

Universidad Politécnica Salesiana

# SMART TECHNOLOGIES

## SMARTTECH-IC 2021

Proceedings of the Second International Conference on Smart Technologies, Systems and Applications

Fabián Narvárez E.  
(Coordinador)





This book contains a group of selected papers which were presented at the Second International Conference on Smart Technologies, Systems and Applications (SmartTech-IC 2021) held on December 1-3, 2021 in Quito, Ecuador. The SmartTech-IC conference aims to attract researchers, scientists and technologists from some of the top companies, universities, research groups, and government agencies from Latin America and around of the world to communicate their research results, inventions and innovative applications in the area of smart science and the most recent smart technological trends. The second edition of SmartTech-IC conference was organized by the Universidad Politécnica Salesiana (Campus Girón), a private institution of higher education with social purposes, nonprofit and co-financed from the Ecuadorian State.

The SmartTech-IC conference has been conceived as an academic platform to promote the creation of technical and scientific collaboration networks. The goal of the conference was addressed some relevant topics related to smart technologies, smart systems, smart trends and applications in different domains in the field of computer science and information systems that represent innovation in current society.



ISBN-13: 978-9978-10-6457-7



9 789978 106457



## SMART TECHNOLOGIES

### SmartTech-IC 2021: Proceedings of the Second International Conference on Smart Technologies, Systems and Applications

© Fabián Narváez E. (Coordinador)

**Autores:** Mario Caiza, Franklin Camacho, Juan P. Cartagena, María Casamitjana, Luis Castano-Londono, Andrés E. Castro-Ospina, Marvin Coto-Jiménez, Franklin De la Cruz, Gloria M. Díaz, Leonardo Duque-Muñoz, Saul Figueroa, Michelle Flores, Rubén D. Fonnegra, July Galeano, Joseline García, Cristian Guarnizo-Lemus, David Jiménez-Murillo, Teresa Magal-Royo, David Marquez-Viloria, Juan D. Martínez-Vargas, André Mejía Páiz, Nicolás Alberto Molina Cerón, Roberto Moya-Jiménez, Fabián Narváez, Leony Ortiz Matos, Ama Owusu-Andanese, Wilson Pavón Vallejos, Carlos A. Ramos-Paja, Bonie J. Restrepo-Cuestas, Valentina Suescún Arias, Daniel F. Terraza-Arciniegas, María Constanza Torres Madroñero, Adriana Trejos, Julian Uribe-Rios, Victor Yeom-Song

1.ª edición: Universidad Politécnica Salesiana  
Av. Turuhuayco 3-69 y Calle Vieja  
Cuenca-Ecuador  
Casilla: 2074  
PBX: (+593 7) 205 0000  
Fax: (+593 7) 408 8958  
correo-e: rpublicas@ups.edu.ec  
www.ups.edu.ec

CARRERA DE BIOMEDICINA  
Grupo de Investigación en Bioingeniería y Biomecatrónica-GIByB

ISBN impreso: 978-9978-10-645-7  
ISBN digital: 978-9978-10-646-4

Edición, diseño,  
diagramación  
e impresión Editorial Universitaria Abya-Yala  
Quito-Ecuador

Tiraje: 300 ejemplares

Impreso en Quito-Ecuador, abril de 2022

Publicación arbitrada de la Universidad Politécnica Salesiana

El contenido de este libro es de exclusiva responsabilidad de los autores



# Organization

## *Honorary Committee*

Juan Cárdenas	Universidad Politécnica Salesiana, Ecuador
María Sol Villagómez	Universidad Politécnica Salesiana, Ecuador
Juan Pablo Salgado	Universidad Politécnica Salesiana, Ecuador
Tatiana Mosquera	Universidad Politécnica Salesiana, Ecuador

## *Organizing Committee*

Fabián Narváez E.	Universidad Politécnica Salesiana, Ecuador
Julio R. Proaño	Universidad Politécnica Salesiana, Ecuador
Paulina A. Morillo	Universidad Politécnica Salesiana, Ecuador
Diego F. Vallejo	Universidad Politécnica Salesiana, Ecuador

## **Program Chair**

### *Smart Technologies:*

César Ferri	Universidad Politécnica de Valencia, Spain
Patricia Acosta Vargas	Universidad de Las Américas, Ecuador

### *Smart Systems*

Teodiano Bastos	Universidade Federal do Espírito Santo, Brazil
Daniel González Montoya	Instituto Tecnológico Metropolitano, Colombia
Esteban Inga	Universidad Politécnica Salesiana, Ecuador

### *Smart Applications:*

Antoine Manzanera	ENSTA-ParisTech, France
Angel Cruz-Roa	Universidad de Los Llanos, Colombia
Gloria Díaz	Instituto Tecnológico Metropolitano, Colombia

## **Program Committee**

Alberto López Delis	Medical Biophysics Center, Cuba
Alexander Águila	Universidad Politécnica Salesiana, Ecuador

Alvaro D. Orejuela	Universidad del Rosario, Colombia
Antoine Manzanera	ENSTA - ParisTech, France
Andrés Felipe Ruiz Olaya	Universidad Antonio Nariño, Colombia
Ana Cecilia Villa	Universidad Politécnica Salesiana, Ecuador
Ángel Cruz Roa	Universidad de Los Llanos, Colombia
Angélica Ramírez	Universidad Militar Nueva Granada, Colombia
Carlos Cevallos	Universidad Politécnica Nacional, Ecuador
Carlos Mera	Instituto Tecnológico Metropolitano, Colombia
Carmen Carrión	Universidad de Castilla - La Mancha, Spain
César Ferri	Universidad Politécnica de Valencia, Spain
Christian Cifuentes	Universidad de los Andes, Colombia
Daniel González Montoya	Instituto Tecnológico Metropolitano, Colombia
David Romo	Universidad Industrial de Santander, Colombia
David Ojeda	Universidad Técnica del Norte, Ecuador
Derfrey Duque Quintero	Universidad Pascual Bravo, Colombia
Diego Carrión	Universidad Politécnica Salesiana, Ecuador
Diego Almeida	Yachay Tech University, Ecuador
Diego Vallejo	Universidad Politécnica Salesiana, Ecuador
Eduardo Pinos	Universidad Politécnica Salesiana, Ecuador
Enrique Arias	Universidad de Castilla - La Mancha, Spain
Erick Reyes	Instituto Tecnológico Metropolitano, Colombia
Esteban Inga	Universidad Politécnica Salesiana, Ecuador
Estefanía Coronado	Fundazione Bruno Kessler FBK, Italy
Fabián Narváez E.	Universidad Politécnica Salesiana, Ecuador
Fernando Urgiles	Universidad Politécnica Salesiana, Ecuador
Fernando Villalba	Yachay Tech Universidad, Ecuador
Germán Arévalo	Universidad Politécnica Salesiana, Ecuador
Germán Corredor	Case Western Reserve University, USA
Gloria M. Díaz	Instituto Tecnológico Metropolitano, Colombia
Gustavo Caiza	Universidad Politécnica Salesiana, Ecuador
Hiram Ponce Espinosa	Universidad Panamericana, Mexico
Hugo Franco	Universidad Central, Colombia
Ignacio Larrabide	Universidad Nacional UNICEN, Argentina
Israel Pineda	Yachay Tech University, Ecuador
Jack Bravo	Universidad Politécnica Salesiana, Ecuador
Jairo J. Espinosa	Universidad Nacional de Colombia, Colombia
José Ignacio Huertas	Tecnológico de Monterrey, Mexico
Jos 'e Sampietro	CELEC EP, Ecuador
Jorge E. Camargo	Universidad Antonio Nariño, Colombia
Juan Pablo DA 'manto	Universidad Nacional UNICEN, Argentina
Juan Sebastián Botero	Instituto Tecnológico Metropolitano, Colombia



Juan David Martínez	Instituto Tecnológico Metropolitano, Colombia
Juan C. Caicedo	Broad Institute of MIT and Harvard, USA
Juan C. Santillán	Universidad Politécnica de Chimborazo, Ecuador
Julio Proaño Orellana	Universidad Politécnica Salesiana, Ecuador
Leony Ortiz	Universidad Politécnica Salesiana, Ecuador
María Blanca Caminero	Universidad de Castilla - La Mancha, Spain
María Constanza Torres	Instituto Tecnológico Metropolitano, Colombia
M. Carmen Juan Lizandra	Universidad Politécnica de Valencia, Spain
María Gabriela Baldeón	University of South of Florida, USA
Mauro Callejas	Universidad Pedagógica y Tecnológica, Colombia
Milton Ruiz	Universidad Politécnica Salesiana, Ecuador
Lenin V. Campozano	Escuela Politécnica Nacional, Ecuador
Lucía Rivadeneira	University of Manchester, United Kindom
Oscar Acosta	Université de Rennes 1, France
Patricia Acosta Vargas	Universidad de las Américas, Ecuador
Pablo Álvarez	Université de Rennes 1, France
Paula Rodríguez	Universidad Nacional de Colombia, Colombia
Paulina Morillo	Universidad Politécnica Salesiana, Ecuador
Ramón Pérez Pineda	Universidad Politécnica Salesiana, Ecuador
Ricardo Flores	Universidad San Francisco de Quito, Ecuador
Ricardo Gutiérrez	Universidad Militar Nueva Granada, Colombia
Rubén D. Fonnegra	Universidad Pascual Bravo, Colombia
Santiago Gonzáles Martínez	Universidad de Cuenca, Ecuador
Sergio Luján	Universidad de Alicante, Spain
Teodiano Bastos	Universidade Federal do Espírito Santo, Brazil
Vladimir Robles	Universidad Politécnica Salesiana, Ecuador
Villie Morocho	Universidad de Cuenca, Ecuador
Ximena López	Universidad de San Marcos, Perú

### **Local Organizing Committee**

Tatiana Mosquera	Universidad Politécnica Salesiana, Ecuador
Janneth Pallascos	Universidad Politécnica Salesiana, Ecuador
Jessica Chávez	Universidad Politécnica Salesiana, Ecuador
María Belén Sánchez	Universidad Politécnica Salesiana, Ecuador
Mónica Ruiz	Universidad Politécnica Salesiana, Ecuador
Julissa Freire	Universidad Politécnica Salesiana, Ecuador
Christian Guachilema	Universidad Politécnica Salesiana, Ecuador
Byron Fernando Velasco	Universidad Politécnica Salesiana, Ecuador



## Sponsoring Institutions



<http://www.ups.edu.ec>



<https://www.itm.edu.co/>



# Table of Contents

---

## Preface

### **Identification of Costa Rican speaker's age group based on acoustic vowel analysis**

*Victor Yeom-Song, Marvin Coto-Jimenez*

### **U-Net Based on MRI Image Patches for Skull Stripping**

*David Jiménez-Murillo, Daniel F. Terraza-Arciniegas,  
Leonardo Duque-Muñoz, Juan D. Martínez-Vargas,  
Andrés E. Castro-Ospina*

### **Python application for analysis of photovoltaics cells operating in both first and second quadrant**

*Bonnie J. Restrepo-Cuestas, Cristian Guarnizo-Lemus, Adriana Trejos,  
Carlos A. Ramos-Paja*

### **Direct diffuse reflectance model implementation using optical parameters applied to the spectral simulation of avocado leaf**

*Nicolás Alberto Molina Cerón, María Constanza Torres Madroño,  
Jully Galeano, María Casamitjana*

### **Evaluation of supervised and semi-supervised classifiers for multispectral image processing: a comparative study**

*Valentina Suescún Arias, María Constanza Torres Madroño,  
Jully Galeano Zea, María Casamitjana*

### **Design and validation of an exoskeleton for hand rehabilitation in adult patients with rheumatoid arthritis**

*Roberto Moya-Jiménez, Teresa Magal-Royo, Michelle Flores,  
Mario Caiza*



**Understanding the Capacitated Vehicle Routing Problem, through the Hamiltonian circuits**

*Saul Figueroa, Franklin De la Cruz, Joseline García, Franklin Camacho*

**Comparative study of PI and fuzzy control techniques for voltage stability of a three-phase inverter in photovoltaic systems, evaluating the voltage in transitory and stable state**

*André Mejía Páliz, Wilson Pavón Vallejos, Leony Ortiz Matos, Ama Owusu-Andanese*

**FPGA-based Low-Cost Multispectral Camera Prototype for Precision Agriculture Applications**

*Julian Uribe-Rios, David Marquez-Viloria, Luis Castano-Londono*

**Outcome Prediction of COVID-19 Patients From Clinical Admission Data Using Machine Learning Models**

*Juan Pablo Cartagena, Rubén D. Fonnegra, Fabián Narváez, Gloria M. Díaz*

# Preface

---

This book contains a group of selected papers presented at Second International Conference on Smart Technologies, Systems and Applications (SmartTech-IC 2021) held on December 1-3, 2021 in Quito, Ecuador. The SmartTech-IC conference aims to attract researchers, scientists and technologists from some of the top companies, universities, research groups, and government agencies from Latin America and around of the world to communicate their research results, inventions and innovative applications in the area of smart science and the most recent smart technological trends. SmartTech-IC 2021 was organized by the Universidad Politécnica Salesiana, a private institution of higher education with social purposes, nonprofit and co-financed from the Ecuadorian State. The SmartTech-IC conference is conceived as an academic platform to promote the creation of technical and scientific collaboration networks. The goal of the conference was addressed some relevant topics related to smart technologies, smart systems, smart trends and applications in different domains in the field of computer science and information systems that represent innovation in current society.

We would like to express our gratitude to all the authors who submitted papers to SmartTech-IC 2021, and our congratulations to those whose papers were accepted. Each submission was reviewed by at least three qualified reviewers chosen from our Program Committee (PC) based on their qualifications and experience. The papers were selected based on a series of criteria including the reviewers average score and comments provided by the reviewers.

We would also like to thank the PC members, who agreed to review the manuscripts in a timely manner and provided valuable feedback to the authors.

December 2021  
Fabián R. Narváez

# Identification of Costa Rican speaker's age group based on acoustic vowel analysis

---

Victor Yeom-Song

University of Costa Rica, San José 11501-2060, Costa Rica

victor.yeom@ucr.ac.cr

<https://orcid.org/0000-0003-4172-1536>

Marvin Coto-Jimenez

University of Costa Rica, San José 11501-2060, Costa Rica

marvin.coto@ucr.ac.cr

<https://orcid.org/0000-0002-6833-9938>

## Abstract

The identification of the speaker's age is a fundamental task in human-computer interaction. In particular, the recognition of children has drawn attention from the research community due to the vulnerability of se persons on the Internet, as well as the potential benefits in digital educational systems. The identification of the age can be determined from the dynamics of the speech information, which has a dependency on language and accent.

This paper presents a study on the identification of Costa Rican adults and children based on acoustic vowel analysis. For this purpose, we chose the five basic vowels of the Castilian Spanish language, and extract a set of acoustic features from segmented speech to explore the accuracy of three classifiers (k-NN, Random Forest and SVM) in distinguishing between adults and children.

The results show promising results in the classification using a single vowel, with the best results from the vowel /u/. This means that an automatic system could be built to identify age as long as enough speech information is received and transcribed, but the accuracy of such a system based on vowels should not be expected to be very accurate in short interactions.

## Keywords

Age recognition, Classification, Vowel analysis.

## Introduction

Speech signals contain non-lexical elements of communication of several kinds, such as the speaker's identity, gender, age, and emotional state [1]. Recently, even the sound of the speech related to SARS-COVID2 has been proposed [2], as a way to support health diagnostics.

Human listeners recognize almost immediately the paralinguistic information at the same time, along with the comprehension of the linguistic content of the speech signals. The case for automatic computer-based systems is quite the opposite, since research has shown the necessity of tackling each problem separately.

Such research has grown considerably over the last two decades, including the analysis and recognition of children's speech [3]. The automatic identification of a child used in human-computer interaction can benefit several application areas, such as the adaptation of the content and responses of personal assistants, child security in communications through the Internet, and personalized educational systems [3].

Due to the differences in speaking styles, current Automatic Speech Recognition (ASR) can benefit from separate training and adjustment according to the user's age. As some authors have shown, ASR systems are not accurate for pre-school children's voices [4].

To develop those systems, it is important to analyze the children's speech in each language and accent, to establish which features can be selected to automatically classify a user's age from the speech information. The challenges in the analysis of children's speech signals began with the proper recording of voices, where proper interaction should be applied for the participation of children.

In this work, we conducted a study on Costa Rican speakers' age classification in two classes: children and adults. We propose to explore the feasibility of performing a proper classification accuracy using acoustic characteristics of vowels, analyzing each one separately. For this

purpose, the transcription and segmentation of speech recordings from children and adults were implemented to extract the audio features from the segments corresponding to vowels. Automatic classification systems were trained using the information of specific vowels to measure their performance on the binary classification task.

## *Related work*

Acoustic analysis of vowels has been largely studied in several languages, to characterize and compare the different accents and forms of verbal communication. Those analyses have also included the comparison between age groups, as well as particular considerations for human-computer interactions [5,6].

In english, the acoustic analysis of parameters such as formants, pitch, and duration of phonemes has been presented first in [7], to characterize children from 5 to 18 years of age. That study can also be considered as the pioneer of the spectral variations and duration in vowels. Recent considerations about the lack of datasets for children's speech have been analyzed in [8].

Apart from the acoustic characterization of vowels, the variability in repetitions has been presented in [9]. In their experiments, researchers found higher variability in the younger population, in terms of vowels' duration, pitch, and formants.

In [4], the acoustic features of children's speech were analyzed as a function of age. The study found that fundamental frequency, duration, and formants of vowels present age-dependent variability, less significant in fundamental frequency but clearly defined in formants and durations.

For the case of the Italian language, a similar study of the changes in acoustic characteristics of children has been presented in [10], confirming the changes in variability in temporal and spectral features. The tendency is to decrease in variability as age increases.

In the Russian language, an analysis of the vowels of children was presented in [11], comparing children's vowels to those from adults. The main differences found in the study were longer duration and higher tones in children.

Recent research in the identification of children and adult speech has made use of pitch and formants, along with several spectral coefficients [12]. With those parameters, the accuracy of classification was above 97%. In the Italian language, the accuracy was higher than 80% according to [13].

For the Portuguese language, a comparison of several classifiers was presented in [14], using Perceptual Linear Prediction coefficients and pitch. The distinction between children and adults had an accuracy as high as 97.4%.

In this work, we present a similar study, extending the application of acoustic analysis of vowels for the case of Costa Rican children's speech, in order to compare several classifiers' performance to detect the voice of an adult or a child. We focus the study on the impact of the number of vowels required to perform the best performance of the classifier, and the best vowel to consider in the analysis. The information provided from this analysis can be incorporated into human-computer systems that contemplate the detection of children in this particular language and accent. The use of vowels can represent a fast response where only speech information is available.

The rest of this paper is organized as follows: Section 2 presents the Background on Acoustic Vowel Analysis and Section 3 presents the Experimental Setup of the work. Section 4 shows the Results and Discussion, and finally, Section 5 presents the Conclusions and future work.

## **Acoustic Vowel Analysis**

Consonants and vowels are the two main classes of speech sounds in Spanish Language. Vowels are speech sounds produced without cons-

triction of the vocal track, and produce voiced sounds. On the contrary, consonants are produced with some blockage of the vocal track, and can produce voiced or unvoiced sounds.

For both consonants and vowels, there are sub-categories that take into account characteristics of the sounds or the vocal track on the emission of the sounds. For example, the five Spanish vowels /a/, /e/, /i/, /o/, /u/ constitute phonemes forming a triangle defined by the features [front, back] and [high, mid, low] [15].

From these categories, the front vowels are /i/, /e/, and the back vowels are /u/, /o/, while /a/ is defined as a central vowel. /i/, /u/ are considered high vowels, /e/, /o/ are mid vowels, and /a/ is low.

The research on the properties of the vowels has also considered that vowels can be more accurately described by their acoustic features, and not only the articulatory specifications. The location of the tongue, lips and mouth represent different sizes and shapes of the vocal cavity, which result in different resonant patterns for each vowel.

The acoustic analysis can be performed with isolated sounds when using proper recordings of each sound. For the purpose of adult or child interaction with technology, the analysis should be implemented in a continuous speech signal, where a segmentation process for detecting temporal frontiers of each sound is done. Fig. 1 illustrates a segmentation process for the detection and acoustic analysis of /a/ vowels.

The acoustic features of vowels usually take into consideration the voiced class of the sounds. For this reason, several statistics of the fundamental frequency of the vowels are often computed, as well as duration.

The resonant differences can be measured using the formants, or characteristic overtones that distinguish each vowel. The first two formants are sufficient to describe a vowel [16].



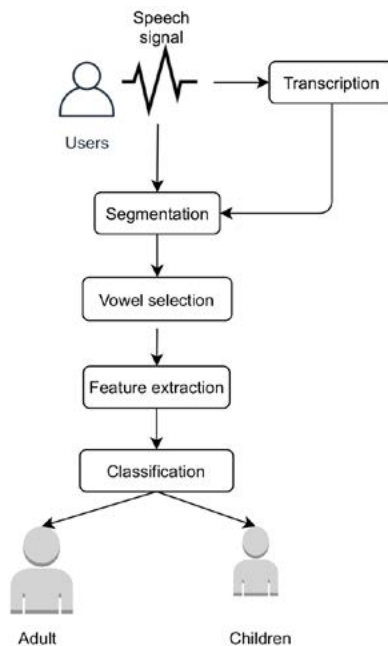


It is important to remark that these kinds of representation of formants are averages, and that vowels are highly context sensitive [16]. For this reason, formant frequencies can be affected based on their linguistic context, and constitute a source of differentiation between categories of native or second language speakers, children and adults, and even gender [17].

## Experimental Setup

The general procedure for the process of the proposed system is illustrated in Fig. [3]. In the following subsections, each of the the main components of the experimental procedure are presented.

**Figure 3**  
**General procedure of the proposed system**



## Database

To obtain the speech information of children and adults, and generate the database consisting on isolated words and their corresponding transcription, several sessions were carried out with four children with ages between 4 and 12 years. The interactions were performed in a professional recording studio using pictographs, which included simple words according to the development of the language of the children, as shown in Fig. 4. The same set of words was repeated for a group of four adults. All the participants are native Costa Rican Spanish speakers.

**Figure 4**  
**Recording session of the children's speech**



The recordings of all the participants were manually edited and then segmented using the Praat system [18], with the EasyAlign extension. With this system, the temporal marks for each isolated sound of the speech can be established, and its corresponding features can be ex-

tracted. For our experiments, we considered the five vowels of the Spanish language: /a/, /e/, /i/, /o/, /u/.

A set of thirty words were collected in each recording, which included 137 /a/, 57 /e/, 33 /i/, 84 /o/ and 25 /u/ from each participant. In the first experiences, the features of a single vowel did not represent significant information to identify a speaker's age. For this reason, we explore statistics of groups of three, five and ten of each vowel. Using this approach, a system should wait until a pre-defined amount of specific vowels are collected from the input speech, and then perform the classification into the adult or children class.

## *Features*

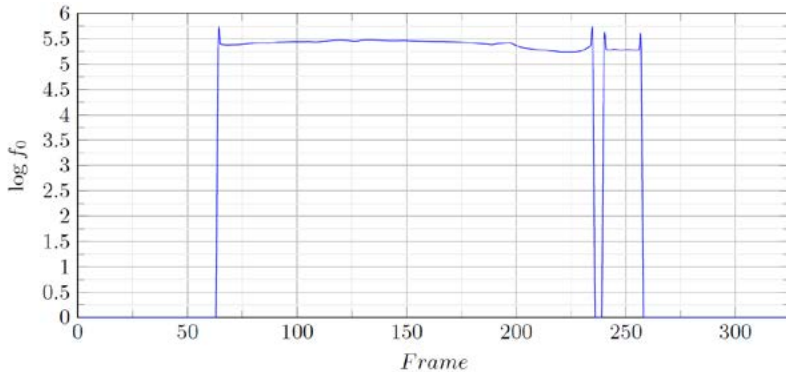
For each vowel, we extract acoustic features, following the availability of software implementations and considering previous references. Given that vowels are voiced sounds, a clear fundamental frequency can be measured for each (with slight variations along the sound), and most of the acoustic analysis takes this feature into consideration. As shown in Fig. 5, a typical recording of a simple word presents segments with positive values of  $f_0$  (voiced sounds), and 0-valued segments (unvoiced sounds).

In our analysis, the following features were extracted from only the voiced sounds corresponding to the five Castilian Spanish vowels, using the Praat/EasyAlign system:

- Average fundamental frequency ( $f_0$ ) in Hz
- Minimum value on the set of vowels  $f_0$  in Hz
- Maximum value of  $f_0$  on the set of vowels, in Hz
- Average duration of the vowel, in seconds.
- Minimum duration of the vowel, in seconds.
- Maximum duration of the vowel, in seconds.
- Average position of the first and second formants of the spectrum (F1 and F2) in Hz.
- Minimum and Maximum of F1 and F2 in Hz.

The implementation of the corresponding algorithms and detailed description of the procedure be found in [19].

**Figure 5**  
**f<sub>0</sub> contour of a word pronounced by a children**



## Classification

The identification task is approached as a supervised learning problem. In this work, we have used three classifiers: a Support Vector Machine (SVM), a Random Forest and a k-Nearest Neighbors (kNN), each one working on distinct mathematical principles. Taking this into account, we may be able to infer characteristics about the acoustic structure of the analyzed vowels through the performance of each classifier; for example, if the kNN classifier performs significantly better than the SVM, this could tell us that the data points are clustered with highly non-linear boundaries between groups. A brief description of the classifiers [20] is presented in the following subsections:

**k-Nearest Neighbors.** The kNN algorithm defines an observed object's class according to its "closeness" to the other data points in the set. This closeness is calculated through a distance function which can vary with the application for which the algorithm is used, and the class assign-

ned to the new observation is defined by the majority of the  $k$  closest data points. Usually an odd value of  $k$  is used to avoid ties in the classification.

In this work, we use the Euclidean distance to calculate the separation of the data points and we use  $k=1$  for the classification of a new instance.

**Support Vector Machines.** The Support Vector Machines constitute a supervised learning algorithm which generates a hyperplane that separates data set into two classes; in the case of vowel classification, for example, the instances may be separated as “is a vowel /a/” or “is not a vowel /a/”. The position of such a hyperplane is set through the maximization of the distance between the hyperplane and a subset of the data used to train the classifier; this subset is what defines the *support vectors* in the algorithm.

**Random Forest.** A Decision Tree classifier takes an input vector, from which it will make a number of checks to generate a final “decision” about the object that the vector describes. This algorithm is described as a tree, in which the nodes represent each check that the algorithm does and the leaves are the possible decisions that the tree can take.

Using a number of decision trees we create an *ensemble*, which takes the decision generated from each tree and combines them to generate a final output; this ensemble is known as a Random Forest. Random Forests have been seen to generate more accurate results than single Decision Trees which could be attributed to a bigger sampling of the dataset.

## Limitations

At the time of writing this work, there are no extensive, well-established databases with Costa Rican voices from children. The work presented here is based on the first advancement to generate such a database, described in Section 3.1. For the difficulties that represent the recording of children, the amount of speakers considered in this study is a limitation on the data available to train the classifiers. Given the number of vowels for each speaker presented in Section 3.1, the simu-

lation of new utterances was performed using random sub-samples of each vowel of size 3, 5, or 10 vowels. Using this procedure, 70 different groups of 3, 5, or 10 vowels were extracted to train the classifiers, and 30 different groups of the corresponding number of vowels for testing.

## Results and discussion

The results are presented according to the classifiers and the number of vowels considered in our experiments: three, five, or ten vowels of the same kind. In Table 1 the results for the /a/ vowel are presented. In this case, the greater amount of vowels represent a better result in the adult/child classification in terms of accuracy, with significant differences from the simplest case of three vowels.

These results present the same pattern in the three classifiers, with the best results of the Random Forest, followed by kNN. The only advantage of the SVM classifier is that the performance is better than the other two classifiers when only three vowels are considered.

**Table 1**  
/a/ vowel classification results. **Bold font indicates best result**

Classifier	Sample Size	Accuracy (%)	Precision	Recall	F1-Score
Random Forest	3	77.10	0.77	0.77	0.77
	5	89.35	0.89	0.89	0.89
	10	96.77	0.97	0.97	0.97
SVM	3	83.23	0.83	0.83	0.83
	5	84.84	0.85	0.85	0.85
	10	94.52	0.95	0.95	0.95
kNN	3	81.60	0.82	0.82	0.82
	5	87.42	0.88	0.87	0.87
	10	96.56	0.97	0.97	0.97



Similar results can be observed in the case of the /e/ vowel, as shown in Table 2 in terms of the best classifier (Random Forest) using 10 vowels. A relevant observation for this vowel is that the differences in accuracy using 3 or 10 vowels tends to be lower than the case of /a/ vowel. The /e/ vowel is the most frequent of the Castilian Spanish language, and for this reason, it is expected to detect and analyze ten /e/ vowels faster than any other in a regular human-computer interaction.

