a file. The resulting file will be stored in the S3 bucket with all the data stored there unlike the key file created in the first action which only stores the last data received. This resulting file from the Kinesis Firehose can be downloaded as a text file so we can analyze the data.

On the third and last action, we activated the AWS Simple Notification Service (AWS SNS). This function will be waiting for a variable to trigger the service, and once the service is triggered, it will send an email with a message that we can specify. For this prototype, we sent an email indicating that the patient is using the orthoses, as we can see in the Fig. 8.

For this prototype, the code will send a signal when it first connects to the IoT Core, and when this happens the email will be sent to any email account that we want to specify in the AWS SNS.

2.5 Testing

After the modifications to the instrumented KAFO, the code, and configure the AWS server, the test with the orthoses were run. In the test, the orthoses send 3500 Json payloads, which we stored in the server. As we said previously, we were sending 100

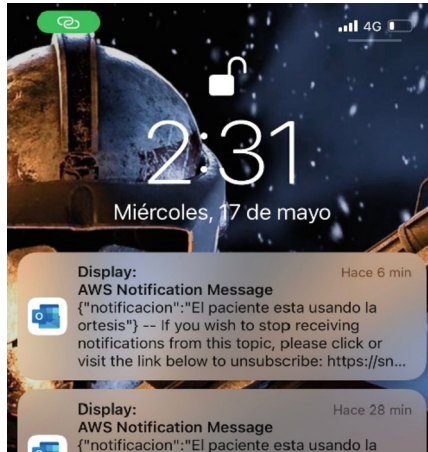


Fig. 8. Screenshot from the cellphone main screen with the email notification telling us that the patient is using the orthoses.

Json payloads for every second, meaning that the test lasted 35 s. In this test the prototype could send data, store it (for further analyses or comparison), graph it in real time, and send the email notification (see Fig. 8). In the Fig. 9 a user is wearing the orthosis and evaluating it walking straight on a line around the laboratory.

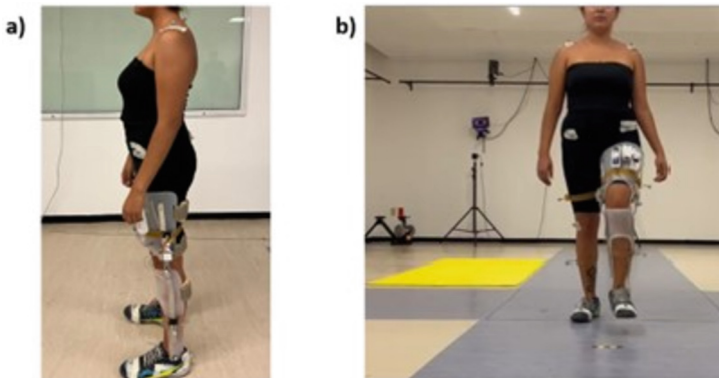


Fig. 9. Testing of the KAFO orthosis. a) standing up, and b) in the swing phase of the gait cycle.

3 Results

From the gait cycle test, we were able to graph all the gathered data, specifically in two types of graphs. One graph is the one made on real time while doing the test and the other we do it with all the data stored in the AWS s3 bucket. For the second graph we downloaded the data and graph it in excel. In the Fig. 10 we can see the comparison between both graphs on the second interval 27–29 where we can realize that they are slightly different. On this time interval we can see that both graphs are almost the same, except for some parts where we lost data (e.g., second 27.6–27.8 where there is a flat line), but as we can see from the graph made from the data stored on the cloud (green one), the data was in fact sent by the device and stored, the problem comes with the server which did not plot the whole data on real time, having this data loss.

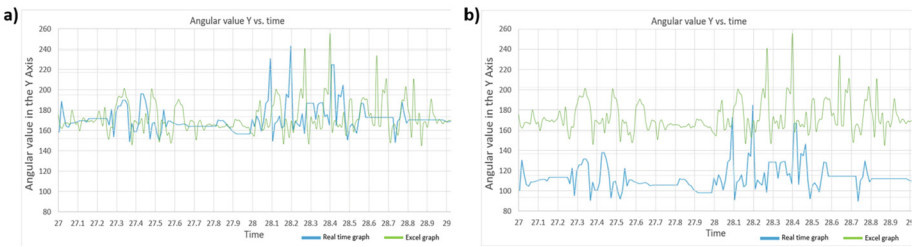


Fig. 10. Comparison graph between the made on real time (blue one) and the one made with the data stored in AWS, downloaded, and graphed in Excel (green one). Both graph comparisons are in the same time interval (second 27 to 30), in a) both are in the same position so we can see the difference in the lines. In b) we displaced downward the real time graph, so now we can see them individually (Colour figure online).

In the Fig. 11, the interval between the second 34.2 to 34.7, in the blue line (the graph made on real time) we can see a straight line, and in the green line (the one made on excel) we can see movement of the patient. This means that in 0.5 s, the graph didn't plot the line showing the values, therefore we lost 50 values in that time interval. As we stated before, we can see from the graph obtained from the data stored in the cloud (AWS) that the data was stored in its entirety, however due to server conditions the graph made on real time is showing incomplete data. We think that the quality of the website that our team developed affected the performance, since it is not a professional server, it is more related to a Alpha testing (first phase of a software testing).

On the other hand, in the test we were able to see the orthoses functionality and see how the system is working during the gait cycle. The system allowed the user to unlock the knee during the gait cycle to be able to bend it and lock it when the user needed to keep the leg straight in the stance phase, independently to the missing data.



Fig. 11. As same in Fig. 10, we see the comparison between the 2 graphs made, one in real time and one downloading all the date stored in AWS. In this case, the Interval of the graphs is from the second 33 to 35(Colour figure online).

4 Conclusions

The main objective was to adapt a previously tested KAFO with an automated knee lock-release system to a remote monitoring system that can assess the patient's gait and if he stops using the orthosis. The lock-release systems use data gathered by a gyroscope sensor (MPU6050) and a pressure sensor. We observe that the instrumented KAFO still perform normally, and that the system was modified so it can send data to a database in real time using IoT, which give advantages e.g., data visualization in graphs, data stored for further analysis and simple notification system use, such as email messages when the orthosis is being used, so the physician or physical therapist can monitor the patient treatment. The loss of data visualization was occasional (up to 25% for the case of the Fig. 11) and the loss as we discussed in the results, was due to problems with the server and the quality of itself. For future work, remains to minimize data loss due to external factor such as internet speed or remote failures, that includes to have a custom server for data storage and improve our website where we will visualize the graphs obtained from the gait cycle.

Acknowledgements. Authors thanks the support to Escuela Nacional de Estudios Superiores Campus Juriquilla, of Universidad Nacional Autónoma de México.


We also thank Dr. José Rafael Alaniz Gómez Coordinator of the biomedical engineering career at Anáhuac Querétaro University.

References

1. Diaz-Hernández, O., Santos-Borráez, A., Gómez, C.F.C., Gisel, I., Arellanes, R.: Sistema de desbloqueo automático en una Órtesis de Rodilla-Tobillo-Pie (ORTP) durante la marcha. https://www.researchgate.net/publication/367350568_Sistema_de_desbloqueo_automatico_en_una_Ortesis_de_Rodilla-Tobillo-Pie_ORTP_durante_la_marcha
2. Chanchotisien, P., Ungkhara, T., Anuraklekha, W., Laitrakun, S.: An IoT-based controlled ankle motion boot for foot and ankle monitoring during the recovery phase. In: 2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Tele-communication Engineering, pp. 149–152. IEEE (2021)
3. Ab Wahid, A.M., Tardan, G., Pangesty, A.I., Rashid, H., Abdullah, A.H.: Development of ankle-foot orthosis with the integration of IoT Controller. *Int. J. Emerg. Technol. Adv. Eng.* **12**(5), 49–55 (2022)
4. Chetan, R., Avinash, N.J., Aditya, K., Gowri, M., Pranav, K.R.: Providing knee movement assistance using android and IOT. In: 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), pp. 140–145 IEEE (2021)
5. Pirker, W., Katzenschlager, R.: Gait disorders in adults and the elderly: a clinical guide. *Wien. Klin. Wochenschr.* **129**(3–4), 81–95 (2017)
6. Gait and balance academy. Understanding Phases of the Gait Cycle. *Protokinetics* (2018). <https://www.protokinetics.com/understanding-phases-of-the-gait-cycle/>
7. Sekiyama, M., Homma, K., Wakita, Y., Kajitani, I., Mori, I., Matsumoto, Y.: Investigation of IoT on ankle-foot orthosis by introducing an existing life-activity measurement system. In: 2021 IEEE International Conference on Advanced Robotics and its Social Impacts (ARSO), pp. 246–253. IEEE (2021)
8. Fan, Y.J., Yin, Y.H., Da Xu, L., Zeng, Y., Wu, F.: IoT-based smart rehabilitation system. *IEEE Trans. Industr. Inf.* **10**(2), 1568–1577 (2014)
9. Bisio, I., Delfino, A., Lavagetto, F., Sciarrone, A.: Enabling IoT for in-home rehabilitation: accelerometer signals classification methods for activity and movement recognition. *IEEE Internet Things J.* **4**(1), 135–146 (2016)



Portable Chest-Worn Device for Simultaneous Detection of Phonocardiogram and Seismocardiogram Using Piezoelectric Films Sensors

Pamela Salas Canales^(✉) , José Mejía Muñoz , and Rafael Gonzalez-Landaeta 

BIOCIM Research Group, Universidad Autónoma de Ciudad Juárez, Ciudad Juárez, Chihuahua 32310, México
a1228162@alumnos.uacj.mx

Abstract. The simultaneous detection of different heart signals provides information about cardiovascular health but typically involves attaching multiple sensors to the body, which can lead to discomfort during long-term measurements. In this work, we propose a different approach to simultaneously detect the phonocardiogram (PCG) and seismocardiogram (SCG). It relies on using two piezoelectric films implemented in a single portable chest-worn device, reducing the number of devices attached to the body. Besides, since these sensors do not require any power supply, the power consumption is considerably reduced. This innovative approach offers a potentially more comfortable alternative for long-term applications, reducing user discomfort. Because both signals have different frequency components, both sensors were strategically coupled to a stethoscope membrane using a custom-built case. The device was tested on a healthy volunteer. It was able to simultaneously detect the PCG with an SNR higher than 34 dB and the SCG with an SNR higher than 62 dB. The electronic circuitry was battery-supplied, and its current consumption was lower than 65 μ A. The possibility of establishing a client/server connection via sockets was also demonstrated, which allows sending of data to the cloud/fog where intelligence algorithms could be implemented.

Keywords: Chest-worn device · Phonocardiogram · Seismocardiogram

1 Introduction

The genesis of the phonocardiogram (PCG) and the seismocardiogram (SCG) comes from the precordial vibrations, which have infrasound and audible components [1]. Regarding the PCG, this signal can offer insights into both the cardiac valve dynamics and arterial blood pressure (ABP) by examining different aspects of the waveform [2, 3]. The recording of heart sounds is commonly done by using microphones of different technologies [4]. However, alternatives based on accelerometers [5], Lead Zirconate Titanate (PZT) [6], and Polyvinylidene Difluoride (PVDF) piezoelectric sensors [7] have also been proposed. On the other hand, the SCG is the recording of the vibrations of the

body produced by the heartbeat [8] and provides information on the mechanical activity of the heart. SCG can be detected by piezoelectric accelerometers, MEMS accelerometers, triaxial gyroscopes, and laser vibrometers [9]. Systems based on accelerometers, gyroscopes, and microphones must deal with noise problems and high current consumption since they have embedded preprocessing stages, including amplifiers and resistors. Even so, the small size of current microphones and accelerometers makes them ideal for implementation in wearable systems with a low form factor, which results in a more comfortable system for long-term measurements.

The simultaneous detection of cardiac signals helps to estimate other physiological parameters [10]. Assessing the temporal relationship between a proximal signal and a distal signal makes it possible to estimate the systolic blood pressure (SBP) without using a cuff. The pulse arrival time (PAT), pulse transit time (PTT), and pulse wave velocity (PWV) [11] are commonly analyzed using electrocardiogram (ECG), photoplethysmogram (PPG) [12, 13], seismocardiogram (SCG) [14], ballistocardiogram (BCG) [15] and phonocardiogram [16] signals and have been used to develop portable devices for cuffless and continuous detection of ABP. The problem with these approaches is that some systems require many sensors attached to the body, which causes discomfort to the subject in long-term monitoring. To tackle this, some devices have been proposed that use the correlation between the PTT and the systolic blood pressure for continuous ABP measurement using a minimum number of sensors. Devices that detect ECG and PPG signals at the bicep and ear or only at the bicep [17], a watch-based device that measures single-lead ECG, tri-axial SCG, and multi-wavelength PPG [18], and a single chest-worn device that measures the PPT using SCG, and PPG [19], have been proposed.

This work proposes a new approach to simultaneously detect PCG and SCG signals from a single point on the thorax. A similar proposal has been presented in [20], using a single CM-01B contact microphone to simultaneously detect respiration, SCG, and heart sounds. Each signal was extracted from the raw data by filtering, which can lead to signal distortion. The CM-01B sensor has an embedded field effect transistor (FET), a source of noise that degrades the Signal-to-Noise Ratio (SNR) of the signals and consumes 100 μ A. Since the signal must be amplified, the total consumption of this system increases, which reduces the system's autonomy. A more sophisticated sensor based on an encapsulated accelerometer contact microphone that uses nano-gap transducers has also been proposed for detecting cardiopulmonary signals [5]. However, this sensor requires a high-precision fabrication technique. In this work, two commercial shielded PVDF (Polyvinylidene fluoride) piezoelectric film sensors are used, which do not require a power supply and reduce the contribution of electromagnetic interference. The two piezoelectric film sensors are attached to a stethoscope membrane using a custom-built case for portability. To the best of our knowledge, using these kinds of sensors to simultaneously detect the PCG and the SCG is a novelty approach that has yet to be proposed. So, the goal of this work is twofold: a new approach for sensing the infra-sound and audible precordial vibrations, and to design a low-power electronic circuit for PCG and SCG detection with high SNR to simplify (in the future) the signal processing algorithms for estimating others cardiovascular parameters. The findings from this study serve as the foundation for creating a portable system that uses the temporal correlation between PCG and SCG signals to estimate the SBP at a single point on the chest.

2 Materials and Methods

2.1 Sensor Coupling

Two sensors were used to simultaneously detect PCG and SCG at a single point on the thorax: one was exclusively for PCG, while the other was for SCG. For each signal, the sensor used was the piezoelectric film sensor SDT1-028K from TE Connectivity, which has a capacitance of 2.78 nF, a minimum electrical impedance of 1 M Ω , and a maximum of 10 M Ω . When it is not attached to any surface, its resonance frequency is up to 10 MHz [21], and the frequency response can be limited by the low-pass response of the signal conditioning stages [7]. In this scenario, the frequency response can be adjusted to detect the frequency components of the S1 and S2 sounds of the PCG, which are in a bandwidth of 100 to 200 Hz and 50 to 250 Hz, respectively. If the sensor is attached to a surface, the bandwidth of the system is limited because the surface acts as a mechanical low-pass filter. Even so, the frequency response is still sufficient to detect the SCG [7], whose bandwidth ranges from 0 to 25 Hz. Both sensors were coupled to a Littmann® stethoscope membrane using a custom-built case (Fig. 1a), which allows the PCG sensor to be placed without direct contact with the membrane and the SCG sensor to be in direct contact with the stethoscope membrane. The case was designed in SolidWorks and 3D printed with resin. The coupling of the sensors to the stethoscope membrane is shown in Fig. 1b.

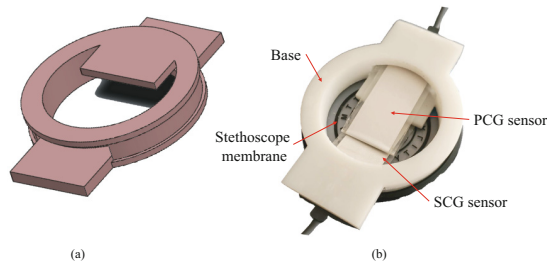


Fig. 1. Sensor coupling: (a) SolidWorks design, (b) 3D printed prototype.

2.2 Signal Acquisition

The circuits for the simultaneous detection of PCG and SCG signals consist of a charge amplifier that transforms charge variations from the sensor to an output voltage, an amplifier that adjusts the signal amplitude and the cutoff frequency, and a DC servo loop circuit to null DC components at the output.

The PCG detection circuit (Fig. 2a) was designed to work in a frequency range of 34 to 600 Hz, with a total gain of 55 dB, to observe as many frequency components of the S1 and S2 sounds as possible. The low-pass frequency response of the system was limited by the gain-bandwidth product of the amplifier (OA2) at a gain of 40 dB. The

high-pass response was limited by the charge amplifier (OA1), which has a sensitivity of 2.12 mV/pC. The SCG detection circuit (Fig. 2b) works in a frequency range of 3.3 Hz to 34 Hz with a total gain of 35 dB. The low-pass frequency response of the system was limited by $R_Y C_4$, while the high-pass response was limited by the charge amplifier (OA4), which has a sensitivity of 2.12 mV/pC. Both detection circuits were implemented using the operational amplifier TLV2322 from Texas Instruments and were supplied by a single 3.7 V/1000 mAh Lithium-Ion battery.

The auscultation point selected was the pulmonary valve located in the second intercostal space to the left of the sternum. This auscultation point was chosen because, in previous tests, it was the point where a greater amplitude of both signals was obtained. The ECG (lead II) was also detected using the AD8232 Single Lead Heart Rate Monitor (Analog Devices). The signals were simultaneously acquired at 1 kSa/s using a National Instruments USB 6341 data acquisition system controlled by an algorithm developed in LabVIEW. The signals acquisition was performed following the CEI-2023–1-853 measurement protocol approved by the UACJ ethics committee. The protocol consisted of detecting the PCG, SCG, and ECG signals from a healthy adult volunteer who remained seated during the process (Fig. 2c) and gave his/her informed consent to participate in the procedure.

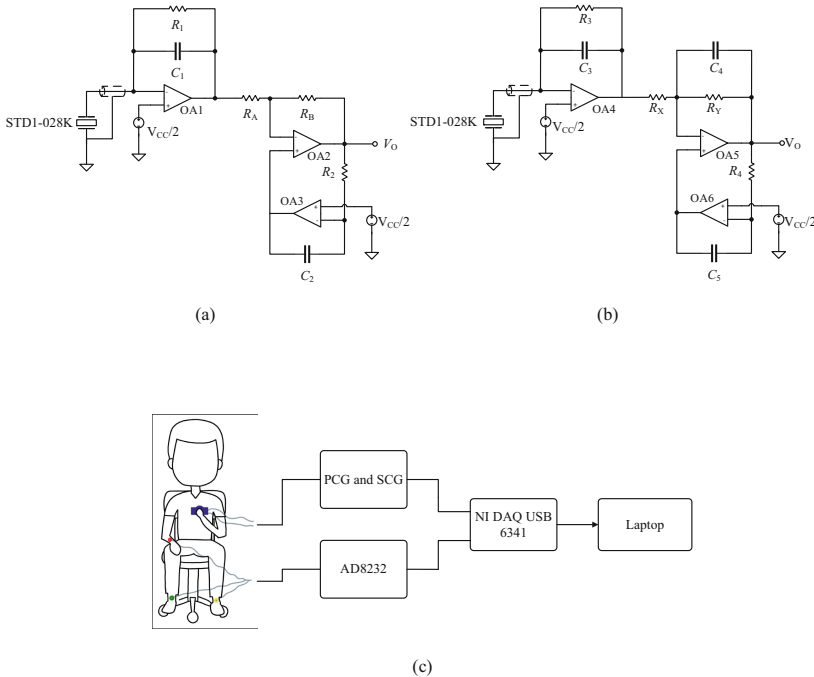


Fig. 2. (a) Designed circuit for PCG detection, (b) designed circuit for SCG detection, (c) Setup for detecting the PCG, SCG, and ECG simultaneously.

A server/client link via TCP/IP sockets connection was implemented to demonstrate the ability of the device to send the acquired signals to the cloud/fog. For this, the signals were acquired using the ESP32 microcontroller, and via TCP/IP sockets connection, the corresponding data to each signal was sent and saved in a local computer, as shown in Fig. 3.

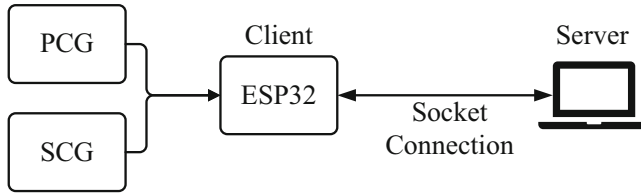


Fig. 3. Client/server connection.

3 Results

The voltage noise spectral density of the circuits of Fig. 2a and 2b are depicted in Fig. 4a and Fig. 4b, respectively. The output noise root mean square (RMS) voltage of the PCG circuit is about 1.4 mV, and of the SCG circuit is about 50 μ V. The current consumption of both circuits working together was lower than 65 μ A (240 μ W). With these characteristics, the system was able to detect the signals with an SNR higher than 34 dB for the PCG and higher than 62 dB for the SCG.

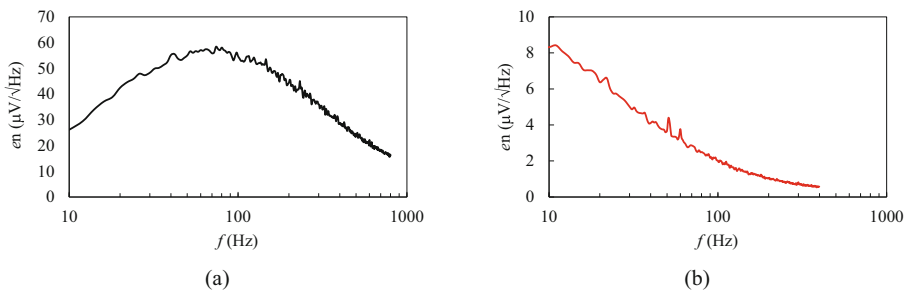


Fig. 4. Voltage noise spectral density of the circuit used for detecting the PCG (a) and the SCG (b).

Figure 5 shows the PCG and the SCG simultaneously detected and compared with the Lead II ECG as a reference. The S1 and S2 of the PCG were detected by the sensor that was not attached to the surface of the membrane. The sensor attached to the surface of the membrane detected the infrasound chest vibrations that give rise to SCG. When measuring with the SCG sensor, the motion artifacts can have a significant impact, which can cause signal distortion; this is because the sensor is located near the chest wall. To get

accurate results, it is crucial to ensure that the subject stays still during the measurement process. By analyzing the PCG and SCG signals displayed in Fig. 5, it becomes feasible to estimate the time correlation between S1 and S2 of the PCG and the main waves of the SCG, which could be an alternative method for estimating SBP using a single chest-worn device.

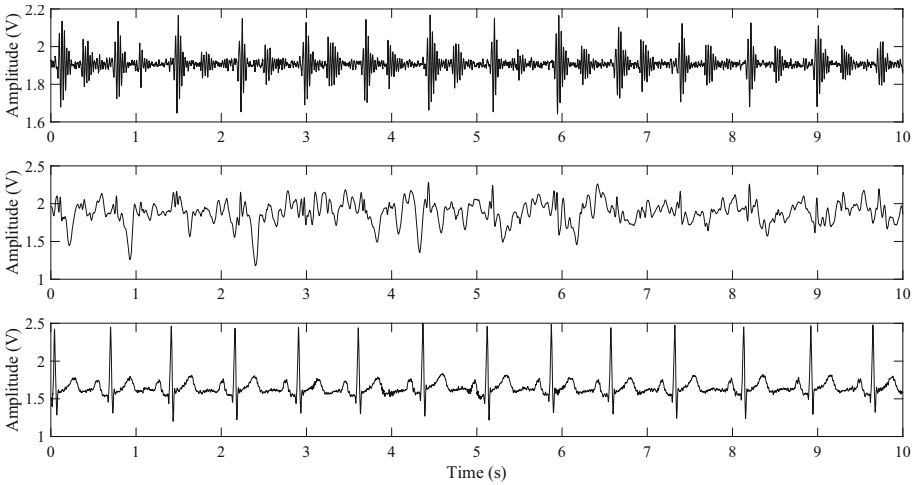


Fig. 5. Raw signals of the PCG (upper trace) and SCG (middle trace) detected with circuits shown in Fig. 2. ECG (bottom trace) detected with the AD8232. All signals are from the same subject and were acquired simultaneously.

Figure 6 shows the PCG and SCG signals acquired with the ESP32 microcontroller and sent via socket connection to a local computer. This demonstrates the possibility of sending information to the cloud/fog without degrading the signal quality. With this approach, artificial intelligence algorithms could extract signal features using fewer hardware resources, which reduces power consumption significantly.

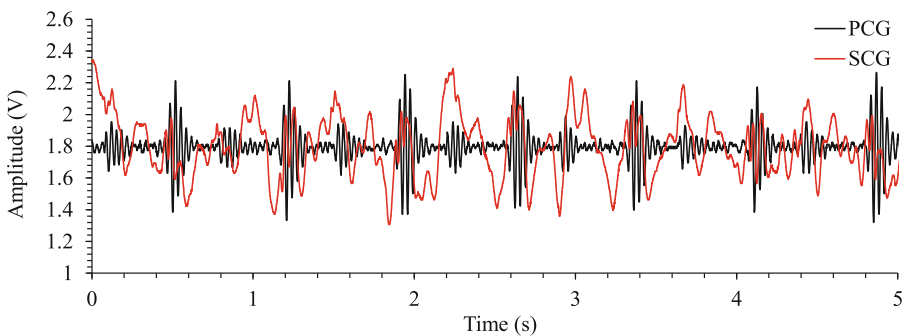


Fig. 6. FCG and SCG signals acquired using the microcontroller and sockets connection.

4 Conclusions

A different approach has been proposed to simultaneously detect the PCG and SCG, which was implemented in a portable and low-power chest-worn device. It relies on using two commercial piezoelectric films, reducing the number of devices attached to the body and the power consumption. The electronic circuitry was battery-supplied, and the total current consumption was lower than 65 μA , making it feasible for long-term measurements. Regarding the PCG, the SNR was higher than 34 dB, making it possible to detect the S1 and S2. For the SCG, the SNR was higher than 62 dB, allowing it to detect the main waves of the signal. The quality of the detected signals simplifies the processing algorithms to detect signal features, which helps estimate, for example, systolic blood pressure. The feasibility of using piezoelectric sensors to detect both signals was demonstrated. However, it must be considered that the coupling of each sensor on the thorax plays an important role due to the different frequency components of each signal. Using piezoelectric sensors helps reduce the power consumption that usually goes to sensors such as microphones and accelerometers. However, it should be noted that these sensors are more prone to movement artifacts. The results obtained in this study are the starting point for developing a portable system that makes it possible to estimate systolic blood pressure by taking measurements at a single point on the thorax.

References

1. Di Rienzo, M., et al.: Wearable seismocardiography: towards a beat-by-beat assessment of cardiac mechanics in ambulant subjects. *Auton. Neurosci.* **178**, 50–59 (2013)
2. Lim, K.H., et al.: Correlation of blood pressure and the ratio of S1 to S2 as measured by esophageal stethoscope and wireless bluetooth transmission. *Pak. J. Med. Sci.* **24**(9), 1023–1027 (2013)
3. Omari, T., Berekci-Reguig, F.: A new approach for blood pressure estimation based on phonocardiogram. *Biomed. Eng. Lett.* **9**, 395–406 (2019)
4. Shyam Kumar, P., Ramasamy, M., Varadan, V.K.: Evaluation of signal quality from a wearable phonocardiogram (PCG) device and personalized calibration. *Electronics* **11**(17), 2655 (2022)
5. Gupta, P., Moghimi, M.J., Jeong, Y., Gupta, D., Inan, O.T., Ayazi, F.: Precision wearable accelerometer contact microphones for longitudinal monitoring of mechano-acoustic cardiopulmonary signals. *NPJ Digit. Med.* **3**(1), 19 (2020)
6. Chen, Z., Chen, D., Xue, L., Chen, L.: A piezoelectric heart sound sensor for wearable healthcare monitoring devices. In: Mucchi, L., Hämäläinen, M., Jayousi, S., Morosi, S. (eds.) *BODYNETS 2019*. LNICS SITE, vol. 297, pp. 12–23. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-34833-5_2
7. Vazquez, K., Cota, J., Sifuentes, E., Gonzalez, R.: High signal-to-noise ratio phonocardiogram using a shielded PVDF film sensor. *IEEE Lat. Am. Trans.* **14**(3), 1139–1145 (2016)
8. Castiglioni, P., Faini, A., Parati, G., Di Rienzo, M.: Wearable seismocardiography. In: 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 3954–3957. IEEE (2007)
9. Taebi, A., Solar, B., Bomar, A., Sandler, R., Mansy, H.: Recent advances in seismocardiography. *Vibration* **2**(1), 64–86 (2019)
10. Soliman, M.M., Ganti, V.G., Inan, O.T.: Toward wearable estimation of tidal volume via electrocardiogram and seismocardiogram signals. *IEEE Sens. J.* **22**(18), 18093–18103 (2022)

11. Tabei, F., Gresham, J.M., Askarian, B., Jung, K., Chong, J.W.: Cuff-less blood pressure monitoring system using smartphones. *IEEE Access* **8**, 11534–11545 (2020)
12. Tjahjadi, H., Ramli, K.: Review of photoplethysmography based non-invasive continuous blood pressure methods. In: 2017 15th International Conference on Quality in Research (QiR): International Symposium on Electrical and Computer Engineering, pp. 173–178. IEEE, Nusa Dua, Bali, Indonesia (2017)
13. Vajanarat, T., Lek-uthai, A.: A comparison of cuff-less blood pressure estimation between pulse arrival time and pulse transit time using photoplethysmography. In: 2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), pp. 13–16. IEEE, Phuket, Thailand (2020)
14. Imtiaz, M.S., et al.: Correlation between seismocardiogram and systolic blood pressure. In: 2013 26th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), pp. 1–4. IEEE, Regina, SK, Canada (2013)
15. Kim, C.-S., Carek, A.M., Mukkamala, R., Inan, O.T., Hahn, J.-O.: Ballistocardiogram as proximal timing reference for pulse transit time measurement: potential for cuffless blood pressure monitoring. *IEEE Trans. Biomed. Eng.* **62**(11), 2657–2664 (2015)
16. Gonzalez-Landaeta, R., Ramirez, B., Mejia, J.: Estimation of systolic blood pressure by random forest using heart sounds and a ballistocardiogram. *Sci. Rep.* **12**, 17196 (2022)
17. Griggs, D., et al.: Design and development of continuous cuff-less blood pressure monitoring devices. In: 2016 IEEE SENSORS, pp.1–3. IEEE, Orlando, FL, USA (2016)
18. Ganti, V.G., Carek, A.M., Nevius, B.N., Heller, J.A., Etemadi, M., Inan, O.T.: Wearable cuff-less blood pressure estimation at home via pulse transit time. *IEEE J. Biomed. Health Inform.* **25**(6), 1926–1937 (2021)
19. Park, J., et al.: Cuffless and continuous blood pressure monitoring using a single chest-worn device. *IEEE Access* **7**, 135231–135246 (2019)
20. Bifulco, P., et al.: Monitoring of respiration, seismocardiogram and heart sounds by a PVDF piezo film sensor. *Measurement* **11**, 786–789 (2014)
21. TE Connectivity SDT Shielded Piezo Sensors. https://www.te.com/commerce/DocumentDelivery/DDEController?Action=showdoc&DocId=Data+Sheet%7FSDT_Shielded_Piezo_Sensors%7FA1%7Fpdf%7FEnglish%7FENG_DS_SDT_Shielded_Piezo_Sensors_A1.pdf%7FCAT-PFS0010. Accessed 10 July 2023



A Novel Intra-Abdominal Pressure Measurement Device: Design and In Vitro Characterization

Alejandra Soto¹(✉) , Ivan Luna¹ , José Camacho¹ , Oscar Núñez² ,
Néstor López³ , and Bersain Reyes¹ 

¹ School of Sciences, Autonomous University of San Luis Potosí,
78295 San Luis Potosí, San Luis Potosí, Mexico
A297684@alumnos.uaslp.mx

² Optical Communication Research Institute, Autonomous University of San Luis
Potosí, 78210 San Luis Potosí, San Luis Potosí, Mexico

³ GM Biomedical Services, 45215 Zapopan, Jalisco, Mexico

Abstract. Intra-abdominal pressure (IAP) measurement methods vary in mechanisms, risks, and associated costs. In this study, it was intended to design and test a novel IAP acquisition system based on a simulation of the Kron's method, as well as to compare it with lectures from a column graduated in cmH₂O, taken as reference. A total of 46 samples of IAP were simulated and acquired, through a bladder simulation system, by placing different levels of pressure in a manual press on a sealed bag of water which, in turn, transmitted a pressure to our device to be compared with the graduated column. The data analysis included linear regression and Bland-Altman, in which, a Pearson correlation coefficient $r = 0.99$, a bias of 0.16 cmH₂O and 95% limits of agreement between -0.444 and 0.758 cmH₂O were found. Also, the obtained results were compared with those of similar devices, denoting fairly close correlation coefficients, and a notable improvement in the bias obtained for our device. Even considering its limitations, the results seem promising and appear to suggest that the proposed device meets its main functionality and is strengthened by different available advantages compared to other devices currently in clinical use.

