UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

REVISTA Ingeniería

Four-monthly scientific journal

2024

Volume 29 - Issue 2 ISSN 0121-750X E-ISSN 23448393

REVISTA Ingeniería

Volume 29 · Issue 2 · Year 2024 · ISSN 0121-750X · E-ISSN 2344-8393

Four-monthly Scientific Journal

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Carrera 7 No. 40-53
Edificio Administrativo
Piso 7 - Facultad de Ingeniería
Bogotá, Colombia
Teléfono: + 57 (1) 323 93 00 ext. 2413
Correo revista:
revista_ing@udistrital.edu.co

http://revistas.udistrital.edu.co/ojs/index.php/reving

Focus and Scope

The *Ingeniería* journal is an open-access academic-scientific online publication specialized in knowledge related to the fields of Engineering and Technology, according to the classification of scientific areas established by the Organization for Economic Cooperation and Development (OECD). Its main goal is to disseminate and promote debates about advances in research and development in the diverse areas of Engineering and Technology. We focus on disseminating original and unpublished articles in English that are relevant at both the national and international level. We seek to provide a platform to share knowledge and foster collaboration between researchers and professionals in the field. The journal does not request any Article Processing Charges (APC), as it is directly funded by Universidad Distrital Francisco José de Caldas. Moreover, the journal operates following a double-blind review process, thus ensuring the impartiality and quality of the published works. The journal is published in a continuous fashion and follows a quarterly numbering.

Open Access Policy

Article Processing Charge

No publication fees are charged to the authors or their institutions, nor are any payments made to expert peer reviewers or associate or adjunct editors. The *Ingeniería* journal is funded by Universidad Distrital Francisco José de Caldas, its Faculty of Engineering, and its Central Research Office.

Indexed

Scopus
SciELO
Publindex Colombia Categoría B
Redalyc
Directorio Latindex
Fuente Académica Premier (EBSCO)
Applied Science & Technology Source Ultimate (EBSCO)
Emerging Sources Citation Index Bibliographic Bases (SC)
Fuente Académica Plus (EBSCO)
OpenAlex
Catálogo 2.0 Latindex
DOAJ
Dialnet
ROAD
Google Schola Metrics
PKP Index
MIAR

Peer-reviewers in this issue

Andrés Fernando barajas
Universidad Francisco de Paula Santander, Colombia

Celso A. Sánchez-Ramírez
Universidad de Santiago de Chile, Chile

Eduardo Giraldo Suarez
Universidad Tecnológica de Pereira, Colombia

Firas Abedi
Al-Furat Al-Awsat Technical University, Iraq

Hermann G. de Meer
University of Passau, Germany

Hernández, Leonel
Institución Universitaria de Barranquilla, Colombia

Ismael López-Juárez
Centro de Investigación y de Estudios Avanzados del IPN (CINVESTAV), Mexico

Javier Ferney Castillo García
Universidad Santiago de Cali, Colombia

Jorge Enrique Espindola Diaz
Universidad Pedagógica y Tecnológica de Colombia

José Luis Saorín
Universidad de La Laguna, Spain

Junyong Zhu
USDA Forest Products Laboratory, United States

Juvenal Villanueva Maldonado
Consejo Nacional de Ciencia y Tecnología: Distrito Federal, Benito Juárez, Mexico

Rita Nagy-Kondor
University of Debrecen, Hungary

Roberto Ciavarella
Enea, Department of Energy Technologies and Renewable Sources, Italy

Deyby Huamanchahua Canchanya
Universidad de Ingeniería y Tecnología - UTEC, Perú

Stanislav H. Ivanov
Zangador Research Institute and Varna University of Management, Bulgaria

Wilmar Yesid Campo Muñoz
Universidad del Quindío, Colombia

Xiangmin Xie
Qingdao University and Qingdao Topscomm Communication Co., Ltd., China

Table of contents

Editorial

The Current Landscape of Access to Electrical Energy
César Leonardo Trujillo-Rodríguez

Biomedical Engineering

A Comparative Analysis between FFT, EMD, and EEMD for Epilepsy Detection
Leandro Dorado-Romero, Maximiliano Bueno-López, Jenny Alexandra Cifuentes

Automated Breast Tumor Detection and Segmentation Using the Threshold Density Algorithm with Logistic Regression on Microwave Images
Azhar Albaaj, Yaser Norouzi, Gholamreza Moradi

Education in Engineering

Scheduling in a Simple Assembly Robotics Cell to Minimize Earliness and Tardiness
John Andres Muñoz Guevara, Jairo Alberto Villegas-Florez, Jhannier Jhoan Jaramillo Tabima

Systems Engineering

Explainable Artificial Intelligence as an Ethical Principle
Mario González Arencibia, Hugo Ordoñez-Erazo, Juan-Sebastián González-Sanabria

Electrical, Electronic and Telecommunications Engineering

Mobile Application for Recognizing Colombian Currency with Audio Feedback for Visually Impaired People
Camila Bolaños-Fernández, Eval Bladimir Bacca-Cortes

Chemical, Food, and Environmental Engineering

Determining the Pulping Conditions and Properties of Unbleached Pulp from Uruguayan Pinus Taeda
Viviana Palombo, Leonardo Clavijo, Maria Noel Cabrera

Civil and Environmental Engineering

Community-Based Early Warning System Model for Stream Overflow In Barranquilla
Iván Andrés Felipe Serna-Galeano, Ernesto Gómez-Vargas, Julián Rolando Camargo-López

Industrial Engineering

Methodology for the Selection of Risk Response Actions while Considering Corporate Objectives in the Metalworking Industry
Álvaro Julio Cuadros-López, Alexander Bustos-Useche, Leonardo Bustos-Useche

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1221

DOI: https://doi.org/10.14483/23448393.22198

Editorial

The Current Landscape of Access to Electrical Energy

El panorama actual del acceso a la energía eléctrica

César Leonardo Trujillo Rodríguez[1]

[1]Universidad Distrital Francisco José de Caldas, Colombia

Universal access to affordable, reliable, sustainable, and modern energy is one of the Sustainable Development Goals (SDGs7) set for 2023. Although the number of people with access to electricity was close to 91 % as of 2021, it is concerning that, according to the energy progress report elaborated by the World Bank and the UN in 2023, 765 million people around the world lack this service.

The goal of taking electrical energy to all people by 2030 is difficult to achieve for several reasons, including high interconnection costs and low demand in certain areas (specially in African countries) stand out. This becomes evident when observing that, in a large number of countries (*e.g.*, Kenya) half of the users consume less than 15 kWh on average, which means losses for the utilities and the impossibility of interconnecting their customers, even if they are close to the generation points.

Electrical microgrids, a solution adopted several decades ago and a supposedly feasible alternative regarding access to electrical energy, has been put into question because many of these systems around the world are sub-utilized for a diversity of reasons, such as communities' lack of knowledge of this technology and operational issues in certain contexts. Isolated solutions for households involving solar panels, switching converters, and batteries follow the same path. However, these solutions provide energy to households at a high cost, in addition to their inability to supply large and productive loads.

The issue, in general terms, is not only reduced to providing all people with electrical energy to cover basic needs; it is necessary to ensure the provision of other services such as food security, access to water, and sanitary attention, etc. The issue is so complex that, in places where food is cooked by means of energy sources such as wood, coal, and kerosene, which are highly polluting, deadly effects are generated. This translates into approximately three million deaths around the world, according to the WHO.

Editorial

At this point, it is necessary to rethink the way to generate solutions regarding access to electrical energy that are sustainable in the long term. Designs adjusted to the needs of users must be elaborated, which not only contemplate aspects such as modularity, scalability, flexibility, interoperability, and efficiency, but also social, economic, and cultural aspects, so that the solution is understood by users and is tailored to their requirements.

Ultimately, it is essential to propose a technical-social approach where both scientific and social research centers converge, as well as local actors and the communities with which the corresponding initiatives will be developed. This is in order to create projects that are sustainable and have a positive impact on communities.

This constitutes an opportunity to build an electrical system that is decentralized, tailored to local needs, and supported by electronic conversion and storage technologies that are highly efficient and robust and facilitate a sustainable future for all humanity.

César Leonardo Trujillo Rodríguez

Alternative Energy Sources Research Laboratory, Department of Engineering, Universidad Distrital Francisco José de Caldas; Electronics Engineer, Master in Electrical Engineering, and Ph.D. in Electronic Engineering.
Email: cltrujillo@udistrital.edu.co

Editorial en español

El panorama actual del acceso a la energía eléctrica

El acceso universal a una energía asequible, fiable, sostenible y moderna es uno de los Objetivos de Desarrollo Sostenible (ODS7) trazados para el 2030. Aunque la cifra de personas con acceso a la electricidad a 2021 se acercaba al 91 %, es preocupante cómo, según el informe sobre progreso energético elaborado por el Banco Mundial y la ONU en el 2023, 765 millones de personas en el mundo carecen de este servicio.

El objetivo de llevar energía eléctrica a todas las personas para 2030 es difícil de alcanzar por diferentes razones, entre las que se destacan los grandes costos de interconexión y la baja demanda en ciertas zonas, principalmente en los países de África. Dicha situación se pone de manifiesto al observar que, en un gran número de países (*e.g.*, Kenia), en promedio la mitad de los usuarios consume menos de 15 kWh al mes, lo que implica pérdidas para las compañías eléctricas y la imposibilidad de conectar a sus clientes, así estén cerca de los puntos de generación.

Las microrredes eléctricas, una solución adoptada desde hace varias décadas y una alternativa supuestamente viable para el acceso a la energía eléctrica, se ha puesto en tela de juicio debido a que muchos de estos sistemas alrededor del mundo se encuentran subutilizados por diversas razones, como el desconocimiento de la tecnología por parte de la comunidad y problemas operativos en ciertos contextos. Transitan el mismo camino las soluciones aisladas para los hogares basadas en paneles solares, convertidores conmutados y baterías, las cuales suplen de energía eléctrica a las viviendas a un alto costo. Esto, aunado a la incapacidad de alimentar cargas grandes y productivas.

El problema, en términos generales, no solo se reduce a llevar energía eléctrica a todas las personas para suplir necesidades básicas; es necesario garantizar la prestación de otros servicios como la seguridad alimentaria, el acceso al agua y la atención sanitaria, *etc.* El problema es tan complejo que, en lugares donde la cocción de alimentos se hace a través de recursos energéticos como la madera, el carbón y el querosene, los cuales son altamente contaminantes, se generan efectos mortales, lo que cada año produce aproximadamente 3 millones de muertes en todo el mundo según la OMS.

En este punto, es necesario replantear cómo se generan soluciones de acceso a la energía eléctrica que sean sostenibles a largo plazo. Se deben realizar diseños ajustados a las necesidades de los usuarios, donde no solo se contemplen aspectos como modularidad, escalabilidad, flexibilidad, interoperabilidad y eficiencia, sino también aspectos sociales, económicos y culturales que permitan que la solución sea entendida por los usuarios y se ajuste a sus requerimientos.

En definitiva, es preciso plantear un enfoque técnico-social en el que confluyan centros de investigación social y científica, así como actores locales y las comunidades sobre las cuales se van a desarrollar las iniciativas correspondientes. Esto, con el fin de lograr proyectos que sean sostenibles y tengan un impacto positivo en la comunidad.

Esto constituye una oportunidad para construir un sistema eléctrico descentralizado, ajustado a las necesidades locales y soportado en tecnologías de conversión electrónica y almacenamiento que sean altamente eficientes y robustas y permitan un futuro sostenible para toda la humanidad.

César Leonardo Trujillo Rodríguez
Laboratorio de Investigación en Fuentes Alternativas de Energía, Facultad de Ingeniería, Universidad Distrital Francisco José de Caldas; Ingeniero Electrónico, Magíster en Ingeniería Eléctrica y Doctor en Ingeniería Electrónica.
Email: cltrujillo@udistrital.edu.co

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1221

DOI: https://doi.org/10.14483/23448393.21311

Research

A Comparative Analysis between FFT, EMD, and EEMD for Epilepsy Detection

Análisis comparativo entre FFT, EMD y EEMD para la detección de la epilepsia

Leandro Dorado-Romero[1]*, Maximiliano Bueno-López[1], and Jenny Alexandra Cifuentes-Quintero[2]

[1]Universidad del Cauca, Popayán, Colombia
[2]Universidad Pontificia Comillas, Madrid, España

Abstract

Context: Epilepsy is a neurological disease that affects more than 50 million people worldwide, causing recurrent seizures, with a significant impact on patients' quality of life due to abnormally synchronized neuronal activity.
Method: This article discusses three methods used for signal analysis in patients diagnosed with epilepsy. Conventional signal decomposition methods, such as the fast Fourier transform, widely used in signal analysis based on time series techniques, have some issues when analyzing nonlinear and non-stationary signals, in addition to difficulties in detecting low-order frequencies.
Results: To overcome these limitations, alternatives such as empirical mode decomposition and one of its variants, called *ensemble empirical mode decomposition*, have been developed. These techniques allow observing different oscillation modes through intrinsic mode functions and instantaneous frequencies.
Conclusions: In this study, the results obtained through the aforementioned techniques were compared, revealing the impact of nonlinear methods on the reconstruction of brain activity.

Keywords: electroencephalogram, empirical mode decomposition, epilepsy, instantaneous frequency, intrinsic mode functions, methodology, nonlinear, non-stationary, oscillation modes, seizures

Article history

Received: 27th/Sep/2023
Modified: 1st/Nov/2023
Accepted: 9th/Apr/2024

Ing, vol. 29, no. 2, 2024. e21311

* **Correspondence:** cesardorado@unicauca.edu.co

Resumen

Contexto: La epilepsia es una enfermedad neurológica que afecta a más de 50 millones de personas en todo el mundo, provocando convulsiones recurrentes, con un impacto significativo en la calidad de vida de los pacientes debido a actividad neuronal anormalmente sincronizada.
Métodos: Este artículo analiza tres métodos empleados para el análisis de señales en pacientes diagnosticados con epilepsia. Los métodos de descomposición de señales convencionales, como la transformada rápida de Fourier, ampliamente utilizada en el análisis de señales basado en técnicas de series de tiempo, presentan algunos problemas al analizar señales no lineales y no estacionarias, así como dificultades para detectar frecuencias de bajo orden.
Resultados: Para superar estas limitaciones, se han desarrollado alternativas como la descomposición empírica de modos y una de sus variantes, llamada *descomposición modal empírica de conjunto*. Estas técnicas permiten observar diferentes modos de oscilación mediante las funciones de modo intrínseco y las frecuencias instantáneas.
Conclusiones: En este estudio se compararon los resultados obtenidos mediante las técnicas mencionadas, revelando el impacto de los métodos no lineales en la reconstrucción de la actividad cerebral.

Palabras clave: electroencefalograma, descomposición epmírica de modos, epilepsia, frecuencia instantánea, funciones de modo intrínseco, metodología, no lineal, no estacionario, modos de oscilación, convulsiones

Table of contents

Page

1 Introduction 2

2 Methods 4
- 2.1 Dataset . . . 4
- 2.2 Signal decomposition methods . . 4
- 2.3 The fast Fourier transform (FFT) . 6
- 2.4 The Hilbert-Huang transform (HHT) . . . 6
- 2.5 Empirical mode decomposition EMD . . . 7
- 2.6 Ensemble empirical mode decomposition (EEMD) . . . 8
- 2.7 Instantaneous frequency (IF) 8

3 Results and discussion 8
- 3.1 Signal decomposition using FFT . 10
- 3.2 Signal decomposition using EMD . 10
- 3.3 Signal decomposition using EEMD 10

4 Conclusions 17

5 CRediT author statement 17

References 17

1 Introduction

Throughout the 20th century, important advances were made in the study of epilepsy, thanks to the use of increasingly sophisticated techniques and tools (1). Among these techniques, electroencephalographic (EEG) monitoring is one of the most widely used for the diagnosis and detection of epileptic seizures (2). EEG monitoring involves recording the electrical activity of the brain

via electrodes placed on the scalp, and it can provide valuable information about epilepsy-related brain activity (3). However, the processing and analysis of EEG data must consider the difficulties associated with large volumes of information (4). Even today, medical professionals identify epileptic seizures by visual inspection, aided by the continuous monitoring of EEG signals, which is time-consuming and subject to human error (5). In this context, several signal analysis techniques have been developed to identify epilepsy-related patterns in EEG records, where the reconstruction of brain activity from the resulting frequency bands is performed (6). Neuroscience has established five frequency bands associated with epilepsy: the delta band (0-4 Hz), the theta band (4-8 Hz), the alpha band (8-14 Hz), the beta band (14-30 Hz), and the gamma band (30-150 Hz) (7).

The identification of epileptic seizures still represents an unsolved challenge in the field of neuroscience, which is why different techniques are still being developed to support the recognition and treatment of this pathology (8, 9). However, due to the complexity and variability of seizures, different methods and tools are required for identification and classification tasks. Therefore, different methodologies have been developed which are based on the use of brain signal recording techniques, such as electroencephalography (EEG), magnetoencephalography (MEG), and functional magnetic resonance imaging (fMRI), as well as signal processing algorithms and machine learning techniques (10–12). These tools have proven to be effective in the identification and classification of seizures, allowing for a better diagnosis and treatment of epilepsy. These tools, derived from signal extraction methods, are mainly the Fourier transform (FT), the wavelet transform (WT), and the Hilbert Huang transform (HHT), as well as their variants (13). Depending on the signals to be analyzed, each of these methods has strengths and weaknesses (14).

The FT and its variants are ideal processing techniques for identifying signals in the time domain, aiming to transform a complicated problem into a solvable one. The FT is a basic tool in the analysis of non-periodic signals that have finite energy (15). However, it exhibits some issues when characterizing signals in short periods of time, as it is difficult to interpret the results (16).

The HHT has gained popularity in the analysis of nonlinear and non-stationary signals. It was initially used to study ocean waves, and its application has now been extended to various fields (17). A fundamental part of this method is empirical mode decomposition (EMD), a technique used for pattern identification in non-stationary and nonlinear signals (18). EMD can decompose signals into basic components known as *intrinsic modes*, which facilitates the identification of key patterns in signal analysis and frequency identification (19, 20).

EMD, based on the HHT, also reports some issues related to the mode mixing problem, which is associated with the mechanism for extracting mono-components from a multi-component signal. As a result, only modes that clearly contribute their maxima and minima can be identified via EMD.

When a mode cannot clearly contribute with extremes, EMD will not be able to separate it into an intrinsic mode function (IMF), and it will remain mixed with another IMF, turning it into noise and causing an inadequate interpretation of the results (21). To solve this problem, a variant known as

ensemble empirical mode decomposition (EEMD) is employed, where a signal called *adaptive noise* is added to each of the IMFs obtained via EMD (22).

In recent decades, considerable progress has been made in the study of epilepsy, including the advancement of sophisticated techniques for seizure detection. In particular, nonlinear and non-stationary signal analysis methods have shown effectiveness in automatic seizure detection, and they are often used to solve problems that cannot be solved via traditional approaches.

This paper compares three different methodologies for signal decomposition: FFT, EMD, and EEMD. The database is sorted, aiming for an organized dataset that aids in finding the most appropriate method, as well as the frequencies of interest for the detection of epileptic seizures. This document is structured as follows: Section 2 presents the signal decomposition methods and a description of the database used. Section 3 shows the results obtained with the proposed methodology, and Section 4 analyzes them. Finally, the conclusions derived from the study are stated in Section 5.

2 Methods

2.1 Dataset

The dataset used in the signal classification study was acquired from a freely available medical research data repository called *Physionet*, which contains signals collected from the scalp of intractable seizure patients at Boston Children's Hospital. This information has been published on the Physionet website (23).

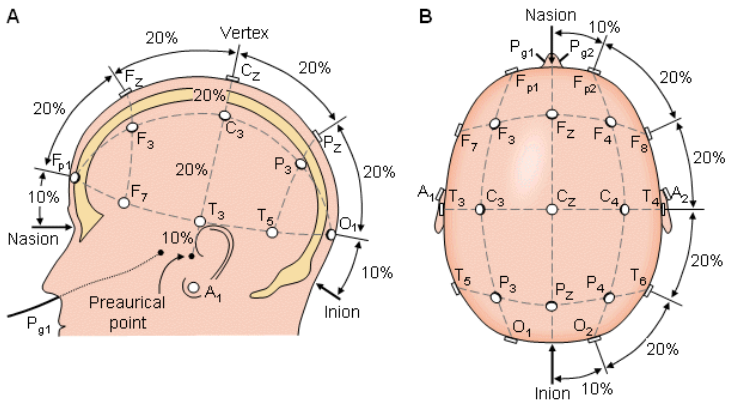
EEG signals were captured by placing electrodes on the scalp of patients using various configurations (24). The patients included both males and females in the ranges of 3-22 and 1.5-19 years old, respectively. The sampling frequency used was 256 samples per second, with a resolution of 16 bits. The nomenclature of the International 10-20 System was used to define the position of the electrodes on the scalp, with 23 channels available for each analyzed patient.

Fig. **1** shows the location of the electrodes according to the International 10-20 System Configuration.

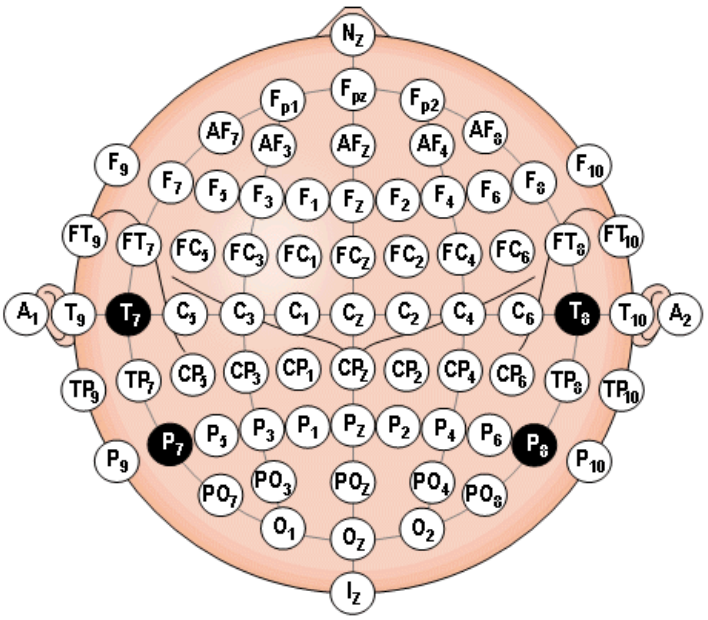
Experts in the field noted the onset and end of epileptic seizures for each of the recordings (26). In total, 941.6 h of interictal activity and 3 h of ictal activity were analyzed, which corresponded to the 181 seizures described in the database. From this database, we selected the signals with epileptic seizures to perform signal decomposition, obtaining 80 samples.

2.2 Signal decomposition methods

For many years, Fourier series analysis was ideal for studying signals in different fields, and it was assumed that conventional methods were sufficient to solve a particular problem (27). With the rise of new technologies, some issues emerged: the composition of signals was affected as more complex signals were being processed, which were assumed to be non-linear and non-stationary. These issues could be solved via traditional methods.



(a) The International 10-20 System seen from (a) the left and (b) above the head



(b) Location of electrodes seen from above the head

Figure 1. Location of electrodes according to the International 10-20 System (25)

Thus, new signal decomposition methods were developed, such as the fast and the short-term Fourier transforms (FFT and STFT), which performed better but had some limitations (28). Subsequently, HHT emerged as a solution to problems that could not be solved with the previous methods, and researchers were refining and perfecting new strategies that generated better solutions. In this context, EMD and EEMD, nonlinear decomposition methods each with a better response than the previous one, were created. A brief description of FFT, EMD, and EEMD is provided below.

2.3 The fast Fourier transform (FFT)

The FT is one of the most widely used tools in engineering applications where the behavior of dynamic systems and periodic signals is studied in the time domain while considering their frequency content (29).

The FFT is an algorithm that allows understanding the behavior of a dataset in the frequency domain, based on the calculation of the discrete FT. Nevertheless, it has disadvantages in spectral analysis, such as aliasing and leakage, as well as with regard to the computation times needed to process data. In addition, it is difficult to detect low frequencies due to its inefficient resolution (17).

Signal processing and learning methods such as machine learning are valuable tools in EEG signal analysis, as they have been developed to diagnose seizures and epileptic attacks. The research conducted by (30) allowed extracting a set of features from the original signals of two datasets to be analyzed using FT and EMD, with the purpose of classifying and evaluating EEG signals. These authors proposed a methodology to compare the behavior of signal analysis methods using classifiers in order to obtain specific features in the time domain.

2.4 The Hilbert-Huang transform (HHT)

The development of HHT arose from the need to describe distorted nonlinear waves combined with variations of the signals that naturally occur in non-stationary processes (31). HHT integrates EMD and HT (developed by Huang). In this case, the EMD method decomposes signals into single components called *IMFs*, from which it is possible to obtain the amplitude $a(t)$ and instantaneous frequencies. This method is widely employed for extracting information from a set of nonlinear and non-stationary data ($f_i(t)$) (32). After this procedure, HT is used to obtain the corresponding Hilbert spectrum (HS).

The HS is a 3D representation of the instantaneous amplitude and the instantaneous frequency for each IMF as a function of time. The HS is defined as follows (22):

$$H_i(f,t) \triangleq \begin{cases} a_i(t) & for \quad f = f_i(t) \\ 0 & otherwise \end{cases} \tag{1}$$

For a general multi-component signal, the HS is defined as the sum of the Hilbert spectra of all the IMFs, as indicated in

$$H(f,t) \triangleq \sum_{i=1}^{N} H_i(f,t) \tag{2}$$

where N is the total number of IMFs.

2.5 Empirical mode decomposition EMD

EMD is a tool for the analysis of nonlinear and non-stationary signals, which was proposed by Huang (33) and aims to decompose a nonlinear and non-stationary signal into a sum of IMFs while satisfying the following criteria (19,20):

1. In the entire data set, the number of extremes and the number of zero crossings must be equal or differ by 1 at most.
2. At any point, the mean value of the envelopes defined by the local maxima and the local minima is zero.

The second condition implies that an IMF is stationary, which simplifies its analysis. However, an IMF can exhibit changing amplitude and frequency modulation (34).

EMD has some issues, such as the presence of oscillations of unequal amplitude in one or more modes and similar oscillations in different modes, a phenomenon known as *mode mixing*. The screening process can be summarized in the following algorithm: decompose a dataset $x(t)$ into IMFs $x_n(t)$ and a residual $r(t)$, such that the signal can be represented as:

$$x(t) = \sum x_n(t) + r(t) \tag{3}$$

EMD has been adjusted to reduce the mode mixing phenomenon and thus ensure a better identification of the different frequencies in a process or system (35).

The EMD algorithm for a signal can be summarized as follows (32,36):

1. Identify all extremes in *x(t)*.
2. Calculate an upper envelope $e_u(t)$ and a lower envelope $e_l(t)$ via interpolation.
3. Determine the local averages as $m(t) = (e_u(t) + e_m(l))/2$.
4. Obtain the residual $r(t) = x(t) + m(t)$.
5. Iterate until the number of zeros equals the number of zero crossings.
6. Subtract the obtained IMF from the original signal.
7. Iterate over the residual until the function becomes monotonic.

The authors of (30), who used FFT and EMD for signal decomposition, state that it is useful to employ EMD and subsequently extract the IMFs defined by the components of amplitude modulation (AM) and frequency modulation (FM). Nonlinear and non-stationary complex signals can be decomposed into a finite number of IMFs in the spectrum of the Hilbert transform. The authors present the different decomposition modes of EMD. Different classifiers are trained and evaluated to find the best methodology.

2.6 Ensemble empirical mode decomposition (EEMD)

EEMD arose as an alternative to eliminate mode mixing, the main issue of the EMD. This approach consists of adding noise to the signal, known as *white Gaussian noise of finite amplitude*, where the real IMF is found as the average of several IMFs (37). Assuming that the added noise is different for each of the IMFs, averaging them over a certain number of attempts should cancel the noise, obtaining a single part and the true IMF.

The proposed algorithm for EEMD is:

1. Add white noise to the signal.
2. Decompose the data added with white noise, using the EMD to obtain the IMFs.
3. Repeat steps 1 and 2 with different white noise in each iteration.
4. Obtain the average of the corresponding IMFs from the decomposition as the final result.

A study conducted by (6) analyzed EEG recordings using EMD and EEMD with several classifiers to identify the IMFs that best represented an original signal. After the IMF selection process, a set of features was created using IMF1, IMF2, and IMF3. According to the authors, the objective was to propose a hybrid method for IMF selection and explore the effect of these IMFs. Their work evaluated the advantages of using EEMD, decomposing versions of signals with added noise to address EMD's mode mixing issues. This approach yielded better results.

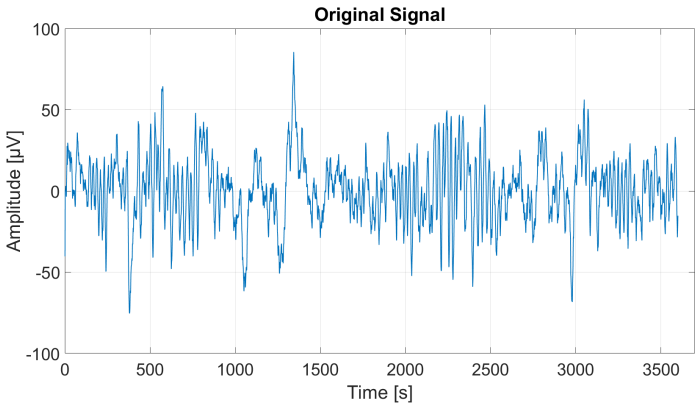
2.7 Instantaneous frequency (IF)

The concept of *instantaneous frequency* has become popular due to its effect on systems related to signal analysis, especially in nonlinear systems, where physical parameters derived from the signals are often characterized. Regardless of their behavior, most signals used to be analyzed via FT, which generates time-invariant frequency and amplitude values (35, 38). As per the analysis proposed by Fourier, the frequency of a signal is derived from its period. However, the frequency of a non-stationary wave is hard to define. It is also possible to define the frequency as the angular velocity associated with the phase change rate. If it is possible to define a phase for a signal, it is possible to calculate its frequency (*i.e.*, IF) (17).

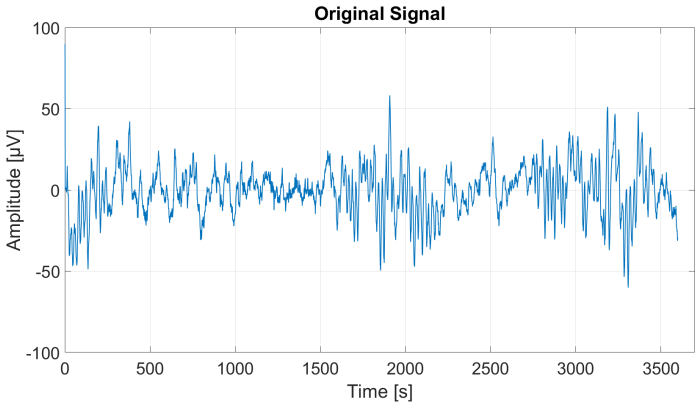
3 Results and discussion

To evaluate the performance of the aforementioned decomposition methods, three signals from patients suffering from epileptic seizures were taken. By analyzing the behavior of the original signal, maximum values could be identified at different points, as shown in Fig. 2, which shows the three original signals corresponding to said patients.

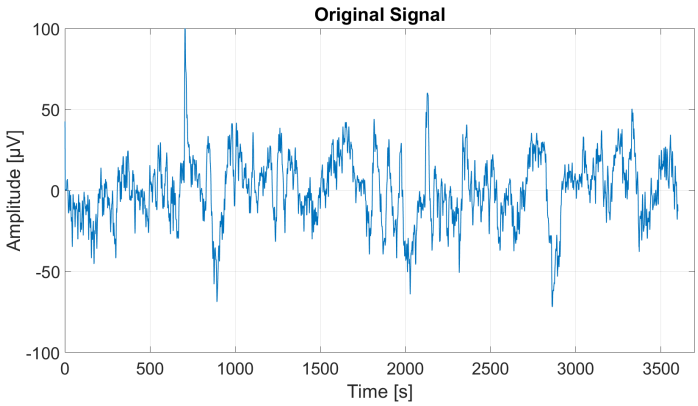
In signal 1, the maximum value was observed at time instant t = 1400 s. In signals 2 and 3, the maximum value was detected at t = 1800 and 700 s, respectively.



(a) Signal 1



(b) Signal 2



(c) Signal 3

Figure 2. Original signals

3.1 Signal decomposition using FFT

The FFT method was first implemented in MATLAB. Fig. 3 shows the results obtained after performing signal decomposition on three seizure signals in only one channel. The objective of this article is to show the frequency detected by each of the methods. Thus, three channels were selected, which allow for multiple frequencies.

Signals 1 and 3 have several similarities; three maxima stand out at the same time instants *i.e.*, t = 200, 1500, and 3000 s. This can be used to determine the points where seizures may be detected. After the instant t = 1500 s, the spectra does not grow as abruptly as before. Signal 2 does not exhibit notable maxima.

3.2 Signal decomposition using EMD

With the EMD method, the IMFs of each signal are obtained. EMD works as a natural filter of the original signals by separating them into frequency components, making it possible to observe the signals of interest. Fig. 4 shows the results obtained by applying EMD to the three signals.

In Fig. 4, IMFs 5 and 6 of signals 1, 2, and 3 clearly show the frequencies of interest for seizure detection. IMF 5 has maximum values of 15-17 Hz, and IMF 6 has maximum values of 8-9 Hz, representing the alpha and beta bands, respectively, *i.e.*, where seizures may occur and where seizures may be initiated. Meanwhile, IMFs 1 and 2 show noise activity, and IMF 4 shows a better separation between modes.

Subsequently, the results corresponding to the IFs of the analyzed signals are obtained. Fig. 5 shows these results.

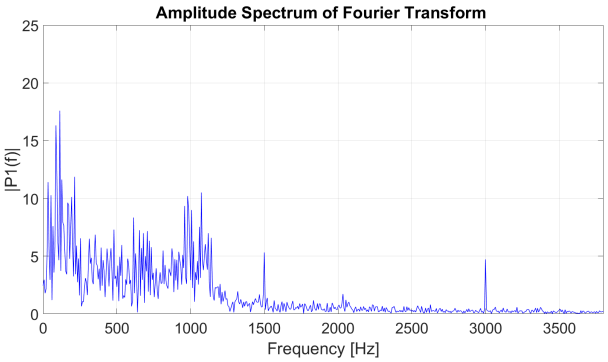
According to Fig. 5, the IF 5 of signals 1, 2, and 3 shows a maximum frequency value that oscillates between 10 and 15 Hz. This is in the alpha band and can be interpreted as an instance where a convulsion occurs.

3.3 Signal decomposition using EEMD

Fig. 6 shows the results obtained with EEMD, where noise is added to each of the signals to obtain a better separation of the IMFs and reduce mode mixing.

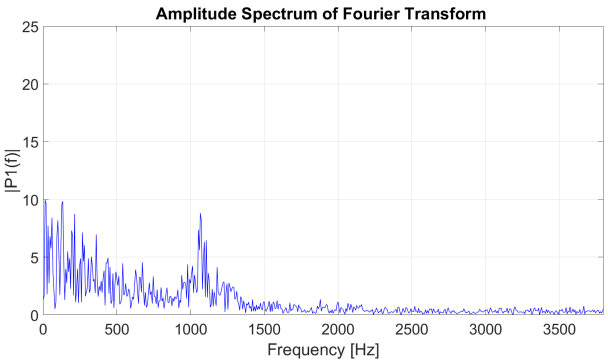
In Fig. 6, IMFs 4 and 5 show the frequency bands where epileptic seizures may be detected, confirming the results obtained with EMD and IF in Figs. 4 and 5, respectively.

Based on the results obtained via the three signal decomposition methods, the following can be stated. Regarding the use of FFT, the results in Fig. 3 do not allow for a clear visualization of the signals of interest due to interference. Although it is possible to observe two maxima in signals 1 and 3, their meaning cannot be determined, which makes it difficult to interpret the results and provide a conclusive



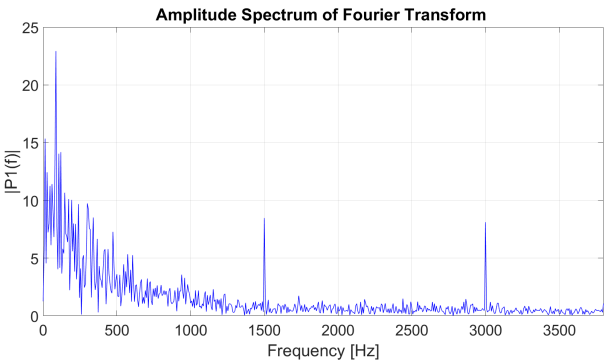
3.1

(a) Signal 1



3.2

(b) Signal 2

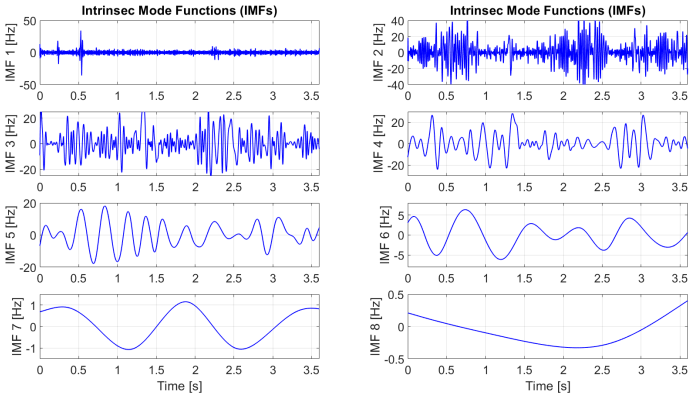


3.3

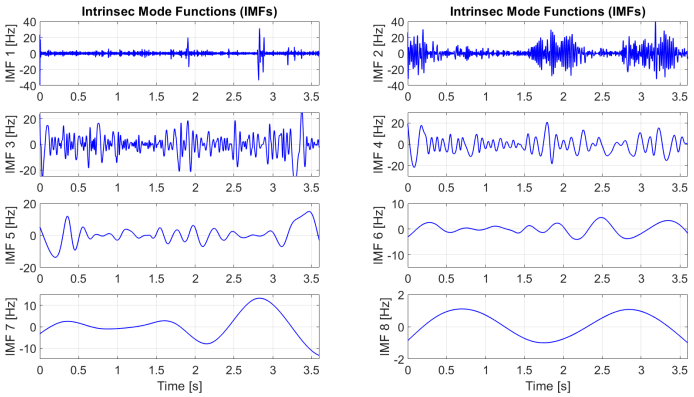
(c) Signal 3

Figure 3. Signal analysis using FFT

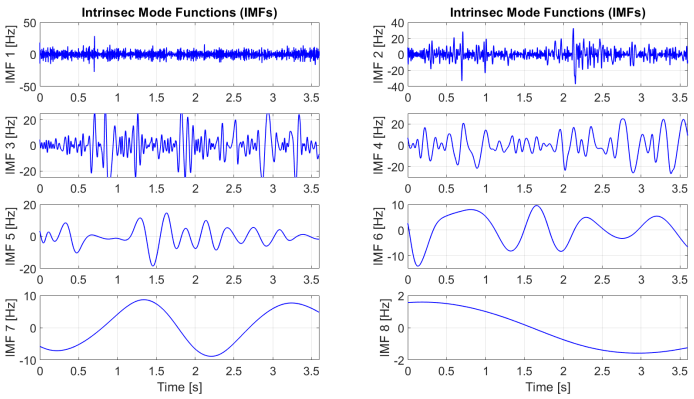
answer. The analysis supports the position of (14), suggesting that each of these methodologies has both advantages and limitations. One of the disadvantages associated with decomposition is its low resolution in the time-frequency domain, which is evident in Fig. 3. In addition, the results of the FFT show a low resolution due to the non-linearity of the signal.



(a) Signal 1

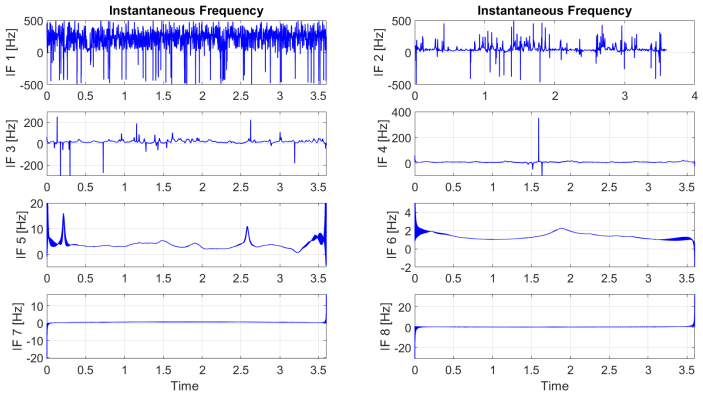


(b) Signal 2

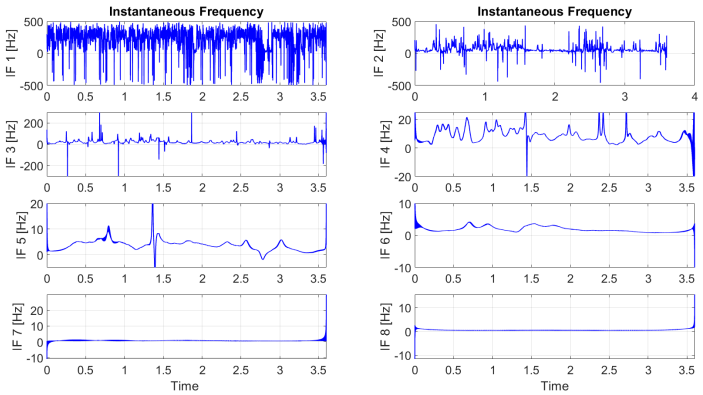


(c) Signal 3

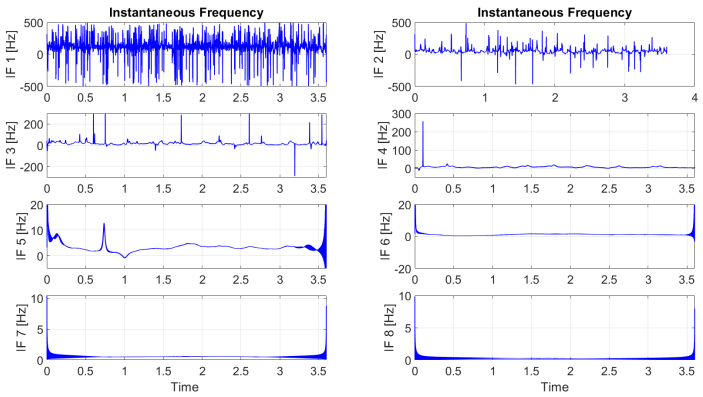
Figure 4. Signal analysis using EMD



(a) Signal 1

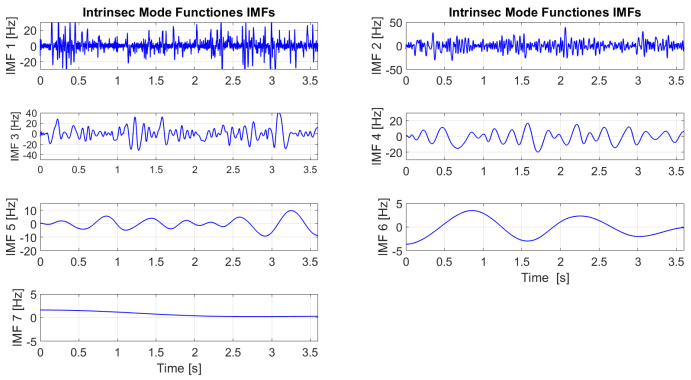


(b) Signal 2

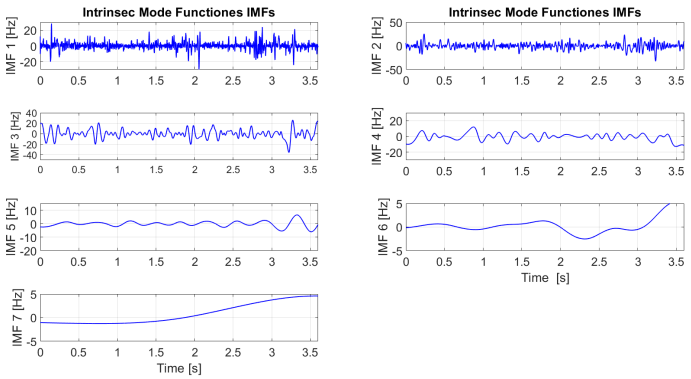


(c) Signal 3

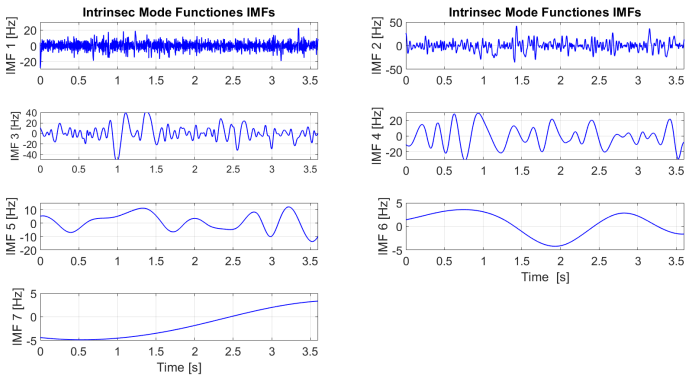
Figure 5. Signal analysis with IFs



(a) Signal 1



(b) Signal 2



(c) Signal 3

Figure 6. Signal analysis using EEMD

Furthermore, as emphasized in (15), the results demonstrate the challenges associated with interpreting signals using short time intervals due to their complex dynamics.

In this vein, FFT is not the best method to decompose non-linear and non-stationary signals where harmonic and non-harmonic components can appear. While FFT is still a valuable tool in many signal processing scenarios, its compatibility with signals exhibiting non-linearity and non-stationarity is notably sub-optimal. Thus, alternative methods that provide improved precision and adaptability to such signals should be explored.

On the other hand, when obtaining IMFs and IFs via EMD, the neuronal activity of each channel shows, in greater detail, the frequencies of interest for epileptic seizures occur. In this case, the alpha and beta frequency bands are those of greatest interest. The beta band is represented in IMF 6, and the alpha band in IMF 5. These frequency bands represent the main brain activity for our study, since they represent the onset and course of a seizure. According to the figures, the frequency increases and the amplitude decreases in IMF 5, whereas, in IMF 6, the frequency decreases and the amplitude increases.

Applying EMD has yielded notably improved results compared to the FFT method. Therefore, the authors of (30) chose to employ FFT and EMD to identify the most appropriate method for detecting epileptic seizures. Upon verifying the results, it became evident that employing EMD resulted in greater clarity and improved interpretation. These findings were further corroborated in (31), where the acquisition of IMFs validated the significance of conducting this procedure, as it enables the decomposition of a finite number of IMFs in the spectrum of the Hilbert transform, integrating both HT and EMD.

EMD's adaptability to non-linear and non-stationary signals provides a more accurate representation of the underlying components, enhancing the precision and interpretability of the results.

Using EEMD confirms the results obtained with the IMFs and the IFs of signals 1, 2, and 3. In IMF 5, the alpha band is of greater interest than the others because it enables the clear visualization of neuronal activity over time. Utilizing EEMD in the analysis yields better results than EMD and FFT. EEMD's ability to address non-linear and non-stationary signal characteristics, coupled with its noise robustness, significantly improved the quality and accuracy of our findings. In comparison with EMD, EEMD's ensemble approach introduces enhanced stability and precision with regard to mode extraction, mitigating issues such as mode mixing.

Validating the results of (6), we found that, by employing EMD while obtaining the IMFs, the EEMD yields results of higher reliability than those provided by EMD alone. These findings reinforce the notion that integrating EEMD into the analysis represents a significant advance in terms of the quality and accuracy of the results, which has important implications for the practical application of these techniques in the field of epileptic seizure detection and diagnosis.

Table I shows the most significant frequency values obtained using each of the signal decomposition methods for epileptic seizure detection. FFT does not allow determining specific frequency values, hindering the interpretation of the results.

Method		Frequency [Hz]
EMD	IMF5	15-17
	IMF6	8-9
IF	IF5	10-15
EEMD	IMF4	2-18
	IMF5	5-10

Table I. Frequency values obtained with the studied decomposition methods

Finally, Table II shows the strengths and weaknesses of the decomposition methods discussed above.

Method	Strengths	Weaknesses
FFT	Efficiency	Limited applicability for nonlinear and non-Stationary signals
	Widely used	Sample size constraints
	Accuracy	Fixed time frequency resolution
EMD	Adaptability to nonlinear and non-stationary signals	Computational intensity
	Intrinsic mode decomposition	Subject to mode mixing
	Higher time-frequency precision	Parameter dependency
EEMD	Robustness to noise	Increased computational load
	Stability through ensemble averaging	Complex parameter tuning
	Enhanced signal extraction	Dependent on domain knowledge

Table II. Strengths and weaknesses of signal decomposition methods

This study compared the efficacy of three signal processing techniques for epileptic seizure signal analysis. FFT excels in revealing frequency components but may miss nonlinear features. EMD captures nonlinear and non-stationary characteristics but is sensitive to noise, and EEMD, as an extension of EMD, offers improved noise robustness and adaptability to non-stationary signals. This comparative assessment aims to elucidate the strengths and limitations of these methods, aiding researchers in selecting the most suitable approach for their specific epilepsy research needs.

4 Conclusions

According to the state of the art and the research background, the brain can exhibit activity at frequencies between 0.5 Hz for Delta waves and 45 Hz for Gamma waves, and, by correctly detecting these frequencies, it is possible to diagnose different pathologies, with epilepsy being one of the most interesting. The use of time-frequency decomposition methods for EEG signals is generally applied in the study of brain processes associated with activity at certain frequencies, and it constitutes one of the main advantages of EMD and its variations (*i.e.*, EEMD).

EMD has shown the ability to separate signals using time-frequency decomposition in various contexts. For example, this method is widely employed to analyze nonlinear and non-stationary signals in fields such as medicine, power systems, image processing, weather forecasting, and climate analysis, among others.

The Fourier transform is a widely used method in different applications. However, it has some issues when it comes to dealing with low-order frequencies, at which epilepsy is detected. With the results presented in this article, it should be possible to establish a way to adequately detect epileptic seizures by calculating instantaneous frequencies.

5 CRediT author statement

Leandro Dorado-Romero: investigation, formal analysis, validation, visualization, software.
Maximiliano Bueno-López: investigation, conceptualization, formal analysis, methodology, supervision, writing (original draft).
Jenny Alexandra Cifuentes-Quintero: investigation, validation, visualization, software, writing (original draft).