**Table 2**  
**/e/ vowel classification results. Bold font indicates best result.**

Classifier	Sample Size	Accuracy (%)	Precision	Recall	F1-Score
Random Forest	3	87.10	0.87	0.87	0.87
	5	91.29	0.91	0.91	0.91
	10	97.74	0.98	0.98	0.98
SVM	3	87.10	0.87	0.87	0.87
	5	92.26	0.92	0.92	0.92
	10	95.81	0.96	0.96	0.96
kNN	3	86.13	0.87	0.86	0.96
	5	90.00	0.90	0.90	0.90
	10	97.10	0.97	0.97	0.97

The case of /i/ vowel is presented in Table 3. The best result for this vowel is consistent with the previous two vowels. For the SVM classifier, this is the worst case, with a maximum accuracy of 88%, and as low as 80% using three vowels. These results make the choice of the /i/ vowel less convenient for the detection of adult and children using acoustic vowel analysis, adding to the fact that the /i/ vowel is less frequent than the /e/ and /a/.

**Table 3**  
*/i/* vowel classification results. Bold font indicates best result

Classifier	Sample Size	Accuracy (%)	Precision	Recall	F1-Score
Random Forest	3	88.07	0.88	0.88	0.88
	5	90.65	0.91	0.91	0.91
	10	<b>95.81</b>	<b>0.96</b>	<b>0.96</b>	<b>0.96</b>
SVM	3	80.00	0.80	0.80	0.80
	5	82.58	0.83	0.83	0.83
	10	88.06	0.88	0.88	0.88
kNN	3	84.52	0.85	0.85	0.85
	5	87.42	0.87	0.87	0.87
	10	95.48	0.96	0.96	0.96

kNN is the best classifier when using the */o/* vowel, but with significant differences from the acoustic analysis of 10 or 3 vowels, as shown in Table 4. The differences in accuracy do not seem to be significant with the Random Forest classifier, and both classifiers even match the Precision, Recall and F1-score measures.

**Table 4**  
*/o/* vowel classification results. Bold font indicates best result.

Classifier	Sample Size	Accuracy (%)	Precision	Recall	F1-Score
Random Forest	3	86.77	0.87	0.87	0.87
	5	91.94	0.92	0.92	0.91
	10	96.77	<b>0.97</b>	<b>0.97</b>	<b>0.97</b>
SVM	3	81.61	0.82	0.82	0.82
	5	86.77	0.87	0.87	0.87
	10	94.19	0.94	0.94	0.94
kNN	3	82.25	0.82	0.82	0.82
	5	91.29	0.92	0.92	0.92
	10	<b>97.42</b>	<b>0.97</b>	<b>0.97</b>	<b>0.97</b>

The /u/ vowel results are shown in Table 5. The analysis of 10 /u/ vowels with the kNN classifier is the best result among all the cases analyzed. In general, the use of this vowel represents better accuracy, precision and recall than all the other four vowels with any of the classifiers when using 10 vowels. For the case of 3 vowels, the results are also the best with Random Forest.

**Table 5**  
**/u/ vowel classification results. Bold font indicates best result**

Classifier	Sample Size	Accuracy (%)	Precision	Recall	F1-Score
Random Forest	3	91.91	0.92	0.92	0.92
	5	93.23	0.93	0.93	0.93
	10	98.71	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
SVM	3	82.26	0.82	0.82	0.82
	5	89.35	0.89	0.89	0.89
	10	96.45	0.97	0.97	0.97
kNN	3	81.61	0.82	0.82	0.82
	5	90.00	0.90	0.90	0.90
	10	<b>99.35</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>

For the application of a personal assistant, or any other human-computer device that use speech, the possibilities of incorporating acoustic vowel analysis for the detection of adults and children are promising, especially if the various cases of vowels can be taken into consideration. For example, being the /e/ the most frequent vowel, the use of Random Forest or SVM seems to be the best choice for a faster detection. But, as the eventual conversation or interaction develops, new adjustments can be done in terms of considering other vowels, and when 10 /u/ vowels can be collected, the classification can be adjusted or verified with kNN.

## Conclusions

In this work, a proposal to identify children and adults using acoustic vowel analysis was presented, for the case of Costa Rican Castilian Spanish. The analysis requires the transcription of the input speech, the subsequent segmentation at the phoneme level, and the analysis of acoustic features of each vowel.

Three classifiers were applied using the statistical analysis of groups of the same vowel, in order to establish the best vowel and classifier for the task of identifying the speaker's age. According to the results, the best vowel in Costa Rican Castilian Spanish to differentiate the classes is the /u/, using the KNN classifier. With /u/ being the least frequent vowel in the language, some considerations should be taken in the implementation of the findings in any application.

For example, using the most frequent vowels of the language, /e/ and /a/, can let the system establish a classification with high accuracy in a short time with the Random Forest classifier, but further and permanent analysis could be implemented to adjust the results as more speech is detected and transcribed.

For future work we expect to test our proposal with narrow age ranges, to study its functionality in detecting more specific ages. In these cases it is expected that more extensive experimentation with classifiers is mandatory, as well as considering combinations of algorithms to obtain better results.

## References

- [1] Safavi, S., Russell, M., Jancovic, P. Automatic speaker, age-group and gender identification from children's speech. In: Computer Speech & Language vol. 50, pp. 141-156 (2018).
- [2] Imran, A., et al. AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. In: Informatics in Medicine Unlocked (2020).

- [3] Safavi, S., et al. Identification of gender from children's speech by computers and humans. In: Proceedings of INTERSPEECH (2013).
- [4] Yildirim, Serdar, et al. Acoustic analysis of preschool children's speech. Proc. 15th International Congress of Phonetic Sciences, Barcelona, Spain (2003).
- [5] Yeung, G., Abeer, A. A Frequency Normalization Technique for Kindergarten Speech Recognition Inspired by the Role of  $f_0$  in Vowel Perception. In: Proceedings of INTERSPEECH (2019).
- [6] Nagano, T., et al. Data Augmentation Based on Vowel Stretch for Improving Children's Speech Recognition. In: IEEE Automatic Speech Recognition and Understanding Workshop (ASRU). IEEE, (2019).
- [7] Lee, S., Potamianos, A., Narayanan, S. Analysis of children's speech: Duration, pitch and formants. In: Fifth European Conference on Speech Communication and Technology (1997).
- [8] Yeung, G., Ruchao, F., and Abeer, A. Fundamental frequency feature normalization and data augmentation for child speech recognition. In: 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, (2021).
- [9] Gerosa, M., et al. Analyzing children's speech: An acoustic study of consonants and consonant-vowel transition. In: 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings (ICASSP). Vol. 1. IEEE, (2006).
- [10] Gerosa, M., Giuliani, D., Brugnara, F. Acoustic variability and automatic recognition of children's speech. *Speech Communication* 49.10, pp. 847-860 (2007).
- [11] Lyakso, E., Frolova, O., Grigoriev, A. The acoustic characteristics of Russian vowels in children of 6 and 7 years of age. In: Tenth Annual Conference of the International Speech Communication Association (2009).
- [12] Zeng, Y., Yi, Z. Robust children and adults speech classification. In: Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), (2007).
- [13] Massarente, E. Classificazione automatica della voce in ambito logopedico: training e testing di un algoritmo per discriminare la voce adulta da quella dei bambini. [Magistrali biennali] (2015).
- [14] Martins, R., et al. Detection of Children's Voices. *I Iberian SLTech* (2009): 77.
- [15] Celdran, E., Elvira-Garcia, W. Description of Spanish Vowels and Guidelines for Teaching Them. In: Rao, R. (ed.) *Key Issues in the Teaching of Spanish Pronunciation*. Oxford: Routledge, (2019).

- [16] Ronquest, R.E. An acoustic analysis of heritage Spanish vowels. Diss. Indiana University, (2012).
- [17] Albuquerque, Luciana, et al. "A comprehensive analysis of age and gender effects in European Portuguese oral vowels." *Journal of Voice* (2020).
- [18] Goldman, Jean-Philippe. "EasyAlign: an automatic phonetic alignment tool under Praat." In: INTERSPEECH, 12th Annual Conference of the International Speech Communication Association. (2011).
- [19] Boersma, P. Weenink, D. Praat: doing phonetics by computer [Computer program]. Version 6.0.37, retrieved May 2020 from <http://www.praat.org/>.
- [20] Stuart R., Norvig, P., Davis, E.. *Artificial Intelligence: A Modern Approach*, 3rd ed. Upper Saddle River: Prentice Hall, (2010).
- [21] Schuller, B., et al. Covid-19 and computer audition: An overview on what speech & sound analysis could contribute in the SARS-CoV-2 Corona crisis. arXiv preprint arXiv:2003.11117 (2020).

# U-Net Based on MRI Image Patches for Skull Stripping

---

David Jiménez-Murillo<sup>1</sup>

Instituto Tecnológico Metropolitano, Medellín, Colombia  
davidjimenez215489@correo.itm.edu.co  
<https://orcid.org/0000-0001-6335-8657>

Daniel F. Terraza-Arciniegas<sup>1</sup>

Instituto Tecnológico Metropolitano, Medellín, Colombia  
danielterrazza212285@correo.itm.edu.co  
<https://orcid.org/0000-0002-5049-8738>

Leonardo Duque-Muñoz

Instituto Tecnológico Metropolitano, Medellín, Colombia  
leonardoduque@itm.edu.co  
<https://orcid.org/0000-0001-7115-3870>

Juan D. Martínez-Vargas

Instituto Tecnológico Metropolitano, Medellín, Colombia  
juanmartinez@itm.edu.co  
<https://orcid.org/0000-0001-7037-6925>

Andrés E. Castro-Ospina

Instituto Tecnológico Metropolitano, Medellín, Colombia  
andrescastro@itm.edu.co  
<https://orcid.org/0000-0003-3893-1137>

## Abstract

Brain extraction or skull stripping of Magnetic Resonance Imaging is the process of removing non brain tissue, a key task for image processing and analysis and for the diagnosis of several neural pathologies. It is a time demanding and prone to human error task that needs to be fully automated. Several attempts of automation have been found in the literature but there is still room for improvement, specially with the emergence of deep learning algorithms like convolutional neural networks, which have outperformed

---

\* These authors contributed equally to this work.



traditional image segmentation techniques. Moreover, they have the advantage of automatically extracting high level features, improving training results. However, they require a lot of data for training, which is a huge gap because medical images are usually unavailable. The U-Net convolutional neural network was developed for tasks where data availability is an issue, like medical image processing, with outstanding performance. In this study, we propose a 2D U-Net trained with small patches extracted from slices of the original MRI as data augmentation technique, in order to achieve state of the art segmentation results and to be able to perform brain extraction in volumes with different dimensions. The new model predictions were quantitatively compared with those of other segmentation tools. The Dice coefficient score obtained by our proposal against manual segmentation performed by an expert was  $0.967 \pm 0.007$  for the test dataset.

## Keywords

Brain extraction, Image segmentation, Skull stripping, U-Net.

## Introduction

Automated brain extraction has become an increasingly key component in neurological image processing because of its direct impact in subsequent brain analysis. Therefore, its accuracy and efficiency are fundamental for diagnostic purposes. Since manual segmentation is an expensive, prone to human error and time consuming process that involves the labor of an expert radiologist, there is the need of developing a computational method that can match consistently with manual segmentations performed by experts.

Several methods have been proposed for brain extraction, which can be classified as manual, semi-automatic and automatic. Manual segmentation is performed by an expert radiologist with the assistance of a Magnetic Resonance Imaging (MRI) processing software in which the voxels considered to belong to the region of interest are manually outlined. Semiautomatic methods are a mixture between automated algorithms and manual segmentation to achieve better results.

Automated brain extraction has had three different approaches [1]: conventional, machine learning and deep learning based methods. Conventional methods use traditional image processing techniques like active contours [2], intensity based distributions and thresholding [3].

Methods based on machine learning techniques like Model-Based Level Set (MLS) [4], Automated Fuzzy logic based Skull Stripping (AFSS) [5] and more traditional algorithms such as Support Vector Machines (SVM) [6], and Principal Component Analysis (PCA) [7] have also been proposed. Deep learning methods have recently gained popularity because of their high performance in tasks like image classification [8], [9], object detection [10], semantic segmentation [11], among others [12], [13]. These methods have the advantage of being able to learn their own set of high level features directly from training data, thus, eliminating the human factor for this task.

However, traditional machine learning methods require the extraction of features, a process that still depends on the subjectivity of the researcher and is, therefore, prone to error. On the other hand, despite the high performance shown by deep learning algorithms, no model has proved to be the final word on the matter, in part because one of the main gaps in the literature has been the lack of an open, large enough MRI database for training them. This means that there is still room to improve previous results by applying data augmentation techniques and selecting the right algorithms for the task.

The U-Net is a Convolutional Neural Network (CNN) architecture specifically designed for the task of classification of medical images with few learning examples [14]. For the purpose of skull stripping, a few U-Net approaches have been proposed. Hwang et al. [15] proposed a 3D patch U-Net with patches of size  $64 \times 64 \times 64$  voxels extracted from MRI volumes of the Neurofeedback skull stripped repository (NFBS). Two methods have been proposed for fetus brain segmentation, Lou et al. [16] proposed multi stage U-Net with deep supervision, in which a 2D U-Net is trained with MRI slices; and Rampun et al. [17] also proposed a 2D U-Net in which the data augmentation problem was overcome by slice rotation, zooming in and out and vertical and horizontal flipping. Finally, Wang et al. [18] used transfer learning for segmenting non-human primate brains, where the pre-trained model was a slice-based 2D U-Net trained with human MRI scans.

The proposed model and contribution of this paper is a 2D patch based U-Net for skull stripping as data augmentation technique. The patches approach allows the U-Net to take advantage of all the structural information in the volumes instead of the need to resizing and other methods that imply resolution down sampling, while providing a number of images suitable for training a deep learning algorithm and allowing the processing of different MRI volume sizes. Then, we quantitatively review its performance against segmentations performed by other third party software as the FSL Brain Extraction Tool (BET) [19] algorithm, and approaches by other authors as a no patch 2D U-Net and a 3D patch based U-Net [15]. The chosen metric was the Sørensen-Dice similarity coefficient between the achieved segmentations and the manual masks given by an expert in the test dataset.

The obtained results are close to the 3D patch U-Net results which is still the architecture with the best performance (Sørensen-Dice score of  $0.986 \pm 0.001$ ), followed closely by the 2D U-Net with  $0.983 \pm 0.002$ . Our model obtained the lowest Sørensen-Dice score for the 2D patch U-Net with  $0.967 \pm 0.007$  which is still close. However, the proposed model has the advantage that it is possible to escalate it to any size of MRI volume. Each of the above mentioned results presented an improvement over the segmentations performed by the BET algorithm whose Dice score was  $0.912 \pm 0.021$ .

## Background

### *Neurofeedback study NFBS Repository*

The Neurofeedback skull stripped (NFBS) repository is a 125 MRI studies dataset; it includes a T1-weighted structural image, a skull stripped image and the brain mask for each subject. The images have  $256 \times 256 \times 192$  pixels and  $1 \times 1 \times 1$  voxel size in millimeters both in the (Y, Z, X) dimensions. This MRI repository was built from anonymized data of the Nathan Kline Institute Enhanced Neurofeedback study

(NFB) [20]. The participants were 77 women and 48 men in the 21-45 age range. Among them, 66 have been diagnosed with some one or more psychiatric disorders in the past. None of the images displays brain abnormalities or incidental findings according to a certified radiologist.

To define which tissues belong to the brain, the authors rely on the standard proposed by [21]. The following tissues were included as brain: Cerebellum, cerebrum, brain stem, internal vessels and arteries, cerebrospinal fluid (CSF) in ventricles, internal cisterns, and deep sulci. Additionally, not included tissues were: skin, skull, eyes, dura mater, external blood vessels and nerves.

Brain masks were created performing a first segmentation using the BEaST method (Brain extraction based on non local segmentation techniques) [21] with standard parameters and with a previous image normalization. These initial masks were then manually edited by an expert to create a library of priors to feed the BEaST method once again. This procedure was repeated until the output segmentation was considered acceptable.

### *Brain extraction tool*

Brain extraction tool (BET) is an automatic brain extraction algorithm developed by Smith (2002) [19] and is part of the FMRIB Software library (FSL) [22]. According to the author, BET “uses a deformable model that evolves to fit the surface of the brain by the application of a set of locally adaptive model forces” [19]. The algorithm begins by setting a lower and upper threshold based on intensity levels taken from the images histogram, then an estimation of the center of gravity and brain volume are computed. Inside that volume a tessellated sphere surface is initialized and then started to be deformed moving towards the brain edge and taking its form.

BET has a parameter named “Fractional intensity threshold” that ranges between 0 and 1. The default value is 0.5, values lower to 0.5 extract a bigger portion of the brain at the expense of accuracy as parts of

the skull can start to appear in the resulting image. As opposite, values bigger than 0.5 make the segmented brain smaller.

## *U-Net*

The U-Net network is a convolutional neural network developed by Ronnenberger et. al. [14] with the goal of efficiently training a network with only a few available annotated samples. This architecture represents a robust method for the accurate and automatic skull stripping with a consistent performance in high and low resolution volumes [23]. Therefore, the introduced model relied heavily on data augmentation strategies. The U-Net architecture can be divided in two stages that give its characteristic ‘U’ shape. The first stage, known as the contracting path or encoder, has the task of capturing context. The second stage is known as the symmetric expanding path or decoder and allows the network to localize the classification information and its output is the segmented image. This type of networks can be trained with very few images and has outperformed other traditional CNN architectures [14].

The contracting path of U-Net consists of repeated arrangements of “two successive  $3 \times 3$  convolutions followed by a ReLU activation unit and a max-pooling layer” as described by Siddique et al. [24]. In the expanding path the feature map is up sampled with  $2 \times 2$  up-convolutions per stage, then concatenated with the feature map of the corresponding layer in the contracting path. Two consecutive  $3 \times 3$  convolutions and a ReLU activation are next. The resulting segmented image is obtained by applying a  $1 \times 1$  convolution filter to reduce the feature map to the desired number of channels [24].

## *Adaptive Moment Estimation (Adam optimizer)*

According to Kingma et al. [25] Adam is an “algorithm for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments” [25]. Unlike sto-

chastic gradient descent (SGD) in which the learning rate is constant for all weights throughout the whole training process, in the Adam optimizer an adaptive learning rate is computed for every parameter.

The Adam algorithm is based on two variations of the SGD algorithm: the Adaptive Gradient Algorithm (AdaGrad) [26] which computes a learning rate for each parameter; and the Root Mean Square Propagation (RMSProp) [27] that also has a per-parameter learning rate and is adapted based on the rate at which the gradients are changing. Adam uses the first and second moments of SGD, hence its name. The  $N$ -th moment of a random variable is the expected value of that variable to the power of  $n$ :

$$m_n = E[X^n] \quad (1)$$

where  $m_n$  is the  $N$ -th moment and  $X$  is the random variable. The first moment is the mean, the second moment is the uncentered variance. The update rule of the Adam optimizer is given by equation (2) as appeared on the paper [25]:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \varepsilon}} \hat{m}_t \quad (2)$$

where  $\hat{m}_t$  and  $\hat{v}_t$  are vectorized estimates for the first and the second moment respectively, also known as the moving averages;  $\theta$  is the weights vector of the model,  $\eta$  is the learning rate and  $\varepsilon$  is a small scalar, usually  $\varepsilon = 1 \times 10^{-8}$  to avoid division by zero.

### **Binary Cross-entropy (BCE)**

The BCE loss function derives from the concept of entropy, which is a measure of the uncertainty of a given probability distribution  $q(y)$ . Nonetheless, in machine learning the probability distribution  $q(y)$  is unknown a priori. Thus, it is approximated by another distribution  $p(y)$ . That

means, it is assumed that data comes from  $p(y)$  but they really are obeying the distribution  $q(y)$ . This is known as the Binary Cross Entropy between the two distributions and can be written as:

$$H_p(q) = - \sum_{c=1}^c q(y_c) \cdot \log(p(y_c)) \quad (3)$$

The BCE loss function looks for the best distribution  $p(y)$  that minimizes the cross-entropy function. BCE employs the  $N$  training samples to compute the loss function and find the best  $p(y)$ . Thus, the binary cross-entropy function is given by (4):

$$H_p(q) = - \frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - p(y_i)) \quad (4)$$

where  $y$  is the vector of labels and  $p(y)$  the probability that a label belongs to any given class and  $N$  the total of training samples.

### *The Sørensen-Dice similarity coefficient*

The Sørensen-Dice similarity coefficient[28], [29] measures the similarity between two samples as described in equation (5):

$$s = \frac{2C}{A + B} = \frac{2 |A \cap B|}{|A| + |B|} \quad (5)$$

where “A” is the predicted mask, “B” is the manually segmented mask or ground truth and “C” is the amount of intersected voxels or pixels for both images  $C = A \cap B$ . The range of dice coefficient is  $[0,1]$  with 1 representing full similarity between the two masks.

## Experimental Setup

Several U-Net architectures were implemented for this study. First, our proposed patched 2D U-Net architecture trained with patches extracted from slices of the original MRI volume as a data augmentation technique. Then, for comparison purposes, we implemented a patched 3D U-Net architecture already found in the literature, and a standard 2D U-Net.

The implementation of U-Net architectures with patches for data augmentation took several steps which are described below:

### *Input data processing*

The 125 volumes of the NFBS dataset were divided into three subsets: 'train', 'validation' and 'test' with a proportion of 80 %, 10 %, 10 % MRI volumes for each subset, respectively.

To train the 3D patch U-Net a corresponding set of patches (sub-volumes) were extracted from the train and validation datasets. The chosen dimensions for each patch were  $64 \times 64 \times 16$  voxels. The patch extraction was performed by a script which ran through every dimension of the original volume with a 64 voxel stride along the X and Y dimensions, 16 voxel stride along the Z dimension and no padding.

For the patched 2D U-Net architectures, the slices were extracted from the 3D U-Net volume patches since they must be trained with images instead of volumes. A total of 16 images of  $64 \times 64$  pixels were extracted from every patch.

### *Model creation*

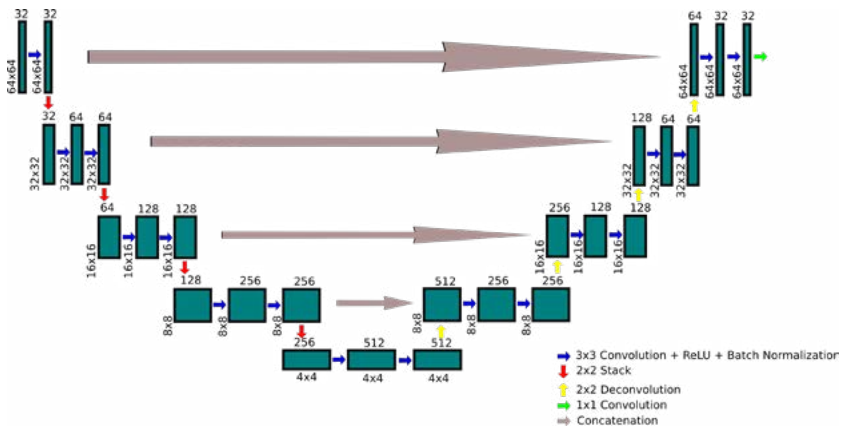
The two U-Net paths, the contracting and expanding paths, were modeled this way: The left path has five convolutional blocks, and each block has 32, 64, 128, 256 and 512 channels respectively as shown in Fig.



1. Each convolutional block is made up of two convolutional layers with a  $3 \times 3$  kernel in the case of both 2D U-Net architectures or  $3 \times 3 \times 3$  kernel for the 3D U-Net. After the convolutional layer follows a ReLU activation function and a max-pooling layer with kernel of size  $2 \times 2$  or  $2 \times 2 \times 2$  except for the final block. To the end of each convolutional layer an additional batch normalization (BN) layer was added.

The expanding path has four convolutional blocks, each block made up of one up convolutional layer with kernel size  $2 \times 2$  and 256, 128, 64 and 32 channels each. Two convolutional layers with  $3 \times 3$  kernel and 256, 128, 64 and 32 respectively, a batch normalization layer was added after each convolution.

**Figure 1**  
**U-Net Architecture**



## Training

The experiments were performed on three different architectures: 2D U-Net, 2D U-Net with patches and 3D U-Net. During the training

stage both architectures were fed with batches of 16 sub-volumes (3D U-Net) or images (2D U-Net). The chosen optimizer algorithm was the Adam optimizer and the loss function was Binary Cross-Entropy. The Sørensen-Dice similarity coefficient was chosen as a metric for validation by implementing equation (5).

The algorithm was set to iterate for 100 epochs with an early stopping criterion of maximum value of the Sørensen-Dice coefficient for the validation set. Training was over if in the lapse of 15 epochs the Sørensen-Dice value was not optimized in at least  $5 \times 10^{-4}$ . Each predicted mask was binarized with a threshold value of 0.5, where values greater than the threshold took the value of 1 and 0 otherwise. The parameters used in each U-Net model are specified on Table 1.

The training stage was performed on a computer equipped with an 8GB VRAM NVIDIA RTX 2060 SUPER GPU, an Intel core i5-5600K CPU and 24GB RAM.

**Table 1**  
**Training parameters of the U-Net architectures**

Architecture	Learning rate	$\beta_1$	$\beta_2$	$\epsilon$	BN momentum
2D U-Net	0.001	0.9	0.999	$1 \times 10^{-7}$	0.9
2D Patch U-Net	0.001	0.9	0.999	$1 \times 10^{-7}$	0.9
3D Patch U-Net	0.0001	0.9	0.999	$1 \times 10^{-7}$	0.9

### *Predictions and volume reconstruction*

For the 3D U-Net architecture, whose segmentation was based on sub-volumes of size  $64 \times 64 \times 16$  voxels, the whole brain volume reconstruction was performed taking into account the stride size parameter upon which the training sub-volumes were extracted. Each one of the predicted sub-volumes was piled next to the other along the X axis until

the original size of the image in this dimension was reached, the same process was performed along the Y and Z axis to achieve a predicted volume for each subject.

The 2D U-Net architecture output were 192 images of size  $256 \times 256$  pixels each. For the predicted volume reconstruction each of the predicted images were piled one after the other along the Z axis until a volume of size  $256 \times 256 \times 192$  voxels was obtained.

The volume reconstruction for 2D patch U-Net architecture predictions was performed as a combination of the methods used for the volume reconstruction of the 2D U-Net and 3D U-Net architectures. Initially,  $64 \times 64$  pixel images were piled up along the Z axis to obtain a  $64 \times 64 \times 16$  sub-volume and then these sub volumes were stacked one next to the other and then one after the other in a similar way to the 3D U-Net reconstruction process to get a reassembled volume of size  $256 \times 256 \times 192$  voxels.

### *BET algorithm segmentations*

The reference brain segmentations for comparison against our model were obtained by running the Nipype python bindings [30]. For the BET algorithm, we tuned the “Fractional intensity threshold” parameter from 0.1 to 0.5 with a step of 0.1. We choose  $frac = 0.2$  because that value provided the best tradeoff between a bigger portion of the brain and absence of skull based on visual examination.

## **Results**

The Sørensen-dice coefficient results obtained with the three u-net architectures after reconstructing the predicted volumes are shown in Table 2. Each predicted brain volume was compared against the manually segmented mask included in the NFBS dataset.

**Table 2**  
**Sørensen-Dice coefficient score obtained by the U-Net architectures**  
**and the BET algorithm against the manual segmentation.**

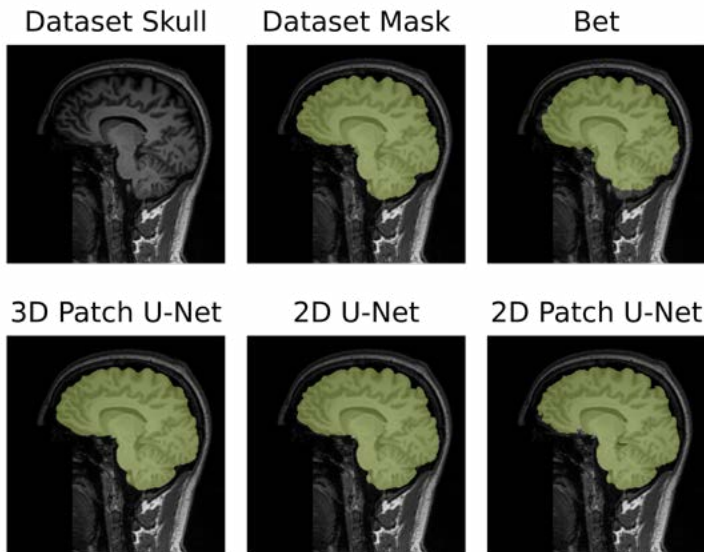
Sample	BET	3D Patch U-Net	2D U-Net	2D Patch U-Net
A00058503	0.926	0.987	0.986	0.972
A00063008	0.915	0.987	0.986	0.968
A00061387	0.924	0.985	0.978	0.952
A00058952	0.899	0.987	0.986	0.969
A00058552	0.908	0.983	0.986	0.958
A00055763	0.882	0.984	0.983	0.967
A00055447	0.904	0.986	0.983	0.972
A00054504	0.892	0.988	0.987	0.977
A00051539	0.950	0.988	0.985	0.962
A00043722	0.888	0.984	0.983	0.963
A00040628	0.917	0.988	0.986	0.969
A00038642	0.942	0.988	0.987	0.972
Total	$0.912 \pm 0.021$	$0.986 \pm 0.001$	$0.983 \pm 0.002$	$0.967 \pm 0.007$

The best performance was achieved by the 3D patch U-Net with a Sørensen-Dice coefficient of  $0.986 \pm 0.001$  among the subjects in the test dataset. Despite the lower quantitative performance of the 2D patch U-Net in relation to the other two architectures, the difference is not meaningful and cannot be perceived. However, there is a notable difference between the proposed models when reviewed against the mask obtained using the BET algorithm as shown in Fig. 2 where the BET algorithm fails to include a considerable portion of the cerebellum.