1 Introduction

The World Society of the Abdominal Compartment Syndrome (www.wsacs.org) defines intra-abdominal pressure (IAP) as the pressure within said cavity, expressed in mmHg, which must be measured without the intervention of

A. Soto—Data collection and statistical analysis.

I. Luna—Modeling of the device and *in vitro* simulation system.

J. Camacho—Device idea and process orientation.

O. Núñez—Signal acquisition electronic system.

N. López—Device Manufacture.

B. Reyes—Oversaw the methodology, data analysis, and writing process.

abdominal muscle contractions and with a transducer set to zero at level from the mid-axillary line, returning values between 5 and 7 mmHg in critically ill adults [1,6].

A significant increase in IAP triggers complications in different organs, inside and outside of the abdominal cavity [2]. Values greater than or equal to 12 mmHg (16 cmH₂O) are considered cases of intra-abdominal hypertension (IAH), from this value on, it is classified as grade 1 for $12 \text{ mmHg} \leq \text{IAP} \leq 15 \text{ mmHg}$ (20 cmH₂O), grade 2 for $15 \text{ mmHg} < \text{IAP} \leq 20 \text{ mmHg}$ (27 cmH₂O), grade 3 for $20 \text{ mmHg} < \text{IAP} \leq 25 \text{ mmHg}$ (34 cmH₂O), and grade 4 for $\text{IAP} > 25 \text{ mmHg}$. Sustained high IAP could evolve in abdominal compartment syndrome, which is a sustained IAP greater than 20 mmHg and it is associated with organ dysfunction or failure.

Monitoring the IAP is of great relevance during clinical practice, since it is a safe, precise, economical, and fast procedure to clinically diagnose IAH [4]. The IAP is usually estimated based on the current pressure in the urinary bladder, using a Foley manometer connected to an end of a urinary catheter. In this way, it is possible to intermittently measure the value of the pressure while the urinal flow is temporarily prevented [5]. Some systems use a needle to measure the pressure, while others use a column of fluid to do it mechanically.

For a given patient, the results of IAP measurement depend on the position of the zero point and the experience of the healthcare personnel carrying out the measurement, in addition to the obstruction of the flow of urine and the risk of urinary infections [3].

A popular procedure to measure IAP, is the indirect Kron's method, which consists of infusing 25 ml of saline solution to the patient's bladder via a Foley-type catheter, which is also connected to a column graduated in cmH₂O, being the volume adjustable. Its main drawbacks include a cumbersome preparation as well as the risk of contamination, making it more expensive compared to other methods or devices [1].

As an alternative to Kron's method, the use of pressure measurement devices at the base of the bladder has been proposed. These devices have common areas for improvement in terms of their design and signal transduction, e.g., reducing leaks on their hermetic sealing membranes, blockage of their catheters by residues, long stabilization time periods and high implementation costs [5]. For this reason, several studies of devices designed with similar technology and purposes, focused on improving the areas mentioned above, have been carried out and have shown positive results [1,7].

To overcome the mentioned areas of improvement, in this study, we propose the design and testing of a novel IAP measuring device based on the bladder pressure. To this end, *in vitro* experiments were performed, in which the measurements of the device were compared to the Kron's method connected to a graduated column, as described in Sect. 2. The results and their interpretation are presented in Sects. 3 and 4, respectively.

2 Materials and Methods

2.1 Device Design and Implementation

A novel device to continuously monitor the IAP based on the indirect method of the bladder was developed. This device prototype, shown in Fig. 1, consists of an encapsulation of three outer parts, and two inner ones, as well as the signal acquisition electronic system that is positioned in the upper part (Fig. 1, label A), which is manually calibrated, and the system inlet in the lower part, hermetically separated by a latex membrane (Fig. 1, label D) preventing contact between both chambers, thus avoiding contamination or infections of the patient's urinary tract, by exposure to microorganisms in the area.

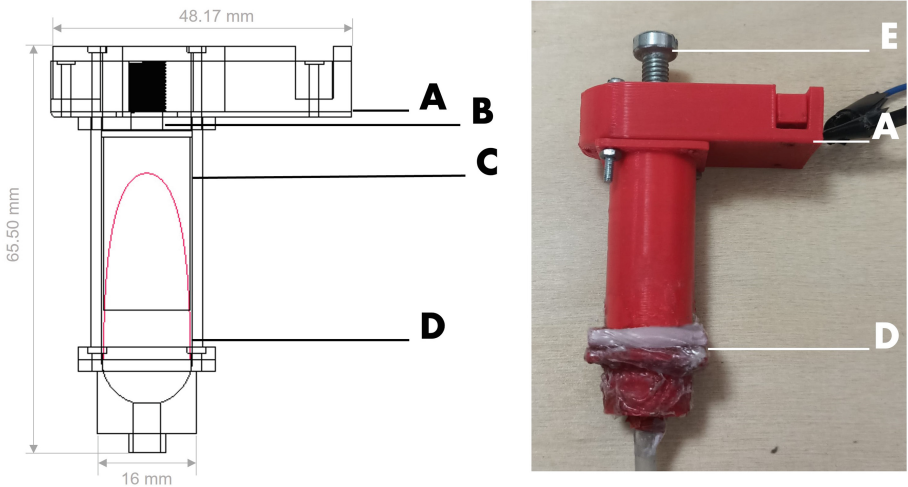


Fig. 1. Designed scheme and real device used *in vitro* tests. **A:** Location of the electronic transducer. **B:** Push heel. **C:** Push cylinder. **D:** Dividing membrane between the entrance and transducer chamber. **E:** Calibration screw.

The design of the device (Fig. 1) was implemented in such a way, the fluid (simulating bladder fluid) enters the bottom of the encapsulation and pushes the latex membrane (Fig. 1, label D), which in turn fills with fluid and pushes a cylinder-shaped part (Fig. 1, label C), that at its other end is shaped like a heel (Fig. 1, label B), that enters in contact with a force sensor (Fig. 1, label A) that is part of the signal acquisition electronic system, which transduces the push of the heel, due to IAP, into a force, by Pascal's Principle. Then, this force is transduced into an electronic signal, and subsequently, by means of a cubic regression relating the voltage read by a force sensor and a value shown in the graduated column, the pressure relationship can be obtained. The DF0-40 force sensor (Leanstar, Jiangsu Province, China) was selected for its measurement range within 500 g, and an almost instantaneous response time (<1 ms), which

are of great relevance for a continuous reading of data. The signal belonging to the simulated IAP is processed by first going through a calibration stage, explained in Sect. 2.2, then through a coupling stage, using an inverting amplifier circuit supplied at 3.3 V, and a negative input voltage of -1 V, according to the recommendations of the manufacturer, and thus passing to an amplification stage with 2.5 as an amplification factor. Finally, the pressure value is calculated using a cubic regression, with voltage as input and a pressure ratio as output, and displayed in cmH_2O , using a PIC18F4553-I/P microcontroller (Microchip Technology Inc., Arizona, United States). The stages of the data acquisition process are shown in Fig. 2.

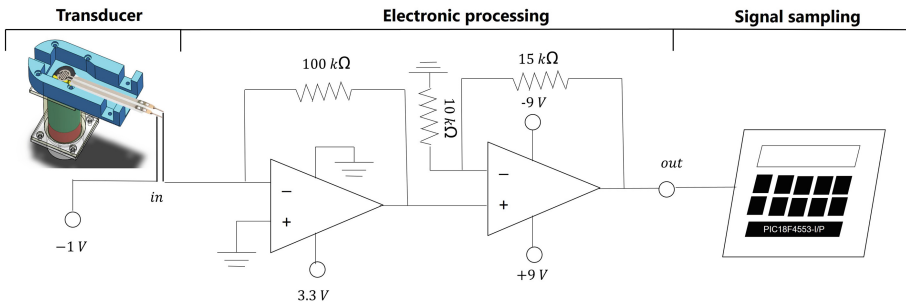


Fig. 2. Electrical scheme of the signal acquisition electronic system.

The different components of the package were designed using the 3D CAD (Computer-Aided Design) software SolidWorks 2019 version (SolidWorks Corp, Suresnes, France). The device prototype was manufactured in Acrylonitrile Butadiene Styrene (ABS plastic) by 3D printing, reducing manufacturing times with minimal costs.

2.2 Data Acquisition

The implemented system for *in vitro* tests, shown in Fig. 3, was based on the Kron's method connected to a graduated column, where saline solution is instilled into the patient's bladder and measurements are made during respiratory expiration or at clinically relevant events. Therefore, for *in vitro* data collection, water was placed in a sealed bag (Fig. 3, label A), which served to simulate the subject's bladder, this was placed in a manual press (Fig. 3, label B) to generate the different levels of simulated IAP. This mechanism was connected to a system of three-way stopcock valve connector (Fig. 3, label C) to the designed device (Fig. 3, label D) and to a column graduated from 0 to 45 cmH_2O (0 to 33 mmHg) (Fig. 3, label G), which worked as the known reference. The discharge duct (Fig. 3, label H) serves to release liquid from the entire system if required. Finally, the reading of the measurement after processing (Fig. 3, label E) was shown on the display of the device (Fig. 3, label F).

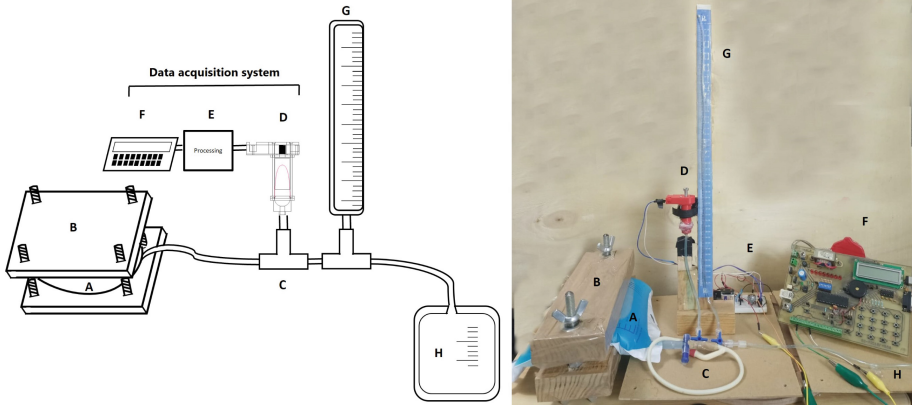


Fig. 3. Schematic and real setup of the system used for *in vitro* tests. **A:** Bag of water. **B:** Hand press. **C:** Three-way stopcock valve connector. **D:** Designed transducer. **E:** Signal processing circuit. **F:** Display. **G:** Water column (reference system). **H:** Collection bag.

To obtain the relationship between the voltage measurements made by the electronic transducer and the graduated column, i.e., to calibrate the device, the manual press was placed at a specific pressure, measured by means of the graduated column (0 cmH₂O was considered as reference). The device has a screw at the top (Fig. 1, label E), which exerts a variable pressure on the sensitive area of the sensor as it is inserted or extracted from the encapsulation, pressing on the external side against the flat part of the calibration screw, and on the internal side against the heel (Fig. 1, label B) mentioned in Sect. 2.1. In this way, by placing a pressure level on the graduated column, the calibration screw was turned to adjust a reference voltage. Finally, the voltage delivered by the device at said pressure was recorded (in the initial calibration process 0 cmH₂O at 0.54 V was selected). It should be noted that the three-way stopcock valves were open to form a closed system between the bag, device and column only. Once this measurement was made, the process was repeated, changing the pressure in the press.

Six sets of data pairs were taken as follows. On each cmH₂O of the column, the voltage was registered, each one denoting an average of 200 V readings, every 1.5 s, waiting around 10 s for each measurement to stabilize, three sets were taken upwards and three downwards, covering a range from 0 to 45 cmH₂O, resulting in 0.54 to 2.5 V, obtaining a total of 276 data. The average of the six measurements on each cmH₂O was used on a regression analysis, from which a cubic polynomial equation was obtained. It is worth mentioning that other polynomial regression orders were explored (Sect. 3). However, a cubic regression was selected and the corresponding equation was used to calculate the equivalent of the voltage on cmH₂O, which was shown on the display of the developed device. Finally, a series of new measurements were carried out to evaluate the accuracy of the designed

device and the calibration approach, taking a reading at each cmH₂O on the above mentioned range (0 to 45 cmH₂O) of simulated IAP. The simulation, and corresponding measuring, of the IAP value was randomly set during this new series.

2.3 Statistical Tests

After device calibration using the cubic polynomial obtained via regression, a linear regression analysis was performed between both methods, the designed device and the graduated column, in which the Pearson correlation coefficient was found. A Bland-Altman analysis was also implemented, from which the mean of differences, called bias, the standard deviation and the upper and lower 95% limits of agreement (LoA) were calculated.

3 Results

Regression analysis was performed to determine the relation between the voltage obtained by the signal acquisition system and the pressure measured by the water column. A lineal regression, a quadratic regression, and a cubic regression were performed, obtaining a Pearson correlation coefficient and a sum of residuals of $r = 0.9937$ and -0.0277 cmH₂O, $r = 0.9942$ and 0.0386 cmH₂O, and $r = 0.9953$ and 0.0144 cmH₂O respectively. The result of the polynomial regression analysis of third order was selected, as shown in Fig. 4, from which Eq. 1 was obtained.

$$IAP(cmH_2O) = -8.6837volts^3 + 37.58volts^2 - 20.19volts + 0.8485 \quad (1)$$

After applying the previous equation to calibrate the device lectures in volts to IAP pressures in cmH₂O, a high correlation was found, as seen in Fig. 5.A, between the graduated column and the designed device with Eq. 1 uploaded, with a Pearson correlation of $r = 0.9997$. Regarding the Bland-Altman analysis, shown in Fig. 5.B, a bias equal to 0.16 cmH₂O and 95% LoA between -0.444 and 0.758 cmH₂O were found. The bias was statistically significant different from a zero bias ($p = 0.0012$) considering 0.05 as the level of significance.

Figure 5.C shown the residuals from the lineal regression between the graduated column and the designed, while their corresponding histogram is shown in Fig. 5.D. Normality test was performed on these residuals and it was found that they are normally distributed, confirming the validity of the linear regression modeling.

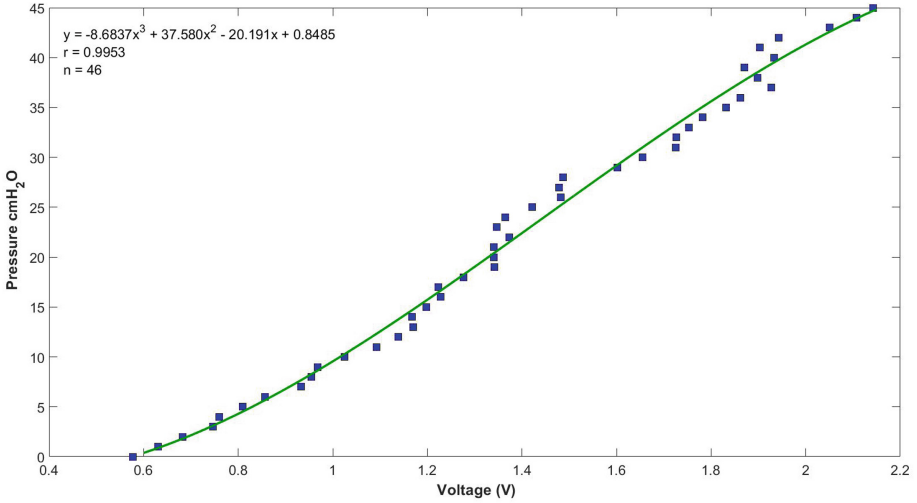


Fig. 4. Device's calibration equation. The blue squares represent the collected data, each one is an average of the 6 measurements done on each cmH₂O, while the green solid line represents the cubic polynomial model. (Color figure online)

4 Discussion

The cubic polynomial Eq. 1 was the correct selection to make the conversion to cmH₂O, since it better represents the relation between the independent variable (voltage) and the dependent variable (cmH₂O) of our method, since from a statistical point of view, it can be observed that the new test of pressure measurements made by the designed device have a high correlation with those of the reference system, i.e., the graduated column.

The Bland-Altman analysis, Fig. 5.B, made evident the presence of a statistically significant bias of 0.16 cmH₂O in the measurements of the designed device, but as it is a systematical error and it remained constant throughout the measurement range explored, it can be corrected by extracting its value from the measurement. It is expected to obtain results with high correlation and a bias closer to zero, being the accepted criteria of research and validation for a novel IAP method of less than 1 mmHg (1.36 cmH₂O) and less than 2 mmHg (2.71 cmH₂O) and 4 mmHg (5.44 cmH₂O) for the bias and LoA respectively, according to the World Society of the Abdominal Compartment Syndrome.

When comparing the obtained results with those from similar devices and *in vitro* experiments, one can see that our method stays at its level and presents some advantages over them, with a high correlation and a lower bias. We can see the advantages of our method with its lower systematical error, comparing it with another similar device [1], where the Pearson correlation coefficient was of $r = 0.99$ and the bias of 0.36 cmH₂O (0.27 mmHg).

In a recent study, another continuous measurement device was reported [7], where preclinical tests were carried out to validate their proposed device using

IAP Measurements. Designed device

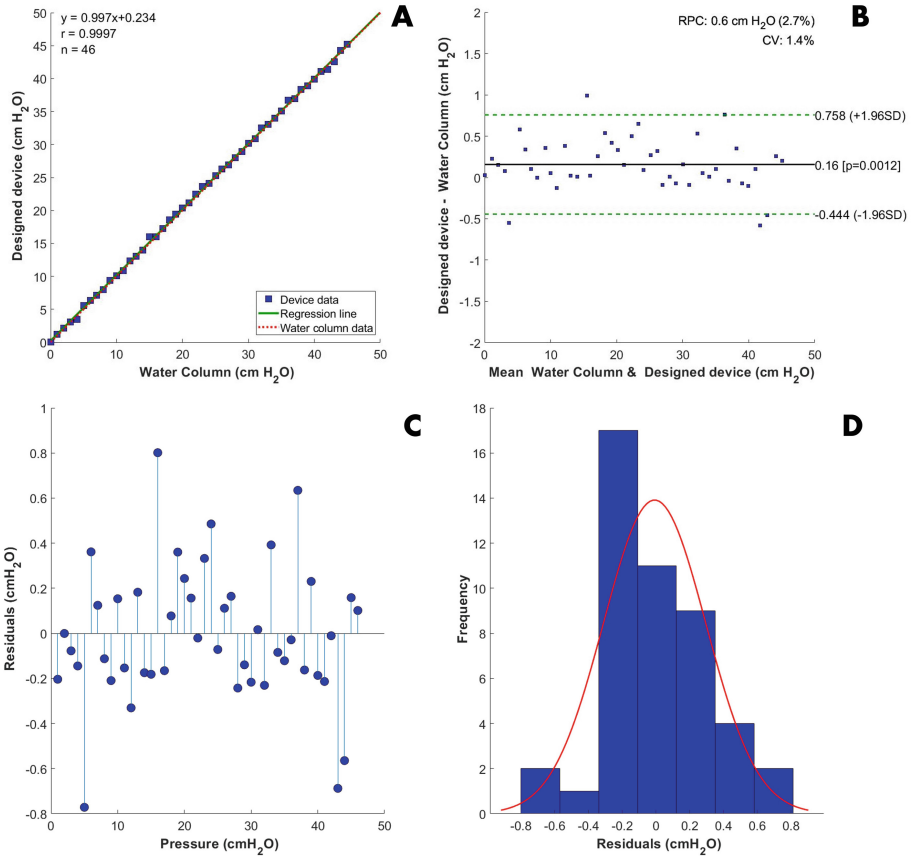


Fig. 5. IAP measurements of designed device. **A:** Lineal regression. **B:** The solid black line denotes the bias, while the green dashed lines denote the 95% limits of agreement (LoA). **C:** Residuals plot from the lineal regression. **D:** Histogram of residuals, the red line represents the normal distribution obtained from the estimated media and variance.

the water column graduated in cmH₂O, in a range similar to ours, from 0 to 35 mmHg. They found a high correlation coefficient of Pearson ($r = 0.99$), furthermore, their Blant-Alman analysis showed a bias of 0.16 cmH₂O (0.12 mmHg), equal to ours. We can consider a good robustness of our device in reproducing the measurements of the gold standard method, denoting that the devices reported in the previously mentioned studies were subsequently used in clinical trials.

Besides high correlation with the reference system, the designed device has other advantages, it does not need recalibration before a subsequent new set of

measurements since it remained during the test, also it does not need to know the exact volume for reproducibility as the results were based on the pressure of the water column, because in the laboratory where this study is carried out, an electronic instrument for reference pressure measurements is not available. In addition, since the main body of the device is manufactured by 3D printing, it facilitates reproducibility for *in vitro* studies by different research groups, as well as making changes to the original prototype and testing them in a matter of hours, contemplating a more appropriate manufacturing technique for possible future clinical implementation, such as the machining.

In comparison with Kron's method, the main advantage of the designed device is that by having a membrane between the fluid and the signal acquisition system it reduces the probability of the patient contracting a urinary tract infection. It also presents the opportunity to show the results directly in mmHg, in contrast with Kron's method in which the conversion must be done manually, reducing the time and effort consumption of healthcare personnel. Furthermore, it has an easiest result reading, since instead of the user determining the level of the water on the graduated column, the value is displayed directly on the device.

5 Conclusions

The designed device has important statistical results for the *in vitro* experiment; it is accurate in its measurements and has the potential to be improved in the future to be of aid in the diagnosis process of IAH.

6 Future Work

The device has its drawbacks that open the opportunity for future improvements. One of them is the redesign of its calibration system, because the one currently implemented is very sensitive, which causes the calibration process to be longer and more difficult, as well as more inaccurate at setting the zero. A bad calibration could cause inaccurate readings that would result in a misdiagnosis. Another area of opportunity to be explored in the future is the stability of the reading, as currently, the device takes approximately 10s to measure a pressure value and get a stable reading, that doesn't fluctuate and is close to the value shown on the reference system.





References

1. Castañón-González, J.A., Satué-Rodríguez, J., Rosales, F.C., Polanco-González, C., Miranda-Ruiz, R., Camacho-Juárez, S.: Nueva técnica y dispositivo para medir la presión intraabdominal. *Cir. Cir.* **81**(2), 112–117 (2013). <https://www.medigraphic.com/cgi-bin/new/resumen.cgi?IDARTICULO=40193>
2. Dabrowski, W., Rola, P., Malbrain, M.L.N.G.: Intra-abdominal pressure monitoring in cardiac surgery: is this the canary in the coalmine for kidney injury? *J. Clin. Monit. Comput.* **37**(2), 351–358 (2023). <https://doi.org/10.1007/s10877-022-00933-y>

3. Iacobovici, L., et al.: Assessment of Intra-abdominal pressure with a novel continuous bladder pressure monitor—a clinical validation study. *Life* **13**(2), 384 (2023). <https://www.mdpi.com/2075-1729/13/2/384>
4. Li, Z., Wang, H., Lu, F.: Monitoring and management of Intra-abdominal pressure in critically ill children. *Crit. Care Nurse* **43**(3), 44–51 (2023). <https://doi.org/10.4037/ccn2023545>
5. L N G Malbrain, M.: Different techniques to measure intra-abdominal pressure (IAP): time for a critical re-appraisal. *Intensive Care Med.* **30**(3), 357–371 (2004). <https://doi.org/10.1007/s00134-003-2107-2>
6. Pal Singh, S., Shankar, R., Bansal, M., Jain, N.: Comparison of intra-abdominal pressure monitoring vs APACHE-II and Ranson’s scoring in predicting the outcome of acute pancreatitis. *Int. J. Surg. Med.* **9**(1), 27–31 (2023). <https://www.ejos.org/?mno=124534>
7. Tayebi, S., Wise, R., Pourkazemi, A., Stiens, J., Malbrain, M.L.N.G.: Pre-clinical validation of a novel continuous intra-abdominal pressure measurement equipment (serenno). *Life* **12**(8), 1161 (2022). <https://www.mdpi.com/2075-1729/12/8/1161>



Analysis of Signals and Energy Storage in HAp-Collagen-Glucose Piezoelectric Samples Containing an Electrolyte

Susana Vargas¹ , Rogelio Rodriguez¹ , and Griselda Rivera²  

¹ Centro de Física Aplicada y Tecnología Avanzada, UNAM, Blvd. Juriquilla 3001, Querétaro 76230, México

² Departamento de Ingeniería Bioquímica. Escuela Nacional de Ciencias Biológicas del IPN, México City, México
argelia_riverav@yahoo.com.mx

Abstract. When a piezoelectric material is subjected to compressive stress, electric charges are induced on the surface producing an electric field E_e and a voltage. When there are free charges (electrolyte), these move towards charged surfaces forming the electric double layer (EDL), producing a field E_i in opposite direction and which reduces the output voltage. The EDL can store large amounts of energy. A model is presented that allows to predict the shape of the output profiles. Numerical values of various parameters were estimated to support these results.

Keywords: HAp · collagen · glucose · chemical capacitance · voltage profiles

1 Introduction

The piezoelectric effect is the property that certain materials have to generate electric charge in response to a mechanical action [1, 2]. When a mechanical stress is applied to a piezoelectric material, there is a shift of the centers of charge, positive and negative [3], which gives rise to an electric field $E_e = \frac{n_e q_e}{\epsilon A}$. In presence of an electrolyte, the ions move towards the charged surfaces forming the electric double-layer (EDL) [4, 5] and the electric field $E_i = \frac{n_i q_i}{\epsilon A}$. From the biological point of view, the piezoelectric effect is fundamental since all materials that form the human body are piezoelectric: HAp, collagen, etc., [6, 7]. It is possible to say that the piezoelectric effects are a fundamental property not only for the proper functioning of the human body and the preservation of life, but of all biological systems [8]. All parts of the body act as sensors, detecting actions of different intensities suffered in the body: it is possible to detect a wide variety of actions with different intensities, from severe which endanger the integrity of the body, to extremely small like a very small touch with a feather, or a hair on the face. How is that the body can detect such small actions?; this will be discussed.

Another essential constituent for the proper functioning of the body is the electrolyte since it is the electrical conduction system that allows the transport of ions; it performs

many important task in the body. The electrolyte fills the entire body, including the interior of cells. The significant ions in the electrolyte are: Na^+ , K^+ , Ca^{2+} , Mg^{2+} , HPO_4^{2-} . In addition to have the ability of electric conduction, a voltage source is required to supply the energy for the conduction; this is provided by the piezoelectric effect. The ions moved by E_e form a double-layer structure which controls the properties of the output signals; this is important because they affect the characteristics of signals that reach the brain who analyzes their shape and intensity and executes the correct actions. Double-layer structures are important because they can store large amounts of energy, maintaining very intense electric fields inside; this will be elucidated when the amount of energy stored in this type of systems is estimated.

1.1 Ionic Transport

Electric charges are induced on the surface of piezoelectric material in very short times when mechanical actions are exerted on it. The resulting net field and the voltage are $E_T = E_e - E_i$ and $V = V_e - V_i$; τ_i is the time required for ions to reach EDL; E_T and τ_i control the shape of the output signals. The arrival of ions at EDL gradually reduces the output voltage leading to a decaying profile; however, the ionic transport is modified by the blocking effect: high concentration of ions around the moving ion blocks their movement reducing transport: the ionic transport depends on concentration. The arrival time τ_i can be obtained by equating the electric force with the viscous force and solving for the velocity, and the inverse is:

$$\tau_i \sim \left(\frac{n_i R_i}{q_i} \right) \quad (1)$$

The blocking effect was introduced as n_i ; q_i , and R_i , are the charge and size of the ions.

An electrolyte in the presence of a charged surface produces a double-layer (DL) [9]. It is known that the accumulation of ions in a finite volume requires considerable amounts of energy due to the enormous repulsive force between charges: there is a minimum distance d_{mda} one charge can approach another that is, there is a maximum charge concentration allowed. The ions move towards the surface trying to be as close as possible to the charged surface due to the attraction with the induced charge; when d_{mda} is reached in the layer closest to the surface (Stern layer) the filling stop; the remaining ions are accommodated in the second layer filling it to reach d_{mda} , and so on; these layers will gradually form a diffuse and no-uniform structure with a thickness λ . The induced charge stabilizes EDL by a strong attraction force F_{DL} and very small separation distance, the Stern thickness λ_{SL} which is in the range [0.3–0.8] nm; this force accommodate large number of ions in a very tight and ordered manner, resulting in a structure that can store high amounts of energy and charges in small volumes ($A \lambda$). λ and $\kappa = \lambda^{-1}$ are given by [10].

$$\lambda = \sqrt{\frac{\epsilon \kappa T}{n_o q^2}} \text{ and } \kappa = (\lambda)^{-1} = \sqrt{\frac{n_o q^2}{\epsilon \kappa T}} \quad (2)$$

Numerical λ values have been reported for different valences and dilute conditions [10]; these are in the range $[2.5 \times 10^{-10} - 10^{-8}]$ m, and for κ $[10^8 - 4 \times 10^9]$ m^{-1} . The induced charge is superficial and EDL is a very thin layer both separated by a distance λ_{SL} ; the attractive force F_{DL} between these layers is intense and depends quadratically on κ , i.e., $F_{\text{DL}} \sim \kappa^2$. The charged surface and DL have opposite sign forming a special capacitor called double-layer capacitor (DLC). The potential ψ and charge density are strong in DL but decay rapidly with the distance [11]:

$$\rho(x) = \rho_o e^{-x/\lambda} = \rho_o e^{-\kappa x} \quad (3a)$$

$$\Psi = \Psi_o e^{-x/\lambda} = \Psi_o e^{-\kappa x} \quad (3b)$$

1.2 Numerical Evaluation of Energy and Other Parameters

To estimate numerically several parameters, some values have to be assigned to various basic physical quantities within their experimental ranges [10]; this estimation will provide an idea of orders of these quantities: λ is in the range $[2.5 \times 10^{-10} - 1.0 \times 10^{-8}]$ then it is assigned the value $\lambda = 10^{-9}$ m; κ is in the range $[10^8 - 4 \times 10^9]$ m^{-1} then it is assigned the value $\kappa = 10^9$ m^{-1} ; ψ_o is in the range $[10 - 100]$ mV then $\psi_o = 0.05$ V; the area will be chosen of 2×2 cm^2 then $A = 4 \times 10^{-4}$ m^2 ; the distance $d = 1$ mm = 0.001 m; the dielectric permittivity of acetonitrile is $\epsilon = 3.32 \times 10^{-10}$ F/m; λ_{SL} is in the range $[0.3 - 0.8]$ nm then $\lambda_{\text{SL}} = 5 \times 10^{-10}$ m; $C_o = \epsilon A/d = 1.33 \times 10^{-10}$ F. The charge on the DL is $Q_{\text{DL}} = \int_0^\lambda \rho(x) A dx = 0.64 \rho_o A \lambda$ and the superficial charge $\sigma_{\text{DL}} = Q_{\text{DL}}/A = 0.64 \rho_o \lambda = 1.6 \times 10^{-2}$ C/m²; this superficial charge corresponds to an area occupied of 10 nm² per ion in a tight packing situation: $\sigma = \frac{1 \text{ ion}}{10 \text{ nm}^2} = 1.6 \times 10^{-2} \frac{\text{C}}{\text{m}^2}$; then $\rho_o = 1.57 \sigma \kappa = 2.51 \times 10^7$ C/m³. The force between charged surfaces $F_{\text{DL}} = \frac{\epsilon}{2} A E^2 = \frac{A}{2\epsilon} \sigma^2 = \frac{A}{2\epsilon} \rho_o^2 \lambda^2 \approx 3.74 \times 10^2$ N this force stabilizes both charged layers and it is responsible for the tight packing of charges. The energy U depends on E^2 , $U = \frac{\epsilon}{2} \int E^2 d\text{Vol}$; for U_{DL} the integral was evaluated from 0 to λ , while for U_{elec} from λ to $d - \lambda$; $U_{\text{DL}} = 1.427 \times 10^{-7}$ J and $U_{\text{elec}} = 2.257 \times 10^{-8}$ J; the results show that 92.1% of the energy is in the DLCs, and 7.9% in the electrostatic. The energy density in both DLCs is: $u_{\text{DL}} = 8.3 \times 10^5$ J/m³ and $u_{\text{elec}} = 0.56 \times 10^5$ J/m³. For the electric field $E = -\frac{d\Psi}{dx}$ the average value is reported: $\langle E \rangle_{\text{DL}}$ in the range $[0 - \lambda]$ and $\langle E \rangle_{\text{elec}}$ in the range $[\lambda - (d - \lambda)]$: $\langle E \rangle_{\text{DL}} = 1.368 \psi_o \psi = 6.84 \times 10^7$ V/m; this value can be compared with that obtained from $E = \psi_o/\lambda = 5 \times 10^7$ V/m, as can be seen these values are close. The DLC is given by $C_{\text{DL}} = \epsilon A \kappa = 1.33 \times 10^{-4}$ F and $C_{\text{elec}} = \epsilon A d = 1.33 \times 10^{-10}$ F. All these values are reported in Table 1. Based on these estimates it is possible to say that the double layer capacitor stores large amounts of energy and charge, with intense electric fields in its interior. This is a very appropriate system to have supercapacitors and store large amounts of energy and charge in very small volumes.