References

[1] R. S. Fisher *et al.*, "Ilae official report: a practical clinical definition of epilepsy," *Epilepsia*, vol. 55, no. 4, pp. 475–482, 2014, https://doi.org/10.1111/epi.12550↑. 2

[2] A. T. Tzallas *et al.*, "Automated epileptic seizure detection methods: a review study," *Epilepsy–Histol. Electroencephalogr. Psychol. Asp.*, pp. 2027–2036, 2012, https://doi.org/10.5772/31597↑. 2

[3] K. M. Fiest *et al.*, "Prevalence and incidence of epilepsy: a systematic review and meta-analysis of international studies," *Neurology*, vol. 88, no. 3, pp. 296–303, 2017, https://doi.org/10.1212/WNL.0000000000003509↑. 3

[4] S. Roy, I. Kiral-Kornek, and S. Harrer, "A deep recurrent neural network for abnormal eeg identification," in *Proc. 17th Conf. Artif. Intell. Med (AIME), Poznan, Poland, Jun 26–29, 2019*, 2019, pp. 47–56, https://doi.org/10.1007/978-3-030-21642-9_8↑. 3

[5] L.-D. Guerrero, L. D. Romero, and M. Bueno-López, "A review of epileptic seizure detection using eeg signals analysis in the time and frequency domain," in *2021 IEEE 21st Int. Conf. Commun. Technol. (ICCT)*, 2021, pp. 1363–1367, https://doi.org/10.1109/ICCT52962.2021.9657835↑. 3

[6] O. Karabiber Cura, S. Kocaaslan Atli, H. S. Türe, and A. Akan, "Epileptic seizure classifications using empirical mode decomposition and its derivative," *Biomed. eng.*, vol. 19, pp. 1–22, 2020, https://doi.org/10.1186/s12938-020-0754-y↑. 3, 8, 15

[7] P. A. Munoz, E. Giraldo, M. B. López, and M. Molinas, "Automatic selection of frequency bands for electroencephalographic source localization," in *Proc. 2019 9th Int. IEEE/EMBS Conf. on Neural Eng. (NER)*, 2019, pp. 1179–1182, https://doi.org/10.1109/NER.2019.8716979↑. 3

[8] J. Engel Jr, "Evolution of concepts in epilepsy surgery," *Epileptic Disorders*, vol. 21, no. 5, pp. 391–409, 2019, https://doi.org/10.1684/epd.2019.1091↑. 3

[9] E. Ceballos Dominguez, M. Subathra, N. Sairamya, and S. Thomas George, "Detection of focal epilepsy in brain maps through a novel pattern recognition technique," *Neur. Comput. and Appl.*, vol. 32, pp. 10 143–10 157, 2020, https://doi.org/10.1007/s00521-019-04544-8↑. 3

[10] N. van Klink *et al.*, "Simultaneous meg and eeg to detect ripples in people with focal epilepsy," *Clin. Neuroph.*, vol. 130, no. 7, pp. 1175–1183, 2019, https://doi.org/10.1016/j.clinph.2019.01.027↑. 3

[11] G. Zazzaro, S. Cuomo, A. Martone, R. V. Montaquila, G. Toraldo, and L. Pavone, "Eeg signal analysis for epileptic seizures detection by applying data mining techniques," *IoT*, vol. 14, p. 100048, 2021, https://doi.org/10.1016/j.iot.2019.03.002↑. 3

[12] S. Kulaseharan, A. Aminpour, M. Ebrahimi, and E. Widjaja, "Identifying lesions in paediatric epilepsy using morphometric and textural analysis of magnetic resonance images," *NeuroImage: Clinical*, vol. 21, p. 101663, 2019, https://doi.org/10.1016/j.nicl.2019.101663↑. 3

[13] L. A. Moctezuma and M. Molinas, "Eeg channel-selection method for epileptic-seizure classification based on multi-objective optimization," *Front. Neurosci.*, vol. 14, p. 593, 2020, https://doi.org/10.3389/fnins.2020.00593x↑. 3

[14] F. McLoughlin, A. Duffy, and M. Conlon, "Evaluation of time series techniques to characterise domestic electricity demand," *Energy*, vol. 50, pp. 120–130, 2013, https://doi.org/10.1016/j.energy.2012.11.048↑. 3, 11

[15] E. Cano, R. Salcedo, and G. Soto, "Análisis de principios y aplicaciones de la transformada wavelet," *Universidad Nacional de Catamarca*, 2010. 3, 15

[16] J. Olbrys and M. Mursztyn, "Measuring stock market resiliency with discrete fourier transform for high frequency data," *Physica A*, vol. 513, pp. 248–256, 2019, https://doi.org/10.1016/j.physa.2018.09.028↑. 3

[17] M. Sanabria-Villamizar, M. Bueno-López, J. C. Hernández, and D. Vera, "Characterization of household-consumption load profiles in the time and frequency domain," *Int. J. Electr. Power Energy Syst.*, vol. 137, p. 107756, 2022, https://doi.org/10.1016/j.ijepes.2021.107756↑. 3, 6, 8

[18] R. Sharma, R. B. Pachori, and S. Gautam, "Empirical mode decomposition based classification of focal and non-focal eeg signals," in *Proc. 2014 Int. Conf. Med. Biometrics*, 2014, pp. 135–140, https://doi.org/10.1109/ICMB.2014.31↑. 3

[19] M. A. Colominas, G. Schlotthauer, P. Flandrin, and M. E. Torres, "Descomposición empírica en modos por conjuntos completa con ruido adaptativo y aplicaciones biomédicas," in *Proc. XVIII Cong Arg. Bioing. y VII Jorn. Ing. Clínica, Mar del Plata, Argentina*, 2011. 3, 7

[20] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mech. Syst. Signal Process*, vol. 35, no. 1-2, pp. 108–126, 2013, https://doi.org/10.1016/j.ymssp.2012.09.015↑. 3, 7

[21] O. B. Fosso and M. Molinas, "Emd mode mixing separation of signals with close spectral proximity in smart grids," in *Proc. 2018 IEEE PES innovat. smart grid technolog. conf. (ISGT-Europe)*, 2018, pp. 1–6, https://doi.org/10.1109/ISGTEurope.2018.8571816↑. 3

[22] M. Bueno-López, M. Molinas, and G. Kulia, "Understanding instantaneous frequency detection: A discussion of hilbert-huang transform versus wavelet transform," in *Proc. Int. Work-Conf. Time Ser. Anal.-ITISE*, vol. 1, 2017, pp. 474–486. 4, 6

[23] G. L. Goldberger AL, Amaral LAN, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," 2000, https://doi.org/10.1161/01.CIR.101.23.e215↑. 4

[24] C. Gómez, P. Arbeláez, M. Navarrete, C. Alvarado-Rojas, M. Le Van Quyen, and M. Valderrama, "Automatic seizure detection based on imaged-eeg signals through fully convolutional networks," *Sci. Rep.*, vol. 10, no. 1, pp. 1–13, 2020, https://doi.org/10.1038/s41598-020-78784-3↑. 4

[25] J. Malmivuo and R. Plonsey, *Bioelectromagn. EEG*, 01 1995, pp. 247–264. 5

[26] A. H. Shoeb and J. V. Guttag, "Application of machine learning to epileptic seizure detection," in *Proc. ICML*, 2010. 4

[27] M. A. Rodriguez, J. F. Sotomonte, J. Cifuentes, and M. Bueno-López, "A classification method for power-quality disturbances using hilbert–huang transform and lstm recurrent neural networks," *J. Electr. Eng. Technol.*, vol. 16, pp. 249–266, 2021, https://doi.org/10.1007/s42835-020-00612-5↑. 4

[28] M. Uyar, S. Yildirim, and M. T. Gencoglu, "An expert system based on s-transform and neural network for automatic classification of power quality disturbances," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 5962–5975, 2009, https://doi.org/10.1016/j.eswa.2008.07.030↑. 5

[29] F. Series, "Fourier series & fourier transforms," *BioMed. Eng.*, 2003. 6

[30] M. Zabihi, S. Kiranyaz, A. B. Rad, A. K. Katsaggelos, M. Gabbouj, and T. Ince, "Analysis of high-dimensional phase space via poincaré section for patient-specific seizure detection," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 3, pp. 386–398, 2015, https://doi.org/10.1109/TNSRE.2015.2505238↑. 6, 7, 15

[31] N. E. Huang, *Hilbert-Huang transform and its appl.* World Scientific, 2014, vol. 16, https://doi.org/10.1142/8804↑. 6, 15

[32] N. E. Huang *et al.*, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. Lond. A math. Phys. Eng Sci*, vol. 454, no. 1971, pp. 903–995, 1998, https://doi.org/10.1098/rspa.1998.0193↑. 6, 7

[33] P. N. Jadhav, D. Shanamugan, A. Chourasia, A. R. Ghole, A. Acharyya, and G. Naik, "Automated detection and correction of eye blink and muscular artefacts in eeg signal for analysis of autism spectrum disorder," in *Proc. 2014 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2014, pp. 1881–1884, https://doi.org/10.1109/EMBC.2014.6943977↑. 7

[34] M. Peel, G. Pegram, and T. McMahon, "Empirical mode decomposition: improvement and application," in *Proc. Int. Congress Model. Simul.*, vol. 1, 2007, pp. 2996–3002. 7

[35] M. Bueno-López, M. Sanabria-Villamizar, M. Molinas, and E. Bernal-Alzate, "Oscillation analysis of low-voltage distribution systems with high penetration of photovoltaic generation," *Electr. Eng.*, vol. 103, no. 2, pp. 1141–1154, 2021, https://doi.org/10.1007/s00202-020-01152-x↑. 7, 8

[36] C. Rehtanz, Y. Li, Y. Cao, and D. Yang, "Interconnected power systems: Wide-area dynamic monitoring and control applications," 2016. 7

[37] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 01, pp. 1–41, 2009. 8

[38] A. R. Messina, V. Vittal, G. T. Heydt, and T. J. Browne, "Nonstationary approaches to trend identification and denoising of measured power system oscillations," *IEEE Trans. Power Syst.*, vol. 24, no. 4, pp. 1798–1807, 2009, https://doi.org/10.1109/TPWRS.2009.2030419↑. 8

Leandro Dorado Romero

He earned a degree in Physical Engineering from the University of Cauca in 2018. He has recently completed his Master's studies in Automation at the same university, with a focus on the analysis of biomedical signals, particularly epileptic signals, with the purpose of automatically detecting them through signal decomposition methods while employing artificial intelligence techniques to obtain meaningful results.
Email: cesardorado@unicauca.edu.co

Maximiliano Bueno López

He obtained his Diploma in Electrical Engineering and a Master's Degree in Electrical Engineering from Universidad Tecnológica de Pereira, Colombia. He earned his PhD in Engineering from Universidad Autónoma de México. Following that, he worked as a postdoctoral researcher at the Norwegian University of Science and Technology-NTNU in the Department of Engineering Cybernetics. Currently, he holds the position of Associate Professor at the Universidad del Cauca in Popayán, Colombia.
Email: mbuenol@unicauca.edu.co

Jenny Alexandra Cifuentes

She holds a PhD in Automatics, as well as in Mechanical and Mechatronic Engineering. She has served as a visiting researcher at the University of Alberta in Canada and at the Technological Research Institute affiliated with Universidad Pontificia Comillas in Spain. Additionally, she has worked as a postdoctoral researcher at the Santander Big Data Institute (IBiDat) affiliated with Universidad Carlos III de Madrid, where she contributed to the formulation and development of research projects applied across various areas of data science. Her research focuses on pattern recognition using deep learning techniques.
Email: jacifuentes@gmail.com

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1221

DOI: https://doi.org/10.14483/23448393.20677

Research

Automated Breast Tumor Detection and Segmentation Using the Threshold Density Algorithm with Logistic Regression on Microwave Images

Detección y segmentación automatizadas de tumores de mama mediante el algoritmo de densidad de umbral con regresión logística en imágenes por microondas

Azhar Albaaj[1,2], Yaser Norouzi[1]*, and Gholamreza Moradi[1]

[1]Electrical Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran.

[2]Ministry of Health in Iraq, Najaf Health Directorate, Al-Hakim General Hospital, Najaf, Iraq.

Abstract

Context: Breast cancer remains a major health burden worldwide, necessitating improved screening modalities for early detection. However, existing techniques such as mammography and MRI exhibit limitations regarding sensitivity and specificity. Microwave imaging has recently emerged as a promising technology for breast cancer diagnosis, exploiting the dielectric contrast between normal and malignant tissues.

Objectives: This study proposes a novel computational framework integrating thresholding, edge segmentation, and logistic regression to enhance microwave image-based breast tumor delineation.

Methodology: The employed algorithm selects optimal features using logistic regression to mitigate the class imbalance between tumor and healthy tissues. Localized density thresholds are applied to identify tumor regions, followed by edge segmentation methods to precisely localize the detected lesions.

Results: When evaluated on a dataset of microwave breast images, our approach demonstrated high accuracy for detecting and segmenting malignant tissues. Density thresholds ranging from 0.1 to 0.8 showcase the highest accuracy in detecting breast tumors from these images.

Conclusions: The results highlight the potential of the proposed segmentation algorithm to improve the reliability of microwave imaging as an adjunct modality for breast cancer screening. This could promote earlier diagnosis and better clinical outcomes. The proposed framework represents a significant advance in developing robust image processing techniques tailored to emerging medical imaging modalities challenged by class imbalance and low intrinsic contrast.

Keywords: automatic segmentation, breast tumor, logistic regression, microwave images, threshold density

Article history

Received: 1st/Apr/2022

Modified: 6th/Nov/2022

Accepted: 5th/Dec/2023

Ing, vol. 29, no. 2, 2024, e20677

* **Correspondence:** y.norouzi@aut.ac.ir

Resumen

Contexto: El cáncer de mama sigue siendo una importante carga sanitaria a nivel mundial, lo que requiere mejores modalidades de cribado para la detección temprana. Sin embargo, las técnicas existentes, como la mamografía y la resonancia magnética, presentan limitaciones en cuanto a sensibilidad y especificidad. Recientemente, la imagen por microondas ha surgido como una prometedora tecnología para el diagnóstico del cáncer de mama, aprovechando el contraste dieléctrico entre los tejidos normales y malignos.
Objetivos: Este estudio propone un novedoso marco computacional que integra el umbralizado, la segmentación de bordes y la regresión logística para mejorar la delimitación de tumores mamarios basada en imágenes de microondas.
Metodología: El algoritmo empleado selecciona las características óptimas utilizando la regresión logística para mitigar el desequilibrio de clases entre los tejidos tumorales y sanos. Se aplican umbrales de densidad localizados para identificar las regiones tumorales, seguidos de métodos de segmentación de bordes para localizar precisamente las lesiones detectadas.
Resultados: Cuando se evaluó en un conjunto de datos de imágenes de microondas de mama, nuestro enfoque demostró una alta precisión para detectar y segmentar los tejidos malignos. Los umbrales de densidad que van desde 0.1 hasta 0.8 muestran la mayor precisión en la detección de tumores mamarios a partir de estas imágenes.
Conclusiones: Los resultados resaltan el potencial del algoritmo de segmentación propuesto para mejorar la fiabilidad de la imagen por microondas como modalidad complementaria para el cribado del cáncer de mama. Esto podría promover un diagnóstico más temprano y mejores resultados clínicos. El marco propuesto representa un avance significativo en el desarrollo de técnicas robustas de procesamiento de imágenes adaptadas a las modalidades emergentes de imagen médica desafiadas por el desequilibrio de clases y el bajo contraste intrínseco.

Palabras clave: segmentación automática, tumor de mama, regresión logística, imágenes por microondas, densidad de umbral

Table of contents

Page

1. **Introduction** 3
2. **Related works** 4
3. **Materials and methodology** 6
   - 3.1. Materials dataset 6
   - 3.2. Proposed methodology 8
   - 3.3. List of acronyms 11
4. **Logistic regression** 11
   - 4.1. Limitations of LR 13
   - 4.2. Deleting the border-connected regions 14
   - 4.3. Segmentation based on the edge . 14
5. **Simulation and results** 15
   - 5.1. Performance measures 16
6. **Conclusion** 20
7. **CRediT author statement** 20

References 20

1. Introduction

Breast cancer is the most prevalent malignancy among females, with around 287 850 novel cases predicted in 2022 in the United States, along with 43 250 mortalities (1). Breast tumors are the uncontrolled proliferation of aberrant cells obstructing normal breast tissue. This inhibits the development and functionality of typical cells in the breast while occupying physical volume within the breast (2). Although this disease is rare in the age group of 25 to 30 years, many cases have been reported at an early age. According to the World Health Organization's statistics, one in every 8-10 women will develop breast cancer (3). As per statistics in Iran, out of every 10-15 women, one is likely to develop breast cancer (4).

The interpretation of mammographies requires highly skilled radiologists, as different experts may report different variations of the same mammography. Mammography has a 68-79 % accuracy and necessitates expert radiologists due to inter-observer variability. Biopsies definitively determine malignancy, but they are expensive, invasive, and time-consuming (3). Motivated by the limitations of traditional breast cancer detection methods, such as MRI and mammography, the use of microwave imaging holds significant promise for early breast cancer detection (5). Microwave imaging offers improved sensitivity in detecting breast tumors by exploiting the electrical differences between benign, malignant, and healthy tissues, thereby enabling the identification of small tumors that may be missed by other modalities (6). Moreover, the non-ionizing nature of microwave frequencies makes them a safer option for screening, especially for high-risk individuals and young women who require frequent monitoring. The low-cost approach of microwave imaging can enhance accessibility to breast cancer screening in resource-limited settings. By avoiding breast compression during imaging, microwave imaging offers a more comfortable experience for patients, potentially leading to higher compliance with regular screenings. The crucial aspect of microwave imaging lies in its capacity to identify tumors in their initial phases, thereby enhancing the prognosis of breast cancer and diminishing mortality rates. Additionally, microwave imaging may offer better visibility of tumors in dense breast tissue, which is often challenging in mammography (7, 8). Advancements in microwave imaging technology and image processing algorithms have further enhanced its accuracy and reliability for breast tumor detection. Hence, exploring and harnessing the capabilities of microwave imaging can complement existing screening techniques, leading to improved breast cancer detection rates, earlier diagnosis. This ultimately allows saving lives (7, 9).

This study embarks on a pivotal journey, introducing an innovative approach to automated breast tumor detection through the lens of microwave images. The fundamental motivation for our research is twofold: to enhance the early detection of breast cancer and to overcome the limitations of existing imaging techniques. We address the inherent challenges of microwave breast imaging, particularly the struggle to distinguish tumors with indistinct boundaries and low contrast, which makes automated segmentation a formidable task. Our work extends the boundaries of medical image processing, with a primary focus on developing a robust algorithm for breast tumor detection and localization.

The central contribution of our research lies in the application of the logistic regression (LR) algorithm for feature selection, mitigating the impact of ineffective attributes. We meticulously delineate the steps of our approach, spanning image acquisition, noise reduction, and tumor region identification, ultimately building a framework that promises precision and resilience. We diligently evaluate our methodology using established metrics such as precision, the area under the receiver operating characteristic (ROC) curve, false and true positive rates, and accuracy. The findings of our study underscore the possibility of selecting the most effective features using LR, thus enhancing our understanding of the algorithm's real potential.

The remainder of this paper is organized as follows. The second section covers the pertinent related works. In the third section, a comprehensive examination of this work's materials and methodology is presented. The fourth section delves into LR. In the fifth section, we provide a detailed overview of the simulation and its results. Finally, some concluding remarks are presented in the sixth section.

2. Related works

The scientific literature on imbalanced classification problems and breast tumor prediction has seen numerous research efforts to develop robust and accurate methods for detecting breast tumors from medical images. Several fully automated methods for segmenting breast tumors in breast microwave images (BMI) have been proposed in recent years, some of which are presented below.

(10) developed an intuitive technique by merging empirical domain-specific expertise with low and high-level image features. Their technique employed a well-trained texture classifier and an active contour model, incorporating local edge information and global statistical data for the automated segmentation of breast cancers in BMI.

(11) presented a segmenting mammography method that employs the threshold and seed selection approaches. The employed algorithm searches for the largest value in each row of the pixel matrix to calculate the mean of these values, determining the threshold for segmentation.

The authors of (12) proposed a breast tumor detection and classification method based on density. They computed statistical properties such as mean, contrast, regularity, standard deviation, entropy, and homogeneity to select seed points, utilizing the mean of these statistical qualities as the first seed pixel.

However, their algorithm faced challenges in accurately detecting multiple tumors in mammography images.

The work by (13) implemented a method using the Wavelet transform, achieving a correct classification of the majority of breast tissue images with microcalcifications, mammary duct improvement, and areas of dense tissue.

(14) designed a breast MRI tumor segmentation algorithm using automated seed selection and particle swarm optimization (PSO) for image clustering. Their approach separated images into groups with similar intensity ranges, identifying the tumor region with the highest intensity value.

In (15), a saliency approach for the detection of malignancies was developed, aiming to identify areas of breast tissue with dark pigment clusters.

The paper by (16) presented a method for early breast cancer detection and classification using expectation maximization-based logistic regression (EM-based LR), achieving an average classification accuracy of 95,90 % and a performance index of 90.72 %. This study uses data from Kuppuswamy Naidu Memorial Hospital and employs the tumor node metasis (TNM) staging system of classification. However, while the paper demonstrates a well-structured and promising approach to breast cancer classification, it could benefit from a deeper explanation of the EM-based LR algorithm and the TNM staging system, as well as from comparisons with other classification techniques.

(17) aimed to improve the performance of LR by modifying the hypothesis used in the classifier and assigning a weighting factor (β) to the sigmoid function. This research delves into examining how the weighting factor's reliance is influenced by factors such as the number of features, the dataset size, and the optimization method applied. LR was selected due to its ability to offer a probabilistic interpretation, a feature not present in alternative classifiers like support vector machines (SVMs). The researchers adopted a hybrid strategy for the automatic localization of breast lesions in microwave imaging. This involves the integration of fuzzy sets, wavelet transform, pulse-coupled neural networks (PCNNs), and SVMs for classifying breast cancer from MRI images. While the approach is promising, it grapples with challenges stemming from its high dimensionality and computational intricacy. The authors employed both gradient descent and advanced optimization methods to minimize the cost function. Notably, they observed a noteworthy improvement in the accuracy of breast cancer detection by suitably selecting the value of β, which is determined by both the number of features and the optimization techniques employed.

(18) strove to develop a method that can accurately localize breast lesions using microwave images and a simplified PCNN. Their study involved the use of the MammoWave microwave imaging device, which employs two antennas rotating around the breast to capture microwave images. The images were processed using the PCNN algorithm for adaptive image segmentation and lesion detection. Furthermore, a non-parametric thresholding method was used to distinguish between breasts exhibiting no radiological observations, benign discoveries, and malignant findings (18). The findings indicate a 81.82 % sensitivity for the proposed method in detecting malignant findings, matching the gold standard for breast lesion localization. One strength of this work is the use of real clinical data acquired from the MammoWave system, which adds credibility to the findings. The study also provides detailed information on the patient population, the imaging procedure, and the gold standard used for breast classification. However, the study acknowledges limitations such as the relatively small dataset and the need for further validation with a larger number of cases.

Finally, (19) proposed a conditional generative adversarial network (cGAN) to segment dense tissues within mammograms. Additionally, they incorporated a convolutional neural network (CNN) for categorizing mammograms based on the standard set of the Breast Imaging-Reporting and Data System (BI-RADS). The cGAN excelled at segmenting dense regions in mammograms, boasting accuracy rates of 98 %, a Dice coefficient of 88 %, and a Jaccard index of 78 %. Employing the segmented masks of dense tissues generated by the cGAN, the classification network achieved remarkable figures regarding precision (97.85 %), sensitivity (97.85 %), and specificity (99.28 %) in breast density classification. This study drew from a collection of 410 images belonging to 115 patients from a breast dataset for screening mammography. In summary, to address the challenges of breast density estimation, including low contrast and significant fluctuations in the fatty tissue background of mammograms, the researchers leveraged cutting-edge deep learning methodologies such as cGANs and CNNs.

The above-presented works explored various image-processing techniques, statistical methods, and machine-learning algorithms to address the challenges of breast tumor detection and segmentation. However, many existing methods suffer from limitations when dealing with imbalanced classification problems in BMI and fail to accurately detect tumors in certain scenarios.

In this study, we present a novel methodology for detecting and segmenting breast tumors from microwave images using a combination of edge segmentation techniques and intensity thresholds. Our approach leverages the advantages of logistic regression and evaluates its effectiveness in selecting the most efficient features for breast tumor detection. By addressing the imbalanced classification problem, our method aims to achieve higher accuracy, sensitivity, and specificity in breast tumor prediction, leading to improved clinical practice and early diagnosis for better patient outcomes.

3. Materials and methodology

3.1. Materials dataset

A team of scholars from the University of Manitoba has made available an open-access test dataset for microwave breast sensing (https://bit.ly/UM-bmid). This dataset was obtained from the numerical models presented in (20). The University of Manitoba Breast Microwave Imaging Dataset (UM-BMID) encompasses details of 3D-printed breast models derived from 1 257 MRI-based scans, serving as a fundamental resource for extensive evaluations of BMI methods. The UM-BMID phantom scans were conducted utilizing a preclinical radar-based system for microwave breast imaging. To this effect, a vector network analyzer (VNA) was employed, which allowed for the generation of a stepped continuous-wave signal at 1 001 points covering the ultra-wideband (UWB) frequency range of 1-8 GHz. In the imaging system, a double-ridged horn antenna served as the essential component for conducting the survey. During the scanning process, 72 antennas were attached to a platform, rotating along a circular path 72 times, with S-parameters measured at each location. By utilizing the well-established UM-BMID dataset and employing advanced microwave imaging technology. Our study builds upon the foundation laid by (20) and provides further insights into microwave

breast sensing for improved tumor detection. To reconstruct images from each phantom scan, a set of beamformers was employed, namely delay-multiply-and-sum (DMAS), delay-and-sum (DAS), iterative delay-multiply-and-sum (itDMAS), and iterative delay-and-sum (itDAS). These beamformers collectively generated 12 reconstructions, each displaying identifiable tumor responses.

The UM-BMID is a publicly accessible repository that contains a vast amount of data. Our work then began by running a Python program to simulate and process image sets, from which a microwave image dataset was then extracted. This dataset comprises 18 distinct groups, each differing in terms of tumor presence or absence, along with varying tumor sizes for typical reconstructions.

To analyze the data, a total of 104 images were subjected to isolation and processing procedures. Out of these, 52 images were diagnosed with tumors, while the remaining 52 images represented normal tissues. Fig. **1** shows the reconstructed images, labeled as *DMAS (a, e)*, *DAS (b, f)*, *itDMAS (c, g)*, and *itDAS (d, h)*. In the initial row of images, reconstructed representations of a phantom scan housing a 10 mm tumor can be observed. The subsequent row displays images from a phantom with a 15 mm tumor. To provide context, a dashed white circle has been added, aiming to roughly delineate the outer edge of the breast phantom in all the reconstructions. Meanwhile, a solid green circle accurately designates the confirmed location of the tumor (21).

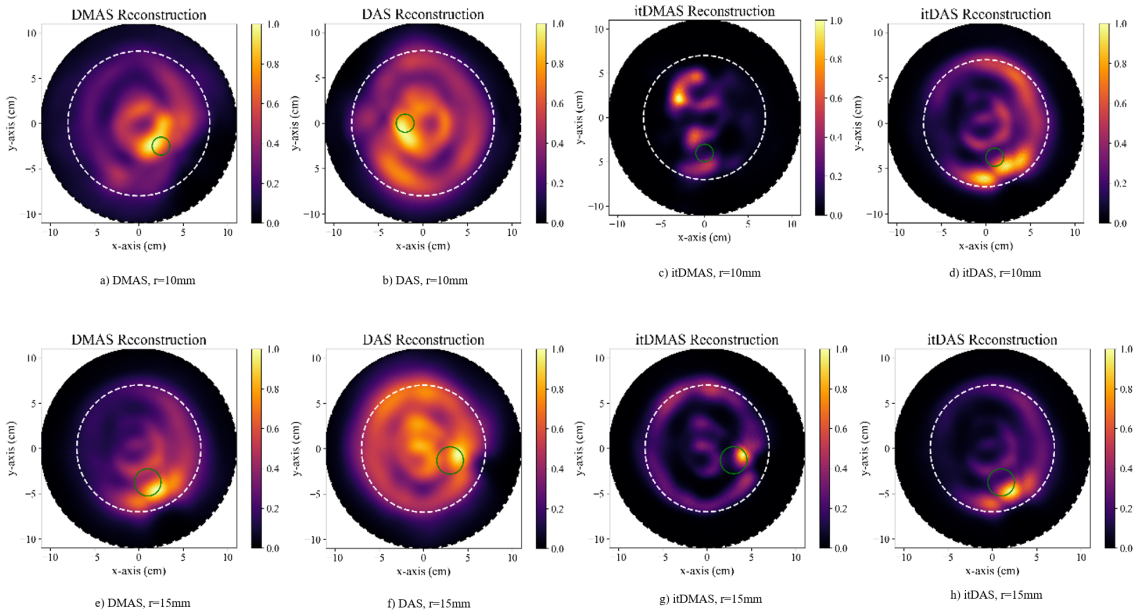


Figure 1. DMAS (a,e), DAS (b,f), itDMAS (c,g), and itDAS (d,h) reconstructed images (https://bit.ly/UM-bmid)

Region of interest (ROI) extraction is a crucial step aimed at identifying potential tumor locations in microwave images. To accomplish accurate tumor segmentation, a density thresholding technique was used to identify tumor regions based on density levels. Subsequently, edge segmentation techniques

were applied to further refine the identification of tumor regions by detecting local pixel intensity gradients.

In addition, LR was employed as a supervised learning technique for feature selection, in order to reduce the number of features used in tumor classification, thus enhancing the algorithm's efficiency. Finally, the algorithm's performance in detecting breast tumors was evaluated using various metrics, including sensitivity, specificity, accuracy, precision, the false positive rate (FPR), and the false negative rate (FNR).

By detailing these steps in the experimental setup, we aim to provide a comprehensive understanding of our proposal for breast tumor detection and segmentation via microwave imaging. These steps collectively contribute to the accuracy and effectiveness of our algorithm in clinical practice.

3.2. Proposed methodology

Our methodology primarily involves the use of LR for the automatic segmentation of breast tumors in microwave images. LR is a statistical model that employs maximum likelihood optimization to determine a weight vector. This weight vector represents a set of weights to be learned, where the constant term is represented by $x_0 = 1$. The LR model then calculates the subsequent distribution of the data using the Bayesian rule. The model also considers the effect of the weight vector on the probability and the subsequent distribution of the data. If the weight vector is greater than 0, the probability and distribution of the data increase as the variable xd increases. Conversely, if the weight vector is less than 0, the data probability and distribution decrease as the variable xd increases. This is mathematically explained in Section 4.1.

Our methodology also includes the deletion of border-connected regions. This operation is commonly used in image processing and computer vision tasks to clean up or isolate objects of interest. It involves removing specific parts of an image or a connected component in a binary image that touches the image's boundary. The steps to delete border-connected regions depend on the software or programming language used. Threshold density is another crucial aspect of the methodology, as it is used to determine the cut-off point for image intensity at the target (tumor) position. This threshold value is crucial for the accurate segmentation and detection of breast tumors in microwave images.

This section presents a detailed description of the proposed methodology. The technical approach of the algorithm involves a combination of edge segmentation techniques and intensity thresholds to identify and localize tumors within microwave images.

The *breast microwave signals* block refers to the raw microwave data acquired directly from the breast before any image reconstruction has been performed. This step represents the original signals obtained from the microwave imaging system. Subsequently, the *microwave image reconstruction (DMAS, DAS, itDMAS, and itDAS images)* block refers to reconstructed microwave images that are provided as inputs to our model. These reconstructed images are obtained through established reconstruction techniques, such as the aforementioned DMAS, DAS, itDMAS and itDAS, which solve the inverse

scattering problem. However, tackling the inverse scattering problem for image reconstruction from raw microwave data is outside the scope of our study. We take the reconstructed microwave images as given inputs, and our novel contributions are in the subsequent numerical processing steps, *i.e.*, preprocessing, segmentation, feature selection, and classification for accurate breast tumor detection and delineation.

Preprocessing

The image preprocessing stage involved a meticulous series of steps to ensure an optimal input for subsequent analysis. Initially, raw BMI were obtained from the UM-BMID. This open-access dataset, curated by scholars from the University of Manitoba, served as a cornerstone for our research endeavors.

Microwave images often contain noise and unwanted artifacts that can interfere with the tumor detection process. To address this issue, preprocessing techniques, including grayscale conversion, were applied to the already reconstructed microwave images of the breast, obtained through established reconstruction methods. This vital preprocessing step prepares the images for subsequent analysis and facilitates more accurate segmentation and classification. By emphasizing that preprocessing occurs after image reconstruction and before tumor segmentation and classification, we clearly delineate it as a distinct intermediate stage in the overall computational framework. Preprocessing transformed images enhance the performance of later steps like edge segmentation and density thresholding for precise breast tumor delineation. The *im2bw* MATLAB function was applied to convert the grayscale image into a binary black-and-white format. This binary representation effectively highlights the tumor regions and suppresses the background noise, making the segmentation process more robust.

Tumor region identification

The core of our methodology lies in the strategic identification of tumor regions within BMI. Leveraging the LR algorithm for feature selection, we meticulously eliminated attributes deemed linear and low-cost, preserving essential features using predefined weights. This process ensures a discerning separation between non-salient and salient features, contributing significantly to the algorithm's robustness.

To achieve precise tumor region identification, a hybrid approach combining threshold density techniques and edge segmentation methods was employed. The definition of intensity thresholds, coupled with advanced edge segmentation, refines the delineation of tumor boundaries. This hybrid strategy enhances the localization accuracy of detected tumors.

Edge segmentation techniques

Sophisticated edge segmentation techniques were employed to identify potential tumor regions within BMI. These techniques focus on detecting abrupt changes in intensity or gradient, indicative of tumor presence. The process begins with the use of the *bwlabel* function in MATLAB, performing

connected component labeling to identify distinct regions in the binary image. Each region, representing a group of connected pixels with similar properties, may signify potential tumor areas. Subsequently, the *regionprops* is employed to extract various properties from these regions, such as the density and area of non-black (white) pixels. This detailed information was seamlessly integrated into our threshold density algorithm, enhancing its ability to accurately delineate potential tumor regions from the surrounding tissue. This comprehensive approach, combining edge segmentation techniques with connected component labeling and properties extraction, contributes to the precision and sensitivity of our methodology, ultimately improving the early detection and diagnosis of breast tumors in BMI.

Tumor discrimination

A nuanced approach was implemented to distinguish tumors from surrounding regions in BMI. In this process, a crucial step involves defining a threshold intensity level, intricately linked to the density of non-black pixels within each region. This threshold serves as a discriminative factor, with high-density regions earmarked as potential tumor areas. The algorithm meticulously analyzes density values within each region, pinpointing the one with the highest density, a key indicator of tumor presence. This identified region is subsequently labeled as the tumor region, setting the stage for further in-depth analysis and precise localization. This methodical discrimination process enhanced the algorithm's ability to differentiate potential tumor regions from the surrounding anatomical structures, contributing to the overall efficacy of our breast tumor detection model in BMI.

Tumor localization

Our algorithm progresses beyond mere identification to achieve precise delineation of tumor boundaries within BMI. To this effect, the organ's function is strategically employed, leveraging the earlier detected tumor region as a reference point for accurate localization. This intricate step goes beyond basic segmentation, ensuring that the algorithm captures the nuanced boundaries of the tumor with a high level of accuracy. By aligning with the organ's function, our approach enhances the reliability of tumor localization, a critical aspect for subsequent diagnostic assessments and effective treatment planning. This detailed localization process significantly contributes to the overall robustness of our model, setting it apart in terms of accuracy and clinical relevance.

On the importance of segmentation

The significance of segmentation lies in its pivotal role within the proposed methodology. This step is critical since it involves the division of the image into distinct regions, allowing the algorithm to isolate comparable features. This division makes it notably easier to discern tumors from the background and other structures present in the image. The binary representation acquired through the thresholding process simplifies the subsequent identification and classification of pixels into tumor and non-tumor regions. This, in turn, greatly facilitates the precise localization of tumors within the overall image, significantly contributing to the accuracy and effectiveness of the entire tumor detection process.

Fig. 2 provides a visual representation of the entire methodology. It shows the sequential steps involved in the breast tumor detection and segmentation processes. The binary image obtained after preprocessing is fed into the edge segmentation techniques, which identify connected regions in the image. The tumor discrimination step determines the high-density region, which is then labeled as the tumor. Finally, the tumor localization step refines the segmentation to precisely locate the tumor boundaries, resulting in accurate tumor segmentation within the microwave image.

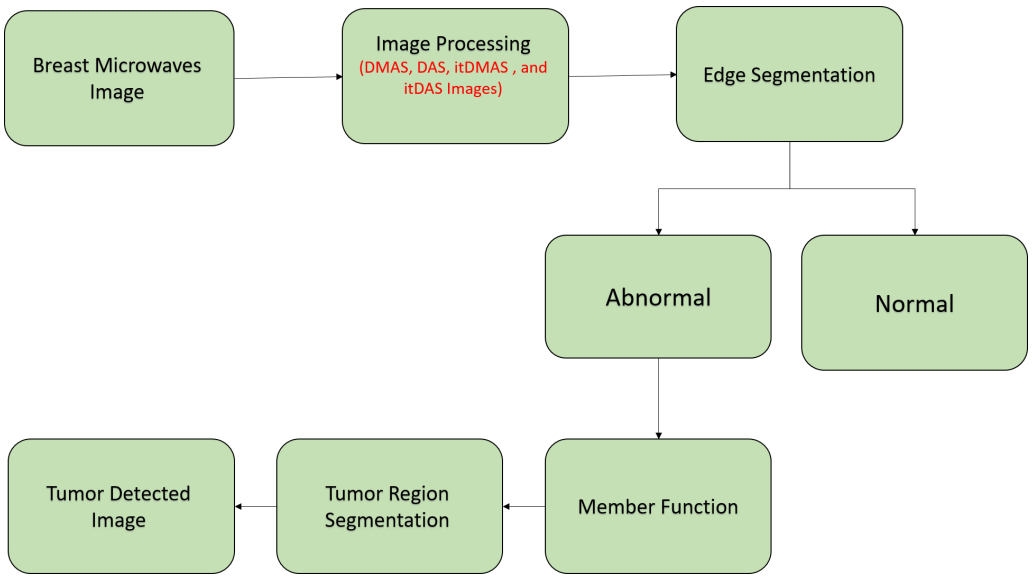


Figure 2. Overview of the proposed method

In conclusion, the proposed methodology combines edge segmentation techniques with intensity thresholds to detect and segment breast tumors from microwave images. By preprocessing the images, identifying potential tumor regions, discriminating tumors based on density values, and accurately localizing the tumor area, the primary objective of the algorithm is to amplify the accuracy and efficiency of breast tumor identification via microwave imaging.

3.3. List of acronyms

Table I presents the compilation of acronyms employed throughout this paper.

4. Logistic regression

LR is a statistical regression model for dichotomous dependent variables, *e.g.*, health and illness. This model is an extended linear model with *Logit* as its link function and a polynomial error distribution. There are two possible outcomes for a random event. For example, *buy or don't buy, registration* or not, and *going bankrupt* or not are variables that have only two statuses, for which the total probability will be eventually one. This method was primarily used in medicine for early-stage

Table I. Nomenclature

Acronym	Meaning
BMI	Breast Microwave Images
BI-RADS	Breast Imaging-Reporting and Data System
BUS	Breast Ultrasound
cGAN	conditional Generative Adversarial Network
CNN	Convolutional Neural Network
DAS	Delay-and-Sum
DMAS	Delay-Multiply-and-Sum
EM	Expectation Maximization
FA	False Alarm
FN	False Negatives
FNR	False Negative Rate
FP	False Positives
FPR	False Positive Rate
GUI	Graphical User Interface
GUIDE	Graphical User Interface Development Environment
itDAS	Iterative Delay-and-Sum
itDMAS	iterative Delay-Multiply-and-Sum
LR	Logistic Regression
MC	Missed Classification
MRI	Magnetic Resonance Imaging
PC	Perfect Classification
PCNN	Pulse Coupled Neural Network
PSO	Particle Swarm Optimization
ROC	Receiver Operating Characteristic
SVM	Support Vector Machines
TN	True Negatives
TNM	Tumor Node Metasis
TP	True Positives
UM-BMID	University of Manitoba Breast Microwave Imaging Dataset
UWB	Ultra-Wideband
VNA	Vector Network Analyzer

disease prediction. However, it is now widely used in all fields of science. LR is a subset of the general linear model and linear regression. It is based on the principle of different assumptions (regarding the connection between dependent and independent variables) rather than the LR model. Two properties of LR highlight the crucial distinction between these two models. First, due to the binary nature of the dependent variable, instead of a Gaussian distribution, the conditional distribution is of the Bernoulli type. Secondly, the forecast values are probabilities that are bounded between 0 and 1, which are derived using the logistic distribution function.

The main objective of feature selection involves reducing the count of attributes employed in a classification, all the while upholding a satisfactory level of accuracy in categorization (22). The selection of features decreases the size and execution time of input data when compared to the estimation method. Feature selection methods are divided into two groups: supervised and unsupervised. A supervised feature selection method evaluates a wide range of subsets of features using an evaluation function or criterion to select only those that are relevant to the decision classes of the data. Traditional feature selection methods are divided into two groups: filter and wrapper. Features are included or excluded in the filter method based on statistical metrics, feature weighting, mutual information, and various other considerations. Comparatively, assessing feature performance is independent of any learning model, making it a simpler and more expeditious process to execute (23).

LR was chosen for classification due to several important considerations:

- *Simplicity and interpretation.* LR offers a straightforward and interpretable approach, which makes it an accessible and understandable option for classification tasks.
- *Efficiency with binary classification.* LR efficiently deals with binary classification tasks, demonstrating its effectiveness in distinguishing between two categories.
- *Likelihood estimation.* LR provides probability outputs for each class, providing valuable insights into the probability of selected classifications, a feature that can be very useful for decision-making scenarios.
- *Low-dimensional data.* For datasets with relatively low-dimensional characteristics, LR proves computationally effective and efficient, allowing for smooth processing.
- *Feature importance.* LR allows identifying the most influential variables, providing valuable insights into the importance of a feature in the classification process.
- *Reduced overfitting.* LR tends to be less prone to overfitting, with smaller data sets and limited features, which enhances its reliability and robustness.

Ultimately, the selection of a machine learning algorithm must be in line with the unique characteristics of the dataset, the research goals, and the specific nature of the classification problem. All this depends on data complexity and volume, as well as on required performance metrics.

4.1. Limitations of LR

Consider the dataset $D = x_i, y_i : i = 1, \ldots, N, x_i \in \mathbb{R}^D$, and $y_i \in \{0, 1\}$. Let $\mathbf{x} = (x_1, \ldots, x_D)$ be the input vector, and let y be the binary class label. y can be either 1 or 0. The probability model underlying LR is as follows (24):

$$p(y = 1|x) = \sigma(s^T x) = \frac{1}{1 + \exp(-s^T x)} = \frac{1}{1 + \exp\left(-\sum_{d=0}^{D} s_d x_d\right)} \quad (1)$$

where $\mathbf{s}$ represents a vector of weights to be learned. Note that $x_0 = 1$ is a constant term. To determine the weight vector $\mathbf{s}$, LR employs maximum likelihood optimization.

$$\lambda(\mathbf{x}) = \ln \frac{p(y=1|x)}{p(y=0|x)} \tag{2}$$

Lemma 1. LR models the subsequent data distribution:

$$\lambda(x) = \sum_{d=0}^{D} s_d x_d \tag{3}$$

Using the Bayesian rule, we calculate the probability:

$$p(y=1|x) = \frac{p(y=1|x)}{p(y=1|x) + p(y=0|x)} \tag{4}$$

$$= \frac{1}{1+\exp(-\lambda(x))} \tag{5}$$

The lemma follows by comparing (1) with (5).

From Lemma 1, note that, in a LR model, a) if $s_d > 0$, then $p(y=1|x)$ and $\lambda(x)$ increase as x_d increases; and b) if $s_d < 0$, then $p(y=1|x)$ and $\lambda(x)$ reduce as x_d increases.

4.2. Deleting the border-connected regions

Deleting border-connected regions typically involves removing specific parts of an image or a connected component in a binary image that touches the image's boundary. This operation is commonly used in image processing and computer vision tasks to clean up or isolate objects of interest. The steps to delete border-connected regions depend on the software or programming language used.

In the work by (25), a computer-aided diagnosis (CADx) system is proposed for the automated segmentation and detection of masses in breast ultrasound (BUS) images. This method employs a two-step segmentation approach based on multi-resolution analysis. In the CADx system, the image-boundary-adjacent regions were directly removed (25). However, the lesion region is occasionally within the border. If the border-connected areas are immediately eliminated, the lesion region will also be eliminated. To avoid incorrectly removing regions, our work proposes a criterion for the center window that is about half the size of the total picture and centered in the image.

4.3. Segmentation based on the edge

When employing edge-based segmentation, the delineation of the image border or edge is determined by the local gradient of pixel intensity. The gradient serves as an approximation of the first-order derivative of the image function. To calculate the magnitude of the gradient for a given function $E(x, y)$, the following method can be applied:

$$|E| = \sqrt{[E_x^2] + E_y^2} \tag{6}$$

The gradient direction is denoted as

$$D = \tan^{-1}\left(\frac{E_y}{E_x}\right) \tag{7}$$

E_x and E_y represent gradients in directions x and y, respectively.

Edge-based approaches are computationally efficient and typically do not necessitate prior knowledge of the image's content. The general issue with this method is that the borders do not always entirely encircle the object. In this segmentation approach, direction and magnitude can be displayed as images. To structure closed boundaries between adjacent regions, a post-processing procedure joining or grouping edges is necessary.

5. Simulation and results

In this work, the MATLAB R2018b software was used to simulate the proposed breast tumor detection and segmentation approach. The GUI was built using GUIDE. For analysis purposes, real-time patient data were collected. Detecting and extracting a tumor from a breast microwave image is relatively straightforward due to the tumor's elevated intensity compared to its background. Fig. 3a depicts the BMI from the open-access dataset, Fig. 3b depicts the BMI for the tumor alone, and Fig. 3c depicts the image of the tumor that was segmented using the proposed methods. In another case, we

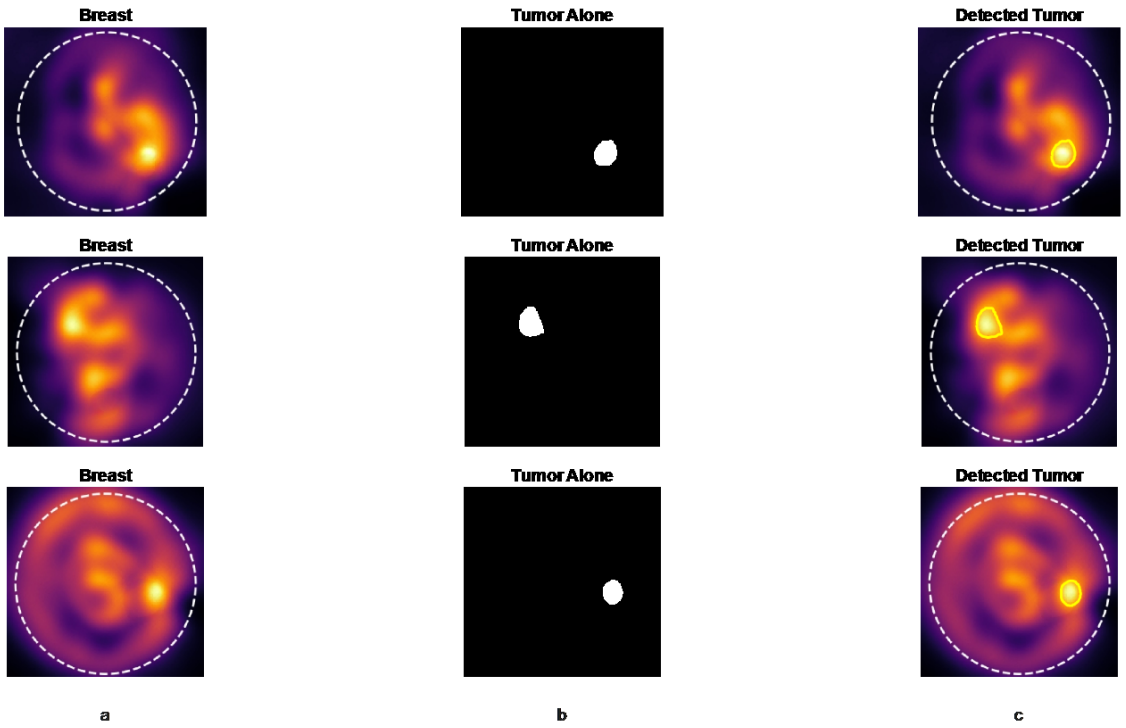


Figure 3. a) Source BMI; b) tumor; c) tumor region segmentation using the proposed method

inserted a microwave image of the breast without any indication of a tumor. As a result, no tumor was detected or identified. This is illustrated in Fig. 4.

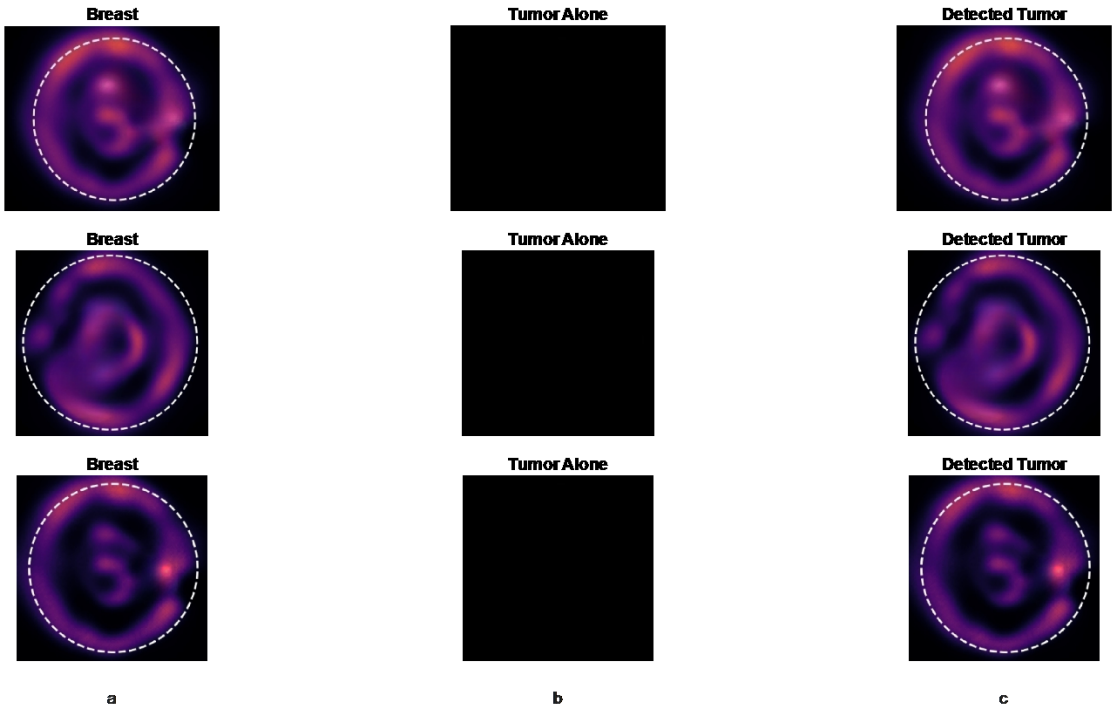


Figure 4. a) Source BMI; b) no tumor; c) no tumor identified

5.1. Performance measures

The performance evaluation of our proposed methodology involved a comprehensive set of metrics, including precision, the area under the receiver operating characteristic (ROC) curve, the false positive rate, the true positive rate, and accuracy. Systematic calculations of these metrics collectively offer a quantitative and qualitative assessment of the algorithm's ability to detect and localize breast tumors. The overall performance of the proposed design was determined by an evaluation matrix comprising true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), as illustrated in Table II. This matrix serves as a key component in evaluating the robustness and effectiveness of our algorithm in breast tumor detection.

Table II. Confusion matrix for the tumor detection algorithm

Logistic Regression – Confusion matrix for the tumor detection algorithm		
Original microwave scan/Algorithm results	Tumor detected (1)	Tumor not detected (Accurately) (0)
Tumor present (1)	True positive (TP)	False negative (FN)
Tumor not present (0)	False positive (FP)	True negative (TN)

This study employed sensitivity (Se), specificity (Sp), accuracy (Acc), and other metrics that are commonly utilized in many medical imaging applications.

The LR statistics were calculated as follows (12):

$$Se = \frac{TP}{TP + FN} \tag{8}$$

$$Sp = \frac{TN}{TN + FP} \tag{9}$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

$$FPR = 1 - Sp = \frac{FP}{FP + TN} \tag{12}$$

$$FNR = 1 - Se = \frac{FN}{TP + FN} \tag{13}$$

In the context of image classification, TP corresponds to images correctly identified as breast tumor images, while the FP refers to non-tumor images inaccurately labeled as breast tumor images. Model accuracy signifies the proportion of accurately classified images among the total number of tested images. In the realm of machine learning and deep learning models, accuracy pertains to the ratio of correctly classified samples (images) to the overall count of test images. This metric serves as a prevalent evaluation measure for assessing the performance of classification models. An increased accuracy value suggests that the model adeptly categorizes test data. TN designates appropriately classified standard images. FN quantifies the erroneously identified non-tumor pixels. The FPR refers to cases where test results indicate that a condition is present when, in fact, it is not. The FNR denotes test outcomes where the results indicate the absence of a condition, despite its actual presence.

We can calculate the statistical expressions for the performance analysis as follows (16):

$$Se = \frac{PC}{PC + FA} \times 100\,\% \tag{14}$$

$$Sp = \frac{PC}{PC + MC} \times 100\,\% \tag{15}$$

$$Acc = \frac{\text{Sensitivity} + \text{Specificity}}{2} \tag{16}$$

$$Performance\ index = \frac{PC - MC - FA}{PC} \tag{17}$$

In a perfect classification (PC) with are no errors, the total number of PCs would be equal to the TPs plus the TNs, since a perfect model would correctly identify all the actual positive cases as positive (TP) and all the actual negative cases as negative (TN). False alarms occur when the model erroneously forecasts the positive class despite the actual class being negative. These instances are also recognized as FPs. The sensitivity in Eq. (14) was calculated in this way. The key difference of this equation is that the denominator contains FPs rather than FNs, so, while the standard sensitivity equation says how

many actual positives were correctly found, this version says how many positive predictions were correct. The issue with using FPs in the denominator is that it penalizes the sensitivity score if the number of false alarms goes up. FNs refer to missed positive classifications when an actual positive case is incorrectly classified as negative. Thus, $MC = FN$.

The performance index (PI) considers the total number of correct positive classifications along with the number of false negatives (MC) and false positives (FA) to assess the algorithm's performance. A PI value closer to 1 indicates a better-performing algorithm, as it suggests a higher percentage of correct classifications compared to the total number of classifications made. A higher PI value would indicate that the proposed algorithm has is better at accurately detecting and segmenting breast tumors from microwave images, making it more effective for clinical applications.

The sensitivity and specificity of the classification statistics derived from the performance analysis were used to determine the efficacy of the proposed breast tumor detection method. High sensitivity and specificity values indicate that the proposed approach is effective for clinical use.

The results show that the 0.1 and 0.6 density thresholds were 100 % accurate at detecting breast tumors from microwave images. 0.1-0.8 were decently accurate, but 0.8 and 0.9 did not quite manage to detect the presence of tumors (Fig. 5).

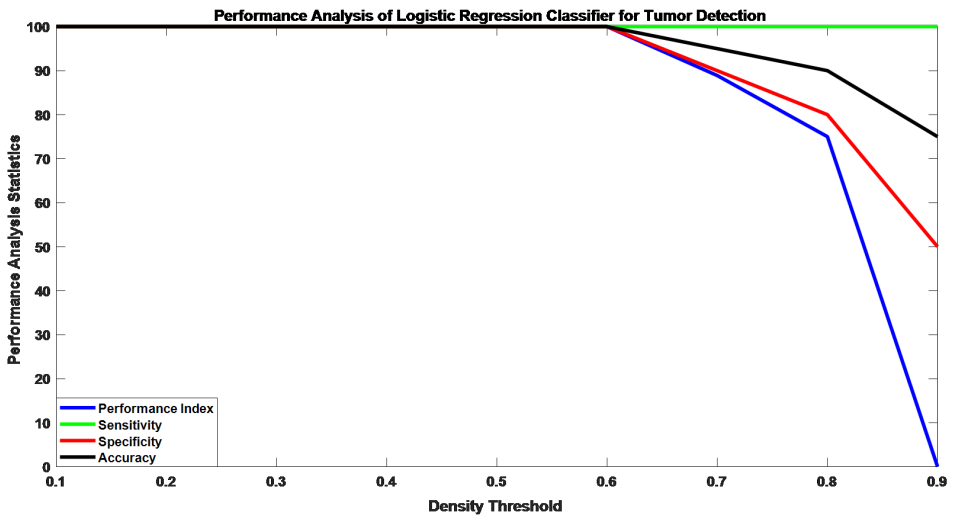


Figure 5. Performance analysis statistics of LR and density threshold for breast tumor detection for 10 samples (images)

The results also showed that, when the density threshold was between 0.1 and 0.6, the PC was 100 % accurate in identifying images with tumors. We also obtained a good ideal classification when the density threshold was between 0.1 and 0.8, and *vice versa*. When the density threshold is high, 0.8-0.9 makes the percentage of missing classification very high, meaning that it is not possible to classify images that contain tumors (Fig. 6).

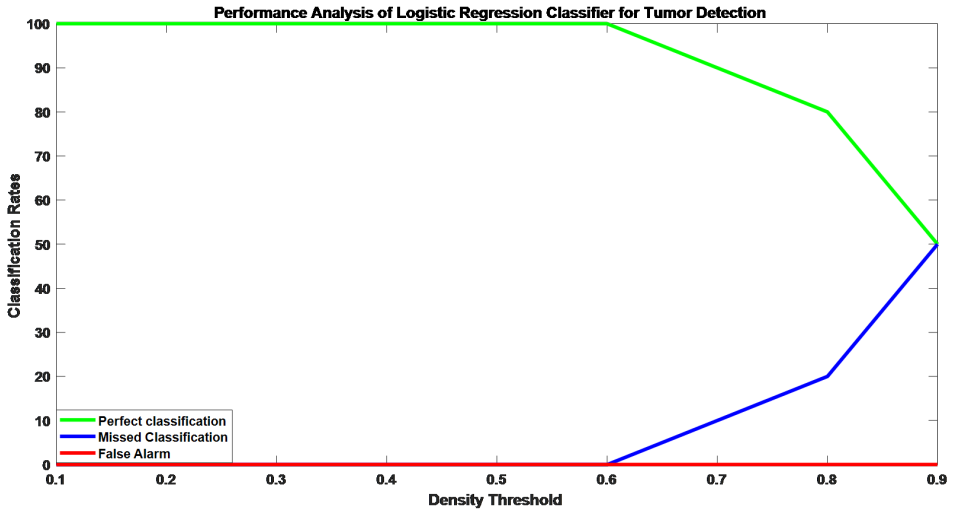


Figure 6. Classification rates predicted via LR and density threshold for breast cancer diagnosis

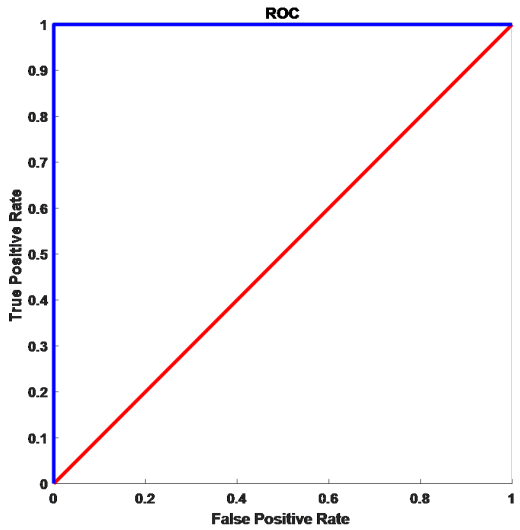


Figure 7. ROC curve of the proposed method

More attention is to be paid on various points in the ROC curve. The point (0, 0) in the lower-left corner represents a technique when the classifier does not generate positive results. The point in the upper right corner depicts the opposing technique, which always yields positive results (1, 1). The (0, 1) point represents the optimal classifier. In general, a perfect ROC curve is achieved when the classifier can completely separate the two classes, resulting in no overlap between the TPR (sensitivity) and the FPR (1-specificity). In this scenario, the TPR is always 1, and the FPR is always 0, regardless of the threshold

selected. Since the TPR is the same as sensitivity and is given as 1 in the curve, and the FPR is equal to 1-specificity and is also given as 0, it implies that all threshold values in Fig. 7 result in PC. Therefore, any threshold value chosen for classification, in this case, would yield a perfect ROC curve.