It is worth noticing in Fig. 2 how the 3D Patch U-Net is smoother in the edges of the manual segmentation because it has more spatial information, also the 2D U-Net has a smoother performance because it has the who-

le slice. Our proposal is the sharpest around the edges. However, with the 2D Patch U-Net we have a trade between performance and flexibility because is not restricted to a one volume size allowing its use in different datasets with different volume sizes, without the need of retraining the network.

**Figure 2**  
**Projection of each model segmented brain over the original head**  
**in the sagittal plane. Sample A00038642**



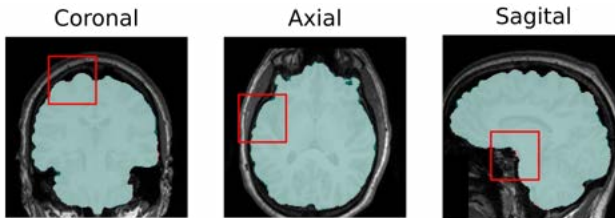
In Table 3 we compared the average of the Sørensen-Dice coefficient between the predicted brain masks for each considered model. In this table is worth noticing the similarity between the 3D patch U-Net and the 2D U-Net segmentations as expected for the two best performing algorithms. The 2D patch U-Net segmentations is quantitatively similar to the best performing algorithms with Sørensen-Dice coefficient 0.967 against the 3D U-Net and 0.972 against the no patch 2D U-Net, emphasizing that there is no clear difference between both pairs.

**Table 3**  
**Sørensen-Dice coefficient comparison**  
**between considered models**

Architecture	BET	3D Patch U-Net	2D U-Net	2D Patch U-Net
BET	1	0.913	0.911	0.895
3D Patch U-Net	-	1	0.984	0.967
2D U-Net	-	-	1	0.972
2D Patch U-Net	-	-	-	1

The projections of the predicted brain masks on the coronal, axial and sagittal planes can be seen in Fig. 3. In this image there are two overlapping brain masks, first the dataset mask with pink color and over it, the 2D Patch U-Net mask in blue. In this image is evident that the proposed 2D patch U-Net algorithm outline is almost identical to that of the manual segmentation in all of the MRI planes. The small regions delimited by red boxes show minimal differences and, in order to see them, it is necessary a large zoom on it, that difference could be associated with human error. Across all samples from test dataset, 2D patch U-Net performance was similar and no visual difference was noticed in the graphic results, confirming that our proposal generalizes enough in order to feed it with unknown data samples.

**Figure 3**  
**2D patch U-Net skull stripped segmentation views in each MRI**  
**plane projected on top of the original volume and the manual mask**  
**as reference. Sample A00038642**



## Conclusions

This work presents a variation in the 2D U-Net architecture with patches for brain segmentation and compares the results with other U-Net architectures and one traditional skull stripping method. Here, the U-Net based approaches outperformed BET achieving at least 5.5 % more with our proposal U-Net 2D patches. Besides, the U-Net does not require the tuning of any parameters as opposed to BET and other algorithms. As we proved in our research, U-Net architecture has better performance than the methods we compare against, like manual, other U-Net based methods and the BET algorithm.

Several reasons led us to take the patches approach to train our network. First, the patch approach allowed the augmentation of the dataset to a number of samples suitable to train a CNN, and thus, generalizing a decision boundary. Also, the use of patches in the 2D and 3D U-Net architecture is beneficial in terms of memory because it allows to train the network minimizing the risk of a memory bottleneck. Finally, the patches approach allows to train the network and make predictions on subjects that belong to different MRI datasets with different volume sizes, an aspect that is key for developing a more general model.

Several limitations were found as this study was developed. For example, there is a critical absence of open MRI datasets of the head with a number of volumes that suits the needs of the more data hungry deep learning algorithms, that is a big gap because a considerable amount of the research time is spent getting together a robust dataset for model creation.

A related limitation is that due to the absence of open datasets, the reproducibility of other works in the literature in order to obtain similar results or recreating models for comparisons, makes this an almost impossible task and narrows the scope of the research.

As future work a more robust process for data augmentation with other techniques should be applied. Also, in pursuit of a more general mo-

del, the network can be trained with volumes from different datasets, and with different training techniques like transfer learning from similar models. This work should be used as a starting point for the development of an integral brain analysis pipeline, for different neurological pathologies, in which the skull stripping is part of the preprocessing stage.

## Acknowledgments

This work is supported by Instituto Tecnológico Metropolitano (*project P20214*) and Daniel F. Terraza-Arciniegas is supported by jóvenes investigadores Minciencias convocatoria 891.

## References

- [1] A. Fatima, A. R. Shahid, B. Raza, T. M. Madni, and U. I. Janjua, “State-of-the-Art Traditional to the Machine- and Deep-Learning-Based Skull Stripping Techniques, Models, and Algorithms,” *J. Digit. Imaging*, vol. 33, no. 6, pp. 1443–1464, 2020.
- [2] Y. Wang, J. Nie, P.-T. Yap, F. Shi, L. Guo, and D. Shen, “Robust Deformable-Surface-Based Skull-Stripping for Large-Scale Studies,” *Lect. Notes Comput. Sci.*, pp. 635–642, 2011.
- [3] B. D. Ward, “Intracranial segmentation,” *Biophys. Res. Institute, Med. Coll. Wisconsin, Milwaukee, WI*, 1999.
- [4] A. H. Zhuang, D. J. Valentino, and A. W. Toga, “Skull-stripping magnetic resonance brain images using a model-based level set,” *Neuroimage*, vol. 32, no. 1, pp. 79–92, 2006.
- [5] S. Kobashi *et al.*, “Fuzzy-ASM Based Automated Skull Stripping Method from Infantile Brain MR Images,” in *2007 IEEE International Conference on Granular Computing (GRC 2007)*, 2007, p. 632.
- [6] J. Sjölund, A. E. Järlideni, M. Andersson, H. Knutsson, and H. Nordström, “Skull Segmentation in MRI by a Support Vector Machine Combining Local and Global Features,” in *Proceedings of the 2014 22nd International Conference on Pattern Recognition*, 2014, pp. 3274–3279.
- [7] X. Han *et al.*, “Brain extraction from normal and pathological images: A joint PCA/Image-Reconstruction approach,” *Neuroimage*, vol. 176, pp. 431–445, 2018.



- [8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
- [9] C. Szegedy *et al.*, "Going Deeper With Convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [10] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587.
- [11] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.
- [12] S. M. Plis *et al.*, "Deep learning for neuroimaging: a validation study," *Front. Neurosci.*, vol. 8, 2014.
- [13] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*. 2017.
- [14] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Lect. Notes Comput. Sci.*, pp. 234–241, 2015.
- [15] H. Hwang, H. Z. U. Rehman, and S. Lee, "3D U-Net for Skull Stripping in Brain MRI," *Appl. Sci.*, vol. 9, no. 3, p. 569, 2019.
- [16] J. Lou *et al.*, "Automatic Fetal Brain Extraction Using Multi-stage U-Net with Deep Supervision," *Mach. Learn. Med. Imaging*, pp. 592–600, 2019.
- [17] A. Rampun, D. Jarvis, P. Griffiths, and P. Armitage, "Automated 2D Fetal Brain Segmentation of MR Images Using a Deep U-Net," *Lect. Notes Comput. Sci.*, pp. 373–386, 2020.
- [18] X. Wang *et al.*, "U-net model for brain extraction: Trained on humans for transfer to non-human primates," *Neuroimage*, vol. 235, p. 118001, 2021.
- [19] S. M. Smith, "Fast robust automated brain extraction," *Hum. Brain Mapp.*, vol. 17, no. 3, pp. 143–155, 2002.
- [20] B. Puccio, J. P. Pooley, J. S. Pellman, E. C. Taverna, and R. C. Craddock, "The preprocessed connectomes project repository of manually corrected skull-stripped T1-weighted anatomical MRI data," *Gigascience*, vol. 5, no. 1, 2016.
- [21] S. F. Eskildsen *et al.*, "BEaST: Brain extraction based on nonlocal segmentation technique," *Neuroimage*, vol. 59, no. 3, pp. 2362–2373, 2012.

- [22] M. Jenkinson, C. F. Beckmann, T. E. J. Behrens, M. W. Woolrich, and S. M. Smith, "FSL," *Neuroimage*, vol. 62, no. 2, pp. 782–790, 2012.
- [23] L.-M. Hsu *et al.*, "Automatic skull stripping of rat and mouse brain MRI data using U-net," *Front. Neurosci.*, vol. 14, p. 935, 2020.
- [24] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications," *IEEE Access*, vol. 9, pp. 82031–82057, 2021.
- [25] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv Prepr. arXiv1412.6980*, 2014.
- [26] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for on-line learning and stochastic optimization," *J. Mach. Learn. Res.*, vol. 12, no. 7, 2011.
- [27] T. Tieleman and G. Hinton, "Lecture 6.5-rmsprop, coursera: Neural networks for machine learning," *Univ. Toronto, Tech. Rep.*, 2012.
- [28] T. J. Sørensen, *A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons*. I kommission hos E. Munksgaard, 1948.
- [29] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [30] K. Gorgolewski *et al.*, "Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python," *Front Neuroinform*, vol. 5, 2011.

# Python application for analysis of photovoltaics cells operating in both first and second quadrant<sup>1</sup>

---

Bonie J. Restrepo-Cuestas

Instituto Tecnológico Metropolitano, Medellín, Colombia

Universidad Nacional, sede Medellín, Colombia

bonierestrepo@itm.edu.co

<https://orcid.org/0000-0001-5276-1651>

Cristian Guarnizo-Lemus

Instituto Tecnológico Metropolitano, Medellín, Colombia

cristianguernizo@itm.edu.co

<https://orcid.org/0000-0002-3853-7976>

Adriana Trejos

Instituto Tecnológico Metropolitano, Medellín, Colombia

adrianatrejos@itm.edu.co

<https://orcid.org/0000-0001-8843-0918>

Carlos A. Ramos-Paja

Universidad Nacional, sede Medellín, Colombia

caramos@unal.edu.co

<https://orcid.org/0000-0003-2231-4177>

## Abstract

Typically, the electrical behavior of photovoltaic (PV) cells is studied when it behaves as a source of energy (direct mode); this means that the voltage and current in their terminal are positive (first quadrant). However, the impact of shadowing a photovoltaic cell in an array causes mismatch losses and provokes the cell to operate in the second quadrant (reverse mode) and can be exposed to overcurrents. Such conditions may cause irreversible damage to the PV cells and are also responsible for the power reduction delivered by the PV system. Under this scenario, it is vital to analyze and evaluate the electrical behavior in direct and reverse operation modes, so this work proposes the

---

1 Supported by Instituto Tecnológico Metropolitano- ITM

development of an educational application in Python for studying the electrical PV cell operation in the first and second quadrants. The software allows representing the effects of irradiance and temperature in the behavior of the current-voltage (I-V) and power-voltage (P-V) curves, using Bishop's model as a circuitual representation of the PV cell. In addition, the GUI allows observing how the variation in each parameter of the Bishop model affects the resulting I-V and P-V curves.

## Keywords

Solar cell, I-V curve, Bishop's model.

## Introduction

Photovoltaic (PV) energy is one of the most promising options to electricity generation in urban and rural areas. The additions in the installed PV capacity in 2020 reached 107 GW and it is expected to keep on growing in an average of 125 GW between 2021 and 2025 [1]. The present environmental and social issues lead to consider PV energy as an important research area; moreover, it is becoming a topic included in engineering programs related to energy among with renewable energy in general [16, 12]. One of the most important issues in PV systems operation is partial shading. Such a condition does not only reduce the power generation but also provokes the shaded cells to operate with negative voltages (second quadrant) which in turn causes the cell to consume power instead of producing it [29]. Moreover, when the cell operates in the second quadrant, its temperature increases causing hot-spots, degradation or irreversible damage [20]. In this way, considering partial shading conditions in the operation of PV systems is important to obtain a more realistic analysis.

Circuit models have been widely used to represent the behavior of the PV cells and modules. Among the most reported models are the single diode model (SDM) and the double diode model (DDM) [26], which provide a good tradeoff between accuracy and complexity. However, such models do not represent properly the behavior of the cell when it is exposed to partial shading and operates in the second quadrant. On the other hand, Bishop's model [9, 21] allows to represent the behavior

of the cell in the second quadrant, but it presents a high complexity at a mathematical level since it includes several unknown parameters and strong non-linear relationships between currents and voltages.

PV circuital models are studied using algorithms based on mathematical equations to relate currents and voltages or employing computational tools. Several works have reported the use of circuital simulations to analyze PV systems. Most of them use well-known software such as Simulink or PSIM [3] aiming to obtain the current-voltage (I-V) and power-voltage (P-V) curves of cells and modules. In [25, 27, 4] the SDM is used to perform different analysis in Simulink. The analysis consists of studying the performance of cells under different radiation and temperature conditions.

On the other hand, some works include graphical user interfaces (GUIs) to provide an intuitive and handy tool. Such is the case of [13], where a Matlab based tool to perform technical and economic analysis for rooftop solar PV systems is introduced. The tool allows to include aspects such as electricity tariffs, battery performance, degradation of the system, among others, to obtain a forecast of the solar power generation; the analysis and computation are based on equations not related to any circuital model. Other reported tools which use graphical interfaces are aimed to study the behavior of cells and modules under different operating conditions. In [15], the SDM model is used to develop a GUI supported in Simulink, for teaching aspects concerning the monitoring and the diagnosis of PV arrays. The contribution is the presentation of an algorithm to detect and localize the fault in a photovoltaic generator using a limited number of voltage sensors. The SDM model is also used in [2] and [18]; both works present GUI for studying the PV module characteristics, i.e., (I-V) and (P-V) curves under different values of radiation and temperature. However, in [18], the GUI can provide solar PV module information including maximum power, voltage and current at the maximum power point, efficiency, fill factor (FF) of the PV module, when parameters such as, module temperature ( $T$ ), solar

radiation ( $I_{rr}$ ), short circuit current ( $I_{sc}$ ), open-circuit voltage ( $V_{oc}$ ), and module area are provided. In [8] a photovoltaic system modeling and characterization software based on a Matlab GUI and Simulink is introduced. In this work, the DDM is used. The first objective is to obtain the parameters of the nonlinear I-V equation by adjusting the curve regarding three particular points: open circuit, maximum power, and short circuit; the tool supports the analysis of arrays in SP configuration considering partial shading conditions. In this paper, we adopted Python as the programming language because it is open source, has a large and growing community that develops libraries and it is highly used by Academia and Industry [23], e.g. SunPower is developing their libraries using Python [19].

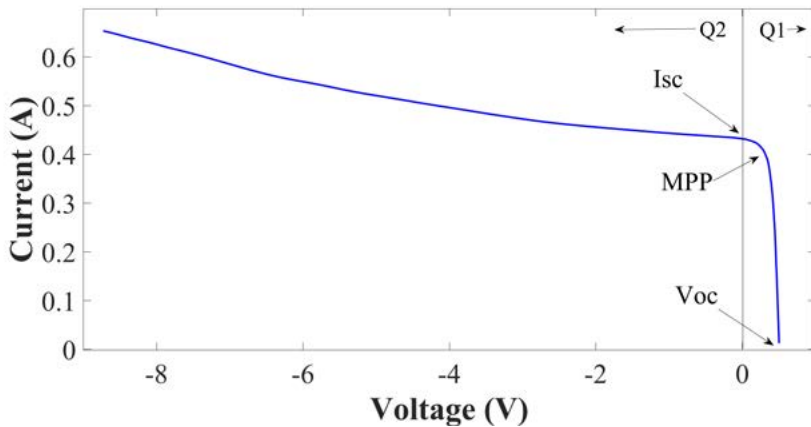
As the literature review shows, most of the works use the SDM to perform their analysis. Only one work uses the DDM, and, to the best knowledge of authors, there is not a work based on Bishop's model. Therefore, this paper introduces a Python GUI-based tool to analyze the behavior of PV cells using Bishop's model. The application has the advantage that the user does not need to use mathematical manipulations or iterative techniques to solve the equation that relates the current and the voltage cell's outputs to understand, appropriately, the effects of partial shading operating conditions. Moreover, the interface provides the user a didactical way to learn the basics of PV cells operation and analyze the behavior of a cell in the second quadrant (affected by mismatching). In addition, it allows evaluating the impact of the variation of each parameter of Bishop's model on the behavior of the I-V and P-V curves. This paper is organized as follows: Section 2 explains the main aspects of Bishop's model. Section 3 describes the GUI. Section 4 provides an analysis of the effect of the parameters in PV cell representation, and Section 5 gives the conclusions of the work.

## Bishop's Model for photovoltaic cells

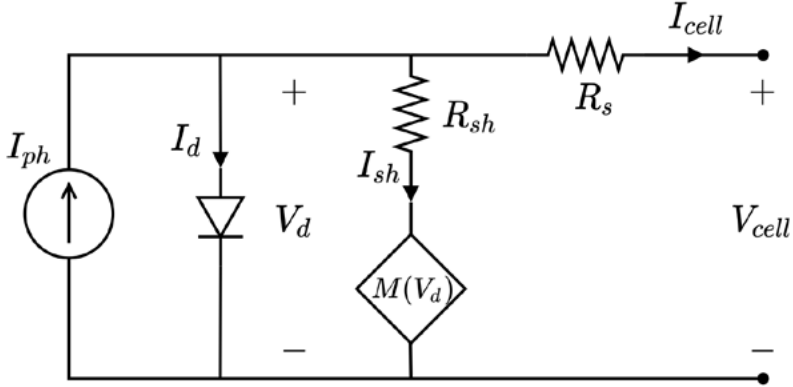
Bishop's model takes the single diode model (SDM), representing the behavior in the first quadrant, and affects the shunt current through an avalanche factor responsible for the behavior in the second quadrant [7, 6, 17]. Figure 1 shows a typical I-V curve for a monocrystalline cell. Here, the first quadrant refers to positive voltage values at cell terminals; this indicates that the cell produces energy. On the contrary, when the cell has a negative voltage at its terminals, it consumes power. This phenomenon can occur in a cell's string when one of the cells is subjected to partial shading [28].

Figure 2 shows the circuit model proposed by Bishop [9]. Here the current source represents the photovoltaic current  $I_{ph}$ , the diode represents the junction of the semiconductor material, the losses due to contact resistance are represented through the series resistance  $R_s$ , and the shunt resistance  $R_{sh}$  is associated with the leakage currents at the edges of the cell.

**Figure 1**  
**First and second quadrant of the I-V curve for a photovoltaic cell**



**Figure 2**  
**Circuit representation of Bishop's model**



The cell output current  $I_{cell}$  is equal to the sum of currents  $I_{ph}$ ,  $I_d$  and  $I_{sh}$ , as shown in eq. (1):

$$I_{cell} = I_{ph} - I_d - I_{sh} \quad (1)$$

The diode current-voltage relationship ( $I_d$ ,  $V_d$ ) can be represented by Shockley's equation as [10],

$$I_d = I_o \left( e^{\left( \frac{V_{cell} + I_{cell} R_s}{n V_{Tn}} \right)} - 1 \right) \quad (2)$$

where  $I_o$  represents the saturation current,  $n$  the ideality factor, and  $V_{Tn}$  is the junction voltage. To represent the behavior in the second quadrant (Q2), Bishop adopts an avalanche factor that affects the current through the shunt resistor [24]. Thus, the  $I_{sh}$  is described by

$$I_{sh} = \frac{V_{cell} + I_{cell} R_s}{R_{sh}} \left\{ 1 + a \left( 1 - \frac{V_{cell} + I_{cell} R_s}{V_{br}} \right)^{-m} \right\} \quad (3)$$



where  $a$  is an Ohmic fraction of the current related to the breakdown of the semiconductor;  $V_{br}$  is the breakdown voltage, and  $m$  is the avalanche exponent. This model requires 8 parameters to evaluate the I-V curve of a cell, which are:  $I_{ph}$ ,  $R_s$ ,  $R_{sh}$ ,  $I_o$ ,  $n$ , (same as in the SDM model for Q1) and  $V_{br}$ ,  $a$ , and  $m$ , (to represent Q2) [30].

From (3) it is observed that  $V_{cell}$  and  $I_{cell}$  have an implicit relationship. Therefore, its evaluation requires iterative techniques such: Newtho-Raphson [22, 14, 11]; or a transformed version as in the case of the Lambert-W function [5].

## Interface development

The software package, named CellSim, is coded in Python using the libraries WxPython, Numpy and Matplotlib. WxPython allows building the GUI using text boxes, buttons and notebooks. The I-V cell relationship is calculated using Numpy (linear algebra library) utilizing the iterative Newton-Raphson algorithm. Voltage and current variables are evaluated in a vector manner to improve the calculation time, from eq. (1) as

$$f(I_{cell}, V_{cell}, \theta) = I_{ph} - I_d - I_{sh} - I_{cell} = 0 \quad (4)$$

where  $\theta = \{I_{ph}, R_s, R_{sh}, I_o, n, a, V_{br}, m, T, I_{rr}\}$  is the set of parameters required for the evaluation of the currents. Finally, Matplotlib is a library to plot data, in this case, it is used to plot the I-V and P-V curves.

In order to analyse how a parameter affects the I-V or P-V curves, the cell's current is estimated first,  $I_{cell}$ , using the Newton Raphson method as described in algorithm 1. This process requires as inputs the set of parameters,  $\theta$ , that fully characterize the solar cell. In line 10 of algorithm 1, the tolerance value of  $10^{-3}$  was obtained through experimental analysis. This value guarantees an adequate convergence and a low computational cost.

Then, it is possible to modify one of the model parameters in a range with a selected number of steps. The GUI returns the family of curves corresponding to the scope and chosen steps. Thus, it is possible to visualize the effect of each parameter on the I-V and P-V curves.

Bishop's model was used to recreate the behavior of a monocrystalline silicon cell with the characteristics given in table 1. The parameters shown in the table 2 are estimated from the curve using Bishop's model.

Algorithm 1 Curve generation based on Newton Raphson	
1:	<b>Input:</b> Parameters $\theta = \{I_{ph}, R_s, R_{sh}, I_o, n, a, V_{br}, m, T, I_{rr}\}$
2:	<b>Input:</b> Voltage range's parameters, initial voltage $V_{ini}$ , final voltage $V_{fin}$ , and number of steps $N_v$ .
3:	<b>Input:</b> Parameter range, initial value $p_{ini}$ , final value $p_{fin}$ , and number of steps $N_p$ .
4:	Generate the voltage vector $\mathbf{V}$ from the voltage range parameters.
5:	Generate Parameter vector $\mathbf{p}$ from the parameter range values.
6:	<b>For</b> $p \in \mathbf{p}$ <b>do</b>
7:	Update $\theta$ according to the value $p$ .
8:	Initialize $\mathbf{I}_t$ as a vector of zero values.
9:	$\mathbf{I}_p = f(\mathbf{I}_t, \mathbf{V}, \theta)$ .
10:	while $ \mathbf{I}_p  \leq 10^{-3}$ do
11:	$\mathbf{I}_t = \mathbf{I}_t - \mathbf{I}_p / \nabla f(\mathbf{I}_t, \mathbf{V}, \theta)$ .
12:	$\mathbf{I}_p = f(\mathbf{I}_t, \mathbf{V}, \theta)$ .
13:	<b>end while</b>
14:	$\mathbf{P} = \mathbf{V} \odot \mathbf{I}_t$ . Element-wise product.
15:	Plot I-V and P-V curves.
16:	<b>end for</b>

**Table 1**  
**Characteristic of the case study**

Parameters	Values
Reference solar irradiance	1000 W/m <sup>2</sup>
Cell reference temperature ( $T_r$ )	47.5°
Short-circuit current ( $I_{sc}$ )	0.68 A
Open Circuit Voltage ( $V_{oc}$ )	0.51 V
Maximum power	0.25 W
Voltage at maximum power point ( $V_{mpp}$ )	0.42 V
Current at maximum power point ( $I_{mpp}$ )	0.61 A

**Table 2**  
**Default Bishop model's parameters adopted for CellSim**

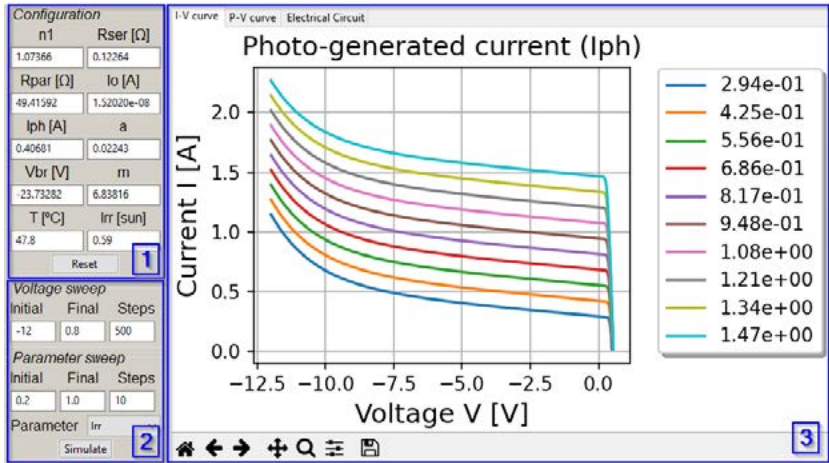
Parameters	Values
Series Resistance ( $R_s$ )	1.073 $\Omega$
Shunt Resistance ( $R_{sh}$ )	0.1226 $\Omega$
Saturation current ( $I_o$ )	$1.52 \times 10^{-8}$ A
Photo-generated current ( $I_{ph}$ )	0.68 A
Fraction of ohmic current ( $a$ )	0.0224
Breakdown voltage ( $V_{br}$ )	-23.73 V
Avalanche breakdown exponent ( $m$ )	6.83

## Features of the GUI design

The graphical user interface (GUI) for CellSim is divided into three regions, as shown in figure 3. In the first region, Bishop's model parameters are presented, as well as temperature and irradiance. Here, the irradiance is measured in sun units, this means that 1 sun is equal to 1000W/m<sup>2</sup>. The default values loaded by the platform are shown in Tables 1 and 2. The user

can modify the values for each parameter. If the user requires to return to the default values, it is necessary to press the Reset button.

**Figure 3**  
**Interface description of CellSim software**



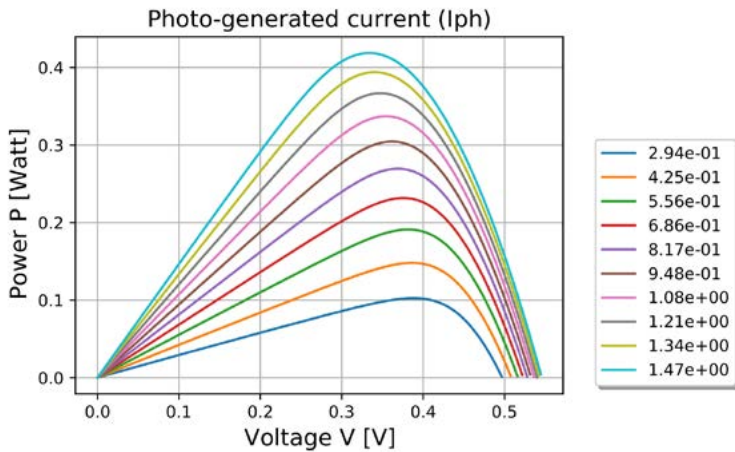
1) Parameters configuration. 2) Definition of Voltage and Parameter ranges for simulation. 3) I-V and P-V curves plotting area.

For the second region, the user sets the initial and final voltage values and the number of steps required to sweep and reconstruct the I-V and P-V curves. The reconstruction of both curves is carried out in both the first and second quadrants.

The user can also select one of the parameters indicated in region 1 to vary it in a defined range (steps, initial and final value), which allows evaluating the impact of the variation of each parameter on the behavior of both curves (I-V and P-V). Figure 4 shows variations in P-V curve when the photo-generated current,  $I_{ph}$ , varies between 0.35 A and 0.55 A, in 5 steps. The toolbar located at the bottom of figures 3 and 4, allows to change some characteristics of the plotted data. Table 3 shows the action that each icon executes.

In the third region, the user has access to the plot area that contains a notebook with three pages. The first page presents the I-V curve (as shown in figure 3), while the second page is for the P-V curve, which is depicted in figure 4. The last page shows the circuitual representation of Bishop’s model and the equation that relates  $I_{cell}$  and  $V_{cell}$  (see Figure 6).