1.3 Shape of the Voltage Profiles

Induced charges q_e produces E_e and V_e at $t = 0$. EDL produces E_i , which reduces the output signal $E_T = E_e - E_i$ and $V = V_e - V_i$. The ions deposit rate decreases as the number

Table 1. Estimated values of some parameters

	Energy Density (J/m ³)	Electric Field (V/m)	Capacitance (F)	Charge Density (C/m ³)
Dark	u _o = 0	E ₀ = 0	C _T = 1.33 × 10 ⁻¹⁰	ρ _o = 0
Light; C ₁	u ₁ = 4.15 × 10 ⁵	E ₁ = 6.84 × 10 ⁷	C ₁ = 1.33 × 10 ⁻⁴	ρ _{o1} = 2.5 × 10 ⁷
Light; C ₂	u ₂ = 0.56 × 10 ⁵	E ₂ = 0.45 × 10 ⁶	C ₂ = 1.33 × 10 ⁻¹⁰	ρ _{o2} = 0
Light; C ₃	u ₃ = 4.15 × 10 ⁵	E ₃ = 6.84 × 10 ⁷	C ₃ = 1.33 × 10 ⁻⁴	ρ _{o1} 2.5 × 10 ⁷
Total	u _T = 8.30 × 10 ⁵			
Attraction Force between Double Layers			F _{DL} = 3.74 × 10 ² N	

of deposited ions increases (dq_i/dt) ~ (q_o - q_i) because the arriving ions have to overcome the field of already deposited ions; solving this equation and since V_i is proportional to q_i, V_i = V_o(1 - e^{-t/τ}) and the output voltage V = V_o - V_i = V_oe^{-t/τ} + V_{DC} is a decaying exponential; because there is more than one type of ions, two exponentials are required:

$$V = V_{o1}e^{-t/\tau_1} + V_{o2}e^{-t/\tau_2} + V_{DC} \tag{4}$$

1.4 Chemical Capacitance C_μ

A capacitor is a device that stores energy in the electric field in a region of space located by metallic objects. Inside a piezoelectric material with an electrolyte, there are two DLCs, and one electrostatic capacitor C_{elec} (Fig. 3); DLCs store most of the energy with a diffuse delimitation of the electric field. DLCs contribute to Chemical Capacitor C_μ which is defined as C_μ = e²(∂n_i/∂μ_i) where n_i is the number density of ions and μ_i the internal chemical potential. Because n_i(μ) = n_{oi}e^{μ/kT}, then, C_μ = εκ² = ελ⁻², i.e., C_μ is proportional to κ², the most important characteristics of DL. The dependence of C_μ with the voltage was calculated using an expression for the chemical potential for ions in a DL [11]:

$$\Delta\mu_j = \mu_j - \mu_o = -\frac{z_j^2 q_i^2}{4\pi\epsilon} \left(\frac{\kappa}{1 + \kappa a} \right) \text{ and } \kappa = \frac{(\mu_o - \mu_j)}{[\beta - a(\mu_o - \mu_j)]} \tag{5}$$

Where β = z_jq_e²/4πϵ. Using Eq. 5b, the fact that ψ is small ψ = ψκ_o [1 - x + ...], because the voltage is V = ψ_o - ψ, and writing the Boltzmann Distribution Law as:

$$n_i = n'_{oi}e^{+(q_i/kT)V} \tag{6}$$

is obtained for C_μ

$$C_\mu = C_{0\mu} e^{+(q_e/kT)V} \tag{7a}$$

The logarithm of Eq. 7a:

$$\text{Ln } C_\mu = b + mV \tag{7b}$$

This is a straight line where $b=\text{Ln}(C_{0m})$ and $m = q_e/kT$. Equations 7a-7b are the same as those obtained by Bisquert et al. [12].

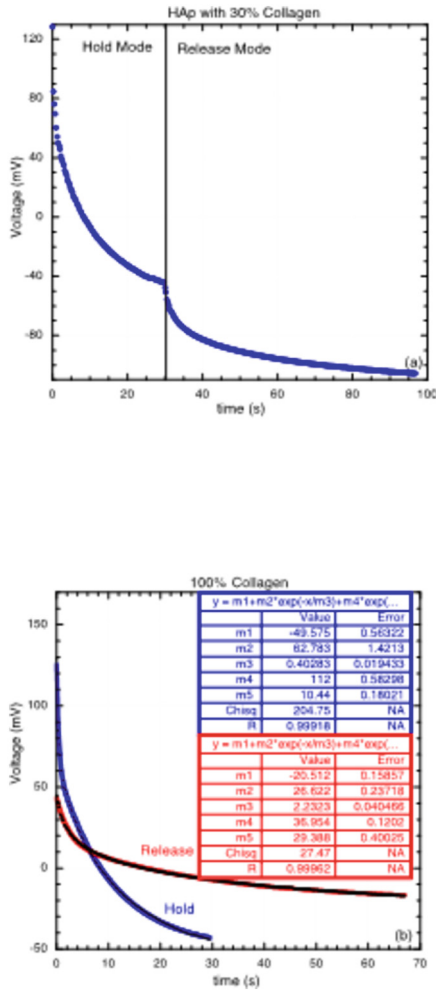


Fig. 1. Voltage profiles of a) HAp-30%C and b) 100%C samples, for H and R modes. Both curves in Fig b were fitted using Eq. 4.

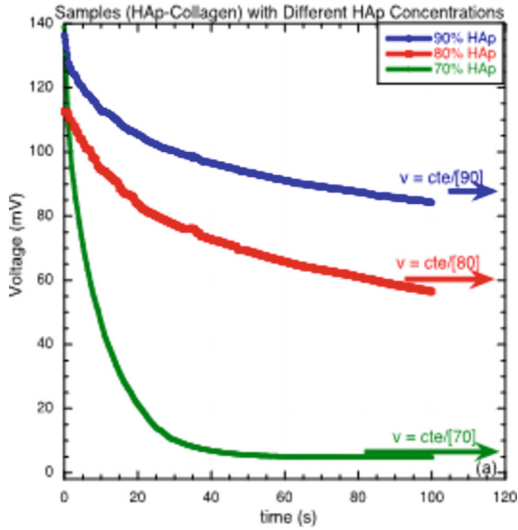


Fig. 2. Voltage profiles of HAp-Collagen samples at different HAp content: 70, 80, 90%.

1.5 Glucose and Chemical Capacitance

Glucose is the primary energy source of living organisms. It is a very active small molecule with five hydroxyls, four secondaries, and one primary which can react with HAp and collagen. High glucose levels in blood produce undesired effects: diabetic foot, neuropathies, etc.

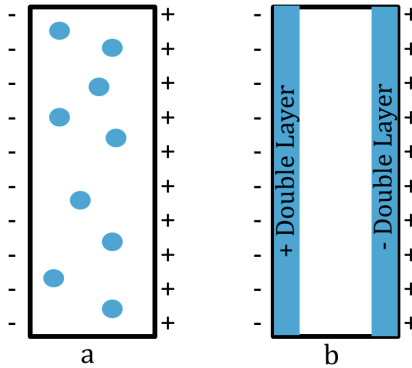


Fig. 3. Schematic diagram showing: a) the ions in the electrolyte and b) the migration toward the charge surface to form the EDL.

Glucose affects the structure of the HAp, interrupting the crystal growth by blocking the exposed functional groups and reducing its piezoelectric response: glucose avoids the deposit of Ca^{2+} and HPO_4^{2-} ions on HAp, preventing its growth:



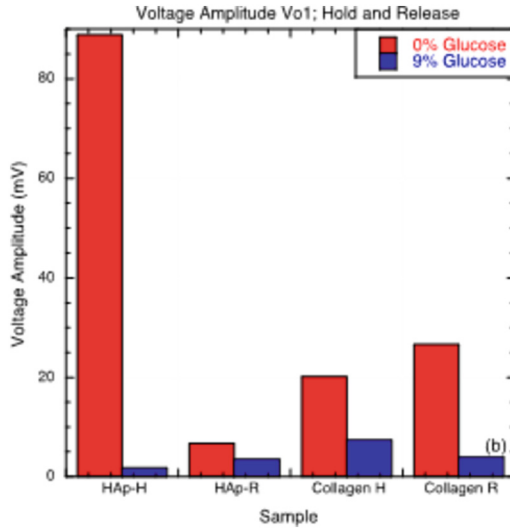


Fig. 4. Effect of the addition of 9% glucose on the output voltage in HAp-C samples.

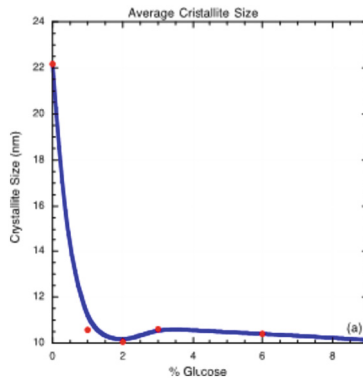


Fig. 5. Effect of increasing glucose from 0 to 9% on the crystal size of HAp.

The reaction with glucose prevents the addition of, for example, Ca^{2+} to continue growth of the HAp. The crystallite size was reduced from 22 to 10.2 nm when glucose concentration was increased to 9% (Fig. 5). The increment in glucose drastically reduces the amplitude of the output voltage (Fig. 6): for HAp in H mode the intensity was reduced from 88.8 to 4.8 mV when glucose was increased from 0 to 3%. In all cases (Fig. 4), there is a reduction in voltage when glucose is increased but a lower scale.

As mentioned, there is a considerable amount of energy stored in DLC; this energy can be recovered in situations of low or no stress: rest, sleep, etc.; during exercise or physical work, a large amount of energy is consumed coming essentially from food; however, part of the energy consumed by the exercise is stored in the DLCs. During sleep, the body recovers part of the energy stored in DLCs and uses it for soft and hard

(bone) tissue restoration purposes; in the electrolyte, there is an abundance of Ca^{2+} and HPO_4^{2-} ions because these ions serve as a mineral reserve to maintain and restore bone tissue.

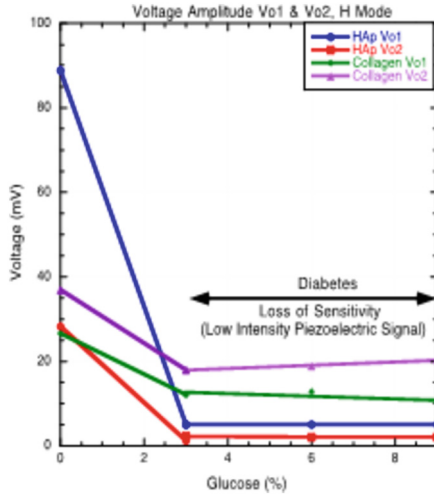


Fig. 6. Voltage amplitudes V_{o1} and V_{o2} as a function of [%G].

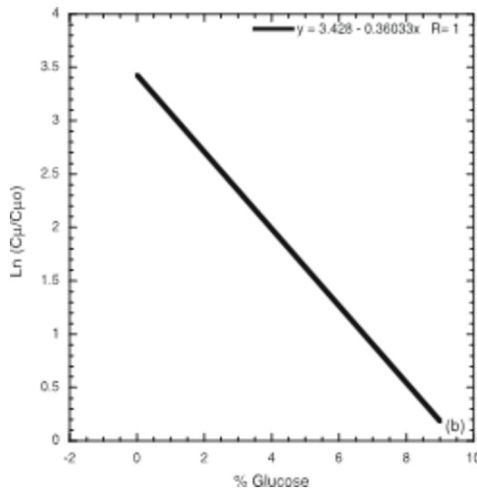


Fig. 7. Dependence of $\text{Ln}(C_{\mu}/C_{\mu o})$ as a function of [%G].

As mentioned, DLC can store large amount of energy; however, this energy can only be recovered in a specific charge and discharge voltage window; this is controlled by decomposition voltage of the electrolyte: electrolysis occurs at low voltages in aqueous solutions (1.5 V) and higher for organic solvents (3 V). To have a supercapacitor at

high voltages operation and/or at high stored energy, as required at industrial level a bank of small supercapacitors stacked in series is a possible option. For n identical DLC capacitors C_i in series, the total capacitance is $\frac{1}{C} = \sum_1^n \frac{1}{C_i} = \frac{n}{C_i}$, i.e., $C = \frac{C_i}{n}$. Using this expressions, the total energy U_T and voltage V_T are [13]:

$$U_T = \frac{1}{2} CV^2 = \frac{1}{2} \left(\frac{C_i}{n} \right) (nV_i)^2 = n U_i \text{ and } V_T = nV_i \quad (9a)$$

Both quantities, energy and voltage, increase linearly with the number of capacitors in the stack. Even though this looks easily feasible, there are several technological problems that need to be solved. Possibly the most difficult of these is related to making and stacking n small identical DLC: these small DLCs have to identical and wit the same internal resistance. When a stack of DLCs is overloaded, capacitors with high resistance and low capacitance are at risk of decomposition and gas production; similar problems occur in batteries made by stacking voltage cells.

2 Experimental

2.1 Materials

HAp was synthesized by the co-precipitation method as reported elsewhere [14]; the reactants were: calcium nitrate tetra-hydrated ($\text{Ca}(\text{NO}_3)_2 \cdot 4\text{H}_2\text{O}$), di-ammonium phosphate ($(\text{NH}_4)_2\text{HPO}_4$), ammonium hydroxide ($\text{NH}_4 \text{OH}$) (all from Aldrich Co); an aqueous solution at 20% of collagen was obtained from Eclat Laboratory. Glucose was obtained from Dextrosa Anidar (d-glucose) by Golden Bell, code 27740. HAp-C samples were prepared at different collagen concentrations: 0, 10, 15, 20, 25, and 30%. Glucose was added to HAp and collagen at concentrations: 0, 3, 6, and 9%. The powder samples were mixed with 20% of polyurethane as agglutinant. Two compression modes were used: hold (H) and release (R); in H mode the voltage was measured with the maximum stress applied (26.5 kPa), while in R mode the stress was removed while the voltage was measured.

2.2 Characterization Techniques

HAp crystal size determination was made using a Rigaku diffractometer model Ultima IV operating at 30 mA and 40 kV and the crystallite size determination was performed using JADE v8.1. The voltages as a function of time were obtained using a digital multimeter USB Steren model MUL605 equipped with data communication software to PC. The multimeter has a resolution of 0.01 mV in the 0–60 mV range with an accuracy of $\pm (0.5\%)$ and an input impedance of 10 M Ω . All samples were moistened for 15 s in a distilled water fog produced by a commercial ultrasonic humidifier (Homedics model UHE-WB01); the moistened disks were placed immediately between electrodes to reduce evaporation and for electrical determination.

3 Results and Discussion

In Figs. 1a–1b show typical voltage profiles of the samples HAp-30%C and 100%C; in Fig. 1a, the H profile is followed by R profile; it is in this way that the voltage profiles were obtained, first applying and then removing the applied stress; in Fig. 1b the profiles are shown separately and both were fitted using Eq. 4; from this figure it is possible to observe that the H profiles are more intense and of short duration (175 mV and 30s) than the R ones (63.57 mV and 70 s). The H profiles are related to acute situations that can represent potential damage to the body; in this case, the brain analyzes very carefully the characteristics of the piezoelectric signal to take the most appropriate actions to protect the body. The R profiles are signals obtained after stress has been removed; then, these signals correspond to a healing situation and they are of low intensity and large duration; in this case, the brain pays less attention since they represent a low risk.

During HAp synthesis, some Ca^{2+} and HPO_4^{2-} ions remain unreacted in the electrolyte; the effect of the ions concentration (70, 80, and 90% HAp) on the profiles are shown Fig. 2. The arrival time, $\tau_i \sim (n_i R_i / q_i)$, is linear with concentration due to blocking effect: high concentration (90%) means low velocity and high τ_i values; on the other hand, low concentration (70%) means high velocity and low τ_i values.

Figure 4 shows the effect of an increase in glucose on HAp and collagen; glucose drastically reduces the voltage from 88 to 1.8 mV for HAp, and from 20 to 7.6 mV for collagen for H mode; for R mode the reductions are moderated. A reduction in voltage corresponds to a loss in sensitivity; patients with high local glucose concentrations show loss of sensitivity, mainly in legs and feet. An analysis of piezoelectric signals allows an early, local, and non-invasive diagnosis of glucose levels in different body parts.

Figure 5 shows a plot of the crystal size as a function of glucose concentration. The main reduction, from 22 to 10.5 nm, occurs when glucose reaches 1%. Equation 8 shows the blocking reaction produced by glucose on the hydroxyl groups of HAp, preventing the reaction with Ca^{2+} and HPO_4^{2-} . This reduction in the crystallite size is linked to the reduction in the voltage and, therefore, to the loss of sensitivity.

It was obtained a relationship between C_μ and V (Eq. 7a), $C_\mu = C_{o\mu} e^{+(q_e/kT)V}$ which is the same as the reported by Bisquert et al. [12]; additionally, these authors report profiles at different temperatures from 0 to 60 °C, where the slope is the same as that obtained here: q/kT . Figure 4 shows a voltage reduction from 88.8 to 4.8 mV when glucose was incremented from 0 to 9% ($V = 0.0888 - (0.084/9)[\%G]$), but Eq. 7a provides a relationship between C_μ and V : $\frac{C_\mu}{C_{\mu o}} = e^{(q/kT)V}$, and because $q/kT = 38.6 \text{ Cs}^2/\text{m}^2\text{kg}$, $\text{Ln}(C_\mu/C_{\mu o}) = 3.43 - 0.36[\%G]$ (Fig. 7) or

$$C_\mu = C'_{\mu o} e^{-\alpha[\%G]} \quad (10)$$

where $\alpha = 0.36$, $C'_{\mu o} = 30.82 C_{\mu o}$, and $[\%G]$ is the percentage of glucose; Eq. 10 is plotted in Fig. 7. The role of glucose is to reduce the chemical capacitance C_μ and, consequently, the energy stored in DLC.

Figure 6 shows a dependence of V_{o1} and V_{o2} for different glucose concentrations for HAp and collagen in H mode; it is possible to observe a reduction in amplitudes when the glucose concentration is increased; particularly when glucose reaches 3% there is a large reduction in V_{o1} and V_{o2} for HAp and collagen: V_{o1} of HAp was reduced

from 88.8 to 4.8 mV when glucose reaches 3%; this concentration can be considered as the threshold for local diagnosing diabetes. When there is excess glucose in the body, it is not distributed homogeneously in the body, but it is concentrated in certain specific places. Piezoelectric profiles make it possible to determine the effect of a local concentration of glucose in the body. This is possibly the biggest difference between this technique and standard techniques such as the standard glycosylated hemoglobin tests (GHT); this measure the average blood glucose throughout the whole body. In the case of piezoelectric signals, 3% of glucose is a local threshold concentration; for this reason, the levels measured by GHT are lower with respect to the obtained here. This value of 3% glucose has been also obtained experimentally in laboratory animals: the glucose uptake in mice skeleton *in vivo* is reported as the percentage of glucose injected dose per gram of tissue weight: %GID g^{-1} ; these determinations were made in the tibia and femur and the average value was $3.5 \pm 0.38\%$ GID g^{-1} [15]. This value is pretty close to that obtained here for the percentage of glucose added per gram of HAp. 3.0% GAg^{-1} ; this threshold value is similar to that obtained in [15].

4 Conclusions

The expression for the arrival time τ_i reported in Eq. (1) was verified for different samples at different concentrations; its effect is observed in Fig. 2; the ions concentration modifies the shape of the output profiles. A simple model was used to explain the shape of output profiles; this is because the time required to induce a charge on the surfaces is very short. It was determined the amount of energy and charge stored in DLCs: practically all the energy is stored in DLC and practically nothing in the electrostatic capacitor. Analytical relationships between C_μ and the output voltage V and glucose concentration [%G] were obtained.







References

1. Shukat, H., et al.: Piezoelectric materials: advanced application in electro-chemical process. *Energy Rep.* **9**, 4306–4324 (2023). <https://doi.org/10.1016/j.egy.2023.03.077>
2. Marino, A., Gnchi, G.G., Sinibaldi, E., Ciofani, G.: Piezoelectric effects of materials on bio-interfaces. *Appl. Mater. Interf.* **9**, 17663–17680 (2017). <https://doi.org/10.1021/acsami.7b04323>
3. Bassett, C.A.L.: Biological significance of piezoelectricity. *Calc. Tiss Res.* **1**, 253–272 (1968)
4. Pollack, S.R., Salzstein, R., Pienkowski, D.: The electrical double layer in bone and its influence on stress-generated potentials. *Calcified Tissue Int.* **36**, S77–S81 (1984)
5. Starr, M.B., Shi, J., Wang, X.: Piezopotential-driven redox reaction at the surface of piezoelectric materials. *Angew. Chem. Int. Ed.* **51**, 5962–5966 (2012). <https://doi.org/10.1002/anie.201201424>
6. Lay, R., Deijs, G.S., Maltröm, J.: The intrinsic piezoelectric properties of materials- a review with a focus on biological materials. *RSC Adv.* **11**, 30657 (2021). <https://doi.org/10.1039/d1ra03557f>
7. Ribeiro, C., Secada, V., Correia, D.M., Lanceros-Mendez, S.: Piezoelectric polymers as bio-materials for tissue engineering applications. *Colloids Surf. B* **136**, 46–55 (2015). <https://doi.org/10.1016/j.colsurfb.2015.08.043>

8. Kim, D., Han, S.A., Kim, J.H., Lee, J-H., Kim, S-W., Lee, S-W.: Biomolecular piezoelectric materials: from amino acids to living tissues. *Adv. Mater.* **32**, 1906989 (2020). <https://doi.org/10.1002/adma.201906989>
9. Debye, P., Hückel, E.: Zur theorie der elektrolyte. II. das grenzgesetz für die elektrische leitfähigkeit. *Physikalische Zeitschrift.* **24**(15), 305-325 (1923)
10. Hiemenz, P.C., Rajagopalan, R.: Principles of Colloid and Surface Chemistry, Third Edition, Marcel Dekker Inc., Chapter 13, Section 7a (1997)
11. Blandamer, M.J., Reis, J.C.R.”le.ac.uk Topic 680, 1.5.21: Chemical Potentials; Salt Solutions; Debye-Hückel Equation”, University of Leicester & Faculdade de Ciências, Department of Chemistry, Course Notes, LibreTexts Chemistry, Topic 680, <http://www.le.ac.uk/chemistry/thermodynamics/pdfs/1000/Topic0680.pdf>
12. Bisquert, J., Fabregat-Santiago, F.: Impedance spectroscopy: a general introduction and applications to dye-sensitized solar cell, from dye-sensitized solar cells. In: Kalyanasundaran, E.K., EPFL Press, Fig. 12.34, pp. 526, 527 (2010)
13. Electrochemical, K.P., Capacitors, D.-L.: Electrochemical energy storage for renewable source and grid balancing. Chap. **19**, 345–407 (2015). <https://doi.org/10.1016/B978-0-444-62616-5.00019-X>
14. Sequeda, L.G., Diaza, J.M., Gutierrez, S.J., Perdomo, S.J., Gomez, O.L.: Obtención de hidroxiapatita sintética por tres métodos diferentes y su caracterización para ser utilizada como sustituto óseo. *Rev Colomb Cienc Quim Farm* **41**(1), 50–66 (2012)
15. Zoch, M.L., Abou, D.S., Clemens, T.L., Thorek, D.L.J., Riddle, R.C.: In vivo radiometric analysis of glucose uptake and distribution in mouse bone. *Bone Res.* **4**, 16004 (2016). <https://doi.org/10.1038/boneres.2016.4>



Intelligent Device for the Control of Electrical Outlet Usage for Medical Devices

Héctor G. Delgado Rivera^(✉), Zaira L. Bojorquez Muñoz,
César F. Hernández Ortiz, Sergio A. Corona Sánchez,
Natalia Montañez Prieto, and José M. Macías Macías

Universidad La Salle Chihuahua, Ingeniería Electromédica, Prolongación Lomas de Majalca #11201, Labor de Terrazas, Chihuahua, Mexico
hectorderi1@gmail.com

Abstract. It is important to verify the correct use electrical outlets in hospitals, since medical equipment must be available to staff and patients. Different institutions in Chihuahua city mention that the use of electrical outlets isn't only used for medical equipment, sometimes the staff prioritizes outlets use for mobile devices or household appliances, which presents a risk to patients when requiring emergency medical equipment use. Due to this problem, an intelligent device was developed to control the electrical outlet use specifically for medical devices, called Health Switch, which is an outlet that controls electrical power cuts, preventing mishandling when trying to connect another type of device that stops the equipment operation in question, thus preventing any possible accident within the facility. In addition, it is observed that by state of Mexico and attention level, there is a 16.8 million MWh consumption in electrical energy, 18.9 million MWh in fossil energy and 12.7 million tons of CO₂ emission per year. This device allows on and off control through a mobile application, making electricity consumption, working time and productivity of the medical center more efficient.

Keywords: medical equipment · socket · energy · electrical · microcontroller · healthcare institutions

1 Introduction

According to the Directorate of Health Information of the Ministry of Health in Mexico, total public and private health care spending in Mexico amounts to 1,589,016.04 pesos (MXN) per year[1]. An analysis of operating costs and energy consumption, based on the information collected from each medical unit in the representative sample, provides a projection of costs at the national level and by state, with energy information for each type of equipment:

1. Power supply.
2. Refrigeration.
3. Lighting.

4. Sanitary hot water.
5. Heating.
6. Air conditioning.
7. Steam.
8. Motors.
9. Pumping system.
10. Compressor system.
11. Medical equipment.
12. Small variable loads.

According to data from the Mexican Secretariat of Energy (SENER), the total energy cost in the health sector represents 4 percent of the total operating expenditure in the health sector [2]. The charges for energy consumed according to the Federal Electricity Commission (CFE) [3] are as follows:

1. 0.843 pesos for each of the first 75 (seventy-five) kilowatt-hours.
2. 1.018 for each of the next 65 (sixty-five) kilowatt-hours.
3. 2,983 for each additional kilowatt-hour.

The energy impact of the health sector in Mexico, based on energy billing, average data, and UAM by state and level of care, is 16.8 million MWh in electrical energy, 18.9 million MWh in fossil energy, and the emission of 12.7 million tons of CO₂ per year.

Table 1 shows the Maximum installed power per area (M.I.P.A.) the Energy consumption per area (E.C.A.), and the Cost per area (C.A.) of care within the medical unit. The emergency room is the area with the highest energy consumption with 20 percent of the total, as it is designed to attend to patients arriving in critical condition at any time, so medical equipment, lighting, and air conditioning, among others, must be in uninterrupted operation.

Table 1. Energy consumption by Area.

Area	M.I.P.A.(kW)	E.C.A.(kW)	C.A.(MXN)
Medical Office	95	71,471	112,898
Emergency	64.5	136,093	214,979
Hospitalization	123.8	156,080	246,552
Specialties	81.1	84,273	133,121
Laboratories	101	167,532	264,640
Imaging	344.2	278,000	439,141
Totals	809.6	893,449	1,411,331

Table 2. Measured and rated power by device category and specific make/model in each.

Category	Spot measurements: standby, average, peak (W)	Rated power (W)
Airway Clearance	12, 233, 235	500
Anesthesia Unit	153, 302, 342	1440
Aspirator A	12, 20, 40	60
Aspirator B	NA, 115, 119	240
Autotransfusion	63,153, 75	NA
Bed A	20, 447, NA	NA
Bed B	30, 94, NA	NA
Circulatory Assist	NA, 4, 10	50
Computer-Infosys	NA, 24, 47	50
Defibrillator	NA, 29, 31	130
EEG	NA, 142, 143	NA
Exam Chair or Table	25, 150, 271	600
Exerciser	NA, 3, 6	NA
Hemodialysis Unit A	67, 87, 131	600
Hemodialysis Unit B	NA, 48, 51	NA
Hemodialysis Unit C	83, 504, 1574	1840
Humidifier A	12, 16, 147	185
Humidifier B	NA, 8, 10	NA
Humidifier C	NA, 40, 45	NA
Incubator-Infant	30, 308, 619	1035
Meter A	6, 16, 16	55
Meter B	NA, 7, 10	NA
Microscope	185, 602, 648	NA
Monitor-Patient A	8, 52, 53	145
Monitor-Patient B	2, 18, 19	NA
Monitor-Patient C	8, 38, 39	156
Monitor-Patient D	2, 5, 8	6
Monitor-Patient E	NA, 17, 17	161
Monitor-Patient F	NA, 49, 52	47
Monitor-Patient G	4, 37, 42	NA
Phototherapy	2, 42, 44	180
Positive Airway A	2, 126, 183	360
Pump A	NA, 8, 8	60
Pump B	NA, 193, 195	372
Pump C	7, 16, 18	14
Pump D	3, 41, 22	150
Pump E	NA, 15, 30	120
Pump F	NA, 57, 96	120
Scanning System	13, 945, 996	NA
Scopes	80, 250, 276	NA
Smoke Evacuation	56, 876, 882	NA
Tester A	NA, 7, 7	NA
Tester B	168, 568, 1026	1920
Ventilator A	NA, 119, 207	NA
Ventilator B	NA, 58, 66	135
Ventilator C + UPS	NA, 35, 80	800
Ventilator D + UPS	71, 92, 194	800
Ventilator E + UPS	NA, 100, NA	800
Ventilator F + UPS	34, 164, 220	863
UPS	NA, 95, 109	660
Warmer-Lab A	1, 22, 94	NA
Warmer-Lab B	1, 19, 92	450
Warmer-Patient A	NA, 650, NA	792
Warmer-Patient B	46, 688, 826	1000
Water Purification A	4, 127, 272	570
Water Purification B	41, 132, 258	NA

There are comparisons of the actual measurements of the devices with their rated powers, including measurements of consumption in standby mode, in normal operation, and at peak consumption whenever possible. Table 2 shows the device categories, measurements, and corresponding power ratings. In general, the rated power is usually higher than the operating and peak power, but the difference varies depending on the equipment category and device brand [4].

The results of the energy consumption of some equipment of the medical unit are presented in Table 2, these are related to the types of equipment (E), the annual operating hours (A.O.H), the hours in which the equipment remains connected in standby representing a percentage of between 7 to 11 percent even when they are in an idle state, as well as the annual energy consumption (A.E.C.), and the annual energy cost (A.E.C.)[2].

Table 3. Energy consumption per equipment.

E	A.O.H.	A.E.C. (kW)	A.E.C. (MXN)
Boilers	1,456	168,258	204,259
Centrifuges	1,602	39,268	70,904
Medical Gases	730	8,432	15,224
X-Ray	759	16,846	30,149
Equipment	2,172	106,810	152,478
Totals	6,719	339,614	473,284

According to data from the Mexican Ministry of Energy (SENER), in most of the country’s medical centers, those responsible for the Medical Care Units (UAM) do not provide their staff and users with any awareness training in energy-saving and efficiency measures, nor in preventive and corrective maintenance operations to improve the conditions of the facilities and maintain consumption levels, as well as rational energy levels [2].

Another of the main problems that usually occur in medical units is the poor management of electrical outlets intended explicitly for medical equipment, which can be a major obstacle at the time of requiring the use of the device, presenting a possible consequence to the health of patients, especially when it is life support, as well as delays in consultation and treatment times.

Worldwide, more than 135 million deaths are registered each year due to cardiovascular causes, with coronary heart disease as the main cause of cardiac arrest. International literature considers that between 0.4–2 percent of patients admitted to a hospital and up to 30 percent of those who die require cardiopulmonary resuscitation techniques [5]. That is why it is of great importance to make good use of specific power outlets for medical equipment, for example, defibrillators, which are vital equipment for cardiopulmonary resuscitation with an 80 percent survival rate when used [6], which has a battery life of 1.5 h and needs up to 4 h to reach its maximum capacity [7], so that if it is disconnected it will not be possible to use it in an emergency of vital importance.

Power factor refers to the effectiveness of a load in utilizing the current supplied by the system. It is represented as the ratio of the actual power (P), which is measured in kilowatts (kW), to the power output (S), which is measured in kilovolt-amperes (kVA):

$$\text{Power Factor} = \frac{P}{S}$$

In this project, relays were used to allow the passage of electric current, so the power factor of the devices is not modified, unlike the triacs which do influence the power factor.

In order to enrich the research, several tests and calculations were performed on different equipment, as mentioned below.

A harmonic is a sinusoidal component of a periodic wave or a quantity having a frequency that is an integer multiple of a fundamental frequency. Harmonic currents are generated by residential and industrial loads with non-linear characteristics. Depending on the method of operation, these sources will produce harmonics at a constant or variable level.