6. Conclusion

The detection of cancer at an early stage is necessary for reducing mortality rates. The efficacy of breast cancer detection is contingent on the precision of tumor segmentation. When segmenting microwave images, the ROI must be extracted since it plays a vital role in the efficacy of tumor detection. This work presents a method for automatically extracting ROI from microwave breast images. These microwave images were preprocessed, which reduces noise and increases contrast. Microwave scans reveal the size and location of a breast tumor. This information allows radiologists to easily diagnose the tumor and prepare its surgical removal. In order to address the challenges in evaluating breast tumors, an automatic detection method using microwave images was developed. Density thresholds ranging from 0.1 to 0.8 showed the highest accuracy in detecting breast tumors from these images. The potential for enhanced robustness lies in incorporating additional data facets into the algorithm, an augmentation that could empower the program to handle diverse breast tumor scans more adeptly. This advanced imaging processing tool has the potential to assist medical professionals in laser ablation procedures. As we look to the future, it is important to address the influence of bright and dense outer tumor edges on the detection algorithm. Implementing a filtering technique to mitigate their impact could further refine the accuracy of the detection process.

7. CRediT author statement

Azhar Albaaj: conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing (original draft), visualization.
Yaser Norouzi: conceptualization, methodology, validation, formal analysis, resources, writing (review and editing), supervision, project administration.
Gholamreza Moradi: conceptualization, methodology, resources, writing (review and editing), supervision, project administration.

References

[1] E. M. Proussaloglou *et al.*, "Updates in the pathology of pregnancy associated breast cancer (PABC)," *Pathol. Res. Pract.*, art. 154413, 2023. https://doi.org/10.1016/j.prp.2023.154413 ↑3

[2] T. H. Aldhyani *et al.*, "Deep learning model for the detection of real time breast cancer images using improved dilation-based method," *Diag.*, vol. 12, no. 10, art. 2505, 2022. https://doi.org/10.3390/diagnostics12102505 ↑3

[3] Z. Khandezamin, M. Naderan, and M. J. Rashti, "Detection and classification of breast cancer using logistic regression feature selection and GMDH classifier," *J. Biomed. Inform.*, vol. 111, art. 103591, 2020. https://doi.org/10.1016/j.jbi.2020.103591 ↑3

[4] A. Najafian-Najafabady, N. Ebrahimi, and S. Vallian, "rs2682818/MiR-618 is a novel marker associated with increased risk of breast cancer in the Iranian population," *Arch. Biol. Sci.*, vol. 73, no. 4, pp. 457–463, 2021. https://doi.org/10.2298/ABS210808039N ↑3

[5] L. Wang, "Holographic microwave image classification using a convolutional neural network," *Micromachines*, vol. 13, art. 2049, 2022. https://doi.org/10.3390/mi13122049 ↑3

[6] M. A. Aldhaeebi *et al.*, "Review of microwaves techniques for breast cancer detection," *Sensors*, vol. 20, no. 8, art. 2390, 2020. https://doi.org/10.3390/s20082390 ↑3

[7] S. Kwon and S. Lee, "Recent advances in microwave imaging for breast cancer detection," *Int. J. Biomed. Imaging*, vol. 2016, art. 5054912. https://doi.org/10.1155/2016/5054912 ↑3

[8] L. Wang, "Microwave imaging and sensing techniques for breast cancer detection," *Micromachines*, vol. 14, no. 7, art. 1462, 2023. https://doi.org/10.3390/mi14071462 ↑3

[9] N. AlSawaftah *et al.*, "Microwave imaging for early breast cancer detection: Current state, challenges, and future directions," *Journal of Imaging*, vol. 8, no. 5, art. 123, 2022. https://doi.org/10.3390/jimaging8050123 ↑3

[10] L. Liu *et al.*, "Automated breast tumor detection and segmentation with a novel computational framework of whole ultrasound images," *Med.l Biol. Eng. Computing*, vol. 56, pp. 183–199, 2018. https://doi.org/10.1007/s11517-017-1770-3 ↑4

[11] A. Melouah and S. Layachi, "A novel automatic seed placement approach for region growing segmentation in mammograms," in *Proc. Int.l Conf. Intel. Info. Processing Sec. Adv. Comm.*, 2015, art. 51. https://doi.org/10.1145/2816839.2816892 ↑4

[12] N. Shrivastava and J. Bharti, "Breast tumor detection and classification based on density," *Multimedia Tools App.*, vol. 79, no. 35-36, pp. 26467–26487, 2020. https://doi.org/10.1007/s11042-020-09220-x ↑4, 17

[13] A. A. Sandino Garzón and R. Herrera García, "Clustered microcalcifications candidates detection in mammograms," *Ingeniería*, vol. 24, no. 2, pp. 159–170, 2019. https://doi.org/10.14483/23448393.12512 ↑4

[14] A. Q. Al-Faris *et al.*, "Breast MRI tumour segmentation using modified automatic seeded region growing based on particle swarm optimization image clustering," in *Proc. 17th Online World Conf. Soft Computing Ind. App.*, 2014, pp. 49-60. https://doi.org/10.1007/978-3-319-00930-8_5 ↑5

[15] H. Shao *et al.*, "A saliency model for automated tumor detection in breast ultrasound images," in *2015 IEEE Int. Conf. Image Proc. (ICIP)*, 2015, pp. 1424-1428. https://doi.org/10.1109/ICIP.2015.7351035 ↑5

[16] H. Rajaguru and S. K. Prabhakar, "Expectation maximization based logistic regression for breast cancer classification," in *2017 Int. Conf. Elec. Comm. Aerospace Tech. (ICECA)*, 2017, pp. 603-606. https://doi.org/10.1109/ICECA.2017.8203608 ↑5, 17

[17] L. Khairunnahar *et al.*, "Classification of malignant and benign tissue with logistic regression," *Inform. Med. Unlocked*, vol. 16, art. 100189, 2019. https://doi.org/10.1016/j.imu.2019.100189 ↑5

[18] M. Dey *et al.*, "Automated breast lesion localisation in microwave imaging employing simplified pulse coupled neural network," *PloS One*, vol. 17, no. 7, art. e0271377, 2022. https://doi.org/10.1371/journal.pone.0271377 ↑5

[19] N. Saffari *et al.*, "Fully automated breast density segmentation and classification using deep learning," *Diagnostics*, vol. 10, no. 11, art. 988, 2020. https://doi.org/10.3390/diagnostics10110988 ↑6

[20] T. Reimer, J. Krenkevich, and S. Pistorius, "An open-access experimental dataset for breast microwave imaging," in 2020 14th Eur. Conf. Antennas Propag. (EuCAP), 2020, p. 1-5. https://doi.org/10.23919/EuCAP48036.2020.9135659 ↑6

[21] T. Reimer, M. Solis-Nepote, and S. Pistorius, "The application of an iterative structure to the delay-and-sum and the delay-multiply-and-sum beamformers in breast microwave imaging," *Diagnostics*, vol. 10, no. 6, art. 411, 2020. https://doi.org/10.3390/diagnostics10060411 ↑7

[22] B. Abdollahzadeh and F. S. Gharehchopogh, "A multi-objective optimization algorithm for feature selection problems," *Eng. Comp.*, vol. 38, no. Suppl. 3, pp. 1845–1863, 2022. https://doi.org/10.1007/s00366-021-01369-9 ↑13

[23] F. Ahmad *et al.*, "A GA-based feature selection and parameter optimization of an ANN in diagnosing breast cancer," *Patt. Analysis App.*, vol. 18, pp. 861–870, 2015. https://doi.org/10.1007/s10044-014-0375-9 ↑13

[24] W. Chen *et al.*, "Density-based logistic regression," in *Proc. 19th ACM SIGKDD Int. Conf. Knowledge Disc. Data Mining*, 2013, pp. 140-148. https://doi.org/10.1145/2487575.2487583 ↑13

[25] R. Rodrigues *et al.*, "A two-step segmentation method for breast ultrasound masses based on multi-resolution analysis," *Ultrasound Med. Biol.*, vol. 41, no. 6, pp. 1737–1748, 2015. https://doi.org/10.1016/j.ultrasmedbio.2015.01.012 ↑14

Azhar Albaaj

received his BSc degree in Electrical Engineering from Kufa University, Najaf, Iraq, in 2012, and his MSc degree in Communications Engineering from Razi Universityghmoradi@aut.ac.ir, Kermanshah, Iran, in 2016. He is a PhD student in Communication Systems at Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran. His research interests are microwave imaging, signal processing, and energy harvesting.
Email: azhar.albaaj@aut.ac.ir

Yaser Norouzi

received his BSc in Electronics and his MSc and PhD in Communication Systems from Sharif University of Technology, Tehran, Iran, in 2002, 2004, and 2008. Since 2010, he serves at the Department of Electrical Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran. His fields of research include signal processing, passive systems, and geolocation.
Email: y.norouzi@aut.ac.ir

Gholamreza Moradi

(senior member, IEEE) received the Ph.D. degree in electrical engineering from Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran in 2002. His main research interests are high frequency characterization of materials, numerical Electromagnetics, microwave imaging, planar microwave/mm wave and THz systems. Dr. Moradi has authored over 10 books in his specialty. He has published and presented almost 200 papers in the refereed journals and international conferences. He is currently a Professor with the Department of Electrical Engineering, Amirkabir University of Technology.

Email: ghmoradi@aut.ac.ir

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1221

DOI: https://doi.org/10.14483/23448393.20895

Research

Scheduling in a Simple Assembly Robotics Cell to Minimize Earliness and Tardiness

Programación de trabajos en una celda robótica de ensamble simple para minimizar los adelantos y las tardanzas

John Andrés Muñoz-Guevara[1], Jairo Alberto Villegas-Florez[1], and Jhannier Jhoan Jaramillo-Tabima[1]*

[1]Facultad de Ciencias Empresariales, Universidad Tecnológica de Pereira, Colombia

Abstract

Background: Robotics Assembly Cells (RAC) have been designed to meet the flexibility requirements demanded by today's globalized market. The objective is to manufacture a vast variety of products at a low cost, which requires equipment with a high level of flexibility, such as robots. The need to schedule a great variety of jobs in an RAC is a very relevant issue, as efficiency and productivity depend on the sequence in which jobs are scheduled. Studies around this matter have developed models with analytical and heuristic approaches, as well as simulation methods and genetic algorithms, seeking to improve performance measures based mainly on time, utilization, and costs.
Method: The purpose of this article is to formulate an exact mathematical model using mixed-integer linear programming (MILP) to optimize small scheduling problems. The objective is to minimize the measure of performance related to the tardiness and earliness of jobs. This optimization aims to mitigate the effects of delays in product deliveries, queue times, and work-in-process inventory in subsequent processes. Doing so facilitates adherence to agreed-upon delivery deadlines and prevents bottlenecks in the assembly cell.
Results: The proposed mathematical model generates optimal solutions to the job scheduling problem in the assembly cell, which serves as a case study. This addresses the need to minimize tardiness to meet delivery deadlines or minimize earliness while avoiding an increase in work-in-process inventories. The model ensures that optimal scheduling decisions are made to optimize both delivery performance and inventory levels.
Conclusions: Due to the NP-hard complexity of the scheduling problem under study, the proposed mathematical model demonstrates computational efficiency in solving scheduling problems with fewer than 20 jobs. The model is designed to handle such smaller-scale problems within a reasonable computational time frame, considering the inherent complexity of the scheduling problem.

Keywords: robotic assembly cell, tardiness, scheduling

Article history

Received: 24th/May/2023

Modified: 17th/Nov/2023

Accepted: 18th/Mar/2024

Ing, vol. 29, no. 2, 2024, e20895

* **Correspondence:** johandmunoz@utp.edu.co

Resumen

Contexto: Las celdas robóticas de ensamble RAC han sido diseñadas para cumplir con los requerimientos de flexibilidad que exige el mercado globalizado actual. El objetivo es fabricar una alta variedad de productos a bajos costos, por lo cual se requiere de equipos con un alto nivel de flexibilidad como los robots. La necesidad de programar una amplia variedad de trabajos en una RAC es un problema muy relevante, dado que la eficiencia y la productividad dependen de la secuencia en la cual se programan los trabajos. Los estudios alrededor de este asunto han desarrollado modelos con enfoques analíticos y heurísticos, así como métodos de simulación y algoritmos genéticos, que buscan mejorar medidas de desempeño basadas principalmente en el tiempo, el grado de utilización y el costo.
Método: El propósito de este artículo es formular un modelo matemático exacto mediante programación lineal entera mixta (MILP) para optimizar problemas pequeños de programación. El objetivo es minimizar las medidas de desempeño de tardanza y adelanto de los trabajos. Esta optimización busca mitigar los efectos de las demoras en las entregas de los productos, los tiempos de cola y los inventarios de producto en proceso en instancias posteriores. Esto permite cumplir con los plazos de entrega pactados y evita bloqueos en la celda de ensamble.
Resultados: El modelo matemático propuesto genera soluciones óptimas al problema de programación de trabajos en la celda de ensamble propuesta, lo que sirve como caso de estudio. Esto aborda la necesidad de minimizar las tardanzas para cumplir con los plazos de entrega o de minimizar los adelantos y evitar un aumento en los inventarios de producto en proceso. El modelo garantiza que se tomen decisiones óptimas de programación tanto para mejorar los tiempos de entrega como los niveles de inventario.
Conclusiones: Debido a la complejidad NP-difícil del problema de programación estudiado, el modelo matemático propuesto demuestra eficiencia computacional en la resolución de problemas de programación con menos de 20 trabajos. El modelo está diseñado para manejar tales problemas de menor escala dentro de un marco de tiempo computacional razonable, considerando la complejidad inherente del problema de programación.

Palabras clave: celdas de ensamble robóticas, tardanza, programación.

Table of contents

	Page
1. Introduction	3
2. Methodology	8
3. Development	9
4. Evaluation	11
5. Discussion	16
6. Conclusions	18
7. CRediT author statement	18
References	18

1. Introduction

The current market demands increasingly personalized products at mass production prices. This poses a significant challenge for industries worldwide. To remain competitive in a globalized market, industries have implemented flexible manufacturing systems (FMS). FMS allows industries to meet market demands and reduce manufacturing costs, ultimately leading to profitability. FMS offer several advantages, particularly in terms of flexibility, which refers to a system's ability to manufacture a wide range of products using the same equipment, with short setup times (1). FMS are generally classified into two main subgroups: flexible assembly systems (FAS) and flexible manufacturing systems (FMS) (2). Much of the research conducted in the field has predominantly focused on FMS, while FAS have received comparatively less attention. Both FAS and FMS are categorized as computer-integrated manufacturing systems, but they exhibit distinct differences in several aspects. Firstly, FAS are capable of processing a significantly greater number of diverse tasks compared to FMS. In FAS, multiple components and parts are joined simultaneously, whereas FMS typically operate on one part at a time. Secondly, the processing times required for each operation in FAS are shorter than in FMS. Consequently, the relationship between transfer and processing times in FAS are generally higher in comparison with that of FMS. Thirdly, the material handling systems of FAS tend to be more complex. These differences highlight the unique characteristics and operational variance observed between FAS and FMS, suggesting the need for dedicated research and analysis to understand and optimize the performance of both systems (3). Table I shows some differences between FAS and FMS, which make issues in the former more difficult (4).

Table I. Differences between FAS and FMS

Features	FAS	FMS
Number of different tasks that can be performed	High	Low
Number of pieces per job	Several	Only one
Processing time per part	Short (seconds)	Long (minutes/hours)
Setup time/processing time ratio	High	Low
Material handling	Complex (assembly operations)	Simple (loading/unloading)
Collaborative work man/machine	High	Low

FAS can be classified into two types: robotic assembly lines (RAL) and robotic assembly cells (RAC) (5). A RAL is a line flow system that comprises assembly stations interconnected in series through an automated material handling system. It is primarily employed for assembling products characterized by a high volume and a low variety. These products typically have stable designs and exhibit minimal fluctuations in demand requirements (6). The RAC is a highly advanced system that incorporates robots, assembly stations, and an automated material handling system, all under computerized control. RACs are designed to efficiently assemble a wide range of products in small batches. One of the key advantages of an RAC is its ability to utilize one or more robots that can perform multiple operations simultaneously or collaboratively. This enables greater flexibility and productivity within the cell (7). Fig. 1 shows the layout of a RAL and a RAC.

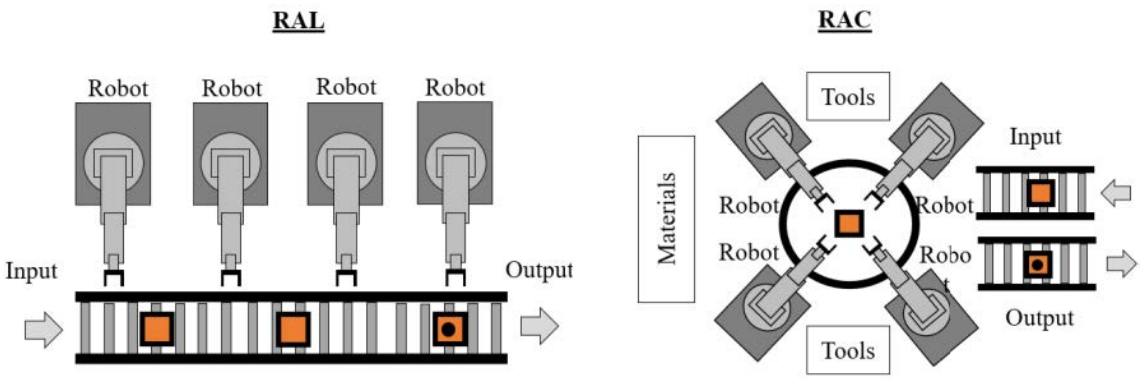


Figure 1. Layout of a RAL and a RAC

Scheduling is a decision-making process that plays a vital role in most manufacturing industries. Its function is to optimize the allocation of limited resources for job processing. These resources include machines, robots, tools, material handling equipment, and materials. A job consists of a series of operations or tasks to be performed by manufacturing systems (8). Few studies have addressed the issue of task scheduling in RACs. These studies can be classified into three approaches: the analytical and heuristic approach, the simulation approach, and genetic algorithms. In addition, studies have used different types of performance measures to evaluate the results of scheduling. Measures based on time, utilization, and costs are the most widely used, and time-based performance measures that has attracted the most attention (9). Most of real-life scheduling issues are difficult to solve with a non-deterministic polynomial time NP-hard model and tend to require a large amount of constraints to generate a reliable solution (10).

Scheduling issues in manufacturing systems are specified by a set of elements. The critical elements are decision variables, constraints, and objective functions. The goal of any manufacturing company is to maximize the utilization of resources, minimize completion times, and meet the due dates of orders (11). In field of scheduling, several time-based objective functions are used to evaluate the performance of a system under different scheduling strategies, such as makespan, flow time, tardiness, idle time, queue time, waiting times, and setup times, among others (12). One fundamental aspect when it comes to managing an RAC is to meet the product's due date, either to deliver the final product to the customer on time or to deliver the semi-finished product to the next process on time, thereby not delaying the system. This job becomes complex when, in a RAC, multiple products with different due dates must be assembled. Here, the problem lies in being able to determine the processing sequence of the tasks that help minimize potential tardiness. Tardiness is defined as the amount of time by which a job exceeds its due date. In scheduling theory, three measures of tardiness are estimated: total tardiness ($\sum T_j$), the sum of the tardiness of all late jobs; maximum tardiness (T_{max}), the value of the job that obtained the longest tardiness; and the number of tardiness jobs (NT) the amount of jobs that were delivered late (13). The other criterion is earliness, defined as the amount of time by which a job is ahead of its due date. This is a very important and frequent industrial problem that is common to most just-in-time (JIT) production environments. JIT consists of delivering products and

services at the right time for immediate use, where the main objective is the continuous search for improving the production process. This is achieved and developed through reduced inventories. JIT scheduling problems are very common in the industry. In a JIT scheduling environment, a job should be finished as close to the due date as possible. An early job completion results in inventory carrying costs, such as those related to storage and insurance (14). Balancing both tardiness and earliness is essential in achieving overall optimization in a RAC. The scheduling algorithm aims to strike a balance between completing jobs within their due dates to avoid tardiness and minimizing excessive earliness to optimize resource utilization. By considering both tardiness and earliness as criteria, manufacturers can effectively manage the assembly process in RACs. Striking the right balance ensures that jobs are completed within the desired time frames, optimizing production efficiency and meeting customer demands without unnecessary delays or excess inventory. Scheduling issues are theoretically classified according to various criteria, such as production volume, the nature of production, production capacity, and manufacturing systems. Each type of scheduling issue has different levels (15, 16).

It has been previously mentioned that one of the challenges in scheduling tasks for a RAC is meeting the due dates for an assigned job. In the scheduling area, the tardiness T_j is calculated as the maximum value of the subtraction between the completion time of the job C_j minus the due date D_j (8). This is calculated via Eq. 1:

$$T_j = max(C_j - D_j; 0) \tag{1}$$

If job j is finished before the due date, the tardiness is 0, since none was generated. Tardiness does no´t take negative values; if that were to happen, we would be dealing with earliness. On the other hand, the eardiness E_j is calculated as the maximum value of the subtraction between the due date D_j minus the completion time of the job C_j (8). This is calculated via Eq. 2:

$$E_j = max(D_j - C_j; 0) \tag{2}$$

It is worth remembering that three measures of earliness are estimated: the total earliness ($\sum E_j$), which is the sum of the earliness of all early jobs; the maximum earliness (E_{max}), which is the value of the job that obtained the maximum earliness; and the number of earliness jobs (NE), which is the amount of jobs that were delivered early.

Scheduling issues for evaluating tardiness and earliness have been solved using heuristic algorithms, exact models, and expert systems (13). The most significant theoretical developments based on heuristic algorithms for the $1||T_j$ problem are the Emmons dominance conditions, known as Emmons theorems, which determine precedence relationships between jobs to generate an optimal sequence. The first Emmons theorem provides the necessary conditions for a shorter job to precede a longer one in an optimal sequence, while the second theorem provides the necessary conditions for a longer job to precede a shorter one in an optimal sequence. These relationships have been extensively used to reduce the solution space in enumeration methods (17). Lawler's decomposition theorem is the second most significant algorithm. In order to explain the decomposition theorem, let us suppose that the jobs are numbered in the order of the earliest due date or (EDD). According to Lawler (18), the $1||T_j$ problem is decomposed with the longest j job in a k position. This author makes the following assumption: the longest j job is completed as tardy as possible in an optimal sequence, which means moving the jobs

from left to right while meeting some conditions. The development of these algorithms gave way to several studies, where dispatch rules were defined based on the execution and delivery times of jobs. The most used dispatch rules to solve the aforementioned problem are:

- *First come, first served (FCFS): jobs that arrive first are scheduled first.*
- *Earliest Due Date First (EDD): jobs with the earliest due date are scheduled first.*
- *Shortest Processing Time (SPT): jobs with the shortest processing time are scheduled first.*
- *Longest Processing Time (LPT): jobs with the longest processing time are scheduled first.*
- *Slack Time Remaining First (STR): jobs with the minimum slack time remaining are scheduled first. As shown in Eq. 3:*

$$STR = D_j - P_j \tag{3}$$

where:
D_j: due date of job j
P_j: processing time of job j

There are many examples where the penalty for a tardiness job remains the same, no matter how long the tardiness is. In this case, the objective is to minimize the number of tardy jobs NT. Moore's algorithm (1968) minimizes the number of tardy jobs for the problem of a single machine, where all jobs have a same release time. This algorithm consists of four very simple steps. In the first step, the jobs are sequenced while following the EDD rule. Then, the tardiness jobs are ordered according to the LPT rule until an optimal solution is reached (19). Most of the exact algorithms for the $1||T_j$ problem use the dynamic programming (DP), the branch and bound method (B&B), or a hybrid DP/B&B approach. The most efficient of these algorithms is the Potts and Van Wassenhove (P-W) algorithm, capable of solving problems with up to 100 jobs. The P-W algorithm decomposes the problem until the generated subproblems are small enough to be solved via DP, and it does not use any lower bound (20). The best exact algorithms to solve $1||T_j$ are the B&B, one that use the latest developments of the Emmons and Lawler decomposition theorems. The performance of these algorithms can be improved by using induced due dates, calculated after certain jobs have been shown to precede/follow others. Szwarc's branching algorithm (2001) is an example of such algorithm, and it can handle problems with up to 500 jobs. In addition, in the last decade, several algorithms based on metaheuristics and artificial intelligence have been developed to solve tardiness minimization problems, considering more variables and constraints for tasks and machines. These algorithms represent a complex computational development (13).

In the literature, earliness and tardiness penalties are studied by various authors from a single-objective point of view. Most works consider distinct or common due dates.

(21) studied the problem considering distinct due dates. They presented an optimal algorithm with polynomial complexity to determine the optimal completion time for each job in a schedule determined by a genetic algorithm (GA). This optimal algorithm was used because it may be interesting to anticipate a job, even paying a penalty if it is shorter than that generated by the tardiness.

(22) presented a hybrid genetic-bees algorithm (GBA)-optimized solution for minimizing the weighted sum of earliness and tardiness penalties in the context of the single machine scheduling problem (SMSP). The bee algorithm (BA) was improved by incorporating GA operators during the global search stage, aiming to progressively enhance the BA's global search capability with new additions.

(23) addressed the SMSP while aiming to formulate schedules for tasks with arbitrary due dates and minimizing the total earliness/tardiness with respect to said due dates. The problem was approached through three distinct formulations: 1) where the start time of the machine was fixed, 2) where the start time belonged to a specified time segment, and 3) where the start time was arbitrary. Given the inherent complexity, there is no exact polynomial algorithm for its solution. For the first two formulations, the PSC-algorithms were presented. Each algorithm incorporates clear indicators of feasible solution optimality and is grounded upon optimal solutions for the single machine problem, specifically designed to minimize the total tardiness of tasks relative to their varied due dates with equal weights.

(24) addressed the SMSP to minimize the total weighted earliness and tardiness relative to a nonrestrictive common due date. This problem is fundamental and has applications in JIT manufacturing. It is connected to a boolean programming problem with a quadratic objective function, referred to as the half-product. This research presented an approach to create a fast fully polynomial-time approximation scheme (FPTAS) for this problem, and the running time was aligned with the best-known running time for an FPTAS aimed at minimizing a half-product with no additive constant.

(25) presented a mixed-integer programming model for a SMSP whose objective was to minimize the total earliness/tardiness duration when the uncertainty of parameters such as processing times and due date was coded with grey numbers. Grey theory and grey numbers serve as tools for representing the uncertainty associated with parameters such as processing times and common due dates. The authors proposed a 0-1 mathematical model to address the problem, coupled with an efficient heuristic method that leveraged the expected processing times for job ordering.

(26) studied the SMSP while considering past-sequence-dependent setup times. This study explored scenarios involving common due date assignments, slack due dated, and different due date assignments. The objective function aimed to minimize the linear weighted sum of earliness-tardiness, the count of early and delayed jobs, and due date costs. The optimal properties of the problem were outlined, and it was proven that the problem is solvable in polynomial time. Additionally, three extensions to the problem were proposed: incorporating assumptions of position-dependent, time-dependent, and position-and-time-dependent processing times.

The objective of this article is to develop a MILP model that aids in scheduling and sequencing jobs within a RAC. The aim is to minimize both tardiness and earliness penalties, considering the presence of various job release times. By formulating this new MILP model, this article seeks to generate

an optimal solution that surpasses the quality of the results obtained through heuristic approaches and alternative algorithms like the B&B method. This approach offers the advantage of reduced computational complexity. The inclusion of different job release times adds a level of complexity to the problem, which the MILP model aims to effectively address.

2. Methodology

This study was conducted at the flexible manufacturing laboratory of Facultad de Ciencias Empresariales, Universidad Tecnológica de Pereira. Here there is a RAC called simple assembly robotized cell (SRAC) because it only has one robot. It was designed with the ability to assemble the same family of products, so the robot does not require a gripper change. There are no transfer or setup times, nor trajectory collisions analyzed since the cell is made up of a single robot. Thus, the number of variables is reduced, making it possible to simplify the mathematical model to perform the first tests.

The SRAC is composed of one robot, which takes the components to be assembled from two deposits called Parts 1 and Parts 2. The combination of different parts allows assembling up to 12 different products without the need for reconfiguring of the robot. The task of joining different parts is carried out in the assembly station, and then, when the process complete, the product is positioned on the pallet to be transported to the next stages of the process. SRAC is the first stage of a three-stage process to complete a finished product. Fig. 2 shows a description of the SRAC.

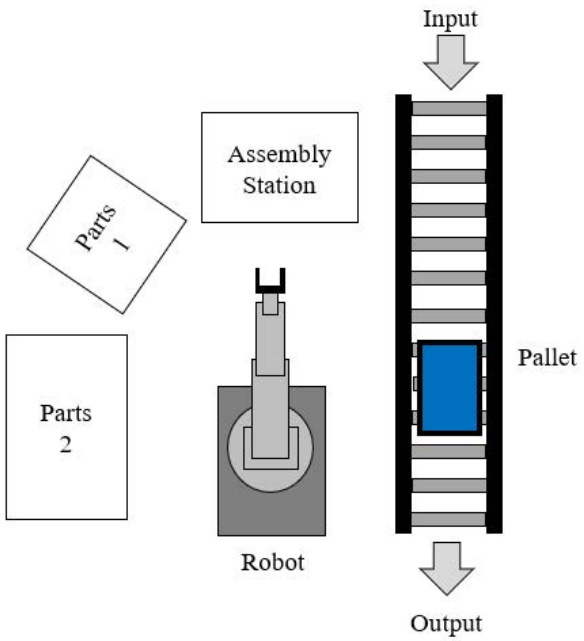


Figure 2. Simple assembly robotic cell (SRAC)

For the first test, the scheduling of 6 jobs was taken as a reference. One job involved a standardized lot of the same product the size of each lot was previously determined to minimize transfer times between jobs. Thus, each time a product needed assembly, all the lot was assembled. Since the SRAC is the first stage in the process, and assuming that the Parts 1 and Parts 2 deposits are always loaded with the required parts, the assembly process does not start until the part from the raw material warehouse arrives at the workstation, which causes job release times to vary. The due dates D_j for the jobs to be assembled at the SRAC were determined through an MRP program that determines due dates according to production orders generated by customer orders. Table **II** shows the requirements of the jobs to be scheduled in the first run:

Table II. SRAC requirements

Cycle time [minutes/unit]	**Jobs**					
	1	2	3	4	5	6
Cycle time [minutes/unit]	10	6	8	8	5	3
Lot size [units]	10	5	5	10	20	20
Processing time P_j [minutes]	100	30	40	80	100	60
Due date D_j [minutes]	150	60	95	230	200	300

Once the information on production requirements was obtained, job scheduling was carried out while following the four dispatch rules described above (FCFS, EDD, SPT, LPT and STR). For the sequencing of the 6 assembly jobs to be processed by the SRAC, there were a total of 6! possible combinations, *i.e.*, 720 different sequencing possibilities. The goal was to find the sequence of jobs that minimizes both earliness and tardiness. An exact mathematical model was formulated, capable of solving the job sequencing problem, with the objective of minimizing the earliness (α) and tardiness (β) penalties. The results obtained with the model were evaluated, validated, and compared against the performance obtained through five dispatch rules.

3. Development

To formulate the MILP problem, the value of the α and β penalties was taken, which depends on the problem to be addressed. If minimizing the work-in-process (WIP) is required after assembly at the workstation, the value of α must be greater than β. On the contrary, if one wishes to minimize job delays to shorten the makespan of the process, then β must be greater than α. To generate the MILP model, the following variables were defined:

j : subscript that identifies the job to be scheduled

i : subscript that identifies the position in the schedule

α_j : earliness penalties for job j

β_j : tardiness penalties for job j

r_j : release time for job j

X_{ij} : binary variable, 1 if job j is performed at position i, 0 otherwise

P_j : processing time for job j

C_i : completion time for job scheduled at position i

D_j : due date for job j

T_j : tardiness for job j, no negativity

E_j : earliness for job j, no negativity

Y_j : binary variable, 1 if job j has tardiness, 0 else

Z_j : binary variable, 1 if job j has earliness, 0 else

Once the job schedules had been obtained by applying heuristic dispatch rules, the following question arose: Is it possible to develop a mathematical model which allows finding a sequence of jobs that minimize the total for earliness and tardiness penalties? The model proposed to answer this question is presented below.

Objective function:

$$\text{Min} = \sum_{j=1}^{N} \alpha_j \cdot E_j + \sum_{j=1}^{N} \beta_j \cdot T_j \tag{4}$$

subject to:

$$\sum_{j=1}^{N} (P_j \cdot X_{ji}) + C_{i-1} \leq C_i \quad \text{for all } j \tag{5}$$

$$\sum_{i=1}^{N} (C_i \cdot X_{ji}) - D_j = T_j \cdot Y_j - E_j \cdot Z_j \quad \text{for all } j \tag{6}$$

$$\sum_{i=1}^{N} (C_i \cdot X_{ji}) \geq r_j + P_j \quad \text{for all } j \tag{7}$$

$$\sum_{i=1}^{N} X_{ji} = 1 \quad \text{for all } j \tag{8}$$

$$\sum_{j=1}^{N} X_{ji} = 1 \quad \text{for all } i \tag{9}$$

$$X_{ji}; Y_j; Z_j : Binary;$$

$$T_j \geq 0; E_j \geq 0;$$

Eq. 4 shows the definition of the objective function. Eq. 5 calculates the completion time of the scheduled job at position i. Eq. 6 calculates the earliness or tardiness of job j. If the earliness is greater than zero, the binary variable Z_j takes a value of 1, and, if the tardiness is greater than 0, the binary variable Y_j takes a value of 1. Eq. 7 establishes that the completion time of job j is greater than the release time plus the processing time when there are different job release times. Eq. 8 determines that a job j can only be scheduled in a single position i. Eq. 10 states that, in a position I, only one job j can be scheduled.

In some cases, the penalty for having an early or tardy job is independent of the time of earliness or tardiness; only having an early or tardy job is penalized. In this case, the objective function is formulated as follows:

Objective function:

$$\text{Min} = \sum_{j=1}^{N} \alpha_j \cdot Z_j + \sum_{j=1}^{N} \beta_j \cdot Y_j \tag{10}$$

4. Evaluation

To evaluate the proposed MILP model, the performance of the earliness and tardiness penalties was evaluated in relation to the dispatch rules (FCFS, EDD, SPT, LPT, and STR). This section lists, the most important assumptions made in this study:

- The workstation can perform only one job at a time.
- Process pre-emption is not allowed.
- The required sequence of machines and the processing times of the jobs are known.
- There is no restriction regarding queue length at the workstation.
- There are no interruptions at the workstation and no machine breakdowns.
- There are no setup times for jobs.
- There are no alternate routings.
- There are no limiting resources other than the workstation.

To evaluate the proposed model, we executed one hundred instances, divided into three stages. Here is a summary of the stages and the parameters used.

- **Stage one.** Constant penalty values were assumed to analyze the system's performance when the penalty for tardiness is greater than that for earliness and vice versa. The penalty values used were α = 5 and β = 10. In the second instance, the penalty values used were α = 10 and β = 5.
- **Stage two.** Different penalty values for each job, denoted as α_j and β_j, were evaluated. These penalty values were generated by drawing data from a uniform distribution with a range of [5-20]. This allowed for a more varied and dynamic evaluation of the system.
- **Stage three:** The objective function was evaluated to minimize the number of tardy and early jobs, considering the penalty values α_j and β_j generated by the same uniform distribution system in stage two.

The instances of this stage were divided into five groups based on the number of jobs: 6, 10, 15, 18, and 20. For each stage, several parameters were considered to create the instances:

- **Job arrivals.** The arrivals of jobs were generated using an exponential distribution.
- **Processing times.** The processing times for each job were drawn from a rectangular distribution within the range of [20-100].

- **Due date.** The due date of each job was determined via an assignment factor (k), which was randomly selected from a uniform distribution within the range of [2-10]. The due date calculation equation was used to determine the specific due date for each job, as shown below.

$$D_j = (r_j + P_j)k \tag{11}$$

By varying the penalty values, the number of jobs and the distribution of job arrivals, process times, and due dates, this study aimed to comprehensively evaluate the proposed model's performance under different scenarios and conditions. The results obtained for the first 50 instances with constant values of α and β are shown in Table III.

Table III. Evaluation with α = 5; β = 10 and α = 10; β = 5

α = 5; β = 10								α = 10; β = 5							
Jobs	**Run**	**FCFS**	**EDD**	**SPT**	**LPT**	**STR**	**MILP**	**Jobs**	**Run**	**FCFS**	**EDD**	**SPT**	**LPT**	**STR**	**MILP**
	1	3 075	1 500	2 125	6 100	2 325	1 400		1	3 075	1 500	2 225	3 950	1 725	1 150
	2	3 500	3 750	3 475	4 950	4 750	2 950		2	2 650	2 850	2 975	2 925	2 825	2 150
6	3	2275	2250	1 625	6 300	1 525	1 525	6	3	2 150	2 025	1 900	4 125	1 700	1 300
	4	3 425	1 825	2 750	5 050	1 625	1 625		4	3 325	1 775	3 100	3 650	1 375	1 175
	5	4 375	2 775	3 575	5 500	3 425	2 575		5	3500	2 100	3 100	3 500	2 125	1 825
	6	14 000	7 350	7 250	16 880	6 395	6 095		6	8 650	3 900	4 300	9 880	3 445	3 295
	7	10 125	10 450	9 250	18 400	9 475	7 500		7	5 475	5 600	5 000	10 550	5 225	4 375
10	8	5 220	3 070	4 620	19 145	3 290	3 070	10	8	6 425	4 800	6 075	9 150	5 125	3 675
	9	6 525	3 725	4 050	12 550	3 525	3 225		9	5 175	3 350	4 650	9 425	3 350	3 300
	10	10 325	5 300	5 825	16 750	5 500	5 100		10	6 325	3 250	3 925	9 525	3 350	3 100
	11	16 100	5 675	12 050	16 950	8 500	5 150		11	15 300	5 825	7 900	16 700	6 700	5 525
	12	19 775	10 675	12 050	26 950	12 500	6 275		12	18 550	5 050	10 200	21 850	6 700	2 175
15	13	14 900	3 600	10 975	24 600	4 050	3 300	15	13	14 725	3 675	12 500	18 075	3 525	2 875
	14	16 500	5 050	10 725	24 200	5 750	2 875		14	12 200	6 075	10 250	14 500	8 050	3 025
	15	12 775	8 050	10 500	18 750	6 525	3 225		15	16 575	5 025	14 225	16 575	6 050	3 525
	16	16 375	4 225	9 125	18 275	4 225	2 350		16	17 400	8 450	14 350	18 650	8 450	4 375
	17	15 950	6 650	11 750	24 475	8 150	4 750		17	16 450	6 700	13 600	20 625	5 650	3 350
18	18	18 450	8 050	12 525	20 750	7 525	5 225	18	18	18 550	7 050	12 775	20 250	8 220	3 575
	19	19 550	6 325	14 075	22 775	6 850	3 825		19	16 750	8 075	14 225	18 450	9 025	5 050
	20	16 450	6 775	15 850	22 150	6 785	4 325		20	18 025	7 520	12 250	20 550	8 050	3 775
	21	32 300	21 175	29 700	38 525	22 650	14125		21	42 850	18 770	28 075	52 620	22 650	6 725
	22	34 820	18 075	28 250	42 250	16 780	10 530		22	47 620	2875	32 425	50 850	26 530	8 860
20	23	36 875	20 450	26 750	41 880	22 720	12 375	20	23	40 885	22 455	34 050	48 355	22 860	10 525
	24	38 200	22 560	34 225	44 530	24 350	10 850		24	46 200	28 360	32 460	52 160	21 225	8 735
	25	36 350	22 050	28 450	42 225	21 335	12 725		25	48 750	24 550	32 285	54 730	25 925	10 220

In this first stage of evaluation, it can be observed that dispatch rules based on delivery deadlines, such as EDD and STR, perform best with responses close to the optimum obtained with the MILP model. These two dispatch rules aim to minimize tardiness, so their performance is better when the value of α is lower than the value of β. When α is greater than β, it can be noted that the gap between the two rules and the optimum value increases. Figs. 3 and 4 show the comparative curves for the first stage of evaluation.

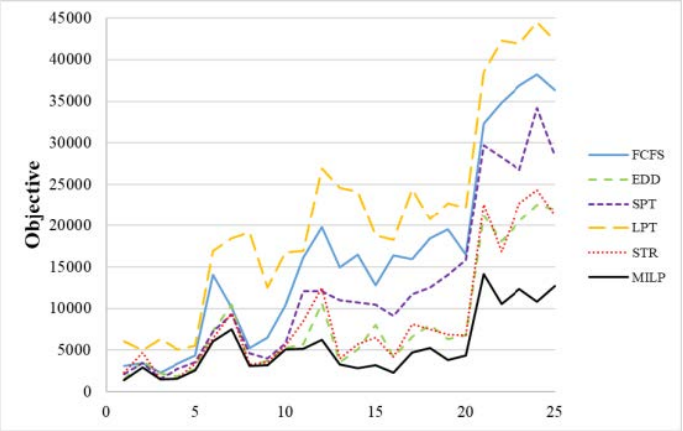


Figure 3. Performance with α = 5 and β = 10

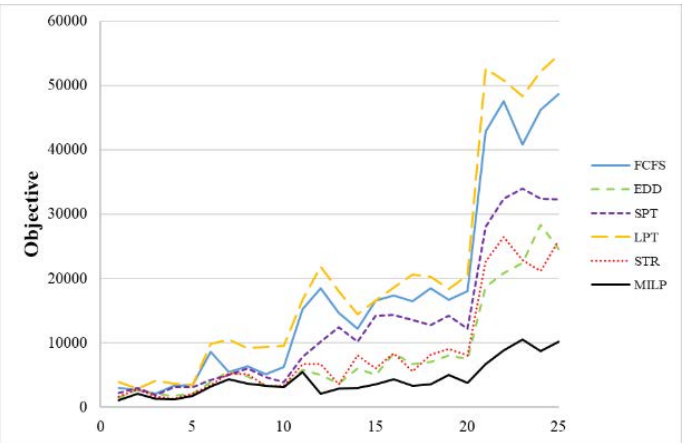


Figure 4. Performance with α = 10 and β = 5

For the second stage of evaluation, random values of α and β were estimated for each job, as described above. The results obtained are shown in Table **IV**.

The evaluation shows that when the penalty values α and β are variable and dependent on each job j, the gap between the optimal performance and the analyzed dispatch rules is much larger compared to when constant penalty values α and β were evaluated. This is because the model is able to prioritize jobs based on their penalty values. For example, a job with a high earliness penalty can be scheduled with some tardiness, or vice versa. Furthermore, the model is capable of scheduling jobs in such a way that both earliness and tardiness are equal to zero. This can be achieved by delaying the start of the job, which generates idle time in the assembly cell and allows minimizing the objective function. Fig. **5** provides an example of the described scenario. This analysis demonstrates the flexibility and effectiveness of the MILP model in handling variable penalty values and optimizing the scheduling of jobs to minimize both earliness and tardiness.

Table IV. Evaluation with α_j and β_j

α_j; β_j							
Jobs	**Run**	**FCFS**	**EDD**	**SPT**	**LPT**	**STR**	**MILP**
6	1	2.335	2.885	3.245	4.535	3.340	1.855
	2	2.160	1.375	1.125	3.560	2.360	1.125
	3	5.275	1.575	1.750	5.775	1.775	1.525
	4	3.665	2.375	2.620	3.950	2.175	2.100
	5	3.025	2.825	2.850	6.350	2.825	1.775
10	6	10.590	4.135	3.930	15.825	4.230	2.410
	7	12.600	8.425	9.050	13.020	7.545	3.345
	8	11.375	8.080	11.400	23.825	8.080	4.675
	9	14.525	9.750	8.625	15.270	10.560	5.065
	10	10.285	6.230	6.890	11.475	6.230	3.225
15	11	25.355	7.075	16.260	39.275	6.375	2.825
	12	25.610	6.700	30.290	29.190	6.150	2.350
	13	40.845	16.460	19.345	39.825	20.365	6.440
	14	32.475	22.535	26.765	38.820	20.880	3.565
	15	18.200	12.500	19.400	35.600	13.500	1.825
18	16	38.350	19.225	24.625	44.750	18.125	5.175
	17	46.475	11.150	12.550	39.550	19.075	4.950
	18	38.600	14.350	17.750	45.925	14.150	2.275
	19	34.860	12.425	10.050	42.560	16.625	3.785
	20	32.245	16.600	22.875	38.630	17.775	4.025
20	21	52.765	24.765	32.665	60.225	26.860	6.865
	22	54.860	26.050	22.890	62.450	30.750	5.750
	23	58.930	22.175	28.330	66.750	22.815	4.270
	24	50.775	25.050	28.780	62.025	26.430	3.085
	25	58.450	20.885	20.250	64.340	22.315	6.315

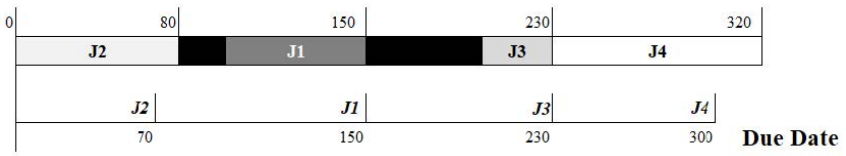


Figure 5. Schedule with MILP model

The SPT rule exhibits a similar performance to EDD and STR, despite not considering delivery deadlines. The performance curves depicted in Fig. 6 provide a visual comparison between the dispatch rules and the MILP model, highlighting their relative performance in terms of minimizing tardiness and earliness penalties. In this figure, it can be observed that the SPT rule achieves a relatively close

performance to EDD and STR. SPT prioritizes jobs with shorter processing times, which can indirectly contribute to minimizing earliness and tardiness. While the SPT rule may not explicitly optimize tardiness or earliness, its performance is still competitive, showcasing its effectiveness as a simple and efficient dispatch rule under certain scenarios.

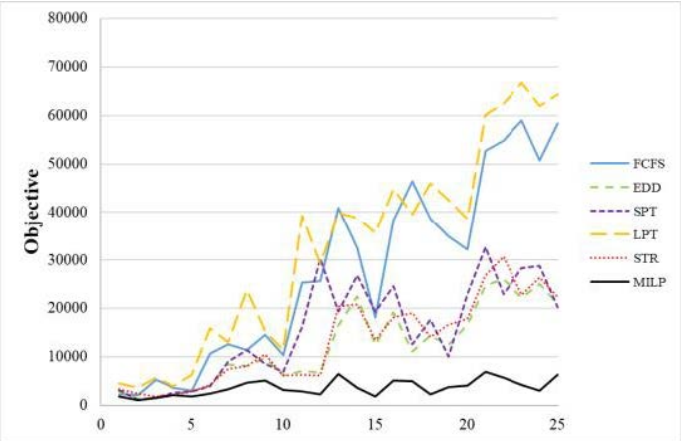


Figure 6. Performance with α_j and β_j

Finally, in the third stage, the MILP model was evaluated to minimize the penalty for getting a job early or late, regardless of the job's number of time units spent in tardiness or earliness. For this evaluation, the objective function described in Eq. 10 was taken as a reference. Table V shows the results of the assessment.

No significant differences can be observed between the dispatch rules used, and the gap with respect to the responses obtained by the MILP model increases as more jobs are scheduled. The evaluation results are presented in Fig. 7. It can be seen that, as the number of jobs increases, the performance of the dispatch rules becomes less competitive compared to the optimal solution obtained by the MILP model. This increasing gap highlights the computational superiority of the MILP model in handling larger scheduling problems and achieving more optimal results.

The MILP model demonstrated as significantly superior performance compared to the dispatch rules. However, it should be noted that the computation time of the model increases exponentially as the number of jobs increases. The model performs efficiently with a maximum of 20 jobs, but, beyond that, the computation time becomes impractical. Fig. 8 depicts the time curves obtained for the solution of the three types of analysis: constant α and β values, variable α_j and β_j values, and $\alpha_j Z_j$ and $\beta_j Y_j$ multiplied by the binary variables. The instances were run on a laptop with an Intel(R) Core (TM) i3-6100U CPU @ 2.30GHz, processor and 6 GB RAM. As expected, the computation time increases as the complexity of the analysis increases, especially when considering variable penalty values and the multiplication of binary variables.

These results highlight the computational limitations of the MILP model and emphasize the importance of considering scalability and computational resources when applying the model to larger

Table V. Evaluation with $\alpha_j Z_j$ and $\beta_j Y_j$

$\alpha_j Z_j$; $\beta_j Y_j$							
Jobs	**Run**	**FCFS**	**EDD**	**SPT**	**LPT**	**STR**	**MILP**
6	1	45	50	45	50	50	40
	2	40	50	75	40	50	35
	3	70	65	50	70	60	45
	4	55	45	65	50	45	40
	5	45	60	60	55	75	35
10	6	100	97	105	100	97	65
	7	110	95	80	95	90	65
	8	100	100	85	100	100	70
	9	75	75	80	100	100	40
	10	100	95	110	95	80	55
15	11	205	150	200	170	150	80
	12	120	110	120	125	110	70
	13	120	145	125	130	150	65
	14	125	120	130	135	120	75
	15	130	125	170	115	120	55
18	16	225	200	215	200	220	80
	17	220	235	250	230	240	95
	18	205	215	185	245	220	65
	19	210	180	215	195	200	70
	20	235	220	210	240	210	105
20	21	285	315	275	330	305	125
	22	310	250	330	265	285	115
	23	260	275	235	320	230	95
	24	290	230	240	285	195	110
	25	335	285	310	275	245	130

scheduling problems. It is crucial to assess the trade-off between computation time and the size of the scheduling problem, in order to ensure a practical implementation. Overall, the MILP model provides excellent performance for small to moderate scheduling problems, delivering optimal solutions within a reasonable computational time frame.

5. Discussion

The formulated MILP model exhibits limitations regarding the computation time necessary for solving problems with more than 20 jobs. Nonetheless, the model's application can be instrumental in evaluating the performance of metaheuristic solution techniques. This approach allows determining the gap between the optimal solution obtained with the MILP formulation and that achieved by the proposed algorithm. Consequently, this enables the calibration and scaling of the proposed algorithms to accommodate a larger number of jobs, with an initial comparison conducted on a smaller scale.

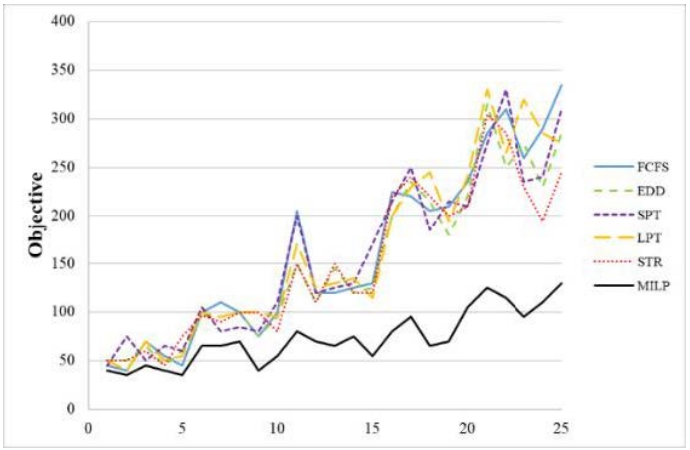


Figure 7. Performance with $\alpha_j Z_j$ and $\beta_j Y_j$

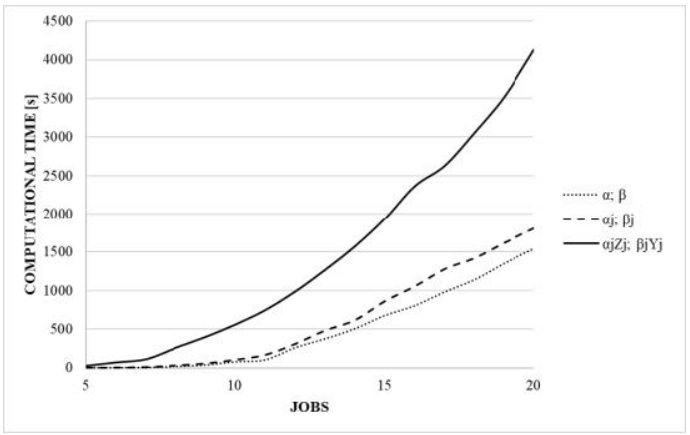


Figure 8. Computation time to solve MILP model vs. the number of jobs

Following the model evaluation, it became apparent that Eq. 12 significantly contributes to the model's complexity, thereby increasing computation times. To address this issue, linearizing the calculation is proposed, introducing the variable S_i, which determines the start time of the job at position i. This variable is then incorporated into Eq. 13 to calculate the variable C_i. This modification results in an approximately 20 % reduction in computation times. While this adjustment improves computational efficiency, its impact is not highly significant, and the formulation remains efficient for up to 20 jobs because, when the number of jobs grows, the number of equations also grows.

$$\sum_{j=1}^{N}(P_j \cdot X_{ji}) + S_{i-1} \leq S_i \quad \text{for all } j \tag{12}$$

$$C_i \geq S_i + B(X_{ji} - 1) \quad \text{for all } j, \text{ for all } i \tag{13}$$

The MILP formulation is very efficient when jobs with variable penalty rates are to be scheduled, which generates a program where the cell incurs a waiting time, so that the jobs do not start before those

required to minimize earliness. Undoubtedly, this characteristic considerably improves performance when evaluating the earliness/tardiness of jobs. However, these waiting times can be considered idle and penalize the cell utilization indicator. This opens the possibility of formulating a bio-objective MILP model, where, in addition to minimizing the earliness/tardiness, cell utilization can be maximized, or the makespan can be minimized.

6. Conclusions

The evaluation results emphasize the importance of utilizing advanced mathematical models (*e.g.*, MILP) for optimizing job scheduling in complex systems. While the dispatch rules may perform well for smaller job sets, their performance deteriorates as the complexity and scale of the problem increase.

The model was developed for an assembly cell with a single robot. We recommend adapting the mathematical model to schedule jobs in assembly cells with more than one robot, analyzing the assignment of tasks to robots, transfer times, and flow shop or parallel systems.

The model is efficient when a reduced number of jobs must be scheduled; if there are more jobs, more computational time will be required to solve the model at an exponential rate.

7. CRediT author statement

John Andrés Muñoz Guevara: conceptualization, investigation, methodology, software, validation, writing (original draft), writing (review & editing).
Jairo Alberto Villegas Flórez: methodology, software, validation, writing (review & editing).
Jhannier Jhoan Jaramillo Tabima: methodology, validation, writing (review & editing).