**Figure 4**  
**Page 2: Power - voltage curve**

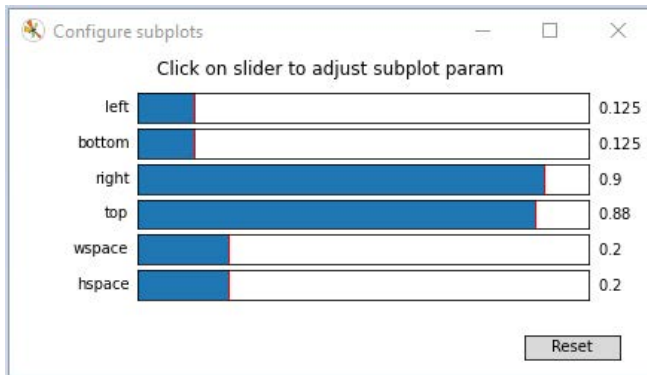


**Table 2**  
**Toolbar menu for I-V and P-V curves**

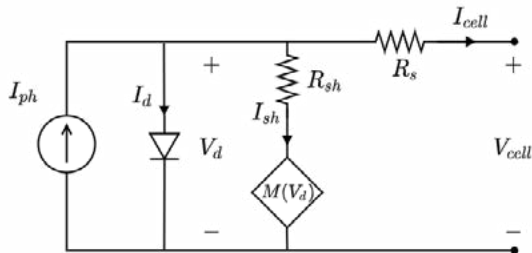
Button	Action
	Reset original view
	Back to previous view
	Forward to next view
	Pan
	Zoom
	Configure subplots
	Save current figure

When the user selects the “configure subplots” option, a new window similar to Figure 5 is displayed. This window allows the user to modify the position of the current figure (left, bottom, right and top). Besides, It allows adjusting the fractions of width and height of the axis (wspace and hspace).

**Figure 5**  
**Configuration subplots**



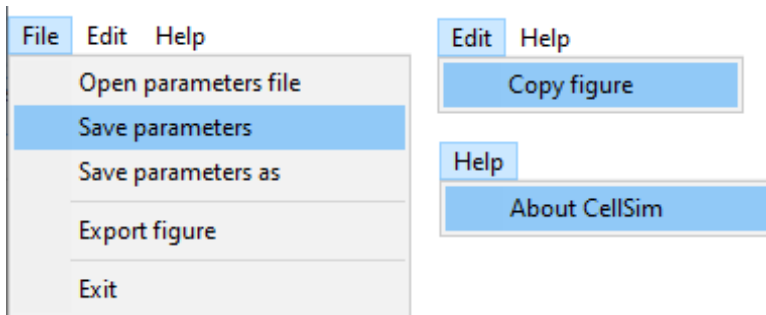
**Figure 6**  
**GUI - Page 3: Electrical circuit and equation of Bishop’s model**



$$I_{cell} = I_{ph} - I_o * \left( e^{\frac{V_{cell} + I_{cell} * R_s}{n_1 * V_{tn}}} - 1 \right) - \frac{V_{cell} + I_{cell} * R_s}{R_{sh}} \left( 1 + a \left( 1 - \frac{V_{cell} + I_{cell} * R_s}{V_{br}} \right) \right)^{-m}$$

Figure 7 illustrates the different actions found within each menu item. The file menu includes the option to save and load the cell's parameter values from files. Besides, the I-V or P-V curve figures can be exported (the currently selected figure in the notebook) as PNG, PDE, PS, EPS and BMP formats. For the Edit menu, the selected figure can be exported to the clipboard. Finally, in the Help menu, it is the option "About CellSim" that opens a window with a description of the developers and the license. CellSim software is available at <https://github.com/cdguarnizo/cellsim>.

**Figure 7**  
**Description of File, Edit and Help menus**



## Conclusions

This paper presented a Python - GUI application named CellSim based on Bishop's model to evaluate solar cell's characteristics. This tool aids in the analysis, design, and education of PV technology. CellSim simulates the characteristics of PV cells and showed the effects of each parameter of Bishop's model, temperature, and irradiation in the reconstruction of I-V and P-V curves. The interaction with the developed platform allows to understand concepts of a PV cell as the maximum power, short-circuit current, and open-circuit voltage. Also, it enables to analyze effects such as partial shading which causes the solar cells to

operate in the second quadrant. Future work can consider implementing a GUI that allows evaluating the impact of partial shading in photovoltaic arrays.

## References

- [1] Solar PV - Renewables 2020 (2020), <https://www.iea.org/reports/renewables-2020>
- [2] Almaktoof, A., Shebani, N., Elfallah, A.: Gui-pv application tool for teaching performance of PV system using matlab-graphical user interface environment. In: 2021 IEEE 1st International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering MI-STA. pp. 447-451. IEEE (2021). <https://doi.org/10.1109/MI-STA52233.2021.9464367>
- [3] Arroyave-Berrio, A.A., Goez-Mora, M.M., Arango-Zuluaga, E.I., Ramos-Paja, C.A., Serna-Garcés, S.I.: Plataforma de modelado y simulación de sistemas de generación fotovoltaicos. *Tecnológicas* (2013). <https://doi.org/http://dx.doi.org/10.22430/22565337.365>
- [4] Banik, A., Shrivastava, A., Potdar, R.M., Jain, S.K., Nagpure, S.G., Soni, M.: Design, modelling, and analysis of novel solar pv system using matlab. *Materials Today: Proceedings* (2021)
- [5] Batzelis, E.I., Routsolias, I.A., Papathanassiou, S.A.: An explicit pv string model based on the lambert w function and simplified mpp expressions for operation under partial shading. *IEEE Transactions on Sustainable Energy* 5(1), 301-312 (2013). <https://doi.org/10.1109/TSTE.2013.2282168>
- [6] Belhachat, F., Larbes, C.: Modeling, analysis and comparison of solar photovoltaic array configurations under partial shading conditions. *Solar Energy* 120, 399-418 (2015). <https://doi.org/10.1016/j.solener.2015.07.039>
- [7] Belhadj, C.A., Banat, I., Deriche, M.: A detailed analysis of photovoltaic panel hot spot phenomena based on the bishop model. In: 2017 14th International Multi-Conference on Systems, Signals & Devices (SSD). pp. 222-227. IEEE (2017). <https://doi.org/10.1109/SSD.2017.8166924>
- [8] Belhaouas, A., Cheikh, M.A., Larbes, C.: Suitable matlab-simulink simulator for pv system based on a two-diode model under shading conditions. In: 3<sup>rd</sup> International Conference on Systems and Control. pp. 72-76. IEEE (2013). <https://doi.org/10.1109/ICoSC.2013.6750837>



- [9] Bishop, J.: Computer simulation of the effects of electrical mismatches in photovoltaic cell interconnection circuits. *Solar cells* 25(1), 73-89 (1988). [https://doi.org/https://doi.org/10.1016/0379-6787\(88\)90059-2](https://doi.org/https://doi.org/10.1016/0379-6787(88)90059-2)
- [10] Boylestad, R.L., Nashelsky, L.: *Electronic devices and circuit theory*. Prentice Hall (2012)
- [11] Díaz-Dorado, E., Cidras, J., Carrillo, C.: Discrete i-v model for partially shaded pv-arrays. *solar energy* 103, 96-107 (2014). <https://doi.org/https://doi.org/10.1016/j.solener.2014.01.037>
- [12] Friman, H.: New trends in the higher education: Renewable energy at the faculty of electrical engineering. *Energy Procedia* 115, 18-28 (2017). <https://doi.org/10.1016/j.egypro.2017.05.003>
- [13] Hassan, M.U., Saha, S., Haque, M.E.: Pvanalytx: A matlab toolkit for technoeconomic analysis and performance evaluation of rooftop pv systems. *Energy* 223 (2021). <https://doi.org/10.1016/j.energy.2021.120>
- [14] Kawamura, H., Naka, K., Yonekura, N., Yamanaka, S., Kawamura, H., Ohno, H., Naito, K.: Simulation of i-v characteristics of a pv module with shaded pv cells. *Solar Energy Materials and Solar Cells* 75(3-4), 613-621 (2003). [https://doi.org/https://doi.org/10.1016/S0927-0248\(02\)00134-4](https://doi.org/https://doi.org/10.1016/S0927-0248(02)00134-4), pVSEC 12, PART III
- [15] Khenfer, R., Mostefai, M., Benahdoug, S., Maddad, M.: Faults detection in a photovoltaic generator by using matlab simulink and the chipkit max32 board. *International Journal of Photoenergy* 2014 (2014). <https://doi.org/https://doi.org/10.1155/2014/350345>
- [16] Lucas, H., Pinnington, S., Cabeza, L.F.: Education and training gaps in the renewable energy sector. *Solar Energy* 173, 449-455 (2018). <https://doi.org/https://doi.org/10.1016/j.solener.2018.07.061>
- [17] Mahammed, I.H., Arab, A.H., Bakelli, Y., Khennene, M., Oudjana, S.H., Fezzani, A., Zaghba, L., et al.: Outdoor study of partial shading effects on different pv modules technologies. *Energy Procedia* 141, 81-85 (2017). <https://doi.org/10.1016/J.EGYPRO.2017.11.016>
- [18] Mandal, A., Alom, A., Khan, K.A., Pal, S., De, A.: Matlab-graphical user interface (gui) to study pv module characteristics. In: 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS). pp. 2399-2403. IEEE (2017). <https://doi.org/10.1109/ICECDS.2017.8389879>
- [19] Mikofski, M., Meyers, B., Chaudhari, C.: Pvmismatch project. <https://github.com/SunPower/PVMismatch> (2018)

- [20] Orozco-Gutierrez, M., Ramirez-Scarpetta, J., Spagnuolo, G., Ramos-Paja, C.: A method for simulating large pv arrays that include reverse biased cells. *Applied energy* 123, 157{167 (2014). <https://doi.org/10.1016/j.apenergy.2014.02.052>
- [21] Piccoli, E., Dama, A., Dolara, A., Leva, S.: Experimental validation of a model for pv systems under partial shading for building integrated applications. *Solar Energy* 183, 356-370 (2019). <https://doi.org/https://doi.org/10.1016/j.solener.2019.03.015>
- [22] Quaschnig, V., Hanitsch, R.: Numerical simulation of current-voltage characteristics of photovoltaic systems with shaded solar cells. *Solar energy* 56(6), 513-520 (1996). [https://doi.org/https://doi.org/10.1016/0038-092X\(96\)00006-0](https://doi.org/https://doi.org/10.1016/0038-092X(96)00006-0)
- [23] RABAI, L.B.A., COHEN, B., MILL, A.: Programming language use in us academia and industry. *Informatics in Education* 14(2), 143-160 (2015). <https://doi.org/10.15388/infedu.2015.09>
- [24] Ramaprabha, R., Mathur, B.: A comprehensive review and analysis of solar photovoltaic array configurations under partial shaded conditions. *International Journal of Photoenergy* 2012 (2012). <https://doi.org/10.1155/2012/120214>
- [25] Reddy, G.S., Reddy, T.B., Kumar, M.V.: A matlab based pv module models analysis under conditions of nonuniform irradiance. *Energy Procedia* 117, 974-983 (2017). <https://doi.org/https://doi.org/10.1016/j.egypro.2017.05.218>
- [26] Saha, C., Agbu, N., Jinks, R., Huda, M.: Review article of the solar pv parameters estimation using evolutionary algorithms. *MOJ Solar and Photoenergy Systems* 2(2), 66-78 (2018). <https://doi.org/10.15406/mojsp.2018.02.00026>
- [27] Vera-Dávila, A.G., Delgado-Ariza, J.C., Sep\_ulveda-Mora, S.B.: Validación del modelo matemático de un panel solar empleando la herramienta simulink de matlab. *Revista de Investigación, Desarrollo e Innovación* 8(2), 343{356 (2018). <https://doi.org/https://doi.org/10.19053/20278306.v8.n2.2018.7972>
- [28] Wang, Y.J., Hsu, P.C.: An investigation on partial shading of pv modules with different connection configurations of pv cells. *Energy* 36(5), 3069-3078 (2011). <https://doi.org/https://doi.org/10.1016/j.energy.2011.02.052>
- [29] Wen, Z., Chen, J., Cheng, X., Niu, H., Luo, X.: A new and simple split series strings approach for adding bypass diodes in shingled cells modules

- to reduce shading loss. *Solar Energy* 184, 497-507 (2019). <https://doi.org/https://doi.org/10.1016/j.solener.2019.03.099>
- [30] Zegaoui, A., Petit, P., Aillerie, M., Sawicki, J., Belarbi, A., Krachai, M., Charles, J.: Photovoltaic cell/panel/array characterizations and modeling considering both reverse and direct modes. *Energy Procedia* 6, 695-703 (2011). <https://doi.org/10.1016/j.egypro.2011.05.079>

# Direct diffuse reflectance model implementation using optical parameters applied to the spectral simulation of avocado leaf

---

Nicolás Alberto Molina Cerón

MIRP Lab - Instituto Tecnológico Metropolitano - ITM, Medellín, Colombia

María Constanza Torres Madroñero

MIRP Lab - Instituto Tecnológico Metropolitano - ITM, Medellín, Colombia

nicolasmolina196140@correo.itm.edu.co

July Galeano

Grupo de Investigación Materiales Avanzados y Energía MatyEr

Instituto Tecnológico Metropolitano - ITM, Medellín, Colombia

María Casamitjana

Corporación Colombiana de Investigación Agropecuaria

AGROSAVIA, Bogotá, Colombia

## Abstract

Spectral systems measure the energy reflected and emitted by a surface along the electromagnetic spectrum. This data can be used for precision agriculture to estimate the productivity and quality of a crop. Using spectral signatures and optical models, we can estimate the amount of water and nutrients in a plant. Although models such as PROSPECT, SAIL, and PROSAIL have achieved outstanding results, there is still a need to develop models that adapt to the characteristics of specific crops such as avocado. This paper describes a diffuse reflectance model to simulate the spectral signature of avocado leaf. We compared three optimization methods: Broyden-Fletcher-Goldfarb-Shanno (BFGS), Sequential Least-Square Programming (SLSQP), and Limited Memory BroydenFletcher-Goldfarb-Shanno with boundaries (L-BFGS-B). Also, we explore three objective functions: Square Mean Error (MSE), Symmetric Mean Absolute Percentage Error (SMAPE), and Mean Absolute Percentage Error (MAPE). The model is evaluated using spectral signatures of avocado Hass taken in RionegroAntioquia. Although additional data are required for model generalization, in this paper, we show the potential

of the proposed optical model for estimating parameters such as leaf thickness, chlorophyll, water, carotenoid, and brown pigments.

## **Keywords**

Diffuse reflectance model, spectral data, vegetation spectral signature, optical parameter.

## **Introduction**

### *Importance of avocado crops*

In recent years, the market and production of avocados have grown exponentially. The avocado has a great variety of vitamins and nutrients such as calcium, magnesium, potassium, sodium, zinc, vitamin A, C, B-12, among many others [1]. The amount of health benefits is relevant, helping to prevent diseases and strengthening health. For this reason, the consumption of this fruit has increased around the world. For the year 2019, the third-largest producers of avocado in the world were Mexico, the Dominican Republic, and Colombia, with a total of 2.17 million, 691,551, and 544,933 tons, respectively, according to statistics provided by food and agriculture organizations [2]

### *Precision agriculture*

Precision agriculture consists of monitoring crops using sensors to obtain information on the state of nutrients, soil condition, temperatures, and all relevant data to make decisions to optimize harvests, improve production and crop health. [3]. Real-time monitoring and good management practices to enhance crop production have led people to use remote sensing in precision agriculture. Precision agriculture is the group of techniques, technologies, information systems, and sensors that allow acquiring continuous data and information along with time and surface throughout a crop. According to a recent systematic literature review, remote sensors are the most used technology in precision agriculture worldwide [4]. This technology consists of capturing the information from a great distance in a specific area; generally, this infor-

mation is acquired by sensors that are in satellites or airplanes, capable of obtaining information from the electromagnetic spectrum, capturing data from the emitted or reflected light by an object [5].

Multispectral sensors can acquire spectral information about an area or object. These sensors can collect high spatial resolution images with tens of bands between the visible and near-infrared range [6]. Instead, hyperspectral sensors collect hundreds of narrow bands along the electromagnetic spectrum [7].

Precision agriculture using remote sensing techniques has been most researched in corn, sugarcane, and wheat crops [4]. Still, avocado is an emergent crop, where applying these tools is necessary for the high market demand. New researchers are implementing remote sensing images on avocado crops to predict production [8].

### *Vegetation spectral signature*

The spectral signature of vegetation in the 400 to 700 nm range is absorbed in high amounts by chlorophyll pigments, especially in the blue and red wavelengths due to photosynthesis. Roughly in the near-infrared range, between 700 to 1300 nm, where high reflectance is noticeably caused by the structural properties of leaves and biomass. Finally, the shortwave infrared zone starts at 1300 nm to 2500 nm, where the absorption is mainly due to water, and other biochemical components present in the leaves [7]. This type of measurement can be affected by the amount of water, nutrients, and light; then, spectral data is used for their characterization [9].

### *Models for simulating spectral signatures*

Optical models such as PROSPECT can simulate the reflectance and transmittance of leaves in the spectral range of 400 to 2500 nm with only a few numerical parameters; on the other hand, SAIL can obtain information about soil and optical properties. Combining PROSPECT and SAIL, a model named PROSAIL is obtained. It is used to estimate biophy-

sical and biochemical variables of plants such as chlorophyll, carotenoids, moisture, nitrogen, water content, and plant phenology [10]. This model has been used to obtain leaf area index, chlorophyll, and nitrogen contents in sugar beet crops [11]. A variation of this model is the N-PROSAIL model and was used to estimate nitrogen levels in winter wheat [12].

Although models such as PROSPECT, SAIL, and PROSAIL have achieved outstanding results, it is still necessary to develop models adapted to the characteristics of specific crops such as avocado because most studies have focused on potatoes and corn crops.

Bearing in mind these needs, in this article, we propose a diffuse reflectance model for the spectral characterization of avocado, and we compare three optimization methods and three objective functions. We present results for eight spectral signatures of avocado plants taken in situ in Rionegro-Antioquia.

## Methods and procedures

For this study, we collected eight spectral signatures of avocado using a FLAME-SVIS-NIR spectrometer of Ocean Insight. Each spectral signature results from averaging ten signatures captured by the spectrometer in the same leave. The spectral data were measured on adult leaves of 8 avocado Hass trees at the Research Center La Selva of AGROSAVIA in Rionegro-Antioquia, located on the same orchard, with the same environmental conditions, selected on a transect along with the orchard. The spectral data were collected at midday under optimal climatological conditions. The spectrometer captures 2049 spectral bands between 339 nm and 1028 nm. For calibration, we measured a black signature,  $r_{black}$  by obstructing the entrance of light to the optical fiber and the spectrum for a white reflectance pattern,  $r_{white}$ . For each signature, the reflectance is calculated using equation (1):

$$R = \frac{r - r_{black}}{r_{white} - r_{black}} \quad (1)$$

Subsequently, wavelengths below 400 nm and above 870 nm are eliminated to reduce the noise level of the signatures. Finally, an averaging filter with a 5-point sliding window is applied to the signal.

For the analysis of light-tissue interaction, different mathematical expressions have been proposed. For this work, the multilayered exponential scheme and is expressed in equation (2) [13][14][15]

$$I(\lambda) = E_0 e^{-(\mu_s + \mu_a)z} \quad (2)$$

An adjustment in terms of reflectance of the equation (2) is expressed as follows:

$$R = \frac{1}{1+k\left(\frac{\mu_a}{\mu_s}\right)} (1 - e^{-(\mu_a - \mu_s)z}) \quad (3)$$

where  $\mu_a(i)$  is the absorption in the layer  $i$  and  $\mu_s(i)$  is the scattering in the layer  $i$ .

The following equation was used to calculate the absorption in the leaves (4): [16][17]:

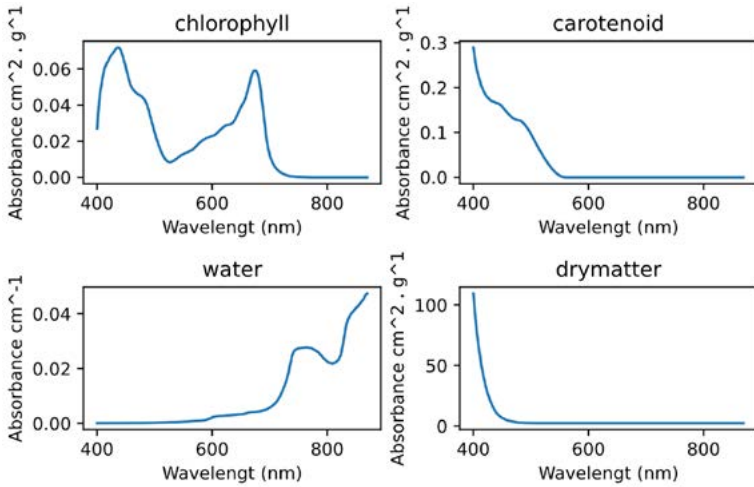
$$\mu_a = \sum_i c_i K(\lambda)_i \quad (4)$$

where  $c_i$  is the amount of the absorption coefficient  $i$ , and  $K(\lambda)_i$ , corresponds to the  $i$ th absorption coefficient.

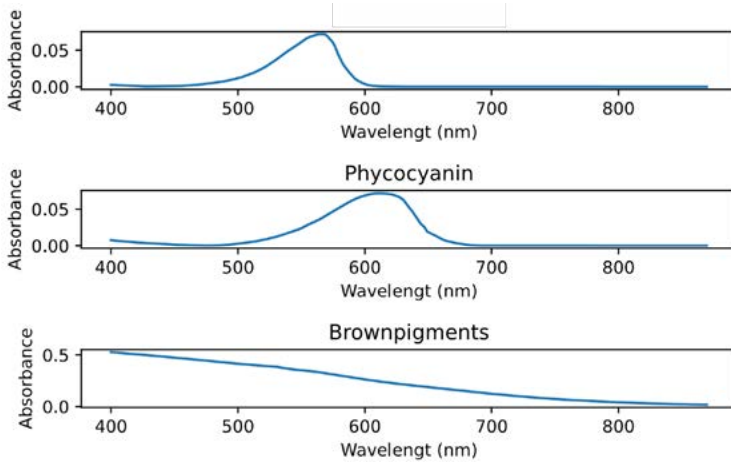
For this study, the following absorption coefficients were used (see Fig. 1): chlorophyll, dry soil, pigments such as carotenoid, which is an important factor in a photosynthetic structure [18], and pigments such as phycocyanin which has blue color existent on old leaves and phycoerythrin red color, that it can be observed on young avocado leaves [19]. The values shown in Fig. 2 were taken as a reference for dispersion.



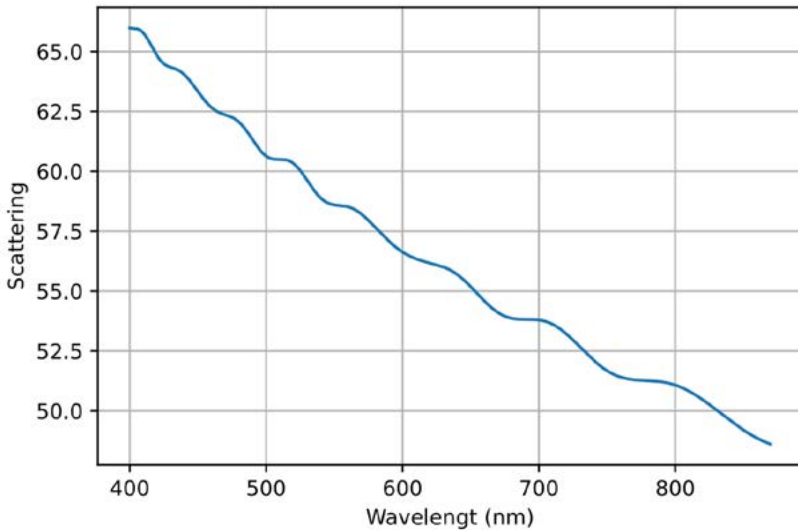
**Figure 1**  
**Specific absorption coefficients [16]**



**Figure 2**  
**Pigments absorption coefficients theoretical values for Phycoerythrin, Phycocyanin and Brown pigments. [20]**



**Figure 3**  
**Light scattering. [21]**



The direct model presented in equation (3), was implemented by using the programming language Python=3.9.6, the packages Matplotlib =3.4.3, Numpy =1.21.2, and Scipy = 1.7.1. The wavelength range for the simulation was between 400 and 900 nm. For model inversion, three minimize objects were created from the Scipy library, and each one was assigned a different method. The selected ones were Broyden-Fletcher-Goldfarb-Shanno (BFGS), Sequential Least-Square Programming (SLSQP), and Limited Memory Broyden-Fletcher-Goldfarb-Shanno with boundaries (L-BFGS-B), the last two being algorithms with restrictions. As target functions, the following functions were used: Square Mean Error (MSE) (5), Symmetric Mean Absolute Percentage Error (SMAPE) (6), and Mean Absolute Percentage Error (MAPE) (7):

$$MSE = \frac{1}{n} \sum_{i=0}^n (Y_i - Y'_i)^2 \quad (5)$$

$$SMAPE = \frac{100\%}{n} \sum_{i=0}^n \frac{|Y_i - Y'_i|}{|Y_i| + |Y'_i|} \quad (6)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \frac{Y_i - Y'_i}{Y_i} \quad (7)$$

With n as the number of wavelengths acquired, as the acquired reflectance, and  $Y'_i$  as the simulated reflectance, both acquired, and simulated reflectance should be normalized from 0 to 1.

Table 1 shows the limits selected for each parameter. Finally, the average signature of the 8 firms was obtained and used as a reference.

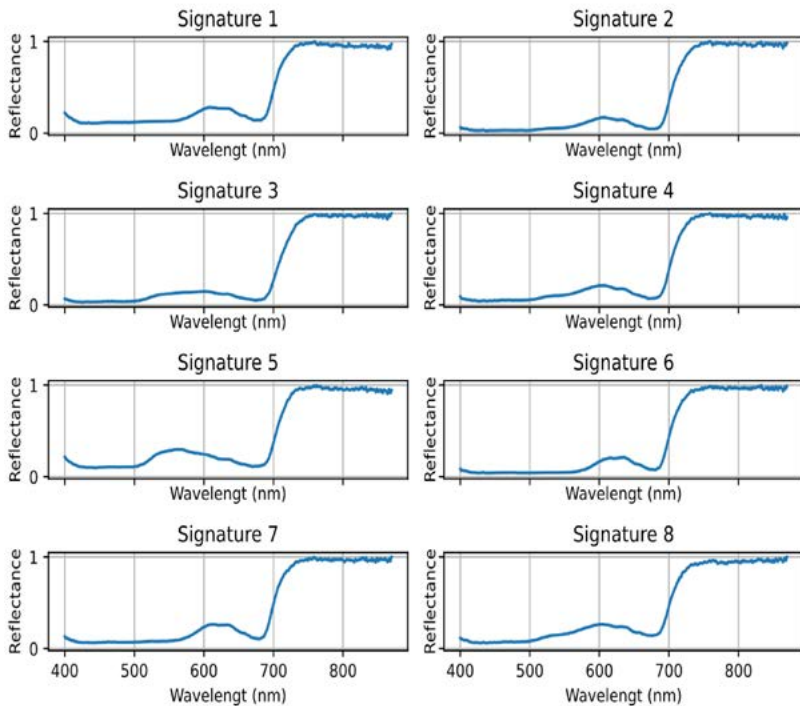
**Table 1**  
**Boundaries selected**

Parameters	Lower boundary	Upper boundary
L (leaf thickness)	0	1
Chlorophyll	0	1000
Water	0	1
Carotenoid	0	1000
Brown pigments	0	1
Dry mater	0	1
K (scaling factor)	0	100
Phycocerythrin	0	1000
Phycocyanin	0	1000

## Results and analysis

The measured spectral signatures of the avocado leaves are presented in Fig. 4. We can note that there is a decrease in the first 400 to 430 nm, then between 520 to 680 nm there is a peak, between 680-750 nm onwards the reflectance begins to increase and finally from 750 nm forward a constant reflectance is maintained.

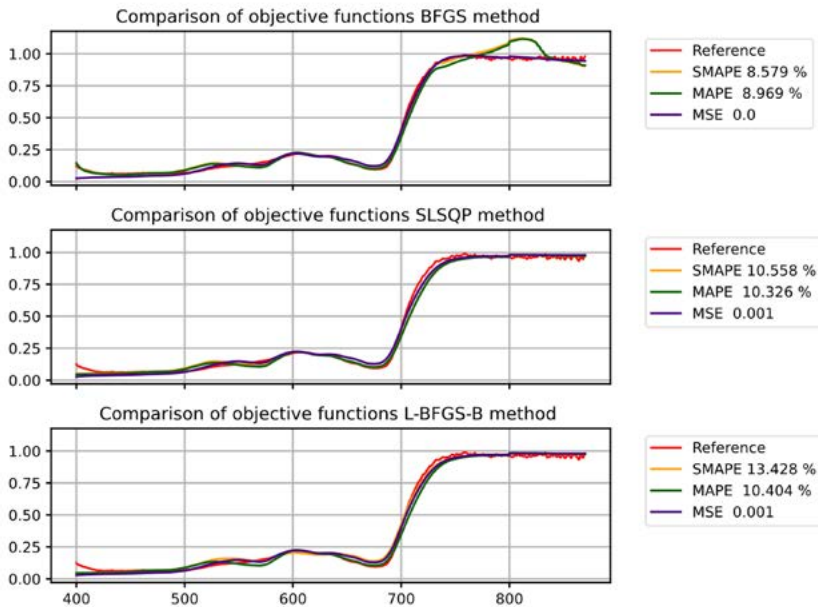
**Figure 4**  
**Avocado spectral signatures taken on leaves**



After applying the BFGS, SLSP, and L-BFGS-B optimization algorithms, we obtained the estimated signatures presented in Fig. 5. For

these resulting firms, it is necessary to apply scale factors. We can note a good fit between the estimated reflectance and the measured signature (red reference signature). Using the MSE, SMAPE, and MAPE, the best results are obtained using the BFGS method. However, in the graph, we can note a deviation of the signatures estimated using the SMAPE and MAPE objective function around 800 nm.