Additionally, the Fourier series can represent a distorted signal, which can be calculated using the following formulas:

a) Percentage of the nth Harmonic Distortion:

$$D_n = \frac{|A_n|}{|A_1|}$$

b) Total Harmonic Distortion in percent:

$$\text{THD} = \sqrt{D_2^2 + D_3^2 + D_4^2 \dots} \times 100$$

Applying these formulas when testing medical equipment and devices in daily use, the following results are obtained:

1. THD in centrifuge: 4.8 percent.
2. THD in electrostimulator: 3.98 percent.
3. THD in computer: 3.43 percent.

Due to these developments, the project aims to address both issues, developing a smart device programmed to perform automatic power cuts depending on the consumption of electronic devices, discriminating the current values of devices that are not within the established parameters, for example, household appliances and non-medical devices, making their use impossible, with the purpose of assigning the use of the switch only to certain medical equipment, as well as through a mobile application manipulate the on and off of medical equipment saving electrical energy and decreasing waiting times for medical care. The use of these smart switches implemented within hospitals or medical centers is of utmost importance when it comes to automating processes and keeping track of the energy consumed.

Another problem in the rehabilitation and therapy area is the use of equipment that requires preheating to reach the appropriate temperature conditions

for the treatments to be performed, therefore, it is necessary to turn them on in advance prior to the consultation.

For example, the Parabath paraffin heater consumes 120W to heat its resistance between 52° - 57° in a time of 30 min [8], or a Chattanooga compressor has a temperature range between 30° to 80° taking one hour to reach this temperature and consuming 1000W to achieve it [9]. Due to the aforementioned, it is more susceptible that the consultation hours are delayed or that the staff chooses to keep the equipment on despite not being used for long periods of time in order not to have economic losses because a therapy session is between 700 MXN to 1250 MXN [10], consequently there is higher energy consumption and even possible affections to the equipment and risks due to overheating.

To solve the above problem, a switch was created using the ESP32 microcontroller[11], and an Android application [12] to control the on and off of the equipment, as well, make efficient the use of the electrical socket for medical equipment.

Ways of control:

1. Remote control through the designed application.
2. Automatic control generated by power outages of devices that have a consumption outside the established parameters.

The remote control is given by the mobile application through Bluetooth connection, changing the high and low states according to the selected. Similarly, you can set specific names for each of the inputs of the outlet. The control in its automatic way looks for the switch to discriminate the values outside the parameters captured by the SCT-013 sensor [13], to prevent the flow of current, therefore it will be impossible to operate the device connected to that input.

2 Materials and Methods

Health switch is composed of three main elements:

1. Electronic Tablet.
2. Microcontroller Operating Algorithm.
3. Mobile Application.

2.1 SCT-013

The SCT-013 sensor works as a transformer, where the current flowing through the cable to be measured acts as the primary winding, which consists of a single turn. Inside, the sensor has a secondary winding that can have up to more than 2000 turns, depending on the model. With this sensor, it is possible to measure currents in the range of 0 to 100 A, and its internal load resistor produces a voltage output with a ratio of 100A/1V. It is used in this case to measure the amperage demanded from the equipment to the outlet and, in this way, to cut off if the values are outside the established parameters.

2.2 Electronic Tablet

The Bluetooth connection, BLE protocol, is between the microcontroller and the mobile device.

A voltage follower is placed on the analog inputs in order to avoid fluctuations of the sensor signal. For switching on and off, the 2N4401 transistors provide the gate function for the connection of the relays that will be the system switches.

Finally, the relay is connected to a protection system between its coil and the necessary sockets between the socket and the board track, since a high voltage is conducted on it (Fig. 1).

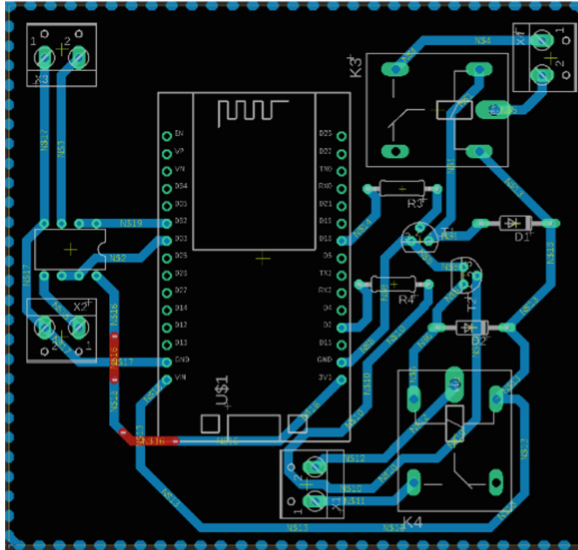


Fig. 1. Eagle Electronic Tablet Designs

2.3 Microcontroller Operating Algorithm

The program for the ESP32 microcontroller has the following functions: Generate automatic current cuts from the evaluation of the parameters acquired by an SCT013 sensor. Analyze and compare the commands received from the mobile application and the value of the current consumed by the equipment in real time.

The libraries EmonLib.h [14] for the use of sensors, and BluetoothSerial.h [15] for Bluetooth communication between the mobile device and the microcontroller are used.

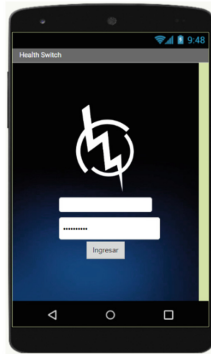
In the first instance, the Bluetooth BLE protocol is used in the application, protocol by which the communication with the ESP32 is established. As well as the ports and storage variables are also given high.

Then in the main loop the current parameters are set, assigning a value, one or zero, to the storage variables depending on the sensed current.

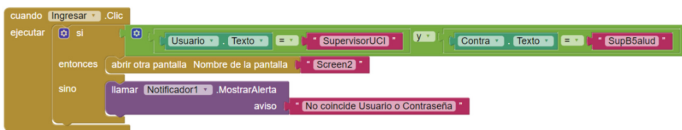
In the second part of the code, the char values are received from the application, making a comparison between the data received from the application and the values of the storage variables. As long as the values from both sources are met, the devices can be switched on or off by the application. However, if the connected devices are out of the programmed current range, the switch function will not be enabled, only the cut-off action. The code is available for download at the following link: <https://drive.google.com/file/d/15EiNdRgh0KmowDp-Py9Sd4ot-AJ0hN2a7/view?usp=sharing>

2.4 Application in App Inventor

App Inventor allows programming based on blocks, carrying out a large number of functions such as: control, logic, text, lists, mathematics, among others [16].



a)



b)

Fig. 2. a) Start-up screen 1. b) Blocks programmed to perform functions on screen 1.

Figure 2 shows screen 1, which corresponds to the login interface, for which it is necessary to enter a user name and password.

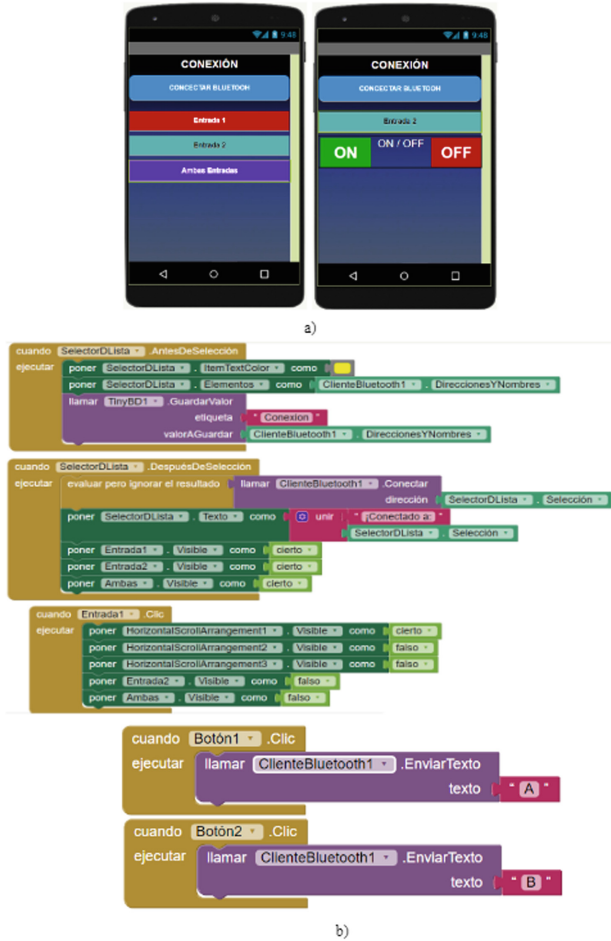


Fig. 3. a) Screen 2. Bluetooth connection and Inputs to choose. b)Blocks programmed to perform functions on screen 2.

Figure 3 shows screen 2, where first the wireless connection is made via Bluetooth with the microcontroller, enabling the options of each socket individually or simultaneously from the sockets; the characters corresponding to the options selected in the application are transmitted to send them to the char variables of the microcontroller code. Specifically, characters “A” and “C” send a logical one to the microcontroller, “B” and “D” transmit a logical zero, while “E” sends a logical one to the conditional variables (data) of both sockets and “F” transmits two logical zeros to the same variables.

However, the switch function does not only depend on the char characters, but the programmed current parameters must also be met.

3 Results

To validate the performance of the device, various household appliances, mobile equipment and different medical equipment were used.

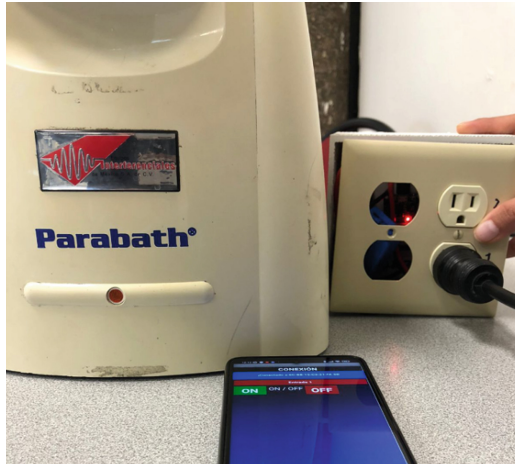


Fig. 4. Paraffin outside the allowed ranges.

In Fig. 4, the paraffin maker was tested by programming the corresponding ranges of current consumption in the socket so that the device does not comply with the parameters, and therefore the automatic cut-off function is performed despite being connected to the socket, in addition to preventing the medical device from being turned on from the mobile application.



Fig. 5. Paraffin within established ranges.

On the contrary, in Fig. 5 the acceptable current consumption conditions were adjusted again so that the relay continues to allow the passage of the current demanded by the parafinder, thus enabling the operation of the on and off buttons of the mobile application.



Fig. 6. Heating outside the allowed ranges.

Figures 6 and 7 show how the operating principle is the same with a household appliance. It should be emphasized that the parameters are adjusted according to the current demand for each piece of equipment and the purposes for which it is to be used. For example, preventing the use of personal use devices (cell phones, computers, coffee makers, etc.) in sockets exclusively for medical equipment or including connection to Brain-Computer Interfaces [31–33].



Fig. 7. Heating within the established ranges.

Figure 8 analyzes the behavior of a computer, electrostimulator and autoclave implementing the designed device, obtaining the same response as in the previous tests.

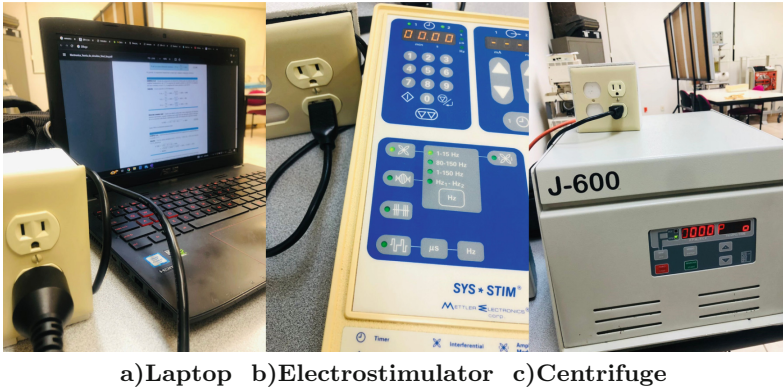


Fig. 8. Current test

4 Conclusions

The developed project, and under the conditions in which it is found, can be a viable solution to the possible negligence of the personnel for the incorrect use of outlets designated to medical devices in particular, this by means of the analysis of the current consumed by the connected devices when they are sensed and compared with the parameters established in the programming code of the microcontroller, to exercise the functions of automatic cut-off and switch.

There are improvements that can increase its accuracy and reliability, reducing risks for the patient and optimizing work times. Among the improvements is Wi-Fi communication to cover a greater connectivity distance.

On the other hand, once the code was completed and the tests corresponding to the current cuts were performed, it was identified that the SCT-013 sensor is capable of measuring the amount of current consumed by the equipment, however, being a basic use sensor, it is recommended to use it exclusively for testing and for applications such as those desired, so it is recommended to be replaced by industrial use sensors such as the Hall effect, in addition, changing the sensors would reduce the size of the prototype, providing, along with its operation, greater aesthetics and presentation.






References

1. Douglas, B., Steven, L., Judy, L., Richard, B., Brett, S.: Evaluation of miscellaneous and electronic device energy use in hospitals. *World Rev. Sci.* **10**, 113–128 (2013)
2. Estudio de Eficiencia Energética En Hospitales. Secretaría de Energía (SENER). Euro-Funding Advisory Group 1–13 (2015)
3. Reanimación Cardiopulmonar en Adultos. Instituto Mexicano del Seguro Social 12–65 (2017)
4. Larson, M.P., Elsenberg, M.S., Cummins, R.O., et al.: Predicting survival from out-of-hospital cardiac arrest: a graphic model. *Ann Emerg Med.* 1652–1658 (1993)
5. CENETEC. Guía Técnica No. 29 Desfibriladores 7–3 (2005)
6. Larson, M.P., Elsenberg, M.S., Cummins, R.O., et al.: Predicting survival from out-of-hospital cardiac arrest: a graphic model. *Ann Emerg Med.* 1652–1658 (1993)
7. CENETEC. Guía Técnica No. 29 Desfibriladores 7–3 (2005)
8. Parabath: Parabath® Paraffin Heat Therapy Owners Guide (2017)
9. Guía del Usuario Unidad Calefactora Hydrocollator. Equipos Interferenciales de México, S.A. de C.V. (2021)
10. Descubre el precio de una sesión de fisioterapia. <https://www.cronoshare.com.mx/cuanto-cuesta/sesion-fisioterapia> (n.d.)
11. Hakki, S.O.Y.: ESP8266 and ESP32 series of SoC microcontrollers. *Programmable Smart Microcontroller Cards*, pp. 110–125 (2021)
12. David, W., et al.: *App Inventor*. O'Reilly Media, Inc. (2011)
13. Ebim Iskandar, M.: Perbandingan Data Sensor Arus SCT 013 dan Sensor Arus ACS 712 Pada Pengukuran Arus Listrik AC. *Diss. Universitas Brawijaya* (2017)
14. de Araujo, N.P.: Protótipo para monitoramento dos parâmetros de transformadores por internet das coisas-IoT (2020)
15. García Aguilera, L.: Control para estabilidad estática de un robot bípedo de 10 GDL (2021)
16. Michel Miola, M., et al.: Desenvolvimento de componentes para App Inventor e servidor para dispositivos IoT (2019)
17. Nicolas-Alonso, L.F.: Brain computer interfaces, a review. *Sensors* **12**(2), 1211–1279 (2012)
18. Ramírez Aguilar, C., et al.: Estudios en Materia de Eficiencia Energética. *SENER* 2 (2019)
19. Luna, V., Ignacio, J.: Monitoreo de concentración de monóxido de carbono usando tecnología long-range. *Ingenius. Revista de Ciencia y Tecnología* **18**, 73–83 (2017)
20. David Morillón, G.: Retos y oportunidades para la sustentabilidad energética en edificios de México: Consumo y uso final de energía en edificios residenciales, comerciales y de servicio (2015)
21. Aguilar, C. Ramírez, Et Al. Evaluación Energética Del Sistema De Climatización En Un Hospital Público De Ciudad Valles, Slp. *Normas Para Publicación En Tectzapic 2 Resiliencia Organizacional Ante Los Procesos Antrópicos Derivados De Las Operaciones Productivas Del Ingenio Alianza Popular Rosas Rivera, M. Del C; Orduña Correa, F.; Izaguirre González R.; Medina Flores, Lr*, vol. 2444, p. 40 (2014)
22. Martí, J., Tiers, T.: Eficiencia energética en hospitales. *Mundo de la domótica: nuevas tecnologías para el hogar, la oficina y la industria.* **59**, 14–15 (2009)
23. Contreras, M.: Erwin Leonel. Criterios para la implementación eléctrica de quirófanos integrados o inteligentes, Tesis de Licenciatura (2012)

24. Lizcaño Jaramillo, P. A. Evaluación de Tecnologías en Salud: Un Enfoque Hospitalario para la Incorporación de Dispositivos Médicos. *Revista mexicana de ingeniería biomédica*, vol. 40, no 3 (2019)
25. Abu Hassan, A., et al.: Development of ESP32-based Wi-Fi electronic nose system for monitoring LPG leakage at gas cylinder refurbish plant. In: 2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA), pp. 1–5. IEEE (2018)
26. Robert, W., ENG, C.: YHDC SCT-013-000 Current Transformer (2016)
27. Iker, H.A.: Diseño de una instalación fotovoltaica para el ahorro de consumo de energía eléctrica en un hospital (2019)
28. Chiappé, Á.: Miguel Enrique. Plan de ahorro de energía en instalaciones hospitalarias (1992)
29. Iñigo, F.P.: Análisis de consumos energéticos y estudio de instalación fotovoltaica para autoconsumo de energía eléctrica en un hospital (2019)
30. Arenas Mimblera, A.: Sistemas de confort para una habitación inteligente de hospital, mediante la tecnología ZIGBEE (2020)
31. Ramírez-Quintana, J., Macias-Macias, J., Corral-Saenz, A., Chacon-Murguía, M.: Novel SSVEP processing method based on correlation and feedforward neural network for embedded brain-computer interface. In: Carrasco-Ochoa, J., Martínez-Trinidad, J., Olvera-López, J., Salas, J. (eds.) *Pattern Recognition. MCPR 2019. Lecture Notes in Computer Science()*, vol. 11524. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-21077-9_23
32. Ramirez-Quintana, J.A., Macias-Macias, J.M., Ramirez-Alonso, G.: A novel deep capsule neural network for vowel imagery patterns from EEG signals. *Biomed. Sig. Process. Control* **81**, 104500 (2023). <https://doi.org/10.1016/j.bspc.2022.104500>. ISSN 1746-8094
33. Macías-Macías, J.M., Ramírez-Quintana, J.A., Chacón-Murguía, M.I., Torres-García, A.A., Corral-Martínez, L.F.: Interpretation of a deep analysis of speech imagery features extracted by a capsule neural network. *Comput. Biol. Med.* **159**, 106909 (2023). <https://doi.org/10.1016/j.compbiomed.2023.106909>. ISSN 0010-4825



Design of a Wrist-Worn Device for Simultaneous Detection of ECG and Cardiac Pulse: A Preliminary Study

Rafael Gonzalez-Landaeta¹  , Aldo Rodrigo Mejía Rodríguez² ,
Guadalupe Dorantes Mendez² , and Dora-Luz Flores³ 

¹ BIOCIM Research Group, Universidad Autónoma de Ciudad Juárez, Ciudad Juárez, Chihuahua 32310, México

rafael.gonzalez@uacj.mx

² Faculty of Sciences, Universidad Autónoma de San Luis Potosí, 78295 San Luis Potosí, México

³ Facultad de Ingeniería, Universidad Autónoma de Baja California, Arquitectura y Diseño, Ensenada, Baja California 22860, México

Abstract. Wearable systems, such as watch/wristband systems must deal with power consumption and power line interference problems without compromising the form factor of the device and the signal-to-noise ratio (SNR). In this study, we present a preliminary design of a wrist-worn device that simultaneously detects ECG and cardiac pulse. Unlike the systems currently available in the market, the cardiac pulse is measured by detecting the magnetic disturbance caused by the blood flow in a localized magnetic field, eliminating the need to use optical sensors, which demand higher currents. The device's circuitry was implemented using surface mount technology (SMT) on a 43.5 x 32.5 mm 4-layer PCB. With these dimensions, the contribution of electromagnetic interferences was lower than 42.5 μV when the device was used in an office setting. The main current consumption was lower than 500 μA , and the SNR was higher than 68 dB for the ECG and higher than 55 dB for the pulse signal, enabling a clear identification of the different waves of the detected signals. The aim is to simplify the signal processing algorithms in such a way that several features of the detected signals can be easily identified using fewer hardware resources.

Keywords: ECG · Cardiac Pulse · Wrist-Worn device

1 Introduction

Monitoring of physiological variables in non-hospital settings has become common today. This practice reduces costs in health systems and allows information on the subject's health status to be obtained during their daily activities. Wearable technology has become popular among consumers [1] since there are systems that provide information on heart rate, breathing, mental health, and physical activity, among others [2, 3].

The most common wearable systems are smartwatches or bracelet-type devices since they can be worn comfortably for long periods. The most popular are the Apple Watch (Apple), Galaxy Watch (Samsung), ScanWatch (Withings), Sense (Fitbit) [4], and the WHOOP band (WHOOP Inc.) [5]; most of these systems detect various physiological signals, from which they indirectly estimate other parameters, such as blood pressure, blood oxygenation, respiration, stress, and sleep disorders, to name a few. Among the signals that are most detected with these systems are the electrocardiogram (ECG) and the photoplethysmogram (PPG). Regarding the ECG, detection is usually done using two dry electrodes, allowing the detection of a single ECG lead (usually lead-I) [6]. However, in the case of the Apple Watch, studies have been carried out where the primary function of such a device has been extended to measure multiple leads (not simultaneously) [7]. Optical sensors (LEDs and a photodiode) are often used to detect PPG. These sensors are in contact with the skin, making it possible to detect changes in blood volume in peripheral arteries.

ECG sensing using two dry electrodes fits very well into the wearable paradigm. However, it is challenging from the electronic design point of view since the contribution of power line interference (60 Hz), V_{EMI} , depends on the imbalance of the impedances of the electrodes (ΔZ_E), the common mode impedance (Z_C) of the front end, the common-mode rejection ratio (CMRR) of the system, and the isolation impedance (Z_{ISO}) [8], that is:

$$V_{EMI} = i_p \frac{Z_B}{Z_C + 2(Z_B + Z_{ISO})} \left(\Delta Z_E + \frac{Z_C}{CMRR} \right) \quad (1)$$

where i_p is the power line displacement current flowing through the body, and Z_B is the patient-ground impedance.

Regarding the detection of PPG, the use of LEDs implies a high-power consumption (tens of mW) [9], which considerably reduces the autonomy of the system. However, there are strategies to reduce the consumption of LEDs, the best known is to use a switched-mode power supply [10], but it requires a greater number of electronic components that make electronics bulky.

In this work, a preliminary study is presented focused on the design of a watch/bracelet-type device capable of simultaneously detecting the ECG and the cardiac pulse signal. The aim is to design a low-power electronic circuitry capable to detect two biosignals with high SNR and with a remarkable immunity to power line interferences. To achieve this, a printed circuit board (PCB) was designed to guarantee the signals' integrity. For the ECG detection circuit, those aspects of the electronic design that allow reducing the contribution of line interference and the loading effect with the electrode-skin impedance were also addressed, such as a system with high differential input impedance (Z_D), and Z_C , a high CMRR, and a high Z_{ISO} . Strategies to reduce electronic noise in the system and maintain low current consumption were also discussed. The method proposed by Phua et al. [11] will be implemented to detect the cardiac pulse. This method detects the disturbance caused by blood on a localized magnetic field. To achieve this, a permanent magnet and a magnetic sensor were placed over a (large-size) artery. This technique requires simple circuits which leads to a lower consumption. This proposal was already presented by Méndez-Lira et al. [12], where it was possible to simultaneously detect the ECG and the cardiac pulse. However, the system consumed

over 10 mA and was developed at the breadboard level, which is far from the ideal design and implementation considerations for a wearable system. The main goal of the proposed system is to track and assess in the future the cardiovascular health, including the automatic detection of the long QT syndrome.

2 Materials and Methods

2.1 ECG Detection Circuit

Figure 1 shows the circuit used for detecting the ECG by two dry electrodes. The front end relied on a fully differential ac-coupling network to achieve a high CMRR and to eliminate the half-cell potential from the electrodes. The input impedance behaves as an inductor, so a high input impedance is ensured at the power line frequency without using very high-value resistors [13]. This reduces the loading effect between the front end and the contact impedance of the dry electrodes, which, according to previous tests, is about 80 k Ω for our custom-built electrodes. The front end has a unity gain, so the instrumentation amplifier (IA) fixed the overall gain. This means that the high gain of the IA amplifies the voltage noise of the front end. To reduce the output noise of the circuit, the noise voltage of the front end must be low. The operational amplifier (OpAmp) used was the LMP2234, a quad micropower amplifier (Texas Instruments) with a noise voltage of 60 nV/ $\sqrt{\text{Hz}}$ @ 1 kHz and current consumption of 34 μA . The IA was the INA826 (Texas Instruments), a rail-to-rail amplifier with a CMRR = 104 dB, a noise voltage of 18 nV/ $\sqrt{\text{Hz}}$ @ 1 kHz, and a current consumption of 200 μA . The ECG circuit was tested in two scenarios: 1) using a TechPatient CARDIO V4 patient simulator (HE Instruments LLC, FL, USA), configured to generate an ECG signal of 1 mV amplitude at 72 bpm. 2) Measuring the lead-I ECG of a volunteer.

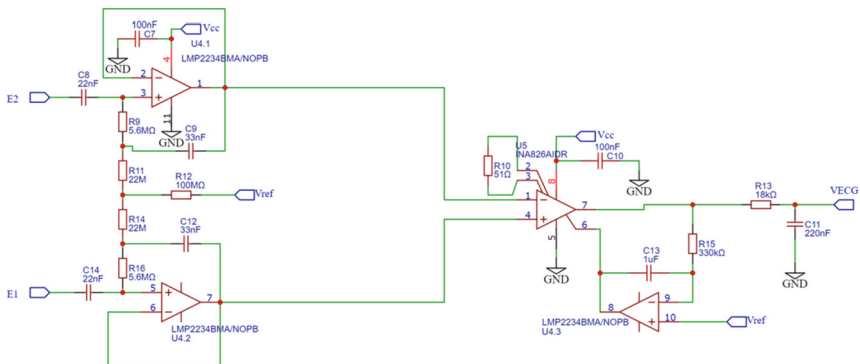


Fig. 1. Electronic circuit for detecting de ECG using two dry electrodes.

2.2 Cardiac Pulse Detection

Magnetic Field Source

The detection of the cardiac pulse is based on the measurement of the disturbance caused by the blood flow on a focused magnetic field. To generate this field, a circular (Diameter = 15 mm) neodymium magnet of 130 mT was used and placed over the radial artery.

Sensor

A TMR2001 (Multi Dimension) tunnel-type magnetoresistive sensor was used to measure the magnetic disturbance. It is a full Wheatstone bridge with a sensitivity of 80 mV/V/mT, an output impedance of 63 kΩ , and consumes 16 uA. To detect the magnetic disturbance, the sensor was placed over the radial artery at a distance of 5 mm from the magnet.

Electronic Circuit

Figure 2 shows the circuit used to detect the cardiac pulse. The differential output of the TMR2001 sensor is connected to the input of this circuit. The front end consists of an IA coupled in ac by a symmetric network that guarantees a high CMRR [14]. With this configuration, zero errors from the sensor in the presence of the constant magnetic field of the magnet are eliminated. The input impedance of this system is 1 MΩ, which reduces loading errors when connected to the magnetic sensor. The IA is the AD627 (Analog Devices, MA, USA); it has a current consumption of 60 μA and a noise voltage of 38 nV/√Hz. This circuit was also tested on the same volunteer.

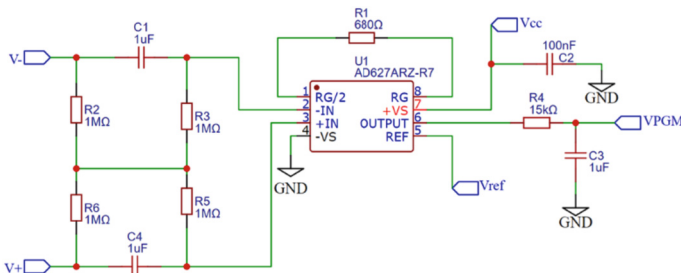


Fig. 2. Electronic circuit for detecting de cardiac pulse.

2.3 Printed Circuit Board Design

All electronic circuits were implemented using surface mount technology (SMT). In addition to the circuits in Figs. 1 and 2, a battery charging circuit that uses a Micro-B SMD USB connector, a 3.3 V voltage regulator circuit, and a 1.65 V reference voltage were also included. Although we are working with low-frequency signals, the PCB design seeks to guarantee the integrity of the signal for the subsequent integration of digital circuits and wireless transmission systems. For this, a 4-layer PCB was designed

with dimensions of $43.5 \text{ mm} \times 32.5 \text{ mm}$. The top layer (Fig. 3a) contains all the circuitry. The inner layer 1 (Fig. 3b) is a ground plane with no segmentations, which helps reduce the impedance and the mutual inductance between circuits. This is very useful when working with digital systems. The ground plane was placed close to the top layer to reduce the path of return currents and reduce induced noises. The inner layer 2 (Fig. 3c) is the power plane. This plane was divided into three parts: a) the one that handles the supply voltage (V_{in}) coming from the USB port, b) the battery voltage V_{BAT} , and the supply voltage of the active circuits (V_{CC}). This prevents noise from the several power supplies from being scattered throughout the circuitry. Finally, the bottom layer (Fig. 3d) was used to trace some signal paths. However, this layer will be used to implement the microcontroller and wireless data transmission circuitry in the near future. Separating the analog and digital circuitry into different layers reduces crosstalk problems.

2.4 Power Supply and Data Acquisition System

All circuits were powered by a 3.7 V/250 mAh Lithium Polymer (LIPO) battery. The POKIT Pro (POKIT) system was used for data acquisition, which can be configured as a portable multimeter, oscilloscope, and data logger. To record the signals, the POKIT Pro was configured as an oscilloscope and acquired the signals at a sampling rate of 1 kSa/s. The signals were displayed on a mobile device with Android, and later the data was downloaded in .csv format.

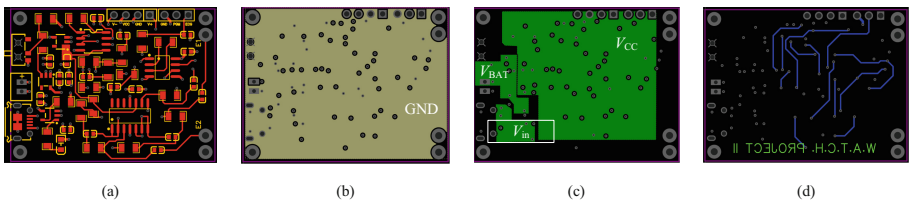


Fig. 3. Design of the 4-layer PCB: a) Top layer, b) Inner layer 1 (Ground Plane), c) Inner layer 2 (Power Plane), and d) Bottom layer.

2.5 Volunteer

To be eligible for participation in the present study, subjects needed to be adults without diagnosed cardiac disease. Individuals of both genders could participate, while pregnant women and children were excluded. Following these criteria, one of the authors of the present work, who provided informed consent, was the test volunteer.

3 Results

The Z_D of the ECG circuit was $3.2 \text{ G}\Omega$, and Z_C was $1.8 \text{ G}\Omega$. Considering the contact impedance of the dry electrodes used, the loading error was lower than 0.002%. The total CMRR was about 80 dB, and the bandwidth was limited between 0.5 Hz and 40 Hz,

with a total gain of 60 dB, approximately. Regarding the cardiac pulse circuit, the total gain was about 50 dB, the CMRR was 70 dB, and the bandwidth was limited between 0.1 Hz and 10 Hz.

The PCB manufactured with all the implemented circuits is shown in Figs. 4a and 4b. The dimensions of the board allowed it to be incorporated into a custom-built case where two dry Ag-AgCl electrodes were also incorporated. In Figs. 4c and 4d, the device worn on the wrist of the volunteer is shown. To detect the ECG, the user must touch electrode 1 using a finger (e.g., index finger) of the hand that does not wear the device. Electrode 2 contacts the skin of the wrist that wears the device; in this way, lead-I is obtained. Figure 4d shows the TMR2001 sensor and the magnet placed on the radial artery, making it possible to detect the disturbance that causes the passage of blood to the localized magnetic field.

The overall consumption was less than 500 μ A. Considering the battery capacity (250 mAh), the autonomy of this system measuring continuously (worst case) would be 500 h. However, this autonomy will be reduced by incorporating the microcontroller and the wireless data transmission system.