References

[1] M. P. Groover, "Flexible manufacturing systems," in *Automation, Production Systems, and Computer-Integrated Manufacturing*, Pallavaram, Chennai, India: Pearson Education India, 2016, vol. 4, ch. 16, pp. 460-462. ↑3

[2] J. Browne, "Classification of flexible manufacturing systems," *The FMS Mag.*, vol. 2, no. 2, pp. 114- 117, 1984. [Online]. Available: https://www.academia.edu/502998/Classification_of_flexible_manufacturing_systems?auto=citations&from=cover_page ↑3

[3] H. Gultekin, "Scheduling in flexible robotics manufacturing cells," PhD thesis, Dept. Ind. Eng., Bilkent Universitesi, 2006. ↑3

[4] K. K. Abd, "Development of an intelligent methodology for scheduling RFAC," in *Intelligent Scheduling of Robotic Flexible Assembly Cells*, Berlin, Germany: Springer, 2016, pp. 31-47. https://doi.org/10.1007/978-3-319-26296-3_3 ↑3

[5] T. Sawik, "Simultaneous loading, routing, and assembly plan selection in a flexible assembly system," *Mathematical Comp. Model.*, vol. 28, no. 9, pp. 19-29, 1998. https://doi.org/10.1016/S0895-7177(98)00142-3 ↑3

[6] G. Levitin, J. Rubinovitz, and B. Shnits, "A genetic algorithm for robotic assembly line balancing," *Eur. J. Operational Res.*, vol. 168, no. 3, pp. 811-825, 2006. https://doi.org/10.1016/j.ejor.2004.07.030 ↑3

[7] K. Abd, K. Abhary, and R. Marian, "Intelligent modeling of scheduling robotic flexible assembly cells using fuzzy logic," 2012. [Online]. Available: https://www.researchgate.net/publication/262398812 ↑3

[8] M. Pinedo, "Advanced single machine models," in *Scheduling; Theory, Algorithms, and Systems*, 4th ed., Berlin, Germany: Springer, 2012, ch. 3, pp. 70-77. ↑4, 5

[9] K. K. Abd, *Intelligent scheduling of robotic flexible assembly cells*. Berlin, Germany, Springer, 2015. ↑4

[10] A. Jain and H. Elmaraghy, "Production scheduling/rescheduling in flexible manufacturing," *Int. J. Prod. Res.*, vol. 35, no. 1, pp. 281-309, 1997. https://doi.org/10.1080/002075497196082 ↑4

[11] J. Blazewicz, "Scheduling computer and manufacturing processes," *J. Operational Res. Soc.*, vol. 48, no. 6, art. 659, 1997. https://doi.org/10.1057/palgrave.jors.2600793 ↑4

[12] R. Ramasesh, "Dynamic job shop scheduling: A survey of simulation research," *Omega*, vol. 18, no. 1, pp. 43-57, 1990. https://doi.org/10.1016/0305-0483(90)90017-4 ↑4

[13] C. Koulamas, "The single-machine total tardiness scheduling problem: Review and extensions," *Eur. J. Operational Res.*, vol. 202, no. 1, pp. 1-7, 2010. https://doi.org/10.1016/j.ejor.2009.04.007 ↑4, 5, 6

[14] J. E. C. Arroyo, R. dos Santos Ottoni, and A. de Paiva Oliveira, "Multi-objective variable neighborhood search algorithms for a single machine scheduling problem with distinct due windows," *Electron. Notes Theor. Comput. Sci.*, vol. 281, pp. 5-19, 2011. ↑5

[15] S. French, *Sequencing and scheduling. An Introduction to the Mathematics of the Job-Shop*. Hemel Hempstead, UK: Ellis Horwood, 1982. ↑5

[16] R. L. Graham, E. L. Lawler, J. K. Lenstra, and A. R. Kan, "Optimization and approximation in deterministic sequencing and scheduling: A survey," *Ann. Discrete Math.*, vol. 5, pp. 287-326, 1979. https://doi.org/10.1016/S0167-5060(08)70356-X ↑5

[17] H. Emmons, "One-machine sequencing to minimize certain functions of job tardiness," *Oper. Res.*, vol. 17, no. 4, pp. 701-715, 1969. https://doi.org/10.1287/opre.17.4.701 ↑5

[18] E. L. Lawler, "A "pseudopolynomial" algorithm for sequencing jobs to minimize total tardiness," *Ann. Disc. Math.*, vol. 1, pp. 331-342, 1977. https://doi.org/10.1016/S0167-5060(08)70742-8 ↑5

[19] S. K. Gupta and J. Kyparisis, "Single machine scheduling research," *Omega*, vol. 15, no. 3, pp. 207-227, 1987. https://doi.org/10.1016/0305-0483(87)90071-5 ↑6

[20] C. N. Potts and L. N. Van Wassenhove, "A decomposition algorithm for the single machine total tardiness problem," *Oper. Res. Lett.*, vol. 1, no. 5, pp. 177-181, 1982. https://doi.org/10.1016/0167-6377(82)90035-9 ↑6

[21] C.-Y. Lee and J. Y. Choi, "A genetic algorithm for job sequencing problems with distinct due dates and general early-tardy penalty weights," *Comp. Oper. Res.*, vol. 22, no. 8, pp. 857-869, 1995. https://doi.org/10.1016/0305-0548(94)00073-H ↑6

[22] B. Yuce *et al.*, "Hybrid genetic bees algorithm applied to single machine scheduling with earliness and tardiness penalties," *Comp. Ind. Eng.*, vol. 113, pp. 842-858, 2017. https://doi.org/10.1016/j.cie.2017.07.018 ↑7

[23] J. Rocholl and L. Mönch, "Hybrid algorithms for the earliness–tardiness single-machine multiple orders per job scheduling problem with a common due date," *RAIRO-Oper. Res.*, vol. 52, no. 4-5, pp. 1329-1350, 2018. https://doi.org/10.1051/ro/2018029 ↑7

[24] H. Kellerer, K. Rustogi, and V. A. Strusevich, "A fast FPTAS for single machine scheduling problem of minimizing total weighted earliness and tardiness about a large common due date," *Omega*, vol. 90, art. 101992, 2020. https://doi.org/10.1016/j.omega.2018.11.001 ↑7

[25] O. A. Arık, "Single machine earliness/tardiness scheduling problem with grey processing times and the grey common due date," *Grey Syst. Theory Appl.*, vol. 11, no. 1, pp. 95-109, 2021. https://doi.org/10.1108/GS-01-2020-0010 ↑7

[26] W. Wang, "Single-machine due-date assignment scheduling with generalized earliness-tardiness penalties including proportional setup times," *J. Appl. Math. Comput.*, vol. 68, no. 2, pp. 1013-1031, 2022. https://doi.org/10.1007/s12190-021-01555-4 ↑7

John Andrés Muñoz Guevara

PhD student of Engineering at Universidad Tecnológica de Pereira; Master in Automatic Production Systems, Universidad Tecnológica de Pereira; Industrial Engineering, Universidad Cooperativa de Colombia. Colombia. He is currently a professor and researcher at Facultad de Ciencias Empresariales, Universidad Tecnológica de Pereira.
Email: johandmunoz@utp.edu.co

Jairo Alberto Villegas Flórez

PhD in Engineering, Universidad Tecnológica de Pereira; Master in Operations Research and Statistics, Universidad Tecnológica de Pereira; Industrial Engineering and Electrical Technology, Universidad Tecnológica de Pereira. He is currently a professor and researcher in the Faculty of Engineering of Universidad Tecnológica de Pereira.
Email: javi@utp.edu.co

Jhannier Jhoan Jaramillo Tabima

Master student of Automatic Production Systems, Universidad Tecnológica de Pereira; Operations and Logistics Management Specialist, Universidad EAFIT de Pereira. Industrial Engineering, Universidad Tecnológica de Pereira. He is currently a professor at Facultad de Ciencias Empresariales, Universidad Tecnológica de Pereira.
Email: jjjaramillo@utp.edu.co

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1221

DOI: https://doi.org/10.14483/23448393.21583

Research

Explainable Artificial Intelligence as an Ethical Principle

Inteligencia artificial explicable como principio ético

Mario González-Arencibia[1], Hugo Ordoñez-Erazo[2], and Juan-Sebastián González-Sanabria[3]*

[1]Universidad de las Ciencias Informáticas, Habana, Cuba
[2]Universidad del Cauca
[3]Universidad Pedagógica y Tecnológica de Colombia

Abstract

Context: The advancement of artificial intelligence (AI) has brought numerous benefits in various fields. However, it also poses ethical challenges that must be addressed. One of these is the lack of explainability in AI systems, *i.e.*, the inability to understand how AI makes decisions or generates results. This raises questions about the transparency and accountability of these technologies. This lack of explainability hinders the understanding of how AI systems reach conclusions, which can lead to user distrust and affect the adoption of such technologies in critical sectors (*e.g.*, medicine or justice). In addition, there are ethical dilemmas regarding responsibility and bias in AI algorithms.
Method: Considering the above, there is a research gap related to studying the importance of explainable AI from an ethical point of view. The research question is *what is the ethical impact of the lack of explainability in AI systems and how can it be addressed?* The aim of this work is to understand the ethical implications of this issue and to propose methods for addressing it.
Results: Our findings reveal that the lack of explainability in AI systems can have negative consequences in terms of trust and accountability. Users can become frustrated by not understanding how a certain decision is made, potentially leading to mistrust of the technology. In addition, the lack of explainability makes it difficult to identify and correct biases in AI algorithms, which can perpetuate injustices and discrimination.
Conclusions: The main conclusion of this research is that AI must be ethically explainable in order to ensure transparency and accountability. It is necessary to develop tools and methodologies that allow understanding how AI systems work and how they make decisions. It is also important to foster multidisciplinary collaboration between experts in AI, ethics, and human rights to address this challenge comprehensively.

Keywords: artificial intelligence, AI, ethics, ethical principles, explainability, transparency

Article history

Received: 24th/Nov/2023
Modified: 17th/Jan/2024
Accepted: 9th/Apr/2024

Ing, vol. 29, no. 2, 2024, e21583

* **Correspondence:** juansebastian.gonzalez@uptc.edu.co

Resumen

Contexto: El avance de la inteligencia artificial (IA) ha traído numerosos beneficios en varios campos. Sin embargo, también plantea desafíos éticos que deben ser abordados. Uno de estos es la falta de explicabilidad en los sistemas de IA, *i.e.*, la incapacidad de entender cómo la IA toma decisiones o genera resultados. Esto plantea preguntas sobre la transparencia y la responsabilidad de estas tecnologías. Esta falta de explicabilidad limita la comprensión de la manera en que los sistemas de IA llegan a ciertas conclusiones, lo que puede llevar a la desconfianza de los usuarios y afectar la adopción de tales tecnologías en sectores críticos (*e.g.*, medicina o justicia). Además, existen dilemas éticos respecto a la responsabilidad y el sesgo en los algoritmos de IA.
Método: Considerando lo anterior, existe una brecha de investigación relacionada con estudiar la importancia de la IA explicable desde un punto de vista ético. La pregunta de investigación es *¿cuál es el impacto ético de la falta de explicabilidad en los sistemas de IA y cómo puede abordarse?* El objetivo de este trabajo es entender las implicaciones éticas de este problema y proponer métodos para abordarlo.
Resultados: Nuestros hallazgos revelan que la falta de explicabilidad en los sistemas de IA puede tener consecuencias negativas en términos de confianza y responsabilidad. Los usuarios pueden frustrarse por no entender cómo se toma una decisión determinada, lo que puede llevarlos a desconfiar de la tecnología. Además, la falta de explicabilidad dificulta la identificación y la corrección de sesgos en los algoritmos de IA, lo que puede perpetuar injusticias y discriminación.
Conclusiones: La principal conclusión de esta investigación es que la IA debe ser éticamente explicable para asegurar la transparencia y la responsabilidad. Es necesario desarrollar herramientas y metodologías que permitan entender cómo funcionan los sistemas de IA y cómo toman decisiones. También es importante fomentar la colaboración multidisciplinaria entre expertos en IA, ética y derechos humanos para abordar este desafío de manera integral.

Palabras clave: inteligencia artificial (IA), ética, explicabilidad, principios éticos, transparencia.

Table of contents

Page

1. Introduction 3

2. Methodology 4

3. Results 5

3.1. AI explainability as an ethical principle 5

3.1.1. Approaches 5

3.1.2. Objective, features, and functions 6

3.2. Ethical challenges and concerns associated with the lack of explainability 7

3.3. Approaches to improving explainability 11

3.4. Solutions to the ethical challenges posed by the lack of AI explainability 12

3.5. Existing regulatory initiatives . . . 12

3.6. Education and awareness: actions towards implementing training programs 13

3.7. The need for multidisciplinary approaches 14

3.8. Evaluation and follow-up 15
4. Conclusions **16**
5. CRediT author statement **17**
References **17**

1. Introduction

Algorithms based on artificial intelligence (AI), especially those using deep neural networks, are transforming the way in which humans approach real-world tasks. Day by day, there is a significant increase in the use of machine learning (ML) algorithms to automate parts of the scientific, business, and social workflow (1). This is partly due to increasing research in a field of ML known as *deep learning* (DL), where thousands – or even billions – of neural parameters are trained to generalize when performing a particular task (2).

As AI systems become increasingly complex and powerful, a fundamental concern arises: *how can we understand and explain the decisions and actions of these systems?* Explainability is a fundamental ethical principle that seeks to address the challenge of understanding and explaining how AI systems make their decisions (3). It is seen as vital to ensuring transparency, distributed accountability, and trust in AI (4).

Explainable artificial intelligence (XAI) is a field of AI that seeks to generate explanations understandable to humans with regard to how AI makes decisions, rather than simply accepting answers or conclusions without clear explanations. The term was coined by the Defense Advanced Research Projects Agency (DARPA) in 2016 to address the lack of transparency and understanding in AI systems, especially in military applications (5). XAI promotes a set of technological tools and algorithms which can generate high-quality explanations that are interpretable and intuitive for humans (6,7).

However, in the ethical dimension of AI, explainability implies that a system must not only be able to explain how it reached a certain conclusion, but also ensure that it is not biased, discriminatory, and unfair. From this point of view, explainability constitutes a fundamental ethical principle.

There are several reasons why AI explainability can be considered to be a strong ethical principle. One of them is the need for responsibility and accountability. If an AI system makes decisions that significantly impact people's lives, it is essential for those affected to understand how these decisions were made and the logic behind them. This allows affected individuals to challenge unfair or erroneous decisions and hold those in charge of their implementation accountable.

AI explainability is important to avoid bias and discrimination. AI models often learn from historical datasets that may be biased and contain harmful patterns. If the results of an AI system are not explainable, it becomes difficult to identify and address these biases, which can result in biased decisions that perpetuate injustice and discrimination. By making decisions and the processes involved transparent, any bias or discrimination that may arise can be examined and corrected.

On the other hand, AI explainability is important to avoid negative impacts on areas such as privacy and security. If AI systems make decisions without clear explanations, there is a risk that incorrect conclusions may be drawn, or that the results may be abused to obtain sensitive information from individuals. By ensuring explainability, users can understand and evaluate how their data are being used and make informed decisions about their privacy.

This leads to gaps in understanding how transparency and explainability can affect public trust in AI systems and how issues such as algorithmic bias and injustice can be addressed. In the field of ethics in particular, there are disagreements on how to tackle these problems (1). According to (21), explainability should be a mandatory requirement for all AI systems, while other scholars advocate for more flexible approaches that balance explainability with other values such as efficiency or privacy (3).

This theoretical contradiction poses challenges in developing clear policies and standards for AI, and it extends to the methodological and practical levels. How can AI explainability be defined and measured? What are the best approaches to this effect? How can ethical concepts be translated into concrete practices and policies? These questions still need clear and consensual answers. Based on these concerns, this research aims to analyze the importance of explainable AI as an ethical principle in the development and application of technologies, striving to ensure transparency, accountability, and trust in AI systems.

Analyzing the importance of explainable AI as an ethical principle is fundamental in today's society. Understanding how decisions are made and how the underlying algorithms work provides transparency and accountability in their use. This is essential to ensuring that decisions made by AI are fair, ethical, and comprehensible to users and society at large.

2. Methodology

This work was carried out while following a qualitative approach, using document research techniques and content analysis. For data collection, a bibliographic review of academic databases was conducted (*i.e.*, Xplore, ACM Digital Library, Google Scholar, and arXiv), looking for relevant literature on XAI published in the last five years. The search keywords were .explicable artificial intelligence", ïnterpretable machine learning", and "transparency in AI".

A systematic review was conducted, obtaining a representative sample of the available literature. Keyword sampling and expert experience were applied, the former to search for documents specifically related to the ethical component of XAI, and the latter to identify leading institutions in the field, aiming to ensure the quality and relevance of the sample. 36 documents were selected, including scientific articles, technical reports, and case studies on the applications of XAI in different domains, all of them in English. As a selection criterion, the sources had to constitute a representative sample of the available literature on the ethical component of XAI.

The analysis was carried out using content coding. Key concepts, definitions, problems, solutions, and documented applications of XAI were identified. Data triangulation was used, contrasting the findings of different sources. Finally, integrative synthesis was employed to consolidate the results. The reported approaches were compared, and relationships between key concepts were established, thus obtaining an updated state of the art on XAI.

3. Results

3.1. AI explainability as an ethical principle

As an ethical principle, the aforementioned concept refers to the ability to understand and justify the decisions made and actions taken by AI systems. In an ethical context, it is critical to explain and understand how AI algorithms generate results, especially when they significantly impact people's lives. This is significant in the field of ethical AI because it allows stakeholders to assess the reliability and fairness of the decisions made by their systems.

XAI allows understanding how the data, algorithms, and models used to reach a certain conclusion have been interpreted. It allows users or affected parties to identify and rectify possible biases or discrimination inherent in AI processes. Moreover, it contributes to accountability and transparency in the development and use of AI systems, enabling developers, users, and other stakeholders to audit and understand AI decision-making processes while ensuring that systems follow some crucial standards, namely:

- *Transparency:* the algorithms and processes employed should be clear and understandable to users and stakeholders (9).
- *Justice:* XAI must be impartial and avoid any kind of bias or discrimination. This involves considering different user groups and ensuring that outcomes are equitable for all.
- *Protection of privacy and personal data:* XAI must respect user privacy and ensure data security (1). To comply with privacy regulations, appropriate measures and policies must be implemented.
- *Accountability* implies taking responsibility for any errors or damages that may arise due to the use of AI. It is important to implement mechanisms to correct and prevent these issues (10).
- *Collaboration and participation* are necessary to the development of XAI. This involves including ethics experts, researchers, users, and civil society organizations the process to ensure a more diverse view of ethical impacts and challenges.

3.1.1. Approaches

There are different approaches and types of XAI in the academic literature (11–14). These approaches include the use of rules and logic, *i.e.*, AI systems use logical rules and deductive reasoning algorithms to provide explanations based on logical inferences. Another perspective is the use of interpretable models, such as decision trees or linear regressions (15).

These intrinsically interpretable models allow AI systems to generate explanations based on their characteristics and coefficients. Moreover, these systems can also keep track of all inferences and reasoning steps taken during the decision-making process, allowing to generate retrospective explanations through inference logs.

In the case of neural networks (NNs), techniques such as the visualization of neural activations or feature elimination can be used to make systems more interpretable and explainable. Furthermore, there are approaches such as meta-explainers, which use separate AI systems to explain the decisions and actions taken by other systems.

To measure and attain explainability in AI systems, there are different approaches. One of them is the transparency approach, where the goal is for the system to provide clear and detailed information about the data used, the algorithms applied, and the processing steps implemented. This allows users to understand how decisions are reached and facilitates the identification and correction of possible errors or biases.

Some researchers are also exploring approaches based on simulation or case-level explanations, where the goal is for an AI system to explain its decisions through concrete examples by relating them to similar, previous cases (34).

3.1.2. Objective, features, and functions

According to the above, the main goal of XAI is to address the issue of opacity in AI systems, *i.e.*, the inability to understand how they work and why they make certain decisions. By providing clear and understandable explanations, the aim is to increase transparency and trust in AI systems, both for users and for decision-makers who rely on them.

The key characteristics of XAI systems include transparency, interpretability, and the ability to provide clear and understandable explanations (14, 16). Therefore, XAI systems must be able to show how they arrived at a specific decision or conclusion, what data they used, and how weights or rules were applied in the decision-making process.

These characteristics are especially relevant in critical applications such as medicine, where it is vital to understand how AI systems arrived at a diagnosis or treatment recommendation. They are also important in applications such as security and justice, where decisions can have a significant impact on people's lives.

The primary function of XAI is to provide explanations that are understandable to humans (17). This involves developing techniques that allow AI systems to elucidate their actions and provide evidence for them.

3.2. Ethical challenges and concerns associated with the lack of explainability

The lack of explainability in AI can lead to unfair and harmful decisions in several ways (4, 9, 10, 16, 18, 19), as shown in Fig. 1.

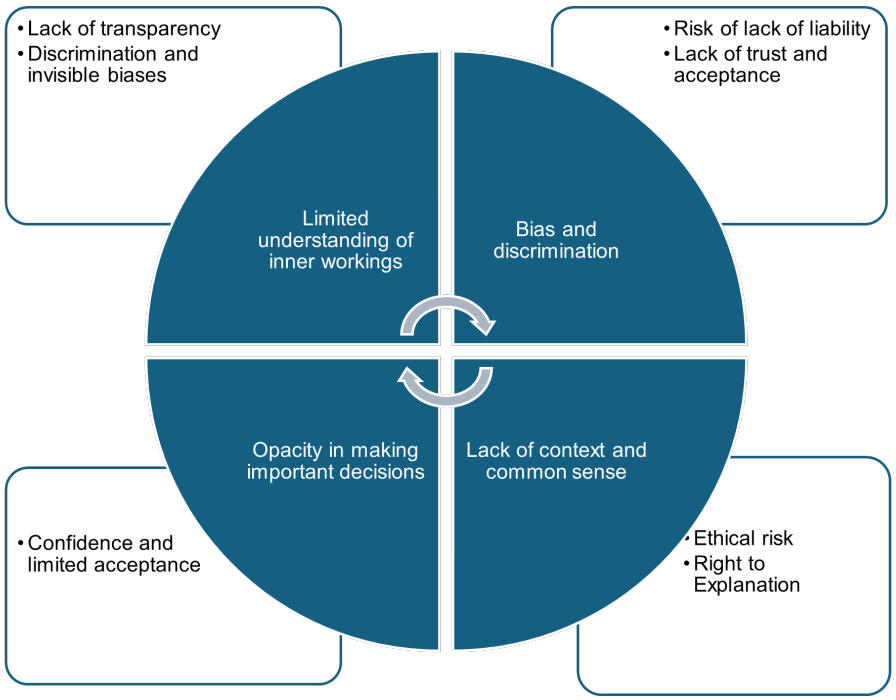


Figure 1. Challenges and ethical concerns related to the lack of explainability in AI

Fig. 1 presents different components related to the ethical challenges and concerns arising from the lack of AI explainability. One of these components is the lack of transparency, which implies the absence of a clear or understandable explanation of how a system or algorithm works. This lack of transparency can lead to discrimination and invisible biases, as the results obtained may disadvantage certain groups or individuals without an easily identifiable cause.

The lack of explainability poses a risk to accountability: if a certain decision or result is not understood, it is difficult to assign responsibilities in the case of issues or errors. This may also entail a lack of trust and acceptance, as people may be hesitant to use or trust opaque systems that cannot be adequately explained. If users do not have full knowledge about the internal functioning of a system or algorithm, their trust in it is likely to be limited. This poses some threats, including the ethical risk of making wrong or unjust decisions, as well as violating people's rights, such as the right to explanation.

Fig. 1 also refers to discrimination, which can be caused by the lack of transparency and explainability in important decision-making processes. This opacity in decision-making can have negative ethical consequences and affect people's rights.

AI has been implemented in vital areas such as medicine, criminal justice, and autonomous driving to improve the efficiency and accuracy of existing systems, with various effects and ethical concerns (10, 18, 20) (Table **I**).

Table I. Applications of AI in vital areas

Vital areas	Ethical concerns
Medical diagnostics	AI systems in medicine may face distrust due to lack of explainability. This raises ethical concerns regarding responsibility and accountability in the event of errors or harm to patients. In addition, discrimination and algorithmic bias have been observed in the development of these systems, such as racial biases in the diagnosis of diseases.
Criminal justice: recidivism prediction	The lack of AI explainability in criminal justice generates distrust and carries ethical risks. Algorithms are regarded as 'black boxes', which limits the ability to challenge AI-based decisions and can perpetuate discrimination and bias in the criminal justice system. This raises ethical concerns regarding justice, equity, and human rights, as algorithms may reproduce and amplify biases in the data used to train them.
Autonomous vehicles	In the context of autonomous vehicles, the lack of transparency carries ethical and legal risks. When an accident caused by an autonomous vehicle occurs, it is vital to be able to understand the decision that led to it. However, the algorithms for autonomous driving are often considered to be black boxes, and they hinder the establishment of liability and guilt. In the field of autonomous driving, ethical concerns encompass the safety of pedestrians and other drivers, as well as responsibility and accountability in accidents – *who* is responsible if an autonomous vehicle is involved in an accident?

One of the main ethical advantages of using XAI is confidence in their results (21). When systems are understandable and explainable, users can comprehend how their results were obtained and rely on their validity and accuracy. That trust is essential, especially in critical applications such as healthcare or justice, where AI-based decisions can have a significant impact.

AI explainability allows detecting and addressing possible biases in data and algorithms. AI systems may be influenced by biases in the training data or introduced by developers (18). The

ability to understand how decisions are made enables the identification and correction of unjust or discriminatory biases, ensuring impartiality and fairness in the decision-making process.

Moreover, explainable AI systems also promote ethical accountability. When they can understand and reflect upon the decisions made by AI tools, people can identify errors, biases, or unintended consequences (10). This facilitates the identification of potential ethical risks and the adoption of measures to mitigate them. Explainability is vital in several cases (22), as presented in Table **II**.

Table II. Cases where AI explainability is vital

Case	Explanation
Disease detection	In disease detection, explainability is vital to ensuring trust and acceptance of the results by healthcare professionals and patients. It allows clinicians to understand how certain conclusions were reached and which patient characteristics influenced the decision. In addition, explainability helps identify possible flaws in the models or data used, and it is relevant when deciding which algorithm to use in clinical practice.
Judicial decision-making	The explainability of AI systems in judicial decision-making is vital to ensuring fairness and transparency. It allows judges and lawyers to understand and question the basis of AI systems' recommendations or conclusions, especially in cases where there may be bias or discrimination. Explainability helps to identify and correct these biases, fostering impartiality and non-discrimination.
Cybersecurity	In cybersecurity, explainability is critical to understanding and addressing cyberthreats and cyberattacks. Security experts need to understand how an attack was made and what vulnerabilities were exploited in order to take corrective action and prevent future incidents. In addition, when it comes to the attribution of responsibility, explainability is needed to identify the culprit and take appropriate legal action.

There are several challenges associated with AI explainability. Some of them are shown in Table **III** (8, 23).

Now, from an economic perspective, explainability can have significant implications. On the one hand, the lack of transparency can undermine consumers' trust in companies and in the use of their data. This can lead to reduced user engagement and, ultimately, to a negative impact on the revenue of businesses that rely on data collection and analysis. On the other hand, transparency can be a

Table III. Challenges associated with the explainability in AI

Challenge	Explanation
Model complexity	Many AI models, such as deep NNs, are highly complex and difficult to understand. Therefore, explaining how a model makes decisions and what features are most relevant to it is a complex task.
Privacy policy	To understand how an AI model makes decisions, it is necessary to analyze the training data used. However, these data may contain sensitive or confidential information. It is necessary to find a balance between explainability and data privacy.
Coherence and consistency	AI models can be inconsistent in their decisions, making it difficult to explain why a particular decision was made at a given time. Ensuring coherence and consistency in decisions is a challenge in ensuring explainability.
Bias and discrimination	AI models can be affected by biases in the training data, which can lead to discriminatory results. It is important to address these biases and ensure explainability, in order to identify and correct any potential discrimination.

competitive advantage for companies, as it can generate greater trust among consumers and therefore increase their loyalty and willingness to share data.

At the policy level, explainability can be vital to ensuring accountability and fair decision-making. If the algorithms used to make decisions – such as those related to loans, employment, or justice – are opaque or biased, there may be negative consequences for certain groups or individuals. Transparency in data collection and use, as well as accountability in cases of incorrect or biased decisions, are key elements to ensure fairness and avoid discrimination.

In cultural and ethical terms, explainability can influence how we perceive data use and automated decision-making. System opacity can lead to mistrust and concerns about a lack of control or knowledge with regard to how our data are used. Likewise, transparency can provide a greater sense of empowerment and autonomy by allowing to understand and question the decisions that affect us.

From a legal perspective, the absence explainability may affect the protection of privacy and individual rights. Transparency in data collection and use can help to ensure compliance with existing laws and regulations, such as the European Union's General Data Protection Regulation (GDPR). In addition, in cases of incorrect or biased decisions, liability may have legal implications and lead to lawsuits or penalties.

3.3. Approaches to improving explainability

So far, the question could be *how to improve AI explainability?* XAI is an important topic from an ethical standpoint. There are different approaches to improving explainability in AI models, including interpretable models, inference rules, and visualization methods (17, 24, 27).

One of the most common approaches to improving explainability is the use of interpretable models. These models, also known as *white-box models*, are those that can be easily understood and explained by humans. Examples of interpretable models include decision trees, linear regression, and decision rules. Another approach is inference rules. These rules allow establishing logical connections between premises and conclusions within a reasoning process. Inference rules can be used to explain predictions by AI models. For example, they can be used to explain why a certain decision was made by a recommendation system.

Visualization methods can also be used to improve the explainability of AI models. These methods allow graphically and comprehensibly representing the decision-making process of a given model. Some examples include flowcharts, tree diagrams, and heat maps. Visual representations help users to more intuitively understand how a certain prediction was reached.

The ethical importance of model interpretability lies in transparency and accountability when using algorithms and models in decision-making processes that can affect people. This includes a) understanding how a certain prediction or decision was reached, b) contributing to increased accountability in the organizations or institutions using AI systems, and c) facilitating trust in the technology. By understanding how algorithms make decisions, people can feel more secure and trust that relevant aspects have been considered, thus avoiding arbitrary decisions.

There are several reasons for the necessity of establishing standards and regulations that promote AI explainability.

Firstly, explainability is crucial for ensuring accountability. If a decision made by an AI system has a negative impact on a person, it is necessary to be able to evaluate and understand the process followed, in order to determine if there has been any bias or discrimination. Without explainability, AI systems cannot be effectively held accountable for their actions.

Secondly, explainability is essential for fostering public trust in AI. Many people are uncomfortable with relying on AI systems whose decisions cannot be understood or explained. The lack of transparency and clarity in AI decision-making can erode trust in this technology and generate resistance to its adoption.

AI explainability is important to upholding ethical principles and people's rights. Automated decisions made by AI systems can affect fundamental rights such as privacy, non-discrimination, and equal opportunities. If we cannot understand how these decisions are made, we cannot ensure that these rights and ethical principles are being respected.

3.4. Solutions to the ethical challenges posed by the lack of AI explainability

To address these challenges, several solutions have been proposed:
Techniques are being developed to enable a better interpretation and understanding of AI models. This includes methods such as the generation of explanations or justifications for decisions made by AI systems and the visualization of the influence of different characteristics on the predictions made. These techniques allow users and stakeholders to understand how a decision is made and to assess its validity and fairness.

Research into approaches to identify and mitigate biases in AI systems is ongoing. This includes systematically analyzing the data used to train the models, identifying potential biases, as well as implementing techniques to reduce their impact. AI model auditing and evaluation can help to ensure that decisions are fair and equitable for all groups involved.

Regulations and standards are being proposed to promote explainability in AI systems. These legal and ethical frameworks can establish clear requirements regarding the transparency and accountability of AI systems, in addition to ensuring fairness and non-discrimination in their implementation.

Approaches such as user-centered design and stakeholder engagement can ensure that decisions on the use of AI are transparent, and that ethical concerns are taken into account. Engaging people affected by AI decisions and enabling human feedback and control can help to avoid unfair or harmful outcomes.

3.5. Existing regulatory initiatives

This section outlines some current initiatives aimed at consigning XAI into valid regulations.

One of the main regulatory initiatives that promotes AI explainability is the European Union's General Data Protection Regulation (GDPR) (26). The GDPR states that individuals have the right to explanations about automated decisions that affect them, including algorithm-based decision-making processes. This regulation requires organizations to provide clear and understandable information about the criteria used to make automated decisions, as well as the consequences and scope of such decisions.

Another regulatory initiative is presented in *Big data: A report on algorithmic systems, opportunity, and civil rights*, a document published by the US White House in 2016. This report emphasizes the importance of explainability in AI systems, especially in critical areas such as healthcare and finance. It proposes regulatory measures that promote transparency, accountability, and explainability in the algorithms used in these sectors.

Another example is the California Consumer Privacy Act (US). This law requires AI systems used by government agencies to be transparent and provide clear and understandable explanations of how decisions affecting citizens are made.

Furthermore, the European proposal for a legal framework on AI (AI Act) aims to establish a comprehensive regulatory framework for AI in the European Union. It includes specific explainability requirements, such as the obligation to provide clear and understandable information about the functioning of AI systems and the expected consequences.

The Institute of Electrical and Electronics Engineers (IEEE) has also developed explainability principles (the IEEE Explainable and Trustworthy Artificial Intelligence initiative) that promote transparency and accountability in AI systems. These principles include generating understandable explanations and documenting the AI decision-making process.

Finally, the UK government has established the Artificial Intelligence Explainability Initiative to promote explainability in AI systems used in the public sector. This initiative involves developing standards and guidelines that foster transparency and allow understanding the decisions made by AI.

The effectiveness of these regulatory initiatives in promoting explainability can be assessed from different perspectives, *e.g.*, the impact they have had on public awareness and widespread concern about the explainability of AI systems. These initiatives have sparked debate and put explainability on the political and business agenda.

Overall, while existing regulatory initiatives represent an important step towards promoting AI explainability, there is a need for a multifaceted approach that combines regulation with technological research and involves multiple actors, such as data scientists, engineers, and policymakers.

3.6. Education and awareness: actions towards implementing training programs

Educating users, developers, and policy makers about this issue is critical to ensuring that AI systems are trustworthy, ethical, and accountable. It is imperative to educate users about AI explainability, as they must understand how such a system works and how decisions are made (1). This will allow them to assess whether these decisions are appropriate and ethical. In addition, users should be informed about the possible bias and discriminatory actions of AI systems and how they can be mitigated.

As for developers, it is critical for them to be aware of the importance of AI explainability from the design and development stages (13). They must acquire skills and knowledge to develop XAI systems. This involves using techniques and algorithms to track and explain the reasoning of their models. Developers must also be trained in ethics and accountability, in order to ensure that AI systems are used fairly and transparently.

Finally, policymakers have a critical role in promoting AI explainability. They must understand its importance and legislate accordingly (27). This implies establishing regulatory frameworks that require the explainability of AI systems in certain contexts, such as automated decision-making in the legal or healthcare fields. They should also encourage research and training and promote collaboration between the public and private sectors, in order to guarantee the responsible development of AI. To implement

training and outreach programs on XAI, different actions can be undertaken (Table **IV**).

Table IV. Actions for training and dissemination regarding AI explainability

Actions	Contents
Courses and workshops	Training programs addressing the fundamentals of XAI, existing techniques, and related ethical and legal challenges can be designed for users, developers, and policymakers.
Conferences and seminars	Educational materials can be developed, such as guides, manuals, or infographics that explain the concepts and techniques related to XAI in a clear and accessible way. These resources could be distributed online or in print.
Research promotion	Research into XAI could be encouraged by funding research projects and programs. This will facilitate the development of more explainable AI techniques and models.
Interdisciplinary cooperation	Collaboration between experts in AI, ethics, law, and other relevant disciplines can be promoted to comprehensively address the implications of XAI.

3.7. The need for multidisciplinary approaches

XAI demands a multidisciplinary approach due to its complexity and its implications for different aspects of society. In this vein,

- *Sociology* can analyze the impact of AI on social structures by studying how it influences work organization, income distribution, and new inequalities. In addition, it can investigate the effects of AI on social interactions and the formation of virtual communities (28).
- *Psychology* can contribute to understanding how people interact and relate to AI systems. Studies on users' perception of AI and how it affects their trust and acceptance could be carried out. Likewise, the possible biases and prejudices of AI models could be analyzed (1).
- *Ethics and law* are fundamental to establishing the principles and regulations guiding the development and use of AI (10, 18, 29). Aspects such as privacy, responsibility, and ethical values must be considered in decision-making processes related to this type of systems.
- *Economics* can assess the economic impacts of AI, focusing on growth and productivity. It can also evaluate potential labor displacements and the restructuring of industries. It is important to analyze how AI can drive innovation and generate economic opportunities, but also how it can lead to inequalities and challenges in the labor market (30).

- *Culture, history, and demographics* can provide a socio-cultural context for AI (19). These disciplines can analyze how these systems can reflect and influence the values, beliefs, and practices of a particular society, as well as the demographic changes that may arise due to their implementation.

3.8. Evaluation and follow-up

Ethical metrics and assessments to measure AI explainability are critical to ensuring transparency and reliability. Some proposals reported in the literature are presented below (18,31,32) (Table **V**).

Table V. Metrics and assessments

Metrics	Assessment
Explainability metrics	These metrics quantify an AI system's degree of explainability. Some common metrics include clarity, comprehensibility, and transparency. An example of this is measuring the amount of information provided by the system to justify its decisions, or its ability of the system to explain the rules or algorithms used.
Model interpretability assessment	This type of assessment focuses on the ability of an AI model to be interpreted and understood by users. Techniques such as data visualization, significant feature extraction, and pattern identification can be used to this effect.
Assessment of social and societal impact	This type of assessment focuses on the impact of an AI system on society in terms of equity, bias, or discrimination. Measures such as differential accuracy can be used in order to assess whether the system shows fair and equitable treatment towards different population groups.

It is important to mention that these metrics and evaluations must be applied ethically, considering all relevant values and ethical principles. Ethical monitoring and updating mechanisms are required to ensure that AI systems remain explainable as they evolve and are updated. Some proposed mechanisms include those shown in Table **VI** (10,21,25).

The importance of the elements in Table VI lies in their contribution to ensuring that AI systems are ethical and meet the established standards. For example, periodic AI auditing processes are essential for regularly evaluating the explainability and ethical impact of AI systems. This enables the identification of potential ethical issues and the timely implementation of corrective actions. Furthermore, this constant review helps to improve the understanding of how a system works, which in turn allows explaining its decisions and actions to users and ethics experts.

Table VI. Ethical requirements and indicators

Indicator	Ethical requirement
AI auditing processes	These processes should be carried out on a regular basis to ensure that AI systems meet the established ethical standards. This involves regularly reviewing and evaluating their explainability and ethical impact.
Continuous feedback	Feedback and communication channels should be established with users and ethics experts in order to gather feedback and concerns about the explainability and ethical impact of AI systems. This feedback can help identify potential issues and improve explainability.
Updated regulatory frameworks and policies	Up-to-date policies and regulatory frameworks that require AI explainability should be established. These policies should consider both technical and ethical aspects.

4. Conclusions

As an ethical principle of AI, explainability is vital due to several reasons:

- First, it promotes transparency and builds trust in how technology works. Users and stakeholders can understand how AI systems make decisions, avoiding the 'black-box' feeling and reducing uncertainty.
- Second, AI explainability makes it possible to ensure accountability and liability for systems and their creators. It is essential to be able to explain and justify the reasons behind the decisions made by AI systems, especially when they have an impact on people's lives. Explainability facilitates the identification of biases, errors, or bad practices. Thus, these issues can be corrected, or the actors involved can be held accountable.
- Third, explainability helps to identify and mitigate biases inherent in the training data or algorithms used in AI systems. Understanding how these biases influence system decisions allows taking steps to ensure fairness and impartiality in the field of AI.
- Fourth, when using AI, it is also critical to ensure compliance with ethical and legal standards. By understanding how decisions are made, one can assess whether relevant ethical and legal principles such as privacy, non-discrimination, and fairness are being met. In addition, explainability facilitates the auditing and monitoring of AI systems.
- Fifth, explainability contributes to improving the acceptance and adoption of AI by society. When people understand how AI works and why it makes certain decisions, they are more likely to trust the technology and feel comfortable using it.
- Finally, explainability reduces the perception of AI as a 'mysterious black box' and aids in overcoming cultural and trust barriers. In short, explainability is essential to ensuring the transparency, accountability, fairness, and acceptance of AI.

5. CRediT author statement

Mario González Arencibia: conceptualization, investigation, methodology, writing (original draft).
Hugo Ordoñez-Erazo: investigation, writing (review and editing).
Juan-Sebastián González-Sanabria: validation, visualization, writing (review and editing).

References

[1] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," *arXiv preprint*, 2017. arXiv:1702.08608. ↑3, 4, 5, 13, 14

[2] M. Huang, V. K. Singh, and A. Mittal, *Explainable AI: interpreting, explaining, and visualizing deep learning*. Berlin, Germany: Springer, 2020. ↑3

[3] A. Hanif, X. Zhang, and S. Wood, "A survey on explainable artificial intelligence techniques and challenges," in *IEEE 25th Int. Ent. Dist. Object Computing Workshop (EDOCW)*, 2021, pp. 81-89. https://doi.org/10.1109/EDOCW52865.2021.00036 ↑3, 4

[4] M. Coeckelbergh, "Artificial intelligence, responsibility attribution, and a relational justification of explainability," *Sci. Eng. Ethics*, vol. 26, no. 4, pp. 2051-2068, 2020. https://doi.org/10.1007/s11948-019-00146-8 ↑3, 7

[5] T. Izumo and Y. H. Weng, "Coarse ethics: How to ethically assess explainable artificial intelligence," *AI Ethics*, vol. 2, no. S1, pp. 1-13, 2021. https://doi.org/10.1007/s43681-021-00091-y ↑3

[6] A. Das and P. Rad, "Opportunities and challenges in explainable artificial intelligence (XAI): A survey," *arXiv preprint*, 2020. arXiv:2006.11371. ↑3

[7] G. Adamson, "Ethics and the explainable artificial intelligence (XAI) movement," *TechRxiv*, Preprint, 2022. https://doi.org/10.36227/techrxiv.20439192.v1 ↑3

[8] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, and F. Herrera, "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fus.*, vol. 58, pp. 82-115, 2019. https://doi.org/10.1016/j.inffus.2019.12.012 ↑9

[9] A. Weller and E. Almeida, "Principles of transparency, explainability, and interpretability in machine learning," *Cogn. Technol. Work*, vol. 3, pp. 1-14, 2020. ↑5, 7

[10] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," *Nat. Mach. Intell.*, vol. 1, no. 9, pp. 389-399, 2019. https://doi.org/10.1038/s42256-019-0088-2 ↑5, 7, 8, 9, 14, 15

[11] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, "Explaining explanations: An overview of interpretability of machine learning," in *IEEE 5th Int. Conf. Data Sci. Adv. Analytics*, 2018, pp. 80-89. https://doi.org/10.1109/DSAA.2018.00018 ↑5

[12] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in *22nd ACM SIGKDD Int. Conf. Knowledge Discovery Data Mining*, 2016, pp. 1135-1144. https://doi.org/10.1145/2939672.2939778 ↑5

[13] A. Weller and H. Aljalbout, "Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models," *JAIR*, vol. 68, pp. 853-863, 2020. ↑5, 13

[14] Z. Lipton, "The mythos of model interpretability," *arXiv preprint*, 2018. arXiv:1606.03490. ↑5, 6

[15] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, "A survey of methods for explaining black box models," *ACM Comput. Surv.*, vol. 51, no. 5, 1-42, 2018. https://doi.org/10.1145/3236009 ↑5

[16] A. Weller and S. V. Albrecht, "Challenges for transparency," in *Proceedings of the AAAI/ACM Conf. AI Ethics Soc.*, 2019, pp. 351-357. ↑6, 7

[17] H. Nissenbaum, *Privacy in context: Technology, policy, and the integrity of social life.* Stanford, CA, USA: Stanford University Press, 2009. https://doi.org/10.1515/9780804772891 ↑6, 11

[18] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," BD&S, vol. 3, no. 2, e2053951716679679, 2016. https://doi.org/10.1177/2053951716679679 ↑7, 8, 14, 15

[19] J. Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms," *BD&S*, vol. 3, no. 1, e2053951715622512, 2016. https://doi.org/10.1177/2053951715622512 ↑7, 15

[20] A. D. Selbst and S. Barocas, "The intuitive appeal of explainable machines," *Ford. Law Rev.*, vol. 87, e1085, 2018. https://doi.org/10.2139/ssrn.3126971 ↑8

[21] A. Weller and L. Floridi, "AIEthics Manifesto," *Min. Mach.*, vol. 29, no. 3, pp. 371-413, 2019. ↑4, 8, 15

[22] V. Dignum, "Responsible artificial intelligence: How to develop and use AI in a responsible way," *ITU J.* (Geneva), vol. 1, no. 6, pp. 1-8, 2021. ↑9

[23] L. Floridi and M. Taddeo, "What is data ethics?" *Phil. Trans. R. Soc. A*, vol. 376, no. 2128, e20180083, 2018. ↑9

[24] Z. Lipton, "The mythos of model interpretability," *arXiv preprint*, 2016. arXiv:1606.03490. ↑11

[25] C. Molnar, *Interpretable machine learning*, 2019. [Online]. Available: https://christophm.github.io/interpretable-ml-book/ ↑15

[26] Unión Europea, *Reglamento general de protección de datos (GDPR)*, 2016. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/ ↑12

[27] European Commission, *Ethics guidelines for trustworthy AI*, 2019. [Online]. Available: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai ↑11, 13

[28] A. Brynjolfsson and A. McAfee, *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. New York, NY, USA: WW Norton & Company, 2014. ↑14

[29] B. Green and S. Hassan, "Explaining explainability: A roadmap of challenges and opportunities of machine learning interpretability," in *24th ACM SIGKDD Int. Conf. Knowledge Discovery Data Mining*, 2019, pp. 2952-2953. ↑14

[30] L. Floridi and J. W. Sanders, "On the morality of artificial agents," *Min. Mach.*, vol. 14, no. 3, pp. 349-379, 2004. https://doi.org/10.1023/B:MIND.0000035461.63578.9d ↑14

[31] A. B. Arrieta, V. Dignum, R. Ghaeini, A. López, V. Murdock, M. Osborne, and A. Rathke, "Transparent AI: An overview,"*Art. Inte.*, vol. 290, pp. 1-43, 2020. ↑15

[32] L. Liao, A. Anantharaman, and K. Pei, "On explaining individual predictions of machine learning models: An application to credit scoring," *arXiv preprint*, 2018. arXiv:1810.04076. ↑15

Mario González-Arencibia

He holds a Bachelor's degree in Political Economy from Universidad del Oriente, a Master's degree in International Economics from Universidad de La Habana, and a PhD in Economic Sciences from Universidad del Oriente. Additionally, he holds a PhD in Economics from Universidad Complutense de Madrid. With 38 years of experience in higher education, he currently serves as a full professor at Universidad de Ciencias Informáticas in Havana, Cuba. He has been involved in digital-era research for over 20 years and has authored over 200 publications, including more than 20 books and articles, covering areas such as economics, politics, and ethics.
Email: mgarencibia@uci.cu

Hugo Armando Ordoñez-Erazo

He is a systems engineer from Fundación Universitaria San Martín. He holds an MSc in Computing and a PhD in Telematics Engineering from Universidad del Cauca. He served as a full-time professor in both undergraduate and graduate programs, and as a researcher at the Department of Computer Science of Universidad de San Buenaventura. He currently works as a full-time professor at Universidad del Cauca. He is a co-author of three books, five JCR articles, and over thirty SJR and SciELO articles. He has supervised 11 graduate students. His research interests include data mining, data analysis, machine learning, information retrieval, and software engineering. His main international collaborations are with Spain, Mexico, Cuba, Ecuador, and Brazil. He is classified as an associate researcher by the Colombian Ministry of Science, Technology, and Innovation (Minciencias).
Email: hugoordonez@unicauca.edu.co

Juan Sebastián González-Sanabria

He's a Systems and Computing Engineer from Universidad Pedagógica y Tecnológica de Tunja (UPTC). He holds two specializations: one in Databases from UPTC and another one in Scientific and Technological Information Management from Universidad Nacional de La Plata. Additionally, he holds a Master's degree in Software and Information Systems from Universidad Internacional de La Rioja. He has attended training courses on Big Data (MIT), Pedagogy for Virtual Education (UTP), and Scientific Journal Editing (Diploma, Universidad de Negocios y Ciencias Sociales). He has taught courses in the Database and Research line, and he served as the director of UPTC's School of Systems and Computing Engineering in the 2014-2016 and 2021-2022 periods. He has been a Technology and Innovation Advisor for UPTC since 2017, and an editor of UPTC's Pensamiento y Acción Journal since 2019. He has authored more than a dozen presentations and scientific articles in various journals, mainly in the areas of data analysis and development.
Email: juansebastian.gonzalez@uptc.edu.co

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1221

DOI: https://doi.org/10.14483/23448393.21408

Research

Mobile Application for Recognizing Colombian Currency with Audio Feedback for Visually Impaired People

Aplicación móvil para el reconocimiento de moneda colombiana con retroalimentación de audio para personas con discapacidad visual

Camila Bolaños-Fernández[1] and Eval Bladimir Bacca-Cortes[1]*

[1]Universidad del Valle (Cali, Colombia)

Abstract

Context: According to the census conducted by the National Department of Statistics (DANE) in 2018, 7.1 % of the Colombian population has a visual disability. These people face conditions with limited autonomy, such as the handling of money. In this context, there is a need to create tools to enable the inclusion of visually impaired people in the financial sector, allowing them to make payments and withdrawals in a safe and reliable manner.
Method: This work describes the development of a mobile application called *CopReader*. This application enables the recognition of coins and banknotes of Colombian currency without an Internet connection, by means of convolutional neural network models. CopReader was developed to be used by visually impaired people. It takes a video or photographs, analyzes the input data, estimates the currency value, and uses audio feedback to communicate the result.
Results: To validate the functionality of CopReader, integration tests were performed. In addition, precision and recall tests were conducted, considering the YoloV5 and MobileNet architectures, obtaining 95 and 93 % for the former model and 99 % for the latter. Then, field tests were performed with visually impaired people, obtaining accuracy values of 96 %. 90 % of the users were satisfied with the application's functionality.
Conclusions: CopReader is a useful tool for recognizing Colombian currency, helping visually impaired people gain to autonomy in handling money.

Keywords: mobile application, convolutional neural network, visually impaired people, Colombian currency recognition

Article history

Received: 12th/Oct/2023
Modified: 10th/Mar/2024
Accepted: 23th/Apr/2024

Ing, vol. 29, no. 2, 2024, e21408

* **Correspondence:** bladimir.bacca@correounivalle.edu.co

Resumen

Contexto: Según el censo realizado por el Departamento Nacional de Estadística (DANE) en 2018, el 7.1 % de la población colombiana tiene una discapacidad visual. Estas personas enfrentan condiciones con autonomía limitada, como lo es el manejo de dinero. En este contexto, es necesario crear herramientas que permitan la inclusión de las personas con discapacidad visual en el sector financiero, permitiéndoles realizar pagos y retiros de manera segura y confiable.
Método: Este trabajo describe el desarrollo de una aplicación móvil llamada *CopReader*. Esta aplicación permite el reconocimiento de monedas y billetes de la moneda colombiana sin conexión a Internet, mediante modelos de redes neuronales convolucionales. CopReader fue desarrollada para ser utilizada por personas con discapacidad visual: toma un video o fotografías, analiza los datos de entrada, estima el valor de la moneda y utiliza retroalimentación auditiva para comunicar el resultado.
Resultados: Para validar la funcionalidad de CopReader, se realizaron pruebas de integración. Además, se llevaron a cabo pruebas de precisión y *recall*, considerando las arquitecturas YoloV5 y MobileNet, donde se obtuvo 95 y 93 % para el primer modelo y 99 % para el segundo. Luego, se realizaron pruebas de campo con personas visualmente discapacitadas, obteniendo valores de exactitud del 96 %. El 90 % de los usuarios quedaron satisfechos con la funcionalidad de la aplicación.
Conclusiones: CopReader es una herramienta útil para el reconocimiento de la moneda colombiana, ayudando a las personas con discapacidad visual a ganar autonomía en el manejo del dinero.

Palabras clave: aplicación móvil, red neuronal convolucional, personas con discapacidad visual, reconocimiento de moneda Colombiana

Table of contents

Page

1. Introduction 2
2. Related works 3
3. CopReader development process and dataset building 6
4. Design and implementation of CopReader 8
   - 4.1. Design of CopReader 8
   - 4.2. Implementation of CopReader . . 9
5. Results and discussion 16
6. Conclusions 21
7. Acknowledgments 22
8. CRediT author statement 22

References 22

1. Introduction

The human visual system is in charge of up to 90 % of the interaction with the environment (1). This system allows gathering, processing, and obtaining images from the surroundings. The brain is part of this system, as it is where images are processed. Another important part includes the eyes and the optic nerve, which gather all the necessary information from the environment and transmit it to the brain. Therefore, a total or partial reduction of visual capabilities is regarded as an impairment, which could

be categorized into two groups: blindness and low vision (2). Visually impaired people have many difficulties when it comes to identifying low- and head-level obstacles and currency, and they lack awareness of their surroundings. These difficulties affect their emotional state, leading to depression, frustration, and anxiety (3).

Nowadays, 2200 million people have a visual impairment, and 11.9 million of them are blind or have severe visual issues (1). In Colombia, 62.17 % of the total population with disabilities have a visual impairment (4), which is why Law 1346 of 2009 adopted conventions regarding people with disabilities. For blind people, thanks to this law, braille language is required in elevator buttons, public signaling, and bus stops, among others. However, one of the most significant issues of visually impaired people is unemployment; 62 % of them have no formal job (5), since most companies do not possess a sufficient infrastructure to ensure inclusiveness. Another difficulty faced by visually impaired people is access to the financial system. In Colombia, the Bank of the Republic implemented two inclusive methods to help visually impaired people to identify Colombian currency: the Braille system (2010) and tactile marks (2016) (6). However, both measures suffer from the wear generated by constant use, making these marks lose their relief.

Additional solutions to Braille systems or printing marks on currency have focused on technological options, such as Orcam (7), a visual-based system with multiple recognition functionalities and audio feedback that is lightweight and compact but very expensive. Another option is using smartphones, which are currently popular. In this context, mobile apps such as CashReader (8) have been developed, and there are many others. These focus on one specific currency, but not the Colombian one.

Therefore, in this work, the development of a mobile application called CopReader is described. CopReader enables the recognition of coins and banknotes of Colombian currency without an Internet connection, by means of convolutional neural network (CNN) models. CopReader was developed to be used by visually impaired people. It is simple to use: it takes a video or pictures, analyzes the input data, estimates the currency value, and uses audio feedback to communicate the result.

This paper is organized as follows: section 2 outlines some related works; section 3 describes the development of CopReader and the dataset used; section 4 delves into the development of CopReader; section 5 presents and discusses the results; and section 6 provides our conclusions.

2. Related works

The development of inclusive technologies based on computer vision and machine learning is an interesting research field. Table **I** summarizes the most representative works related to this research. These works were compared while considering the recognition goal, the platform used, the input data, the algorithm employed, the size of the dataset, the type of feedback, and the accuracy obtained. These criteria are important technical aspects to develop this kind of inclusive technology.

Table I. Related works

Ref.	Rec. Goal	Platform	Input Data	Algorithm	Dataset	Feedback	Accuracy
(9)	Objects	Smartphone	Video	YoloV3	N/A	Audio	N/A
(12)	Banknote	Emb. Syst.	Photo	Perceptron M.	13 200 img.	Glove	96 %
(10)	Objects	Smartphone	Video	YoloV3, RCNN	80 000 img, 80 classes	Audio	85.5 %
(13)	Objects	Glasses	Photo	YoloV3	330 000 img.	Audio	96.3 %
(14)	Addresses	Smartphone	Photo	Perceptron M., MobileNet, CNN	600 000 img.	Video	94.6 %
(15)	Car license plates	Emb. Sys.	Photo	YoloV2	500 img.	Video	94.9 %
(16)	Coins	Smartphone	Photo	Perceptron M.	48 img.	Audio	82 %
(17)	Coins	Smartphone	Photo	AlexNet	8 320 img.	Audio	72.2 %
(18)	Coins	Smartphone	Photo	SIFT Matching	700 img.	Audio	65.7 %
(11)	Objects and banknote	Smartphone	Photo	CNNdroid	N/A	Audio	66.6 %
(19)	Banknote	Smartphone	Photo	MobileNet	12 160 img.	Audio	96.6 %
(20)	Coins	Emb. Sys.	Photo	AlexNet	1 600 img.	Video	N/A
(21)	Banknote	Emb. Sys.	Photo	SVM, AlexNet	504 img.	Prot.	99.7 %
(22)	Banknote	Emb. Sys.	Photo	kNN, CNN	80 img.	Prot.	95.6 %
(23)	Banknote	Emb. Sys.	Video, Photo	CNN	70 000 img.	Prot.	94.3 %

The works presented in Table I show that it is important to recognize banknotes, coins, and objects in general, as this provides visually impaired people with more autonomy in daily activities. Works such as (9, 10), and (11) focus on low and head-level objects, which are potentially dangerous for the aforementioned population. The studies on banknote and coin recognition shown in Table I are especially designed for some currencies, and they do not support others.