**Figure 5**  
**Simulation of avocado signatures with BFGS, SLSQP,**  
**and L-BFGS-B methods, the red signature is the average signature**  
**of the 8 samples, for the objective functions yellow for SMAPE,**  
**green for MAPE, and purple for the MSE**



Tables 2, 3 and 4 show the resulting values for each optimization methods and for each of the SMAPE, MAPE and MSE objective functions.

**Table 2**  
**BFGS method**

Parameters	SMAPE	MAPE	MSE
L (leaf thickness)	0.995	0.964	1.03
Chlorophyll	1163.708	1132.848	786.788
Water	67.073	65.131	1.993
Carotenoid	487.099	486.586	899.967
Brown pigments	10.111	13.413	-7.14
Dry mater	-1.142	-1.149	0.163
K (scaling factor)	7.112	7.244	8.48
Phycocerythrin	636.785	693.147	451.978
Phycocyanin	-26.324	-32.994	81.92

**Table 3**  
**SLSQP method**

Parameters	SMAPE	MAPE	MSE
L (leaf thickness)	0.408	0.408	0.937
Chlorophyll	997.706	991.518	786.783
Water	1	1	0.141
Carotenoid	441839	484.619	899.991
Brown pigments	0	0	0
Dry mater	0	0	0.051
K (scaling factor)	7.919	7.994	7.822
Phycocerythrin	578.33	642.686	451.958
Phycocyanin	0	0	81.896

**Table 4**  
**L-BFGS-B method**

Parameters	SMAPE	MAPE	MSE
L (leaf thickness)	0.386	0.408	1
Chlorophyll	890.247	1000	786.782
Water	1	1	0.426
Carotenoid	767.671	498.2	899.991
Brown pigments	0	0	0
Dry mater	0	0	0.048
K (scaling factor)	6.194	7.854	7.835
Phycocerythrin	514.868	675.47	451.958
Phycocyanin	197.3	0	81.896

## Conclusions

A diffuse reflectance model to simulate the spectral signature of avocado was proposed. This model was subjected to 3 optimization methods, BFGS, SLSQP, and l-BFGS-B, having as objective functions three metrics to calculate the error. These were SMAPE, MAPE, and MSE to compare the results with the average signature of 8 spectral signatures taken in Rionegro-Antioquia.

Although the performance of the BFGS method has the lowest error results, the values of its parameters in some cases become negative. Thus, this method is only comparative to SLSQP and l-BFGS-B, whose error percentages are slightly higher.

Simulating the spectral signature of the avocado leaf or any other type of crop could help develop and characterize certain diseases or crop deficiencies considering the behavior and values of its components.

Future work will involve capturing more spectral data to allow a better generalization of the proposed model to identify stress levels in the avocado trees. Likewise, it is expected to test other optimization methods such as genetic algorithms to improve the inverse model.

## Funding

This work was financed with resources from the Patrimonio autónomo fondo nacional de financiamiento para ciencia, tecnología e innovación Francisco José de Caldas through RC 80740-475-2020, supported by the Ministerio de Ciencia, Tecnología e Innovación - Minciencias, Instituto Tecnológico Metropolitano - ITM, Corporación Colombiana de Investigación Agropecuaria - AGROSAVIA, and BLACK-SQUARE S.A.S

## References

- [1] Araújo, R. G., Rodríguez-Jasso, R. M., Ruiz, H. A., Pintado, M. M. E., & Aguilar, C. N.: Avocado by-products: Nutritional and functional properties. *Trends in Food Science & Technology*, 80, 51–60 (2018).
- [2] Agronegocios, <https://www.agronegocios.co/agricultura/colombia-estercero-en-el-mundo-en-produccion-y-area-cosechada-de-aguacate-hass-3142547>, last accessed 2021/08/21
- [3] Shafi, U., Mumtaz, R., García-Nieto, J., Hassan, S. A., Zaidi, S. A. R., & Iqbal, N. Precision agriculture techniques and practices: From considerations to applications. *Sensors*, 19(17), 3796. (2019).
- [4] Weiss, M., Jacob, F., & Duveiller, G.: Remote sensing for agricultural applications: A metareview. *Remote Sensing of Environment*, 236, 111402. (2020).
- [5] Deng, L., Mao, Z., Li, X., Hu, Z., Duan, F., & Yan, Y.: UAV-based multispectral remote sensing for precision agriculture: A comparison between different cameras. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146, 124–136. (2018).
- [6] Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., & Sousa, J.: Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Pro-



- cessing and Applications for Agriculture and Forestry. *Remote Sensing*, 9(11), 1110. (2017).
- [7] Segarra, J., Buchaillet, M. L., Araus, J. L., & Kefauver, S. C.: Remote Sensing for Precision Agriculture: Sentinel-2 Improved Features and Applications. *Agronomy*, 10(5), 641. (2020).
- [8] Wu, D., Johansen, K., Phinn, S., Robson, A., & Tu, Y. H.: Inter-comparison of remote sensing platforms for height estimation of mango and avocado tree crowns. *International Journal of Applied Earth Observation and Geoinformation*, 89, 102091. (2020).
- [9] Wang, L., Huang, S., Huang, Q., Leng, G., Han, Z., Zhao, J., & Guo, Y.: Vegetation vulnerability and resistance to hydrometeorological stresses in water- and energy-limited watersheds based on a Bayesian framework. *CATENA*, 196, 104879. (2021).
- [10] Berger, K., Atzberger, C., Danner, M., D’Urso, G., Mauser, W., Vuolo, F., & Hank, T.: Evaluation of the PROSAIL Model Capabilities for Future Hyperspectral Model Environments: A Review Study. *Remote Sensing*, 10(2), 85. (2018).
- [11] Jay, S., Maupas, F., Bendoula, R., & Gorretta, N. Retrieving LAI, chlorophyll and nitrogen contents in sugar beet crops from multi-angular optical remote sensing: Comparison of vegetation indices and PROSAIL inversion for field phenotyping. *Field Crops Research*, 210, 33-46. (2017).
- [12] Li, Z., Jin, X., Yang, G., Drummond, J., Yang, H., Clark, B., ... & Zhao, C. Remote sensing of leaf and canopy nitrogen status in winter wheat (*Triticum aestivum* L.) based on NPROSAIL model. *Remote Sensing*, 10(9), 1463. (2018).
- [13] Kubelka, P.: New Contributions to the Optics of Intensely Light-Scattering Materials Part I. *Journal of the Optical Society of America*, 38(5), 448. (1948).
- [14] Ishimaru, A.: Wave propagation and scattering in random media and rough surfaces. *Proceedings of the IEEE*, 79(10), 1359–1366. (1991).
- [15] Karabourniotis, G., Liakopoulos, G., Bresta, P., & Nikolopoulos, D. The Optical Properties of Leaf Structural Elements and Their Contribution to Photosynthetic Performance and Photoprotection. *Plants*, 10(7), 1455. (2021).
- [16] Jacquemoud, S., Ustin, S. L., Verdebout, J., Schmuck, G., Andreoli, G., & Hosgood, B.: Estimating leaf biochemistry using the PROSPECT leaf optical properties model. *Remote Sensing of Environment*, 56(3), 194–202. (1996).

- [17] Feret, J.-B., François, C., Asner, G. P., Gitelson, A. A., Martin, R. E., Bidel, L. P. R., ... Jacquemoud, S.: PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment*, 112(6), 3030–3043. (2008).
- [18] Hirschberg, J. Carotenoid biosynthesis in flowering plants. *Current Opinion in Plant Biology*, 4(3), 210–218. (2001).
- [19] Hsieh-Lo, M., Castillo, G., Ochoa-Becerra, M. A., & Mojica, L. Phyco-cyanin and phycoerythrin: Strategies to improve production yield and chemical stability. *Algal Research*, 42, 101600. (2019).
- [20] Secretaría de Agricultura, ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA) México. Estudio de factibilidad para el establecimiento de cultivo de nopal (opuntia) en tierras ociosas en los estados de Aguascalientes, San Luis Potosí, Guanajuato y Zacatecas con fines alimenticios, energéticos y ambientales. OCTUBRE 2015.
- [21] Mätzler, C. MATLAB functions for Mie scattering and absorption, version 2. (2002).

# Evaluation of supervised and semi-supervised classifiers for multispectral image processing: a comparative study

---

Valentina Suescún Arias

MIRP Lab, Instituto Tecnológico Metropolitano - ITM,  
Medellín, Colombia  
valentinasuescunarias@gmail.com

María Constanza Torres Madroñero

MIRP Lab, Instituto Tecnológico Metropolitano - ITM,  
Medellín, Colombia  
mariatorres@itm.edu.co

July Galeano Zea

Grupo de Investigación Materiales Avanzados y Energía MatyEr,  
Instituto Tecnológico Metropolitano - ITM, Medellín, Colombia  
julygaleano@itm.edu.co

María Casamitjana

Corporación Colombiana de Investigación Agropecuaria  
AGROSAVIA, Bogotá 250047, Colombia  
mcasamitjana@agrosavia.co

## Abstract

Spectral image classification is a topic of interest in remote sensing. Most of the studies focus on hyperspectral image (HSI) classification using supervised learning and semi-supervised learning with models such as Support Vector Machine (SVM) and Random Forest (RF). These models are accompanied by different pre-processing techniques to obtain better results in classification, such as Minimum Noise Fraction (MNF) and extended morphological attribute profiles (EMAPs). Due to the high cost of HSI imaging, it is common to use multispectral imaging (MSI) to perform land cover classification. However, performing this type of classification requires having a complete ground truth, which is difficult to generate due to the great human effort required to label an image.

Semi-supervised models are used to decrease the need for a fully labeled ground truth, as this learning implements labeled and unlabeled data. Therefore, in this paper, we propose implementing different supervised and semi-supervised learning configurations on MSI to evaluate the effectiveness of these methods on images with few spectral bands and limited labeled data. We use a reduced Indian Pines - AVIRIS image in this comparison. We selected ten spectral bands to transform the HSI image into an MSI. Taking advantage of the ground truth available for Indian Pines, we compared the performance of several supervised (SVM, RF, Stack Classifier + MNF) and semi-supervised classifiers (SVM, RF, Stack classifier + EMAPs), obtaining accuracies of up to 76.09% and 91.37%. This paper aims to identify the capabilities and limitations of supervised and semi-supervised classifiers for MSI obtained by commercial off-the-shelf cameras, where labeled data is limited.

## Keywords

Hyperspectral images, Multispectral images, Band Selection, Supervised classifier, Semi-supervised classifier.

## Introduction

Hyperspectral images (HSI) are widely used in different fields such as precision agriculture and land cover classification. Even so, the large dimension of HIS can affect the accuracy of classification [4], unlike multispectral imaging (MSI), which has a limited number of spectral bands. The main difference between these two types of images is the number of spectral bands, bandwidth, spatial resolution, acquisition method, and cost. MSI can be considered a special case of HIS in which the wavelength range of the bands cannot be considered continuous but discrete [1].

During the last years, different supervised classifiers have been applied for HSI classification, such as Support Vector Machine (SVM), due to its remarkable capacity to handle high dimensional data with a limited number of training samples, and Random Forest (RF), since it is not sensitive to over-training [5]. However, the classification task can be affected by the noise contained in the images, and it is necessary to use of denoising techniques such as Minimum Noise Fraction (MNF). This technique creates a set of images organized by quality by selecting components that increase the signal-to-noise ratio (SNR), which allows retaining helpful information [10].

One of the limitations of supervised learning is obtaining an MSI with fully labeled data. It is a relatively expensive task, so it is essential to look for alternatives that use both labeled and unlabeled data as semi-supervised learning [8]. However, since semi-supervised learning does not have many labeled samples for training, its accuracy may not be the best. Several works have proposed using the extended morphological attribute profiles (EMAPs) to increase the accuracy of semi-supervised classification since this technique allows extracting both spatial and spectral information [6, 9].

This article implements two supervised and two semi-supervised classification configurations on MSI obtained from the Indian Pines HSI published by M Graña, MA Veganzons, B Ayerdi [11]. This work aims to identify and evaluate the capabilities of supervised and semi-supervised classifiers for MSI in the spectral bands that can be obtained by commercial off-the-shelf cameras like the Micasense RedEdge-MX, where labeled data is limited. The results obtained demonstrate that these supervised and semi-supervised techniques can be used in the future in fully labeled or low-labeled MSI acquired from different crops, such as avocado, bean and corn fields.

## Materials and Methods

The different pre-processing and classification methods are compared and evaluated using the Indian Pines HSI obtained by the AVIRIS sensor (Fig. 1a). This HSI has 145x145 pixels and 200 spectral bands (after eliminating the water absorption bands) with a wavelength of 0.4-2.5  $\mu\text{m}$ . The image has an atmospheric correction and ground truth classification map with sixteen assigned classes as presented in (Fig. 1b).

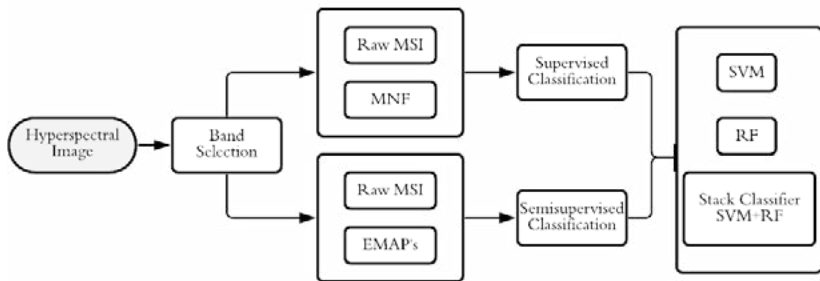
Figure 1



(a) RGB color remote sensing image of the Indian Pines. (b) Ground Truth of the labeled area with sixteen classes of land cover: Alfalfa, Corn-notill, Corn-mintill, Corn, Grass-pasture, Grass-trees, Grass-pasture-mowed, Hay-windrowed, Oats, Soybean-notill, Soybean-mintill, Soybean-clean, Wheat, Woods, Buildings-Grass-Trees-Drives, and Stone-Steel-Towers.

Fig. 2 shows the working scheme of the implemented methodology, which consists of the following main steps: band selection, pre-processing, supervised classification, and semi-supervised classification.

**Figure 2. Flow chart of the proposed methodology. MSI refers to Multispectral Image, MNF refers to Minimum Noise Fraction, and EMAPs refers to Extended Morphological Attribute Profiles**



## Band selection

To evaluate how different classifiers behave on images with a limited number of bands, we selected ten bands from Indian Pines. The selected bands have wavelengths similar to the commercial camera Micasense RedEdge-MX, which is widely used for remote sensing applications. Table 1 presents the selected bands.

**Table 1**  
**HSI Indian Pines Spectral Bands Selection based**  
**on the bands obtained by the commercial camera Micasense RedEdge-MX**

Spectral Bands - Micasense [nm]	Spectral Bands - HSI Indian Pines [nm]
Coastal Blue (444)	Band 6
Blue (475)	Band 9
Green (531)	Band 14
Green (560)	Band 17
Red (650)	Band 26
Red (668)	Band 28
Red Edge (705)	Band 34
Red Edge (717)	Band 35
Red Edge (740)	Band 38
Red Border (740)	Band 48

## Supervised Classification

Supervised classification is performed on the ten-band and the ten-band images with MNF pre-processing. A total of four classifiers are applied: 1. SVM with Gaussian Radial Basis Function (RBF) kernel that has been extensively used in remote sensing for its superior and stable performance, 2. RF with 81 trees in the forest (n estimators), and a Stack Classifier (SVM+RF) in which SVM and RF are used as meta classifier.

**Supervised classification on RAW MSI.** For the Raw MSI (Fig. 3a), conventional supervised classification is performed with the above.

First, the image is divided into 80% for training and 20% for testing, then each of the models is trained with the training set, and, finally, a prediction and accuracy is obtained with the testing set. Fig. 4 summarizes this framework.

**Supervised classification on MNF MSI.** The MNF transform is applied to the ten-band Raw MSI. The output is an MSI with the same amount of input bands. The difference is that these contain less noise and preserve the helpful information components to improve the classification process. Fig. 3b shows the RGB composition for the image obtained by MNF.

**Figure 2**

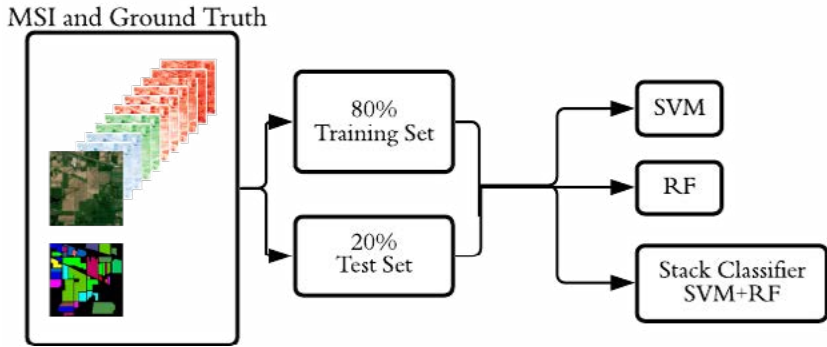


(a) RGB color remote sensing image of the Indian Pines. (b) RGB color Multispectral remote sensing image of the Indian Pines after applying MNF preprocessing.

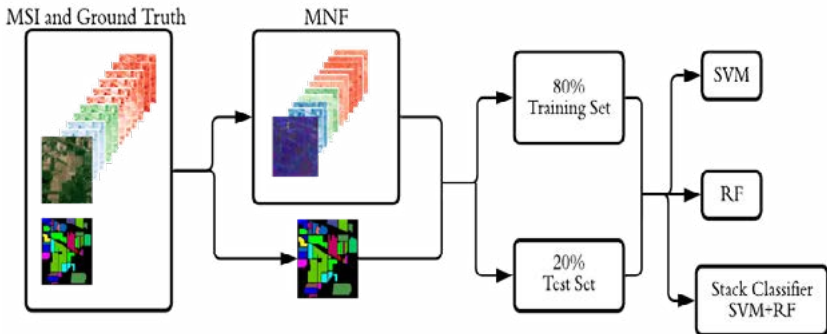
The MSI obtained from the MNF transform, as in the classification on the original image, is divided into 80% for training and 20% for testing. The four models are trained with the training set and, finally, prediction and accuracy are performed with the test set. Fig. 5 presents the flow chart for the supervised classification using MNF.



**Figure 3**  
**Flow chart of the Supervised Classification on the Raw MSI**



**Figure 4**  
**Flow chart of the Supervised Classification on the MNF MSI**



### *Semi-supervised classification*

The semi-supervised classification is performed on the ten-band image and the ten-band image with EMAPs pre-processing. The four classifiers used in supervised classification (SVM, RF, and Stack Classifier - SVM+RF varying the meta classifier) are applied. The difference between the two types of classification is that the percentage

of labeled data in the MSI and its ground truth are varied for the semi-supervised task, starting at 5% and increasing five by five until reaching 50%.

**Semi-supervised classification on RAW MSI.** Before starting the semi-supervised classification process, the percentage of labeled samples (5%, 10%, 15%,...,50%) of the raw MSI is selected. With the labeled samples, the classification models are trained. Then, a prediction is made with the samples that are not labeled. From the obtained predictions, the ones that obtain a confidence score of over 80% are chosen and are added to the training set to repeat the process (see Fig. 6).

**Semi-supervised classification on EMAPs MSI.** EMAPs are constructed from the RAW MSI. These profiles consist of a series of openings (Eq. 1) and closings (Eq. 2) operations by reconstruction for each band that makes up the MSI.

Let the MSI  $X_h$  and the structuring method  $B$ . Then EMP is the reconstructed 3D image containing the spatial and spectral information:

$$R_{x_o} (X_h \ominus B) \tag{1}$$

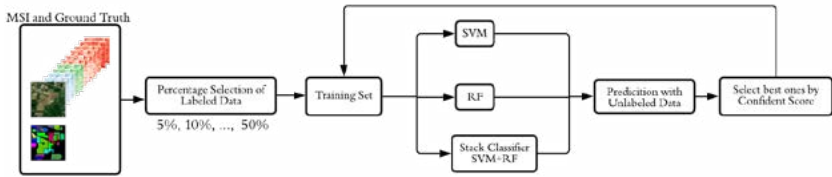
$$R_{x_c} (X_h \oplus B) \tag{2}$$

$$EMP (R_{x_o}, R_{x_c}) \tag{3}$$

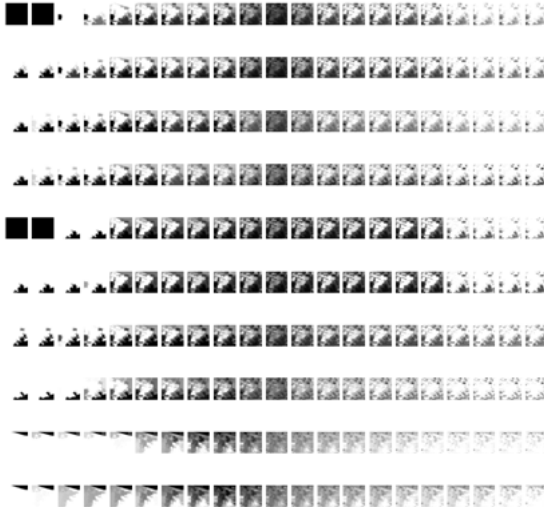
For each of the bands, twenty-one EMPs are obtained, obtaining as a final result an EMP of 145x145x210, as shown in Fig. 7.

With the EMP, the percentage of labeled samples (5%, 10%, 15%,...,50%) is selected, the classification models are trained with the labeled samples, and a prediction is obtained with those not labeled shown in Fig. 8.

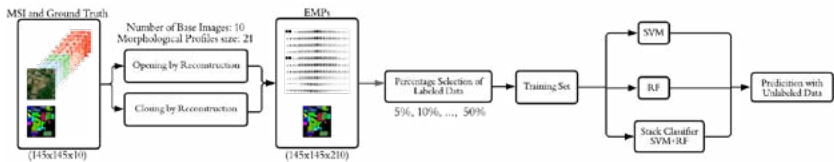
**Figure 5**  
**Flow chart of the Semi-Supervised Classification on the Raw MSI**



**Figure 6**  
**Some obtained EMPs for each band of the MSI**



**Figure 7**  
**Flow chart of the Semi-Supervised Classification on the EMAPs MSI.**



## Results

### *Supervised Classification*

Table 2 shows the results obtained for the Indian Pines MSI with and without MNF preprocessing with the four classifiers implemented. The best results for Indian Pines without pre-processing were obtained by SVM with an overall accuracy of 76.09%. Fig. 9a presents the classification map using SVM for the raw data. The best results were obtained for the pre-processed image using the stack classifier SVM+RF with a 73.12%. Fig. 9b presents the classification map using SVM+RF for the MNF image.

**Table 2**  
**Classification performance of the SVM, RF, and Stack Classifier**

Image	SVM	RF	SC1	SC2
Raw MSI	76.09	74.97	74.58	71.63
MNF MSI	72.04	72.09	73.12	72.09

(SC1: SVM+RF Meta classifier: SVM; SC2: SVM+RF Meta classifier: RF) for the RAW MSI and MNF from the Indian Pines. The highest accuracy is highlighted in bold. The accuracy values are given as percentages

### *Semi-supervised classification*

Tables 3-4 show the results obtained for the semi-supervised classification with the percentage of labeled data from 5% to 50% on the Indian Pines MSI with and without EMAPs pre-processing with the four classifiers. For the raw image, the best results for all percentages of labeled data were SVM, with accuracies between 60% and 72%. The best results were obtained using 50% of labeled data. This classification map is presented in Fig. 9c. On the other hand, the results obtained from the EMAPs are the best, with accuracies higher than 80%. RF and the stack

classifiers obtained better results than SVM in this case. The classification map obtained by RF using 50% of labeled data is presented in Fig. 9d.

**Table 3**  
**Classification performance of the SVM, RF, and Stack Classifier**

Model	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
SVM	60.58	64.29	66.82	68.97	69.26	69.95	70.43	71.51	71.60	72.09
SRF	55.02	60.97	63.21	65.26	65.65	66.43	68.39	69.56	68.58	70.34
SC1	59.80	56.68	61.46	59.51	65.07	64.48	65.85	67.90	68.00	67.51
SC2	58.34	61.26	63.80	64.68	64.58	66.04	66.24	68.39	68.58	67.70

(SC1: SVM+RF Meta classifier: SVM; SC2: SVM+RF Meta classifier: RF) for the Raw MSI of the Indian Pines with several percentages of labeled data. The highest accuracy for each percentage column is highlighted in bold. The accuracy values are given as percentages

**Table 4**  
**Table 4: Classification performance of the SVM, RF and Stack Classifier**

Model	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
SVM	82.52	85.86	87.05	87.75	88.18	88.23	88.62	88.86	89.25	89.25
SRF	88.44	87.06	88.50	89.01	89.66	90.25	90.60	90.81	91.27	91.37
SC1	83.66	86.67	88.41	89.28	89.56	90.20	90.47	90.89	91.38	91.29
SC2	83.54	86.98	88.31	89.23	89.82	90.25	90.56	90.88	91.29	91.27

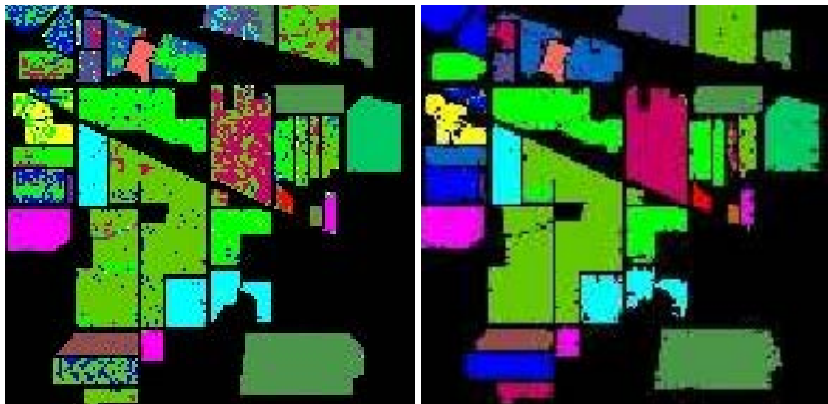
(SC1: SVM+RF Meta classifier: SVM; SC2: SVM+RF Meta classifier: RF) for the EMAPs MSI of the Indian Pines. The highest accuracy for each percentage column is highlighted in bold. The accuracy values are given as percentages

Figure 8



(a)

(b)



(c)

(d)

(a) Classification map obtained by Supervised SVM without MNF. (b) Classification map obtained by Supervised Stack Classifier SVM+RF (Meta classifier: SVM) with MNF. (c) Classification map obtained by Semi-supervised SVM - 50% labeled data without EMAPs. (d) Classification map obtained by Semi-supervised RF - 50% labeled data with EMAPs

## Discussion of Results

Comparing the results presented in Tables 2, 3, and 4, we can note that, for the MSI without MNF pre-processing, a slight increase in accuracy is observed for all supervised models compared to the values obtained on the image with MNF pre-processing. In other studies related to MNF, this technique does not tend to obtain the best classification accuracies compared to other feature extraction methods, as shown in [2], in which MNF obtains regular results.

The results of the semi-supervised models, from 5% of labeled data, are obtained accuracy higher than 50%. However, the accuracy obtained with the MSI with EMAPs pre-processing is higher than those obtained on the original MSI, starting with an accuracy of 82.52% and ending with an accuracy of 91.37%. The best classification results depend on the percentage of labeled data. The classification maps presented in Fig. 9 evidence the better performance of semi-supervised RF using EMAPs than the other evaluated approaches. Another reason why a better semi-supervised classification result is obtained with EMAPs pre-processing, as mentioned in [3], it is important to use informative features like EMAPS, which provides much spatial information besides spectral information.

The SVM and RF models give excellent results for MSI image supervised and semi-supervised classification, evident in tables 2, 3, and 4. In terms of performance and overall accuracy, the results obtained are similar to those obtained in [12], where SVM and RF have the best classification accuracy compared to other classification models.

## Conclusions

This paper describes the results obtained for supervised and semi-supervised classification on the Indian Pines image with a considerable reduction in spectral bands, going from a HSI of two hundred bands to an MSI of ten bands. A comparative analysis of accuracies obtained with highly used and cited models such as SVM and RF on the MSI with and without MNF pre-processing is shown in the supervised classification.

The results obtained with the classifiers on the image without pre-processing show an increase in the accuracies compared to the percentages obtained on the image with MNF processing. The MNF transform is a widely used technique for dimension reduction of HSI; however, in this case, since it is an image of only ten spectral bands, its performance may not be as optimal. It is reported in the literature that applying the traditional version of MNF does not provide the expected results, and in most cases, it tends to decrease the classification accuracy [13].

As for the semi-supervised classification, a comparison of the accuracies obtained on the MSI with and without EMAPs pre-processing is performed. The accuracy percentages obtained on the image with EMAPs pre-processing are higher than those obtained on the image without EMAPs pre-processing. These results demonstrate that using EMAPs on spectral images, especially those containing fewer bands [7], is a viable alternative to obtain accuracies higher than 80% for land cover classification.