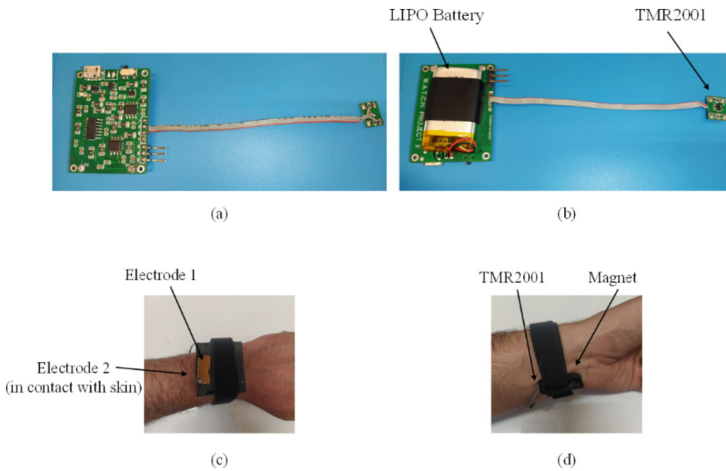


Fig. 4. a) Top layer of the PCB with all the electronic components assembled, b) Bottom layer of the PCB, the LIPO battery, and TMR2001 sensor, c) the custom-built case where the PCB was included, d) the TMR2001 sensor and the magnet placed near the radial artery.

Figure 5 shows the signals obtained with the developed device. The upper trace shows the result obtained with the patient simulator. All ECG waves are clearly depicted and without distortion. The middle trace shows the lead-I obtained from the volunteer, where a larger noise contribution is observed. This is due to electrode-skin contact because the skin was not prepared before the test. Even so, the SNR of the signal was higher than 68 dB, and the ECG waves could be distinguished, allowing the estimation of intervals and amplitudes of interest using simple signal processing algorithms. Considering the dimensions of the PCB ground plane, C_{ISO} was estimated according to [15]; in our case, $C_{ISO} \approx 1.4$ pF (Z_{ISO} @ 60Hz ≈ 1.9 G Ω). Assuming $\Delta Z_E = 5$ k Ω , $Z_C = 1.8$ G Ω , CMRR

= 80 dB, and using the values proposed in [16], that is: $C_B \approx 200$ pF ($Z_B @ 60\text{Hz} \approx 13$ M Ω) and $i_p = 100$ nA, from (1), $V_{EMI} = 42.5$ μV when the device was used in an office, which implies a very low contribution from the power line interference. The lower trace displays the cardiac pulse signal obtained by measuring the magnetic disturbance of the blood. The signal had an SNR > 55 dB, where the systolic and diastolic amplitudes and the dicrotic notch are clearly shown.

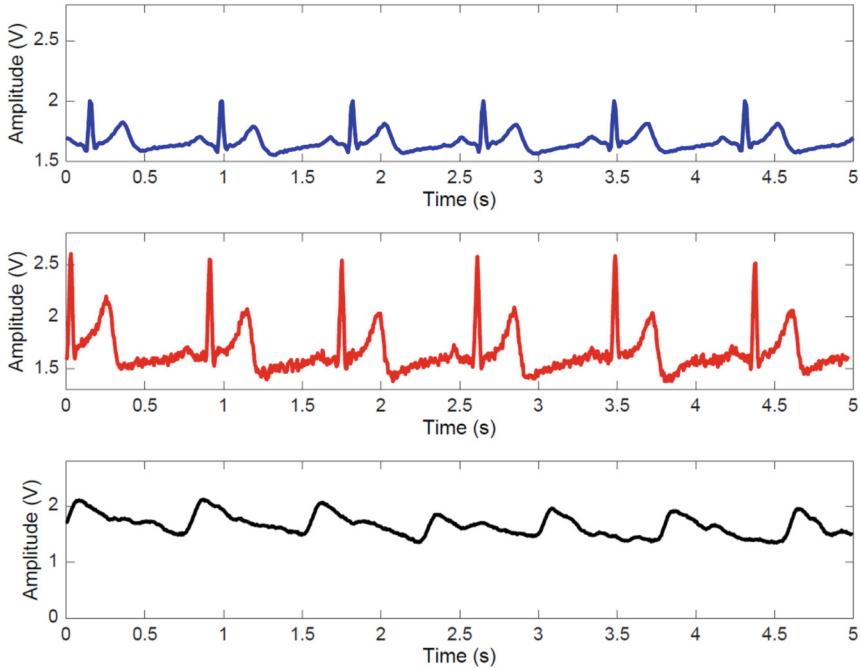


Fig. 5. Cardiac signals detected with the proposed device: From the patient simulator (upper trace), Lead-I ECG from a volunteer (middle trace), and cardiac pulse signal from a volunteer (bottom trace).

4 Conclusions






The preliminary design of a device capable of simultaneously detecting ECG and cardiac pulse was presented. The current consumption of the system was lower than 500 μA , and it was able to detect signals with an SNR greater than 55 dB. Thanks to the dimensions of the electronic circuitry, the interference contribution was lower than 42.5 μV . This suggests that the designed system can be implemented in a wearable device and is suitable for use in regular spaces where electromagnetic interferences are normally present. With this preliminary design, it is possible to develop a wearable system where simple signal processing algorithms can be implemented in such a way that it is possible to assess and track the cardiovascular health of a subject in non-hospital settings.

References

1. Skyrme, T., Dale, S.: IDTECH: Wearable Technology Forecasts 2023–2033. <https://www.idtechex.com/en/research-report/wearable-technology-forecasts-2023-2033/928>
2. Ates, C., Ates, H.C., Nguyen, P.Q., Gonzalez-Macia, L.: et al. End-to-end design of wearable sensors. *Nat. Rev. Mater.* **7**, 887–907 (2022). <https://doi.org/10.1038/s41578-022-00460-x>
3. Min, W., Wu, M., Luo, J.: Wearable technology applications in healthcare: a literature review. *Online J. Nurs. Inform.* **23**(3) (2019). https://www.himss.org/resources/wearable-technology-applications-healthcare-literature-review#_ENREF_30
4. Mannhart, D., et al.: Clinical validation of 5 direct-to-consumer wearable smart devices to detect atrial fibrillation: BASEL wearable study. *clinical Electrophysiology* **9**(2), 232–242 (2023)
5. Qermane, K., Mancha, R.: WHOOP Inc: digital entrepreneurship during the Covid-19 pandemic. *Entrepreneurship Educ. Pedagogy* **4**(3), 500–514 (2021)
6. Knecht, D.S., et al.: Technical characterization of the single-lead electrocardiogram signal from four different smartwatches and its clinical implications. *Europace*, **25**(Suppl 1) (2023)
7. Li, K., Elgalad, A., Cardoso, C., Perin, E.C.: Using the apple watch to record multiple-lead electrocardiograms in detecting myocardial infarction: where are we now? *Tex. Heart Inst. J.* **49**(4), e227845 (2022)
8. Spinelli, E.M., Mayosky, M.A.: Two-electrode biopotential measurements: power line interference analysis. *IEEE Trans. Biomed. Eng.* **52**(8), 1436–1442 (2005)
9. Ebrahimi, Z., Gosselin, B.: Ultra-low power photoplethysmography (PPG) sensors: a methodological review. *IEEE Sens. J.* (2023)
10. Pelaez, E.A., Villegas, E.R.: LED power reduction trade-offs for ambulatory pulse oximetry. In: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2296–2299. IEEE (2007)
11. Phua, C.T., Lissorgues, G., Mercier, B.: Non-invasive acquisition of blood pulse using magnetic disturbance technique. In: 13th International Conference on Biomedical Engineering: ICBME 2008 3–6 December 2008 Singapore, pp. 786–789. Springer Berlin Heidelberg (2009). https://doi.org/10.1007/978-3-540-92841-6_193
12. Méndez-Lira, O.A., Gutiérrez-Chávez, A., Cota-Ruiz, J.D., Díaz-Román, J.D., González-Landaeta, R.E.: Sistema Vestible para la Detección Simultánea y No Invasiva del ECG y el Flujo Sanguíneo. *Revista mexicana de ingeniería biomédica* **39**(3), 249–261 (2018)
13. Pallas-Areny, R., Colominas, J., Rosell, J.: An improved buffer for bioelectric signals. *IEEE Trans. Biomed. Eng.* **36**(4), 490–493 (1989)
14. Casas, O., Spinelli, E.M., Pallas-Areny, R.: Fully differential AC-coupling networks: a comparative study. *IEEE Trans. Instrum. Meas.* **58**(1), 94–98 (2008)
15. Catacora, V.A., Guerrero, F.N., Spinelli, E.M.: Size constraint to limit interference in dnl-free single-ended biopotential measurements. *J. Med. Biol. Eng.* **42**(3), 332–340 (2022)
16. Haberman, M., Cassino, A., Spinelli, E.: Estimation of stray coupling capacitances in biopotential measurements. *Med. Biol. Eng. Comput.* **49**, 1067–1071 (2011)



Heart Monitoring Vest (MAC) for Cats and Dogs

Ana Lorena García Zárate¹  , Anali Elizabeth González Espinosa¹ ,
Christian Roberto Ibáñez Nangüelú¹ , and Norma Celeste Estrada Calvo² 

¹ Faculty of Biomedical Engineering, Universidad Politécnica de Chiapas, Suchiapa,
Chiapas 29150, México

191143@ib.upchiapas.edu.mx

² Veterinary Clinic “Fauna”, Tuxtla Gutiérrez, Chiapas 29076, México

Abstract. Heart failure in dogs and cats is a relatively common disorder that affects animals of all ages. There are different cardiac pathologies that are susceptible to developing this condition, so it is important to identify and intervene quickly [1]. Cardiac arrhythmia occurs when the electrical impulses of the heart do not function properly, resulting in irregular heartbeats that can be too fast (tachycardia) or too slow (bradycardia). In cats, visible symptoms are unlikely to occur, but if present, they may include palpitations, chest pain, fainting, dizziness, and even myocardial infarctions. Therefore, detecting heart problems in cats often happens when the disease is already advanced and may be difficult to treat or require specialized tests that can cause stress to the patient. The aim of this project is to develop a preventive monitoring vest for the physiological variables of cats, as a means to detect potential heart issues before they become severe. It is important to note that our vest does not predict the progression of cardiac arrhythmia treatment. The Heart monitoring vest (MAC) for cats and dogs serves as a tool for detecting an arrhythmia before a regrettable consequence occurs, the vest can be used if the veterinarian detects any anomalies in their physiological variables, with the intention of assisting veterinarians in diagnosing this condition. It should be noted that the vest is designed for cats but can also be used by dogs.

Keywords: Arrhythmia · heart attacks · stress · preventive

1 Introduction

The temperature, pulse, and heart rate in dogs, as well as in cats, are vital to know in order to determine any anomalies in the health of our pet, especially regarding the heart. Without a doubt, continuous monitoring of physiological variables in dogs and cats entails a significant investment of time and money, particularly for the acquisition of devices capable of efficient monitoring. Therefore, this document will discuss the possible complications and the method that will be used as a monitoring tool designed to be used non-invasively in clinical practices in cats and dogs, in order to provide an accessible tool for veterinarians as well as for the owners of these animals. Arrhythmia is one of the most revealing symptoms of heart disease in cats. It is essentially an irregularity in the pattern of heartbeats. Sometimes, detecting an arrhythmia is not easy

because studying the heart rhythm of our pet is not common. For timely detection of a heart disease or heart murmur, it is advisable for the veterinarian to examine the animal's heart once a year.

Cats can also suffer from heart diseases, although it is more difficult to detect in them due to their calm lifestyle and well-known talent for sleeping [2]. The main objective of our project is to create awareness and thus achieve early detection of arrhythmias and prevent regrettable consequences, not only assisting veterinarians but also pet owners involved.

Electrocardiography is a tool used in patients with signs of heart disease to assess the electrical activity of the heart and diagnose cardiac pathologies in a non-invasive manner. Heart diseases are quite common in dogs and cats, particularly in older dogs and cats [3]. Symptoms may include coughing, fatigue, shortness of breath, weight loss, and syncope, which refers to fainting episodes [4]. It is true that the lifespan of our pets is relative, and we cannot magically extend it. However, if we want to enjoy their company during that time, they must be healthy [5]. Therefore, it is important to be alert in order to identify health problems in cats and inform a veterinarian as soon as possible.

2 Methodology

2.1 Materials

- 1) ECG Module AD8232: a module that allows us to collect real electrocardiogram data.
- 2) MLX9014 temperature sensor: used to measure the body temperature of the feline. It is chosen for having a non-contact infrared sensor.
- 3) MAX30102 oxygen saturation sensor: used for blood oxygen saturation measurement (SpO₂), and heart rate (H.R).
- 4) Piezoelectric Sensor: for measuring respiratory rate (R.R.).
- 5) Bluetooth module: Responsible for transmitting the information and it has the following feature: wireless connection with a range of 10 m.
- 6) Thonny Software and Hardware IDE: Compatible with the sensors, Python high-level programming language. Given previous experience, two Raspberry Pi Pico boards are used.
- 7) ECG patches.
- 8) Vest (made of high elastic polyester fabric, lycra, and nylon).

2.2 Methods

The vest was studied with 7 cats under different conditions from April to July 2023 in the city of Tuxtla Gutiérrez, Chiapas. There were a total of 4 male cats and 3 females. One of the cats had a respiratory disease, two cats were anesthetized, one was of advanced age, and two were shaved. The measurements were taken starting from 1 year of age. Additionally, the vest was also tested on 4 dogs of different breeds, 3 males and 1 female. One of the dogs had a cardiac disease. To validate this project, measurements were initially taken from the cats or dogs using biomedical equipment such as an infrared thermometer, vital signs monitor, and pulse oximeter to measure some of the physiological variables of the project. Next, the preparation of the cat or dog was done, including the fitting of

our vest. If the cat was going to be shaved or anesthetized, these circumstances were noted to check if the vest detected changes in the signal. The physiological signals of the cat or dog were acquired, processed, and transmitted via Bluetooth. The signal and measurements were visualized in an Android application called MAC, created specifically for this project. Temperature measurements were recorded at intervals of 10 min, and the mode was calculated to perform a paired T-test correlation (a method used to test if the mean between pairs of measurements is equal to zero [6]) to analyze the obtained error range.

3 Results and Discussion

The Heart monitoring vest (MAC) Fig. 1 and Fig. 2.; is adjustable to fit medium-sized cats up to small-medium-sized dogs. It houses sensors that, through the Android Studio application, allow observation of temperature, heart rate, respiratory rate, and oxygen saturation parameters. The application is integrated with a menu that includes the electronic ID card for the cat or dog (Fig. 3. A), which can be edited later. It also features a medication tracking section and a parameters section that includes an ECG, allowing real-time visualization of the signal on a graph (Fig. 3. B). Additionally, it displays the values obtained on the Raspberry Pi device via Bluetooth. Our application includes an alarm system that activates when the parameters are outside the established range.

In the View of the vest from the inside the following list of components can be observed (Fig. 1):



1. Green ECG lead.
2. Red ECG lead.
3. MAX30102 Sensor.

Fig. 1. View of the vest from the inside.

In the Exterior view of the vest. The following list of components can be observed (Fig. 2):



1. Board O.
2. Board E.
3. Piezoelectric sensor.

Fig. 2. Exterior view of the vest.

Board O contains the circuit for measuring oxygen saturation, while Board E contains the circuit for measuring temperature, respiratory rate, and ECG.

The mobile application designed for the device includes the following (Fig. 3):

- a. View of the main menu of our application
- b. In this image, you can observe the graph of the signal obtained from the ECG and the one processed via Bluetooth

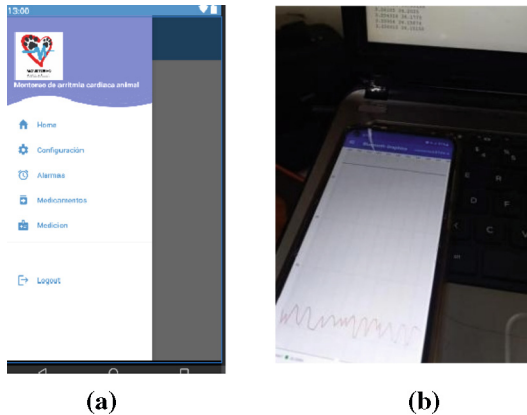


Fig. 3. a) Main menu of the application, b) ECG graph obtained via Bluetooth.

The tests were conducted on 5 cats of different ages and conditions, the data can be seen in Table 1. Additionally, tests were also performed on 4 medium-sized dogs, yielding the following results, the data can be seen in Table 2:

We obtained results indicating that our circuit has a lower range of error when the cat's fur is shaved in the areas of the leads, the data can be seen in Table 3. Additionally, we found that it detects value changes very well when the cat is in a normal or anesthetized.

From Fig. 4. The result of the method yielded an error range greater than $0.07\text{ }^{\circ}\text{C}$ when compared to the measurements of the infrared thermometer. It was concluded that the error range is acceptable, considering that safety standards state that a thermometer has a margin of error of $\pm 1\text{ }^{\circ}\text{C}$.

From Fig. 5., the result of the method yielded an error range greater than 2 rpm when compared to the measurements of a veterinary vital signs monitor used to measure respiratory rate. It was concluded that the error range is acceptable, considering that the accuracy of the vital signs monitor is 1 rpm.

From Fig. 6., the result of the method yielded an error range greater than 2 bpm when compared to the measurements of a pulse oximeter used to measure heart rate. It was concluded that there is room for improvement in this measurement, as most pulse oximeters have a margin of error of 1 bpm.

From Fig. 7, the result of the method yielded an error range greater than 2% when compared to the measurements of a pulse oximeter used to measure oxygen saturation. It was concluded that the error range is acceptable, considering that safety standards state that an oximeter has a margin of error of $0.28 \pm 3.1\%$.

Table 1. Measurement Results in Cats - Part One.

NAME	SALEM ZÁRATE	NUEVE	NUEVA	CERO
AGE	3 years old	9 years old	8 years old	1 years old
SPECIES	Felis silvestris catus	Felis silvestris catus	European domestic cat	Felis silvestris catus
SIZE	Large	Large	Small	Mediumsized
FUR	Low	Shaved	Abundant	Shaved
GENDER	Male	Male	Female	Male
TEMPERATURE 38 °C– 39 °C)	38.5°	38.45°	38.96°	38.69°
ECG	Some noise	Noiseless	Noise	Noiseless
RESPIRATORY RATE (20 A 42 RPM)	61 rpm	30 rpm	42 rpm	34 rpm
HEART RATE (140–200 BPM)	189 bpm	154 bpm	176 bpm	180 bpm
OXYGEN SATURATION (95 A 100%)	96%	92%	98%	96%
CONDITIONS	Cat with feline asthma	Anesthetized and shaved	Normal	Shaved
VEST	Yes	Yes	Yes	Yes

3.1 Discussion

The Heart Monitoring Vest (MAC) in cats and dogs is a functional monitoring system to assess the cardiac status of the patient in real time. It does not predict the progress of cardiac arrhythmia treatment. The obtained results indicate that temperature, pulse, and heart rate in both dogs and cats are limited. This is done with the purpose of detecting any anomalies in the health of our pets, particularly regarding the heart. Therefore, it is of utmost importance to remain vigilant in identifying health problems in our pets. This is why a vest is designed to monitor cardiac arrhythmias in cats using an ECG circuit. To visualize the monitoring of physiological variables such as Oxygen Saturation (SpO₂), Heart Rate (HR), Respiratory Rate (RR), and Temperature (T). All these variables are analyzed and sent in real time via Bluetooth to an Android mobile device. The device features visual alarms to ensure that the patient is not disturbed.

Currently, there are various devices and techniques that assist in monitoring vital signs in animals, such as apnea monitors, 3-channel Holter devices.[8], chest auscultation techniques, rectal thermometers for body temperature readings, pulse oximeters, among others. It is important to mention that these techniques and devices are not very specific or reliable, and are designed solely for the measurement of a physiological variable.

Table 2. Measurement Results in Cats - Part Two.

KITTY CASTILLO	NUEVA	NUEVE	SILVESTRE	MISH
8 YEARS OLD	8 years old	9 years old	2 years old	4 years old
FELINE ANGORA	European domestic cat	Felis silvestris catus	Felis silvestris catus	Feline Maine Coon
SMALL	Small	Large	Small	Large
ABUNDANT	Abundant	Shaved	Low	Abundant
FEMALE	Female	Male	Male	Female
38.96°	36.57°	37.47°	38.02°	38.4°
NOISE	Interferences	Noiseless	Noiseless	Some noise
28 RPM	25 rpm	65 rpm	33 rpm	42 rpm
163 BPM	129 bpm	174 bpm	186 bpm	190 bpm
95%	90%	96%	97%	98%
NORMAL	Anesthetized	Normal and Shaved	Normal	Normal
VEST	Yes	Yes	Yes	Yes

Table 3. Measurement Results in Cats - Part Two.

NAME	POLO	GIS ALBERTO	TOBY ESPINOSA	MOLY
AGE	2 años	5 años	4 años	6 años
SPECIES	Pit bull dog	Dog	Maltese dog	Pug dog
SIZE	Medium-sized	Medium-sized	Medium-sized	Medium-sized
FUR	Low	Low	Abundant	Low
GENDER	Male	Male	Male	Female
TEMPERATURE (37.5 °C- 39 °C)	37.68°	37.97°	36.3°	38°
ECG	Noiseless	Noiseless	Noiseless	Noise
RESPIRATORY RATE (10 A 40 RPM)	35 rpm	39 rpm	39 rpm	22 rpm
HEART RATE (60 - 180 LPM)	175 bpm	173 bpm	182 bpm	80 bpm
OXYGEN SATURATION (90 A 100%)	94%	98%	96%	92%
CONDITIONS	Normal	Normal	Stressed	Heart disease

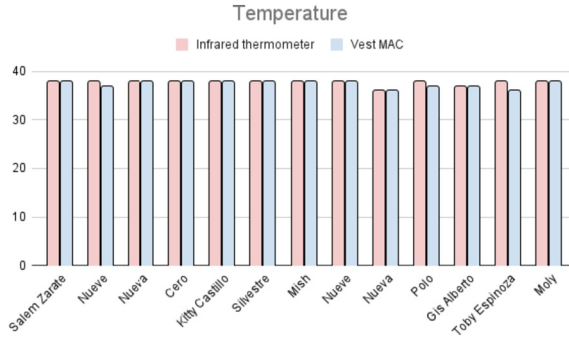


Fig. 4. Paired T-test statistical method results (temperature).

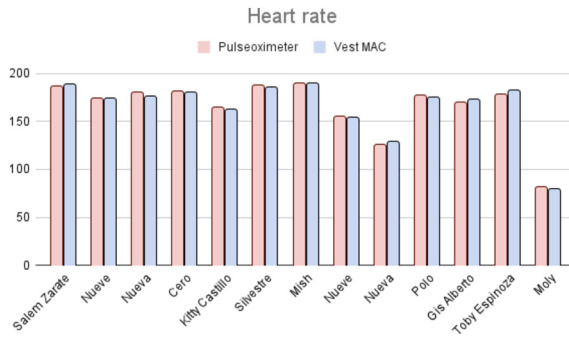


Fig. 5. Paired T-test statistical method results (temperature).

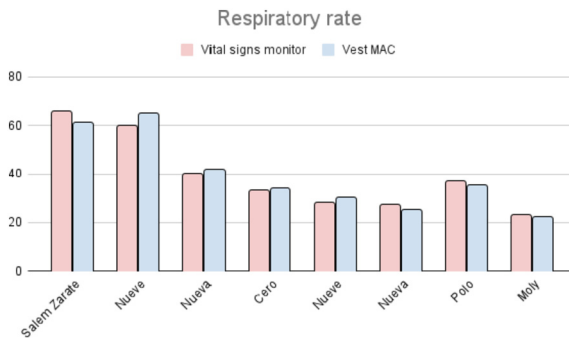


Fig. 6. Paired T-test statistical method results (heart rate).

In addition, there are non-invasive portable devices to acquire physiological signals for canines in real time, where the acquired physiological variables are heart rate, respiratory rate, and temperature. Based on these analyzed signals, the actual physical condition of the animal is determined [9].

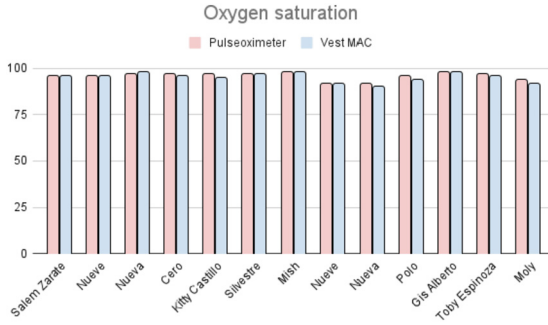


Fig. 7. Paired T-test statistical method results (oxygen saturation).

The advantages of this device compared to other existing devices are that, in addition to being portable, easy to handle and non-invasive, this device helps us monitor multiple physiological parameters in real time and automatically, without the need to use other equipment for measurement. It also alerts us when heart rate and respiratory parameters are outside the normal range. This allows us to know when bradycardia (slow heart rate) or tachycardia (fast heart rate) is occurring.

4 Conclusions

Regularly bringing our pets for monitoring enables the detection of diseases in early stages. This is something we should be aware of if we want to provide optimal quality of life for the animal. Some technical considerations of each component are relevant in the design of each stage in the development of this project, due to the information processing needs for each of the sensors.

Two low-cost and compact Raspberry Pi Pico computers were used, each configured to connect with a specific set of sensors. One of them is configured to obtain SpO2 values, while the other processes information from temperature, respiratory rate, and ECG sensors. Both devices send the gathered information to the mobile device through Bluetooth communication. The acquisition of values. The software used for the design of the mobile application was Android Studio. The vest design includes pockets that house two boxes containing our electronic circuits. Additionally, the vest is made of flexible and compact material. It is equipped with a rechargeable battery, making it portable with a duration of approximately 8 h. This type of battery was chosen for its current storage capacity of 10000 mA.

Some improvements for the vest include adding more piezoelectric sensors to expand the acquisition of information for heart rate and thus provide better results. Due to the instability and slightly higher error range of the heart rate data obtained from the MAX30102 sensor, it is recommended to use another sensor specifically for heart rate measurement.

Acknowledgement. The authors A. Garcia and A. González would like to acknowledge and express their gratitude to BS. Lorena Zarate Díaz, Mrs. Beatriz Espinosa Mazariegos, and Professor Agustín Eredi González Dardon for allowing us to pursue our studies and for the financial support provided to this project.

The authors A. Garcia and A. González would like to thank Mr. Jorge Alberto López Méndez for his time and kind assistance in resolving veterinary-related queries. The authors A. Garcia and A. González would like to express their appreciation to Ms. Norma Celeste Estrada Calvo for her time and helpful guidance in resolving project-related inquiries.

The authors A. Garcia and A. González extend their gratitude to advisor Mr. Christian Roberto Ibáñez Nangüelú for his moral support, time, and kind guidance in resolving project-related queries.

The authors A. Garcia and A. González express their gratitude to Mr. Daniel León Orantes for his support and dedicated time in creating the project's logo.

The authors A. Garcia and A. González are grateful to the staff at "Fauna" Veterinary Clinic for their support, time, and the space provided for conducting project validation tests.

The authors A. Garcia and A. González would like to thank all those individuals who allowed us to conduct validation tests with their pets.

The authors A. Garcia and A. González would like to express their gratitude to Engineer Jorge Jara Jimenez for his guidance in project management and development.






The authors A. Garcia and A. González would like to thank Engineer Enrique Tadeo Santoyo Espinosa for his guidance in project integration.

References

1. Arias, D.O., Rodríguez, R.R., Aprea, A.N.: Métodos complementarios de diagnóstico: Pequeños animales y equinos. Editorial de la Universidad Nacional de La Plata (EDULP), Argentina (2020)
2. La Frecuencia Cardíaca En Gatos. ¿Tu Mascota Tiene Problemas Del corazón?. <https://mascotafiel.com/frecuencia-cardiaca-en-gatos/,2023/06/20>
3. Brložnik, M., Likar, S., Krvavica, A., Avbelj, V., Domanjko Petrič, A.: Sensor corporal inalámbrico para monitorización electrocardiográfica en perros y gatos. *Revista de práctica de animales pequeños* **60**(4), 223–230(2018)
4. Electrocardiografía. <https://www.veterinariavillava.com/la-clinica/electrocardiografia/,2023/06/20>
5. La importancia de llevar a la mascota al veterinario con frecuencia. <https://wakyma.com/blog/importancia-llevar-mascota-ve,2023/06>
6. La Prueba T pareada. https://www.jmp.com/es_mx/statistics-knowledge-portal/t-test/paired-t-test.html#:~:text=%C2%BFQu%C3%A9%20es%20una%20prueba%20t,%20o%20no%20igual%20a%20cer
7. Montes de Oca, G., Rodríguez, G., Rodríguez, A., Almeida, R., González, R., Pina, N.: Analizador de un sistema de monitoreo holter de 3 canales:
8. Excorde 3c (2016)
9. Yerney, W., Ávila, W.A., Carrero, W.M.A., Hurtado, W.: Dispositivo de electrónico invasivo para medir de forma remota señales fisiológicas en caninos. *Revista Colombiana de Tecnologías de Avanzada* **3**, 49–56(2020)



A Low-Cost, IoT-Connected Force Platform for Fall Risk Assessment in Older Adults

Enrique Hernandez-Laredo¹ , Lorena Parra-Rodríguez² ,
Ángel Gabriel Estévez-Pedraza³  , and Rigoberto Martínez-Méndez³ 

¹ Unidad Académica Profesional Tianguistenco, UAEMex, 52640 Santiago Tianguistenco, Estado de México, México

² Research Department, Instituto Nacional de Geriatria, 10200 Mexico City, Mexico

³ Facultad de Ingeniería, UAEMex, Toluca, Estado de México 50100, México

aestevezp@uaemex.mx

Abstract. This article presents the development of an IoT-enabled control and data logging interface for the modified Wii Balance Board (mWBB) v1.0. The objective was to enhance usability and functionality for geriatric balance assessment. Hardware modifications, firmware programming, and a user-friendly web-based interface were employed. Preliminary testing on healthy older subjects demonstrated the feasibility and accuracy of the platform in measuring Center of Pressure (CoP) signals. Preliminary results showed successful integration of IoT capabilities, enabling wireless communication, real-time data transmission, and remote access to balance data. The mWBB v1.0 represents a cost-effective and accessible solution for balance assessment in elderly populations. The developed interface allows professionals to easily register subject data, initiate tests, and visualize comprehensive results. The platform's IoT integration expands its potential applications in research, clinical practice, and IoT-driven healthcare systems. Future directions include exploring additional balance metrics, investigating different populations, and integrating advanced data analysis techniques, with the intent to further analyze geriatric results in subsequent research. The mWBB v1.0 and its interface provide a valuable tool for enhancing balance assessment in the elderly, offering opportunities for remote monitoring and personalized interventions.

Keywords: IoT-enabled interface · Geriatric balance assessment · Center of Pressure (CoP)

1 Introduction

Confronting global aging trends, the risk of falls among the elderly has emerged as a significant public health concern. According to the World Health Organization (WHO), adults over 65 years old experience the highest number of fatal falls. As it is estimated that over 30% of seniors living in the community experienced a fall, and this proportion increases to 50% for those in residential care facilities [1]. Moreover, the population pyramid is gradually inverting worldwide, with countries like Japan, Italy, and Germany

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024

J. de J. A. Flores Cuautle et al. (Eds.): CNIB 2023, IFMBE Proceedings 96, pp. 374–385, 2024.

https://doi.org/10.1007/978-3-031-46933-6_39

already having more than 20% of their population aged 65 or over [1, 2]. In Mexico, the population growth rate in that age group is expected to increase. Although the country currently has a predominantly young population, data project that by 2030 the proportion of older people will increase to 14.9%. A complete transition to an aging population is expected, representing approximately 25% of the entire population at older ages [3]. These statistics underscore the growing need to develop research focused on the diagnosis and care of the problem of elderly falls.

A fall significantly reduces a person's quality of life, shortens their lifespan, and impacts not only their physical but also their mental health. The ripple effect extends to caregivers, often family members, affecting their mental health as well [4]. Indeed, the economic burden of managing falls is high, with the US alone spending approximately \$50 billion annually [5].

Determining fall risk and implementing preventive measures could mitigate this public health problem. Interventions can focus on improving the skills and senses needed to prevent falls, such as muscle mass, vision, and balance [4]. On the other hand, current clinical methods for assessing fall risk, which rely on functional tests of gait, strength, balance, posture, physical examinations, and questionnaires, are often time-consuming, require the intervention of specialists and expensive equipment, and people tend to show resistance to taking preventive measures due to lack of time, knowledge, or resources [6]. In addition, the accuracy, performance, and predictive validity of these instruments have been questioned due to their lack of precision [7].

In addition to the clinical methods, numerous studies have led to various technological attempts to assess the risk of falling. These include image analysis techniques, wearable inertial sensors, and force platforms, the latter being considered as the golden standard. However, these solutions often come with drawbacks like high costs, complex setup and usage, and privacy concerns, limiting their widespread adoption [8].