Another important aspect is the platform used. Most of the reviewed works use smartphones to implement the proposed solution, but employing embedded systems is also popular. (13) propose the use of smart glasses, albeit with no onboard processing capabilities. These glasses also need remote backend server support, which means a permanent Internet connection. The works that use smartphones run their solutions offline and take advantage of the current capabilities of these devices. Using embedded systems such as the Raspberry Pi (12, 15, 20) is a popular option, as it is very flexible. However, short response times are needed, not 2 s or even 3.6 s, as reported in these works. These values are not enough, as per the needs stated by surveyed test subjects, who need to obtain a solution just by placing the currency in front of the camera.

Visually impaired people usually walk around their environment carefully, but their motions are performed at normal speed. Therefore, a high data acquisition rate is a significant requirement. Most of

the works reported in the literature (as shown in Table **I**) employ a shot of the scene to be processed, but some of them (9, 10, 23) use video to recognize objects.

Object recognition (*i.e.*, objects in general, coins, or banknotes) is a challenging task, where changes in the background, lighting conditions, occlusions, and/or partial information hinder the obtainment of high accuracy measures. Thus, it is common to use deep learning algorithms, such as Yolo, RCNN, MobileNet, AlexNet CNNdroid, or the Multilayer perceptron. Other machine learning algorithms have been used, *e.g.*, SVM (21) or classical feature matching based on SIFT (Scale-Invariant Feature Transform) (18). Normally, visually impaired people need portable and reliable solutions, which technically means the use of an algorithm with high confidence but low parametrization. Therefore, the works reviewed used Yolo (9, 10, 13, 15), MobileNet (14, 19), AlexNet (17, 20, 21), or R-CNN (10, 11, 14, 22, 23).

Using deep learning or machine learning algorithms implies using a dataset that has a high number of samples and is well organized. Therefore, the dataset used and the accuracy achieved in a given work are related. Proposals such as (9, 11, 15, 21), and (22) could be over-trained, and their real-time performance (which is not reported by the authors) could be very different. However, in general, the works reviewed in Table **I** use large datasets to ensure diversity.

User feedback is a very important aspect of human-machine interaction. Some cases, such as (14, 15), and (20) are solutions oriented to slight visual impairments. Nevertheless, for people with high levels of visual impairment or blindness, audio feedback is preferable – or, alternatively, the use of gloves (12) to transmit the result of recognition tasks via an established touch code. Some other works (21–23) present prototypes that are tested on embedded systems but are not fully developed for actual interaction with visually impaired people.

The accuracy of any proposed solution is important, as it indicates whether the approach is reliable. Works such as (11, 16, 17), and (18) report low accuracy values, since the authors use classical matching techniques (18), which are not enough to deal with the complexity of the problem complexity. In other cases, the authors use datasets with a very small number of images (11, 16). Successful currency recognition is a complex problem if one only uses partial information, if there is occlusion, if blurry images are captured, if the banknotes are dirty or damaged, or if they coins are worn. Furthermore, in the case of (17), hyperparameter optimization showed improved accuracy. It is worth noting that the works that used Yolo, MobileNet, R-CNN, and AlexNet reported accuracy values ranging from 85.5 to 96.6 %.

The works shown in Table **I** mention useful properties to develop our proposal, the CopReader mobile application: 1) using a smartphone is preferable, since visually impaired people are familiar with using these devices in conjunction with mobile applications such as Google Talkback (24); 2) the input data should include photos and videos; 3) the deep learning algorithm should be reliable and have a small number of hyperparameters; 4) user feedback should include audio, an inclusive alternative for most visually impaired people.

3. CopReader development process and dataset building

Fig. 1 depicts the development of CopReader, which consists of two stages. The first stage is building the dataset and developing, training, and validating the neural network model. This process is iterative, meaning that the model is continuously validated, the results are analyzed, and the model is re-trained if it does not meet the expectations. The second stage involves using CopReader with the trained neural network model.

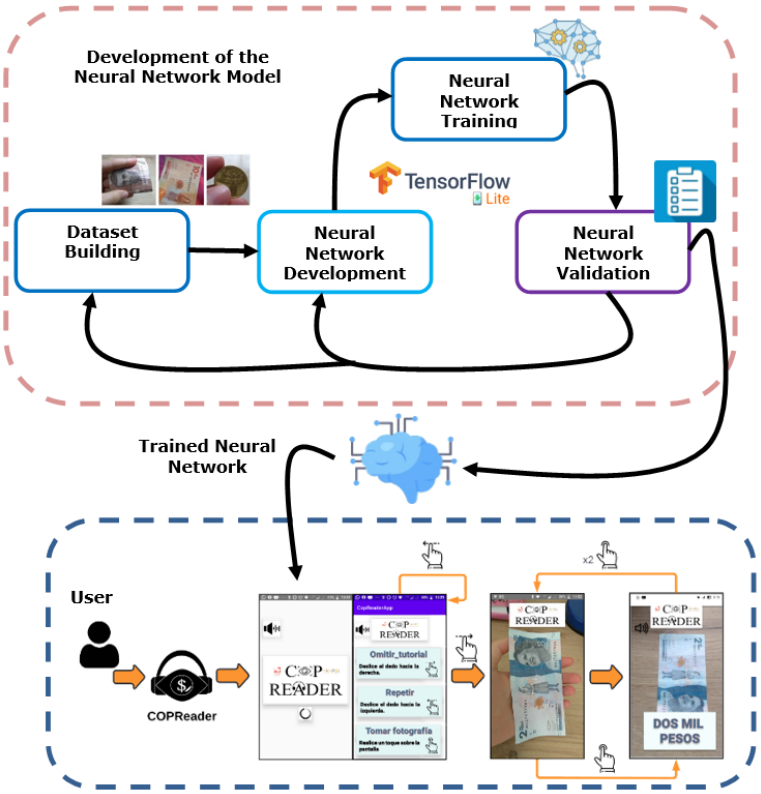


Figure 1. CopReader development process

Building the dataset is crucial for applications based on neural network models. In this specific case, to recognize Colombian currency, the dataset must be created from scratch. To this effect, it is important to clarify the classes of coins and banknotes that CopReader will recognize. Fig. 2 shows the coins and banknotes currently used by Colombia's Bank of the Republic. The following classes were defined: $50 pesos coin, $100 pesos coin, $200 pesos coin, $500 pesos coin, $1000 pesos coin, $2000 pesos banknote, $5000 pesos banknote, $10 000 pesos banknote, $20 000 pesos banknote, $50 000 pesos banknote, and $100 000 pesos banknote. Fig. 2 shows the one face of the banknotes and coins, but the dataset included both faces.

Image acquisition was performed using a 13 Mpixel camera and considering different conditions, such as lighting, occlusion, the state of the currency, and the point of view. Figs. 3a to 3f show examples

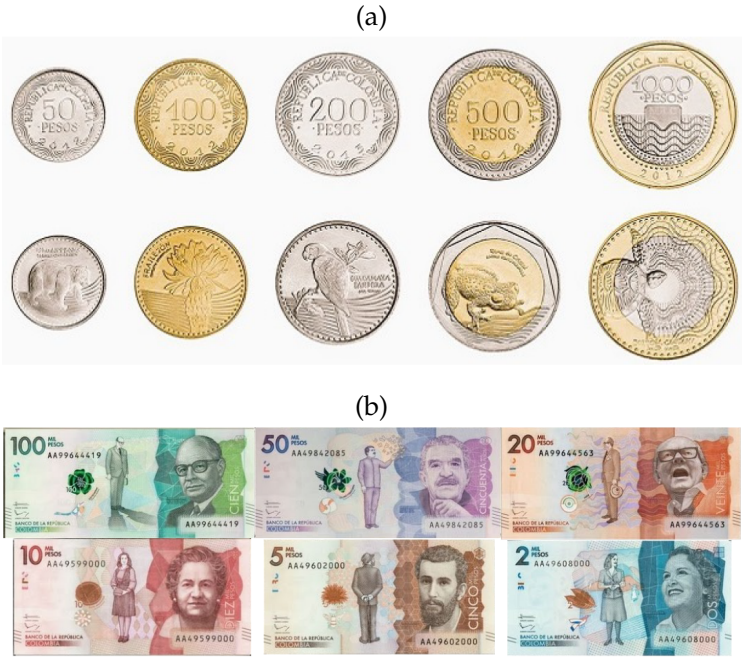


Figure 2. a) Coins and b) banknotes used in Colombia

of these conditions, *i.e.*, crumpled banknotes, dirty coins, written banknotes, blended banknotes, incomplete banknotes, and occluded coins. There are examples for each class of coins and banknotes.

One common cause of recognition errors is incorrectly captured input images. To account for this issue, another class was added, *i.e.*, 'without currency'. This new class is composed of images from (Figs. 3g to 3i) that are blurry, moved, or out of focus. It is important to remember that CopReader uses a vision-based data input to recognize coins and banknotes of Colombian currency; since it has no other type of data input, it cannot detect counterfeit bills – this typically requires the use of UV lamps or texture information.

The initial dataset had 400 3120 x 460 pixel images in each of the 11 classes, for a total of 4400. These images were uploaded to the RoboFlow software tool (25) for proper labeling, as well as to apply data augmentation techniques regarding brightness (±10 %), darkening (±10 %), rotation (90°), out-of-focusup 2 pixels, and horizontal flip. The dataset included 8640 images, with 720 images per class. 80 % of the images were used for training, 10 % for validation, and 10 % for testing.

The related works presented in Table I show that the most commonly used neural network models that exhibit high accuracy values and have suitable configuration hyperparameters are Yolo (9,10,13,15) and MobileNet (14, 19). Thus, in this work, we selected the Yolo V5 nano and MobileNet V2 models for the development, training, and validation phase. The models' pipeline was implemented in Google Colaboratory (26). More details in this regard are provided in section 4.2. Testing was conducted with visually impaired and non-impaired individuals.

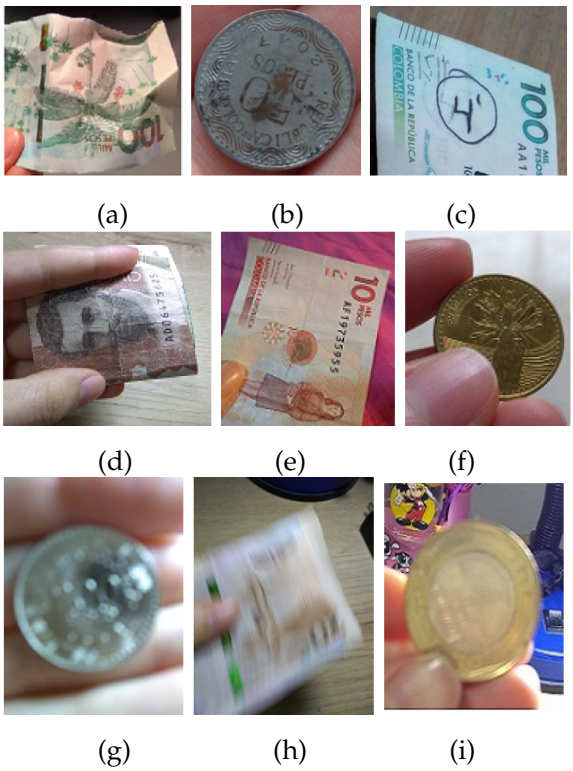


Figure 3. Images included in the dataset: a) crumpled, b) dirty, c) written, d) blended, e) incomplete, f) occluded, g) blurry, h) moved, and i) out-of-focus coins and banknotes

4. Design and implementation of CopReader

CopReader was developed while considering the analysis presented in section 2 and by following the RUP (Rational Unified Process) methodology (27) to document the software engineering process. The RUP methodology includes the following deliverables: functional and non-functional requirements, a conceptual diagram, real use cases, and sequence, relational, class diagrams. However, due to space limitations, real use cases and sequence diagrams are not included in this manuscript.

4.1. Design of CopReader

CopReader was developed while considering the following functional and non-functional requirements:

- *Functional requirements:*
  - Visually impaired users should be provided with an initial tutorial about how the application works.
  - The user should be able to omit the tutorial by sliding the screen to the right.
  - The user should be able to repeat the tutorial by sliding the screen to the left.

- The user should be able to activate the mobile phone's camera by sliding the screen to the right. The input image/video will then be analyzed by the neural network model to estimate the currency in front of the camera.
- The user should receive audio feedback regarding the coin or banknote in front of the camera.
- The user should be able to repeat the audio feedback by sliding the screen to the left.
- The user should be able to repeat the image/video acquisition process by touching the screen twice.

- *Non-functional requirements:*
  - CopReader was developed in Android Studio 4.0
  - CopReader use OpenCV 3.4 and the TensorFlow 2.0 library
  - The minimum camera resolution allowed should be 8 Mpixels.

After defining the functional requirements, the conceptual diagram of the app, shown in the inferior part of Fig. 1, was elaborated. CopReader is a mobile application that requires simple interactions, as it is aimed at visually impaired people. The app assumes that users have their headphones enabled. Then, it presents a tutorial on how to use it. This tutorial is read by a common mobile application called *Talkback*, typically installed in mobile phones as an accessibility tool. This is useful for new users, who are able to repeat the tutorial by sliding the screen to the left; otherwise, they can omit it by sliding the screen to the right. Afterwards, when users slide the tutorial screen to the right, the data acquisition process starts: the neural network model processes the input data and provides an estimation of the coin or banknote in front of the camera. This estimation is communicated to the user via audio feedback, which can be repeated by sliding the screen to the left. Users can repeat the currency estimation process by touching the screen twice.

Fig. 4 shows the class diagram of CopReader. The *MainActivity* and *Initialization* classes start the mobile application. The *Camera* class provides the data input to the *Classifier* class, which in turn processes the input data, yielding an estimation of the currency in front of the camera. Finally, the results are stored in the *Result* class, which generates the audio feedback for the user.

4.2. Implementation of CopReader

CopReader's GUI is very simple since it is aimed at visually impaired people. The bottom part of Fig. 1 shows the final implementation. This section explains how the neural network models were trained, validated, and selected to be used within the app.

The YoloV5 nano neural model was trained over 150 epochs, using a batch size of 32 and random weight initialization. Fig. 5a presents the precision-recall training results, showing values close to 1 and a mAP0.95 (*i.e.*, a mean average precision at an intersection over a union threshold of 0.95) of 0.89, indicating a high generalization capability. In addition, the box (*box_loss*) and object (*obj_loss*) loss values shown in Fig. 5a, which are related to the training and validation phases, do not show

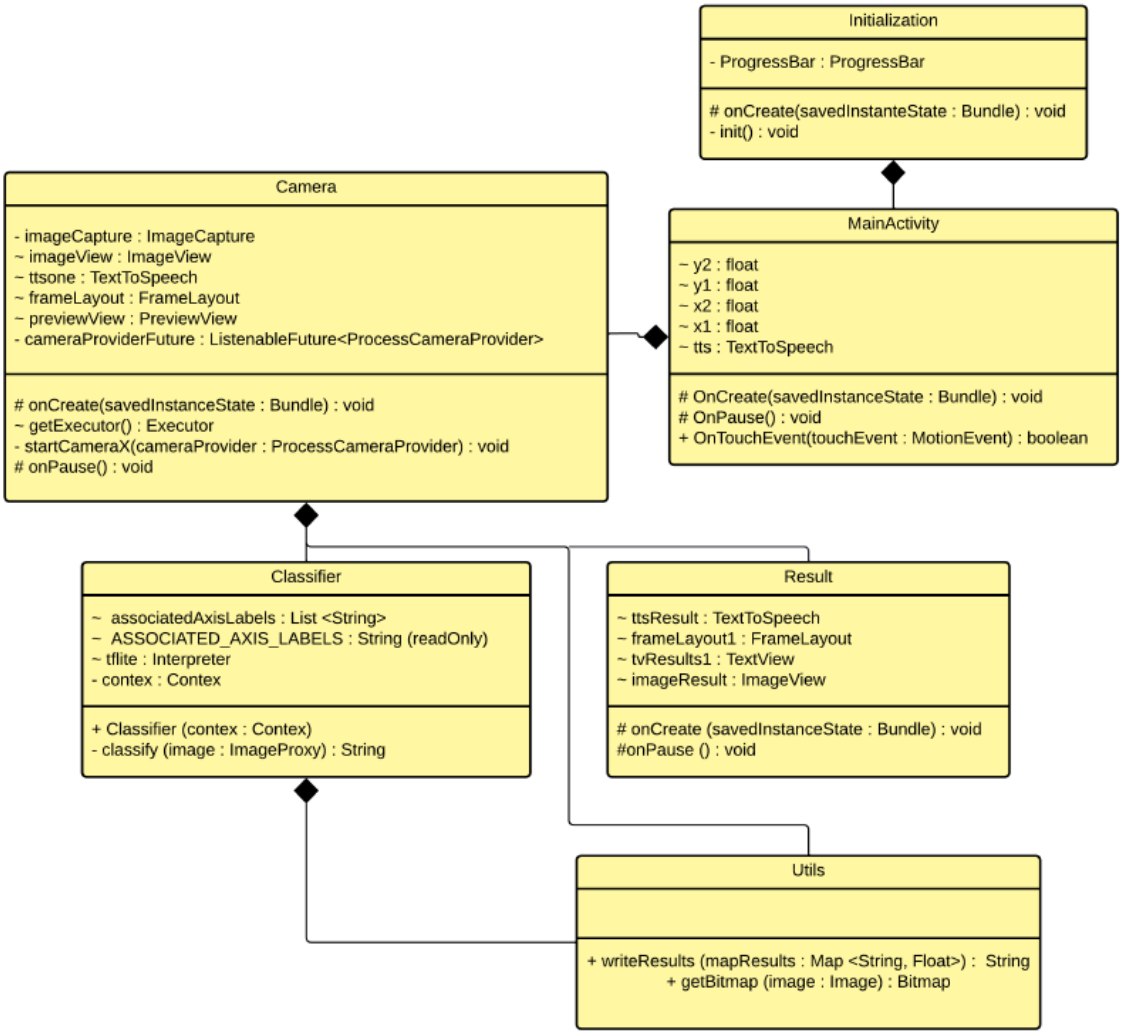


Figure 4. CopReader class diagram

evidence of overtraining. Fig. 6a shows the validation results for the confusion matrix of this neural model without data augmentation. There are no false negatives with a background, but there are false positives with a background of up to 20 %. Moreover, the $100 000 pesos banknote exhibits a high value of false negatives (9 %).

To reduce the number of false negatives and false positives, a neural model with data augmentation was trained and validated. Fig. 5b shows the precision-recall results for this case, which are close to 1 and report a mAP0.95 of 0.85. As shown in Fig. 6b, the confusion matrix validation results show a 2 % reduction in the number of false negatives with respect to the model without data augmentation. Table II summarizes the training results for YoloV5. It can be observed that the accuracy metric is similar for both models, but the other metrics differ by up to 7 %. However, the mAP0.95 metric differs by 4 %,

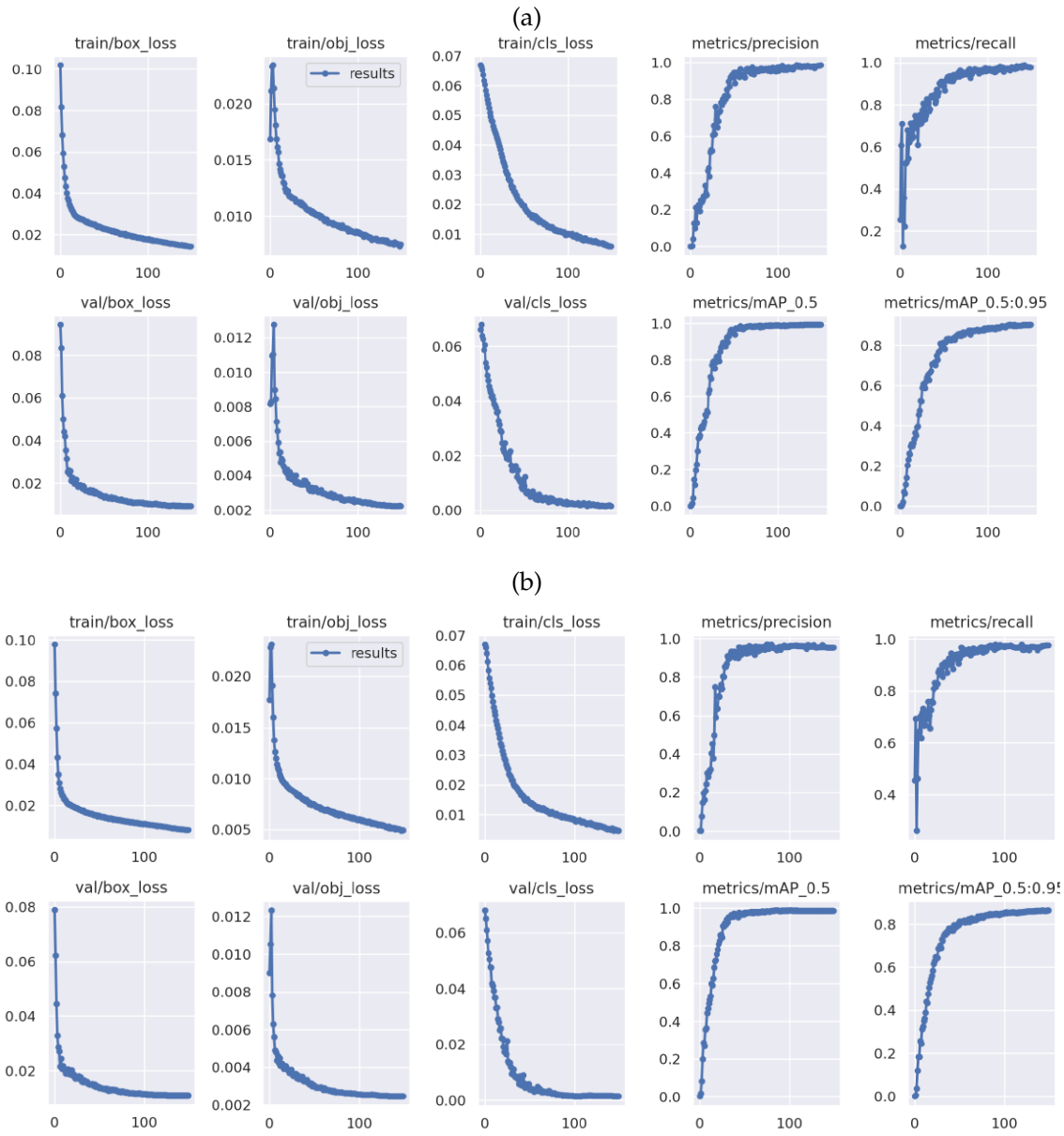


Figure 5. Training results for YoloV5 a) without and b) with data augmentation

with the second neural model (with data augmentation) exhibiting a better robustness against false negatives and false positives. Another reason for this mAP0.95 decrease could be that the Roboflow software tool does not rotate the delimiter box when rotating the sample image.

The second neural model selected was MobileNetV2. The first training attempt is presented in Fig. 7a. This attempt used 40 epochs, a batch size of 32, and random weight initialization. The obtained accuracy for validation was at least 0.95, with loss values lower than 0.2.

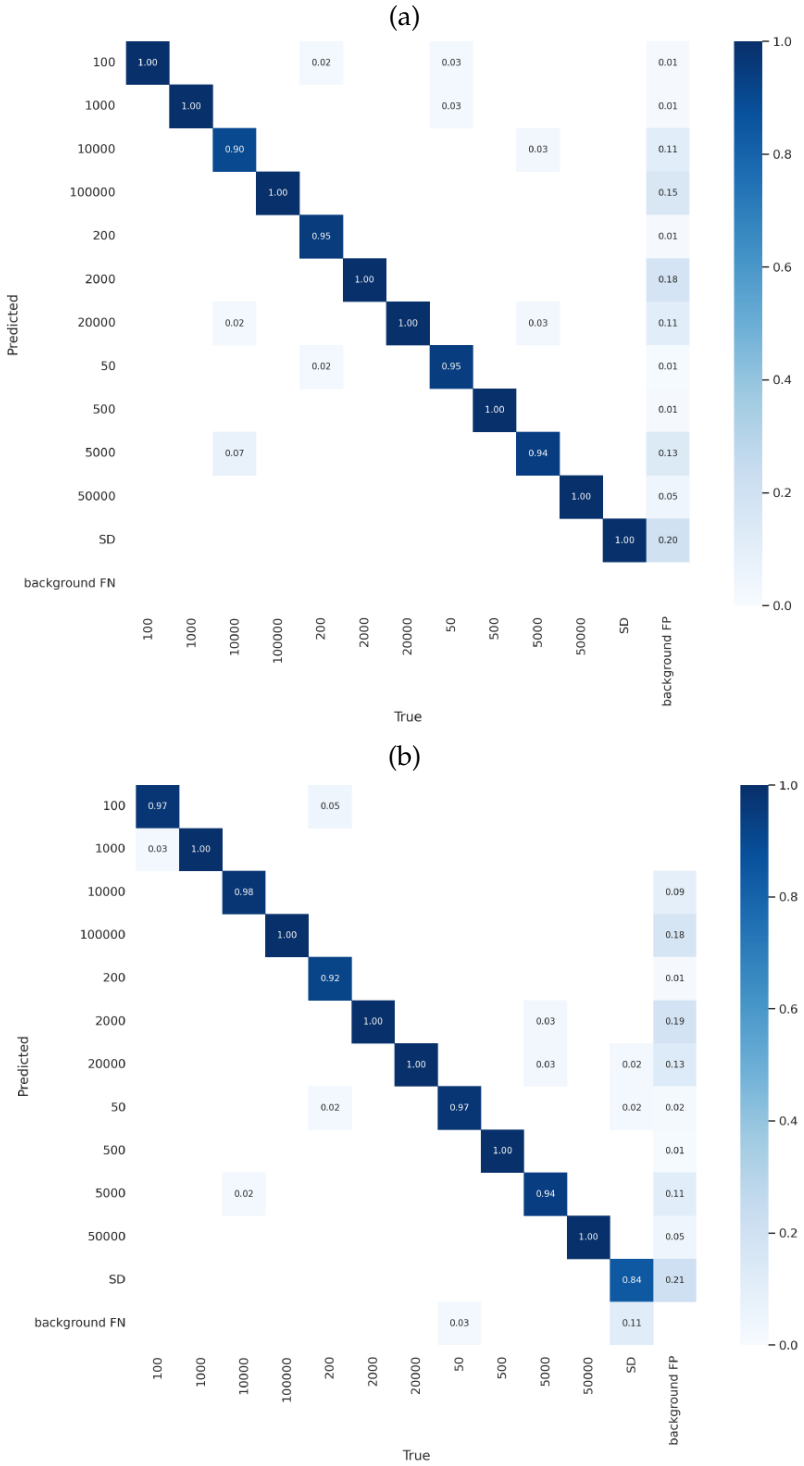


Figure 6. a) Confusion matrix for YoloV5 a) without and b) with data augmentation

Table II. Summary of training results

Metrics	YoloV5	MobileNetV2		
	Without data aug.	With DATA AUG.	Without data aug.	With data aug.
Accuracy	0.92	0.90	0.98	0.99
Precision	0.95	0.88	0.98	0.99
Recall	0.93	0.92	0.98	0.99
mAP	0.89	0.85	0.98	0.99

Fig. 8a shows the confusion matrix validation results. Eight out of 12 classes do not report false negatives, and two exhibit up to 5 % of false negatives. However, there are classification errors, such as those observed with the $200 (9 %) and $5000 (8 %) pesos classes. To reduce the number of false negatives and false positives, data augmentation techniques were applied, as described in section 3. Figs. 7b and 8b show the training results for this model with data augmentation, considering 30 epochs, a batch size of 32, random weight initialization, and L2 regularization, indicating accuracy values over 0.95 and loss values lower than 0.1. Fig. 8b shows the confusion matrix validation results. Note that the false negatives were reduced by 11 % with respect to MobileNetV2 without data augmentation, and the misclassification of the $200 pesos was reduced to 4 %. Table II summarizes the training results obtained for MobileNetV2. In general, the studied metrics are close to 1 (0.99), and the last training of MobileNetV2 reduced the number of false negatives by up to 6.2 %.

In light of that shown in the table II, to select the most suitable neural model, it is important to consider that there is a meaningful difference between the results obtained. In this regard, an analysis of variance (ANOVA) was performed. Two hypotheses were proposed: 1) the group means are equal (null hypothesis), and 2) they are different (alternative hypothesis). To apply this type of analysis, the following conditions must be satisfied: a normal data distribution, the homogeneity of variances, and an independent group of measures. Fig. 9 shows the residual histogram. Using these residuals, the Levene test was used to check for homogeneity, with the significance level (*alpha*) set at 0.05. The p-value computed with these residuals was Pr = 0.05015. Therefore, since Pr $>$ *alpha*, the data were homogeneous. As the experiments performed did not depend on each other, they were independent.

Afterwards, the ANOVA test was performed. Given a significance level of 0.05, the f- and p- values were computed, resulting in 41.29 and 0.03, respectively. As f $>$ p and p $<$ 0.05, the null hypothesis was rejected (28). This means that there were significant differences in the group of data. Therefore, since the performance metrics of MobileNetV2 were higher, this model was selected for implementation on the smartphone. Moreover, MobileNetV2 has a lower size (3.4 Mb) which is adequate for execution in embedded systems. Before migrating this model to the device, the following steps were required: first, the neural model had to be converted to TensorFlow Lite; second, the model had to be optimized for execution on a smartphone; and third, we had to test whether this optimization could impact the performance metric. This was not the case, since the results obtained with the optimized model also achieved a 99 % accuracy.

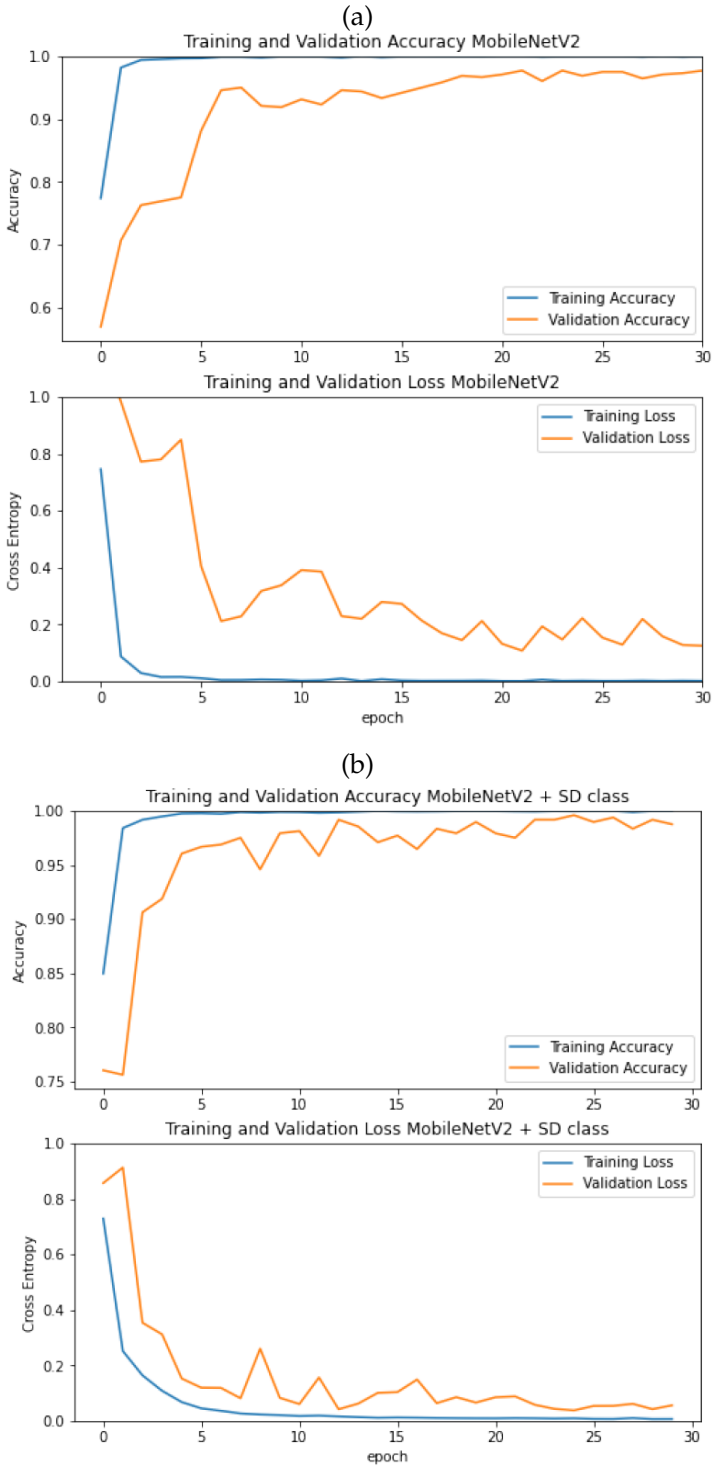


Figure 7. Training results for MobileNetV2 a) without and b) with data augmentation

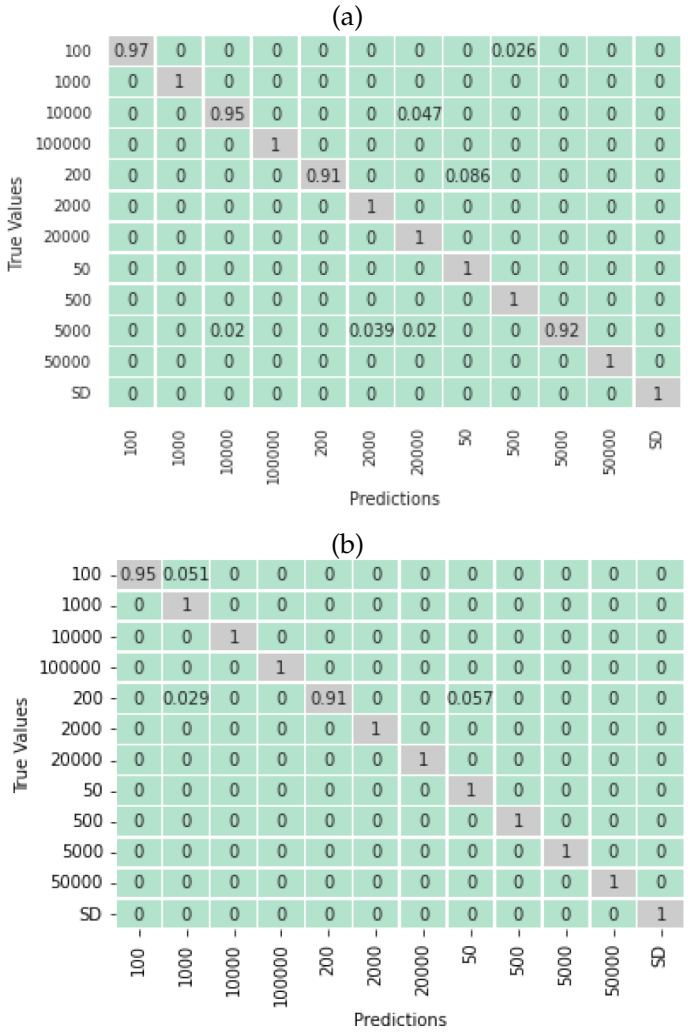


Figure 8. a) Confusion matrix for MobileNetV2 a) without and b) with data augmentation

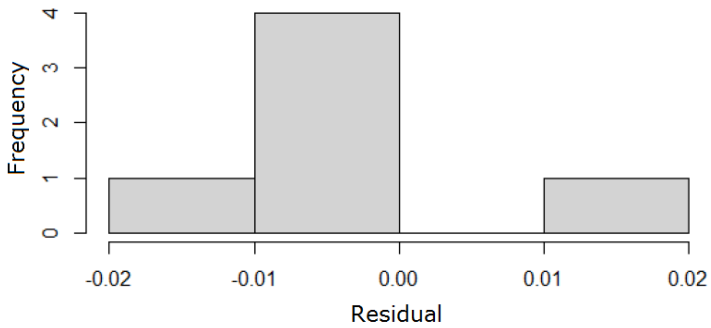


Figure 9. Residual histogram

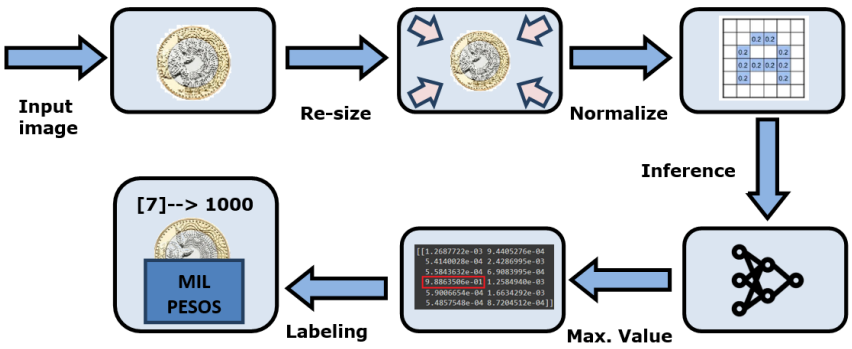


Figure 10. Smartphone process pipeline

Fig. 10 shows CopReader's smartphone process pipeline. Here, the input image is re-sized and then normalized prior to the inference process with MobileNetV2. Afterwards, the maximum value of the recognition inference is selected, associated with a label, and prepared for audio feedback.

5. Results and discussion

CopReader was validated using three quantitative tests: 1) a test with visually non-impaired users to verify its functionalities and measure performance metrics such as precision-recall; 2) a test with visually impaired users to evaluate its usefulness; 3) a test with an improved version of the app, considering the suggestions made by the visually impaired subjects, where new performance metrics were obtained.

The experimental conditions for these tests are described below:

a. The mobile devices used were a Samsung Galaxy A10s with a MediaTek Helio P22, 2 GHz processor and a Moto G5S Plus with a Snapdragon 625 octa-core, 2 GHz processor.

b. The selected number of visually non-impaired users was 10.

c. The selected number of visually impaired was 10, where nine persons were blind and one had low vision.

d. The visually non-impaired users were surveyed at Universidad del Valle, Cali, Colombia.

e. The visually impaired users were surveyed in the Hellen Keller room of the Jorge Garcés Borrero public library, with ethics committee endorsement no. 030-22 of Universidad del Valle.

In the first and second tests, 40 different coins and banknotes from each class were used by 20 different users (ten with no visual disability and ten with a visual impairment). Afterwards, all 20 users were asked to answer a survey with six simple questions. Fig. 11 shows the confusion matrix of this test, where a total precision of 0.92 and a recall of 0.9 were obtained. This figure also shows a slight increase in false negatives, as well as decreasing precision and recall values. However, these results are greater than 0.9, a good result in comparison with the related works reviewed in Table I.

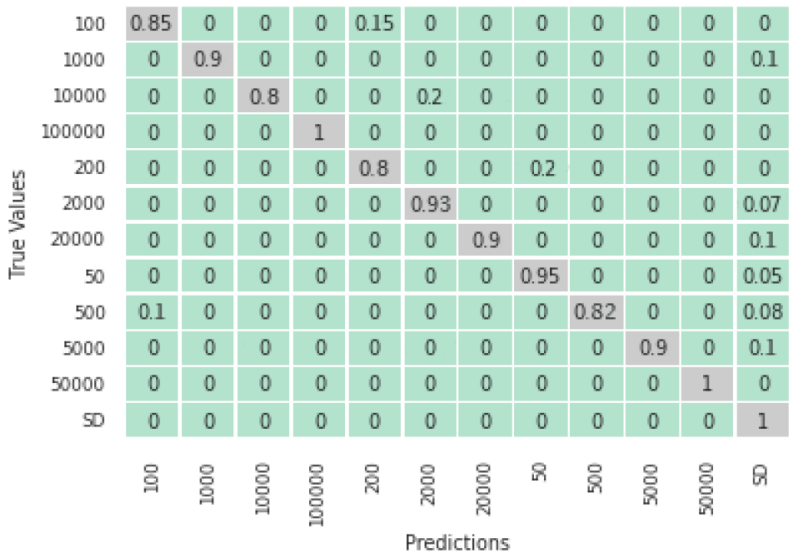


Figure 11. Confusion matrix for test 1 involving visually non-impaired users to verify the app's functionalities

All visually non-impaired users (ten in total) were asked to answer the following six questions:

1. Was the tutorial clear enough to understand how to use the mobile application?
   Answers: totally agree (5), agree (4), disagree (3), and totally disagree (2).
2. What level of satisfaction did you feel in relation to the mobile application's gestures?
   Answers: totally satisfied (5), satisfied (4), dissatisfied (3), and totally dissatisfied (2).
3. What level of satisfaction did you feel in relation to the picture acquisition process?
   Answers: totally satisfied (5), satisfied (4), dissatisfied (3), and totally dissatisfied (2).
4. What level of satisfaction did you feel in relation to the audio feedback?
   Answers: totally satisfied (5), satisfied (4), dissatisfied (3), and totally dissatisfied (2).
5. What level of satisfaction did you feel in relation to CopReader in general?
   Answers: totally satisfied (5), satisfied (4), dissatisfied (3), and totally dissatisfied (2).
6. How likely are you to recommend CopReader?
   Answers: totally likely (5), likely (4), unlikely (3), and totally unlikely (2).

These questions were inspired by the SUS usability scale (29). This scale, however, was not directly used, since it was not designed for visually impaired people. Figs. 12a to 12f show the results of the survey. It can be observed that users agree, totally agree, are satisfied, or are totally satisfied about the different properties of CopReader, such as the tutorial explanation at the start of the mobile app, the simple navigation gestures, the image acquisition process, and the operation of CopReader in general. The users are also likely to recommend this app. However, one person was not satisfied with it, stating that many gestures (more than 2) are required to identify the currency value in front of the smartphone camera.

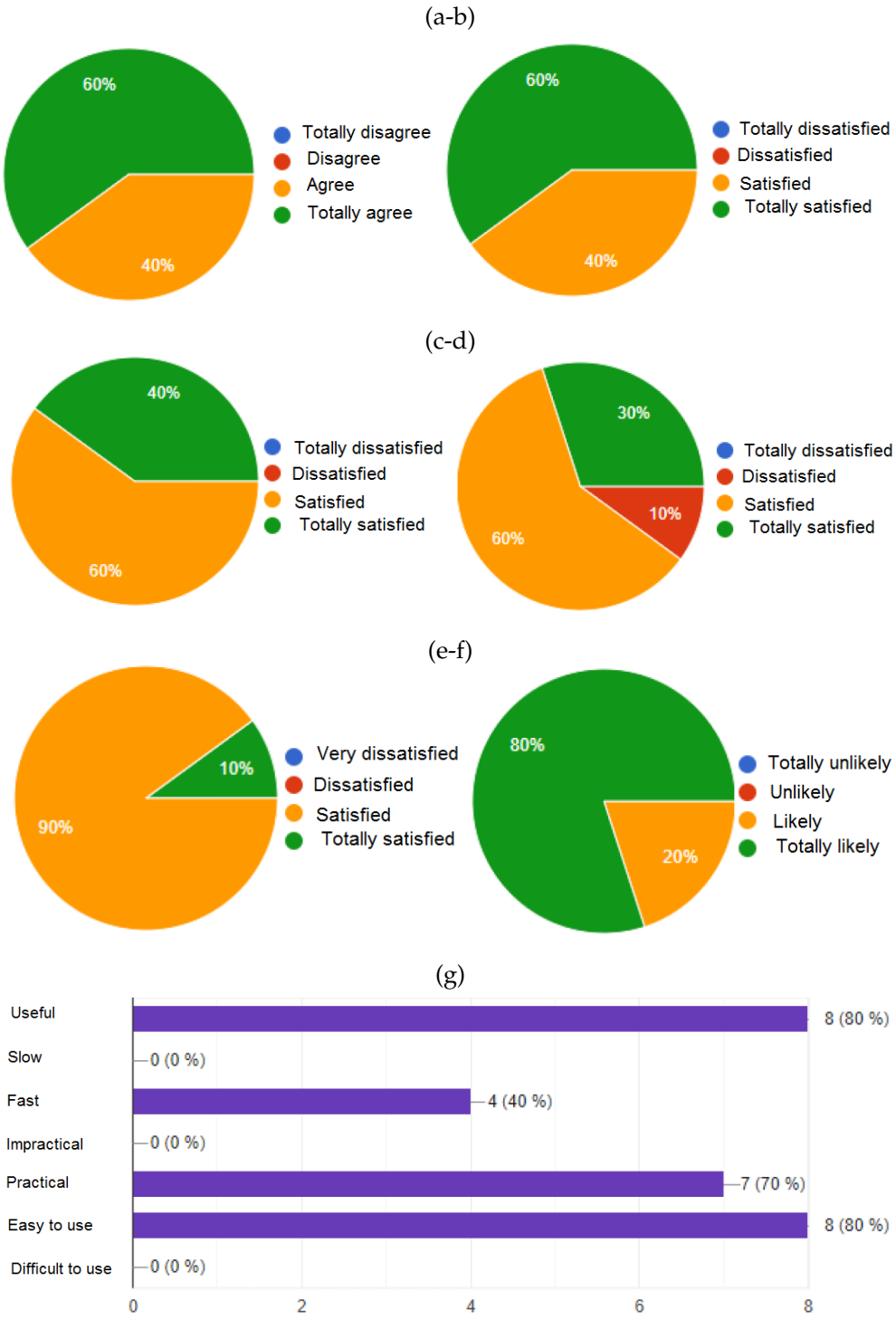


Figure 12. Survey results: a) tutorial explanation, b) use of gestures, c) image capture, d) audio feedback, e) mobile app in general, f) recommendation likelihood, g) mobile app ranking by keywords

In the last question of the survey, the users were asked to select a keyword related to CopReader. These results are shown in Fig. 12g. The users generally thought that CopReader is useful, fast, practical, and easy to use. Negative words, such as *slow, impractical*, or *difficult to use*, were not selected.

In the second test, the visually impaired subjects were divided into nine blind users and one with low vision (Fig. 13a). It is worth noting that these users used the Talkback app in conjunction with CopReader. They were surveyed with the same questions, but an additional one was introduced:

7. How likely are you to use the mobile application again?
   Answers: totally likely (5), likely (4), unlikely (3), and totally unlikely (2).

Figs. 13b to 13h show the results of the survey. The users highlighted that CopReader has a tutorial description, which is important. They felt satisfied and totally satisfied with the audio feedback and the way in which the app works. They are likely to recommend it to their acquaintances, and they will use it again. However, Figs. 12c and 12b show that 20 % and 30 % of users were dissatisfied, stating that it was difficult to locate the currency in front of the camera.

The visually impaired users were also asked for keywords regarding the CopReader mobile application. Fig. 13i shows the results, where users described CopReader as useful, practical, easy to use, and fast. Once again, negative terms such as *slow, impractical*, or *difficult to use* were not selected.

Table III. Statistical survey results

Question	Visually non-impaired users				Visually impaired users			
	Mode	F_{MO}(%)	$\bar{x}$	Cv	Mode	F_{MO}(%)	$\bar{x}$	Cv
1	5	60.0 %	4.6	8.7 %	5	60.0 %	4.6	8.7 %
2	5	60.0 %	4.6	8.7 %	4	50.0 %	4.1	22.0 %
3	5	60.0 %	4.6	8.7 %	4	50.0 %	3.9	28.2 %
4	4	60.0 %	4.2	19.0 %	5	90.0 %	4.8	4.2 %
5	4	90.0 %	4.1	22.0 %	4	90.0 %	4.1	22.0 %
6	5	80.0 %	4.8	4.2 %	5	60.0 %	4.6	8.7 %
7	-	-	-	-	4	80.0 %	4.2	19.0 %

All the surveys were validated using the deviation coefficient shown in Eq. (1) and the corresponding values shown in Table III.

$$Cv = \frac{|5 - \bar{x}|}{\bar{x}} \cdot 100 < 25\,\% \tag{1}$$

In this equation, $\bar{x}$ represents the arithmetic mean, where $\bar{x} > 4$ is desirable. To decide whether the results in Table III are significant, the deviations must be lower than 25 %. The results mode is also expected to correspond to the maximum value (5). Based on Table III, the deviation coefficient (Cv) is lower than 25 % in most results, the mean is greater than 4, and the mode is 5 in 54 % of the results.

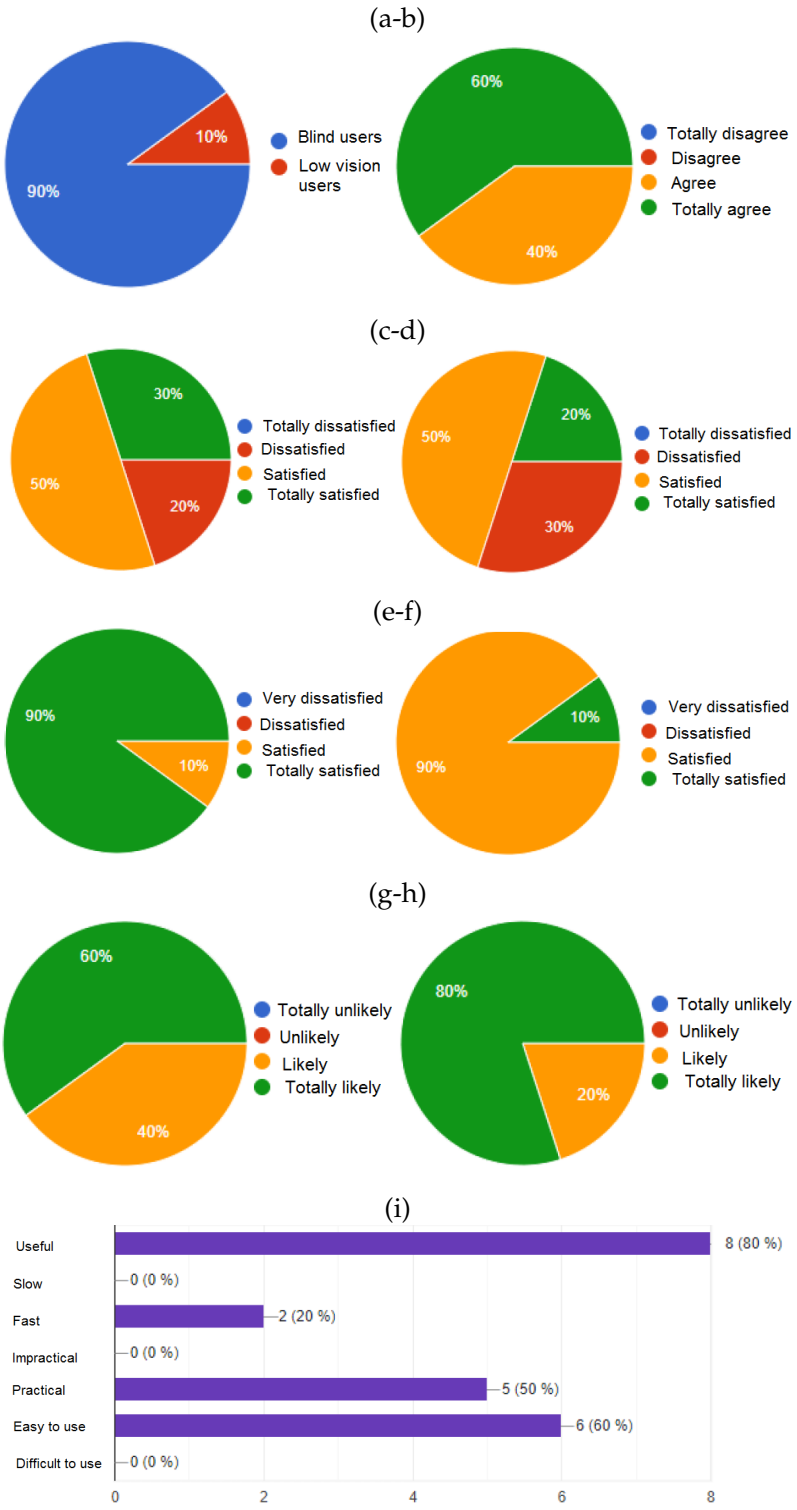


Figure 13. Results of the survey with the visually impaired subjects: a) population surveyed, b) tutorial explanation, c) use of gestures, d) image capture, e) audio feedback, f) mobile app, g) recommendation likelihood, h) likelihood of reuse, i) ranking by keywords

It is worth noting that, for the visually non-impaired users, the results for question 5 are close to the 25 % threshold, as well as questions 2 and 5 for the visually impaired users. For the latter, the deviation coefficient of question 3 is greater than 25 %. These questions are related to the CopReader gestures (question 2), the image acquisition process (question 3), and the overall functioning of the app (question 5). The users expressed that CopReader requires too many gestures (more than two) to know the value of the currency in front of the smartphone camera. Moreover, they suggested that it would be nice for the app to process videos instead of photos, as it could be difficult to locate the coin or banknote in front of the smartphone. This statistical analysis led to the third test, in which a new version of CopReader equipped with additional features was used.

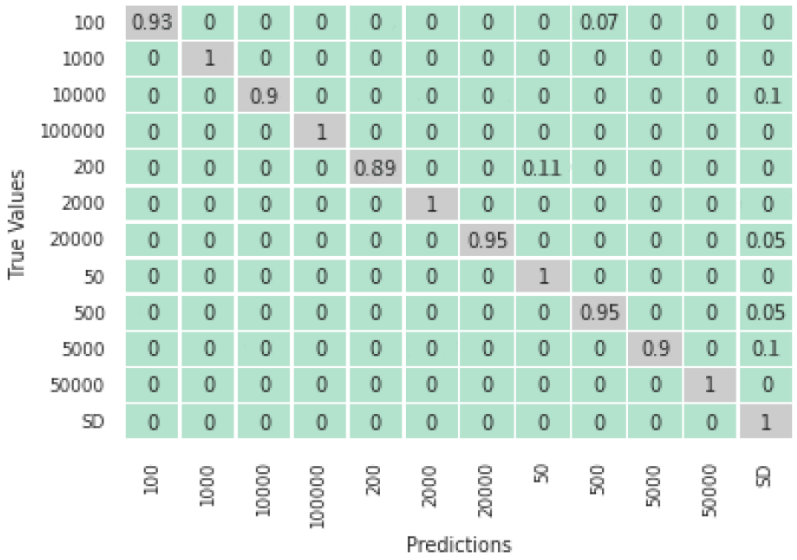


Figure 14. Confusion matrix for the improved version of CopReader

For this test, considering the drawbacks observed in the previous surveys, the improved version of CopReader included video processing, a common suggestion made by the visually impaired users. This version also addressed the issue with the number of gestures needed to achieve currency recognition and obtain audio feedback while solving the difficulty of positioning the currency in front of the smartphone. This new version of CopReader employes the same pipeline shown in Fig. 1. However, instead of processing images, it processes video in YUV_420_888 format at 24 frames per second. The app was tested using 40 different coins and banknotes for each class. The resulting confusion matrix is shown in Fig. 14. In this test, the precision obtained was 0.96, and the recall value was 0.97.

6. Conclusions

This work described the development and testing of CopReader, a mobile application for visually impaired people that provides a tutorial regarding its operation, working in conjunction with Talkback to interact with the users via audio feedback, capturing images and videos and processing them to estimate the currency value of coins and banknotes placed in front of the smartphone camera.

To process the images and videos using neural network models such as YoloV5 and MobileNetV2, a dataset was built which included 8640 images, with 720 images per class. The current coins recognized by CopReader have values of $50, $100, $200, $500, and $1000 pesos, and the banknotes ae worth $2000, $5000, $10 000, $20 000, $50 000, and $100 000 pesos. YoloV5 and MobileNetV2 were trained with this dataset, obtaining precision values greater than 89 % and recall values higher than 92 %.

CopReader was validated using three different tests. First, precision and recall metrics were obtained through interaction with visually non-impaired users. This was also done with visually impaired users, and an improved version of CopReader was tested. In the first and second tests, precision and recall values of 0.92 and 0.90 were obtained, respectively. A survey was conducted with all users, according to which they agree or totally agree, or are satisfied or totally satisfied with the tutorial, the navigation gestures, and the data acquisition process. Moreover, the users are likely to recommend CopReader to others, and they found the app useful, fast, practical, and easy to use.