Future work includes improving the accuracy of the already implemented supervised and semi-supervised classifiers and exploring other pre-processing techniques.

**Acknowledgements** This work was financed with resources from the Patrimonio autónomo fondo nacional de financiamiento para ciencia, tecnología e innovación Francisco José de Caldas through RC 80740-475-2020, supported by the Ministerio de Ciencia, Tecnología e Innovación - Minciencias, Instituto Tecnológico Metropolitano - ITM, Corporación Colombiana de Investigación Agropecuaria - AGROSAVIA, and BLACKSQUARE S.A.S.

## References

- [1] Amigo, J.M.: Hyperspectral and multispectral imaging: Setting the scene. In: *Data Handling in Science and Technology*, vol. 32, pp. 3–16. Elsevier (2020). <https://doi.org/10.1016/B978-0-444-63977-6.00001-8>
- [2] Beirami, B.A., Mokhtarzade, M.: Superpixel-based minimum noise fraction feature extraction for classification of hyperspectral images. *Traitement du Signal* 37(5) (2020). <https://doi.org/10.18280/ts.370514>



- [3] Du, P., Bai, X., Tan, K., Xue, Z., Samat, A., Xia, J., Li, E., Su, H., Liu, W.: Advances of four machine learning methods for spatial data handling: A review. *Journal of Geovisualization and Spatial Analysis* 4, 1–25 (2020). <https://doi.org/10.1007/s41651-020-00048-5>
- [4] Gao, L., Zhao, B., Jia, X., Liao, W., Zhang, B.: Optimized kernel minimum noise fraction transformation for hyperspectral image classification. *Remote sensing* 9(6), 548 (2017). <https://doi.org/10.3390/rs9060548>
- [5] Ghamisi, P., Plaza, J., Chen, Y., Li, J., Plaza, A.J.: Advanced spectral classifiers for hyperspectral images: A review. *IEEE Geoscience and Remote Sensing Magazine* 5(1), 8–32 (2017). <https://doi.org/10.1109/MGRS.2016.2616418>
- [6] Kang, X., Zhuo, B., Duan, P.: Semi-supervised deep learning for hyperspectral image classification. *Remote sensing letters* 10(4), 353–362 (2019). <https://doi.org/10.1080/2150704X.2018.1557787>
- [7] Kwan, C., Ayhan, B., Larkin, J., Kwan, L., Bernabé, S., Plaza, A.: Performance of change detection algorithms using heterogeneous images and extended multi-attribute profiles (emaps). *Remote Sensing* 11(20), 2377 (2019). <https://doi.org/10.3390/rs11202377>
- [8] Levatić, J., Ceci, M., Kocev, D., Džeroski, S.: Semi-supervised classification trees. *Journal of Intelligent Information Systems* 49(3), 461–486 (2017). <https://doi.org/10.1007/s10844-017-0457-4>
- [9] Liu, M., Cao, F., Yang, Z., Hong, X., Huang, Y.: Hyperspectral image denoising and classification using multi-scale weighted emaps and extreme learning machine. *Electronics* 9(12), 2137 (2020). <https://doi.org/10.3390/electronics9122137>
- [10] Luo, G., Chen, G., Tian, L., Qin, K., Qian, S.E.: Minimum noise fraction versus principal component analysis as a preprocessing step for hyperspectral imagery denoising. *Canadian Journal of Remote Sensing* 42(2), 106–116 (2016). <https://doi.org/10.1080/07038992.2016.1160772>
- [11] M Graña, MA Veganzons, B.A.: Hyperspectral remote sensing scenes - grupo de inteligencia computacional (gic). [http://www.ehu.es/ccwintco/index.php/Hyperspectral Remote Sensing Scenes](http://www.ehu.es/ccwintco/index.php/Hyperspectral%20Remote%20Sensing%20Scenes). Last accessed 22 May 2021.
- [12] Nhaila, H., Elmaizi, A., Sarhrouni, E., Hammouch, A.: Supervised classification methods applied to airborne hyperspectral images: comparative study using mutual information. *Procedia Computer Science* 148, 97–106 (2019). <https://doi.org/10.1016/j.procs.2019.01.013>, <https://www.sciencedirect.com/science/article/pii/S1877050919300134>

# Design and validation of an exoskeleton for hand rehabilitation in adult patients with rheumatoid arthritis

---

Roberto Moya-Jiménez

Universidad Técnica de Ambato, Av. Los Chasquis 180207 Ambato, Ecuador  
rc.moya@uta.edu.ec  
<https://orcid.org/0000-0002-3918-2935>

Teresa Magal-Royo

Universitat Politècnica de Valencia, Camino de Vera s/n 46022 Valencia, España  
<https://orcid.org/0000-0002-7640-6264>

Michelle Flores

Universidad Central del Ecuador, Av. Universitaria 170129 Quito, Ecuador  
<https://orcid.org/0000-0002-0433-8741>

Mario Caiza

Universidad Central del Ecuador, Av. Universitaria 170129 Quito, Ecuador

## Abstract

The development of this article considers the own research carried out, which is oriented to the design of a functional prototype of a hand exoskeleton for the rehabilitation of older adults with Rheumatoid Arthritis (RA). The results of each study have allowed to follow a line of development until reaching the creation of a portable hand exoskeleton composed of an adapted exoglove that allows the autonomous flexion-extension of the fingers with some opposition tension, which facilitates the physiotherapist the management of rehabilitation processes through medical/therapeutic exercises that contribute to reduce the pain caused by the disease. The hardware components of the prototype are connected to a mobile application through the microcontroller of the electronic architecture that complements the system. The degree of usability of the presented model considers a patient-centered approach, therefore a methodology was designed and developed for the validation of the ARTH-aid System ExoGlove, involving a generic group of 30 patients and an expert group of physiotherapists assigned to a residence for aging adults in Quito, Ecuador.

## Keywords

Hand rheumatoid arthritis, exoskeleton, Physical rehabilitation, rapid prototyping, Computer-aided design.

## Introduction

With the advances of robots within the medical industry, the interest in exoskeletons has been increasing due to the demand for assistive technologies, and to satisfy the concerns of the growing aging society [9]. Likewise, several studies have shown that the use of exoskeletons in assisted therapies provide significant and favorable results in patients' lives [3, 4, 7].

In this context, the investigation was intensified for the creation of a functional prototype that aligns with the expectations of improving the quality of life of patients. In this regard, among the innovative technologies that were investigated [9], it was identified that an exoskeleton specifically designed for hand rehabilitation should consider ergonomic, functional, and therapeutic aspects needed in the treatment of RA.

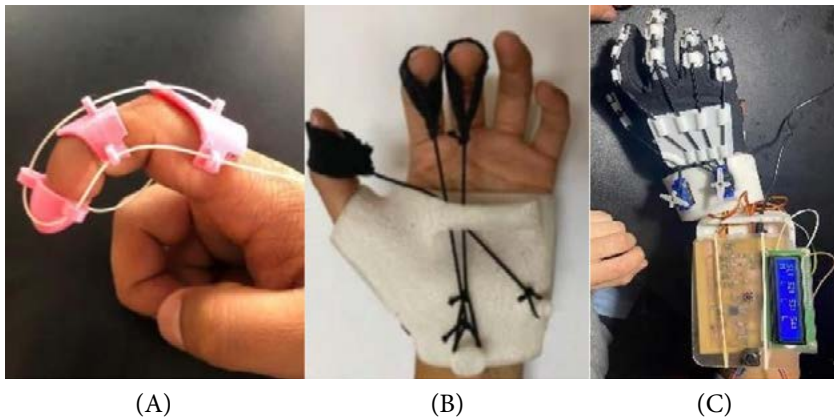
In "Hand Exoskeleton Design for the Rehabilitation of Patients with Rheumatoid Arthritis" the projects and prototypes of hand exoskeletons of the last five years existing on the market were exhaustively reviewed [7]. Of the 36 exoskeletons, alternatives made of rigid modules or joints were discussed, but also flexible gloves that are introduced in the early stages of RA where the patient has not been invasively compromised yet.

Due to the scarcity of evidence from well-controlled projects with conclusive results regarding the effectiveness of the proposed interventions, it is imperative to focus rehabilitation techniques on training adapted to the patient's needs in terms of the practice of the exercises needed to reduce the suffering caused by the condition [6,7].

In reference to this, there are external factors in current hand rehabilitation therapies which are not at all favorable for this process. These factors include the low availability of specialists, assistants, and equipment, as well as the difficulties in transporting the patient to the centers where these treatments are performed [11]. These factors increase the cost and increase the difficulty of the patient's recovery process.

According to the preliminary studies for the verification of the efficacy of various prototypes, the purpose is to establish a rehabilitation for patients in the early stages of the disease considering the results and previous data. The advances made in the application of rapid prototyping (RP) technologies in the construction of customized devices for the physical rehabilitation of hands [5,8] highlight important findings of the study which suggest, in terms of mobility, that the articular arc of movement is significantly improved, joint stiffness is reduced, muscle stretch is increased, and there is a reduction of pain during finger movement, especially in those users suffering from mobility problems in phalanges and metacarpal bones.

**Figure 1**  
**Assembly of previous prototypes**



(A) finger module [6]; (B) prototype for tension fingers [8]; (C) exo-glove prototype [9]

Following and extending the scope of this line of research, it will be described the development of a user-centered methodology that, on one hand will allow to evaluate the factors related to the patient-device interaction; and on the other hand will consider the efficiency of the prototype, which must be accepted by the user and by the physiotherapist to guarantee the degree of usability of the system in a selected group of users.

By virtue of this, the prototype obtained in this article is configured based on requirements acquired from explorations in this field. Consequently, it will be possible to design larger scale tests on a group of patients with hand rehabilitation needs.

## Methodology

For the development and construction of the glove, an experimental methodology was used to gather information from the approach to the final user. The work was carried out with the express permission of the medical management of “Santa Catalina Labouré”, an adult home located in Quito, Ecuador <sup>1</sup>. The residence has 70 patients, and the organization is dedicated to the care of the elderly.

The home also has a population of expert physiotherapists and a rehabilitation area that includes a specific place for the treatment of patients with RA. With the support of the physiotherapist and based on a clinical history, was established a sample of 30 patients (10 men and 20 women) whose ages range from 65 to 104 years, and who attend to the interests of the research, which means that all of them present mild or severe joint pain due to the disease.

With this generic group, it should be noted that the main reason for purposive sampling is to generate estimates or inferences for this specific population, but it is not related to a totally probabilistic character in which a population is represented. This decision has been taken in consideration of the particularities and characteristics of the case study.

Although the intervention group is not copious, with a confidence level of 95% (based on the total population of the residence) it is high enough to detect most of the usability problems.

---

1 Hogar de Ancianos “Santa Catalina Labouré”. In: <https://hdlcecuador.com/index.php/2020/12/15/comunidad-santa-catalina-laboure/>

For the first phase of the methodology, the starting point was the experience of previous research and the recommended principles for validation with users were also considered, as detailed below [8]:

- Early diagnosis
- Provide an opportune treatment
- Indicate appropriate management by the physiotherapist
- Focusing on the object of remission or low activity level
- Individualized management
- Periodic follow-up of the process
- Consider comorbidity
- Adapt to the reality of medical/therapeutic practice

Here also takes place the registration of the patients which allows to classify them as shown in table 1:

**Table 1**  
**Classification by age range. Source: Own elaboration (2021)**

Age ranges	No. patients
65-74	4
75-84	13
85-94	11
95-104	2
Total	30

In general, it has been clinically demonstrated that about 70% of patients start the disease in an incipient stage where the patient has not yet been invasively involved [2]. It has also been pointed out that as age advances, the consequences of RA in patients are more evident, with higher rates of polyarticular involvement in women. Therefore, classification by age enables a more accurate estimation of the impact of the exoskeleton on the patient, and in this case, there is a greater number of women to be analyzed.

To provide an appropriate treatment, it is considered that functional disability limits the activities of daily living of the older adult, who often requires the assistance of others [10]. Therefore, to the patient-group is applied a questionnaire adapted from the HAQ (Health Assessment Questionnaire) with 17 items grouped in 6 areas that contain daily activities that will permit the evaluation of the patient's dependence (or independence) to execute the activities. The questionnaire also includes corrective questions that can modify (correct) the score of the areas they affect.

**Table 2**  
**HAQ adapted questionnaire. Source: Own elaboration (2021)**

*Durante la última semana, ¿ha sido usted capaz de...*

Vestirse y asearse	1	¿Vestirse solo, incluyendo abrocharse los botones y atarse los cordones de los zapatos?					
		<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>
	2	¿Ducharse y lavar su cabello / cabeza?					
		<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>
3	¿Lavarse y secarse todo el cuerpo?						
	<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>	
4	¿Utilizar anillos?						
	<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>	
Levantarse	5	¿Levantarse de una silla sin brazos?					
		<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>
6	¿Acostarse y levantarse de la cama?						
	<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>	
Comer	7	¿Cortar un filete de carne?					
		<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>
	8	¿Abrir un cartón o paquete de comida?					
<b>Incapaz de hacerlo</b>		0	1	2	3	<b>Sin dificultad</b>	
9	¿Agarrar una taza o servirse la bebida?						
	<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>	









Alcanzar	10	¿Alcanzar un objeto pesado de una estantería elevada?					
		<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>
	11	¿Agacharse y recoger ropa del suelo?					
		<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>
Preñión	12	¿Abrir tarros?					
		<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>
	13	¿Abrir y cerrar los grifos?					
		<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>
	14	¿Sujetar llaves de cerraduras?					
<b>Incapaz de hacerlo</b>		0	1	2	3	<b>Sin dificultad</b>	
15	¿Exprimir una toalla?						
	<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>	
Otras	16	¿Hacer tareas de casa como barrer o lavar los platos?					
		<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>
	17	¿Presionar botones del control remoto u otro dispositivo?					
		<b>Incapaz de hacerlo</b>	0	1	2	3	<b>Sin dificultad</b>

In the first instance, the following corollary has been obtained after data analysis in the Excel statistical program:

1. 50% of women have + 6.66% difficulty in dressing and grooming, and most of these women are between 70 and 85 years of age.
2. More than 60% of patients (both men and women) have no difficulty getting out of chairs or out of bed.
3. When it comes to eating, men have the greatest difficulties (3% more than women). Likewise, this percentage is present in 30% of men when reaching for objects.
4. Only 40% of patients (among men and women) indicate a high level of difficulty in grasping objects.
5. To perform other activities, such as housework, men have a -3% difficulty in reaching for objects.



**Figure 2**  
**Screenshot of data obtained from the HAQ questionnaire**

		Nivel de dificultad								Preg. corr.	Utens. de ap.	FOTOS
-	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	
1	M	73	1	0	1	1	1	1	1	vertizo y are-azro (2)	Bartón, muletaz	
2	F	82	1	0	1	1	1	1	1	vertizo y are-azro (2)	Bartón, muletaz	
3	M	73	1	0	1	1	1	1	1	vertizo y are-azro (2)	Arianta baña	
4	M	78	3	3	3	3	3	3	3	vertizo y are-azro	Bartón, muletaz	
5	M	100	3	3	3	3	3	3	3	vertizo y are-azro	Bartón, muletaz	
6	M	85	0	0	0	0	1	1	1	vertizo y are-azro (1)	Abriar tarraz	
7	M	76	0	0	0	0	0	0	0	vertizo y are-azro (1)	Abriar tarraz	
8	M	84	3	3	3	3	3	3	3	vertizo y are-azro	Bartón, muletaz	

Source: Own elaboration (2021)

These clarifications lead to the reflection and understanding of the responses, so that a correct measurement of the maximum degrees of flexion and extension, as well as the pressure applied, is required. This is considering that the movement of the fingers is measured in terms of the maximum degrees of flexion-extension and requires standardization to ensure the correct operation of the prototype by all physiotherapists.

**Table 3**  
**General technique for measuring MCP flexion-extension of the fingers**

Articulation	Description of the technique	Type of bending and/or movement
<b>All fingers of the hand - metacarpophalangeal joint</b>	Position: patient seated, elbow at 90°, with the hand and forearm resting on a table, with the wrist in the 0 position and the thumb in the 0 position.	<ul style="list-style-type: none"> <li>• Normal values: MCP flexion of the fingers of the hand: 0-90°.</li> <li>• Normal values: MCP extension of the fingers of the hand: 0-30° MCP extension of the fingers of the hand: 0-30°.</li> </ul>
<b>Thumb - metacarpophalangeal joint</b>		<ul style="list-style-type: none"> <li>• Normal values: MCF thumb flexion: 0-50°.</li> </ul>

Source: Own elaboration (2021)

To enter phase two, which corresponds to therapy and measurements, it is necessary to establish an organized routine with weekly exercises (two or three days a week) to be realized with the exoskeleton and it is essential that the device records the range of motion of the fingers with a frequency of once a week. At this point it should be noted that, in general, goniometers are not accurate, and their measurements are subject to the way in which the goniometer is positioned on the patient [8] and the prototype must solve this factor of measurement to obtain precision data that can guarantee the validation of the prototype.

With the focus on motor recovery and passive/active hand training, the exercises assigned to each patient were approved by the physiotherapist and they follow the guidelines of the traditional exercises recommended by the American Physical Therapy Association (APTA) [1]. In this regard, the exercises are simple to do by the users and are efficient according to the expert's prescription, who recommends mainly those low-impact exercises that encourage muscle tone and flexibility.

The therapies are defined in a time interval (5 minutes minimum and 10 minutes maximum) that takes care of the patient's integrity avoiding exposing him/her to overexertion and fatigue. At the rehabilitation level, it is required that the treatment sessions with the patient have an appropriate posture and effort control, as well as a periodic measurement of the flexion-extension range of movement.

After the weeks of therapy, a third and final phase of evaluation of the 30 patients is proposed, focusing on identifying the impact of the system on the users and determining how it contributes to the early improvement of mobility and reduction of articular pain. The summative evaluation plus the application of satisfaction questionnaires to both groups of users will verify the effectiveness of the prototype. The reason for evaluating both patients and experts is due to: (1) the senility of the patients, from an ethical point of view, the number of clinical participants involved is limited and (2) to avoid locating technical errors. It is necessary to obtain feedback in these two groups.

The outcomes are transformed into quantitative metrics, with the purpose of improving the design of the system (if necessary) and refining the specifications for rehabilitation use.

## Results

The developing of the ARTH-aid System Exo Glove is a result of the requirements necessary for the treatment of the disease and brings together the main features obtained from the research conducted, as well as from the approach to a group of patients and experts attending to the needs of the research. The results of the application of the methodology designed for validation will help to make decisions to improve the activities of patients in relation to the accompanying expert and the use of the exoskeleton intended for this process.

Table 2 shows each of the necessary requirements for the construction of the prototype.

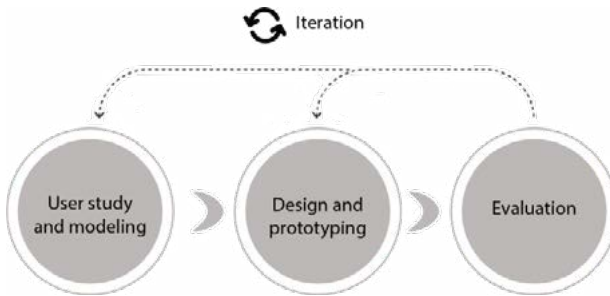
**Table 2**  
**Requirements of the ARTH-aid System Exo Glove.**

Requirements		Description	
<b>Of use</b>	A)	Handling	It is transportable and can be applied in any context.
	B)	Versatility	Must be used in different user's hands
	C)	Minimum weight	The structure complies with an indicated weight limit.
	D)	Maintenance	Modules requiring care are independent and can be separated from each other
	E)	Easy control interface	Proper fitting and removal procedure
<b>Of therapy</b>	F)	Strength control	Must activate resistance in both flexion and extension of the fingers.
	G)	Flexibility	It must provide the necessary number of GDLs (3).
	H)	Gripping capacity	It is necessary to provide the possibility to perform different types of gripping (the most common)
	I)	Open palm	Since some grasping skills are anticipated, the user's palm and fingertips should be left as free as possible.
	J)	Simplified activation	Activation of the flexion-extension opposition is regulated.
	K)	Response time	Movements are immediately captured by the electronic system.
	L)	Sensibility	Flexosensors are precise in their measurements or have a low margin of error.
	M)	Materialization	Manufacturing allows customization of the device using rapid prototyping tools
	N)	Robustness	It is constructed based on an alloy of 3D printing technology and textile materials.

Source: Own elaboration (2021)

The development started with an innovation for combining digital tools within the use of RP by using 3D printing in PLA material as the main source. The ARTH-aid System Exo Glove consists of an exoglove that allows the polyarticular movement of the user's fingers of any age group and gender; and includes the use of a module based on the technology of the boa to generate the necessary tension to the bending movements. The system is complemented by the "BlueTerm" software application linked to an electronic architecture. The design process is summarized in the diagram below.

**Figure 3**  
**Design process**



Source: Own elaboration (2021)

The prototype is designed with 3 GDL, which were verified in previous studies and in such a way it helps physiotherapists in the execution of a previously defined therapy as well as to collect the data of the therapy and the evolution of each patient. The system makes possible that all this information can be stored in a registry and provides the expert with the possibility to create a historical reference of the patient's evolution over time.

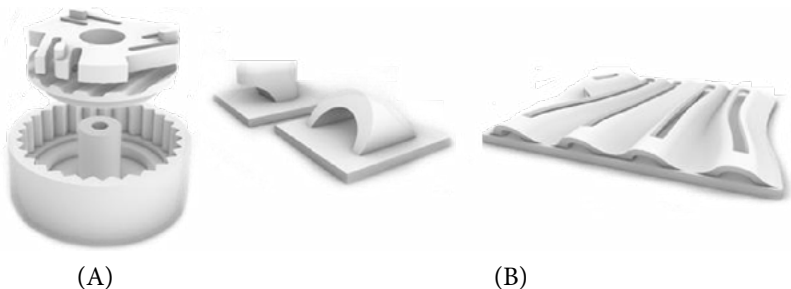
### ***System modules***

According to the main objective of the mechanical design, which is to achieve the independent and organic movement of the finger, the

modules were designed with a specific mechanism so that the patient can adapt to the opposition of force generated by the tension wires that are attached along each finger.

- Knob: this module oversees controlling the tension of the threads that generate opposition to finger flexion. This option was chosen instead of servomotors for its simplicity of control and space saving.
- 3D modules: they are attached to the glove and allow the placement of a tensioning thread along each finger.

**Figure 4**  
**Glove components**



(A) Control knob; (B) Tension wire modules. Source: Own elaboration (2021)

To provide a control adjustment of the system in relation to the efforts required for each patient, the construction of the exo-glove uses gauges to detect the improvement of mobility and establish a therapeutically controlled rehabilitation by the expert(s).

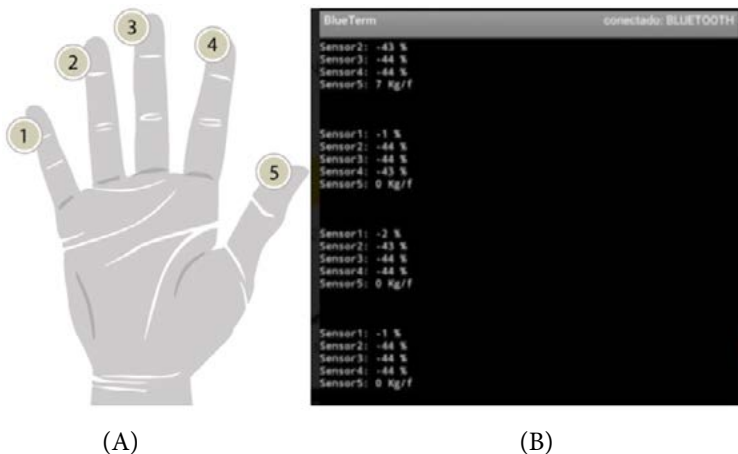
Sensors: four flexosensors and a force sensor are in the palmar area to determine the degrees of the flexion arc and the pressure generated by the patient. In essence, these elements work as a resistance, therefore, the more flexion increases the value of the resistance. Thanks to a voltage divider that is interfaced to a microcontroller, it is possible to read the sensor and convert its value into an equivalent value in per-

centage to analyze the data during the movement and behavior of the user's fingers while performing the task.

The choice of thin and low-cost sensors allows the model to be accessible and does not compromise the aesthetics and robustness of the glove, and it can be easily used in different contexts. It has an overall low power consumption and a 3.7-volt rechargeable battery. The size is suitable for placement in the protective case of the device.

The software application was not originally planned for this research but given the possibilities of adapting it to the physiotherapist's needs, it was decided to use the Bluetooth option with the "BlueTerm" mobile application, which offers greater ease of data registration for each patient. The sensor reading indicates the pressure produced by the thumb and is transformed into a value in kg/f dimensions while the remaining four sensors (1 to 4) indicate the percentage of flexion.

Figure 5



(A) Finger numbering for each sensor. (B) Screenshot of the serialized reading in the mobile application.

Source: Own elaboration (2021)

## *Construction of the ARTH-aid System Exo Glove*

The formal design of the system is based on the verification of the effectiveness of several previous prototypes, considering the GDL required for the treatment of the disease, the mechanical analysis of the loads, the main stresses applied and the resistance according to the shape and the material.

The first step was the manufacture of the glove, with a material that presents a certain degree of rigidity to protect the sensors and ensure the correct functioning of the sensors. The glove, which is flat, is provided with internal pockets to place the sensors so that they can be changed when required. In addition, all the 3D modules that are part of the glove's mechanical system are attached to the counter palm.

**Figure 6**  
**Assembly and physical testing of the prototype.**



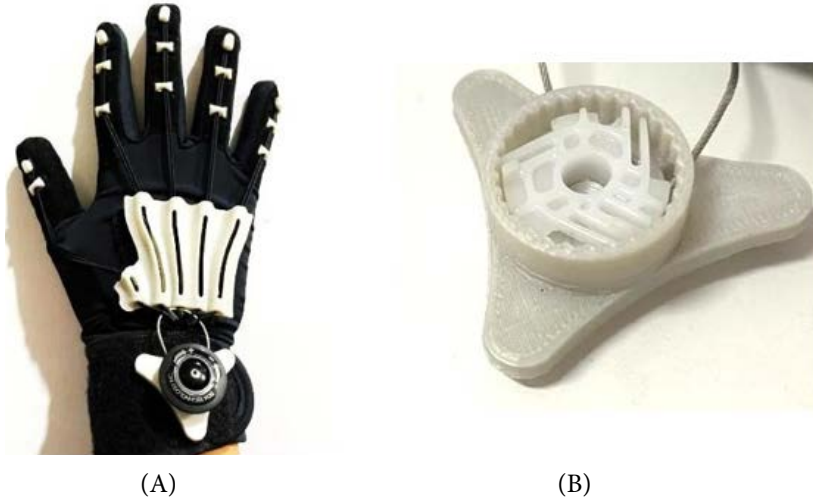
Source: Own elaboration (2021)

Concerning the assembly of the electronic components, they were all soldered on a 12 x 12 (cm) PCB. Giving an approximate weight of 100 grams, including the weight of the glove. With the board complete, and the programming ready, a protective case was built to contain the components and protect them from bumps or falls that could cause da-



mage to the prototype. The case is the most rugged part of the prototype; however, it is not very robust since all these aspects were designed in such a way that they do not compromise the aesthetics of the prototype and that they fulfill their function.

**Figure 7**



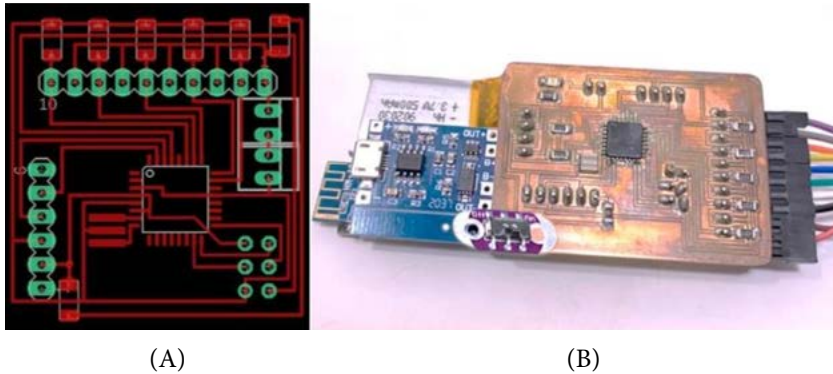
(A) Exoskeleton and (B) 3D printed module adapted for force control  
Source: Own elaboration (2021)

In general, the design of the modules and the electronic architecture is designed to avoid any inconvenience to patients during the use of the glove and to integrate all the components without further complication, which also benefits the manufacturing and installation of the system and contributes to the regular use of the glove, which consequently will have a positive impact on the rehabilitation of the elderly.

The functionality of the ARTH-aid System Exo Glove is oriented towards the patient and the physiotherapist so that it can be understood and does not require a considerable amount of time to configure and

place the system. During all sessions, the set of sensors used for this purpose are collected in the mobile application.

**Figure 8**  
**Electronic architecture**



(A) PCB layout and (B) PCB with components  
Source: Own elaboration (2021)

## Conclusions

The design and application of the methodology according to the rationale of several inquiries, definitely presents several advantages for the rehabilitation of patients and also allowed using the experimental data to apply it in the generation of a complete system that is adaptable to the patient's needs and satisfies the requirements without losing the formal aspects (design and ergonomics), the medical aspects and the functional issues (mechanisms, electronic instrumentation and control systems).