In this context, a modified Wii Balance Board (mWBB) was proposed as a reliable and low-cost force platform. The initial version (mWBB v0.0) although effective, required numerous improvements, particularly in terms of usability and data acquisition. The specifics of mWBB v0.0, its design, validation, and application in balance assessment in the elderly population are detailed in previous work [9–11].

In this paper, a revised version, the mWBB v1.0, is proposed to overcome these challenges. Focusing primarily on the design, construction, and preliminary testing, it provides a platform for subsequent in-depth geriatric analysis. The device is designed to be user-friendly, requiring the individual to simply stand on it like a bathroom scale. The current version measures the Center of Pressure (CoP), calculates several indices according to Prieto [12], and applies a statistical model [11] to provides a balance alteration assessment in just 60 s. Moreover, mWBB v1.0 is enhanced with Internet of Things (IoT) capabilities, allowing it to send the results to a national database linked to the National Institute of Geriatrics, thus contributing to the integration of a comprehensive national information source for future specialized analyses.

2 Methodology

The methodology follows a Design Science Research (DSR) approach and is divided into two main sections. The first one is “Design Criteria”, which explains the reasoning behind the design choices made for the development of the mWBB v1.0, focusing on the requirements to improve the data acquisition and transmission capabilities. The second one is “Integration process”, where the specific changes made to the electronic components of the mWBB v1.0 and programming logic are detailed.

2.1 Design Criteria

The design process for the mWBB v1.0 was heavily influenced by professional healthcare feedback and the areas of technical improvement of the mWBB v0.0. The primary requirements were derived from the need to gather the data more effectively, transmit it to a database, and increase the user-friendliness of the device. The mWBB v0.0 had incorporated a 16×2 LCD display to show the test results. However, considering the new IoT capabilities, this was deemed unnecessary as the device could now connect to any smart device like a computer or a smartphone, from which the application could be controlled, and results could be viewed. Detailed descriptions of the mWBB v0.0, including its design and validation, can be found in [9–11].

2.2 Integration Process

Hardware Modifications

The hardware redesign of mWBB v1.0 involved both the removal of components and the establishment of direct communication with the analog-digital converters (ADCs). Figure 1 Shows the diagram depicts the specific changes made to the original Wii Balance Board and how the different electronics components are interconnected in mWBB v1.0, providing a clearer view of the system’s improved architecture.

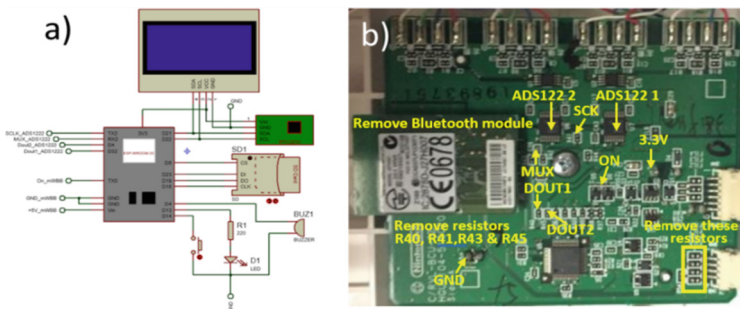


Fig. 1. (Left, a) The electronic diagram of the redesigned mWBB v1.0 system, showing the connections between the different components. (Right, b) A photograph indicating the modifications made to the original Wii platform PCB

The specific changes in the mWBB components from version 0.0 to version 1.0 are itemized in Table 1. Each row corresponds to a different component of the mWBB, and the columns compare the components of the two versions.

Table 1. Comparison of components between mWBB version 0.0 and mWBB version 1.0

Component	Version 0.0	Version 1.0
Sensor platform	Wii Balance Board®	Wii Balance Board®
Display for messages	Liquid Crystal Display (LCD), 16 × 2 characters	OLED technology, 128 x 64 pixels
Voice module and speaker	APR9600 Voice Integrated Circuit and buzzer	Only buzzer
Microcontroller	dsPIC30F6014A (8 kB RAM)	ESP32 DevKitV1 (520 kB RAM)
Memory module	MicroSD® Adapter	MicroSD® Adapter
Internal time control	DS1307 RTC Module	3231 RTC Module
Interface buttons	3 push buttons for time and date adjustment	Removed
LED indicators (state of an evaluation)	RGB LED	Removed

A comparative illustration of the mWBB v0.0 and the redesigned mWBB v1.0 with the changes is provided in Fig. 2. The new design emphasizes a minimalist hardware architecture with essential peripherals for a more streamlined, user-friendly experience.

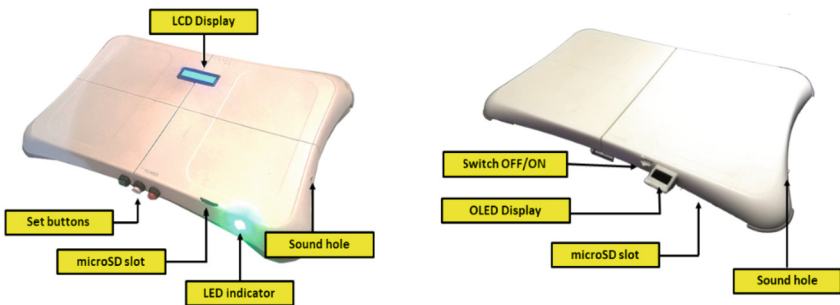


Fig. 2. Comparison between the mWBB v0.0 (Left), and the mWBB v1.0 (Right)

Firmware

The firmware section outlines the key elements of the programming executed on the ESP32 microcontroller for the mWBB v1.0. The firmware was developed using Visual Studio Code (version 1.73.0) with the PlatformIO extension (Core 6.1.5, Home

3.4.3) for the Espressif ESP32 Dev Module utilizing the 5.2.0 framework. Figure 3. Shows a schematic representation of the organization and partitioning of the ESP32 microcontroller firmware.

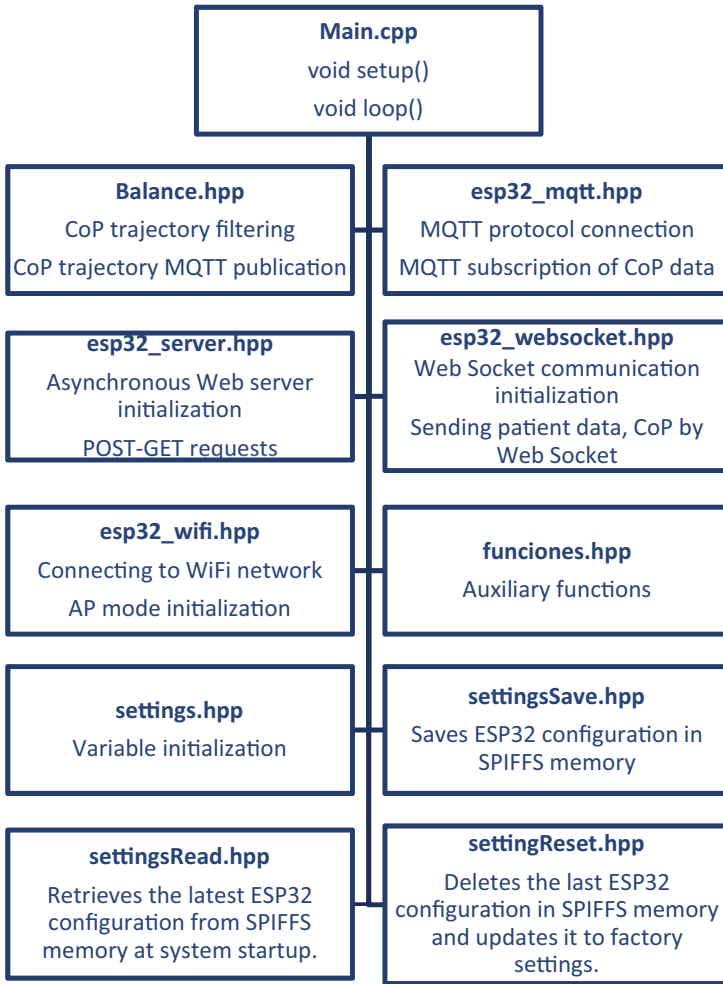


Fig. 3. Organization and partitioning of the mWBB v1.0 firmware

The primary program (main.cpp) is divided into two sections. The first is the “void setup ()”, where the communication protocols, the Serial Peripheral Interface Flash File System (SPIFFS), the peripheral modules, and the calibration for the mWBB v1.0 sensors are initialized. The second section, “void loop ()”, maintains the Wi-Fi connection, updates user and CoP metrics data, and keeps the mWBB v1.0 device data current on the web server via the WebSocket protocol.

To conduct a balance test, a user request is made via the interface. If the user’s weight (over 20 kg threshold) is detected on the platform, the ESP32 starts recording sensor

information at a consistent 50 Hz sampling rate for 2 min. Once complete, an alarm is triggered until the user steps off the platform. The signal from the mWBB v1.0 sensors is processed to calculate the CoP trajectories according to [13], filtered by an Infinite Impulse Response (IIR) digital filter of order 7 Butterworth topology with 5 Hz cutoff frequency, and stored locally on a MicroSD card (two separate.txt files are created for each balance test conducted, one for raw data and one for filtered CoP data) and then sent to the server, which returns the results of the 78 CoP metrics reporting by [12] to the ESP32 for display on the interface.

The ESP32 also handles user requests, such as initiating a balance test, Wi-Fi setup, MQTT server connection, and changing user/password and device identifier, using its second core. All these requests are executed via POST and GET methods.

Designed with user interface upgrades, a broader range of communication protocols, and expanded CoP metrics, the ESP32's firmware caters to enhanced performance and usability compared to its predecessor.

User Interface

The user interface consists of web pages stored in the ESP32's SPIFFS memory. These pages, written in HTML, CSS, and JavaScript, are retrieved by the ESP32's asynchronous server and displayed on Wi-Fi-connected devices. Figure 4. Illustrates the communication logic between the ESP32 microcontroller and its user interface, as well as the files involved for the main actions.

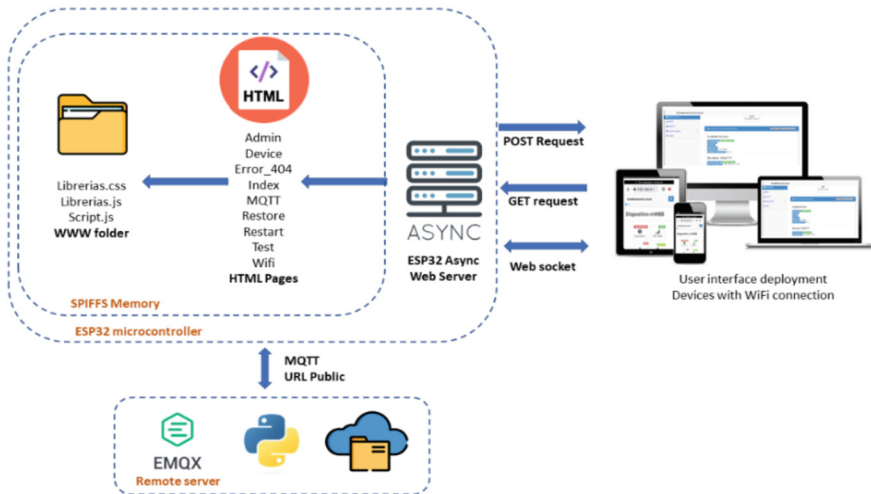


Fig. 4. Connections between web files and user interface

A JavaScript file (scripts.js) handles the logic and WebSocket communication between the web pages and the server. The web page “test.html” allows healthcare professionals to register patient data, download test history, initiate balance test and view its results as CoP metrics and statokinesigrams. The “index.html” page provides device information and displays the status of Wi-Fi and MQTT connections. The “wifi.html”

page simplifies Wi-Fi configuration, enabling easy network connection or “Access Point mode” for local data storage. The “mqtt.html” page manages the MQTT connection, allowing users to control the service. The “manage.html” page provides access to device information, including serial numbers, and allows for resetting and changing user credentials. In case of accessing a non-existing web address, the “404.html” page is displayed, with a return button to navigate back to “index.html”.

Server Connection and Configurations

This section presents the data architecture and data pipeline between mWBB v1.0 and remote server. The microcontroller ESP32 acts as the data flow manager, initiating the asynchronous web server upon device activation. Bilateral communication via Web-Socket protocol allows the ESP32 to receive sensor data, perform balance tests, and update the user interface. An overview of the data architecture and flow is depicted in Fig. 5.

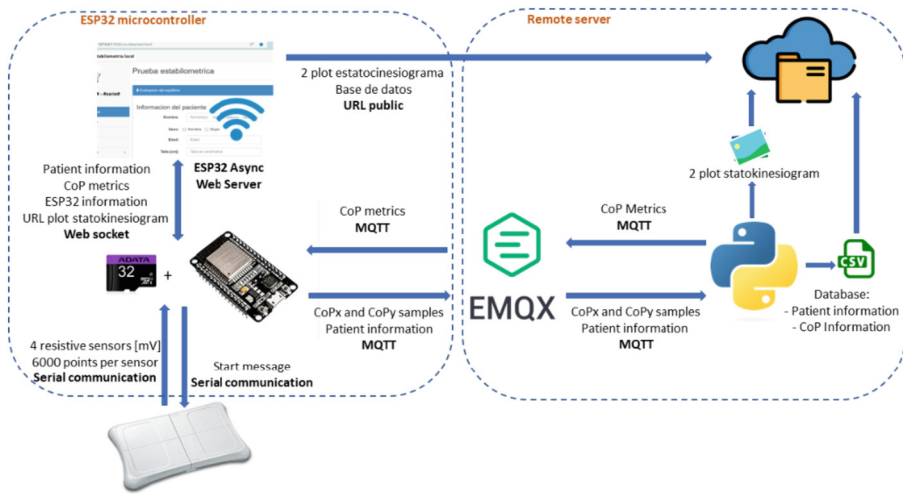


Fig. 5. Connections between web files and user interface

As previously stated, once a balance test is completed, the ESP32 filters and sends the sensor signals to the remote server through MQTT protocol. A Python script running on the remote server processes the data, calculating the 78 metrics of the CoP [12] and generating statokinesigram plot images. The server manages a database containing person information, CoP metrics, and CoP trajectory data.

The communication between the ESP32 and the Python script is facilitated by the EMQX MQTT broker (www.emqx.io), an open-source solution. To configure the server, the EMQX broker needs to be installed and executed on the remote server.

The remote server is based on DELL PowerEdge M630 equipment with LINUX UBUNTU 20.04 operating system, which is located at National Institute of Geriatrics. It hosts the necessary files for the interface to function, including backend system and MQTT Broker.

Accessing the mWBB v1.0 control interface is possible from any location with a stable internet connection. The server setup allows users to interact with the interface, view test results, and manage device configurations.

3 Experimental Results

Preliminary testing of the mWBB v1.0 was conducted on four healthy older subjects to assess its functionality and evaluate the obtained signals. Each subject performed a balance test using the platform. Figure 6. Shows an image of one of the subjects using the mWBB v1.0 platform.



Fig. 6. Subject performing a balance test using the mWBB v1.0 platform

During the test, the platform captures signals from its sensors, which are processed to calculate CoP trajectory data in the medial-lateral plane (CoPx) and anteroposterior plane (CoPy) directions, then filtered and stored in separate txt files, and finally, 78 CoP metrics are calculated based on temporal, frequency, and hybrid measurements [12]. The literature shows that some of these metrics are predictors of falls [11, 14] and are also related to recurrent falls [15], intrinsic factors of falls [16] and the Berg Balance Scale [17]. All the CoP indicators are displayed in the user interface together with the statokinesiograms, which are graphs generated from the correspondence rule of the CoP displacement values, as shown in the Fig. 7. This figure illustrates a comparative analysis between two balance tests for the same subject, one with open eyes and the other with closed eyes. The graphic on the right displays a greater displacement of CoPx and a similar displacement for CoPy, indicative of poorer balance control with eyes closed. As seen from this analysis, the subject's balance becomes more unsteady when the eyes are closed, consistent with expectations. This observation is instrumental in understanding

the patient's ability to maintain balance under different sensory conditions, and it serves as an assessment tool for potential balance alteration and fall risk.

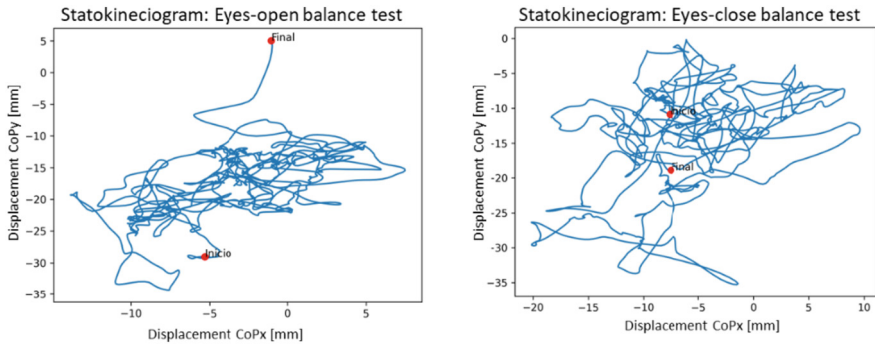


Fig. 7. Signals obtained in a balance test using the mWBB v1.0 platform. The larger displacement in CoPx with eyes closed, as compared to eyes open, indicates poorer balance control in the tested subject.

In accordance with Prospeni's work [18] and based on the information shown in the statokinesiograms, the person evaluated presents balance alteration and fall risk, since their CoP displacements excursion with values greater than 10 mm. These preliminary tests showed the ability of the mWBB v1.0 to measure balance and assess fall risk. And based on the reliability, ease of use and validation reported for the mWBB v0.0 [9–11], it is expected that the signals obtained can provide valuable information for analysis and validation of the current version device performance to be used as a promising longitudinal balance assessment tool in older adults.

4 Discussion

The Wii Balance Board is a system designed for entertainment purposes that was introduced in 2010 to the biomedical field for the evaluation and training of human balance [19], mainly because of its ability to measure CoP at low cost compared to laboratory-grade force platforms. However, its main limitations were unstable sampling frequency (time jitter) and data loss, which detracts from signal quality [20]. In response to this technical problem, the mWBB v0.0 version was developed, which incorporated hardware adaptations [10] that allowed a stable sampling frequency at 50Hz, while achieving a valid and reliable balance evaluation [9, 10]. In addition, the mWBB v1.0 also includes a user interface and IoT connectivity to improve usability, increase its processing and storage capabilities.

In order to better understand the technical specifications of the newly redesigned mWBB v1.0 system, Table 2 presents a detailed overview of the device's general technical characteristics. These specifications provide important insight into the functionality, operation, and capacity of the mWBB v1.0 system.

The cost of the mWBB v1.0 has been reduced by 20% compared to the v0.0 prototype [10], thanks to a series of strategic modifications. The primary factor contributing

Table 2. General technical characteristics of the mWBB v1.0 system

Characteristic	Specification
Weight (including battery)	3.5 kg
Dimensions	30 cm × 48 cm × 7.9 cm
Measurement range of CoP displacement	142.2 mm (Mid-lateral plane), 136.6 mm (Anteroposterior plane)
CoP indices calculated	78
Maximum supported weight	150 kg
Storage	MicroSD® class 10 card of 4 GB - 64 GB
Sampling frequency	50 Hz
Power supply	4 AA batteries
Time for stabilometric evaluation	2 min (divided into 2 tests of 1 min separated by 10 s, it can be used to perform tests in different conditions such as open vs. closed eyes.)
Resolution	1/100 mm
Autonomy in power supply	48 h (constant use)

to this cost reduction is the adoption of a more affordable ESP32 microcontroller, coupled with the exclusion of peripherals like the APR9600 Voice Integrated Circuit. As a result, the mWBB v1.0 is now priced at approximately \$30 USD, and this cost could be further diminished through mass production or by utilizing an external manufacturing base, independent of Nintendo®. This proposal presents a highly promising and cost-effective IoT solution for evaluating balance in older adults, particularly when compared to commercial platforms that carry exorbitant price tags upwards of \$2,900 USD [21].

While the technical specifications and cost-efficiency of the mWBB v1.0 are essential, its primary application lies in evaluating the risk of falls and balance alteration among the elderly population. The stable sampling frequency, precise CoP measurement, and enhanced usability make it an effective tool for assessing balance control, a key factor in fall risk. Coupled with its low cost, the mWBB v1.0 offers a practical solution for widespread screening and prevention, particularly in resource-limited settings. Future research will further explore its clinical efficacy in various populations and settings, potentially integrating it with comprehensive fall prevention programs.

5 Conclusion

In this article, we presented the development of the mWBB v1.0 control and data logging interface, transforming it into an IoT device. The integration of these improves the usability, functionality and connectivity of the platform.

The mWBB v1.0 offers a user-friendly interface and wireless communication, allowing for remote monitoring and data analysis. Preliminary testing on healthy subjects

demonstrated its feasibility and accurate measurement of CoP signals. This work contributes to the field of geriatric balance assessment by providing an accessible and cost-effective solution that leverages IoT technology.

The unique design and IoT capabilities of mWBB v1.0 position it as a promising tool in the field of geriatric balance assessment, with potential applications in fall risk evaluation and prevention. Through accurate CoP signal measurement and the possibility of real-time data transmission and analysis, it can offer healthcare providers and researchers deeper insights for developing personalized interventions that may proactively address fall risks in the elderly. Its affordability facilitates broader adoption, which could positively influence how fall risks are managed across various healthcare settings.

The mWBB v1.0 opens up new opportunities for remote monitoring, longitudinal and cross-sectional studies, and personalized interventions in balance assessment. Its modular design and IoT capabilities enable real-time data transmission, cloud-based storage, and remote access to balance data.

Future research can explore new balance metrics, investigate different population groups, and integrate advanced data analysis techniques using the mWBB v1.0. Its versatility and adaptability make it a valuable tool for research, clinical practice, and IoT applications.

In conclusion, the mWBB v1.0 control and data logging interface, as an IoT-enabled device, represents a significant advancement in geriatric balance assessment. Its user-friendly interface, wireless connectivity, and cost-effectiveness make it a valuable tool for improving balance assessment and management in elderly populations.












References

1. Kalache, A., Fu, D., Yoshida, S., Al-Faisal, W.: *Beattie: World Health Organisation Global Report on Falls Prevention in Older Age*, 1st edn. World Health Organisation, Geneva (2007)
2. Nader, S., Niloofer, D., Melika, A., Shamarina, S., Masoud, M.: Global prevalence of falls in the older adults: a comprehensive systematic review and meta-analysis. *J. Orthop. Surg. Res.* **17**(1), 334 (2022)
3. Gonzalez, K.: *La situación demográfica de México 2015*. 1st edn. Consejo Nacional de Población, México (2015)
4. Vaishya, R., Vaish, A.: falls in older adults are serious. *Indian J Orthop. falls in older adults are serious.* *Indian J. Orthop.* **54**(1), 69–74 (2020)
5. Cost of Falls Center for Disease Control and Prevention Page. <https://www.cdc.gov/falls/data/fall-cost.html>. Accessed 24 June 2023
6. Park, S.H.: Tools for assessing fall risk in the elderly: a systematic review and meta-analysis. *Aging Clin. Exp. Res.* **30**(1), 1–16 (2018)
7. Gallouj, K., Altintas, E.: I remember the fall: memory of falls in older adults. *Clin. Gerontol.* **6**(1), 1–9 (2020)
8. Panjan, A., Sarabon, N.: Review of methods for the evaluation of human body balance. *Sport Sci. Rev.* **19**(6), 131–163 (2010)
9. Estévez-Pedraza, Á.G., Hernandez-Laredo, E., Millan-Guadarrama, M.E., Martínez-Méndez, R., Carrillo-Vega, M.F., Parra-Rodríguez, L.: Reliability and usability analysis of an embedded system capable of evaluating balance in elderly populations based on a modified wii balance board. *Int. J. Environ. Res. Public Health* **19**(17), 11026 (2022)

10. Estévez-Pedraza, Á.G., Martínez-Méndez, R., Portillo-Rodríguez, O., Parra-Rodríguez, L.: Portable device for the measurement and assessment of the human equilibrium. *Ann. Biomed. Eng.* **49**(2), 933–945 (2020)
11. Estévez-Pedraza, Á.G., Martínez-Méndez, R., Portillo-Rodríguez, O., Parra-Rodríguez, L.: A novel model to quantify balance alterations in older adults based on the center of pressure (CoP) measurements with a cross-sectional study. *PLoS ONE* **16**(8), e0256129 (2021)
12. Prieto, T.E., Myklebust, J.B., Hoffmann, R.G., Lovett, E.G., Myklebust, B.M.: Measures of postural steadiness: differences between healthy young and elderly adults. *IEEE Trans. Biomed. Eng.* **43**(9), 956–966 (1996)
13. Leach, J.M., Mancini, M., Peterka, R.J., Hayes, T.L., Horak, F.B.: Validating and calibrating the Nintendo Wii balance board to derive reliable center of pressure measures. *Sensors*. **14**(10), 18244–18267 (2014)
14. Thapa, P.B., Gideon, P., Brockman, K.G., Fought, R.L., Ray, W.A.: Clinical and biomechanical measures of balance as fall predictors in ambulatory nursing homeresidents. *J. Gerontol. A Biol. Sci. Med. Sci.* **51**(5), 39–246 (1996)
15. Stel, V.S., Smit, J.H., Pluijm, S.M., Lips, P.: Balance and mobility performance as treatable risk factors for recurrent falling in older persons. *J. Clin. Epidemiol.* **56**(7), 659–668 (2003)
16. Pajala, S., Era, P., Koskenvuo, M., Kaprio, J., Törmäkangas, T., Rantanen, T.: Force platform-balance measures as predictors of indoor and outdoor falls in community-dwelling women aged 63–76 years. *J. Gerontol. A Biol. Sci. Med. Sci.* **63**(2), 171–178 (2008)
17. Berg, K.O., Maki, B.E., Williams, J.I., Holliday, P.J., Wood-Dauphinee, S.L.: Clinical and laboratory measures of postural balance in an elderly population. *Arch. Phys. Med. Rehabil.* **73**(11), 1073–1080 (1992)
18. Prosperini, L., Fortuna, D., Giannì, C., Leonardi, L., Pozzilli, C.: The diagnostic accuracy of static posturography in predicting accidental falls in people with multiple sclerosis. *Neurorehabilit. Neural Repair.* **27**(1), 45–52 (2012)
19. Clark, R.A., Mentiplay, B.F., Pua, Y.H., Bower, K.J.: Reliability and validity of the wii balance board for assessment of standing balance: a systematic review. *Gait Posture* **61**(1), 40–54 (2018)
20. Goble, D.J., Cone, B.L., Fling, B.W.: Using the Wii Fit as a tool for balance assessment and neurorehabilitation: the first half decade of wii-search. *J. Neuroeng. Rehabil.* **11**(1), 1–9 (2014)
21. BIOSIGNALSPLUX Force Platform Page. <https://www.pluxbiosignals.com/products/force-platform>. Accessed 28 June 2023



Multifrequency Bioimpedance Characterization of DNA Concentration: *Sensitivity Analysis*

Nancy Nallely Gómez-González¹ , Catalina Montoya-Solórzano² ,
Leticia Arias-González³ , Liliana Becerra-Martínez⁴ ,
Jaime Eduardo López-Madrigal³ , Virginia Sánchez-Monroy³ ,
Jacobo Esteban Munguía-Cervantes⁵ , Nadia Mabel Pérez-Vielma⁶ ,
Modesto Gómez-López³ , and César Antonio González-Díaz³  

¹ Universidad de Guadalajara, 44100 Guadalajara, Jalisco, Mexico

² Universidad Autónoma de Manizales, 170001 Manizales, Caldas, Colombia

³ Escuela Superior de Medicina, Instituto Politécnico Nacional,
11320 Ciudad de México, Mexico

cgonzalezd@ipn.mx

⁴ Universidad Veracruzana, 91090 Xalapa, Veracruz, Mexico

⁵ Centro de Nanociencias y Micro y Nanotecnologías, Instituto Politécnico Nacional,
07369 Ciudad de México, Mexico

⁶ Centro Interdisciplinario de Ciencias de la Salud Unidad de Santo Tomas, Instituto Politécnico Nacional, 11320 Ciudad de México, Mexico

Abstract. In the past few decades, advancements have been made in the development of biosensors for the detection and analysis of Deoxyribonucleic Acid (DNA) and genes. The objective of this study was to explore Electrical Bioimpedance Spectroscopy (EBiS) for DNA concentration. EBiS patterns in the complex plane for different concentrations of salmon sperm DNA were obtained, and analyzed to identify potential sensitivity parameters for characterization and calibration curves design. According to the results, characteristic EBiS patterns are associated to the analyte vehicle and its parameters change as a function of the present DNA concentration. It seems possible to determine DNA concentration from measuring reactance (X) and resistance (R) at frequencies centered around 200 and 5 kHz, respectively. Exploration of X and R parameters at such frequencies seems pertinent to design calibration curves to estimate DNA concentration of 73 to 1200 ng/ μ L in a sample. Further studies are warranted to confirm the observations.

Keywords: Bioimpedance · DNA · Biosensor

1 Introduction

The need for a simple and inexpensive technique for DNA quantification and gene detection, especially for less developed countries, is convenient and highly useful. Use cheaper tools for DNA analysis is a task in which work worldwide continues, several biosensors have been developed to examine various analytes, the size of which includes individual

ions, small molecules, nucleic acids, proteins, and entire viruses or bacteria [1]. Lab on a chip technology is an important challenge to create minimal systems for label free DNA quantification. Multifrequency bioimpedance measurements such as Electrical Bioimpedance Spectroscopy (EBiS) is a minimal invasive and simple technique to explore electrical properties of biological materials, and have shown potential value as an emerging technique to detect gene expression as well as total DNA quantification [2].

DNA molecule has a natural electronegative charge condition [3], thus DNA concentration could be associated intrinsically to its electrical properties expressed as EBiS measurements. Studies indicate a relationship of DNA concentration with the magnitude and phase of EBiS in a given frequency bandwidth for the detection of the presence of DNA in the content of a sample. Where currents injected at multifrequency in the region of lesser dispersion of the gamma band generate a relaxation effect of the dipole in a water molecule allowing the interaction with a DNA molecule due to its electronegative charge. The described interaction behaves as a function of DNA concentration and can be correlated with changes in volumetric impedance, being of great importance for DNA detection [4, 5]. Sensitivity observed might depend on the genetic material concentration as well as its base pair length (molecular weight) [6].

In this work multifrequency bioimpedance patterns in the complex plane are used to explore DNA concentration, and to identify sensitivity parameters for further calibration curves design, this in order to determine the viability of the development of a DNA biosensor device by means of EBiS measurements. With this new development, a saving of time, reduction of expensive equipment and expenses of consumables is foreseen.

2 Methodology

2.1 Experimental Design

EBiS spectra were measured in five total DNA samples at different concentrations, as well as in Milli-Q water as vehicle reference. Resistance (R) and reactance (X) results were obtained at every frequency and Nyquist plots were drawn. An electrophoresis assay was developed for qualitative characterization of DNA concentrations, in addition, DNA samples concentrations were quantified through UV-Vis Spectrophotometry. Every measurement was developed in triplicate. The block diagram of the experimental design is shown in Fig. 1.

2.2 DNA Procedure

Samples were prepared from commercially available Deoxyribonucleic acid sodium salt, from salmon sperm, Sigma-Aldrich, reference: D1626, No. Cat: 438545-06-3 (Sigma-Aldrich Quimica S de RL de CV, Toluca, México). A master solution was prepared by dissolving the solute Milli-Q water (Milli-Q® IQ 7003/7005 Water Purification Systems) at initial concentration adjusted to 1200 ng/ μ L, then; lineal dilutions were obtained in a final volume of 50 μ L (1:2, 1:4, 1:8 and 1:16). A qualitative test was performed to assess the presence of DNA at each concentration by electrophoresis using an agarose gel, then, final DNA concentrations were measured by the use of the NanoDrop Spectrophotometer

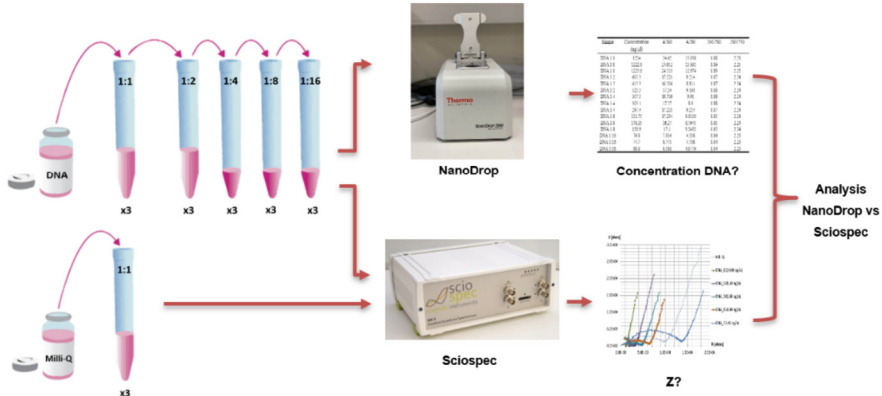


Fig. 1. Block diagram of the experimental design, assays were developed in triplicate.