After these tests, the users suggested that it would be nice for the app to process video instead of photos, as it could be difficult to place coins or banknotes in front of the smartphone. These suggestions helped to improve CopReader. The number of gestures required was reduced to increase the app's acceptance. Then, a third test was performed with an improved version of CopReader, equipped with video processing. This new version processes video in YUV_420_888 format at 24 frames per second and obtains precision and recall values of 0.96 and 0.97, respectively.

The precision results obtained by related works (16–18), range from 65.7 to 82 %. In comparison, the precision and recall values of CopReader are greater than 95 %, making our proposal an interesting option for visually impaired people to increase their autonomy in daily life activities. However, since our app only uses vision-based information, it has difficulties in detecting counterfeit bills.

7. Acknowledgments

This work was conducted within the framework of the research project called *GlobalViewAid: An innovative navigation system for persons with visual impairments* of the Gelbert Foundation (Genève, Switzerland), in collaboration with the EPFL and the University of Genève, identified with the ID No. 21162 of Universidad del Valle, Cali, Colombia.

8. CRediT author statement

Camila Bolaños-Fernández: data curation, software development, and validation.
Bladimir Bacca-Cortes: conceptualization, methodology, resources, supervision, and writing.

References

[1] M. Ramamurthy and V. Lakshminarayanan, "Human vision and perception," in *Handbook of Advanced Lighting Technology*, Cham: Springer International Publishing, 2015, pp. 1-23. ↑2, 3

[2] J. C. Suárez, "Discapacidad visual y ceguera en el adulto: revisión de tema," *Med. U.P.B.*, vol. 30, no. 2, pp. 170-180, 2011. ↑3

[3] J. Alberich, D. Gómez, and A. Ferrer, *Percepción Visual*, 1st ed., Barcelona, Spain: Universidad Oberta de Catalunya, 2014. ↑3

[4] D. Parra, "Los ciegos en el censo del 2018," Instituto Nacional para Ciegos INCI, 2020. [Online]. Available: https://www.inci.gov.co/blog/los-ciegos-en-el-censo-2018 ↑3

[5] M. del R. Yepes, "La intermediación en la inclusión laboral de la población con discapacidad visual," Instituto Nacional para Ciegos INCI2. [Online]. Available: https://www.inci.gov.co/blog/la-intermediacion-en-la-inclusion-laboral-de-la-poblacion-con-discapaci\dad-visual ↑3

[6] Banco de la República, "Marca táctil en los billetes," 2016. [Online]. Available: https://www.banrep.gov.co/es/node/31529 ↑3

[7] Orcam, "OrCam MyEye," 2023. [Online]. Available: https://www.orcam.com/en-us/myeye-store ↑3

[8] M. Doudera, "Cash reader: Bill identifier," mobile app, Google Play, 2024. https://play.google.com/store/apps/details?id=com.martindoudera.cashreader&hl=en&gl=US ↑3

[9] C. Aramendiz U., D. Escorcia G., J. Romero C., K. Torres R., and C. Triana P., "Sistema basado en reconocimiento de objetos para el apoyo a personas con discapacidad visual (¿Que tengo enfrente?)," *Investig. y Desarro. en TIC*, vol. 11, no. 2, pp. 75-82, 2020. https://revistas.unisimon.edu.co/index.php/identic/issue/view/261 ↑4, 5, 7

[10] S. Vaidya, N. Shah, N. Shah, and R. Shankarmani, "Real-time object detection for visually challenged people," in *2020 4th Int. Conf. Intell. Com. Control Syst. (ICICCS)*, 2020, pp. 311-316. https://doi.org/10.1109/ICICCS48265.2020.9121085 ↑4, 5, 7

[11] M. Awad, J. El Haddad, E. Khneisser, T. Mahmoud, E. Yaacoub, and M. Malli, "Intelligent eye: A mobile application for assisting blind people," in *2018 IEEE Middle East North Africa Comm. Conf. (MENACOMM)*, 2018, pp. 1-6. https://doi.org/10.1109/MENACOMM.2018.8371005 ↑4, 5

[12] J. Tamayo, "Sistema de reconocimiento de billetes para personas con discapacidad visual mediante visión artificial," undergraduate thesis, EIA University, 2018. https://repository.eia.edu.co/entities/publication/321a4983-afec-4865-9915-4ede5b26a435 ↑4, 5

[13] J.-Y. Lin, C.-L. Chiang, M.-J. Wu, C.-C. Yao, and M.-C. Chen, "Smart glasses application system for visually impaired people based on deep learning," in *2020 Indo – Taiwan 2nd Int. Conf. Comp. Analytics Net. (Indo-Taiwan ICAN)*, 2020, pp. 202-206. https://doi.org/10.1109/Indo-TaiwanICAN48429.2020.9181366 ↑4, 5, 7

[14] O. Stephen, Y. J. Jang, T. S. Yun, and M. Sain, "Depth-wise based convolutional neural network for street imagery digit number classification," in *2019 IEEE Int. Conf. Comp. Sci. Eng. (CSE) and IEEE Int. Conf. Embedded Ubiq. Comp. (EUC)*, 2019, pp. 133-137. https://doi.org/10.1109/CSE/EUC.2019.00034 ↑4, 5, 7

[15] Joshua, J. Hendryli, and D. E. Herwindiati, "Automatic License Plate Recognition for Parking System using Convolutional Neural Networks," in *2020 Int. Conf. Inf. Manag. Technol. (ICIMTech)*, 2020, pp. 71-74. https://doi.org/10.1109/ICIMTech50083.2020.9211173 ↑4, 5, 7

[16] S. M. M. Roomi and R. B. J. Rajee, "Coin detection and recognition using neural networks," in *2015 Int. Conf. Circuits, Power Comput. Technol. [ICCPCT-2015]*, 2015, pp. 1-6. https://doi.org/10.1109/ICCPCT.2015.7159434 ↑4, 5, 22

[17] N. Capece, U. Erra, and A. V. Ciliberto, "Implementation of a Coin Recognition System for Mobile Devices with Deep Learning," in *2016 12th Int. Conf. Signal-Image Technol. Internet-Based Syst. (SITIS)*, 2016, pp. 186-192. https://doi.org/10.1109/SITIS.2016.37 ↑4, 5, 22

[18] J. Xu, G. Yang, Y. Liu, and J. Zhong, "Coin Recognition Method Based on SIFT Algorithm," in *2017 4th Int. Conf. Inf. Sci. Control Eng. (ICISCE)*, 2017, pp. 229-233. https://doi.org/10.1109/ICISCE.2017.57 ↑4, 5, 22

[19] S. Mittal and S. Mittal, "Indian Banknote Recognition using Convolutional Neural Network," in *2018 3rd Int. Conf. Internet Things: Smart Innov. Usages (IoT-SIU)*, 2018, pp. 1-6. https://doi.org/10.1109/IoT-SIU.2018.8519888 ↑4, 5, 7

[20] A. U. Tajane, J. M. Patil, A. S. Shahane, P. A. Dhulekar, S. T. Gandhe, and G. M. Phade, "Deep Learning Based Indian Currency Coin Recognition," in *2018 Int. Conf. Adv. Commun. Comput. Technol. (ICACCT)*, 2018, pp. 130-134. https://doi.org/10.1109/ICACCT.2018.8529467 ↑4, 5

[21] N. A. J. Sufri, N. A. Rahmad, N. F. Ghazali, N. Shahar, and M. A. As'ari, "Vision Based System for Banknote Recognition Using Different Machine Learning and Deep Learning Approach," in *2019 IEEE 10th Control Syst. Grad. Res. Colloq. (ICSGRC)*, 2019, pp. 5-8. https://doi.org/10.1109/ICSGRC.2019.8837068 ↑4, 5

[22] U. R. Chowdhury, S. Jana, and R. Parekh, "Automated System for Indian Banknote Recognition using Image Processing and Deep Learning," in *2020 Int. Conf. Comput. Sci., Eng. Appl. (ICCSEA)*, 2020, pp. 1-5. ↑4, 5

[23] R. Tasnim, S. T. Pritha, A. Das, and A. Dey, "Bangladeshi Banknote Recognition in Real-time using Convolutional Neural Network for Visually Impaired People," in *2021 2nd Int. Conf. Robotics, Electr. Signal Process. Tech. (ICREST)*, 2021, pp. 388-393. https://doi.org/10.1109/ICREST51555.2021.9331182 ↑4, 5

[24] Google Play, "Android Accessibility Suite," Talkback, 2023. [Online]. Available: https://play.google.com/store/apps/details?id=com.google.android.marvin.talkback&hl=en_US&gl=US ↑5

[25] B. Dwyer, J. Nelson, and J. Solawetz, "Software Roboflow (Version 1.0)." Roboflow Inc., 2022. [Online]. Available: https://roboflow.com/ ↑7

[26] E. Bisong, "Google colaboratory," in *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, Berkeley, CA: Apress, 2019, pp. 59–64. ↑7

[27] Taryana Suryana, "Rational Unified Process (RUP)," *Ration. Unified Process*, vol. 3, no. September, pp. 1-6, 2007. ↑8

[28] D. Trenkler, "A handbook of small data sets: Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J. & Ostrowski, E. (1994): Chapman & Hall, London. xvi + 458 pages, including one diskette with data files (MS-DOS), 40 British Pounds, ISBN 0-412-39920-2," *Comput. Stat. Data Anal.*, vol. 19, no. 1, p. 101, Jan. 1995. ↑13

[29] Wikipedia, "System usability scale," 2024. [Online]. Available: https://en.wikipedia.org/wiki/System_usability_scale ↑17

Camila Bolaños-Fernández

She received a degree in Electronic Engineering in 2022 from Universidad del Valle, Cali, Colombia. Currently, she serves as a software developer at SuperGIROS. Her main interests are web and mobile app development, and embedded systems applications.
Email: mariafdz51@gmail.com

Bladimir Bacca Cortes

He received a degree in Electronic Engineering in 1999 from Universidad del Valle (Cali, Valle del Cauca, Colombia); a Master's degree in Automation from Universidad del Valle in 2004; and a PhD in Technology from Universitat de Girona (Cataluña, Spain) in 2012. He is currently a full professor at Universidad del Valle, in the School of Electrical and Electronics Engineering. He is a member of the research group on Perception and Intelligent Systems. His research interests are mobile robotics, simultaneous localization and mapping systems, and computer vision.
Email: bladimir.bacca@correounivalle.edu.co

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1221

DOI: https://doi.org/10.14483/23448393.21172

Research

Determining the Pulping Conditions and Properties of Unbleached Pulp from Uruguayan *Pinus Taeda*

Determinación de las condiciones de pulpeo y propiedades de pulpa marrón a partir de *Pinus taeda* Uruguayo

Viviana Palombo*, Leonardo Clavijo, and María Noel Cabrera

Grupo de Ingeniería de Procesos Forestales, Instituto de Ingeniería Química, Facultad de Ingeniería, Universidad de la República (Montevideo, Uruguay)

Abstract

Context: In Uruguay, numerous *Pinus* taeda plantations are at final-turn age, but they still do not have commercial destination and are exported as green-wood logs. For the development of this sector, it is necessary to strive towards a comprehensive processing of this resource.
Method: This work focused on analyzing the use of *Pinus* taeda wood available in the country to produce brown kraft pulp with a kappa number of 80, which can be used to make packaging paper. As raw materials, we employed by-products from the mechanical wood-transformation industry (wood chips and thinning wood) and final-turn wood. Pulping tests were carried out, varying the alkali charge and the H-factor while aiming for a kappa number of 80. The viscosity, pulping, and rejection yield were measured in the pulps, and the pH and residual alkali content were evaluated in black liquor.
Results: Based on the results, the best pulping conditions were an active alkali charge of 14 % (Na_2O) and an H-factor of 1260 for sawmill chips and thinning wood, as well as an active alkali charge of 14 % (Na_2O) and an H-factor of 1080 for final-turn wood. The pulp obtained with sawmill chips exhibited the most resistant fibers, and the final-turn wood pulp was the most sensitive to the refining process. Under the aforementioned conditions, the paper properties of laboratory-produced pulp are comparable with those of commercial pulp.
Conclusions: This indicates that it is technologically possible to produce brown kraft pulp from Uruguayan *Pinus taeda* wood, with adequate quality to produce packaging paper.

Keywords: kraft pulp, *Pinus taeda*, paper properties

Article history

Received: 21st/Aug/2023
Modified: 10th/Apr/2024
Accepted: 6th/May/2024

Ing, vol. 29, no. 2, 2024, e21172

* **Correspondence:** vpalombo@fing.edu.uy

Resumen

Contexto: En Uruguay existen numerosas plantaciones de *Pinus taeda* que están en edad de turno final, pero que aún no tienen un destino comercial y se exportan como troncos verdes. Para el desarrollo de este sector, es necesario buscar un procesamiento integral de este recurso.

Método: Este trabajo se enfocó en analizar el uso de la madera de *Pinus taeda* disponible en el país en la producción de pulpa *kraft* marrón con un número *kappa* de 80, que puede utilizarse para fabricar papel de embalaje. Como materia prima, se emplearon subproductos de la industria del aserrado (*chips* y madera de raleo) y madera de turno final. Se realizaron ensayos de pulpeo, variando la carga de álcali y el factor H con el objetivo de alcanzar un número *kappa* de 80. Sobre las pulpas se midieron la viscosidad, el rendimiento del pulpeo y el rechazo, y se evaluaron el pH y el contenido de álcali residual sobre el licor negro.

Resultados: Con base en los resultados, las mejores condiciones de pulpeo fueron una carga de álcali activo del 14 % (Na_2O) y un factor H de 1260 para los *chips* de aserradero y la madera de raleo, así como una carga de álcali activo del 14 % (Na_2O) y un factor H de 1080 para la madera de turno final. La pulpa obtenida con chips de aserradero presentó las fibras más resistentes, y la pulpa de madera de turno final fue la más sensible al proceso de refinado. Bajo las condiciones mencionadas, las propiedades papeleras de la pulpa producida en laboratorio son comparables con las de la pulpa comercial.

Conclusiones: Esto indica que es tecnológicamente posible producir pulpa *kraft* marrón a partir de madera de *Pinus taeda* uruguaya, con la calidad adecuada para producir papel de embalaje.

Palabras clave: pulpa *kraft*, propiedades papeleras, *Pinus taeda*

Table of contents

Page

1. Introduction 3

2. Materials and methods 4
- 2.1. Raw materials 4
- 2.2. Drying and classification 4
- 2.3. Physicochemical characterization of the wood 5
- 2.4. Brown kraft pulp production . . . 5
- 2.5. Black liquor characterization 6
- 2.6. Pulp characterization 6
- 2.7. Disaggregation of commercial pulps 6
- 2.8. Paper properties 6
- 2.9. Calculation of uncertainties and statistical analysis 7

3. Results and discussion 7
- 3.1. Chemical characterization 7
- 3.2. Sawmill chips 7
  - 3.2.1. Kraft pulping 8
  - 3.2.2. Black liquor 9
  - 3.2.3. Thinning and final-turn wood 9
  - 3.2.4. Commercial pulp samples . 10
- 3.3. Paper properties 11
  - 3.3.1. Drainability 11
  - 3.3.2. Air permeability 11
  - 3.3.3. Opacity 12
  - 3.3.4. Tear resistance index 12
  - 3.3.5. Tensile strength index . . . 13
  - 3.3.6. Bursting strength index . . 15

4. Conclusions 15

5. Acknowledgements 15

6. CRediT author statement 16

References 16

1. Introduction

In Uruguay, the forest industry has been bolstered since the approval of the second Forest Law in 1987, increasing the area of forest plantations from 80 000 ha in 1988 to more than 1 000 000 in 2018 (1). Since 2007, the production of bleached kraft pulp from eucalyptus wood has positioned itself as one of the leading markets in the country. However, the processing of pine wood is focused on the production of sawn boards and plywood in small and medium-sized production plants and in a context where wood supply widely exceeds the demand, resulting in the accumulation of this raw material (1, 2). Likewise, the production yields are approximately 50 %, generating 1 000 000 m^3 of chips and cutouts on an annual basis. In addition to the accumulation of final-turn wood, 1 400 000 m^3 of thinning wood are produced annually and are underutilized (3). Currently, this biomass is used for power generation, but, due to changes in the country's energy matrix, including the installation of wind farms and the energy input from eucalyptus kraft pulp production mills, this alternative will not be economically sustainable in the future (4).

On the other hand, the consumption of packaging paper is increasing worldwide, with a compound annual growth rate (CAGR) of 2.3 % in 2021 (5). This is due to increases in international trade and human consumption. The rise in trade implies a greater consumption of packaging paper (6) and, consequently, an increase in the consumption of brown pulp, with the growth forecast being 4.6 % by 2030 (7). Additionally, there is a trend towards replacing non-removable plastic packaging with packaging made from more sustainable materials like paper.

Kraft cooking is the most used pulping method worldwide due to several advantages: the good quality of the pulp produced, the fact that the recovery cycle allows operating with low replacement and on large production scales, and its low environmental impact in modern mills (6,8). This method focuses on dissolving lignin while preserving the highest possible fraction of cellulose and hemicellulose in the pulp. Since the objective of this work is to obtain unbleached pulp, a greater portion of lignin and hemicellulose remains in the pulp, which increases the cooking yield and improves the properties of the paper (8, 9).

The main operating parameters in the kraft process are the alkali charge and the H factor. The first indicates the amount of sodium hydroxide and sodium sulfide that is added to the digester per ton of wood, and the second is a combination of time and the temperature at which the process is carried out, according to Arrhenius-type kinetics, in order to reach a predetermined value regarding the *kappa* number of the pulp at the digester discharge (9). For pine wood used in the production of packaging paper, the alkali charge typically varies between 15 and 20 % (as Na_2O), and the H factor is between 750 and 2040 (10–15).

In kraft cooking, sodium hydroxide and sodium sulfide are used as active chemicals. Sodium hydroxide provides the alkaline medium for lignin solubilization, while sodium sulfide degrades lignin and makes the process more selective, as fewer amounts of hemicelluloses are solubilized. It is worth adding that the wood is chipped prior to being fed to the digester, so the shape and size of the chips

determine the vaporization and impregnation performance, as well as the heat and mass transfer in the digester (8, 16). The feasibility of using wood that was chipped for another purpose (chip subproducts of sawn wood) should be evaluated.

Paper properties are a set of physical, chemical, and physicochemical measurements that allow characterizing a sample of paper or cardboard. To carry out these measurements, the pulp sample is subjected to a refining process, and manual sheets are subsequently formed. The refining extension must be enough to produce fibrillation of the fibers in order to enable sheet formation, but it cannot be excessive to the point of damaging the integrity of the fibers (17).

The degree of refinement is quantified by determining drainability, which measures the interaction between fibers in forming a framework. This measurement is carried out using the Canadian Standard Freeness (CFS) method. In the industrial production of packaging paper, elongation properties are the most relevant, especially tensile strength, which depends on the length of the fibers and the links between them, and the bursting resistance, which evaluates the overall paper strength. Tear resistance is also measured as a strength property. To characterize paper, grammage, thickness, air permeability, and opacity are also usually determined (11, 18).

The aim of this work was to evaluate the quality of the available *Pinus taeda* wood in Uruguay to produce unbleached kraft pulp, determining the pulping conditions for residues from mechanical wood processing, thinning, and final-turn wood. In addition, selected pulping conditions were used to evaluate the paper properties of the selected pulps, aiming to establish whether it is technologically possible to market them as inputs for the production of packaging paper.

2. Materials and methods

2.1. Raw materials

Pinus taeda chips were provided by a sawmill located in the department of Rivera (Uruguay), using 24-year-old clearcut logs with an average diameter at the fine point of 40 cm. We also tested 20-year-old final-turn clearcut *Pinus taeda* logs with apical diameters of 54.3, 40.0, and 38.6 cm from the department of Tacuarembó (Uruguay), as well as 16-year-old wood from commercial thinning, with apical diameters of 28.2, 30.9, and 22.0 cm, also from the department of Rivera. In the last two cases, the material was chipped, and a mixture composed of three specimens and six branches was produced.

To compare the papermaking properties of this wood against available commercial samples, two samples of brown kraft pulp from *Pinus radiata* were used.

2.2. Drying and classification

The chips of the three types of wood were processed separately. First, they were classified in a tray sieve (Regmed 87 model CC-2, Technical Industry of Precisao Ltda., Brazil) according to the SCAN-CM 40:01 standard.

The 'accepted chips' were dried in a Nardi tray dryer with induced draft air circulation (12 Hz) at 40 °C for 72 h, reaching a final humidity of 8 %. They were kept in a dry environment and not exposed to solar radiation until use.

2.3. Physicochemical characterization of the wood

The basic density was determined using the TAPPI T 258 OM-16 standard. The extractives were quantified using water and ethanol as solvents (19). The extractive-free biomass was characterized in terms of carbohydrates, aliphatic acids, and acid-soluble and -insoluble lignin according to the National Renewable Energy Laboratory's procedure (NREL) (20), using an HPLC instrument (Shimadzu Corporation, model LC20, Kyoto, Japan) with a refractive index and UV-visible detectors. With regard to the aforementioned procedure, it is worth noting that only the Aminex HDX-87H column was used, with a flow of 0.6 mL/min, a temperature of 45 °C, and 0.005 M of sulfuric acid as the mobile phase. Soluble lignin levels were measured by spectrophotometry (Shimadzu Corporation, model UVmini-1240) at a wavelength of 240 nm, considering an absorptivity of 12 L/g.cm according to the NREL procedure (20). The ash content was determined by following the TAPPI T 211 standard.

2.4. Brown kraft pulp production

Pulping tests were carried out on the chips obtained from the sawmill. Once the best conditions for these chips were established, tests were carried out in limited condition ranges on the chips from thinning and final-turn wood.

A thermostatic oil bath with eight 316 stainless steel rotating cylinder vessels was used for pulping. Each vessel had a volume of 300 mL (Fibretec model FBI PT 2002B, India). The maximum cooking temperature was fixed at 170 °C, the wood liquor ratio was 4.5:1, and the percentage of sulfidity was 30 %. The alkali charge varied between 14 % and 19.5 % (Na_2O), which was determined by following the ISO AWI 23774 standard, and the H-factor was set at 720, 900, 1080, or 1260. The H-factor is a process control variable that expresses the cooking time and temperature in batch reactors as a single variable. Geometrically, it is the area under the relative rate of reaction vs. the time curve (8).

After cooking, the pulp was washed by hand in cloth bags. In the first instance, it was washed using three spin cycles, and a total 2 L of deionized water were used per cycle. After the last cycle, the cloth bag with the pulp remained submerged in deionized water for 24 h. After that time, a fourth washing cycle was conducted, and the pulp was stored at 4 °C until analysis.

The pulp was refined using a disc refiner (Kumagai Riki Kogyo Co., Ltd., Tokyo, Japan) until the defibrillation point was reached. This value was determined by observation under a microscope. The refiner disks were 30 cm diameter, and the gap between them was 1.12 mm.

The refined pulp was homogenized at 2000 rpm (Kumagai Riki Kogyo Co., Ltd., Tokyo, Japan) and sorted using a separating sieve (Kumagai Riki Kogyo Co., Ltd., Tokyo, Japan) with a 0.15- mm slotted plate for 20 min and under 8 kg/cm^2 of water pressure.

2.5. Black liquor characterization

For the black liquor obtained after cooking, the final pH and residual alkali were determined according to the ISO AWI 23775 standard.

2.6. Pulp characterization

The pulp water content was determined by drying at 105 °C, as per the ISO 638 standard. We also calculated the cooking yield was, defined as the relation between the mass of the dry pulp obtained and the mass of the dry chips fed to the reactor. The percentage of rejected pulp was determined, defined in turn as the mass of dry rejects retained in the separating sieve in relation to the total dry pulp mass.

For the accepted pulp, the *kappa* number (ISO 302) and the viscosity (ISO 5351) were determined. The standard for viscosity determination applies to pulps with low lignin content (less than 5 %); otherwise, dissolution in cupriethylene diamine is not completed. In order to use this standard, the literature proposes dissolution, followed by filtering the material before carrying out the measurement measuring (21).

2.7. Disaggregation of commercial pulps

Commercial samples were received in sheets with 90 % dryness. These samples were separated into 2 cm^2 pieces and left in distilled water for 24 h at room temperature, with a ratio of 24 g of dry pulp per L of water. The hydrated pieces were disintegrated with 2 L distilled water at 40 °C in a disintegrator pulper (Universal Engineering Corporation, model UEC-2008, India) for 15 min.

2.8. Paper properties

The pulp was refined at 6000, 9000, and 12 000 revolutions in a PFI mill (model no. 209, Hamjern, Oslo, Norway; TAPPI Test Method T 248 sp15, 2015). The drainability was determined by measuring the Canadian Standard Freeness (CSF), following the TAPPI T 227 om-17 standard. Ten manual sheets were formed per sample, according to the ISO 5264-2 and ISO 5269-2 standards, and conditioned at 23 °C and 50 % humidity for at least 4 h according to the ISO 536.

The sheets were used to determine grammage, thickness, opacity, and air permeability, as per ISO 536, ISO 534, ISO 2471, and ISO 5636-5, respectively. As for the mechanical properties, tensile strength, tear resistance, and burst strength were determined by following TAPPI T 494 om-01, ISO 1974, and ISO 2758.

The resistance indices were calculated as the relationship between the resistance and the average weight of the sample at a given degree of refinement.

2.9. Calculation of uncertainties and statistical analysis

The statistical error for the structural components of the biomass was calculated as twice the standard deviation. The error for the H-factor corresponds to the contribution during ± 1 min at maximum temperature. The error in alkali loading and residual alkali corresponds to ± 0.1 mL of titrant during titration. The errors of the other parameters were calculated according to the recommendations of their corresponding standard methods. To compare the paper properties, the reported error was taken as the standard deviation. Regarding statistical treatment, an ANOVA was performed while using Tukey's test with a confidence value of 95 %. The tables corresponding to the paper properties present an analysis that considers each sample and the number of revolutions as partitions (lowercase and uppercase, respectively).

3. Results and discussion

3.1. Chemical characterization

85 % of the sawmill by-product chips had a suitable size for use in the kraft pulping process, whereas 11 % were larger and could be re-chipped.

The sawmill chips exhibited a basic density of 466 ± 2 kg/m^3, while the final-turn wood reported 442 ± 10 kg/m^3 and thinning wood showed 432 ± 6 kg/ m^3. These values are within the range reported in the literature for *Pinus taeda* wood (10,22). The thinning and final-turn wood samples exhibited lower basic density values, which may be due to the fact that the samples contained branches. Branches have higher lignin content because of the reaction wood that is generated during their growth. Lignin is a hydrophobic compound, so, for the same mass of dry solid, the volume saturated in water is lower.

Table I shows the chemical composition of the three types of wood, expressed as grams per 100 grams of dry wood. Carbohydrates (cellulose and hemicelluloses) are presented as their constituent monomers. The contents of xylose, mannose, and galactose are expressed as xylose equivalents, as the chromatographic column used does not allow separating these three monomers.

The three types of wood exhibited similar carbohydrate content (glucose, arabinose, xylose, mannose, and galactose). However, the sawmill chips had lower extractives and lignin contents, favoring the pulping process. The higher content of lignin and extractives in the thinning and final-turn wood could be explained by the presence of branches in the composition of the starting mixture, as mentioned above.

3.2. Sawmill chips

As previously mentioned, the selection criteria for pulping were adjusted using sawmill chips and performing kraft pulping under the conditions mentioned in section 2.4. Once the best conditions were found, a narrow search of conditions was performed for the thinning and final-turn wood.

Table I. Chemical characterization of the different Pinus taeda woods used in this work

	Sawmill chip	Thinning wood	Final-turn wood
Ash	0.37±0.02	0.35±0.02	0.33±0.02
Extractives in water	1.56±0.38	2.16±0.10	3.08±0.32
Extractives in ethanol	1.90±0.08	3.39±0.04	2.27±0.08
Insoluble lignin	27.9±1.4	29.7±2.6	28.8±0.8
Soluble lignin	3.1±0.6	4.1±1.2	4.8±2.4
Glucose	47.2±4.2	45.4±1.0	47.2±3.6
Arabinose	2.8±1.5	2.9±0.6	2.3±0.3
Xylose+mannose+ galactose (as xylose)	9.4±0.5	9.8±0.3	9.5±0.3

3.2.1. Kraft pulping

Table II shows the results obtained for the different pulping conditions of the sawmill chips.

Table II. Properties of the pulps obtained after applying different conditions to the sawmill wood chips

Test number	H factor (±17)	Alkali charge (%Na2O) (±0.5)	Yield (%)	Kappa number	Rejects (%)	Viscosity (mL/g)
1	733	13.4	67.1±2.0	117.5±5.8	5.3	209±16
2	721	17.1	54.6±1.4	112.9±8.0	5.3	319±14
3	733	19.5	48.1±1.6	62.5±3.8	2.5	659±12
4	973	16.5	53.0±1.8	101.6±0.8	1.3	413±14
5	909	18.7	53.3±2.0	82.0±3.0	3.0	520±22
6	1,092	12.9	58.1±1.6	115.2±0.8	3.7	348±12
7	1,077	16.0	60.7±1.2	85.1±5.8	1.1	626±14
8	1,082	17.1	56.0±3.4	76.6±0.6	4.6	789±14
9	1,084	19.1	46.2±1.2	61.2±0.2	0.7	957±6
10	1,268	14.0	52.5±1.4	83.7±1.2	0.3	741±8
11	1,266	14.5	51.1±4.2	78.0±1.4	1.2	1,018±6
12	1,26	15.0	52.9±1.4	67.8±1.0	0.5	869±12
13	1,265	16.1	41.4±1.4	52.6±2.4	0.1	1,284±10
14	1,258	16.8	44.7±1.8	45.5±2.4	0.4	1,199±12

The yield values obtained (50-65 % for a *kappa* number greater than 68 and 40-50 % for a *kappa* number of less than 68) are similar to those reported in the literature and are consistent with the

characteristics of the process (10–15). Note that, the higher the degree of delignification, the lower the residual lignin content – thus, a lower *kappa* number can be obtained. This also entails a greater loss of carbohydrates, which further contributes to the decrease in yield. If the variation of the *kappa* number is analyzed within the same H-factor range, the expected trend is obtained. This is because, when the alkaline load decreases, the *kappa* number and the yield increase.

The percentage of rejection for the samples with a *kappa* number between 70 and 80 was within the expected range (0.9-6.9 %). However, the tests with more severe pulping conditions reported a lower rejection percentage, corresponding to a higher degree of delignification (which can be seen as a lower *kappa* number). Nevertheless, there were samples (such as number 10) that deviated from the general trend. It is worth stating that a greater delignification decreases the rigidity of the matrix, which makes the mechanical refining stage more effective, resulting in a lower number of fiber bundles retained in the screening stage.

In the viscosity analysis, the expected trend was not found, since the increment in the alkali load produced greater carbohydrate degradation and a decrease in the degree of polymerization. However, the viscosity value increased. This could be explained by the experimental procedure used; due to the high lignin content, the tested samples did not completely dissolve in the cupriethylene diamine. In the samples with lower lignin contents (*i.e.*, a lower *kappa* number), the dissolved material in the cupriethylene diamine was higher, since cellulose was more available. Consequently, the viscosity value increased. Therefore, only samples with similar *kappa* numbers should be compared. In these cases, the viscosity value increased when the alkali load decreased, and the H-factor increased, which was the expected trend. It can be concluded that, for samples with a similar *kappa* number, those that were produced with a lower alkali load exhibited a higher viscosity (8,21).

Considering the results reported in Table **II**, tests 7 (H-factor: 1077; alkali load: 16.0 % Na_2O) and 10 (H-factor: 1268; alkali load 14.0 % Na_2O) were selected as the best conditions, given that, out of the samples with a kappa number in the range of 80, they reported the highest yields and viscosity values, as well as a lower rejection content.

3.2.2. Black liquor

In all the tests carried out, the final pH of the black liquor showed values greater than 11. Thus, there was no risk of lignin precipitation. The samples that reported values of residual active alkali (RA) within the recommended range (2-4 g/L as Na_2O) were numbers 6, 7, and 11. Considering the RA recommended by the literature and the volume used in each test, the recommended value was 0.48 to 0.96 g as Na_2O. The samples with an RA of less than 0.96 g as Na2O exhibited a higher degree of polymerization, and, consequently, they yielded pulp of better quality (8).

3.2.3. Thinning and final-turn wood

The thinning and final-turn wood samples were analyzed using test 10, which was selected because it demonstrated the lowest active alkali consumption among the two best conditions. Under these

conditions, a *kappa* number of 72.0±0.9 was reached with the final-turn wood. In order to reach the target *kappa* number of 80, the same active alkali charge was used, and the H-factor decreased (H-factor: 1077; alkali load: 14 % as Na_2O). This could be explained by the presence of compression wood in the sample composition (23).

Table III shows the results for the thinning and final-turn wood, as well as for the sawmill chips obtained with the same active alkali load and the highest H-factor.

Table III. Results for the kraft pulping of thinning and final-turn wood when compared to sawmill chip wood under similar pulping conditions

	Alkali charge (%Na2O)	H factor	Kappa number	Yield (%)	Rejection (%)	Viscosity (mL/g)
Final-turn wood	14.0±0.5	1,077±17	79.0±2.0	48.9±1.2	0.3	938±14
Thinning wood	14.0±0.5	1,268±17	80.0±2.4	50.7±2.8	0.9	874±6
Sawmill chip wood	14.0±0.5	1,268±17	83.7±1.2	52.5±1.4	0.3	741±8

The pulping yields, using 14 % active alkali (Na2O), were similar for the three types of wood. Likewise, in all cases, the percentage of rejection was within the expected range, as reported by other authors (24, 25).

On the other hand, the viscosity values of the pulps obtained from thinning and final-turn wood were higher than those from sawmill chips. In the case of the final-turn wood, the high viscosity was likely due to the lower H-factor used, which implies less severe cooking conditions and, therefore, less attacking of the cellulose chains. As for the thinning wood, a higher rejection value was obtained. Thus, more severe conditions would be needed to match the degree of cooking of the sawmill chip wood. The thinning wood's lower degree of cooking was also reflected in a higher viscosity value.

3.2.4. Commercial pulp samples

The commercial samples, called *commercial sample 1* (CS1) and *commercial sample 2* (CS2), exhibited relatively low *kappa* numbers (CS1: 33±3 and CS2: 40±4, respectively). In addition, comparable viscosity values were obtained (973±4 mL/ g for CS1 and 882±4 mL/ g for CS2) with respect to the first two samples in Table III. Laboratory pulps with lower viscosity than the commercial ones were obtained at similar *kappa* numbers. This behavior has been previously observed in our work group, and it could be attributed to the fact that laboratory-scale equipment is considerably smaller than industrial pulping reactors; the former has a higher metal-surface- to-pulp ratio than the latter. Therefore, there was a greater deterioration of the fiber in contact with the metal wall.

3.3. Paper properties

This section outlines the paper properties achieved with the different studied pulps. The pulps were refined at different revolutions, as indicated in Section 2. The different properties measured at each revolution are shown in the figures, as well as the trends observed. The notation used is as follows: AA14 (sawmill chip wood; alkali charge: 14 % as Na_2O; H-factor: 1260), AA16 (sawmill chip wood; alkali charge: 16 % as Na_2O; H-factor: 1080), FTW (final turn wood; alkali charge: 14 % as Na_2O; H-factor: 1080), TW (thinning wood; alkali charge: 14 % as Na_2O; H-factor: 1260), CS1 (commercial sample 1), CS2 (commercial sample 2).

3.3.1. Drainability

The drainability was quantified by measuring the CSF, as shown in Fig. 1.

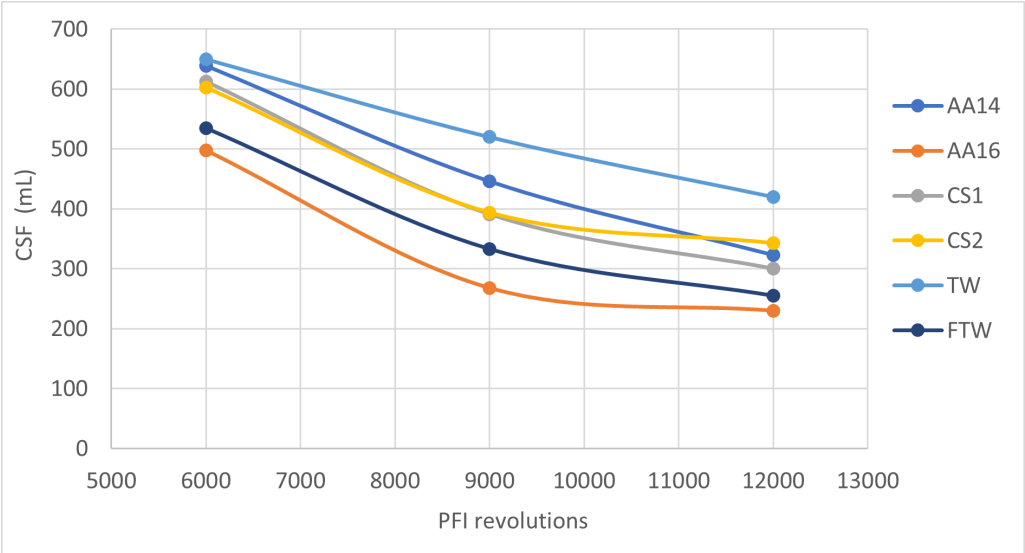


Figure 1. Drainability (mL) according to CSF measurements for four laboratory pulps and two commercial brown pulps

The results showed the expected trend: drainability values decreased when refining increased due to the fibrillation phenomenon (17). Sample AA14 exhibited the highest drainability values for all the revolutions tested. Thus, it was considered to be the fiber with the least damage.

3.3.2. Air permeability

Fig. 2 shows the air permeability values, which increased for all samples as the number of revolutions increased. The reason for this is that the higher the refinement, the greater the fibrillation. Therefore, the sheets produced with more refined pulps were more compact and less porous. When partitioning by sample, a statistically significant difference between the three refining points was found, except for the thinning sample, which reported no significative difference between 9000 and 12 000 revolutions. When partitioning by degree of refinement, it was observed that, at 9000 and 12 000

revolutions, sample AA14 was the least porous, which agreed with the previous result, exhibiting the lowest drainability values. This implies that sample AA14 is the most resistant fiber to mechanical treatment.

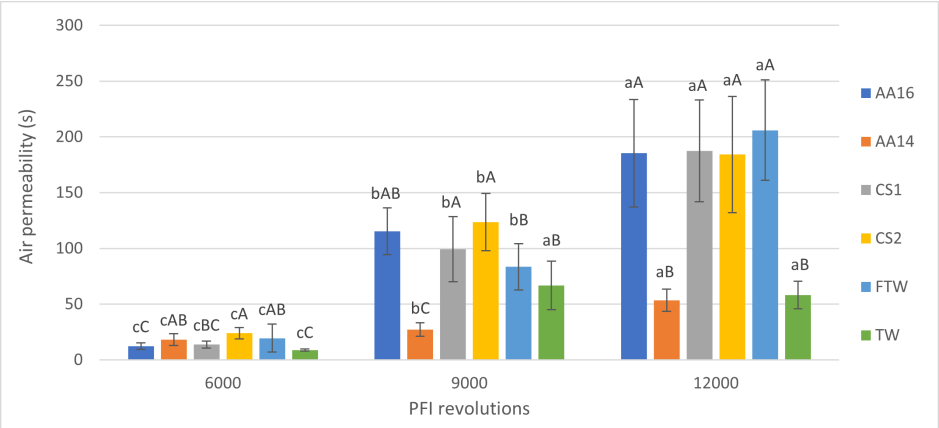


Figure 2. Air permeability (s) values for the sheets evaluated with the different samples at different PFI revolutions. Lowercase letters are used to compare different samples with the same PFI revolutions. Capital letters are used to compare the same pulp at different PFI revolutions. The same letter indicates no statistical differences between the samples.

3.3.3. Opacity

The opacity value (Fig. 3) decreased as the number of revolutions increased. This behavior was expected, as the sheets were more compact when the opacity decreased. This could be explained by the external fibrillation generated during the refinement of the pulp produced with a more closed framework, wherein fines were generated, filling the spaces between the fibers. When partitioning by samples, the statistical analysis showed statistically significant differences for samples AA14 and TW at 6000 revolutions, as well as for samples CS2 and FTW at 12 000 revolutions, for which no trend was evident.

3.3.4. Tear resistance index

According to the values presented in Fig. 4, the tear strength index value decreased as the number of revolutions increased, possibly due to the deterioration of the fiber due to mechanical action. At 6000 revolutions, samples AA14 and AA16 reported a higher tear strength index value. The same trend was maintained for the different degrees of refinement, meaning that these samples had the longest fiber length. This could be explained by the fact that samples AA14 and AA16 corresponded to the older specimens.

Except for sample AA14, no statistically significant difference was observed at 9000 and 12 000 revolutions when partitioning by sample; increasing the refining process did not cause perceptible damage to the analyzed pulps.

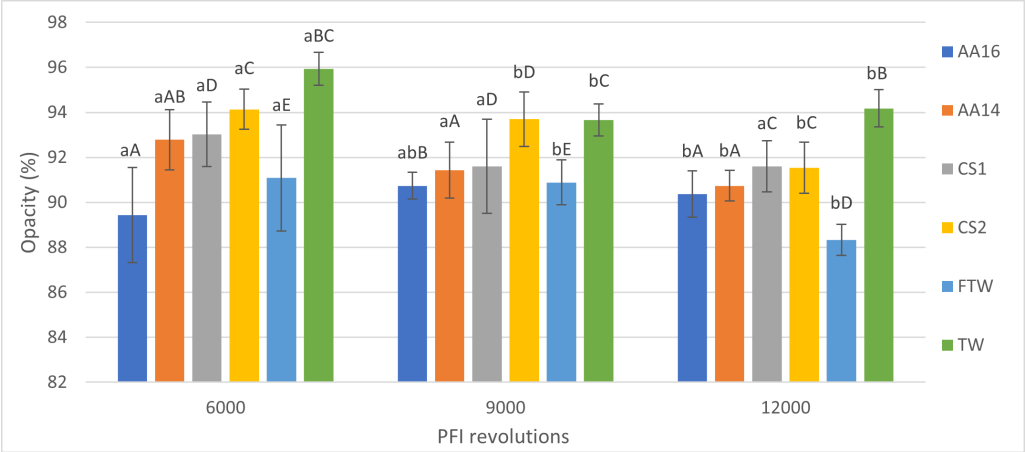


Figure 3. Opacity (%). Lowercase letters are used to compare different samples with the same PFI revolutions. Capital letters are used to compare the same pulp at different PFI revolutions. The same letter indicates no statistical differences between samples.

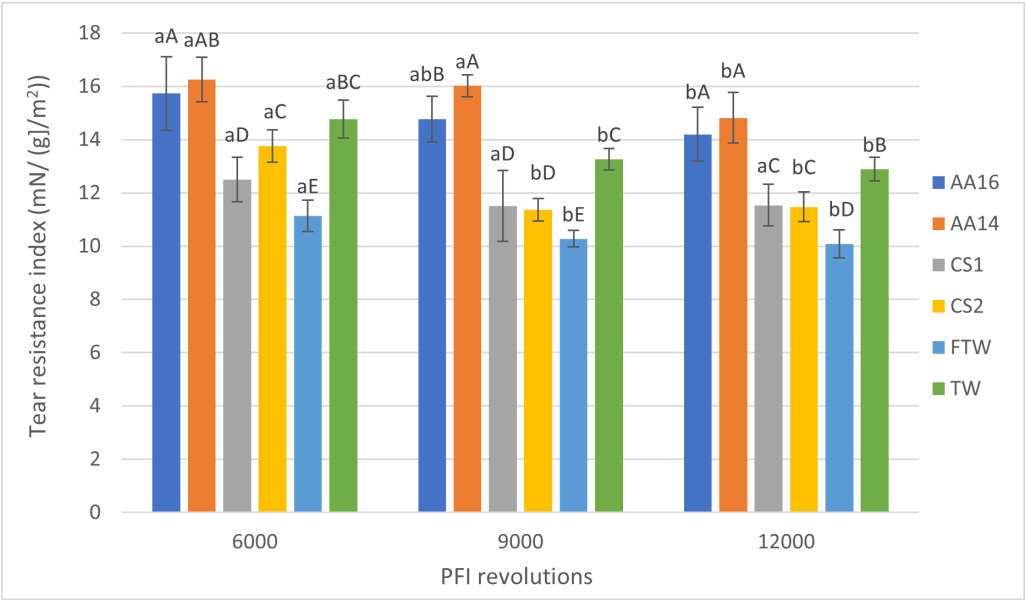


Figure 4. Tear resistance index (mN/[g/m^2]). Lowercase letters are used to compare different samples with the same PFI revolutions. Capital letters are used to compare the same pulp at different PFI revolutions. The same letter indicates no statistical differences between samples.

3.3.5. Tensile strength index

Fig. 5 shows the tensile strength indices for different levels of refinement. These values increased with the number of revolutions because the external fibrillation phenomenon produced during refinement generates a greater number of resistant bonds (26). When partitioning by sample, a similar

behavior was observed in the opacity measures. hus, the trend was not clear. When partitioning by degree of refinement, at 6000 revolutions, there was no statistically significant difference between the samples. However, for higher refinement revolutions, the samples were differentiated, with the final-turn pulp having the best response to the process.

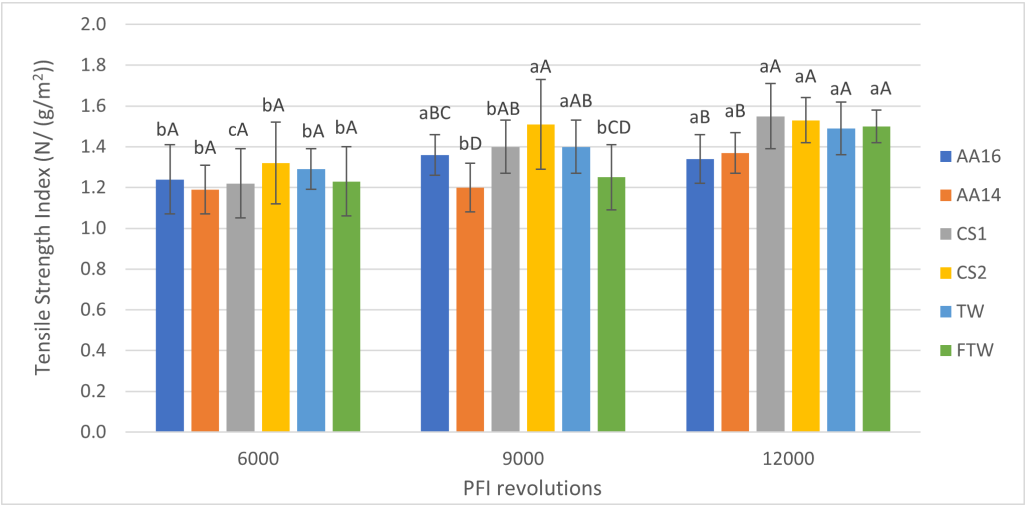


Figure 5. Tensile strength index (N/[g/m^2]). Lowercase letters are used to compare different samples with the same PFI revolutions. Capital letters are used to compare the same pulp at different PFI revolutions. The same letter indicates no statistical differences between samples.

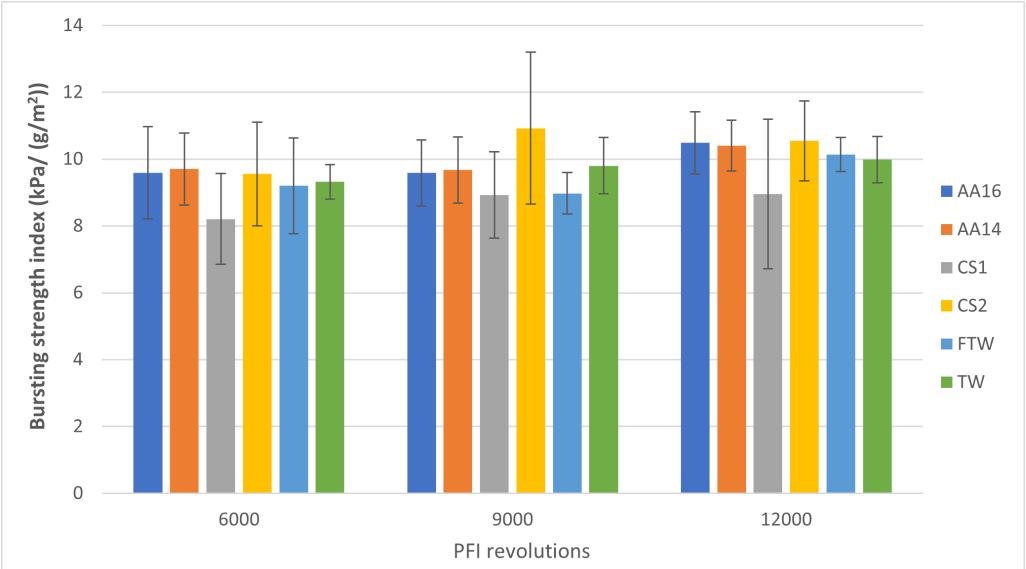


Figure 6. Bursting strength index (kPa/[g/m^2]). Lowercase letters are used to compare different samples with the same PFI revolutions. Capital letters are used to compare the same pulp at different PFI revolutions. The same letter indicates no statistical differences between samples.

3.3.6. Bursting strength index

The values obtained for this index are shown in Fig 6. Although the distribution of errors of the experimental measurements did not allow for statistical analysis, the resistance values obtained with the experimental samples were similar to each other as well as to those obtained from commercial pulps. This measure is similar to that reported by other authors (27).

4. Conclusions

Based on the experimental data, sawmill by-product chips were found to be a viable raw material for use in a kraft pulp mill. Out of the sawmill chips used, 85 % had a suitable size for use as raw, while 11 % were larger and could be re-chipped.

The three types of wood tested exhibited similar chemical compositions and small differences in their basic density and lignin contents, which could be explained by the presence of branches in the composition of the composite mixture.

It was possible to obtain pulp with a *kappa* number of 80 under an active alkali charge of 14 % (Na_2O) and an H-factor of 1260 for sawmill chips and thinning wood, as well as 14 % (Na_2O) and an H-factor of 1080 for final-turn wood. These pulping conditions resembled those reported in the literature, in addition to the yields and the rejection values obtained. However, the viscosity values were lower than expected.

This difference could be explained by the differences between the methods used, as our measurements were outside the standard method conditions.

Sample AA14, which corresponds to sawmill chips pulped with a 14 % acetic alkali charge (Na_2O) and an H-factor of 1080, exhibited the best quality in terms of paper properties, with the highest CSF and tear index values. Likewise, the final-turn sample was shown to be the most sensitive to the refining process, reporting the lowest tear resistance indices and the most compact sheets.

In this vein, it can be stated that the brown kraft pulp produced with the *Pinus taeda* wood available in Uruguay exhibits adequate quality and can be marketed as an input to produce packaging paper.

The results obtained in this research confirm the technological feasibility to produce *Pinus taeda* brown kraft pulp mill. In the future, an economic and environment study could be conducted to determine the feasibility of installing an industrial-size mill in Uruguay.

5. Acknowledgements

The authors would like to thank DANK S.A. and Eng. Carlos Montero for kindly providing the pine chips and logs; the Graduate Academic Commission (Comisión Académica de Posgrado, CAP)

for the postgraduate scholarship of Viviana Palombo (BDMX_2020_1#49478137); and the Technological Laboratory of Uruguay (Laboratorio Tecnológico del Uruguay, LATU) for providing support equipment.

6. CRediT author statement

Viviana Palombo: Conceptualization, data curation, formal analysis, investigation, writing (original draft, review & editing).
Leonardo Clavijo: conceptualization, investigation, writing (review & editing).
Maria Noel Cabrera: conceptualization, investigation, writing (review & editing).