The approach to patients and experts within the Ecuadorian context has been satisfactory and has raised a wide interest to extend the scope of this line of research and the prototype. A deeper understanding of the functioning of the hand as a tool for life has been the driving force for the creation of the exoskeleton and with the use of different

technologies linked to rapid prototyping it is possible to optimize the design and get closer to a model that can be scaled even to the self-manufacturing of the exoskeleton.

In the future the validation methodology is intended to be applied with the prototype created and further researchers who wish to venture into the field of rehabilitation of RA patients are encouraged to use the methodology in collaboration with an expert group to guide the process. There are several factors that can support or inhibit the rehabilitation of the hand of patients with RA and therefore expert guidance is essential to provide a complete system that responds to the main needs of the end user.

## Acknowledgments

This research has been developed thanks to the predoctoral research carried out for the doctoral thesis “Estudio sistémico de estructuras exoesqueléticas de la mano para la rehabilitación de la atrofia muscular debido a la artritis en personas adultas” de la Universitat Politècnica de Valencia.

## References

- [1] APTA. (2019). American Physical Therapy Association. Obtenido de: <https://www.apta.org/> (21/06/2021).
- [2] Batlle Gualda, E., Mínguez Vega, M., Bernabéu Gonzáles, P., y Panadero Tintero, G. (2013). Enfermedades Reumáticas. Valencia: Ibáñez & Plaza Asociados S.L. ISBN: 978-84-8882310-6.
- [3] Borghetti, M., Sardini, E., y Serpelloni, M. (2013). Sensorized glove for measuring hand finger flexion for rehabilitation purposes. IEEE Transactions on Instrumentation and Measurement, Vol. 62(12), 3308-3314. doi: 10.1109/tim.2013.2272848.
- [4] Chávez, & et al. (2010). Exoesqueletos para potenciar las capacidades humanas y apoyar la rehabilitación. Colombia: Revista Ingeniería Biomédica, Vol.4(7), 63-73.

- [5] Kang, B. B., In, H., & Cho, K. (2012). Force transmission in joint-less tendon driven wearable robotic Hand. In Proceedings 12th International Conference on Control, Automation and Systems (ICCAS), 1853-1858.
- [6] Moya, R., & Magal-Royo, T. (2019). Diseño y prototipado de un dispositivo de rehabilitación para la artritis reumatoide de mano. *Tsantsa. Revista De Investigaciones Artísticas*, Vol. (7), 233-240. ISBN, 1390-8448.
- [7] Moya-Jiménez R., Magal-Royo T., Ponce D., Flores M., Caiza M. (2020). Hand Exoskeleton Design for the Rehabilitation of Patients with Rheumatoid Arthritis. In: Rodríguez Morales G., Fonseca C. E.R., Salgado J.P., Pérez-Gosende P., Orellana Cordeiro M., Berrezueta S. (eds) *Information and Communication Technologies. TICEC 2020. Communications in Computer and Information Science*, vol 1307. Springer, Cham. DOI: [https://doi.org/10.1007/9783-030-62833-8\\_2](https://doi.org/10.1007/9783-030-62833-8_2)
- [9] Moya-Jiménez, R., & Magal-Royo, T. (2020). Diseño y prototipado de un dispositivo de rehabilitación para la artritis reumatoide de mano con técnicas de prototipado rápido. *Revista Ibérica de Sistemas e Tecnologias de Informação*, (E33), 152-163.
- [10] Moya-Jiménez, R., Magal-Royo, T., & Ponce, D. (2021). Tecnologías digitales innovadoras para el diseño de exoesqueletos para la mano aplicados en la rehabilitación de pacientes con artritis crónica. In *Proceedings INNODOCT/20. International Conference on Innovation, Documentation and Education* (pp. 537-543). Editorial Universitat Politècnica de València. DOI: <http://dx.doi.org/10.4995/INN2020.2020.11866>
- [11] Smolen, J. S., Landewé, R., Bijlsma, J., Burmester, G., Chatzidionysiou, K., Dougados, M., y Aletaha, D. (2017). EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Annals of the rheumatic diseases*, 76(6), 960-977. doi:10.1136/annrheumdis-2016-210715.
- [12] Viaux C.G.G. (2007) Penetración de las Tecnologías de la Información en las Realidades Hospitalarias de los Países Latinoamericanos y del Caribe. In: Müller-Karger C., Wong S., La Cruz A. (eds). *IV Latin American Congress on Biomedical Engineering*, Vol. 18. DOI: 10.1007/978-3-540-74471-9\_263.
- [13] 10.1007/978-3-540-74471-9\_263.

# Understanding the Capacitated Vehicle Routing Problem, through the Hamiltonian circuits<sup>1</sup>

---

Saul Figueroa

Yachay Tech University, Ecuador  
saul.figueroa@yachaytech.edu.ec  
<https://orcid.org/0000-0001-9395-5903>

Franklin De la Cruz

Yachay Tech University, Ecuador  
franklin.de@yachaytech.edu.ec  
<https://orcid.org/0000-0002-0429-5219>

Joseline García

Yachay Tech University, Ecuador  
joseline.garcia@yachaytech.edu.ec  
<https://orcid.org/0000-0001-9190-7852>

Franklin Camacho

Yachay Tech University, Ecuador  
fcamacho@yachaytech.edu.ec  
<https://orcid.org/0000-0001-7802-5687>

## Abstract

Graph theory is helpful to understand everything about networks and modeling real problems. Therefore, it is important to tackle every knowledge needed about graphs to understand how the Capacity Vehicle Routing Problem (CVRP) works. Then, it is explained how to solve the base problem, which in fact is the Vehicle Routing Problem (VRP), by using the concepts of Hamiltonian's paths and circuits. A meta-heuristic method called Simulated Annealing (SA) will be used for its implementation, which is a method to find optimal

- 
- 1 Supported by Vice Chancellery of Research and Innovation of Yachay Tech University which has partially funded this work through the project *Qualitative Decisions* code MATH19-13.

solutions to combinatorial optimization problems such as the VRP. To this, it is presented a Python implementation of this SA algorithm to execute tests. With all this information in mind, a flexible distribution of zones is presented in relation to the number of vehicles available. The main characteristic of these zones is the non-intersection of the Hamiltonian paths presented by the route found with the proposed implementation. Finally, there will pose some strategies to tackle the CVRP considering new constraints in the implementation presented.

## Keywords

Hamiltonian Circuits, VRP, Simulated Annealing, Capacitated Vehicle Routing Problem, Zone Mechanism.

## Introduction

The Königsberg bridge problem asks if the seven bridges of the city of Königsberg can be traveled one by one without travel twice for the same bridge. Leonhard Euler solved this problem in 1736, and his resolution was the origin of graph theory [1]. Since that time, graph theory has been used to solve everyday problems.

A common problem in companies is leading with transportation operations. For example, delivering goods to customers need to have an optimal load distribution. These types of problems are known as Vehicle Routing Problems.

This problem has been studied for a long time, for example, in 1981 Golden et al wrote “Capacitated Arc Routing Problems” [9]. Recently, Rojas et al., in [2], proposed a model for carriers companies to optimize route planning. They presented a distribution model for route planning, formulated as a Capacitated Vehicle Routing Problem for Carriers (CVRP<sub>fC</sub>), which is proposed to address the distribution scenario of an actual situation with sequence restrictions associated with its delivery and vehicle-storing locations. In the same way, in [10] and [11] are presented two different ways to get an optimal solution for the CVRP. In the first one, it is presented an implementation of an Ant colony optimization algorithm with support from the SA algorithm. In the second one, there is established a relation between the route to follow and the customers, providing a decision step during the process to reduce the number of customers to serve according to the price they are able to pay.

In this work, the authors introduce this problem by planting the most straightforward problem related to the Vehicle Routing Problem (VRP). It is about finding the minor path to travel for a given set of nodes. Here we will use the concept of hamiltonian path to visit all the vertices of the graph once with the restriction that we must use the smallest edge weight to do it. For the implementation, it will be used a specific distribution of zones based on the number of vehicles available. The main characteristic of this distribution is that there is non-intersection of the Hamiltonian paths. This ensures that the polygon traced by the routes of each zone does not have an intersection with any of the other polygons traced by its respective zone. As a consequence, an effective strategy is generated to further optimize the solution to this problem. In addition, it will be use a meta-heuristic method called Simulated Annealing which help to find the best solution for VRP.

In order to explain the basis of Capacitated Vehicle Routing Problem (CVRP), in section 2, we will describe the basics notions about graph theory, its properties, and types. Starting with simple graphs, directed graphs, graphs representation, until hamiltonian paths, circuits, and theorems such as Dirac's theorem and Ore's theorem. Also, the base problem is posed. In section 3, it is presented a strategy to solve the base problem by finding the smallest circuit given vertices and a number of vehicles to travel by applying the simulated annealing algorithm. The next section, 4, is shown the results obtained for three tests provided. Finally, in section 5, the author concludes that the implementation proposed solves the base problem and led to understand the CVRP. In addition, some mechanisms and new strategies for future works are presented in the subsection 5.1.

## **Preliminaries**

### *Basic Notions*

Graphs are discrete structures consisting of vertices and edges that connect these vertices. A graph  $G = (V, E)$  consists of  $V$ , a nonempty

set of vertices (or nodes) and  $E$ , a set of edges. Each edge has either one or two vertices associated with it, called its endpoints. It's called a simple graph when there are not multiple edges connecting two vertices; each edge connects two different vertices. Also, no edge connects a vertex to itself; it means there are no loops. A directed graph (or digraph)  $G = (V, E)$  consists of a nonempty set of vertices  $V$  and a set of directed edges (or arcs)  $E$ . The directed edge associated with the ordered pair  $(u, v)$  starts at  $u$  and ends at  $v$ .

A path is a sequence of edges that begins at a vertex of a graph and travels from vertex to vertex along the edges of the graph. A simple path in a graph  $G$  that passes through every vertex exactly once is called a Hamilton path, and a simple circuit in a graph  $G$  that passes through every vertex exactly once is called a Hamilton circuit. Both Dirac's and Ore's theorem provide sufficient conditions for a simple connected graph to have a Hamilton circuit. For more information on graphs, as well as the proof of the following theorems, to see [3].

**Theorem 1 (Dirac's Theorem).** *If  $G$  is a simple graph with  $n$  vertices with  $n \geq 3$  such that the degree of every vertex in  $G$  is at least  $n/2$ , then  $G$  has a Hamilton circuit.*

**Theorem 2 (Ore's Theorem).** *If  $G$  is a simple graph with  $n$  vertices with  $n \geq 3$  such that the  $\deg(u) + \deg(v) \geq n$  for every pair of nonadjacent vertices  $u$  and  $v$  in  $G$ , then  $G$  has a Hamilton circuit.*

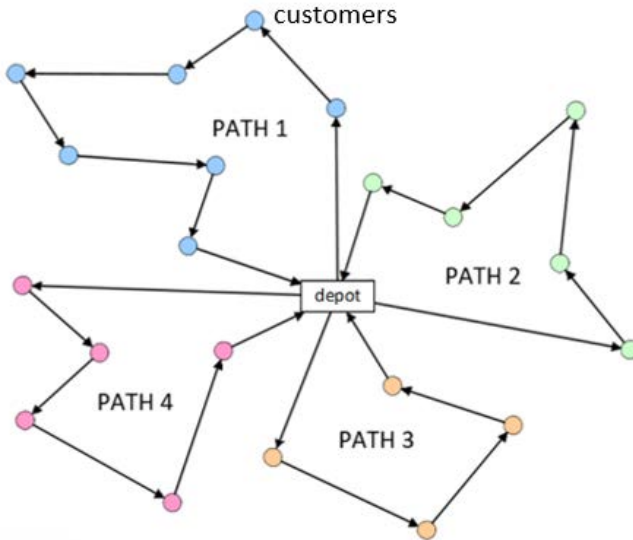
## What is Vehicle Routing Problem and why it is important

**Base Problem:** The Vehicle Routing Problem (VRP) separates the zone into different subsets or zones. The first node visited is the one that is near the center inside each zone. Then a path is done visiting each node of a determined zone. This path should be minimized as possible. The criterion is to minimize the total distance traveled, adding all the paths mentioned before in the sum. Hamiltonian in graph theory refers



to the fact that each node has to be visited exactly once. Moreover, this base problem (VRP) can be illustrated as a graph representation, where all the previously described components take part (see Fig. 1).

**Figure 1**  
**Base problem representation**



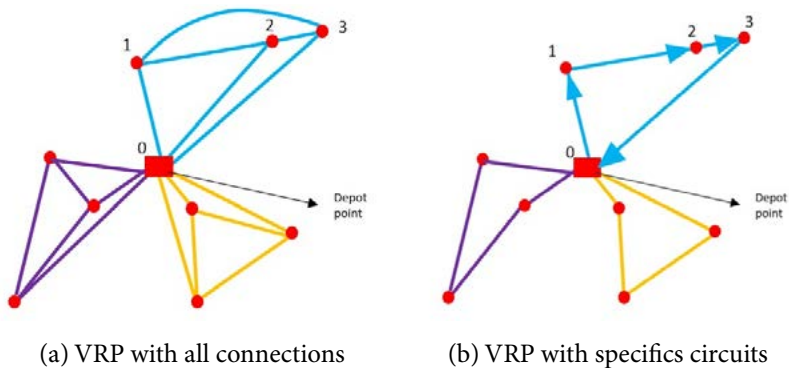
**VRP Importance:** Transportation is an important part of logistic operations in a company, which is aimed to deliver the goods required by customers to the right place at the right time. This is the reason why costs associated with transportation represent a significant part of the value of each product. The distribution costs may represent more than 25% of the value of the sales product. Therefore, transportation service is an alternative for reducing the cost distribution of a company which is associated with the efficiency of its route planning. By this, the solution to the base problem will allow an efficient and optimal route planning [2].

## Methodology

### *Hamiltonian Circuits and VRP Relation*

It is possible to pose the base problem Vehicle Routing Problem through a graph divided into different regions (zones). Every Hamiltonian circuit in each zone would consist of a possible solution to the problem. This could be represented by various sub-graphs with some possible hamiltonian circuits that share the same starting and ending vertex, which is commonly denominated “depot point” at the VPR. This representation can be seen below in Fig. 2a.

**Figure 2**  
**VRP graph representations**



If a hamiltonian circuit that minimizes the total distance traveled for each vehicle is found, then the problem would be solved. This graphical solution is illustrated in Fig. 2b. In addition, each region (subgraph) can be analyzed by the two theorems proposed in the previous section. These let assure the existence of hamiltonian circuits in the problem. For instance, let consider the blue subgraph from Fig. 2a. The respective adjacency list will be given by:

**Table 1**  
**Adjacency list for Dirac's Theorem**

Vertex	Adjacent vertices	Degree of Vertex
0	1 2 3	3
1	0 2 3	3
2	0 1 3	3
3	0 1 2	3

Due to the fact that  $\frac{\text{number of vertices}}{2} = \frac{4}{2} = 2$  and by the Theorem 1, This subgraph has a Hamilton circuit.

In the same way, let consider the adjacency matrix from the same subgraph, given as:

$$G_R = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

where each 1 value in the position  $u, v$  represent the connection between vertex  $u$  and  $v$  with  $i, j = 0, 1, 2, 3$ . Then, the following table shows the degree of  $u$  and  $v$  for every pair of nonadjacent vertices:

**Table 2**  
**Adjacency matrix for Ore's Theorem.**

i,j	deg(u)	deg(v)	deg(u)+deg(v)
0	3	3	6
1	3	3	6
2	3	3	6
3	3	3	6

Thus, by the results in Table 2 and Theorem 2, the subgraph  $G$  has a Hamilton circuit. In the same way, each subgraph has a hamiltonian circuit, so it is guaranteed that at least one hamiltonian circuit exists for each zone in the VRP.

### *Simulated Annealing to solve the VRP*

In computational science, one of the best methods to find solutions to combinatorial optimization problems is the implementation of meta-heuristics. In [4], Bayram and Sahi state that this method is a general optimization technique whose objective is to search for new solutions from an initial one. Moreover, it seeks the best approximation to the optimal solution of a given problem. Since the study problem qualifies as a combinatorial optimization problem, it will be necessary to use one of the many meta-heuristic algorithms that exist today, Simulated Annealing.

Simulated Annealing (SA) was proposed in 1983 by Kirkpatrick and colleges as a technique for finding globally optimal solutions, based on an analogy with a metallurgical process known as “Annealing”, to see [5]. This process consists of heating a solid at high temperatures until its atomic structure is dispersed. This is followed by a controlled cooling process, in which the temperature is slowly reduced to restructure the atoms of the solid. Finally, a stable configuration is achieved with a minimum of energy, for better reading, see [6, 7].

Initially, the algorithm generates a random solution and an initial temperature. Then, within a loop (cooling process), start looking for better solutions that are obtained through random modifications to the initial solution. Subsequently, considering the objective function (a function that evaluates the solutions found), in each iteration, the algorithm verifies whether the new solution obtained is favorable or not in order to replace it with the current solution. For this, the SA method makes use of a probability criterion called the Metropolis Criterion, which will indicate whether the new solution should be accepted or not. Even if the solution is not optimal, it may still be accepted if the tempe-

perature is too high. This last component is fundamental in the algorithm since it allows the new solutions not to be stuck in local minimum or maximum but to escape from these to obtain globally optimal solutions.

**SA components description** To understand how this algorithm solves the VRP, it is necessary to define the components and mechanisms to be used.

*Solution S:* Since the base problem can be represented as a directed graph. The solutions obtained represent the order in which the vertices are visited. In other words, it is a set of vertices that form an optimal Hamiltonian circuit.

*Objective Function  $f(S)$ :* Due to the fact that the principal objective is to obtain the minimum distance to be traveled in each zone, SA must have a function that is responsible for calculating the total distance of each solution obtained to select the optimal solution.

*Iteration Variable Temp:* In the same way, as in the metallurgical annealing process, the algorithm has a loop where the cooling of the system will be performed. For this, a variable that refers to the temperature of the system is needed. Therefore, it will decrease for each iteration with the help of the cooling rate.

*Neighbors  $S_{\text{new}}$ :* This component represents the new solutions obtained from a modification of the current solution. This modification consists of a small change in the order in which the vehicle visit each vertex. In this way, a new Hamiltonian circuit is created. Then, it will be evaluated later by the objective function to verify if they should be accepted or not, considering the probability criterion.

*Cooling Rate:* This factor allows the SA method to simulate the cooling process mentioned above. The cooling ratio consists of a value that is continuously multiplied by the current temperature in each iteration. This value can vary between (0, 1), but it is recommended to use values close to one, as 0.999, since it is important to decrease the temperature as slowly as the metallurgy process.

*Metropolis criterion*  $P(S_{new})$ : Once the neighbor mechanism obtains a new solution, the algorithm needs to check if this solution is better than the previous one. So then, it is necessary to compare the objective function evaluating both solutions. Therefore, there are two possible options, either the new solution is better or worst.

This criterion is taken from the Metropolis Algorithm, which also studies the annealing process. The main function of this criterion is to let the SA method explore and consider even the worst solution to avoid being stuck in some local minimum or maximum, [6]. This escape tool is given by the following probability criterion:

$$P(S_{new}) = \begin{cases} 1 & \text{if } f(S_{new}) < f(S) \\ e^{\left(\frac{-\Delta f}{Temp}\right)} & \text{otherwise} \end{cases} \quad (1)$$

Where:

$S, S_{new}$  are the current solution and the new solution respectively.

$f(S), f(S_{new})$  are the solutions evaluated in the objective function.

$\Delta f$  is the difference between the new and current solution evaluated by the objective function ( $\Delta f = f(S_{new}) - f(S)$ ).

Temp is the current temperature.

Then,  $P(S_{new})$  indicates that a good new solution is accepted with a probability of one. On the other hand, if  $S_{new}$  is a bad solution, it will be accepted with a probability of less than one. Note that the higher the temperature is, the high the probability of accepting bad solutions is.

Having considered the previous interpretation of the SA components to the VRP, a representation of the algorithm is given in the pseudo-code (Algorithm 1), where  $T_f$  represents the final temperature when the best solution is found. Moreover, if the current temperature decreases sufficiently, it will be less or equal to  $T_f$ , so the iteration stops, and a good approximation to an optimal global solution will be given.

In Algorithm 1, the probability criterion only takes action if the new solution  $S_{new}$  is a bad one, since if the  $S_{new}$  is a better one, then it is

totally accepted. In the same way, the random threshold condition lets the algorithm explore new solutions randomly and arbitrarily allow to accept new bad solutions.

**Implementation details of the SA for the VRP** In order to implement the SA to solve the VRP, the authors will use the programming language Python, due to it is an easy software to use, and have multiverse libraries that facilitate the implementation.

---

**Algorithm 1: Simulated Annealing**

---

**Input:** A random solution  $S$ , initial temperature  $Temp > 0$ , final temperature  $T_f$  and a cooling rate  $0 < \alpha < 1$   
**Output:** Optimal solution  $S_{new}$  and a final total distance traveled  $f(S_{new})$

```

f(S) ← Total Dist(S);           /* Evaluate the initial solution */
while  $Temp > T_f$  do
     $S_{new} ← Neighbour(S)$ ;
     $f(S_{new}) ← Total Dist(S_{new})$ ;
    if  $f(S_{new}) < f(S)$  then
         $S ← S_{new}$ ;           /* Accept the new solution */
    else
         $P ← exp(-\Delta f(S) / Temp)$ ;
        if  $P > random(0, 1)$  then
             $S ← S_{new}$ ;
        end
    end
     $Temp ← \alpha \times Temp$ ;   /* Update the temperature */
end

```

---

For this implementation, some relationships must be established between the input (number of vertices) and the algorithm components, such as the temperature, cooling rate, and the number of zones (sectors) to use in the problem.

*Temperature Relation:* Due to the fact that the complexity of the problem increases at the same time that the number of vertices increase, it is necessary to set a higher temperature that allows the algorithm to find a good approximation to the global optimal solution surely. For this,

a cubic relation might be enough to fit the correct temperature according to the input data. Then, we can represent that relation as:

$$\text{Temperature} = (\text{Number of vertices})^3 \quad (2)$$

*Cooling Rate Relation:* In the same way, the cooling rate needs to be closer and closer to one if the number of vertices increases. To this, the following expression represents that relation:

$$\text{Cooling Rate} = 1 - \frac{1}{(10) \cdot (\text{Number of vertices})^2} \quad (3)$$

As the right term is a low number, subtract one and then obtain a closer value to one (0.999, 0.99998...). The more decimals the cooling rate has, the more the algorithm iterates, implying a good optimal solution.

*Number of Zones Relation:* The value of this component could be implemented in different manners. However, to set an easy relation, the number of zones must be established as follow:

$$\text{Number of Zones} = (\text{Number of vertices}) \quad (4)$$

Thus, some intervals of vertices will work with a certain number of zones (due to the floor operation). For example, there will be four zones for the interval of the number of vertices [16, 24]. Then, for [25, 36], the number of zones will result in five. Therefore, this mechanism helps to divide the zones to be visited for each vehicle in an easy manner.

**Zone Mechanism** Due to the fact that the VRP involves traveling in some specific regions for each vehicle, it is important to define a mechanism to find those objective zones. Once the relation between the number of vertices and zones has been set, some vertices for each zone need to be added. For this, each zone could be represented as a set whose elements will be the vertices that are contained in that specific region. Furthermore,

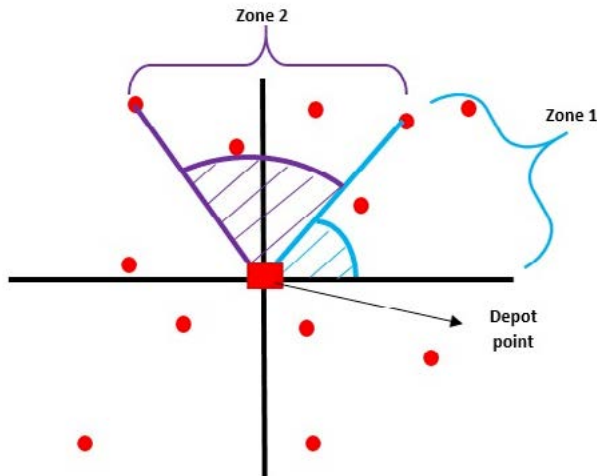


each zone is traveled by only one vehicle and the number of vehicle needed will determinate the number of zone required to the problem.

Then, in order to add a vertex to a zone, a good strategy will be to use a Cartesian plane and the degrees that can be obtained between the origin of this plane (the depot point) and each vertex by applying some trigonometric calculations such as  $\sin$ ,  $\cos$ ,  $\tan$ . As the total angle of the Cartesian plane is 360, dividing it over the number of zones that are needed will result in segmented regions, and thus, those vertices that are within the range of the angle of each zone will belong to that specific region.

In Fig. 3, the zone mechanism is illustrated, those vertices (red points) whose angle are within that specific region, belong to that zone. In the same way, those vertices in the blue and purple zones represent the objective regions to travel for. Then, each zone will be generated by the next range of degree. Starting from the final angle of the previous zone and ending in the next added degree to that starting angle.

**Figure 3**  
**Zone Mechanism representation**



For this implementation, every first zone starts at zero degree. However, this is just one of the whole possible mechanisms to divide the total area to cover over the number of zones needed. More possible mechanism are also presented in section 5.

## Results

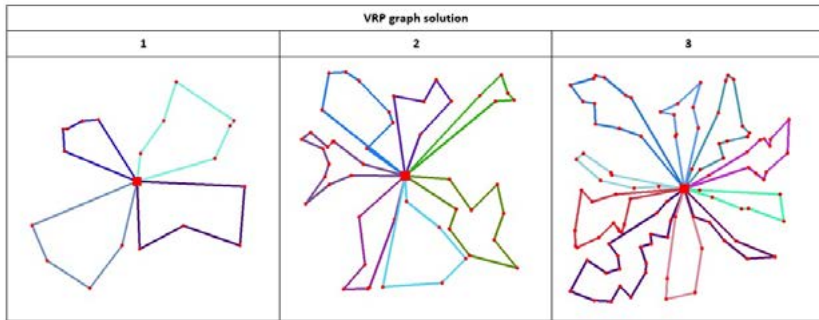
Since the objective of this implementation is to thoroughly understand the base problem and to study its potential and limitations, it is important to perform several tests with various input data. This, will generate a graph representation of the final solution. With the Python implementation completed, let us see how the program solve the VRP.

In Fig. 4, there are three examples for the VRP and the respective results. Every zone has its respective temperature and cooling rate factor which depend on the number of vertices in that specific zone.

**Figure 4**  
SA solving VRP for number of vertices: 20, 50, 100

No	Objective Function	Results:																							
1	2739.332	Parameters				Solution																			
		Temp	$\alpha$	No Vertices	Routes																				
		343	0.9979592	7	0	6	5	4	3	2	1	0													
		216	0.9972222	6	0	1	4	3	2	5	0														
		125	0.9960000	5	0	4	2	3	1	0															
125	0.9960000	5	0	3	1	4	2	0																	
2	4745.489	Parameters				Solution																			
		Temp	$\alpha$	No Vertices	Routes																				
		216	0.9972222	6	0	3	4	1	5	2	0														
		216	0.9972222	6	0	5	3	1	2	4	0														
		512	0.9984375	8	0	1	7	2	5	6	3	4	0												
		2197	0.9994082	13	0	4	10	2	8	12	6	5	1	9	3	11	7	0							
		343	0.9979592	7	0	4	6	1	2	3	5	0													
		216	0.9972222	6	0	3	5	1	4	2	0														
1000	0.9990000	10	0	6	9	8	3	7	2	5	4	1	0												
3	6141.986	Parameters				Solution																			
		Temp	$\alpha$	No Vertices	Routes																				
		2197	0.99941	13	0	12	6	8	2	11	9	5	7	10	3	1	4	0							
		2197	0.99941	13	0	7	12	9	11	1	6	4	10	3	8	5	2	0							
		729	0.99876	9	0	8	3	6	7	2	1	5	4	0											
		1728	0.99931	12	0	3	5	4	7	11	1	2	8	9	6	10	0								
		512	0.99844	8	0	4	5	3	7	1	2	6	0												
		2744	0.99949	14	0	9	4	8	13	1	11	7	2	3	12	10	6	5	0						
		8000	0.99975	20	0	5	11	13	17	18	15	2	3	10	14	8	9	6	4	19	7	16	12	1	0
		216	0.99722	6	0	1	5	3	4	2	0														
343	0.99795	7	0	6	4	3	2	5	1	0															
343	0.99795	7	0	4	6	2	3	1	5	0															

**Figure 5**  
**Graph results for each example**



In Fig. 5, the solutions presented are the optimal for each zone. One important fact to recall, is that the number of vehicle which participate in the problem are the same of the number of zones. Thus, one vehicle works with a unique zone. This is useful for lower numbers of vertices to cover in the problem. However, it could not be as good enough for higher numbers of vertices, since the number of zone will increase and the angle for each zone decrease, this led the problem to have more stretched zones and it could affect other parameters in more advanced problems related to the VRP.