2000 (Thermo Fisher Scientific - Waltham, MA USA) in triplicate assays. The 260/280 optical density rate was used to determine the purity and concentration of DNA, which was within acceptable values (1.8 to 2.0) [7]. Table 1 shows the initial and diluted DNA concentration in ng/uL as well as purity ratios 260/280.

2.3 Bioimpedance Measurements

A system to measure EBiS in 2 uL volume samples was implemented through the use 10x10 mm gold film microelectrodes in an interdigitated array (200 um line width and 50 um spacing), and embedded on a glass surface (Fig. 2). EBiS measurements were developed by the use of the bioimpedance analyzer Sciospec ISX-3 (Sciospec Scientific Instruments GmbH, Leipziger, Bennewitz, Germany) in a frequency bandwidth from 100 Hz to 10 MHz at 126 steps distributed on a logarithmic scale with a maximum

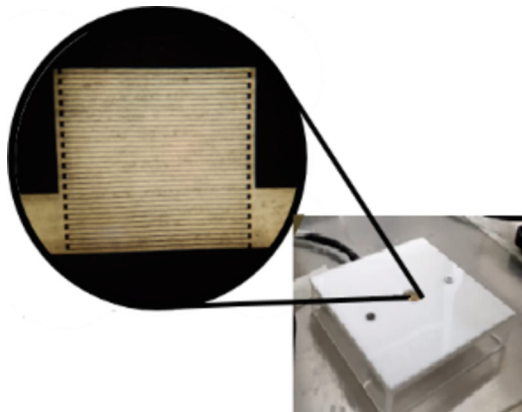


Fig. 2. Interdigitated array microelectrodes consisting of a gold film embedded on a glass surface. EBiS measurements were performed on 2 μL DNA drop.

voltage of 100 mv. ScioSpec programming and data storage was done using a computer (HP mini 110-1150LA PC, HP Inc.).

3 Results

3.1 DNA Dilutions

Figure 3 shows the electrophoresis assay for the five DNA concentrations. Qualitative decrease in fluorescence as a function of DNA dilution is evident. Table 1 shows the DNA concentration values in $\text{ng}/\mu\text{L}$ and the absorbance rates found for each dilution. The purity of the DNA was confirmed by the optical density ratio 260/280 in the range greater than 1.8 nm.

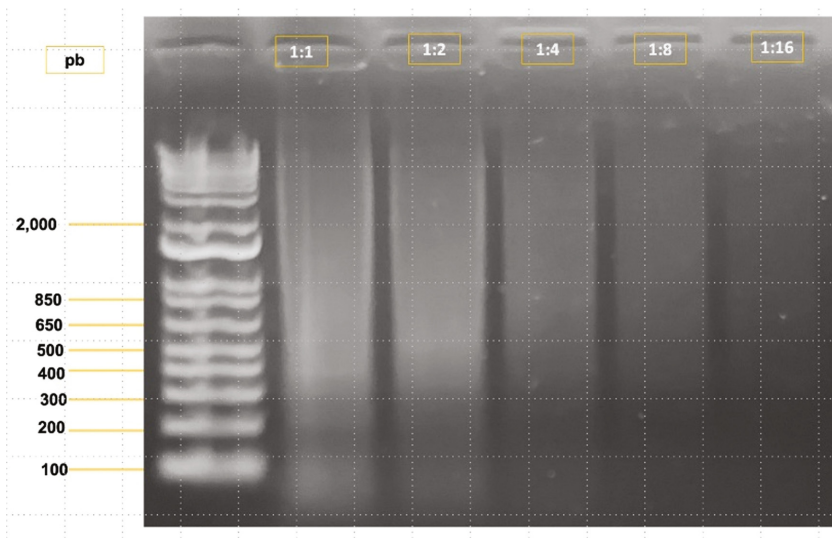


Fig. 3. Electrophoresis assay for five DNA concentrations. Diminished fluorescence as a function of DNA dilution is evident. Figure 3 Electrophoresis assay for five DNA concentrations. Diminished fluorescence as a function of DNA dilution is evident. The base pairs (bp) can be seen on the left side.

3.2 Bioimpedance Spectra

Figure 4 shows the Nyquist plot for DNA dilutions and Milli-Q water as contrast. A relationship between changes of the EBiS patterns in the complex plane and the concentration of DNA is evidenced. Milli-Q water curve has similar pattern for DNA samples.

Table 1. Final concentrations in ng/ μ L and absorbance index for five DNA dilutions samples. Data are shown in mean values and standard deviation (triplicate assays).

Sample	Concentration	A260 mean	A280 mean	260/280 mean
	mean (ng/ μ L)			
DNA 1:1	1224.06	24.60	13.00	1.89
<i>StdDev</i>	1.50	0.07	0.03	0.01
DNA 1:2	581.30	16.99	9.07	1.87
<i>StdDev</i>	5.48	0.42	0.23	0.01
DNA 1:4	302.66	17.07	9.24	1.88
<i>StdDev</i>	5.06	0.31	0.66	0.01
DNA 1:8	151.00	16.87	9.12	1.85
<i>StdDev</i>	2.70	0.52	0.28	0.00
DNA 1:16	73.43	8.43	4.59	1.84
<i>StdDev</i>	4.15	0.44	0.23	0.00

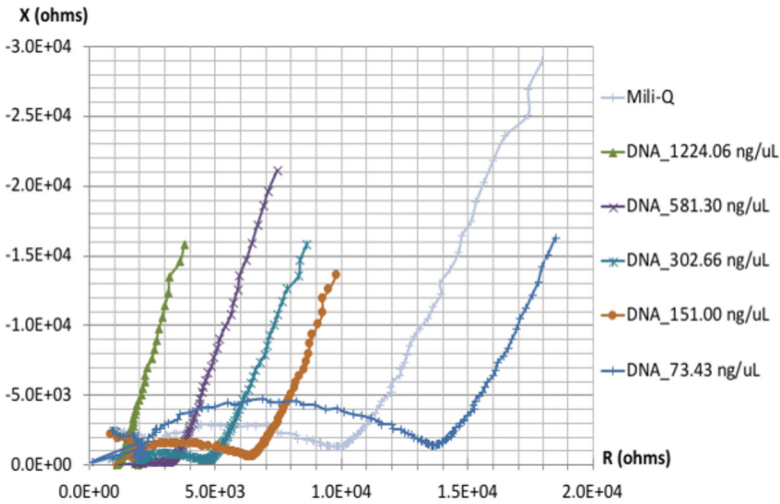


Fig. 4. Bioimpedance Nyquist plot for Milli-Q water and five DNA dilutions. Every point is the mean value of three R and X measurement at its corresponding frequency.

4 Discussion

The interaction of water molecules and DNA at frequencies below the gamma band could be correlated with changes in bioimpedance, based on this foundation, in the present study, DNA samples were analyzed and their impedance spectra in the complex plane were observed. Nyquist plots show similar curves with apparent different analytical parameters accordingly DNA concentration, it seems the fundamental frequency of the

semicircle observed is centered around 200 kHz, where the reactance values are inversely proportional to the DNA concentration. In addition, an interesting inflexion point is evident around 5 kHz where resistance values seems decrease as DNA concentration increase. Exploration of reactance and resistance parameters at such frequencies seems pertinent to design calibration curves to estimate DNA concentration. Further studies are warranted to confirm the observations. Milli-Q water curve has similar pattern for DNA samples, it seems the observed characteristic patterns are associated to the analyte vehicle and the quantity DNA change the curve parameters. Despite having obtained consistent results, new measurements are required to corroborate the information obtained.

5 Conclusions

Characteristical EBiS patterns are associated to Milli-Q water as analyte vehicle, and its parameters change as a function of DNA concentration. It seems possible to determine DNA concentration in the range of 73 to 1200 ng/ μ L from measuring reactance and resistance at frequencies centered around 200 and 5 kHz, respectively, sensitivity analysis for reactance and resistance at such frequencies seems pertinent to design calibration curves to estimate DNA concentration.



Acknowledgments. NNGG, CMS and LBM thank to “Programa Delfín” their support to developed this research during Summer Scientific 2023. The Instituto Politécnico Nacional-México supported this work under project SIP-20230467.

References

1. Gnaim, R., Sheviriyov, J., Golberg, A., Ames, G., Oziel, M., González, C.A.: Label-free cDNA detection based on radiofrequency scattering parameters: a new approach for an inexpensive gene sensor. *J. Med. Devices* **14**(1), 014502 (2020)
2. Ames-Lastra, G., et al.: Bioimpedance spectroscopy as a potential technique to detect label-free PCR products. *J. Phys. Conf. Ser.* **2008**(1), 012016 (2021)
3. Paleček, E., Bartošík, M.: Electrochemistry of nucleic acids. *Chem. Rev.* **112**(6), 3427–3481 (2012)
4. Ames-Lastra, G., et al.: Design proposal of a DNA biosensor by means of relative bioimpedance measurements and genetic amplification: preliminary results. *J. Phys. Conf. Ser.* **2008**(1), 012020 (2021)
5. Hernández-Salinas, C.A., Corzo-Cruz, A., Sánchez-Monroy, V., Munguía-Cervantes, J.E., Gon-zález-Díaz, C.A.: Correlation of the DNA concentration of human samples to electrical bioimpedance measurements: a pilot study. *J. Electr. Bioimpedance* **13**(1), 132–135 (2022)
6. Ames, G., et al.: Impedance measurements sensitive to complementary DNA concentrations. In: Bertemes-Filho, P. (ed.) 17th International Conference on Electrical Bioimpedance. ICEBI 2019. Springer, Singapore, pp. 154–157 (2020). https://doi.org/10.1007/978-981-13-3498-6_23
7. Shim, S.M., et al.: Multilaboratory assessment of variations in spectrophotometry-based DNA quantity and purity indexes. *Biopreservation Biobanking.* **8**(4), 187–192 (2010)



Design and Heat Transfer Simulation of a Thermal Block for the Integration of a DNA Biosensor Based on PCR

Gerardo Ames Lastra¹  and César A. González Díaz²  

¹ Universidad Autónoma Metropolitana, Unidad Iztapalapa, Av. San Rafael Atlixco 186, Leyes de Reforma Ira Secc, Iztapalapa, 09340 Ciudad de México, CDMX, Mexico

² Instituto Politécnico Nacional, Escuela Superior de Medicina, Salvador Díaz Mirón esq. Plan de San Luis S/N, Miguel Hidalgo, Casco de Santo Tomas, 11340 Ciudad de México, CDMX, Mexico

cgonzalezd@ipn.mx

Abstract. In this work, we present the design of a thermal block and the heat transfer simulation for the integration of a DNA biosensor based on PCR. The thermal block contains the PCR tubes which needs to be heated and cooled in order to catalyze the Polymerase Chain Reaction for genetic amplification. The simulation is compared with thermographic images from an experiment performed with the integrated system. Even though there are minimal error in the temperatures of the simulation and the thermographic images, the design resulted adequate for the application.

Keywords: Thermal Block · PCR · Heat Transfer Simulation · COMSOL · Genetic Amplification

1 Introduction

One of the most revolutionary advances in molecular biology, which has been introduced into a vast variety of areas of medicine, among other disciplines, is the Polymerase Chain Reaction discovered by Karl Mullis in 1983 and which earned him the Nobel Prize of Chemistry in 1993. PCR is an enzymatic reaction catalyzed by thermal cycling with specific temperature steps to amplify a specific DNA sequence; this technique has been used in a wide variety of genomic analysis applications. In sum, the reaction is carried out by repeating temperature cycles consisting of 3 stages: DNA denaturation, which is carried out at around 95°C; the alignment of the primers, which is carried out around 60°C; and elongation by DNA Polymerase, which occurs around 70°C [1]. Each PCR reaction requires containing the template DNA, the primers, nucleotides, the reaction buffer with the MgCl₂ enzyme cofactor and the DNA polymerase. The aforementioned reagents are mixed in a container, usually but not limited to PCR tubes, and processed in the thermal steps described. Initially, thermocycling was done manually by introducing the PCR samples into individual and separate thermal baths; the user introduced the PCR tubes to a first thermal bath previously set at the denaturation temperature and kept it

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024

J. de J. A. Flores Cuautle et al. (Eds.): CNIB 2023, IFMBE Proceedings 96, pp. 392–402, 2024.

https://doi.org/10.1007/978-3-031-46933-6_41

for the time defined for this stage. Subsequently, the PCR tube was removed and placed in another pre-set thermal bath at the alignment temperature, and so on, repeating the process as many times as defined [1]. Of course, this process was tedious and susceptible to human errors since users had to remain standing for hours and pending to complete the process, so technology had the challenge of solving these problems. Therefore, devices were introduced that automated the thermocycling process. On the one hand, there were instruments that did not change the heating and cooling scheme, but rather automated the change of the thermal bath by means of robotic arms that submerged and removed the PCR tubes in the predisposed baths.

On the other hand, the development of thermal cyclers was not limited to automating the manual tasks carried out by the user within the laboratory. Rather, scopes were made that changed the management of thermal energy introduced into the samples so that thermal baths were no longer used, but temperature control systems with different operating principles and actuators. The challenges to overcome were the high reaction volumes, high energy consumption and low reaction efficiency due to slow temperature transitions as well as poor thermal stability in the reaction stages. Next, some of the different scopes of thermal cyclers developed over time with the aim of overcoming these obstacles described in a bibliographic review article published by Sailaja V. et al. [2] will be mentioned.

The idea that was shared was that the miniaturization of thermocycling equipment could overcome the obstacles mentioned above because it would solve the reaction volume issue, that is, it would reduce the volume of reagents needed for a single assay. Similarly, reducing the size of the thermal mass allowed for faster transitions due to the reduction in thermal capacitance introduced by the thermal blocks used to transfer temperature from the actuators to the sample. The first miniaturized thermal cyclers appeared in 1993, designed with micro-engineered chemical reactors with integrated heaters [2]. D.S. Lee and colleagues used an infrared heater and water shock to heat and cool, respectively, the sample placed inside capillary tubes, reporting [3]. Other scopes used a tungsten lamp as the heat source and forced air convection as the cooling mechanism, reporting temperature ramps of up to 65 C/s and 80 C/s for heating and cooling, respectively [3]. They also reported the evaluation of the use of natural convection and forced convection for sample temperature management, reporting a performance of 2 C/s and 6 C/s, respectively. Lagally et al. They used an integrated microfluidic device consisting of sub-microliter volume PCR chambers; used nitrogen flow on top of the device for cooling and thin-film heaters below the PCR chambers for heating, reporting cooling and heating ramps around 10 C/s [4]. Niu et al. developed a PDMS (polydimethylsiloxane) glass device for PCR, using a zigzag-shaped Platinum heater and a forced convection cooling fan; reporting heating ramps of 10 C/s and cooling ramps of 4.6 C/s [5].

These achievements overcame some of the obstacles that arose for the development of PCR devices (thermal cyclers). However, they introduce other obstacles to the use and development of these technologies, such as sample handling, the complexity of device design due to temperature management systems that were based on the movement of fluids (flow) either gaseous or liquids, which also hinders their portability, and the energy consumption that the actuators represented. The development of thermal cyclers was

not limited to miniaturization, but efforts were focused on the use of different actuators that allowed thermal control in an efficient and simple manner. Thermoelectric Cooling (TEC) systems have become popular in thermocycler designs due to their convenient features of having no moving parts, compact size and light weight, their ability to cool below temperature environment, precise temperature control, high reliability, acoustic noise-free operation, operation in any orientation, and that are friendly to the environment [2]. Today, a large part of commercial thermal cyclers includes Peltier modules (TEC-based actuators) in their designs, although they report being energy-intensive designs (between 200 and 1000 Watts) intended primarily for high-resource laboratories. On instruments such as the Agilent SureCycler 8800, Bio-Rad C1000, Eppendorf Mastercycler Pro, Life Technologies GeneAmp 9700, Roche Lightcycler 96, and Thermo Scientific Arktik, manufacturer literature indicates maximum temperature ramps in the range of 1.5 to 5 C/s and precision of $\pm 0.25^\circ\text{C}$. Some companies have improved temperature ramps by optimizing the designs of the thermal blocks that contain the samples, such as the Bio-Rad S1000 with temperature ramps between 2.5–6 C/s, with power consumption of 700 W. Other companies have changed the thermal block material by seeking higher thermal conductivity and lower thermal mass using silver thermal blocks such as Eppendorf Mastercycler Pro S with temperature ramps of 4.5–6 C/s and power consumption of 950 W. Companies such as Thermo Scientific Piko have improved their ramps temperature ranges between 4.5–5 C/s by reducing system size and reaction volumes. Chai Technologies company optimized the temperature ramps of its OpenPCR instrument at 1 C/s with relatively low power consumption of 180 W by reducing the number of samples introduced for a single assay, which decreases block size and thermal mass. Companies such as Cepheid SmartCycler and Streck Philisa improved the performance of their thermal cyclers to ranges of 2.5–10 C/s with power consumption of 350 W and 12–15 C/s, respectively, by redesigning the PCR tubes to improve the ratio surface-volume of the samples, favoring the conduction of temperature in the samples. Others have improved their speeds at the expense of energy consumption and portability, such as the Idaho Technology company with a convective cycling design, now included in the Roche Lightcycler 2.0 system and the Rotor-GeneQ system, demonstrating temperature ramps between 15–20 C/s in standard sample volumes (between 0.1 and 0.5 mL) and 1.9–3.6 C/s in capillary volumes, with power consumption of 800 W [2]. So, designs with Peltier modules present different ways to optimize the characteristics of a thermal cycler.

In this work, the design of a thermal block is reported as a part of the thermocycler module integrated in the design of a DNA biosensor based on PCR and bioimpedance measurements, whose implementation is documented in [6]. This work aims to document the design process of the thermal block and the multiphysics simulation in which it was supported during the design process.

2 Methods

2.1 Thermal Block Design

As previously mentioned, in order to catalyze the polymerase chain reaction, it is necessary to subject the sample to thermocycling that involves heating and cooling the sample in a wide range of temperatures that can go from room temperature, or even lower than this, up to 100°C . To do this, a thermal block was designed with a geometry such that the PCR tubes, which contain the sample, could be placed in and that the external surface of the PCR tubes is in full contact with the thermal block to optimize heat transfer from the actuators towards the sample. Aluminum was selected as the block material due to its good thermal conductivity (around 209 W/mK [7]) and its cost. To define the dimensions of the block, the geometry of the PCR tubes was measured and based on these, the block was designed in a computer-aided drawing software, looking for the largest possible contact surface with the actuators without the dimensions were unnecessarily large, because we are looking for rapid heat transfer. A hole for the LM35 temperature sensor was contemplated at the bottom, in the center of the block with a given height so that the sensor would sit symmetrically in the center of the two samples. To sustain the pressure exerted from inside the PCR tubes due to the evaporation of the sample, a lid made of the same material as the block was designed, held by a screw to the bottom part of the thermal block, with holes designed to fit with the top of the PCR tubes. In addition, one of the holes was drilled to allow the connection with the electrodes and perform the bioimpedance measurements, since it is a thermal block integrated into a DNA biosensor, as previously mentioned. The final external dimensions of the thermal block are depicted in millimeters in Fig. 1.

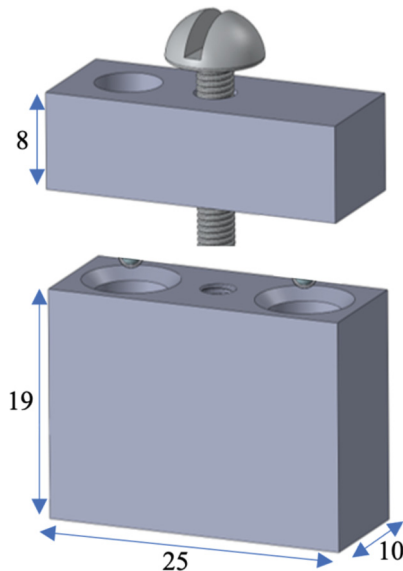


Fig. 1. Thermal block design with its external dimensions in millimeters.

The thermocycling system was designed using thermoelectric cooling technology (TEC) modules as actuators because they are systems that do not contain moving parts, do not emit acoustic noise, are friendly to the environment, have a compact structure, are small in size and allow precise temperature control, according to the data sheet of the TEC modules used [8]. TEC1–4905 Peltier modules were selected, with dimensions of 25 mm × 25 mm × 4 mm, maximum voltage of 6.8 V, maximum current of 5.3 A. One Peltier was placed on each side of the thermal block with the side to be heated in contact with the thermal block, according to the direction of the electric current. In addition, a heatsink and a fan whose speed is controlled by a PWM signal were placed on each external side of the Peltiers, to dissipate the heat accumulated on the external side of the Peltiers to the environment whenever the sample cools; since the side in contact with the thermal block will cool down and the external side will heat up. It is important that the heat dissipation is efficient, otherwise the heat accumulated on the hot side will start to heat the cold side and the temperature control will not work properly.

2.2 Heat Transfer Simulation

In order to assess whether the proposed heat block design was suitable for the intended application, a heat transfer versus time simulation was performed to see if the heat block design would allow heat transfer from the peltiers to the samples in a uniform and efficient way. For this, the multiphysics simulation software COMSOL Multiphysics was used.

To define the geometry of the simulated system, the 3D models of the different components that make up the physical system were imported, these being the Peltiers, the thermoblock, the sample containers and the sample itself. Figure 2 shows the assembled thermoblock system with PCR tubes containing 25 μ l of water simulating the PCR mix. The materials of each one of the components were assigned from the COMSOL library, being alumina for the Peltiers, aluminum for the thermoblock, polypropylene for the sample containers and pure water for the samples; In addition, an air bubble surrounding the entire system was considered to simulate more realistic conditions and not neglect the energy losses to the environment that would occur in a real scenario. Because the materials were selected from the COMSOL library, physical properties such as thermal conductivity, heat capacity, thermal diffusivity, and density of each component material are already taken into account. The finite element mesh was automatically performed from the COMSOL software selecting the element size as extra fine; The element size parameters were 0.007 m as maximum element size, 0.0003 m as minimum element size, 1.35 for the maximum element growth factor, curvature factor of 0.3, and resolution for narrow regions of 0.85. As initial condition of the system, the ambient temperature was considered, at 25°C. The first step of the simulation consisted of defining the temperature of the Peltiers at 95°C, corresponding to the denaturation temperature in a PCR assay, thus evaluating the heat transfer from an initial state to the temperature of the first stage of the assay. by PCR; For the second simulation step, the Peltier temperature was defined at 60°C, corresponding to the alignment temperature in a PCR assay and the last result of the first simulation step was used as the initial condition, and, for the third step of the simulation, the temperature of the Peltiers was defined at 72°C, corresponding to

the elongation temperature in a PCR assay, and the last result of the second step of the simulation was considered as the initial condition.

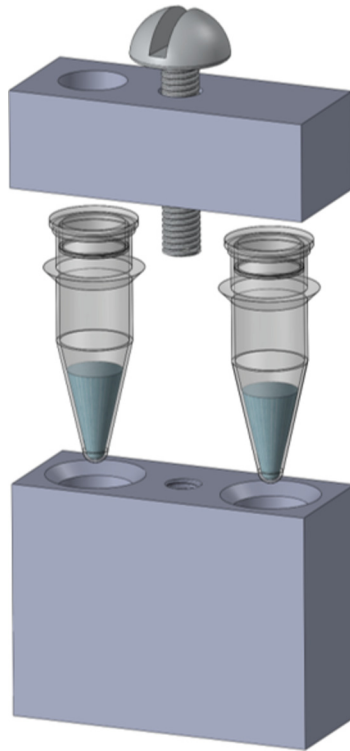


Fig. 2. 3D model of the block assembled with the PCR tubes containing 25 μl of pure water simulating the sample.

3 Results

Figures 3, 4 and 5 show the temperature of the thermal block after 40 s elapsed in each step of the simulation. For the figures, 3 types of COMSOL results were integrated in one single image, the first one shows the surface temperature of each domain, a second one that shows the temperature of a cut or defined plane to be able to observe the loss of energy to the environment and a third result that places the annotation of a specific data. For the latter, it was defined that the annotation of the temperature was made at a given point, that point was located in the center of the water domain (simulating the PCR sample) to evaluate the temperature of the sample in the elapsed time. All three types of results were integrated in one image, each figure was obtained in such way (Figs. 3, 4 and 5).

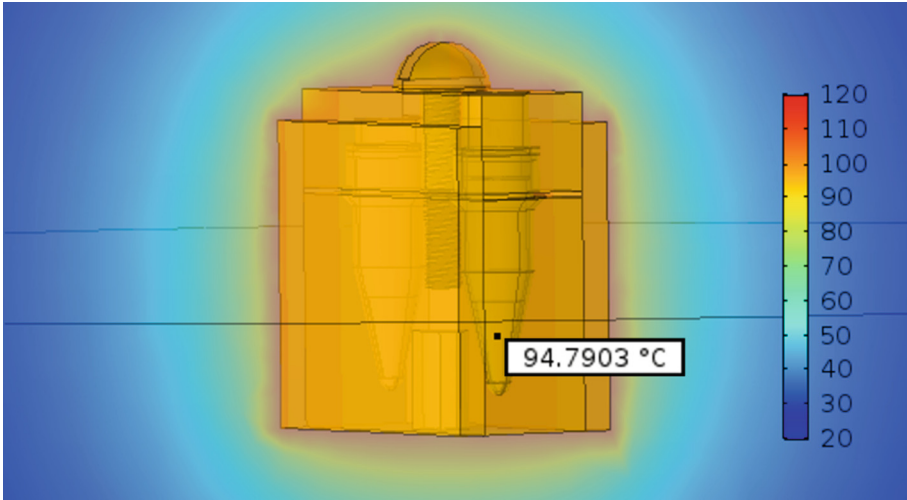


Fig. 3. Step 1 of the simulation with 95°C defined as the peltier’s temperature. Surface temperature, temperature of a defined plane and an annotation for the sample temperature re integrated in the image. The annotation corresponds to the average of the temperature of the water domain (sample).

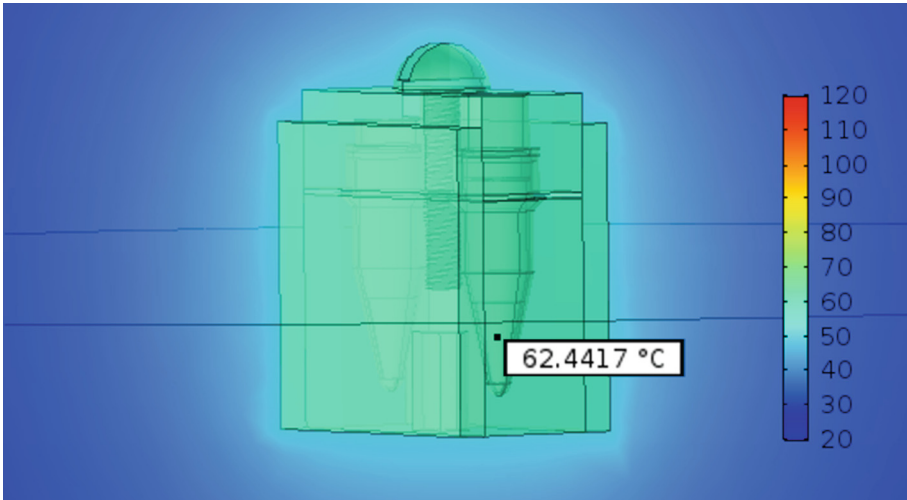


Fig. 4. Step 2 of the simulation with 62°C defined as the peltier’s temperature. Surface temperature, temperature of a defined plane and an annotation for the sample temperature re integrated in the image. The annotation corresponds to the average of the temperature of the water domain (sample).

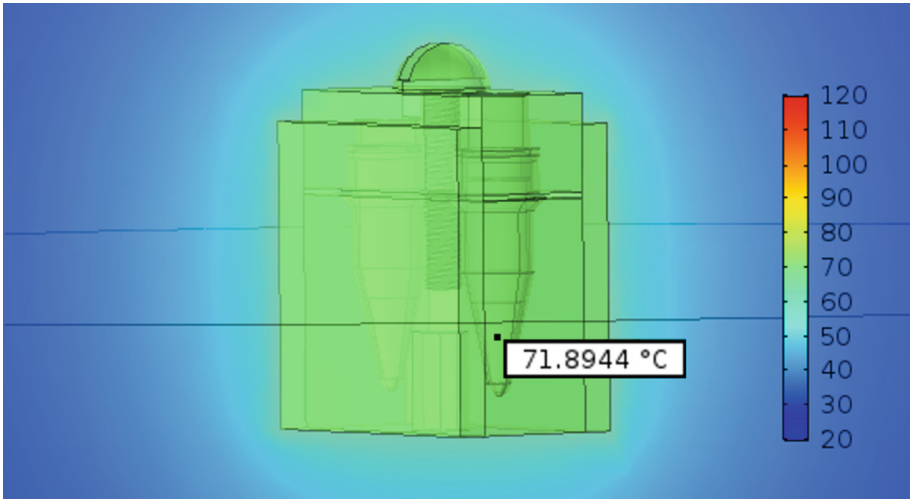


Fig. 5. Step 3 of the simulation with 72°C as the peltier's defined temperature. The annotation corresponds to the average temperature of the water domain (sample).

In addition, a probe was defined that averages the temperature of the water domain (PCR sample) to plot it against time and to be able to evaluate the heat transfer performance from the peltiers to the sample. Figure 6 shows the average temperature of the water sample over the whole simulation.

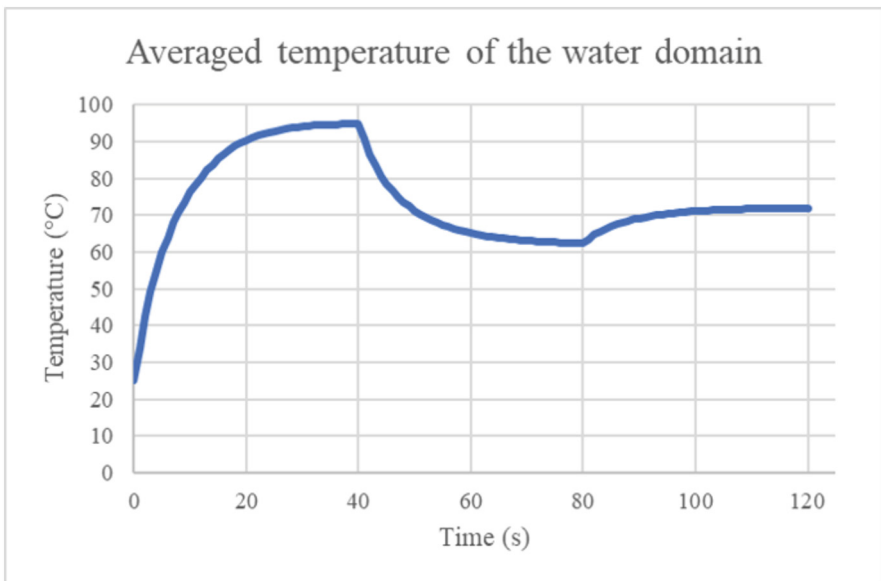


Fig. 6. Plot of the average temperature of the water domain (sample) against time.

In order to compare the simulation with the real thermal block (fabricated around from the computer-aided design files), thermographic images of the thermal block were taken in operation with the thermal cycler, but without lid and PCR tubes, with the support of a camera. Fluke TI25 Infrared. In addition, the same simulation described above was carried out with a 3D model suitable for the conditions of the thermographic images (without cap and PCR tubes). Figure 7 shows the images obtained from the simulation (left side) and the thermographic images (right side).

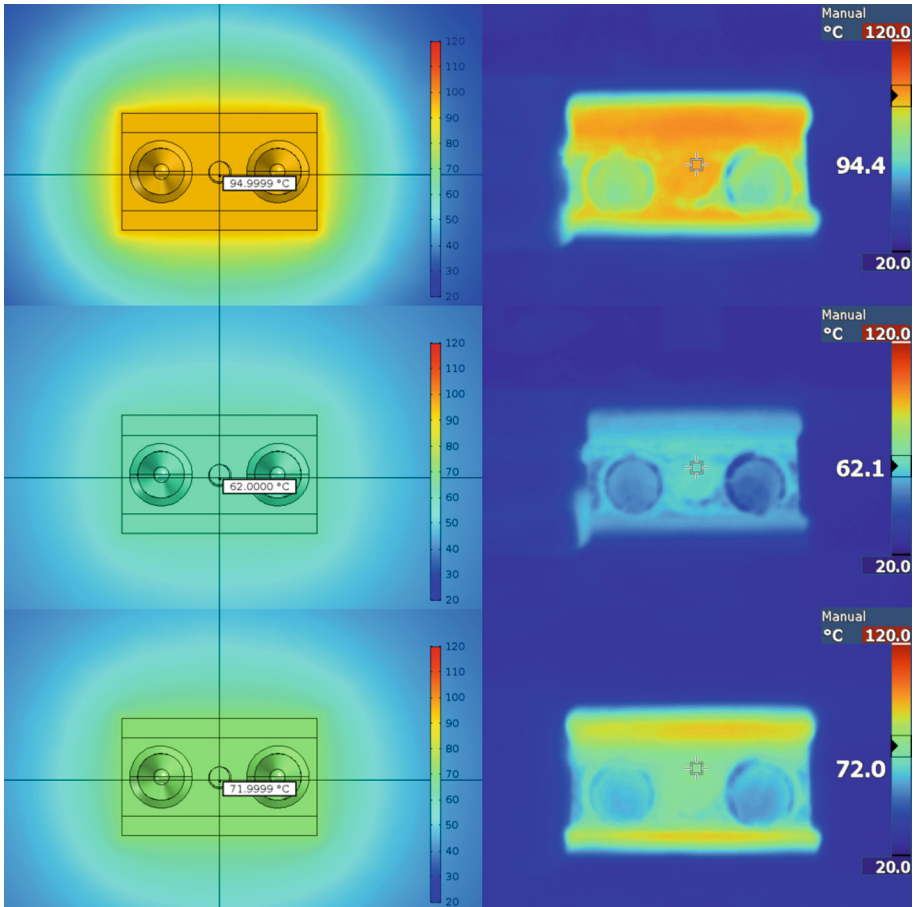


Fig. 7. Comparison of the images obtained from the simulation (left) with those obtained with the thermographic camera.