References

[1] M. Boscana and L. Boragno (2018). "Estadísticas Forestales 2018," Montevideo, Uruguay: Dirección General Forestal - Ministerio de Agricultura, Ganadería y Pesca. [Online] Available: http://www.mgap.gub.uy/sites/default/files/dgf_boletin_estadistico_2018_0.pdf ↑3

[2] Uruguay XXI, "Informe del sector forestal en Uruguay," 2016, [Online]. Available: http://www.uruguayxxi.gub.uy/inversiones/wp-content/uploads/sites/3/2014/09/Sector-Forestal-Uruguay-XXI-2014.pdf ↑3

[3] A. Dieste, *Programa de promoción de exportaciones de productor de madera*. Uruguay: Dirección Nacional de Industria, Ministerio de Industria, Energía y Minería, Consejo sectorial Forestal-Maderero, 2012. ↑3

[4] Ministerio de Industria, "Balance energético nacional," 2023. [Online] Available: https://ben.miem.gub.uy/descargas/1balance/1-1-Libro-BEN2022.pdf. ↑3

[5] "Data Bridge - market research," 2022. [Online] Available: https://www.databridgemarketresearch.com/news/global-bleached-kraft-pulp-market ↑3

[6] FAO, "Pulp and paper capacities – Survey | Capacités de la pâte et du papier - Enquête | Capacidades de pulpa y papel - Estudio 2019–2024,"2020. [Online]. Available: https://doi.org/10.4060/cb1212t ↑3

[7] "Molded Fiber Pulp Packaging Market Size & Share by 2033." https://www.futuremarketinsights.com/reports/moulded-fibre-pulp-packaging-market#:~:text=MoldedFiberPulpPackagingMarketSnapshot&text=Theoverallten-yearcompound,ofCOVID-19inducedrestrictions. (accessed Sep. 26, 2023). ↑3

[8] H. Sixta, Handbook of pulp, 1st ed. Germany: Wiley-VCH Verlag GmbH & Co, 2006. https://doi.org/10.1002/9783527619887 ↑3, 4, 5, 9

[9] K. Kaihonen, *et al.*, "Know Pulp," 2022 [Online]. Available: https://www.knowpulp.com ↑3

[10] M. A. Vivian *et al.*, "Wood quality of Pinus taeda and Pinus sylvestris for kraft pulp production," *Sci. For.*, vol. 43, no. 105, pp. 183–191, 2015. ↑3, 7, 9

[11] M. Akgül, Y. Çöpür, and S. Temiz, "A comparison of kraft and kraft-sodium borohydrate brutia pine pulps," *Build. Environ.*, vol. 42, no. 7, pp. 2586–2590, 2007. https://doi.org/10.1016/j.buildenv.2006.07.022 ↑3, 4, 9

[12] S. H. Yoon and A. Van Heiningen, "Kraft pulping and papermaking properties of hot-water pre-extracted *loblolly pine* in an integrated forest products biorefinery," *Tappi J.*, vol. 7, no. 7, pp. 22–27, 2008. ↑3, 9

[13] C. Fuenmayor and S. W. Park, "Evaluación del efecto de la carga alcalina y factor H en el pulpeo kraft de *Pinus Caribaea var*. hondurensis de las plantaciones de CVG-PROFORCA – Venezuela," *Copérnico*, vol. 7, pp. 03–14, 2007. https://www.researchgate.net/publication/303051113 ↑3, 9

[14] F. Huang and A. Ragauskas, "Extraction of hemicellulose from loblolly pine woodchips and subsequent kraft pulping," *Ind. Eng. Chem. Res.*, vol. 52, no. 4, pp. 1743–1749, 2013. https://doi.org/10.1021/ie302242h ↑3, 9

[15] S. D. Mansfield, K. K. Y. Wong, and A. R. Dickson, "Variation in the response of three different Pinus radiata kraft pulps to xylanase treatments," *Wood Fiber Sci.*, vol. 32, no. 1, pp. 105–115, 2000. ↑3, 9

[16] P. Martinez, "Efecto del sobre espesor de las astillas de *Pinus radiata* en el proceso de cocción kraft", Master's dissertation, Universidade Federal de Viçosa, Minas Gerais, Brasil, 2012. [Online]. Available: https://www.locus.ufv.br/bitstream/123456789/5929/1/texto%20completo.pdf ↑4

[17] S. Gharehkhani *et al.*, "Basic effects of pulp refining on fiber properties - A review," *Carbohydr. Polym.*, vol. 115, pp. 785–803, 2015. https://doi.org/10.1016/j.carbpol.2014.08.047 ↑4, 11

[18] S. K. Gulsoy and F. Ozturk, "Kraft pulping properties of european *black pine* cone," *Maderas Cienc. Tecnol.*, vol. 17, no. 4, pp. 875–882, 2016. https://doi.org/10.4067/S0718-221X2015005000076 ↑4

[19] A. Sluiter, R. Ruiz, C. Scarlata, J. Sluiter, and D. Templeton, "Determination of extractives in biomass - NREL/TP-510-42619," National Renewable Energy Laboratory, 2008. [Online]. Available: https://www.nrel.gov/docs/gen/fy08/42619.pdf ↑5

[20] A. Sluiter *et al.*, "Determination of structural carbohydrates and lignin in Biomass - NREL/TP-510-42618," National Renewable Energy Laoratory, 2008, [Online]. Available: http://www.nrel.gov/docs/gen/fy13/42618.pdf ↑5

[21] G. Uçar and M. Balaban, "Accurate determination of the limiting viscosity number of pulps," *Wood Sci. Technol.*, vol. 38, no. 2, pp. 139–148, 2004, https://doi.org/10.1007/s00226-003-0218-0 ↑6, 9

[22] P. A. Rigatto, R. A. Dedecek, and J. L. Monteiro de Matos, "Influence of soil attributes on quality of Pinus taeda wood for cellulose Kraft production," *Viçosa*, vol. 28, no. 2, pp. 267–273, 2004. https://doi.org/10.1590/S0100-67622004000200013 ↑7

[23] B. Hortling, T. Tamminen, and O. Pekkala, "Effects of delignification on residual lignin-carbohydrate complexes in normal pine wood and pine wood enriched in compression wood. 1. Kraft pulping," *Nord. Pulp Pap. Res. J.*, vol. 16, no. 3, pp. 219–224, 2001. https://doi.org/10.3183/npprj-2001-16-03-p219-224 ↑10

[24] A. Ahmed, G. C. Myers, and S. AbuBakr, "Packaging grade kraft pulp from small- diameter softwood," in *TAPPI Pulping/Proc. Product Quality Conf.*, 2000, pp. 1–9. ↑10

[25] A. Hussein, W. Gee, P. Watson, and S. Y. Zhang, "Effect of precommercial thinning on residual sawmill chip kraft pulping and pulp quality in balsam fir," *Wood Fiber Sci.*, vol. 38, no. 1, pp. 179–186, 2006 ↑10

[26] H. R. Motamedian, A. E. Halilovic, and A. Kulachenko, "Mechanisms of strength and stiffness improvement of paper after PFI refining with a focus on the effect of fines," *Cellulose*, vol. 26, no. 6, pp. 4099–4124, 2019. https://doi.org/10.1007/s10570-019-02349-5 ↑13

[27] M. Bäckström, M. C. Kolar, and M. Htun, "Characterisation of fines from unbleached kraft pulps and their impact on sheet properties," *Holzforschung*, vol. 62, no. 5, pp. 546–552, 2008. https://doi.org/10.1515/HF.2008.081 ↑15

Viviana Palombo

is a chemical engineer graduated from Universidad de la República (Montevideo, Uruguay) in 2018. In 2021, she obtained her MSc degree in Chemical Engineering, with a thesis focused on the valorization of Taeda Pinus waste to produce brown kraft pulp for packaging paper. Since 2022, she has been pursuing her Doctorate studies in Chemical Engineering, which focus on studying post-treatment methods for the removal of hemicelluloses from bleached eucalyptus kraft pulp, in order to obtain dissolving pulp. Since 2017, she has been part of the Forest Processes Engineering group attached to the Institute of Chemical Engineering of Universidad de la República, currently holding a position as Assistant.
Email: vpalombo@fing.edu.uy

Leonardo Clavijo

received a degree in Chemical Engineering in 2003 and an MSc in Pulp and Paper Engineering in 2010 from Universidad de la República's Faculty of Engineering (Uruguay). Since 2011, he has been an associate professor at the Institute of Chemical Engineering within the same Faculty, and, since 2016, he has held the position of head of the Forest Processes Engineering group. His main research lines are related to the chemical modification of wood, with an emphasis on pulp production processes, forest biorefineries, and lignin valorization.
Email: lclavijo@fing.edu.uy

María Noel Cabrera

has a BSc in Chemistry and Chemical Engineering, an MSc in Pulp and Paper Engineering, and a PhD in Chemical Engineering. She is an aggregate professor at the Forest Processes Engineering group, Faculty of Engineering, Universidad de la República, Uruguay. Her research focuses on pulp and paper technology, lignocellulosic biorefineries, and biomass valorization.
Email: ncabrera@fing.edu.uy

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1221

DOI: https://doi.org/10.14483/23448393.21846

Research

Community-Based Early Warning System Model for Stream Overflow in Barranquilla

Modelo de sistema de alerta temprana para desbordamiento de arroyos en barranquilla basado en la comunidad

Iván Andrés Felipe Serna-Galeano[1], Ernesto Gómez-Vargas[1], and Julián Rolando Camargo-López[1]*

[1]Universidad Distrital Francisco José de Caldas, Bogotá, Colombia

Abstract

Context: This work aims to design and create an early warning model based on the community as an alternative for the mitigation of the disaster caused by the streams that overflow in Barranquilla (Colombia). This model is based on the contributions in social networks, which are consulted through the API of each social network and filtered according to their location.
Methods: With the information collected is performed cleaning and debugging, and then with natural language processing techniques tokenize vectorize the texts, seeking to operate mathematically to find the vector similarity between processed texts, thus generating a classification between texts associated with stream overflow and texts that are not associated with overflow.
Results: The texts classified as stream overflow are processed again to obtain a location or assign a default one, to consequently georeferenced these data in a map that allows to associate the risk zone and visualize it in a web application, monitoring and reducing the possible damage generated to the population.
Conclusions: To choose the best classifier, 3 classification algorithms were selected (random forest, extra tree, and k-neighbor), which presented better performance and R2 in reference to the data processed in the regressions performed. the three algorithms were trained, and the k-neighbor algorithm was found to be the best.

Keywords: Stream overflow, social network, Machine learning, Natural language processing

Article history

Received: 14th/Feb/2024

Modified: 17th/Mar/2024

Accepted: 18th/Apr/2024

Ing, vol. 29, no. 2, 2024, e21846

* **Correspondence:** jcamargo@udistrital.edu.co

Resumen

Contexto: Este artículo busca crear un modelo de alerta temprana como alternativa para mitigar el desastre provocado por los arroyos en Barranquilla Colombia, este modelo está basado en la comunidad por medio de redes sociales, consultando el api de cada red social, en donde se filtra por área de localización.
Métodos: Con la información recolectada y depurada por medio de técnicas de procesamiento de lenguaje natural se convierten los textos en vectores, con el fin de clasificar en base a la similitud vectorial entre los textos procesados, generando así una clasificación entre los textos asociados al flujo de desbordamiento y los textos que no lo son. asociado con el desbordamiento.
Resultados: Los textos clasificados como desbordamiento de arroyos son nuevamente procesados para obtener una ubicación o asignar una predeterminada, para posteriormente georreferenciar estos datos en un mapa que permite asociar la zona de riesgo y visualizarla en una aplicación web, monitoreando y reduciendo los posibles daños generados a la población.
Conclusiones: Para elegir el mejor clasificador se seleccionaron 3 algoritmos de clasificación random forest, extra tree y k-neighbor), los cuales presentaron mejor rendimiento y R2 en referencia a los datos procesados en las regresiones realizadas. Se entrenaron los tres algoritmos y se descubrió que el algoritmo k-neighbor era el mejor.

Palabras clave: arroyos, redes sociales, aprendizaje automático, procesamiento de lenguaje natural

Table of contents

Page

1. Introduction 2
2. Methodology 3
   - 2.1. Model structure 3
   - 2.2. Data collection 5
   - 2.3. Data cleaning 6
   - 2.4. Classifying the collected data . . . 6
   - 2.5. Information processing 6
     - 2.5.1. Exploratory data analysis and preprocessing 6
     - 2.5.2. Training 7
3. Results 7
4. Discussion 11
5. Conclusions 14
6. Acknowledgements 14
7. CRediT author statement 14

References 14

1. Introduction

Floods are natural risk events produced by excess water from rivers in areas that have been invaded in normally dry conditions (1). The city of Barranquilla (Colombia) presents a severe case of overflowing streams, rivers, and creeks that cross or border the city, causing flooding in the urban area, which brings material damage and even, in some situations, human losses; this problem is caused by various factors such as its location near the tributary of the Magdalena River and the sea (2), as well as

the low topographic slope that is around 5 % (3). In addition, social and sanitary problems associated with garbage and poor planning cause the rainwater system to collapse quickly (4); the latter makes it a pivotal point to deal with emergencies during the rainy season (2). One way to mitigate the adverse effects of these climatic events is to design and create an early warning model that allows monitoring and alerting affected communities to reduce the impact of overflows, as detailed in (5). Currently, there are several methods of warning and tracking. For example, in Barranquilla, an early warning system was created using sensors and updating the information in a web application (4). This type of system has also been added to supply the system with solar energy (5). Currently, there are several warning and monitoring methods. In Barranquilla, an early warning system was created using sensors and updating the information in a web application (4); improvements have also been added to this system, seeking to supply the system with solar energy (5). Also, for the case study in Barranquilla, hydrological and hydraulic models have been developed to predict the areas where the flow will be high and fast to issue early warnings to the community (6).

On the other hand, in Barranquilla, monitoring of atmospheric phenomena has been implemented, making use of radars that generate alerts before an imminent storm or rain (7); as for data collected from social networks, in Japan, a system was created, which obtains information from Twitter and the processing of this information generates Tsunami alerts (8), likewise in the departmental risk plan of the region of Atlántico (Colombia), this risk scenario has been identified, as well as its causes and antecedents (9). Therefore, it is critical to find an additional solution to those proposed so far, where a model with input information from social networks will be advantageous and increasingly necessary, as this type of information stands out for being fast and updated. Additionally, it provides active monitoring that will generate alerts to the community. Given the above, this article will show in the first section the proposed flow for the model with the design of the system model structure, where the interaction between API, the database, and the processing of the information that generates the stream event prediction, in the second section the methodology used, which includes the data cleaning process, a previous analysis in which the models that best fit the behavior of the data will be reviewed. The third section describes the information processing and training of the selected algorithm. Finally, the results of this research process and the discussion and conclusions will be presented.

2. Methodology

2.1. Model structure

For the model design and structure, a flow verification was initially established, where the precipitation percentage was validated through a meteorological API. If the required percentage was met, the proposed model began by collecting information from the API of the social network X (formerly known as *Twitter*), as described in Figure. **1**. The collected data were stored in a specific database and then processed using the trained algorithm selected for the classification. The database was then updated with each classification result (Figure. **2**). In cases where the text contained a location (address), this information was extracted to update the database; otherwise, a default address was assigned. Afterwards, the geocoding process began, yielding coordinates (latitude and longitude) and that were

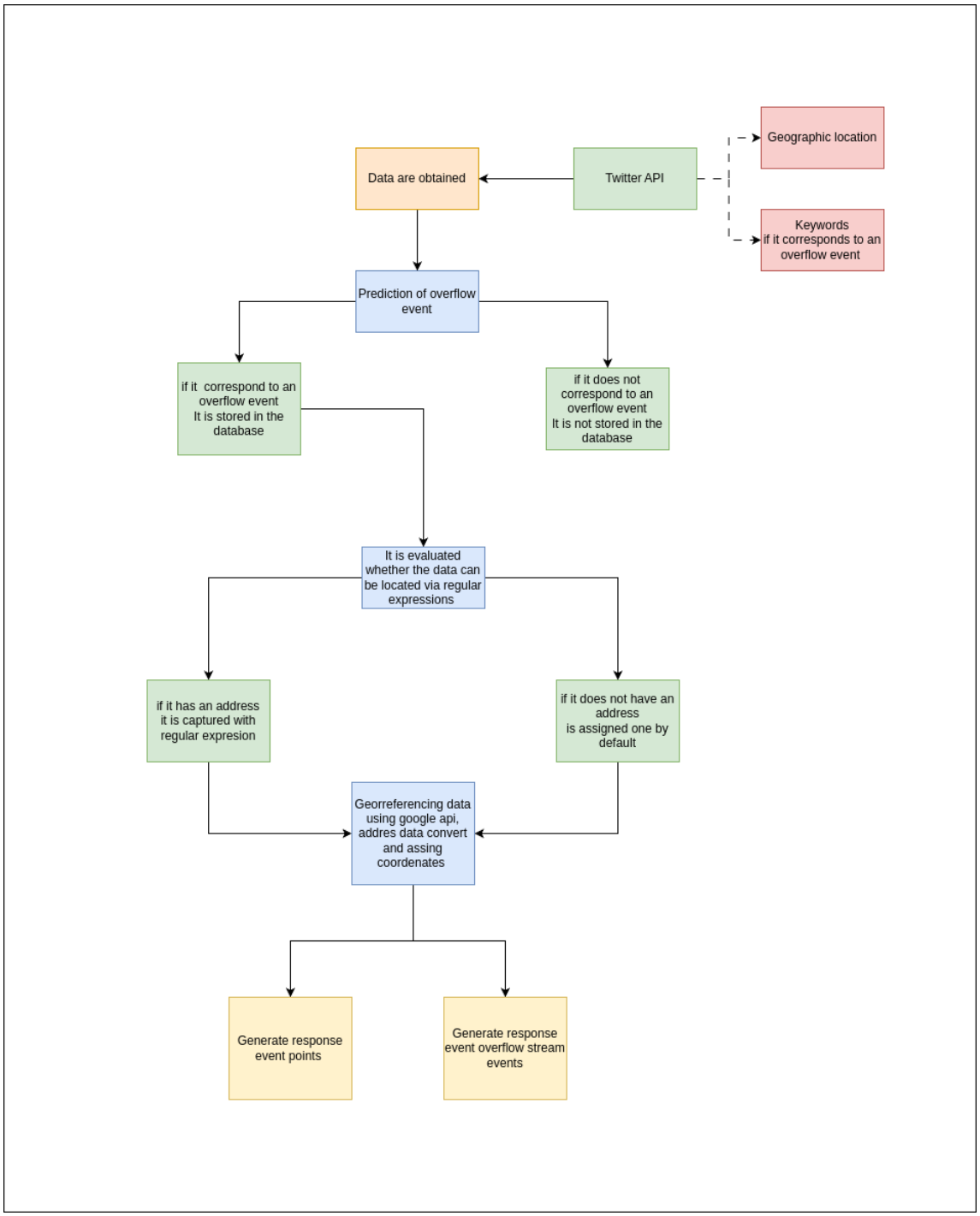


Figure 1. Model flow

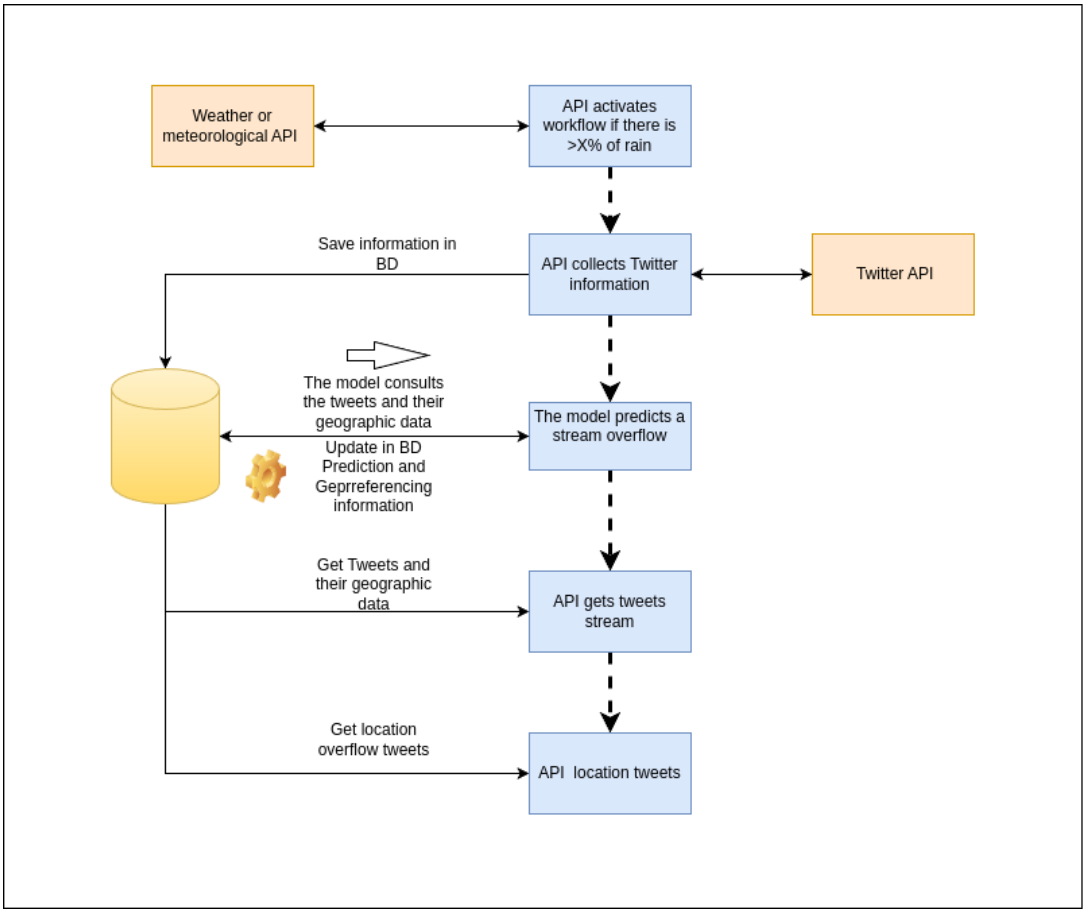


Figure 2. Processing flow

assigned to the data. The next step in the proposed model was to identify whether the recorded events occurred in stream areas. To this effect, it was necessary to make a spatial crossing using a geoprocess, which sought to intercept the location of the events detected as streams vs. the polygons where they circulate. A record of streams was established as roads in Barranquilla (Colombia). The result of the geoprocess was the frequency of events in each stream polygon and the location of each river stream event. This information was displayed on a map (web map) along with the record of the events found in order to generate alerts to the population in risk areas.

2.2. Data collection

The information needed to feed the database was obtained by creating a research account in the X social network (formerly known as *Twitter*) (10). Initially, a search was performed using the X API. The search was filtered by keywords, which were selected based on information from publications such as news and informative posts on the Internet. The words used to perform the search were the following: (Arroyo, Emergencias, Arroyos Barranquilla, Arroyos en la calle, Arroyo en la carrera, creciente, reporte lluvias, calle y carrera). in a time window from 2006 to December 2022. Once the best keywords were

selected, a filter by location was added, taking the city of Barranquilla as the center and adding a 10 km radius to generate the coverage area.

2.3. Data cleaning

Data cleaning is a process that consists of correcting and removing incorrect or duplicated information through computerized methods, as indicated in (11). In this case, the collected data were stored in a SQL database, for which the database engine "Postgresql" was used, and the collected information was structured in database tables. For the data cleaning process, the Python programming language was used, as well as the "Spacy," "Numpy," and "Skylearn" libraries, which contain functionalities that allow connecting the programming language with the database, obtaining each data stored in the database and performing the data cleaning. In the data cleaning process, the methodology mentioned in (12) and (13) was followed where the quality of the data was evaluated. Then the methodology was replicated, where a grammatical analysis of the texts contained in each data was made, finding data with special characters, links, and emojis, after which the data transformation continued, where special characters, emojis, and links were removed using the libraries, leaving the data normalized. Finally, duplicated data found in the database was removed. The debugging performed on the database made it possible to give consistency to the information, thus reducing possible errors generated at the training time.

2.4. Classifying the collected data

Each tweet captured in the search was labeled in three categories, set as a classification criteria variable as follows: event will contain Y (if stream) or N (not stream) data and refers to whether the tweet indicates risk of a stream event, the sarcasm column which will contain Y values (if sarcasm) or N values (not sarcasm) and refers to phrases with a mocking tone that seek to say the opposite, the location column which will contain Y values (if it has an address) or N values (no address) and refers to the geographic location detected in the tweet. For manual classification, each tweet is discriminated against by a user in charge of labeling each data in its corresponding categories. For this classification, the tweet's semantics and context were taken as criteria. This was done manually, considering that it is a supervised process and given that the initial training depends on human interaction, it may contain bias according to the criteria of the person performing the initial classification.

2.5. Information processing

2.5.1. Exploratory data analysis and preprocessing

The exploratory analysis consisted of conducting a series of initial studies and tests necessary to obtain basic approximations to data processing (14) using the information previously stored in the database, which had been previously cleaned and normalized. As in data cleaning, the exploratory analysis was carried out using the Python programming language and the Numpy (15), Scikit learn (16), Spacy (17), and Nltk (18) libraries. Then, we proceeded to establish a count of the number of times that words are repeated (frequency) in the database; this count is essential to determine the words with the highest frequency in the information collected. Another critical step is to filter the collected data by the

one containing the event label equivalent to (”Y”); in this way, the words with the highest frequency during a stream event were found. Additionally, the number per word found helps make histograms and helps in classification to account for data labeled as events, sarcasm, and location.

Additionally, in the exploratory analysis process, the aim was to generate graphs and statistics that would allow the behavior of each variable to be elucidated (14), thus obtaining the regression models with the best behavior in the trend and behavior of the data (19). Given the above, the pre-trained algorithm ”es_core_news_lg” from the Spacy library (17) was used to deliver the new algorithm. It should be noted that the pre-trained algorithm has an accuracy of 100 % in tokenization, 99 % in part of speech, and 98 % in morphological analysis (17). Consequently, we proceeded to vectorize the data from the tokenized data to process vectors and not words. We performed mathematical operations on the vectors and regression, comparing the different statistical models and their behavior. We verified and chose the statistical models with an R2 closer to 1 and a lower Root Mean Squared Error (RMSE).

2.5.2. Training

For training, the information was standardized by changing the values of the labels for each data by ”1” and ”0”, where ”1” corresponds to ”Y” (Yes) and ”0” to ”N” (No). The information obtained was exported in CSV format to train the cleaned data. This copy was made so as not to manipulate the information initially collected and to make it easier to read the corpus file (input data for processing). Taking into account the data structure and its pre-processing shown above, we proceeded to generate a matrix from the corpus; we used the embedding method (vectorization method), which consists of converting the words or sentences (linguistic units) into vectors, for this we used the pre-trained algorithm es_core_news_lg from the Spacy library (18), which would generate for the entire corpus an array of vectors equivalent to a matrix. Given the vectorized data set, this array was split using the Split method, leaving 70 % of the total vectorized data for training and 30 % for testing.

After training the algorithms that showed the best behavior in the regressions performed for 1, 2, and 3 variables (event, sarcasm, and location), the algorithm was trained with the information previously divided, classified, and vectorized, obtaining as a result a trained classification machine learning algorithm, which was saved in PKL format and loaded in Python. The different tests were conducted using 30 % of the divided data, while the training was tested using the Scikit learn library (16).

3. Results

According to the information collected, it was possible to obtain 63259 data. In Table **I** it is possible to see the data found per year and per keyword, where an exponential increase in the amount of information related to streams is observed, it is also possible to find that words such as ”stream” or ”rain” have a greater number of coincidences with respect to the other keywords.

For data cleaning, it was found that the initial search criteria (streams, flood, and overflow) obtained similar results to other search criteria previously applied, causing duplication of information in the

Table I. Tweets by year

Year	Stream on the avenue	Stream on the street	Streams	Flood	Stream emergency	Rain	*La Felicidad* Stream	Stream	Barranquilla Streams	Country Stream	Streams in Barranquilla	40th Avenue
2006	0	0	0	0	0	0	0	0	0	0	0	0
2007	0	2	6	36	0	209	0	999	0	0	0	5
2008	0	1	33	130	0	499	0	500	0	0	0	21
2009	0	43	486	494	0	497	0	492	13	0	7	207
2010	16	73	489	488	7	490	19	496	492	3	364	496
2011	36	493	495	494	14	496	52	495	494	69	492	494
2012	112	489	480	471	26	494	124	484	494	28	496	497
2013	92	496	484	477	55	487	134	490	495	39	495	498
2014	122	484	491	479	35	491	56	478	478	18	494	497
2015	187	480	493	488	62	494	174	450	474	65	475	494
2016	459	496	481	493	117	490	492	491	495	127	491	488
2017	294	475	494	481	68	479	364	477	490	43	490	496
2018	488	457	481	495	208	483	404	435	493	71	495	475
2019	440	480	462	477	493	489	64	488	485	24	486	496
2020	306	493	496	490	252	442	297	496	465	44	465	470
2021	400	500	499	495	101	495	352	496	496	22	496	499
2022	498	497	500	499	45	500	186	499	493	29	494	497
Total	**3450**	**5959**	**6870**	**6987**	**1483**	**7535**	**2718**	**7766**	**6357**	**582**	**6240**	**6630**

Note: The data were retrieved between 2006 and 2022

database, which generated a considerable increase in information cleaning times. Another critical factor in this section was the purification of special characters and stopwords to ensure a text that is possible to analyze and train (corpus). This reduced by about 40 % the volume of information initially obtained, as it went from 63259 tweets collected to 36720 purified. Additionally, it was found that 8600 were related to a stream event in the city of Barranquilla after manual classification. Of these, 1600 had a text that could be associated with an address and used for geo-coding.

After comparing the different models tested, it can be noted that the random forest model has the best data adjustment; after being adjusted, this regression model presents an R-squared of 7899, as shown in Table **II**. Where the statistics mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), coefficient of determination or R-squared (R2), root mean log error (RMSLE), and mean absolute percentage error (MAPE) are found. Additionally, it is possible to note that this adjusted model presents an RMSE of 0.1795, as shown in Table **II**, corroborating that it has the best statistics for the required adjustment.

To To confirm the fit, the same analysis was validated, considering the classification variables such as location and sarcasm, finding similarity of results and the behavior of the data; for this purpose, error prediction graphs were also created, as shown in Figure **3**, where the adjusted error and the predicted error are denoted, which shows no significant variation.

Table II. Data regression and results by statistical model

Algorithm	MAE	MSE	RMSE	R2	RMSLE	MAPE
Random Forest	0,0489	0.0323	0.1795	0.7899	0.1257	0.1306
Extreme Gradient	0.0706	0.0328	0.1807	0.7867	0.1251	0.1951
Extra Trees	0.0476	0.0363	0.1903	0.7639	0.1330	0.1255
Light Gradient	0.0816	0.0384	0.1959	0.7498	0.1363	0.2221
K Neighbors	0.0693	0.0406	0.2015	0.7356	0.1425	0.1743
Decision Tree	0.0450	0.0450	0.2119	0.7070	0.1469	0.1213
Gradient Boosting	0.1197	0.0543	0.2327	0.6474	0.1581	0.3455
Bayesian Ridge	0.1821	0.0793	0.2815	0.4844	0.1955	0.4836
Least Angle	0.1812	0.0794	0.2817	0.4835	0.1953	0.4878
Rige	0.1812	0.0795	0.2817	0.4834	0.1953	0.4979
AdaBoost	0.2378	0.0996	0.3155	0.3516	0.2385	0.3698
Lasso	0.2998	0.1499	0.3871	0.0247	0.2717	0.7895
Elastic Net	0.2998	0.1499	0.3871	0.0247	0.2717	0.7894
Lasso Least Angle	0.2998	0.1499	0.3871	0.0247	0.2717	0.7895
Linear	0.2994	0.1499	0.3871	0.0246	0.2717	0.7894

Note, after performing the corresponding data regressions, the Random Forest statistical model showed a better performance; its R2 was the closest to 1, implying a better fit.

The most frequently found words coincided in most cases with the initial search criteria, which improved the rate of search and information acquisition during the training of the neural processing algorithm (NPL). Figure **4** presents a Word-Cloud showing the words with the most repetitions in the information collected.

Figure **5** presents a word frequency plot of the social network ”X” during a stream event. Ten new algorithms were trained using classification variables set for this model to select the classification algorithm. From the ten tests, it was possible to choose the three best-performing algorithms during the exploratory analysis (KNC, RF, EXTRA). These three algorithms were tested with the classification variables as follows: ”event,” ”event, sarcasm,” ”event, sarcasm, location.” According to the results (Table **III**), it was found that the classification variable sarcasm has great importance in the training of the model since it allows discriminating ambiguities from two-way phrases in several Latin American areas; the use of sarcasm as a communication tool is used every day.

The ”Extra Tree” of the trained algorithms showed a statistically significant performance that performed best in the statistics. The ”Extra Tree” obtained up to 93.66 % accuracy on one classification variable (event). However, the same algorithm obtained 78.99 % of R2 in the regression. When two variables (event, sarcasm) were analyzed, the algorithm obtained an accuracy of 98.66 %, and with three variables (event, sarcasm, location) obtained 98.24 %. The second-best performing algorithm was the K-Neighbors Classifier, which obtained 94 % accuracy with one variable (event), while the regression

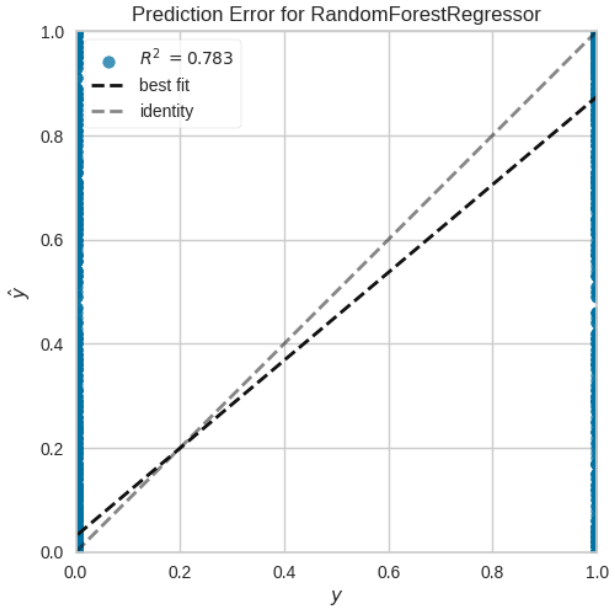


Figure 3. Error Prediction with Random Forest Model Fitting (Note, the fitted model does not differ drastically from the identified error)

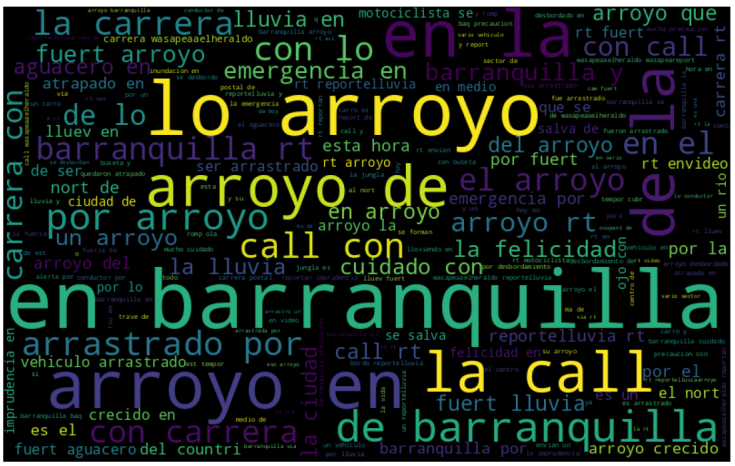


Figure 4. WordCloud of words with the highest frequency by geographic area and by keyword

showed an R2 close to 73.56 %. Under two variables (event, sarcasm), the second algorithm obtained 94.79 % accuracy, while under three variables (event, sarcasm, location), the algorithm obtained 93.91 % accuracy (Table **III**). The two best algorithms were then tested using 400 data. The best result was obtained with the K-Neighbors Classifier, which only produced 88 errors out of 400 data.

Finally, according to the data collected and processed through the K Neighbors Classifier algorithm, the process was completed where, according to the model, it returns the frequency of tweets for each

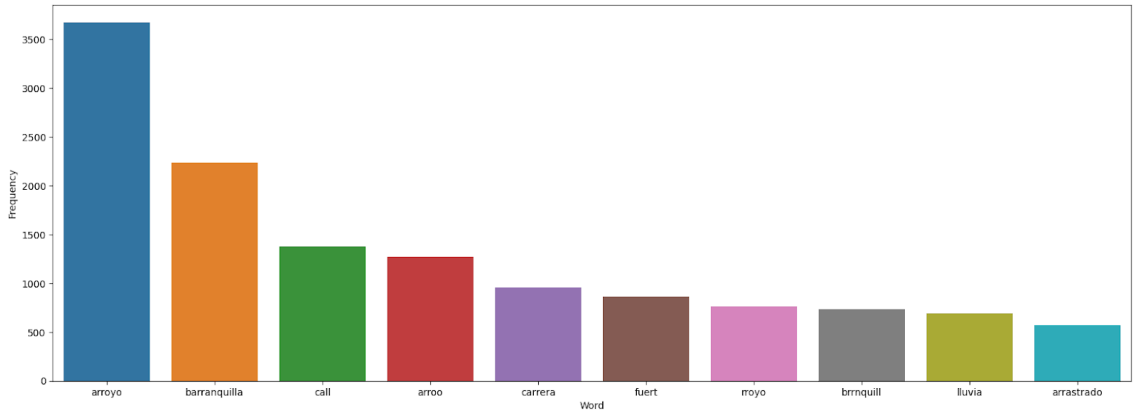


Figure 5. Frequency and intensity of word use during a stream (note: there are words that predominantly appear during a stream event, such as stream, Barranquilla, and rainfall report). The units shown above quantity *vs.* item

Table III. Results of testing in the training algorithms

Algorithm	Fitted precision	Stream	Sarcasm	Location	Fails
KNC	0.9479	X	X	–	88
KNC	0.9391	X	X	X	92
KNC	0.9400	X	–	–	94
RF	0.9822	X	X	–	97
RF	0.9763	X	X	X	104
RF	0.9771	X	–	–	102
EXTRA	0.9366	X	–	–	125
EXTRA	0.9866	X	X	–	105
EXTRA	0.9400	X	X	X	125

stream and the location of each tweet issued that is related to stream overflows, finding that the proposed model works in the best case with an error rate of 0.22, generating an alert to the population according to the location and proximity of these events as shown in Figure 6.

4. Discussion

An early warning model was built from information on stream overflow events in the city of Barranquilla, Colombia, obtained from the social network "X." To improve information filtering, a filter by coverage area was added, which allowed obtaining information only from the required area. This search method immensely helped to delimit the collection of required information and

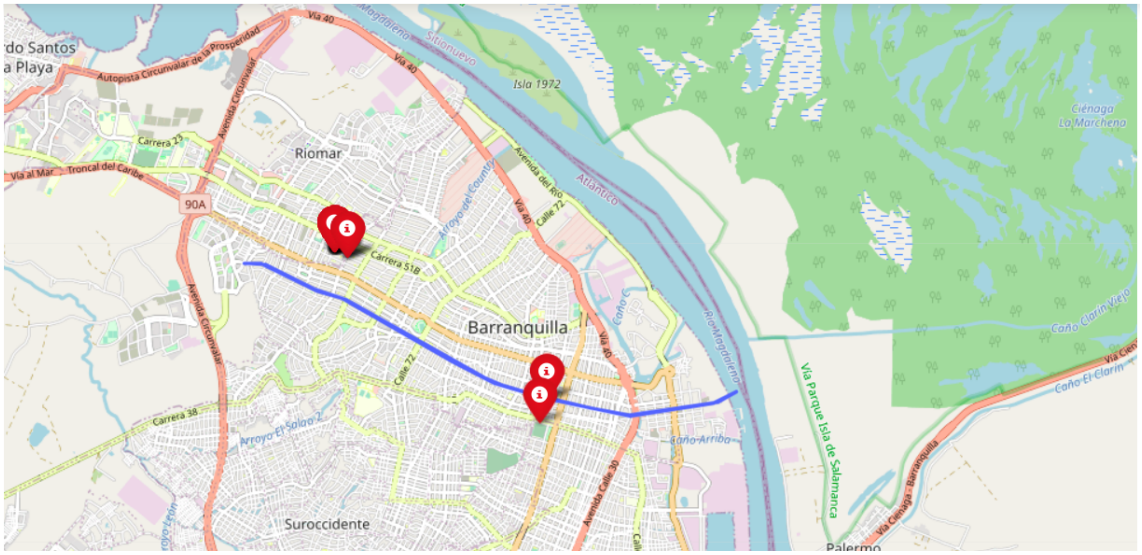


Figure 6. Web map generated by the early warning model for stream overflow events in Barranquilla based on community data. The map shows an example of identified overflow events, as they match the location associated with the analyzed Twitter posts

reduce the data that generated noise in the data cleaning and subsequent training. Therefore, it is possible to consider data cleaning as the totality of operations performed on the data to eliminate anomalies and obtain an accurate and unique representation (20). One of the major drawbacks was the high degree of duplicity in the information, since when searching by area of coverage and keywords, the information contained repeated results; this significantly increased the amount of information in the database. In the same sense, before training the algorithm, it was necessary to label the data obtained; this classification was done manually since the classification criteria can be subjective.

Some of the most representative diagrams obtained during the exploratory analysis were the bar diagram, scatter, and intensity plots. These allowed during the exploratory study to determine the behavior of each variable (14) and from the graphs and statistics obtained to generate initial hypotheses of the behavior of each of the tested algorithms, unlike statistical inference where the hypotheses are born preliminarily and are subsequently challenged and tested from the confirmatory tests (21). Other exploratory analysis techniques also allow for comparing the data distribution by applying a statistical model. Thus, it is possible to determine which statistical model best fits the trend and behavior of the data (19). In this work, it was possible to obtain different regression graphs and models that allowed for the analysis of the behavior of the data and its trend; likewise, these results allowed for the identification of the models with the best behavior. However, this preprocessing required a pre-trained algorithm, which had to be focused on the Spanish language; for this study, an algorithm trained with news in Spanish was used (17).

During the processing and training of the algorithm using tweets, it is critical to highlight the importance of natural language processing, whose objective is the interaction between the computer

and human language (22). It can also be indicated that natural language processing is the ability of machines to process information communicated in human language (23), i.e., natural language processing aims to analyze, understand, and generate language that humans use naturally (24). Thus, text processing seeks to detect and write rules that form structural patterns and then find those patterns in linguistic units such as letters, words, and sentences (22) to be embedded in vectors (24). Additionally, natural language processing has different levels of complexity; each of these represents a type of analysis to be performed to extract specific information; among these levels, it is possible to find morphological, lexical, syntactic, semantic, and pragmatic (25–27).

The morphological level is responsible for analyzing the composition of words, the linguistic level is responsible for establishing the individual meaning of each word, the syntactic level is responsible for investigating the function of each word within the sentence, the semantic level is responsible for establishing the meaning of the sentence from the interaction of words, and the pragmatic level is responsible for analyzing the comprehension of a text (25). The techniques immersed in natural language processing are sentence detection, word segmentation, and discrimination, grammatical tagging, also known as Part of Speech (POS), morphological segmentation, and stopword elimination (28).

Among the best-known applications in the field of natural language processing are content classification and summarization, automatic contextual extraction, sentiment analysis, speech-to-text conversation, and, finally, machine translation (24, 29, 30). The advances in data augmentation and its usefulness in applying different problems (31). This system is also used as a text classifier using various algorithms (32), where the main challenge is to discriminate from ambiguities (33), obtaining. As a result, there is a binary answer in the classification.

Natural language processing methods are usually used to classify texts. These methods allow for identifying the data's tendencies, so they are widely used for classification depending on the probability of similarity to a target variable. Among the most used methods to classify texts are K-Neighbor, Logistic Regression, Naive Bayes, SVM, Random Forest (34), and, in some cases, regressions such as Linear regression (35). Algorithms such as Extra Tree and random forest are among the family of random forest methods. This type of algorithm combines the randomness of the subspace and bagging, which trains multiple decision trees slightly different from the data set (36). On the other hand, the K-Neighbor Algorithm is a statistical model that uses the Euclidean distance to determine which data is closer to the data to be classified. Depending on the count found, the target data will be classified accordingly (37). For the case of this study, the K-Neighbor classification algorithm (KNC) was chosen since it presented the best behavior when analyzing and sorting actual data. During the last years, several research has focused on this field, which has allowed rapid progress in the subject to the point that nowadays, it is possible to find natural language applications in cellphone and virtual assistants or different types of call centers (33, 38, 39).

Finally, this study applied training techniques such as tokenization, which seeks to divide sentences into semantic units; vectorization, which aims to convert the union of semantic units into vectors and, in

some cases, the identification of the part of speech (POS) (40), which seeks to identify the function of the previously tokenized word in the context, i.e., whether it is a verb, adverb, adjective, connector and thus establish the weight within the sentence. This type of preprocessing helps to condition the algorithm's training so that the necessary coincidences and structures are found to predict the classification variables later. Having the information vectorized, mathematical operations are applied to the vectors, seeking as objective the vector similarity between the target variable and the text to be classified. Likewise, although the algorithms with the best statistics were chosen, it was not the ones that showed the best adjustment when evaluating the selected one since overfitting in the models is denoted, which caused unreliable classifications. Given the above, the k-neighbor algorithm was the one that presented the best classification results.

5. Conclusions

The proposed model allowed the collection of information relevant to the case study, which allowed the detection of possible events from the collected data and georeferencing them, whereby by changing the different prioritized filters (location and keywords), it is possible to adapt to a problem unique to the region. On the other hand, the exploratory analysis from the regression allowed us to correctly determine a group of algorithms with better behavior, as well as the words with higher frequency in a stream event. An essential factor to mention was the impact of the column or variable "sarcasm," as it obtained a significant weight in the exploratory analysis and training since, at the time of vectorizing the data, the selected text favors the resolution of ambiguities in the tweets collected and filtered due to the semantics of each text analyzed, which is strongly altered by the popular vocabulary of the study area.

6. Acknowledgements

The authors would like to thank the Twitter team for providing a Twitter academic license, which was useful for data search and algorithm training. These processes involved long time intervals, but the collected information constituted a sufficient and necessary input for executing this research project.

7. CRediT author statement

All authors contributed equally to the research.

References

[1] Riesgo por Inundación - IDIGER». Accedido: 25 de febrero de 2024. [Online]. Available: https://www.idiger.gov.co/rinundacion ↑2

[2] H. D. Van Strahlen Bartel, "Estudio de la problemática de los arroyos urbanos de la cuenca El Rebolo (Barranquilla, Colombia) y propuesta de soluciones," Master's thesis, Universitat Politècnica de València, 2017. [Online]. Available: http://hdl.handle.net/10251/90068 ↑2, 3

[3] H. Ávila, "Perspectiva del manejo del drenaje pluvial frente al cambio climático-caso de estudio: ciudad de Barranquilla, Colombia," *Rev. Ing.*, vol. 36, pp. 54-59, 2012. http://www.scielo.org.co/pdf/ring/n36/n36a11.pdf ↑3

[4] J. A. Sepúlveda Ojeda, "Aplicación web para la visualización de sensores del sistema de alertas tempranas de los arroyos de Barranquilla-Colombia," *Rev. Espacios*, vol. 38, no. 47, p. 17, 2017. http://hdl.handle.net/11323/2024 ↑3

[5] L. J. Pérez Flórez and J. S. Hernández Miranda, "Diseño del modelo económico energético para un sistema de alerta temprana (MEESAT) para los arroyos de Barranquilla," undergraduate thesis, Universidad de la Costa, 2015. http://hdl.handle.net/11323/4899 ↑3

[6] M. Acosta-Coll, F. Ballester-Merelo, and M. Martínez-Peiró, "Early warning system for detection of urban pluvial flooding hazard levels in an ungauged basin," *Nat. Hazards*, vol. 92, pp. 1237-1265, 2018. https://doi.org/10.1007/s11069-018-3249-4 ↑3

[7] M. A. Coll, "Sistemas de alerta temprana (SAT) para la reducción del riesgo de inundaciones súbitas y fenómenos atmosféricos en el área metropolitana de Barranquilla," *Sci. Tech.*, vol. 18, no.o 2, pp. 303-308, 2013. https://revistas.utp.edu.co/index.php/revistaciencia/article/view/8661/5411 ↑3

[8] A. Chatfield and U. Brajawidagda, "Twitter tsunami early warning network: A social network analysis of Twitter information flows," in *23rd Australasian Conf. Info. Syst.*, 2012, pp. 1-10. https://core.ac.uk/download/pdf/301388984.pdf ↑3

[9] Gobernacion del Atlántico, "Plan departamental de gestión del riesgo Atlántico (Colombia)," 2021. [Online]. Available: http://repositorio.gestiondelriesgo.gov.co/handle/20.500.11762/392?locale-attribute=es ↑3

[10] X, "Use Cases, Tutorials, & Documentation," X Developer Platform. https://developer.twitter.com/en (accessed July 8, 2023). ↑5

[11] Y. Valdés Hernández and D. Marmol Lacal, "DBAnalyzer 2.0, sistema para analizar bases de datos libre," undergraduate thesis, Universidad de las Ciencias Informáticas, 2008. https://repositorio.uci.cu/jspui/bitstream/ident/TD_1287_08/1/TD_1287_08.pdf ↑6

[12] E. Estoque Cabrera, L. Baró Galán, and M. E. Escobar Pompa, "Implementación de algoritmos para la limpieza de datos," undergraduate thesis, Universidad de las Ciencias Informáticas, 2015. https://repositorio.uci.cu/jspui/handle/ident/8774 ↑6

[13] I. Zeroual and A. Lakhouaja, "Data science in light of natural language processing: An overview," *Procedia Comput. Sci.*, vol. 127, pp. 82-91, 2018. https://doi.org/10.1016/j.procs.2018.01.101 ↑6

[14] G. D. Buzai and C. A. Baxendale, "Análisis exploratorio de datos espaciales," *Geogr. Sist. Inf. Geográfica*, no. 1, pp. 1-11, 2009. https://ri.unlu.edu.ar/xmlui/bitstream/handle/rediunlu/702/Buzai_An%C3%A1lisis%20Exploratorio%20de%20Datos%20Espaciales.pdf?sequence=1&isAllowed=y ↑6, 7, 12

[15] «NumPy - documentation». [Online]. Available: https://numpy.org/ ↑6

[16] «scikit-learn: machine learning in Python — scikit-learn 1.4.1 documentation». [Online]. Available: https://scikit-learn.org/stable/ ↑6, 7

[17] «Spanish · spaCy Models Documentation», Spanish. [Online]. Available: https://spacy.io/models/es ↑6, 7, 12

[18] «NLTK :: Natural Language Toolkit». Accedido: 25 de febrero de 2024. [En línea]. [Online]. Available: https://www.nltk.org/ ↑6, 7

[19] Z. Jianqiang and G. Xiaolin, "Comparison research on text pre-processing methods on twitter sentiment analysis," *IEEE Access*, vol. 5, pp. 2870-2879, 2017. https://doi.org/10.1109/ACCESS.2017.2672677 ↑7, 12

[20] H. Müller and J. C. Freytag, "Problems, methods, and challenges in comprehensive data cleansing", 2005. [Online]. Available: https://tarjomefa.com/wp-content/uploads/2015/06/3229-English.pdf ↑12

[21] P. Carranza and J. Fuentealba, "Una introducción al análisis exploratorio de datos por medio de Google Analytics," *Yupana Rev. Educ. Matemática UNL*, vol. 7, pp. 53-65, 2013. https://doi.org/10.14409/yu.v1i7.4262 ↑12

[22] M. M. E. Torres and R. Manjarrés-Betancur, "Asistente virtual académico utilizando tecnologías cognitivas de procesamiento de lenguaje natural," *Rev. Politécnica*, vol. 16, no. 31, pp. 85-96, 2020. https://doi.org/10.33571/rpolitec.v16n31a7 ↑13

[23] A. Gelbukh, "Procesamiento de lenguaje natural y sus aplicaciones," *Komputer Sapiens*, vol. 1, pp. 6-11, 2010. https://www.gelbukh.com/CV/Publications/2010/Procesamiento%20de%20lenguaje%20natural%20y%20sus%20aplicaciones.pdf ↑13

[24] L. Deng, "Deep learning: from speech recognition to language and multimodal processing," *APSIPA Trans. Signal Inf. Process.*, vol. 5, no. 1, Jan. 2016. https://doi.org/10.1017/ATSIP.2015.22 ↑13

[25] F. Ramos and J. Vélez, "Integración de técnicas de procesamiento de lenguaje natural a través de servicios web," undergraduate thesis, Universidad Nacional del Centro de la Provincia de Buenos Aires, 2016. [Online]. Available: https://www.ridaa.unicen.edu.ar/handle/123456789/644 ↑13

[26] P. Johri, S. K. Khatri, A. T. Al-Taani, M. Sabharwal, S. Suvanov, and A. Kumar, "Natural language processing: History, evolution, application, and future work," in *Proc. 3rd Int. Conf. Computing Informatics Networks: ICCIN 2020*, 2021, pp. 365-375. http://dx.doi.org/10.1007/978-981-15-9712-1_31 ↑13

[27] M. Maldonado, D. Alulema, D. Morocho, and M. Proano, "System for monitoring natural disasters using natural language processing in the social network Twitter," in *2016 IEEE Int. Carnahan Conf. Sec. Tech. (ICCST)*, 2016, pp. 1-6. https://doi.org/10.1109/CCST.2016.7815686 ↑13

[28] D. Moreira *et al.*, "Análisis del estado actual de procesamiento de lenguaje natural," *Rev. Ibérica Sist. Tecnol. Informação*, no. E42, pp. 126-136, 2021. https://dialnet.unirioja.es/servlet/articulo?codigo=8624557 ↑13

[29] A. Gutiérrez Domínguez, "Aplicación de técnicas de procesamiento de lenguaje natural (NLP) en Twitter para la evaluación de políticas agrarias y del medio rural," Master's thesis, 2022. [Online]. Available: http://hdl.handle.net/10251/186767 ↑13

[30] Z. Zong and C. Hong, "On application of natural language processing in machine translation," in *2018 3rd Int. Conf. Mech. Control Comp. Eng. (ICMCCE)*, Sep. 2018, pp. 506-510. https://doi.org/10.1109/ICMCCE.2018.00112 ↑13

[31] B. Li, Y. Hou, and W. Che, "Data augmentation approaches in natural language processing: A survey," *AI Open*, vol. 3, pp. 71-90, Jan. 2022, https://doi.org/10.1016/j.aiopen.2022.03.001 ↑13

[32] M. B. Hernández and J. M. Gómez, "Aplicaciones de procesamiento de lenguaje natural," *Rev. Politécnica*, vol. 32, 2013. https://revistapolitecnica.epn.edu.ec/ojs2/index.php/revista_politecnica2/article/view/32 ↑13

[33] A. Yadav, A. Patel, and M. Shah, "A comprehensive review on resolving ambiguities in natural language processing," *AI Open*, vol. 2, pp. 85-92, Jan. 2021. https://doi.org/10.1016/j.aiopen.2021.05.001 ↑13

[34] A. A. Turdjai y K. Mutijarsa, «Simulation of marketplace customer satisfaction analysis based on machine learning algorithms», en *2016 International Seminar on Application for Technology of Information and Communication (ISemantic)*, ago. 2016, pp. 157-162. https://doi.org/10.1109/ISEMANTIC.2016.7873830 ↑13

[35] Y. Takefuji and K. Shoji, "Effectiveness of ensemble machine learning over the conventional multivariable linear regression models". [Online]. Available: http://202.240.109.17/publications/pdf/reg_vs_ml.pdf ↑13

[36] Y. Al Amrani, M. Lazaar, and K. E. El Kadiri, "Random forest and support vector machine based hybrid approach to sentiment analysis," *Procedia Comput. Sci.*, vol. 127, pp. 511-520, 2018. https://doi.org/10.1016/j.procs.2018.01.150 ↑13

[37] D. S. Osorio, J. J. M. Escobar, L. Chanona-Hernández, G. Sidorov, and C. J. Núñez-Prado, "Clasificación bi-clase de canciones infantiles aplicando inteligencia artificial y procesamiento de lenguaje natural," *Res. Comput. Sci.*, vol. 151, no. 5, pp. 31-38, 2022. ISSN 1870-4069 ↑13

[38] S. Quarteroni, «Natural language processing for industry: ELCA's experience», *Inform.-Spektrum*, vol. 41, n.o 2, pp. 105-112, 2018. GF https://doi.org/10.1007/s00287-018-1094-1 ↑13

[39] N. Kaur, V. Pushe, and R. Kaur, "Natural language processing interface for synonym," *Int. J. Comput. Sci. Mob. Comput.*, vol. 3, n.o 7, pp. 638-642, 2014. ISSN 2320–088X ↑13

[40] E. Kouloumpis, T. Wilson, and J. Moore, "Twitter sentiment analysis: The good the bad and the omg!," in *Proc. Int. AAAI Conf. Web Soc. Media*, 2011, pp. 538-541. https://doi.org/10.1609/icwsm.v5i1.14185 ↑14

Iván Andrés Felipe Serna-Galeano

Cadastral and geodetic engineer from Universidad Distrital Francisco José de Caldas and Master's student in Information and Communications Sciences at the Department of Engineering of Universidad Distrital Francisco José de Caldas in Bogotá, Colombia.
Email: iasernag@udistrital.edu.co

Ernesto Gómez-Vargas

Full professor at the Department of Engineering of Universidad Distrital Francisco José de Caldas in Bogotá, Colombia. Electronics engineer, specialist in Mobile Telecommunications, and master in Teleinformatics from Universidad Distrital Francisco José de Caldas; PhD in Engineering from Pontificia Universidad Javeriana, Bogotá.
Email: egomez@udistrital.edu.co

Julián Rolando Camargo-López

Full professor at the Department of Engineering of Universidad Distrital Francisco José de Caldas in Bogotá, Colombia. Electronics engineer from Universidad Distrital Francisco José de Caldas, specialist in Design and Construction of Telematic Solutions from Universidad Autónoma de Colombia, and master in Information and Communications Sciences from Universidad Distrital Francisco José de Caldas.
Email: jcamargo@udistrital.edu.co

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1221

DOI: https://doi.org/10.14483/23448393.21108

Research

Methodology for the Selection of Risk Response Actions while Considering Corporate Objectives in the Metalworking Industry

Metodología para la selección de acciones de respuesta a riesgos considerando los objetivos estratégicos en la industria metalmecánica

Álvaro Julio Cuadros-López[1], Alexander Bustos-Useche[1], and Leonardo Bustos-Useche[1]*

[1]Universidad del Valle, Cali, Valle del Cauca, Colombia

Abstract

Context: Projects in metalworking companies are affected by risk. Proper risk management depends on the responses provided to improve the project plan. However, multiple potential actions may result in constraints due to multiple factors. The purpose of this article is to propose a hybrid approach to solve the problem of selecting risk response actions while considering strategic objectives, fuzzy logic, and simulation.

Method: First, 334 risks were identified through a literature review and a discussion with experts. These were then filtered, resulting in 70 operational risks. Subsequently, the ten critical risks were prioritized using the risk matrix. Then, using Monte Carlo simulation and correlation analysis, the activities most affected by the risks were identified. Finally, potential response actions were designed for each case, and fuzzy logic and quality function deployment were applied to evaluate them.

Results: The selected responses were framed within the strategic objectives, *i.e.*, customer satisfaction, business profitability, and implementation of new technologies. This, while considering some corporate attributes that the actions had to meet finishing the project on time, having low costs, and meeting the scope. The selected actions had a better profile than others seeking to minimize time or costs.

Conclusions: EPCC projects are complex and often suffer from gaps in scope, time, and cost. Risk analysis and the selection of responses in the planning phase help to improve performance. This study developed a risk response plan for a project executed in Brazil. Risks were identified, classified, and mitigated using simulations, resulting in an 11-day reduction in the project's estimated duration.

Keywords: fuzzy logic, Monte Carlo simulation, project risk management, risk response actions

Article history

Received: 8th/Aug/2023

Modified: 10th/Apr/2024

Accepted: 14th/May/2024

Ing, vol. 29, no. 2, 2024, e21108

* **Correspondence:** johandmunoz@utp.edu.co

Resumen

Contexto: Los proyectos en empresas metalmecánicas se ven afectados por el riesgo. Una correcta gestión de riesgos depende de las respuestas que se brinden para mejorar el plan del proyecto. Sin embargo, múltiples acciones potenciales pueden resultar en restricciones por múltiples factores. El propósito de este artículo es proponer un enfoque híbrido para resolver el problema de seleccionar acciones de respuesta a riesgos considerando objetivos estratégicos, lógica difusa y simulación.
Método: Primero, se identificaron 334 riesgos mediante una revisión de la literatura y una discusión con expertos. Estos fueron filtrados, lo que resultó en 70 riesgos operacionales. Posteriormente, se priorizaron los 10 riesgos críticos utilizando la matriz de riesgos. Luego, mediante simulación Monte Carlo y análisis de correlación, se identificaron las actividades más afectadas por los riesgos. Finalmente, se diseñaron potenciales acciones de respuesta para cada caso, y se aplicó lógica difusa y despliegue de funciones de calidad para evaluarla
Resultados: Las respuestas seleccionadas se enmarcaron en los objetivos estratégicos, i.e., satisfacción del cliente, rentabilidad del negocio, e implementación de nuevas tecnologías. Esto, teniendo en cuenta algunos atributos corporativos que las acciones debían cumplir: finalizar a tiempo el proyecto, tener costos bajos y cumplir con el alcance. Las acciones seleccionadas tuvieron un mejor perfil que otras opciones que buscaban minimizar tiempo o costos.
Conclusiones: Los proyectos EPCC son complejos y a menudo sufren de desfases en alcance, tiempo y costo. El análisis de los riesgos y la selección de las respuestas en la fase de planificación ayudan a un mejor desempeño. Este estudio desarrolló un plan de respuesta a riesgos para un proyecto desarrollado en Brasil. Los riesgos fueron identificados, clasificados y mitigados mediante simulaciones, lo que resultó en una reducción de 11 días en la duración estimada del proyecto.

Palabras clave: lógica difusa, simulación Monte Carlo, gestión de riesgos de proyectos, acciones de respuesta a riesgos.

Table of contents

Page

1. **Introduction** 3
2. **Literature review** 4
3. **Methodology** 6
   - 3.1. Identification of operational risks . 8
   - 3.2. Prioritization of operational risks . 9
   - 3.3. Assessment of risks impact 9
   - 3.4. Analysis of response actions 11
4. **Results – a case study in a construction project** 13
   - 4.1. Identification of operational risks . 14
   - 4.2. Prioritization of operational risks . 14
   - 4.3. Assessment of risks impact 14
   - 4.4. Framework for the proposed model 15
5. **Discussion** 19
   - 5.1. Theoretical implications 19
   - 5.2. Practical implications 20
   - 5.3. Limitations and discussion of future research 20
6. **Conclusions** 20
7. **CRediT author statement** 21

References 21

1. Introduction

Industrial projects involve several specialties such as construction, metalworking, electricity, hydraulics, environment, and safety. The metalworking discipline is responsible for transforming steel into goods, building machines and structures using raw metal materials. This process includes cutting, burning, welding, machining, forming, and assembly tasks. The products range from laminates, pipes, metal structures, and wires to industrial machinery, such as elevators and boilers.

The lifecycle of these projects follows the generic EPCC process (engineering, procurement, construction, and commissioning). The engineering component includes basic and detailed engineering as well as planning; procurement includes logistics, transportation, receipts, purchases, and invoices; construction includes mechanical, electrical, and civil installations; and commissioning includes testing and delivery, after-sales service, and modernization (1). By nature, these projects are quite complex and plagued by uncertainty (2,3).

This is an industry that generally operates on projects commissioned by external clients. The work is of job shop nature, so the characteristics of the activities are unique. The work is subject to unpredictable factors, such as activity durations, special project requirements, and uncertainty (4). The complexity of this type of project impacts performance, and these projects may even finish with time delays or cost overruns. Project durations may extend up to 25 % beyond the planned timeline due to scope changes and unconsidered risks (5). Other estimates consider cost overruns of 30 % and delays of 40 % (6).

Bodies of knowledge on project management promote several practices to successfully complete projects. One of them, the Project Management Body of Knowledge (PMBOK) proposes several areas to plan and control, *i.e.*, scope, time, cost, communication, quality, resources, and risks (7). Planning and control require decision-making processes to select paths of action in several situations during a project's lifecycle.

The selection of corrective actions is required at two moments during the project lifecycle. The first moment is the planning phase, when a potential threat to the baseline is identified. To prevent these threats, corrective actions become part of the baseline. The second moment is the execution phase, when an event has occurred. In this case, these actions take place after the event, aiming to correct the progress of the project. Research has typically studied the first of these moments, *i.e.*, the selection of risk response strategies (RRS) (8) or actions (RRA) (9).

This selection process is conducted during the planning phase, after designing a preliminary baseline of scope, time, and costs (7). Risk management is the process involved in this type of decision-making, and its phases include identification, prioritization, evaluation, response design, implementation, and control (7,10,11).

On the other hand, organizations develop projects to implement business strategies and enable the creation of business value (7). They may choose among low-cost, scale-based, specialization, newness,

flexibility, quality, service, and sustainability strategies (12, 13). According to the selected corporate strategy, they define objectives and projects. However, when monitoring and controlling them, they often fail to consider their strategic logic. Corrective actions are usually selected to meet time and cost goals (7).

This paper proposes a hybrid approach to the selection of corrective actions during the planning phase. This approach combines risk management, fuzzy quality function deployment (FQFD), and Monte Carlo simulation (MCS). Response actions are selected through FQFD, and pre-mitigation and post-mitigation analysis is performed via MCS.

The remainder of this document is structured as follows. Section 2 provides an overview of the relevant literature, presenting risk management methods as well as their advantages and disadvantages. Section 3 details the steps involved in constructing the risk response model. In Section 4, the model is applied to a real EPCC project carried out in São Paulo (Brazil). Different response plans were also used to evaluate alternative selection approaches. Section 5 presents the conclusions of this work, states its contributions, and proposes future lines of research.

2. Literature review

The risk management methodology involves a series of activities that may be conducted through several methods. These are presented below.

Identification of potential risks. The first step in the methodology is to identify the risks that may positively and negatively affect the project. There are different risk categories within the scope of a project, *e.g.*, operational, economic, social, political, financial, regulatory, nature-related, environmental or technological. *Operational risks*, for instance, are those related to the actual operation of the project. The product of this phase is a list with many potential risks. Table **I** presents the various methods for identifying potential risks.

Qualitative analysis. After obtaining the list of potential risks, the methodology determines their priority level to identify critical risks for the project. Some qualitative analysis methods are shown in Table **II**.

Quantitative analysis. Once the critical risks have been identified, a numerical evaluation of the impact on project objectives is conducted. This effect is usually oriented to the project's time or cost objectives (7, 19). Some quantitative analysis methods are shown in Table **III**.

Response planning. After evaluating the project's risk level, response plans must be developed while considering several corrective actions, thereby minimizing the threats to the project objectives and maximizing the opportunities it can offer (7). Those actions are usually classified into four general categories: acceptance, transference, avoidance, and reduction. A decision-making process has to be conducted to evaluate and select the appropriate action. Some response planning methods are shown in Table **IV**.

Table I. Techniques used for risk identification

Techniques	Advantages	Disadvantages
Interview	It encourages those who are afraid of revealing risks in front of others to speak up.	If the interviewer does not know the topic, the identified risks may be of little value.
Workshop	It involves project stakeholders and gets them to contribute. It is personal and direct.	It requires time. If the people involved are not the right ones to contribute, it is of no use.
Survey	It facilitates the collection of responses and their analysis.	The quality of the risks identified depends on the quality of the survey. There may be a significant number of people who do not answer them.
Delphi method	It helps to achieve consensus. Many experts can be included to capture their knowledge. It is anonymous.	It takes time if there are many rounds or if there are many experts. Its quality depends on whether the people selected are actual experts on the subject.
RBS (risk breakdown structure)	It helps to think quickly about various aspects of the project and about risks in other areas.	It is limited to identifying risks from the categories of an RBS.
Expert judgment	It allows identifying risks from parts of the project that are not known.	If there is no similar project, there is no previous practical experience that can be contributed.
Brainstorming	Quick, easy, and creative.	It requires time, a good moderator, and engaging the right people.
WBS analysis	It ensures that the entire scope of project work is analyzed for risk identification.	It allows identifying risks exclusively related to the scope.
Fault tree analysis	It shows how resistant a system is to one or more failures.	If drawn manually, it can require quite a bit of effort. It is not good at finding all possible faults.
Checklists	It is most useful when the list is created in a project with a similar context.	They are generic, not specific. If they are old, they may not be useful. If they are long, they can be intimidating. There is a tendency to ignore what is not on the checklist.
Cause and effect analysis (Ishikawa)	It allows identifying and visualizing the causes and sub-causes of risks as well as their relationships.	If the risk has many causes and sub-causes, it may be difficult to perform.
SWOT analysis	It helps to identify negative, positive, internal, and external risks.	It aids in identifying generic risks, not ones that are very detailed.
Project documents and lessons learned	This approach provides, detailed information about the context of the project, its organization, personnel, and experience.	It only identifies the risks that may arise from analyzed documents.

Sources: (7, 14, 15).

Table II. Techniques used for qualitative risk analysis

Techniques	Advantages	Disadvantages
Probability and impact matrix	It is easy to use and learn, and it is widely used. It only requires knowledge of basic probability. Good for communicating visually. It shows the risk level of a project in a matrix. It can be performed and updated in a template, for which there is specific software.	Some consider it subjective, but there are no disadvantages regarding its cost/benefit.
Double probability, and impact matrix	It allows visualizing positive and negative risks in a single matrix.	Some may get confused by analyzing both types of risk at the same time, in addition to threats and opportunities.
Risk categorization	It is inexpensive, scalable, useful, and quick to review. It helps not to forget the entire risk categories.	If the quality of the RBS is not good, it is not very useful. There may be categories that, if not in the RBS, will not allow identifying risks.
Risk urgency	It makes the team focus on what is most urgent first.	Sometimes, people want to mark too many risks as urgent, compromising the richness of the tool.
Quality of information on risks	It is relatively easy to implement.	It is not always easy to obtain more specific data.
Expert consultation	It provides instant experiences and adds value to the analysis.	If the individual is not an expert in the risks of similar projects, the analysis will not be very useful.

Sources: (7, 16–18).

3. Methodology

This study proposes a model for selecting risk responses while following the risk management methodology. Its first phase identifies the set of potential risks that may arise during the project. The second phase identifies critical risks. Afterwards, the impact of risks on the schedule is assessed (pre-mitigation analysis). The next phase identifies potential responses, which are later evaluated. Subsequently, a post-mitigation analysis is conducted to validate the improvement of the schedule. At the end, a project baseline is designed, which may be executed. This framework is shown in Fig. **1**.

Table III. Techniques used for quantitative risk analysis

Techniques	Advantages	Disadvantages
Probability distributions	Uncertainty can be modeled through distributions. They establish the range that the inputs can take to predict the probability of occurrence of each value.	It is necessary to understand spreadsheets, probability distributions, and how to use them with simulation software.
Models and Monte Carlo simulation (MCS)	It is an inexpensive and powerful way to analyze many scenarios in order to observe their effects. It allows forecasting more realistic dates and costs to see if the cost or date needs to be changed. It helps to determine the necessary contingency. It shows the probability that certain activities are on the most frequent critical path. It is used by project managers due to its simplicity and versatility in many aspects of projects and their management.	It applies only to cost and time, not to the other areas of risk. If the model or data is not good, the results will not be useful. It requires the use of software or add-ins.
Expert consultation	It provides instant experiences and adds value to analytics. It helps to determine the input data needed for use with tools such as MCS.	The answers can be objective or subjective.
PERT estimates	They provide a more realistic and accurate timeline, as they incorporate uncertainty in durations and/or costs. They are easy to understand and load into software.	Three estimates must be entered for each higher-risk task instead of one. This is more expensive and takes longer to both create and maintain. In general, its use in low-risk projects is not justified.
Tornado graph and sensitivity analysis	It allows focusing on the critical variables that most influence the project's objectives, as well as knowing which risks require planning for responses.	It analyzes the deviations of one parameter at a time, rather than combinations of several.
Expected monetary value analysis	It is simple and does not require software to perform calculations.	It only calculates the expected value of uncertain events, but, in general, that alone is not enough to make decisions.
Decision trees	They are an easy-to-understand and visual tool. There is software to draw them, which presents the alternatives and results in a professional way. It allows selecting the option with the best profit or the lowest cost.	If there is no basis regarding the estimates of each alternative's probability of occurrence, cost, and gain, the result is unrealistic. It is impractical if there are many risk events, since the total number of outcomes increases exponentially.
Bayesian networks	They are a powerful communication tool. Cause-and-effect relationships are easily visualized without the need for probability calculation. They provide the possibility of combining objective and subjective data (expert judgment).	The model is as good as the modeler is and the experts' perception of reality. *A priori* assumptions that are too optimistic or pessimistic can invalidate the results or skew the network.

Sources: (7, 20, 21).

Table IV. Techniques used for response planning

Techniques	Advantages	Disadvantages
Zonal-based approach (ZBA)	It considers several zones which are used to design specific strategies. It is regarded as the first approximate tool for selecting risk response strategies.	Working with only two criteria is considered a limitation.
WBS-based approach (WBSA)	It is based on the analysis of the project or the activities of the WBS.	There is a lack of mathematical methods for discriminating between the quality of the solutions generated.
Case-based decision analysis (CBDA)	Responses are generated from historical cases regarding risks and responses in previous projects.	It has limited practical applications. There is a lack of information. The target case information may not exist in case-base historical data.
Decision tree method (DTA)	Selection is conducted through scenario analysis. Responses are generated through the WBSA or the ZBA approach.	It only considers make or buy decisions. It is limited in practical applications.
Trade-off approach (TOA)	Analyzing the project through the WBSA or the ZBA provides specific risks and strategies.	This approach only considers two factors or makes trade-offs based on qualitative analysis. Thus, there is a lack of mathematical solutions to the problem.
Optimization-based methods (OBA)	These methods commonly apply specific analysis (e.g., WBSA) to generate risks and responses. This technique has low complexity due to the use of multicriteria decision analysis (MCDA).	Only two criteria are considered. Mathematical models usually solve project examples and not real industry projects.

Sources: (16–24).

3.1. Identification of operational risks

This process generates a list of all the risks that may affect the project. Two strategies were combined to identify risks. A review of scientific literature and a review of historical projects in the studied organization were conducted. The risks obtained were integrated into a single list that was later refined with the support of experts. These experts discarded repeated risks, those that were implicit in others, and those that did not apply to the context. To better understand the risks, they were described in terms of event, impact, and causes (7). Subsequently, the list was categorized using the PMI's structure of first- and second-level categories (7).

Experts are defined as people with education, knowledge, skills, or specialized training, who provide a judgment based on their experience in some area of knowledge, industry, or discipline (7). The experts considered in this study hold decision-making positions in the organization's projects. Their profession and experience in projects must be considered. These experts also took part in the other steps of the methodology.

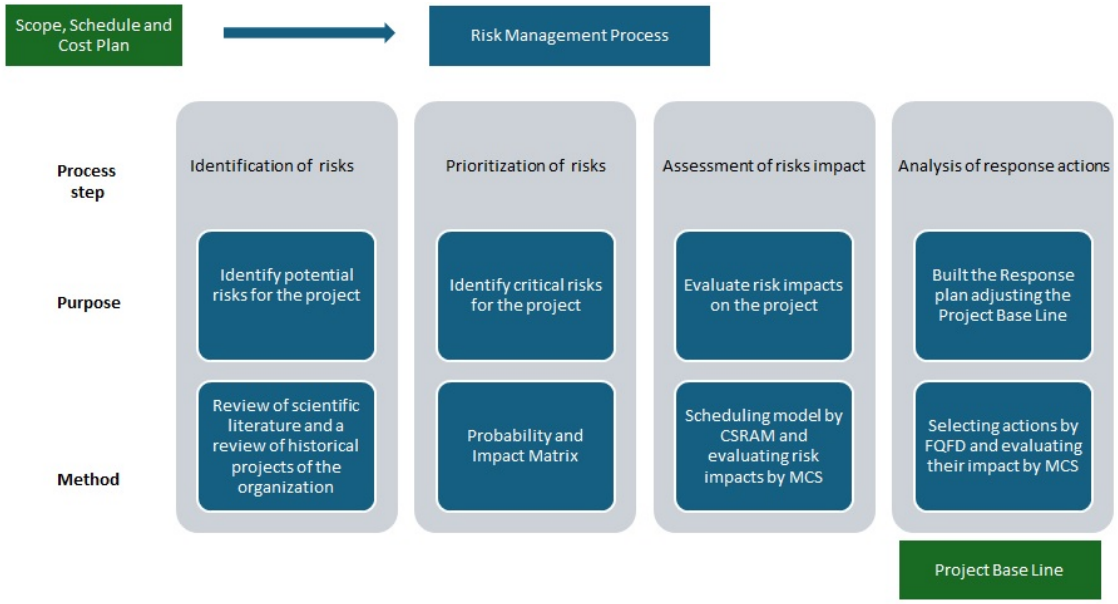


Figure 1. Framework for the proposed model

3.2. Prioritization of operational risks

This involved a qualitative analysis to identify critical risks from a potential risks list. These risks were identified through a survey applied to experts in relation to their analysis of the probability and impact of each risk.

3.3. Assessment of risks impact

This involved a quantitative analysis that measured the impact of critical risks on the project. To this effect, MCS was used. This analysis was performed in pre-action mode and once again in post-action mode. Thus, the impact of response actions could be evaluated.

MCS generates thousands of possible outcomes or scenarios based on probability distributions. This allows modeling costs and times at the activity level (7). Input values are chosen randomly for each iteration, and the outputs represent the range of possible outcomes, such as the project completion date. Thus, the impact of risks on the objectives of the project can be evaluated (25).

According to (35), MCS replaces the estimated duration of any activity of a project with a randomly generated number extracted from a statistical distribution. In general, MCS operates in three steps. The first step is the generation of random numbers in the interval [0,1]. In the second step, the probability value is located in the distribution function of the task, in order to determine the value that corresponds to said probability. Finally, the generated value is used as the input number in the iteration, i.e., the value of the duration or cost of the task.

Using MCS to evaluate the impact of risks on projects is not a new approach. Research varies according to how the information is obtained; the relationship between tasks is studied, as well as the impact of risks on tasks and between the risks themselves. An example of research on the topic is the one by (26). Regarding the impact of risks on project scheduling, MCS is regarded as easy to apply and is indeed the most widely applied method (7, 27).

That said, the methodological structure that we decided to use in this work is CSRAM (28), as it facilitates data collection – most of the data were obtained qualitatively through expert judgment. In addition, it considers the influence of risk on tasks and the correlation between risks. The required data results from the steps presented below.

Deterministic model design. This step involves the elaboration of a network diagram showing task (most probable time t_m) and project duration (critical path method), in addition to minimum (optimistic time t_o) and maximum task duration (pessimistic time t_p).

Influence analysis. This step first defines the risk influence degree of the activities. Influence is categorized as *effective* (E) and *very effective* (VE). By default, any other relationship is rated as *ineffective* (IE) Risks and tasks are related through a matrix specifying how each risk affects each activity. Moreover, potential scenarios regarding the impact of risks on task performance are established. The probability limits are categorized as *better than expected* if the task is positively affected by the risk; *as expected* when the risk had no impact on the task; and as worse *than expected* when the uncertainty of the risk negatively affects the task. For all risks, the better-than-expected limit was set at 0.1, and the worse-than-expected limit was defined as 1.0. The limit for what was expected was between 0.3 and 0.5. The sum of these three values must be equal to 1. Finally, the correlation between risk factors was included. This is typically done through a rating of 1 or 0 given by experts. If the initial risk materializes (1), then the correlated risk also receives that value. However, if the initial risk does not occur (0), there is a random probability that the correlated risk materializes.

Stochastic modeling and simulation. Several computational tools were available for running the model, but we decided to run the deterministic model on MS Project and the stochastic model on Palisade @Risk. First, we defined the inputs, outputs, and number of iterations for the MCS. Task duration was considered as an input related to risk impact and was included using a duration coefficient (**?**). Two types of analysis were conducted: Eqs. (1) and (2) were used if the coefficient was greater or lower than zero, respectively.

$$AD_i = t_m + (t_p - t_m) \cdot DC_i \tag{1}$$

$$AD_i = t_m + (t_m - t_o) \cdot DC_i \tag{2}$$

where AD_i represents the duration of activity i; t_m denotes the most probable time; t_p is the pessimistic time; t_o represents the optimistic time; and DC_i is the duration coefficient for activity i.

In Eq. (3), activity duration oscillates between the most probable and the maximum value. In Eq. (4), the duration oscillates between the minimum and the most probable duration. The duration coefficient depends on the state of the risks. It is a random number that simulates a scenario within the probability

limit imposed on each risk. It also depends on the value associated with the degree of influence of the risk-activity factor. This coefficient aims to combine how risks materialize and how they influence activities, and it is given by Eq. (3).

$$DC_i = \sum_{i=1}^{N} rn_i \cdot iv_{ij} \tag{3}$$

where rn_i is a random number; iv_{ij} represents the degree of influence of every risk j on activity i; and N represents the total number of risks.

The duration coefficient increases or decreases the task duration based on the most probable time. If the risk turns out to be worse than expected, the value of the degree of influence is positive. If the risk materializes as expected, the value of the degree of influence is zero. If the risk occurs better than expected, the value of the degree of influence is negative. These qualitative analyses are converted to quantitative values, considering that very effective ratings represent 70 % of the total impact, as shown in Eq. 4, and that the effective ones account for the remaining 30 %, which is indicated in Eq. (5).

$$iv_{ij} = \frac{0{,}7}{vet_i} \tag{4}$$

$$iv_{ij} = \frac{0{,}3}{et_i} \tag{5}$$

where vet_i represents the number of very effective terms assigned to activity i; and et_i denotes the number of effective terms assigned to activity i.

In addition, considering that a risk may or may not materialize, a binomial variable was included. A new formula to obtain the activity duration coefficient was proposed, replacing Eq. (6).

$$DC_i = \sum_{i=1}^{N} E_j * rn_i * iv_{ij} \tag{6}$$

where E_j is 1 if the risk occurs and 0 if it does not.

3.4. Analysis of response actions

This analysis comprised three steps: identification of potential responses, fuzzy analysis to select the response, and simulation to re-evaluate the risk profile.

First, potential response actions were designed to project risks through workshops with a group of experts. Later, the potential actions were prioritized through a fuzzy quality function deployment (QFD) analysis. QFD was designed by Mitsubishi in 1972 for the manufacturing industry (29). It has spread to other applications such as services and product development, but it is rarely employed in project management (30). It is a tool that links customer requirements to technical specifications. It translates ordinary language into technical terms, facilitating cooperation between marketing, engineering, and manufacturing. It ensures that all requirements are considered and not forgotten (29), and it brings the voice of the customer (VOC) to the product's technical characteristics (29).

This involves the so-called *house of quality*, a matrix that relates customer needs/attributes, known as whats (left side), to engineering characteristics, *i.e.*, *hows* (top side). The what-how relationship level is in the center of the matrix. It also includes needs-relative importance (left side) and requirements co-relationships (upper side), and it allows carrying out benchmarking with competitors or with earlier product versions (right side) as well as assigning requirements importance weights (bottom side) (29).

To consider the uncertainty related to real-world problems and decision-making for modeling imprecise data, the researchers expanded it with fuzzy logic, known as fuzzy QFD (FQFD) (30). The FQFD was applied as proposed by Osorio (31,32) with the following steps:

Identifying customer needs (*whats*). A workshop was conducted with the team of experts to define the project's basic requirements.

Determining the relative importance of the *whats*. A linguistic scale was defined to allow each member of the group to determine the level of importance or weight of each *what*. This scale was associated with a fuzzy number (Table V).

Table V. Fuzzy linguistic scale

Linguistic terms	Triangular fuzzy numbers
Very low (vl)	(0,1,2)
Low (l)	(2,3,4)
Medium (m)	(4,5,6)
High (h)	(6,7,8)
Very high (vh)	(8,9,10)

Sources: (31).

To obtain the weight, the average of each rating given by the experts was calculated using Eq. (7), the result of which was a fuzzy triangular number.

$$w_i = \frac{1}{n} \cdot (w_{i1} + w_{i2} + \cdots + w_{in}) \tag{7}$$

where w_i represents the triangular number (w_{ia}, w_{ib}, w_{ic}) for every *what*; q represents the total number of *whats*; and n represents the number of experts.

Identifying the company's strategic objectives (*hows*). These are the criteria with which all possible alternatives were evaluated. The identification process was carried out through interviews with two of the team's experts who hold management positions in the organization. They determined the objectives based on their experience in the company and after consulting with the company's management.

Determining the correlation between the *whats* and the *hows*. This correlation was established by using the aforementioned linguistic scale and the following equation:

$$r_{ij} = \frac{1}{n} \cdot (r_{ij1} + r_{ij2} + \cdots + r_{ijn}) \tag{8}$$

where Correlation $= \{r_{ij} \mid i = 1, \ldots, q;\ j = 1, \ldots, c\}$; c represents the total number of *hows*; and r_{ij} is the triangular number $(r_{ija}, r_{ijb}, r_{ijc})$ consolidated between each q_i and each c_j.

Determining the weight of the hows. The average value of the evaluations provided by the experts in previous steps was calculated according to Eq. (9).

$$W_j = \frac{1}{q} \cdot ((r_{j1} \cdot w_1) + \cdots + (r_{jq} \cdot w_q)) \tag{9}$$

where W_j is the triangular number (W_{ja}, W_{jb}, W_{jc}).

Determining the impact of alternatives on the *hows*. The experts determined the level of importance of each alternative regarding the strategic objectives using the same scale as in Table **V** and according to Eq. (10).

$$CA_{hj} = \frac{1}{n} \cdot (ca_{hj1} + \cdots + ca_{hjn}) \tag{10}$$

where $CA = \{CA_{hj},$ where $h = 1, \ldots, p;\ j = 1, \ldots, c\}$; p represents the number of alternatives; and ca_{hjn} represents the fuzzy evaluation of the expert n for the alternative h in relation to the variable j.

Prioritizing the alternatives. The fuzzy affinity index (ID) was calculated for each alternative. The result was a fuzzy triangular number obtained through Eq. (11).

$$ID_h = \frac{1}{c} \cdot ((CA_{h1} \cdot W_1) + \cdots + (CA_{hc} * W_c)) \tag{11}$$

$$ID = \{ID_h \mid h = 1, \ldots, p\}$$

Finally, to obtain a non-fuzzy number, the Facchinetti approach was employed. The triangular fuzzy numbers were defuzzified to get a priority ranking that could be ordered from the highest to the lowest value.

$$IDF = \frac{ID_{ha} + 2 \cdot ID_{hb} + ID_{hc}}{4} \tag{12}$$

Finally, a post-action analysis was carried out through simulation to re-evaluate the risk profile, which included risks responses, new tasks, changes in project logic, and new task durations ($t_o, t_m,$ and t_p). This was done through workshops with experts, who evaluated the new conditions to carry out the project.

4. Results – a case study in a construction project

Our method was applied in a real construction project executed in São Paulo (Brazil) by a Colombian engineering company. The research project had the support of five experts from the organization: one mechanical engineer, one electrical engineer, and three industrial engineers with more than five years of experience in project management.

4.1. Identification of operational risks

This process was carried out in two stages. First, a keyword search was carried out, and, later, a detailed review to identify the papers that should be used in the study. In this literature review, 53 articles were found. They were reviewed in detail, and those that dealt with non-operational risks or were related to sectors other than engineering, construction, metalworking companies, logistics, and hydroelectric projects were discarded. Finally, 11 articles remained, corresponding to studies in Ecuador, Colombia, Spain, Peru, the United Kingdom, South Africa, France, China, Australia, and India (33,34). The literature review allowed identifying 252 operational risks, and the review of projects in the organization and workshops with project managers returned 82 risks.

The final list included a total of 334 operational risks. It was filtered with the experts, resulting in 70 classified risks. Following the PMI categorization approach, the first-level risks were technical, management, commercial, and external in nature. These four categories included 22 second-level categories.

The largest number of potential risks was found in the management category, which consolidated 44 % of the risks identified (31), followed by the technical category with 24 % (17 risks). The external and commercial risk categories grouped 19 % (13) and 13 % (9), respectively.

4.2. Prioritization of operational risks

For this analysis, a survey for the expert group was designed. This survey, elaborated in Excel Visual Basic, included the risks list and the qualifications defined by the experts. The critical risks were as follows: failure to deliver plans and manufacturing lists on time (R27); poor management of the communication between those involved in the project (R65); little documentation of activities (R3); loss of intellectual property by sharing design, production, or technology capabilities with other companies (R68); lack of learning techniques to increase knowledge (R2); failure of IT systems (R5); poor project management (R66); reactive risk management (R67); optimistic bias (time, cost) (R63); and pandemics (R55).

4.3. Assessment of risks impact

First, the deterministic model was defined, and information from the influence analysis was collected. The next step was to calculate the duration coefficient of each task, in order to model the duration of the tasks that served as input in the simulation. The output of the model was the duration of the project. With the @Risk add-in of Microsoft Excel, the model was elaborated, and several tests were carried out to evaluate the convergence of the results. The simulation was carried out with 2000, 5000, and 10 000 iterations. In the first case, the average duration was 70.68 days, with a standard deviation of 2.2 days. In the second case, this value was 70.68 days, also with standard deviation of 2.2 days. In the third case, the average duration was 70.62 days, with the same standard deviation. With the little variation obtained, we decided to use the model with 5000 iterations for the project. The results are shown in Figure 2.

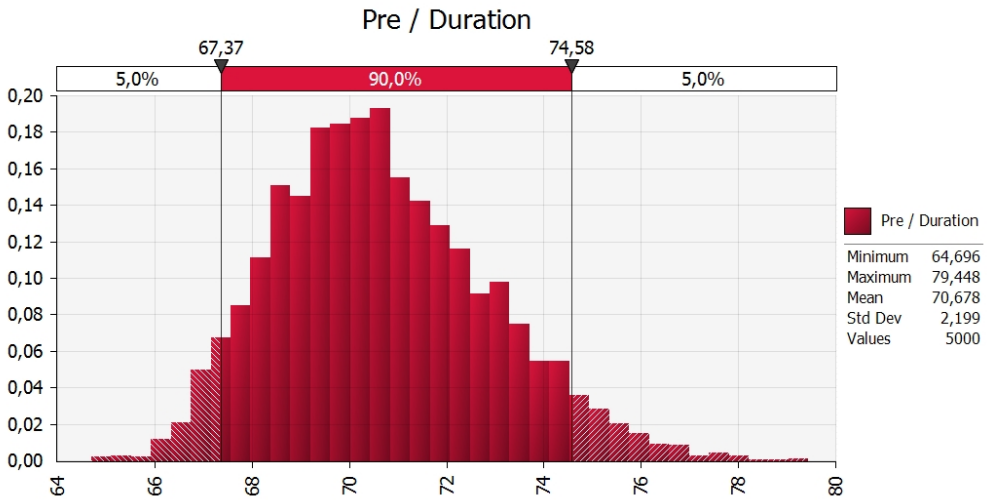


Figure 2. Frequency distribution with 5 000 iterations

An important finding was that the original 63 days schedule had a 0the 5000 iterations. On the other hand, a sensitivity analysis was conducted, considering activity duration changes. We elaborated a tornado graph by calculating the duration of the project plus the randomness related to the coefficient of activity duration. This graph highlighted the activities that most affected the duration of the project, with the first four being *1.8.0 Dispatch, 1.3.1 Legalization of the contract, 1.3.3 Validation of PTF*, and *1.4.1.1 Modulation through Forline*.

4.4. Framework for the proposed model

Following the proposed method, the potential response actions were initially designed. The expert group designed responses for the critical risks (step 3.2) and the tasks with the greatest variability (step 3.3). The group designed several actions that could respond to each case. Table **VI** presents the number of potential actions designed by the project team.

A set of 40 potential actions was used to design the response plan and adjust the project. Each action was thoroughly designed. To decide which action to implement, FQFD was applied. The experts defined the components of the matrix. A set of basic project requirements was established, representing customer needs (*whats*). The project requirements were as follows: finish the project on time (Q1), have low costs (Q2), and comply with the scope (Q3). A set of company strategic objectives was established to represent the *hows*.

For the workshop, the expert group also invited the manager of the company. The group defined some strategic objectives while considering the corporate strategy and goals, *i.e.*, satisfy the customer (C1), make the business profitable (C2), and implement new technologies (C3). The correlation between *whats* and *hows* was established while considering several values, following fuzzy logic as seen in Table **VI**.

Table VI. Potential actions

Risk	Potential actions	Activity	Potential actions
R27	2	A 1.8.0	3
R65	3	A 1.3.1	3
R3	2	A 1.3.3	3
R68	3	A 1.4.1.1	3
R2	2		
R5	2		
R66	3		
R67	3		
R63	3		
R55	5		

Table VII. What-how relationship

	C1			C2			C3		
Q1	7	8	9	6	7	8	4	5	6
Q2	5	6	7	6	7	8	4	5	6
Q3	3	4	5	5	6	7	4	5	6

The Q and C weights were also established. The project requirements were assigned the same weight, while the strategic objectives received different weights. Table **VIII** shows the weight of the *hows*.

Table VIII. Weight of the *hows*

C1			C2			C3		
32	44	58	36	49	64	25	37	50

The following is the application of the method for the first risk (R27, failure to deliver manufacturing plans and lists on time). This risk impacts 22 tasks and may cause delays in the company's production plant. The proposed response actions were hiring more personnel to speed up the delivery of plans (A) and applying two shifts to the engineering area (B).

The FQFD analysis of each alternative, considering its impact on strategic objectives, resulted in 220 (option A) and 313 points (option B). Option B was selected, and two shifts were defined for the engineering area, entrusted with modeling, design, and blueprint elaboration. Two groups of five individuals were assigned to the shifts from 6 am to 2 pm and from 2 pm to 10 pm. The same method was applied for the remaining risks. The final response plan included 17 possible actions.

The last step was to carry out a post-action simulation of the project. Considering the responses to the prioritized risks, the new conditions were included in the model. For the non-prioritized, remaining, and secondary risks, the group decided to apply a contingency reserve from the MCS, and, for the unknown risks, the group used a management reserve defined by the company (1 %). The simulation was carried out with 5000 iterations, resulting in a duration between 55 and 64 days, with an average of 60 days (Fig. 3).

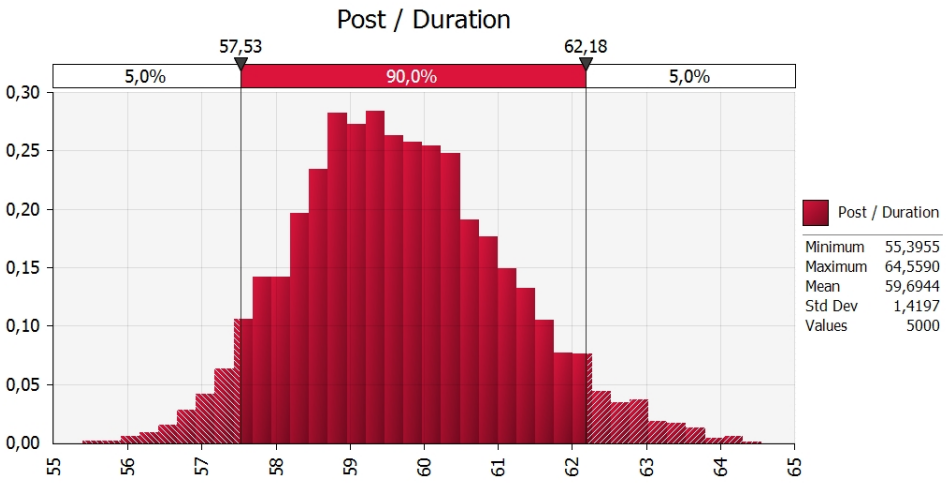


Figure 3. Post-action frequency distribution with 5 000 iterations

Compared to the pre-action simulation, the 63-day project had a 96 % probability of completion. In graphical terms, the frequency curve shifted to the left, as indicated in Fig. 4 where the initial simulation is seen in red and the post-action one is seen in blue.

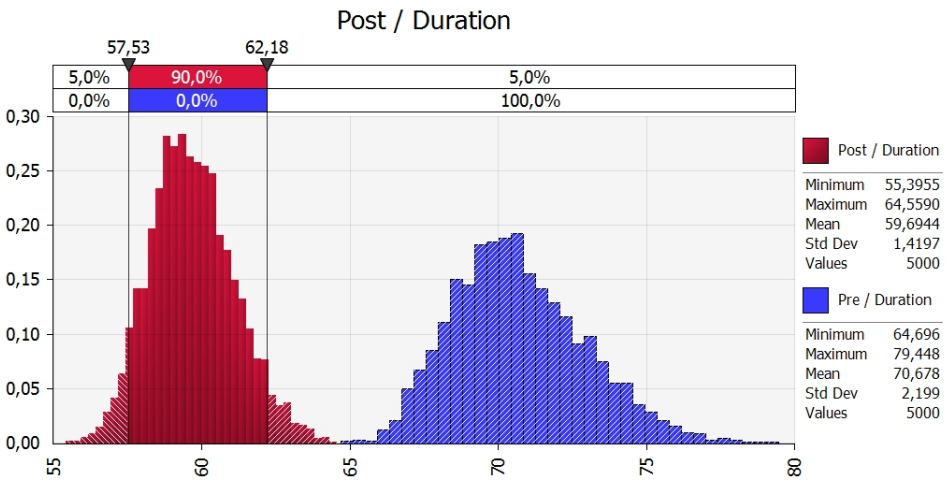


Figure 4. Pre- and post-action frequency distribution

The new project design includes new tasks but was accelerated by 11 days. To achieve this, increased project budgets were established. The project went from an initially estimated cost of $1 030 772 694 to $1 135 813 141. However, part of the budget reserves may become part of profitability if they are not required.

The development of new response plans aimed to test the behavior of the model. Following the literature, the team developed plans with the lowest cost and shortest duration. Other plans logically and comprehensively integrated actions for the project. Finally, the team conducted simulations using the same parameters as the previous application. Table **IX** summarizes the information for each plan, including the FQFD-designed base plan.

Table IX. Response plans

Response plan	Response budget	Estimated project cost	Mean duration	Standard deviation
Plan 1	$ 8 463 631	$ 1 039 236 325	63	3.1
Plan 2	$ 19 999 996	$ 1 050 772 690	70	1.4
Base plan	$ 59 278 522	$ 1 090 051 216	60	1.3
Plan 3	$ 66 776 004	$ 1 097 548 698	59	3.1
Plan 4	$ 89 263 520	$ 1 120 036 214	59	3.3
Plan 5	$ 64 589 764	$ 1 095 362 458	61	2.1
Plan 6	$ 79 463 853	$ 1 110 236 547	65	2.3

The first response plan aimed to minimize costs, resulting in a total cost of $8 463 631. It changed the project network because of the emergence of new tasks and the adjustment of some task durations. The simulation resulted in an average duration of 63 days, which is the maximum term of the project. The simulation also revealed that there was a 50 % chance that the project duration would exceed 63 days. As a result, the team decided that this option was not viable.

After exploring different potential actions, a new plan was designed to cost less than the model's proposed plan. The new plan would cost $19 999 996 and include changes to the project network, such as adding new tasks, removing others, and changing the duration of some. The initial simulation showed that the project would take an average of 70 days to complete, which is longer than the maximum allowable time. Additionally, based on the simulation data, there was a 0 % probability of delivering the project within 63 days or less, meaning that the project would not be completed within the required timeframe under any scenario.

By shifting the directive to prioritize minimal time, two combinations of actions were found which involved modifications to the project design. Plans 3 and 4 resulted in an average duration of 59 days and deviations of 3.1 and 3.3 days, respectively. In both cases, the probability of meeting the maximum term of 63 days was high, *i.e.*, 90 and 89 %, respectively. However, these plans cost more than the previously designed, $59 278 522 plan.

Two additional plans proposed by the organization's engineering team were tested. Plan 5 resulted in an expected duration of 61 days, within the maximum time allowed. Considering the resulting deviation of this simulation, this plan had an 83 % probability of completion before 63 days for $64 589 764. This plan was in the middle of the previous plans, which sought to minimize cost and time. However, it was more expensive than the base plan and had a lower probability of success.

Finally, plan 6 turned out to be an inconvenient option. Its average duration was 65 days, longer than required, and it had a 19 % chance of being delivered. The cost, $79 463 853, was one of the highest among all the options.

In summary, plans 3, 4, and 5 were convenient, resulting in a >83 % probability of meeting the required time. Nevertheless, considering the cost of the proposals, it was convenient to implement the base plan. In order of convenience, plan 5 could follow it, given its cost, despite offering an 83 % probability of delivery within the term, which was the lowest among the options.

5. Discussion

5.1. Theoretical implications

This paper proposes a hybrid approach for selecting RRAs in EPCC projects, which combines risk management, mathematical techniques, and business strategy. This broadens the scope of study regarding construction project risk management and offers a new viewpoint for choosing RRAs. The theoretical implications of this research are threefold.

First, this study expands research on the selection of RRAs for EPCC projects by proposing a strategic perspective. Prior studies used selection criteria based on operational factors such as cost, time, or quality, without taking the organization's strategic direction into account. This study adds strategic criteria to the existing literature on RRA selection research and uses FQFD to evaluate them. By integrating techniques like WBS and mathematics in the selection process, this study also broadens the body of literature.

Second, this study enriches research on project scheduling in risk situations in EPCC projects. Unlike other studies, uncertainty is not applied to task duration, but to the way in which risks can materialize. This development contributes to the theoretical understanding of variability and expands the boundaries of RRA selection research. It can also inspire academics to conduct in-depth research on the relationship between uncertainty and task duration ranges.

Third, this research contributes to the modeling of other approaches for managing uncertainty in EPCC projects. Our model considers the design of individual actions selected via the strategic approach, whose impact is evaluated with a post-mitigation simulation. This enables a comparative pre- vs. post-mitigation evaluation during project design and the selection of the best strategy to respond to risks. Current research selects the best actions from criteria such as time or cost, using optimization

models that do not consider the relationships between actions.

5.2. Practical implications

This model provides relevant and necessary information for decision-making by the project team. It is consistently integrated into the project baseline planning process. In addition, it allows for the discussion of decisions by the management (with a strategic focus) and the project team (with an operational focus). It also allows delving deeper into the analysis of risks in EPCC projects, their behavior and impact, and the way to design better responses. Consequently, it offers a way to improve planning and increase the probability of success.

5.3. Limitations and discussion of future research

Some limitations suggest avenues for future research. Our model establishes task durations while considering uncertainty as a deviation from the most probable time. Future research could incorporate uncertainty in another way, within optimistic and pessimistic time ranges. Additionally, variability may be considered based on the effect of resource usage, which ultimately affects the task duration. Likewise, new ways of understanding the relationship between the risks themselves can be explored, which can entail additional effects on the project.

Since response actions can generate secondary risks, future research may consider aggregating these impacts into decision-making. This requires understanding whether risk behavior can change during the project's lifecycle. In this way, models can more accurately assess impacts.

The model relates strategic and project objectives to assess potential responses to risks. More research is needed on how to relate these goal levels in order to validate their usefulness. Future works could study how to relate them to the requirements formulation of the project. Other criteria, such as stakeholder or sustainability considerations, should also be explored.

Our model analyzes and makes decisions regarding potential risks in a project. Future research should relate comprehensive decision-making to both the project and the project portfolio. Another future direction for research is to apply the proposed model to more EPCC projects, aiming to verify its broad applicability.

6. Conclusions

EPCC projects are complex and often suffer from gaps in basic objectives regarding scope, time, and cost. Analyzing risks and selecting responses in the planning phase helps for a better execution and increases the likelihood of success. Usually, the problem is not studied outside the criteria of time, cost, quality, and optimization, which can affect the selection of responses. This study developed a practical model based on risk management which considers the impact of and relationships between risks and strategic and operational criteria to select RRAs. The validation of said model in a real project carried out in Brazil allowed verifying its applicability. Experimentation with response plans

focused on decreasing project times or costs, resulting in cross-effects. The plans that decreased time resulted in the highest costs, while those that decreased cost lengthened the project. The model led to a cost-and-time response plan with the confidence that decisions were framed within the organization's strategic priorities.

The main contributions of this study are as follows. First, it introduces a new perspective for evaluating RRAs, enriching theoretical research on the selection of RRAs in EPCC projects. Second, it effectively integrates the relationship between risks and incorporates effect variability into the problem, instead of assessing duration. Third, the proposed model can integrate the risk management process with the scope, time, and cost definition process, serving as a guide for planners to enhance project performance.

7. CRediT author statement

Alvaro Cuadros conceptualized, supervised, and validated the study. Alexander and Leonardo Bustos designed the methodology, conducted the research, and validated the results. All authors collaborated in writing and improving the manuscript.

References

[1] S. Mittal and N. Gorowara, "Knowledge Integration in engineering, procurement and construction projects: A conceptual study," *Psychol. Educ. J.*, vol. 58, no. 1, pp. 5733–5738, 2021. https://doi.org/10.17762/PAE.V58I1.2209 ↑3

[2] H. Erol, I. Dikmen, G. Atasoy, M. Talat Birgonul, and M. T. Birgonul, "An analytic network process model for risk quantification of mega construction projects," *Expert Syst Appl*, vol. 191, art. 116215, 2022. https://doi.org/10.1016/J.ESWA.2021.116215 ↑3

[3] E. Oliveira and C. Santos, "Application of a risk management methodology in industrial projects: A case study in the metalworking sector," in *Educ. Excellence Innov. Manag. Vision 2020 App.*, 2019, pp. 5647–5662, Online]. Available: https://repositorium.sdum.uminho.pt/bitstream/1822/61217/1/Application%20of%20a%20Risk%20Management%20Methodology%20in%20Industrial%20projects.pdf. ↑3

[4] A. Birjandi and S. M. Mousavi, "Fuzzy resource-constrained project scheduling with multiple routes: A heuristic solution," *Autom. Constr.*, vol. 100, pp. 84–102, Apr. 2019. https://doi.org/10.1016/j.autcon.2018.11.029 ↑3

[5] J. Silva *et al.*, "Improvement of planning and time control in the project management of a metalworking industry - Case study," *Procedia Comput. Sci.*, vol. 196, no. 2021, pp. 288–295, 2021. https://doi.org/10.1016/j.procs.2021.12.016 ↑3

[6] PricewaterhouseCoopers, "En la ruta de la Competitividad. Principales hallazgos de la 1ra Encuesta Nacional de Madurez en Gerencia de Proyectos," PwC, Bogotá, Colombia, 2011. ↑3

[7] S. Changali, A. Mohammad, and M. Van Niewlan, "The construction productivity imperative," 2015. [Online]. Available: https://www.mckinsey.com/capabilities/operations/our-insights/the-construction-productivity-imperative ↑3, 4, 5, 6, 7, 8, 9, 10

[8] PMI, *A guide to the Project Management Body of Knowledge (PMBOK Guide)*, 6th, 2017. ↑3

[9] L. Bai, Q. Xie, J. Lin, S. Liu, C. Wang, and L. Wang, "Dynamic selection of risk response strategies with resource allocation for construction project portfolios," *Comput. Ind. Eng.*, vol. 191, p. 110116, May 2024. https://doi.org/10.1016/j.cie.2024.110116 ↑3

[10] Y. Zhang and F. Zuo, "Selection of risk response actions considering risk dependency," *Kybernetes*, vol. 45, no. 10, pp. 1652–1667, Nov. 2016. https://doi.org/10.1108/K-05-2016-0096 ↑3

[11] APM, *Association for Project Management Body of Knowledge*, 6th., 2012. ↑3

[12] OGC, *Managing successful projects with PRINCE2*, 6th ed., 2017. ↑4

[13] M. R. Gleim, H. McCullough, N. Sreen, and L. G. Pant, "Is doing right all that matters in sustainability marketing? The role of fit in sustainable marketing strategies," *J. Retail. Consum. Serv.*, vol. 70, 2023, art. 103124. https://doi.org/10.1016/j.jretconser.2022.103124 ↑4

[14] W. Stevenson, *Operations Management*, 14th ed., New York, NY, USA: McGraw-Hill Education, 2021. ↑5

[15] R. F. Jacobs and R. B. Chase, *Administracion de operaciones: produccion y cadena de suministros*, 15th ed., New York, NY, USA: McGraw-Hill Education, 2019. https://ucreanop.com/wp-content/uploads/2020/08/Administracion-de-Operaciones-Produccion-y-Cadena-de-Suministro-13edi-Chase.pdf ↑5

[16] PMI, "The standard for portfolio management," 2017. [Online]. Available: https://www.pmi.org/pmbok-guide-standards/foundational/standard-for-portfolio-management ↑6, 8

[17] L. E. Dounavi, E. Dermitzakis, G. Chatzistelios, and K. Kirytopoulos, "Project management for corporate events: A set of tools to manage risk and increase quality outcomes," *Sustainability*, vol. 14, no. 4, pp. 1–37, 2022. https://doi.org/10.3390/SU14042009 ↑6, 8

[18] M. Hamid, A. M. Abdelalim, M. Abdel, H. Hassanen, and A. M. Abdelalim, "Risk identification and assessment of mega industrial projects in Egypt," *Int. J. Manag. Commer. Innov.*, vol. 10, no. 1, pp. 187–199, 2022. https://doi.org/10.5281/zenodo.6579176 ↑6, 8

[19] S. Bakri *et al.*, "Identification of factors influencing time and cost risks in highway construction projects," *Int. J. Sustain. Constr. Eng. Technol.*, vol. 12, no. 3, pp. 280–288, 2021. https://doi.org/10.30880/ijscet.2021.12.03.027 ↑4, 8

[20] T. Yuan, P. Xiang, H. Li, and L. Zhang, "Identification of the main risks for international rail construction projects based on the effects of cost-estimating risks," *J. Clean. Prod.*, vol. 274, p. 122904, 2020. https://doi.org/10.1016/j.jclepro.2020.122904 ↑7, 8

[21] A. Nurdiana, M. Agung Wibowo, Y. Fundra Kurnianto, M. A. Wibowo, and Y. F. Kurnianto, "The identification of risk factors of delay on the road construction project in indonesia," in *Int. Conf. Maritime Archipelago (ICoMA 2018)*, 2019, pp. 384–387. https://doi.org/10.2991/ICOMA-18.2019.82 ↑7, 8

[22] M. H. Kotb and M. M. Ghattas, "Risk identification barriers in construction projects in MENA," *PMI World J.*, 2018. https://doi.org/10.13140/RG.2.2.20614.83525. ↑8

[23] S. Maulana and F. D. Ariyanti, "Application of lean project management method in environmental drainage development case study: x area Bekasi City," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1096, no. 1, art. 12085, 2021. https://doi.org/10.1088/1757-899X/1096/1/012085 ↑8

[24] J. Chilumo *et al.*, "Risk management practices on performance of building construction projects," *J. Entrep. Proj. Manag.*, vol. 4, no. 6, art. 202, 2020. https://stratfordjournals.org/journals/index.php/journal-of-entrepreneurship-proj/article/view/659 ↑8

[25] J. Crispim, L. H. Silva, and N. Rego, "Project risk management practices: The organizational maturity influence," *Int. J. Manag. Proj. Bus.*, vol. 12, no. 1, pp. 187-210, 2018. https://doi.org/10.1108/IJMPB-10-2017-0122 ↑9

[26] E. Prihartanto and M. D. Bakri, "Identification the highest risk of performance based contract in Bojonegoro-Padangan road projects," in *Reg. Conf. Civil Eng. RCCE*, Surabaya, Indonesia, 2017. http://dx.doi.org/10.12962/j23546026.y2017i6.3243 ↑10

[27] D. K. Sudarsana, "A concept model to scale the impact of safety risk in a construction project using a semi quantitative method," *Civ. Eng. Archit.*, vol. 9, no. 1, pp. 263–269, 2021. https://doi.org/10.13189/cea.2021.090122 ↑10

[28] Ayuningtyas, D. and Dita Rarasati, A, "Work acceleration strategy development on design-build project to improve risk based quality performance," *Glob. J. Sci. Eng.*, vol. 02, pp. 10–15, 2020. https://doi.org/10.37516/global.j.sci.eng.2020.007 ↑10

[29] L. Wu, H. Bai, C. Yuan, and C. Xu, "FANPCE technique for risk assessment on subway station construction," *J. Civ. Eng. Manag.*, vol. 25, no. 6, pp. 599–616, 2019. https://doi.org/10.3846/JCEM.2019.10373 ↑11, 12

[30] B. Barghi, S. S. sikari, and S. Shadrokh sikari, "Qualitative and quantitative project risk assessment using a hybrid PMBOK model developed under uncertainty conditions," *Heliyon*, vol. 6, no. 1, art. 3097, Jan. 2020. https://doi.org/10.1016/J.HELIYON.2019.E03097 ↑11, 12

[31] M. Kaut, H. Vaagen, and S. W. Wallace, "The combined impact of stochastic and correlated activity durations and design uncertainty on project plans," *Int. J. Prod. Econ.*, vol. 233, art. 108015, 2020. https://doi.org/10.1016/J.IJPE.2020.108015 ↑12

[32] M. Eckhart, B. Brenner, A. Ekelhart, and E. Weippl, "Quantitative security risk assessment for industrial control systems: Research opportunities and challenges," *J. Internet Serv. Inf. Secur. (JISIS)2*, vol. 09, no. 03, pp. 52–73, 2019. https://doi.org/10.22667/JISIS.2019.08.31.052. ↑12

[33] N.-T. Nguyen, Q.-T. Huynh, and T.-H.-G. Vu, "A Bayesian critical path method for managing common risks in software project scheduling," in *Ninth Int. Sym. Info. Comm. Tech. - SoICT 2018*, 2018, pp. 382–388. https://doi.org/10.1145/3287921.3287962 ↑14

[34] X. Xu, J. Wang, C. Z. Li, W. Huang, and N. Xia, "Schedule risk analysis of infrastructure projects: A hybrid dynamic approach," *Autom. Constr.*, vol. 95, pp. 20–34, 2018. https://doi.org/10.1016/j.autcon.2018.07.026 ↑14

[35] M. Vanhoucke, *Integrated Project Management Sourcebook*, vol. 2, CHam, Germany: Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-27373-0 ↑9

Álvaro Julio Cuadros López

He graduated as an industrial engineer in 1993, as a specialist in Technology Management in 1997, as a Master of Administration in 2010, and as a PhD in Engineering in 2019. He currently serves as a professor-researcher in the area of Project Management at the School of Industrial Engineering of Universidad del Valle. His publications and presentations are related to the areas of project risk modeling, project planning and control model, maturity models, and project management bodies of knowledge. He is a member of the Project Management Institute (PMI) and the Ibero-American Project Engineering Network (RIIPO).
Email: alvaro.cuadros@correounivalle.edu.co

Alexander Bustos Useche

He graduated as a mechanical engineer in 2018 from Universidad del Valle, Cali, Valle del Cauca, Colombia. He graduated as a Master of Project Management in 2021 from the same university. He currently serves as a project director at INAGROMIN SAS, a company dedicated to the design and manufacture of integral solutions for the transport and storage of materials in the agricultural and mining sectors.
Email: alexander.bustos@correounivalle.edu.co

Leonardo Bustos Useche

He graduated as an industrial engineer in 2008 from Universidad del Valle, Cali, Valle del Cauca, Colombia. He received a Master's degree in Project Management from the same university in 2021. He currently serves as an audit analyst at Automann Heavy Duty Canada, a company dedicated to the design, manufacture, and commercialization of parts for trucks and trailers.
Email: bustos.leonardo@correounivalle.edu.co