4 Discussion

From Figs. 3, 4 and 5, we can see that the temperature is distributed symmetrically from the peltiers, located on the sides of the thermal block, towards the sample. The heating of the thermal block is uniform and conducted efficiently. From the annotation, we can see that the sample is heating up to the defined temperatures with a slight error in the order of decimal places (around 0.2°C for the initial step at 95°C, 0.4°C for the second step at 62°C and 0.1°C for the third step at 72°C). This error is also observed in the comparison figures of the images obtained from the simulation and the images obtained from the thermographic camera (see Fig. 7), which resulted in 0.5°C for the first step at 95°C and 0.1°C for the second and third step at 62°C and 72°C, respectively. This error, which we consider to be insignificant, may be due to the assumptions that the simulated model contemplates, such as zero energy losses by radiation (reasonably neglected), the ambient temperature of the simulation was different from the ambient temperature of the real conditions of the experiment. In the simulation, the air is stagnant (without flowing) and the experiments were carried out with the laboratory windows open, so air flow could have occurred. This could explain that in the thermographic images, the environment of the thermal block is observed in a strong blue color, indicating a temperature close to 25°C (room temperature). Also, the infrared temperature measurements from metallic objects can be difficult because of the emissivity of aluminum, or any metallic shiny object, is dependent on the surface finish of the object.

From the graph of the average temperature of the water domain, we can observe a heat transfer from an overdamped system, this is expected because the temperatures of the peltiers in each step of the simulation were defined in a fixed way, that is, it does not change with respect to time or the temperature of the system in contact. This is a big difference with the conditions of the real experiment, because the temperature control, reported by Ames et al. in [6], the temperature control system has as feedback the temperature read by the LM35 sensor, located symmetrically in middle and at the height of the samples. This means that the temperature of the peltiers could be reaching temperatures above the defined temperature in each step of the simulation. Data suggested by the temperature values also reported in Ames et al. [6]. This could also explain the temperature error observed in the results of this work.

However, we consider that the differences in temperature and heat transfer time are negligible, since the errors are minimal and the samples reach the desired and required temperatures to catalyze the polymerase chain reaction and cause the desired genetic amplification, as reported in [9]. Confirming that the thermal block design is suitable for the application for which it was designed.

5 Conclusion

In this work, the design of a thermal block is reported. Such thermoblock is integrated in a DNA biosensor based on the Polymerase Chain Reaction and relative bioimpedance measurements documented in a previous work. The computer aided design of the block and a heat transfer simulation carried out in order to evaluate the design of the block and whether the heat transfer would be uniform and efficient, in the sense of bringing the sample to the desired temperatures, were reported.



Discrepancies were found in the values obtained in the simulation and in the real experiments, in terms of temperature and heating time, which may be due to the differences in environmental conditions between the simulation and the experiments, and to the difference in the temperature control system of the real system with the defined temperatures of the simulations. However, these differences are not critical and can be neglected, since the temperatures reached by the sample are very close to the desired and necessary temperatures to catalyze the polymerase chain reaction and cause the desired genetic amplification, confirming that the design is suitable for the intended application.

References

1. Bronzino, J.D.: *The Biomedical Engineering Handbook* (2000)
2. Sailaja, V., Raju, K.N., Sailaja, C.A.V.: A review on heating and cooling system using thermo electric modules. *J. Electr. Electron. Eng.* **14**(2), 49–57 (2019)
3. Lee, D.S., Tsai, C.Y., Yuan, W.H., Chen, P.J., Chen, P.H.: A new thermal cycling mechanism for effective polymerase chain reaction in microliter volumes. *Microsyst. Technol.* **10**(8–9), 579–584 (2004)
4. Lagally, E.T., Medintz, I., Mathies, R.A.: Single-Molecule DNA amplification and analysis in an integrated microfluidic device molecules followed by capillary electrophoretic (CE). *Anal. Chem.* **73**(3), 565–570 (2001)
5. Niu, Z.Q., Chen, W.Y., Shao, S.Y., Jia, X.Y., Zhang, W.P.: DNA amplification on a PDMS-glass hybrid microchip. *J. Micromech. Microeng.* **16**(2), 425–433 (2006)
6. Ames-Lastra, G., et al.: Design proposal of a DNA biosensor by means of relative bioimpedance measurements and genetic amplification: preliminary results. *J. Phys. Conf. Ser.* **2008**(1), 012020 (2021)
7. Cortés, L.L., Duarte, S.G., Lángo, E.M., González, D.E.: Conductividad Térmica De Metales. *Simp. Meteorol.* **1**(1), 1–4 (2010)
8. Thermonamic. TEC1–4905 Datasheet
9. Ames-Lastra, G., et al.: Bioimpedance spectroscopy as a potential technique to detect label-free PCR products. *J. Phys. Conf. Ser.* **2008**(1), 012016 (2021)



Magnetic Induction Spectroscopy for DNA Quantification: Preliminary Results

Jaime Eduardo López-Madrugal¹ , Liliana Becerra-Martínez²,
Leticia Arias-González¹, Nancy Nallely Gómez-González³,
Catalina Montoya-Solórzano⁴, Virginia Sánchez-Monroy¹,
Nadia Mabel Pérez-Vielma⁵, Modesto Gómez-López¹,
and César Antonio González-Díaz¹ 

¹ Escuela Superior de Medicina-Instituto Politécnico Nacional, 11320 Ciudad de México, Mexico

cgonzalezd@ipn.mx

² Universidad Veracruzana, 91090 Xalapa, Veracruz, Mexico

³ Universidad de Guadalajara, 44100 Guadalajara, Jalisco, Mexico

⁴ Universidad Autónoma de Manizales, 170001 Manizales, Caldas, Colombia

⁵ Centro Interdisciplinario de Ciencias de la Salud Unidad Santo Tomas, 11320 Ciudad de México, Mexico

Abstract. Biosensors on the basis of label-free and non-functionalization principles have been studied in the last decade. In particular, sensors for total Deoxyribonucleic Acid (DNA) and Genes, a necessary technology within molecular biology assays, have so far represented a major technological challenge. The objective of this study was to explore Magnetic Induction Spectroscopy (MIS) for DNA detection. The inductive phase measurements were obtained for five different DNA concentrations (1:1, 1:2, 1:4, 1:8 and 1:16), and analyzed to identify its potential value as a non-invasive tool for DNA quantification. According to the results, integration of the inductive phase curve in the frequency range of 1 to 10 MHz correlates with statistical significance to DNA concentration ($R = 0.9723$, $P < 0.05$). Sensitivity analysis for such integrated parameter seems pertinent for calibration of Magnetic Induction Spectroscopy measurements to estimate DNA concentrations. Further studies are warranted to confirm the observations.

Keywords: Magnetic · Induction · DNA · Biosensor

1 Introduction

Deoxyribonucleic Acid (DNA) quantification is a necessary technique within molecular biology assays, and emerging easy-access technologies for such purpose have been studied [1]. Electrical Bioimpedance (EBI) measurements is a technique that could utilizes electrical properties of DNA and has shown correlate highly with DNA concentration [2, 3]. Bioelectrical measurements by electromagnetic induction with non-contact electrical coils are considered as a valuable alternative to contact electrode measurement

[4] and [5]. Magnetic Induction Spectroscopy (MIS) is a new technique to determine the electrical properties of biological materials without contact with the sample. Several studies have been developed to show the feasibility of MIS as non-invasive monitoring technique [6] and [7].

Our study relates to the use of MIS for measuring different DNA concentrations. We are particularly interested in using the inductive phase occurring through magnetic induction as a means to detect the presence and quantity of total DNA. To the best of our knowledge multifrequency inductive phase as a measure of DNA concentration was not studied before. In this study we explore the use of MIS for DNA detection. Multifrequency inductive phase measurements were obtained for different DNA concentrations, and analyzed to identify its potential value as a non-invasive tool for DNA quantification.

2 Methodology

2.1 Experimental Design

MIS measurements were developed in five total DNA samples at different concentrations, as well as in Milli-Q water as vehicle reference. Inductive phase was obtained at every frequency and Bode plots were drawn. An electrophoresis assay was developed for qualitative characterization of DNA concentrations, in addition, DNA samples concentrations were quantified through UV-Vis Spectrophotometry. Data were analyzed through a non-parametric correlation statistical test using the Excel 2021 (18.0) with a statistical significance $p < 0.05$. Figure 1 shows a block diagram of the experimental design.

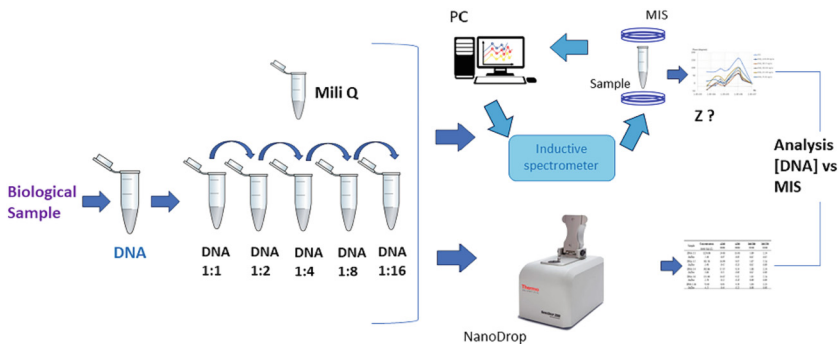


Fig. 1. Block diagram of the experimental design.

2.2 DNA Procedure

Samples were prepared from commercially available Deoxyribonucleic acid sodium salt, from salmon sperm, Sigma-Aldrich, reference: D1626, No. Cat: 438545-06-3 (Sigma-Aldrich Quimica S de RL de CV, Toluca, México). A master solution was prepared by dissolving the solute in Milli-Q water (Milli-Q® IQ 7003/7005 Water Purification Systems) at initial concentration adjusted to 1.2 mg/ml, then; lineal dilutions were obtained

in a final volume of 50 μL (1:1, 1:2, 1:4, 1:8 and 1:16). An electrophoresis was carried out on agarose gel to evaluate qualitatively the presence and integrity of DNA at every concentration, then, final DNA concentrations were measured by the use of the NanoDrop Spectrophotometer 2000 (Thermo Fisher Scientific - Waltham, MA USA) in triplicate assays. DNA concentration and purity was determined by examining the optical density 260/280 ratio, which was within acceptable values (1.8 to 2.0) [8]. Table 1 shows the initial and diluted DNA concentration in $\text{ng}/\mu\text{L}$ as well as purity ratios 260/280.

2.3 MIS Measurements

An inductive spectrometer was designed to measure MIS in DNA samples. The system uses a direct digital synthesizer Keysight, model 2567 B (Keysight Inc. Santa Rosa, CA, USA) coupled to an arrangement of inductor-sensor cooper coils coaxially centered and geometrically disposed around a polypropylene PCR tube as container, and to induce magnetic fields in a total volume of 50 μL at eight magnetic field frequencies (0.005, 0.02, 0.05, 0.10, 0.20, 1, 2 and 10 MHz). An electronic circuit was designed on the basis of the integrated chip AD8302 (Ana-log Devices Inc. Wilmington, MA USA) to detect the inductive phase in the sensor-coil with respect to the inductor-coil. Signals were digitalized through an USB data acquisition (DAQ) devices with 8 analog inputs, 16-bit resolution, and 100 ks/s sample rate (Measurement computing Inc. Norton, MA, USA). A personal PC was used to program the system and store the data.

3 Results

3.1 DNA Dilutions

Figure 2 shows the electrophoresis assay for the five DNA concentrations. Qualitative decrease in fluorescence as a function of DNA dilution is evident. Table 1 shows the DNA concentration values in $\text{ng}/\mu\text{L}$ and the absorbance rates found for each dilution. The purity of the DNA was confirmed by the optical density ratio 260/280 in the range greater than 1.8 nm.

3.2 MIS Spectra

Figure 3 shows the MIS spectra for inductive phase in final DNA dilutions. Spectrum for Milli-Q water sample is included. Inductive phase patterns show an inflexion point at 1 MHz, and a relation MIS spectra slopes at subsequent frequency regarding DNA concentration is evident. Figure 4 shows a zoom of the MIS spectra in the range 1 to 10 MHz to highlight the correlation of the spectra's slopes with respect DNA concentration. As a first approach to a sensitivity analysis, integration of the inductive phase curve from 1 to 10 MHz was estimated (see Fig. 5). A Spearman correlation test of the inductive phase-integration values in the range 1 to 10 MHz regarding DNA concentration showed a coefficient correlation of $R = 0.9723$ with statistical significance ($p < 0.05$).

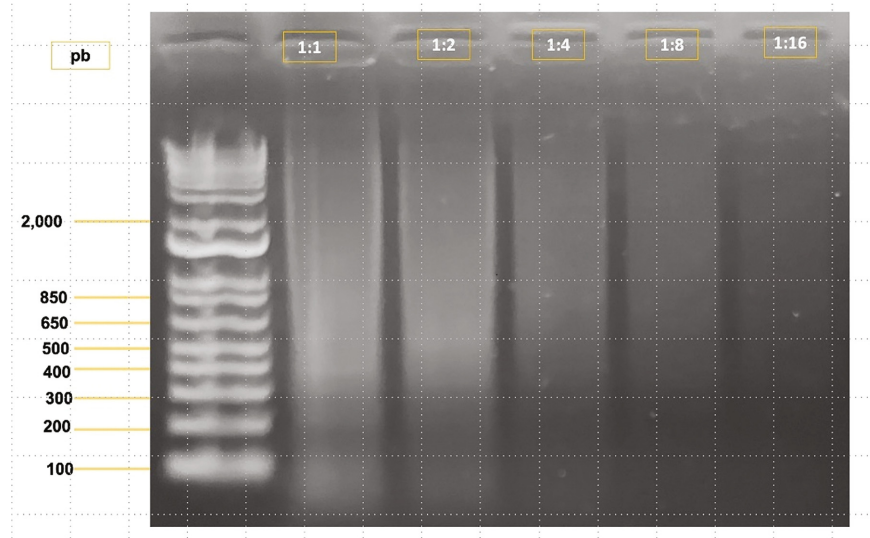


Fig. 2. Electrophoresis assay for five DNA concentrations. Diminished fluorescence as a function of DNA dilution is evident. Molecular ladder in the witness lane (left side) are in pb.

Table 1. Final concentrations in ng/uL and absorbance ratio for five DNA dilutions samples. Data are shown in mean values and standard deviation.

Sample	Concentration	A260 Mean	A280 mean	260/280 mean
	mean (ng/ μ L)			
DNA 1:1	1224.06	24.60	13.00	1.89
<i>StdDev</i>	1.50	0.07	0.03	0.01
DNA 1:2	581.30	16.99	9.07	1.87
<i>StdDev</i>	5.48	0.42	0.23	0.01
DNA 1:4	302.66	17.07	9.24	1.88
<i>StdDev</i>	5.06	0.31	0.66	0.01
DNA 1:8	151.00	16.87	9.12	1.85
<i>StdDev</i>	2.70	0.52	0.28	0.00
DNA 1:16	73.43	8.43	4.59	1.84
<i>StdDev</i>	4.15	0.44	0.23	0.00

4 Discussion

Milli-Q water curve has similar pattern than DNA samples, it seems the observed characteristic patterns are associated to the analyte vehicle and the quantity DNA change the curve parameters, particularly in the inflexion point observed at 1 MHz. MIS sensitivity seems evident in the range 1 to 2 MHz, at such frequencies the interaction of the water

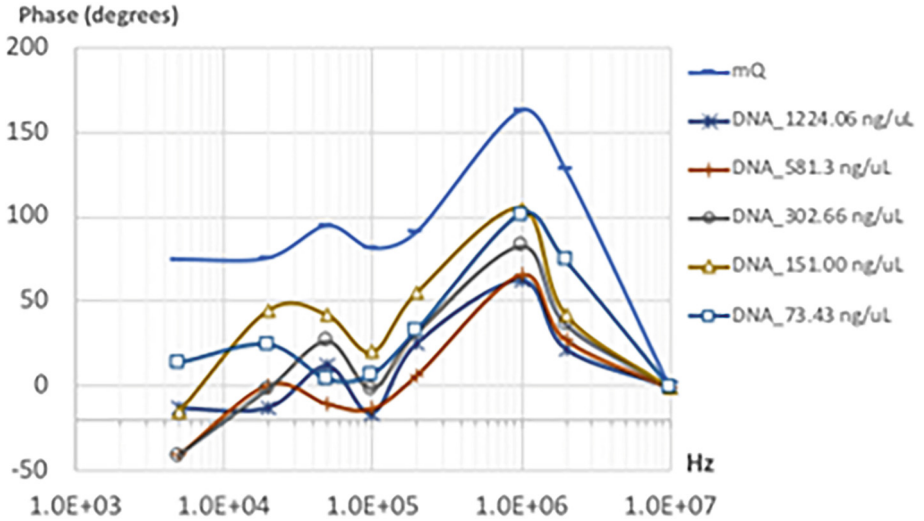


Fig. 3. MIS (inductive phase) spectra in five DNA concentrations. Spectrum for mQ (Milli-Q water) sample is included. Data are shown in mean value and standard error.

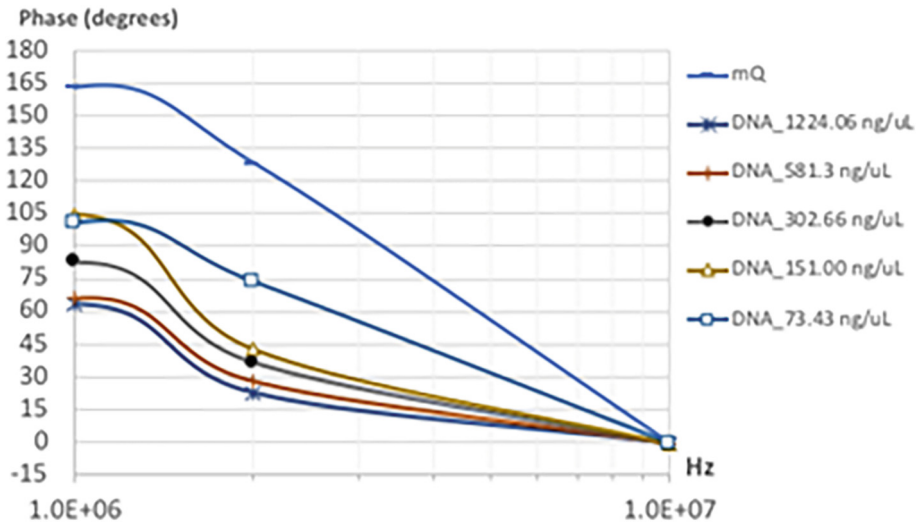


Fig. 4. Zoom of MIS (inductive phase) spectra in five DNA concentrations in the range 1 to 10 MHz. Spectrum for mQ (Milli-Q) sample is included. Data are shown in mean value and standard error.

dipole with electronegative charge of DNA molecule could be relevant, and volumetric conductivity reflects inductive phase changes, indeed inductive phase has the potential for being a non-invasive method for detecting DNA concentrations in the range of 73

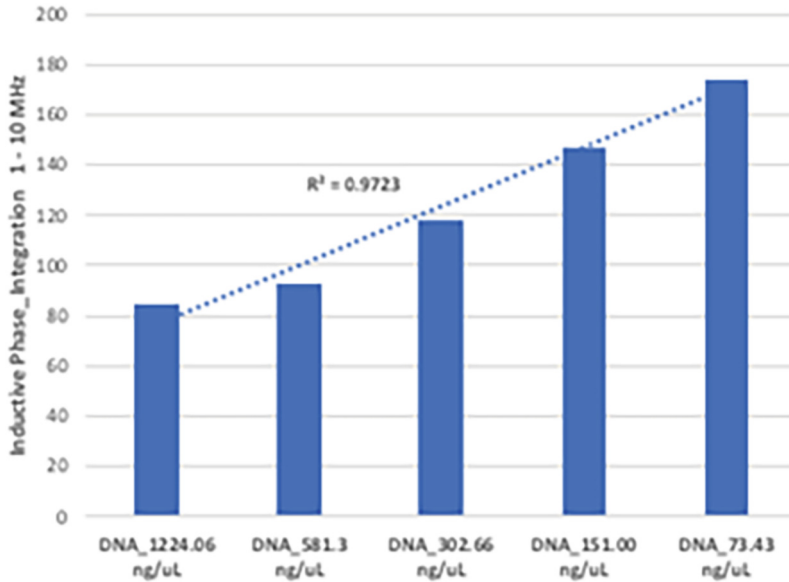


Fig. 5. Integration of the inductive phase curve in the range 1 to 10 MHz for five DNA concentrations. Data are shown in mean value and standard error. Significant statistically correlation inductive phase-integration vs DNA concentration was found ($R = 0.9723$, $P < 0.05$).

to 1200 ng/ μ L. Obviously, this is a preliminary study and much research remains to demonstrate the observations.

5 Conclusions

MIS measurements show technical feasibility for non-invasive detection of total DNA. Integration of the inductive phase curve in the frequency range of 1 to 10 MHz correlates to DNA concentration with statistical significance ($R = 0.9723$, $P \leq 0.05$), thus sensitivity analysis for such parameter seems pertinent for calibration of MIS measurements to estimate DNA concentration.

Acknowledgments. LBM, NNGG and CMS thank to “Programa Delfín” their support to developed this research during Summer Scientific 2023. This work was supported by the “Instituto Politécnico Nacional-México” under project SIP-20230467.

References

1. Hernández-Salinas, C.A., Corzo-Cruz, A., Sánchez-Monroy, V., Munguía-Cervantes, J.E., González-Díaz, C.A.: Correlation of the DNA concentration of human samples to electrical bioimpedance measurements: a pilot study. *J. Electr. Bioimpedance* **13**(1), 132–135 (2022). <https://doi.org/10.2478/joeb-2022-0018>

2. Ames-Lastra, G., et al.: Design proposal of a DNA biosensor by means of relative bioimpedance measurements and genetic amplification: preliminary results. *J. Phys. Conf. Ser.* **2008**(1), 012020 (2021). <https://doi.org/10.1088/1742-6596/2008/1/012020>
3. Hernández, A., et al.: Correlation between the concentration of DNA and electrical bioimpedance. *J. Phys. Conf. Ser.* **2008**(1), 012021 (2021). <https://doi.org/10.1088/1742-6596/2008/1/012021>
4. Netz, J., Forner, E., et al.: Contactless impedance measurement by magnetic induction- a possible method for investigation of brain impedance. *Physiol. Meas.* **14**, 463–471 (1993)
5. Griffiths, H., Steward, W.R., et al.: Magnetic induction tomography – a measuring system for biological materials. *Ann. NY Acad. Sci.* **873**, 335–345 (1999)
6. Al-Zeibak, S., Saunders, N.H.: A feasibility study of in vivo electromagnetic imaging. *Phys. Med. Biol.* **38**(1), 151–160 (1993). <https://doi.org/10.1088/0031-9155/38/1/011>
7. González, C.A., Rubinsky, B.: Frequency dependence of phase shift in Edema: a theoretical study with magnetic induction. In: *Proceedings of the 2005 27th Annual International Conference of the Engineering in Medicine and Biology Society*, vol. 1617238, pp. 3518–3521. IEEE-EMBS 2005. Institute of Electrical and Electronics Engineers Inc. (2005). <https://doi.org/10.1109/iembs.2005.1617238>
8. Shim, S.M., et al.: Multilaboratory assessment of variations in spectrophotometry based DNA quantity and purity indexes. *Biopreserv. Biobank.* **8**, 187–192 (2010)

Author Index

A

- Abarca-Castro, Eric Alonso 275
Aguilar-Gameros, Samantha 14
Aguirre-Ortega, Monica 14
Aizpuru, Ana P. Leyva 3
Alavez, Nohemi Barranco 258
Aljama-Corrales, Tomas 283
Ames Lastra, Gerardo 392
Anaya Campos, Leonardo Eliú 128
Arámburo-Castell, María Fernanda 201
Arellanes, Ileri Gisel Rosas 303
Arellano-Arteaga, Kevin Javier 67
Arenas-Ballesteros, Gabriel 137
Arias-González, Leticia 386, 403
Arquer Ruiz, María del Carmen 128

B

- Barreto, Alexica Celine Marquez 145
Bayareh-Mancilla, Rafael 211
Becerra-Martínez, Liliana 386, 403
Bojorquez Muñoz, Zaira L. 343
Buendia, Jesus Moises Martinez 84

C

- Calvo, Norma Celeste Estrada 365
Camacho, José 321
Camarillo-Cisneros, Javier 137
Campos-Delgado, Daniel Ulises 100
Campos-Lopez, Zarah 32
Cañedo-Figueroa, Carlos Eduardo 50
Cantillo-Negrete, Jessica 241
Cera-Domínguez, Edgar 137
Charleston-Villalobos, Sonia 283
Chávez, Sergio Israel Velázquez 41
Chavez-Ramirez, Adrian 14
Chávez-Sánchez, Manuel 174
Cigarroa-Mayorga, Oscar Eduardo 211
Cisneros, Javier Camarillo 145
Contreras-Rodríguez, Alison 283
Corkidi, Gabriel 59
Corona Sánchez, Sergio A. 343

- Corral-Barrón, Rolando 50
Cortés Sáenz, David 222
Cota-Ruiz, Juan 32
Cruz-Guerrero, Inés Alejandro 100
Cruz-López, Pilar 275
Cruz-Ramos, José Alfonso 67

D

- Darszon, Alberto 59
de la Cruz Gallardo, Paola Aidee 24
de la Peña Martínez, Rodrigo Ruiz 24
De La Rosa Hernández, Gerardo Adrián 183
Débora-Ramirez, Jorge Alonso 50
Del Val Saucedo, Eutzhel André 24
Delgado Rivera, Héctor G. 343
Díaz-Guerrero, Dan Sidney 59
Diaz-Hernandez, Octavio 303
Diaz-Roman, Jose 32
Dorantes-Méndez, Guadalupe 292
Dueñas, Astrid J. Ríos 3

E

- Equihua-Guillen, Fabian 93
Espinosa, Anali Elizabeth González 365
Estévez-Pedraza, Ángel Gabriel 374
Eusebio-Martín, Esteban 164

F

- Felix, Kristian Segura 41
Felix-Beltrán, Olga Guadalupe 201
Figueroa, Carlos E. Cañedo 3
Figueroa, Carlos Eduardo Cañedo 24
Flores, Dora-Luz 357

G

- García, Eduardo Pinto 41
García, Francisco Flores 41
García, Rodrigo Domínguez 145
García-Gil, Gerardo 67
Garcia-Lara, Adrian 93

García-Valdez, Andrea Alejandra 192
 Gómez, Cynthia Fernanda Carrillo 303
 Gómez-Celaya, Luis Alberto 50
 Gómez-González, Nancy Nallely 386, 403
 Gómez-López, Modesto 386, 403
 González-Díaz, César Antonio 386, 392, 403
 Gonzalez-Landaeta, Rafael 313, 357
 Gordillo-Castillo, Nelly 32, 222
 Guzman-Pando, Abimael 137

H

Hernández Ortiz, César F. 343
 Hernández, Carlos Alberto Gallegos 84
 Hernández, Haydee Olínca 59
 Hernández-Herrera, Paul 59
 Hernandez-Laredo, Enrique 374
 Hernández-Nava, Gerardo 112
 Hernández-Rodríguez, Yazmín Mariela 211
 Hernández-Vázquez, Natanael 155
 Huizar, Alina Andrea García 120
 Hurtado-Esquivel, Christian Isaac 258

J

Jaquez, José Irving Hernández 41
 Jimenez-Angeles, Luis 232
 Jurado, Roxana Quintana 145

L

Lázaro-Picazo, Marisela 275
 Leyva-López, Sheila 112
 Lira, Natalia Gabriela Sámano 3, 24
 López, Geshel David Guerrero 41
 López, Néstor 321
 López-Armas, Gabriela del Carmen 67
 López-Madrigal, Jaime Eduardo 386, 403
 López-Mata, José Carlos 232
 Luna, Ivan 321

M

Macías Macías, José M. 343
 Macías-Castro, Diana Karen 249
 Mariaca-Gaspar, Carlos 283
 Martínez, Paulina 222
 Martínez-Arboleyda, Daia Natalia 100
 Martínez-Hernández, Verónica Janeth 292
 Martínez-Méndez, Rigoberto 374
 Maya-Venegas, Victor 283
 Mederos-Madrado, Boris 32

Medina-Pérez, Víctor Manuel 67
 Medina-Ramos, Luis Antonio 211
 Mejía Muñoz, José 313
 Mejia-Muñoz, Jose 32
 Mena-Camilo, Enrique 112
 Mendez, Guadalupe Dorantes 357
 Meneses Peñaloza, Alicia 128
 Miramontes, Andrés Kaleb Aguiñaga 145
 Miranda-Hernández, Abigail 232
 Montañez Prieto, Natalia 343
 Montero, Efraín Muñoz 84
 Montoya, Fernando 59
 Montoya-Solórzano, Catalina 386, 403
 Morales Gómez, Jesús Alberto 183
 Morín-Castillo, María Monserrat 201
 Mosco-Vargas, Abel 283
 Munguía-Cervantes, Jacobo Esteban 386
 Muñoz, José Manuel Mejía 120

N

Nangüelú, Christian Roberto Ibáñez 365
 Núñez, Oscar 321

O

Oliveros-Oliveros, José Jacobo 201
 Olvera-Montes, Nemecio 283

P

Padilla-Becerra, María Cristina 249
 Pando, Abimael Guzmán 3, 24, 50
 Parra-Rodríguez, Lorena 374
 Perez Sanpablo, Alberto Isaac 128
 Pérez, Iliana González 84
 Pérez, Tomás Efraín Sánchez 84
 Pérez-Vielma, Nadia Mabel 386, 403
 Pizarro, Cynthia D. Márquez 3

Q

Quiñones Uriostegui, Ivett 128
 Quiroz Compean, Griselda 183

R

Ramirez-Alonso, Graciela 14
 Ramirez-Fernandez, Odin 93
 Reyes, Bersain 283, 321
 Reyes-Lagos, José Javier 275
 Ríos, Álvaro Anzueto 258
 Rivera, Griselda 331
 Rivera, Karen P. Rodríguez 3

Rodríguez Reyes, Gerardo 128
Rodríguez, Aldo Rodrigo Mejía 357
Rodríguez, Jesús J. Martínez 3
Rodríguez, Rogelio 331
Rodríguez-Domínguez, Poleth 50
Rodríguez-García, Martín Emiliano 241
Román-Godínez, Israel 155, 164, 174, 192,
249
Romero-Ramos, Andrea 14
Rubio, Mario Francisco Jesús Cepeda 41
Ruiz, Antonio Gómez 303
Ruiz, Raúl Martínez 84
Rusanova, Iryna 67

S

Saenzpardo, Raquel Portillo 145
Salas Canales, Pamela 313
Salazar-Colores, Sebastián 112
Salido-Ruiz, Ricardo Antonio 155, 164,
174, 192, 249
Salinas, Luis Antonio Jiménez 84
Sánchez-Monroy, Virginia 386, 403
Santiago-Fuentes, Laura Mercedes 275
Santos-Arce, Stewart René 155
Santos-Borráez, Alejandra 303

Soto, Alejandra 321
Stockton, Aarón Morales 145

T

Talavera-Peña, Ana Karen 275
Toriz-Vázquez, Alfonso 211
Torres Treviño, Luis Martín 183
Torres-Ramos, Sulema 155, 164, 174, 192,
249
Trujillo-Romero, Citlalli Jessica 258

V

Valadez-Roque, Gabriel 232
Valdés-Cristerna, Raquel 241
Varela-Rodríguez, Hugo 50, 137
Varela-Rodríguez, Luis 137
Vargas, Susana 331
Velasco, Arturo Abundes 84
Verdugo-Valerio, Aldair 50

Z

Zárate, Ana Lorena García 365
Zuñiga-Aguilar, Esmeralda 93
Zúñiga-Mondragón, Isaac 67