## Conclusions and future work

A new strategy was proposed to attack the problem “Vehicle Routing Problem” using Hamilton circuits. This is based on minimizing the distance traveled by the vehicle given certain vertices.

Considering the complexity of the problem, we decided to use a meta-heuristic method called Simulated Annealing because it allows solving combinational optimization problems, such as VRP and traveling salesman problem.

One of the limitations of VRP is the increase in the number of inputs (total number of vertices). It is because the number of possible

solutions increases exponentially. This is why during the implementation of the algorithm to solve this problem. There were established three main relations that allowed Simulated Annealing to find good approximations for each zone proposed in the base problem. However, these relations are not unique. Some several mechanisms and relationships can be considered for future work.

### *Future Work*

There are a lot of searchers that tackle the base problem. Nevertheless, they have gone further, introducing other restrictions such as the capacity of the vehicle that must be optimized too, and the last two vertices of each hamiltonian circuit for each zone must be the same. Despite having new restrictions it is possible to implement a solution based on graph theory and Simulated Annealing as was made in the base problem. Some topics to take in considerations for future work are:

- Compare the implementation of the SA to solve the base problem with other meta-heuristic methods found in recognized researches.
- Find new relations and mechanism to implement in the SA to guarantee the best solution. For example, try to change the zone mechanism to do not divide the total area to work over degrees. One mechanism could be to start at two consecutive or closed vertices to get a better optimized zone and take advantage of the route to do not waste more fuel.
- CVRPFc could be solved in the same way of the VRP. For the implementation, the last two elements must be part of the final route of each zone. The most important parameter will be the capacity of the vehicles. In this case, there are some considerations to take into account, one of them could be to considering that each vertices visited, subtract the same capacity to the vehicle total capacity, so then the problem will be reduced to find the hamiltonian circuit that minimize the total weight which fit the

capacity of the vehicle in each zone and to minimize the total distance traveled at the same time.

- Other consideration could be to pick different products with varied capacity for every vertex and once the vehicle is full, return to the last two vertices. In this case, the numbers of zone will vary according of the vertices product capacity.

## References

- [1] Hopkins, B., Wilson,R.-J.: The truth about Königsberg The College Mathematics Journal, **35**(3), 198--207 (2004)
- [2] Rojas-Cuevas, I.-D., Caballero-Morales,S.-O., Martinez-Flores,J.-L., Mendoza-Vazquez,J.-R.: Capacitated vehicle routing problem model for carriers, Journal of Transport and Supply Chain Management, **12**(1), 1–9 (2018)
- [3] Rosen, K.-H.: Discrete mathematics and its applications Mathematical and computational Applications. 7th edn. McGraw-Hill, New York (2007)
- [4] Bayram, H., Sahi, R.: A new simulated annealing approach for travelling salesman problem. Mathematical and computational Applications, **18**(3), 313—322 (2013).
- [5] Kirkpatrick, S., Gelatt, C.-D., Vecchi, M.-P.: Optimization by simulated annealing. Science **220**(4598), 671—680, (1983).
- [6] Chopard, B., Tomassini, M.: An introduction to metaheuristics for optimization. Springer (2018).
- [7] Delahaye, D., Chaimatanan, S., Mongeau, M.: Handbook of metaheuristics. Springer, (2019).
- [8] Liu, R., Jiang, Z., Fung R.-Y, Chen F., Liu, X.: Two-phase heuristic algorithms for full truckloads multi-depot capacitated vehicle routing problem in carrier collaboration. Computers operations research **37**(5), 950-959, (2010)
- [9] Golden B.-L, Wong R.-T.: Capacitated arc routing problems. Net-works. **11**(3), 305–315 (1981)
- [10] Wan, F, Gou, H., Pan, W., Hou, J., Chen, S.: A Mathematical Method for Solving Multi-Depot Vehicle Routing Problem. Available at SSRN 3943419, 2–10, (2021)
- [11] Afsar, H., Afsar, S., Placios, J.: Vehicle routing problem with zone-based pricing **152**, (2021)

# Comparative study of PI and fuzzy control techniques for voltage stability of a three-phase inverter in photovoltaic systems, evaluating the voltage in transitory and stable state

---

André Mejía Páiz

Universidad Politécnica Salesiana, Quito, Ecuador

Newcastle university, UK

amejiap2@est.ups.edu.ec

a.b.asiedu-asante2@newcastle.ac.uk

Wilson Pavón Vallejos

wpavon@ups.edu.ec

<https://orcid.org/0000-0002-9319-8815>

Leony Ortiz Matos

<https://orcid.org/0000-0002-7883-5513>

Ama Owusu-Andanese

## Abstract

This paper deals with a comparative study of two different control methods implemented to improve the voltage quality at the output of a three-phase photovoltaic DC/AC inverter. The study applies a classical PI control and a heuristic control by fuzzy logic, both independently controlling the duty cycle (PWM). For the comparison of the two controllers, three experimental cases have been considered, in each one the environmental conditions of temperature and irradiance of the photovoltaic panel have been modified. As a result of the investigation it was concluded that the fuzzy controller presents better quality, because its steady state error values are below 0.133%, on the other hand, its characteristic values such as peak voltage, over peak time and settling time, are more favorable than the classical PI control. The design of the controllers and the system has been implemented in Matlab-Simulink software, suitable for modeling dynamic systems. Finally, the use of fuzzy control is recommended for its advantages of stability at medium voltage levels, suitable for industrial applications.

## Keywords

Photovoltaic, Voltage control, Fuzzy logic, Stability, micro-grid.

## Introduction

The world is increasingly concerned about the depletion of fossil fuels and the environmental problems caused by the combustion of fossil fuels, the application of renewable energies cooperates in an important way to the sustainability of the planet, for which clearly, efforts should be focused on improving the quality of clean technologies, such as photovoltaic and wind generation. Company-sponsored studies on renewable energy have gained increasing attention because clean energy needs to be prioritized in industries in order to generate large-scale impacts and environmental awareness[1]. Renewable energy sources, specifically solar and wind, have become more common to meet the demand for electricity consumption. According to the International Energy Agency, solar energy will have provided 11% of global electricity generation by 2050 [2]. Some statistics show that solar electric power consumption has increased by 20% to 25% annually over the last 20 years [3].

Emission-free PV systems are an important part of the growth and development of a sustainable smart grid. Consequently, distributed PV systems have increased considerably around the world, making the integration of these types of energy sources a challenge for electric utilities [4]. With the rapid growth of society and the economy, the need for electrical power sources is increasing every day to meet production and living needs [5].

Traditional electric power, introduces environmental pollution with carbon emissions. For which it is expected that future low voltage grids can cooperate with high levels of energy such as small-scale solar photovoltaic as shown in Figure 1. Grid-tied photovoltaic systems are often installed to improve the performance of the power grid as it allows the reduction of power losses and improves the grid profile [2]. On the other hand, the high insertion of photovoltaic energy can cause several problems such as: voltage fluctuations, harmonics, radiation, noise, among others. It is for this reason that the electrical energy delivered to the grid must be of the best possible quality [6].

Solar energy has gained importance over time with its advantages such as being an inexhaustible energy source, accessible anywhere in the world, an operation that involves little maintenance. This type of energy is converted into direct current (DC) electricity by photovoltaic (PV) modules. Knowing that most of the loads are alternating current (AC), it is necessary to use an inverter which converts the DC voltage of a PV system into AC suitable for a public grid connection or simply for a load [7]. An overview of the different types of PV inverters is available in [4] and [5]. Due to the privileged efficiency and small size, PV inverters without isolation transformers become more sought after in grid-connected solar systems [6]. Moreover, they commonly do not decrease the direct current injection, which implies the saturation of the distribution transformers in the power grid resulting in poor power quality, losses and overheating in the electrical system. For such drawbacks, standards and regulations have been formulated to restrict the injection of DC current from the inverter to the power system[8].

## Park Transforms

Three-phase alternating current signals are transformed into two direct current elements, defined by Park's transforms (dq0 transform), making it simpler and more accurate to regulate two direct current elements as opposed to dealing with alternating current components [9].

$$T_{\theta} = \frac{2}{3} \begin{bmatrix} \cos(\theta) & \cos\left(\theta - \frac{2\pi}{3}\right) & \cos\left(\theta + \frac{2\pi}{3}\right) \\ -\sin(\theta) & -\sin\left(\theta - \frac{2\pi}{3}\right) & -\sin\left(\theta + \frac{2\pi}{3}\right) \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad (1)$$



The inverse of equation (1) is:

$$T_{\theta}^{-1} = \frac{2}{3} \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 1 \\ \cos\left(\theta - \frac{2\pi}{3}\right) & -\sin\left(\theta - \frac{2\pi}{3}\right) & 1 \\ \cos\left(\theta + \frac{2\pi}{3}\right) & -\sin\left(\theta + \frac{2\pi}{3}\right) & 1 \end{bmatrix} \quad (2)$$

For which the angle  $\theta$  is a reference angle in the reference frame.

$\theta = \omega t$  where  $\omega$  is the rotational velocity in frame dq0, t is the time since the initial alignment.

The dq0 transformation allows the mapping of three-phase signals from an abc reference frame to new quantities in a rotating dq0 reference frame. Denoted as:

$$X_{abc} = [X_a, X_b, X_c]^T \quad (3)$$

$$X_{abc} = [X_d, X_q, X_0]^T \quad (4)$$

Where X represents the quantity to be transformed either voltage, current or flux. The transformation dq0 with respect to the reference frame with angle  $\theta$  can be defined as  $X_{dq0} = T_{\theta} X_{abc}$  where the indices d, q and 0 represents direct, quadrature and zero respectively [10].

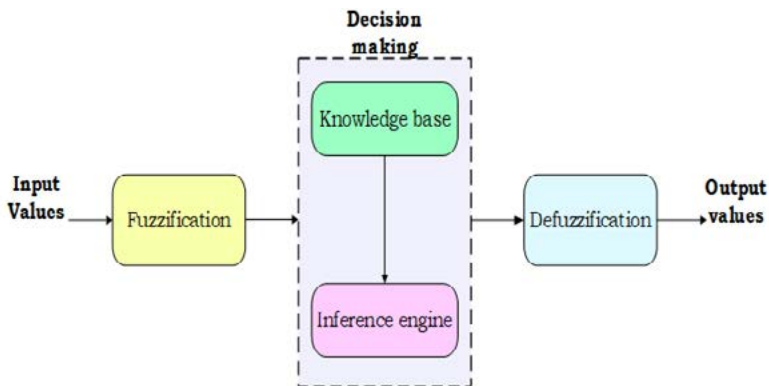
## Fuzzy logic

Fuzzy logic control is one of the most effective methodologies of fuzzy set theory. Fuzzy logic can model human behavior in terms of intelligence and meaning in intelligent machines. By applying program-

mer-defined rules in the realization of system processes, fuzzy control is considered flexible whereby rules can be easily added and modified to improve the performance of the process or system [11].

This fuzzy control methodology does not need a complex control framework, but uses human compression of the system in the form of normal linguistic variables, e.g. large, short, high, low, etc. and characterized by simple logical forms (AND, OR, NOT) using the fuzzy sets [12].

**Figure 1**  
**Structure of a fuzzy logic process**



The main component elements as shown in Figure 1 in FLC (fuzzy logic control) are the fuzzy unit at the input terminal, the knowledge, inference engine, and the defuzzifier at the output terminal.

## **Controllers PID**

PI and PID controllers are generally employed in the control of power electronic converters and are also used in the control of inverters that interact with the power grid. The constants of PI and PID controllers are usually determined using different methodologies and the mathematical model of the system, with methods such as Ziegler-Nichols.

PID controllers are in the time domain and are implemented based on the following equation [13].

$$F(t) = K_p e(t) + K_i \int e(t) dt + K_D \frac{de(t)}{dt} \quad (5)$$

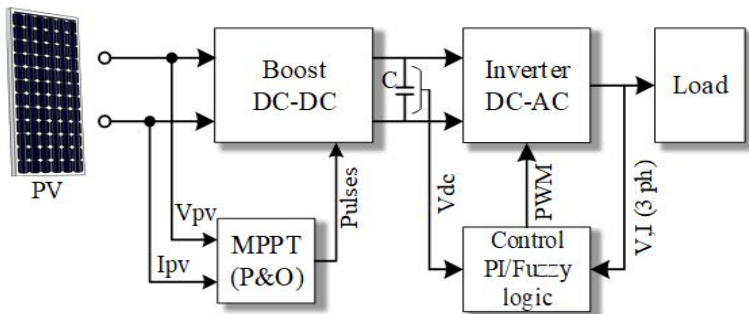
Where  $K_p$  is the proportional gain,  $K_i$  is the integral gain and  $K_D$  is the derivative gain.

The performance of controllers can be affected by various factors such as noise. Grid-interactive inverters operate over a wide operating range and their operating parameters are affected by environmental conditions such as solar radiation and temperature [2].

## System Description

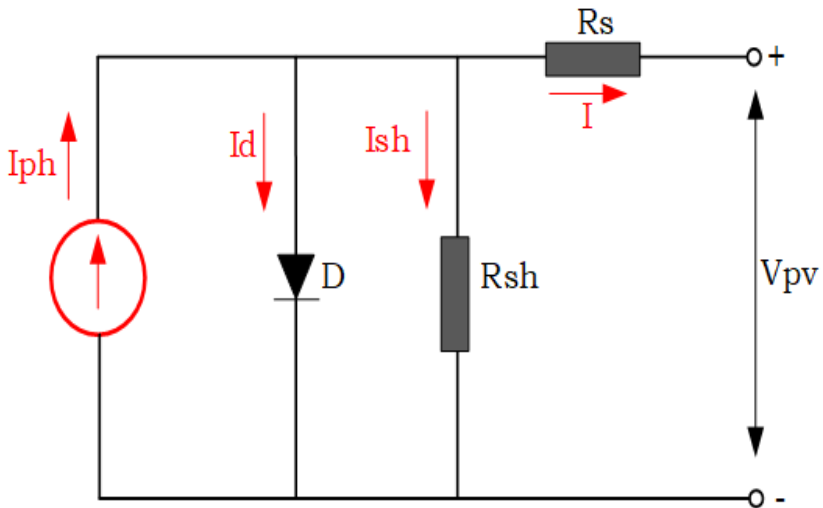
In order to carry out the comparative study of this research, the respective control strategies will be developed to improve the voltage quality at the inverter output. Therefore, it is important to mention preliminary points, such as the structure of the photovoltaic system with its DC converter and the DC/AC inverter, as detailed in Figure 2.

**Figure 2**  
Photovoltaic system structure



Photovoltaic panels by means of irradiance and temperature deliver voltage and current to the DC-DC converter. Therefore, it is necessary to know the energy delivered by the solar system. Next, the modeling of the panel [14], which is based on the equivalent circuit of a single diode as shown in Figure 2, is presented.

**Figure 3**  
Equivalent circuit of a PV panel



The equations representing the equivalent circle model of a photovoltaic panel are:

$$I = I_{ph} - I_d - I_{sh} \quad (6)$$

$$I = I_{ph} - I_o \left( e^{\left( \frac{V + R_s I}{nV_t} \right)} - 1 \right) - \left( \frac{V + R_s I}{R_{sh}} \right) \quad (7)$$

$$I_{PH} = \frac{G}{G_{ref}} I_{sc} + k_i (T_c - T_{ref}) \quad (8)$$

$$I_o = \frac{I_{sc} e^{\left(\frac{E_{go}}{V_{to}} - \frac{E_g}{V_t}\right)}}{\left(e^{\left(\frac{V_{oc}}{N_s V_{to}}\right)} - 1\right)} \left(\frac{T_c}{T_{ref}}\right)^3 \quad (9)$$

$$E_g = E_{go} - \frac{\alpha_{gap} T_c^2}{\beta_{gap} + T_c} \quad (10)$$

$$I_{mg} = I_{pm} \left( \frac{I_{mm}}{1000} G_i + \left( \frac{dI_{scm}}{dT} \right) (T_c - T_{ref}) \right) \quad (11)$$

$$V_{mg} = N_{sm} \left( N_s V_T \ln \left( 1 + \frac{I_{scm} - I_{mm}}{I_{scm}} \left( e^{\frac{V_{ocm}}{N_s V_T}} - 1 \right) \right) - I_{mm} R_{sm} \right) \quad (12)$$

$$I_{mg} = N_{pm} \left[ I_{mpo} (C_o E_e + C_1 E_e^2) (1 + \alpha_{Imp} (T_c - T_{ref})) \right] \quad (13)$$

$$V_{mg} = N_{sm} \left[ \begin{aligned} &V_{mpo} + C_2 N_s \delta(T_c) + C_3 N_s (\delta(T_c) \ln(E_e))^2 \\ &+ \beta_{Imp} E_e (T_c - T_{ref}) \end{aligned} \right] \quad (14)$$

Where:

**Table 1**  
**Parameter nomenclature**

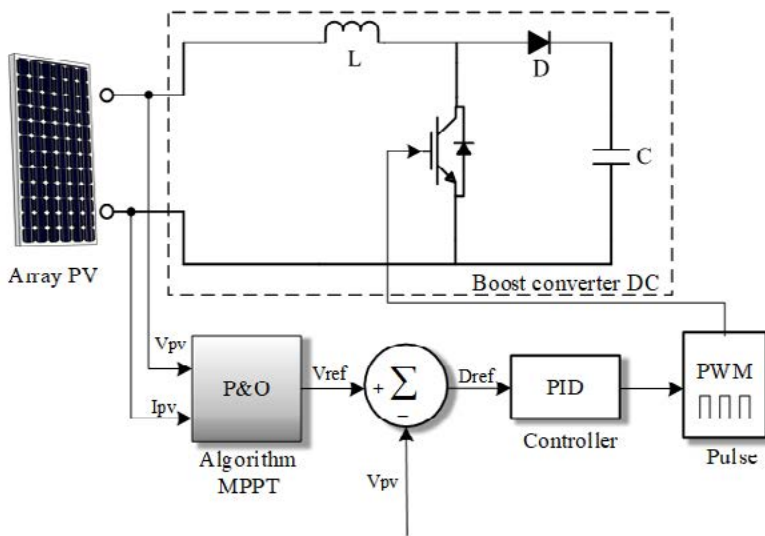
Parameters	Description
$I_o$	Reverse saturation current
$n$	Diode factor
$R_s$	Series resistance
$R_{sh}$	Shunt resistance
$I_{ph}$	Photogenerated current
$G$	Arbitrary irradiance value
$T_c$	Cell temperature
$V_t$	Thermal voltage
$G_{ref}$	Irradiance in short-circuit current
$T_{ref}$	Temperature at short-circuit current
$k_i$	Temperature coefficient
$STC$	Short-circuit current
$I_{sc}$	Short-circuit current
$V_{oc}$	Open circuit voltage

The panel current is given by equation (6) obtained from the summation of currents at the node as illustrated in Figure 5. The I-V characteristics of the solar cells is given by an implicit nonlinear equation, see equation (7). In (8) the photogenerated current is defined, which can be evaluated for any arbitrary value of irradiance, leaving the short-circuit current of the PV cell as the unknown. Thus resorting to (9) to obtain the saturation current of the circuit. Additionally, it is important to obtain the semiconductor energy values at different temperatures of the cell defined in (10). To find the maximum panel power operating points (MPP), it involves obtaining the open circuit voltage and current values of the solar

panel array which is represented in (11) and (12). The MPP voltage and current of the cells is finally defined in (14) and (15).

The controller must follow the MPP, for that is the I-V and P-V characteristic of the array. Fig. 3 shows the approach for MPP tracking (MPPT), where the perturbation and observation (P&O) algorithm receives the current and voltage of the PV array. The P&O block send the reference to the PID controller so that it can send the control signal to the PWM. It should be noted that, if the pulses increase, the array voltage can reduce to reach the MPP [15].

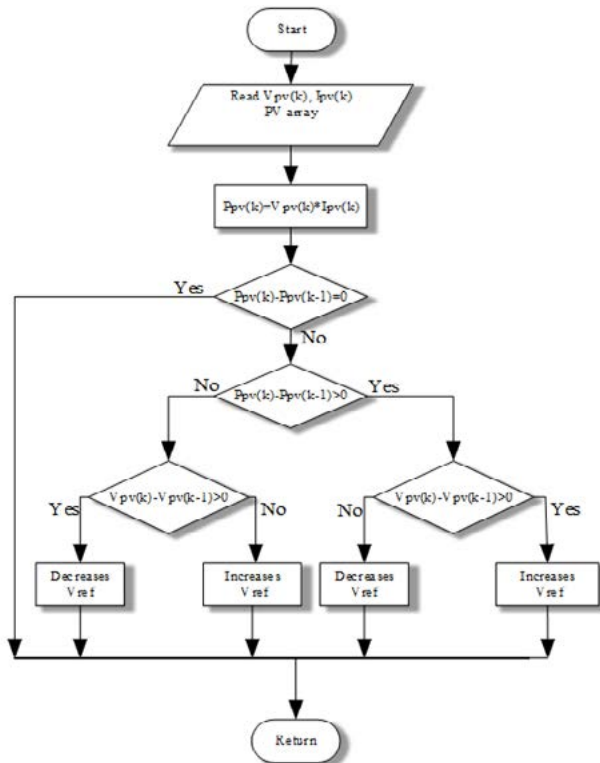
**Figure 4**  
**MPPT for DC/DC Boost converter**



The P&O algorithm has an iterative approach, detailed in Fig. 4, which receives the current and voltage of the PV array. The output power is calculated as the product of these two parameters. The P&O perturb the system by a small increment, and the output PV voltage is

compared to its previous cycle measurement. Then, if the perturbation leads to an output voltage increase, the subsequent perturbation is in the same direction. In this way, the MPPT incessantly searches for the location of MPP. Once the PV MPP is guaranteed, the DC-DC converter delivers DC voltage to the three-phase DC/AC inverter, as shown in Figure 2. For this next stage, the PI control strategy and fuzzy logic will be designed, which generates inverter duty cycle and pulse-width modulation (PWM) for transistors switching.

**Figure 5**  
**P&O algorithm**

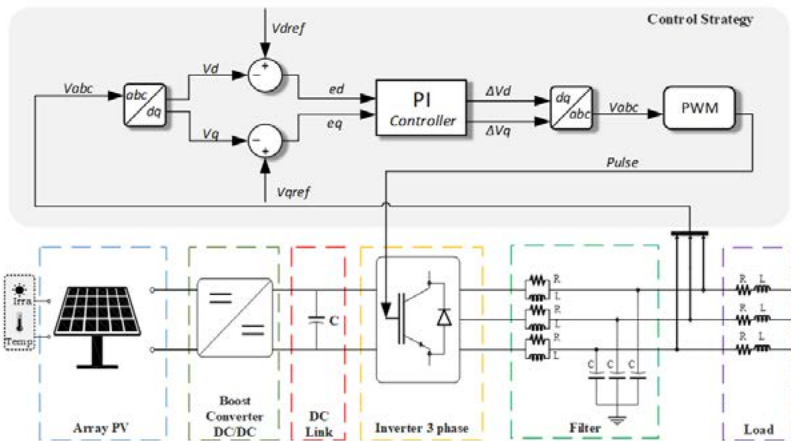




Traditional PI controller is characterized by slow response, large steady state error, as a result, the output voltage is not adequate. The PI controller is expressed by Eq. 11 in the continuous time domain. Figure 5 depicts the control strategy for PV inverter, where  $dq$  variables are converted into  $abc$  coordination frame, which are the PWM signals and the pulses for transistors switching. The switching of the IGBTs converts the PV voltage to alternating current output at the inverter output as a pure sine wave.

$$u(t) = K_p e(t) + K_i \int_{\tau} e(\tau) d\tau \tag{15}$$

**Figure 6**  
**PI control strategy for PV inverter**



The fuzzy logic control is one of the most effective control techniques, because fuzzy logic algorithm is suitable for the converter with complex nonlinear models and parameter variation, for example, RCL filter and inverters in electrical networks.

The voltage control inputs are  $Vd_{ref}, Vq_{ref}, Vd, Vq$ ; while the system outputs are  $\Delta e_d, \Delta e_q$ . It is the same both implemented controllers. For voltage control, the fuzzy logic controller has three parameters. The first input is the voltage error, the second is the error change rate, and the last is the reference voltage output. The control variables are calculated for each sampling, having a real time control.

$v_{d_{ref}}(k), v_{q_{ref}}(k)$ , and  $v_d(k), v_q(k)$  are the reference and the voltage measured at instant  $k$ . In the fusification stage, the variables  $e(k)$  and  $\Delta e(k)$  become fuzzy variables, which can be identified by the membership functions. Eqs. 17, 18 represents the error in direct shaft, while the Eqs. 19 and 20 demonstrate the quadrature shaft.

$$e(k) = v_{d_{ref}}(k) - V_d(k) \quad (16)$$

$$\Delta e = \frac{e(k) - e(k-1)}{T_s} \quad (17)$$

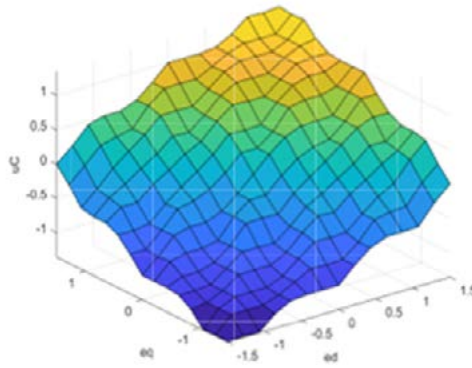
$$e(k) = v_{q_{ref}}(k) - V_q(k) \quad (18)$$

$$\Delta e = \frac{e(k) - e(k-1)}{T_s} \quad (19)$$

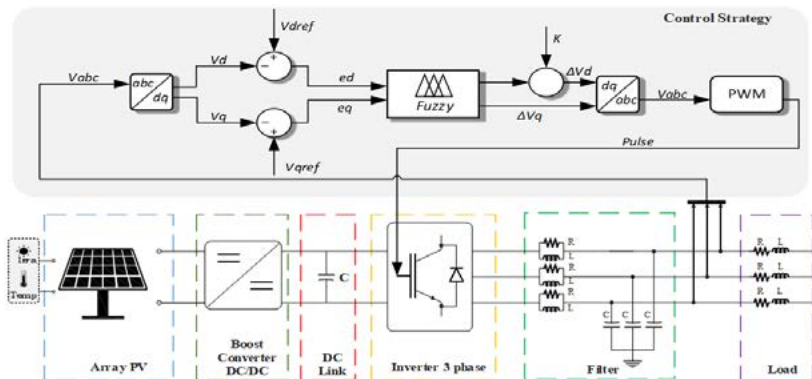
For better visualization of the behavior of the merging rules, they are presented in a three-dimensional plane, Fig. 7, which allows a better controller operation understanding. If the quadrature and the direct axis error decrease, the control signal should be insignificant. On the other hand, if the quadrature and the direct axis error is high, then the control signal should be a high value. In other words, if the errors in both quadrature and direct axis increase, the control signal should increase considerably. For Figure 8, we discuss the fuzzy control design, which is applied to the three-phase inverter. By transforming the three-phase voltages from the synchronous reference frame (abc) to the static reference frame (dq). Consequently, the two-phase coordinate system is de-

veloped from the three-phase coordinate system with the fundamental angular frequency. In most cases, three-phase variables such as current add up to zero. Therefore, the zero sequence components of the system will be omitted. Once the inverter transistors have been switched, the desired sine wave is obtained at the inverter output.

**Figure 7**  
3D fuzzy rule surface for PWM control



**Figure 8**  
Fuzzy logic control strategy for PV inverter



## System configuration

A referenced PV system is implemented to test the proposed system. Also, environmental irradiance and temperature conditions change to verify the system robustness. The PV system is the Soltech 1STH-215-P module with 40 arrays in parallel and 10 in series. Table 2 details the module specifications, and the PV details are described in Table 3. The articles [16] are the referenced case study, and for practical purposes the working frequency is 60 Hz.

**Table 2**  
**PV module specifications**

Parameter	Quantity	Unit
Maximum power	213,15	[W]
Open circuit voltage	36,3	[V]
Maximum voltage point	29	[V]
Cells per module	60	-
Short circuit current	7,84	[A]
Maximum current point	7,35	[A]

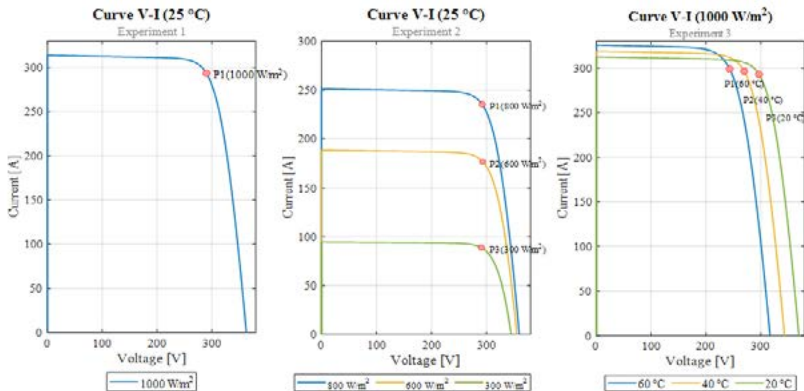
**Table 3**  
**Isolated system parameters**

Parameter	Quantity	Unit
Array PV	1	[Kw]
DC Link (Capacitor)	470	[ $\mu$ F]
DC voltage supply	325	[V]
f (network frequency)	60	[Hz]
Switching frequency	6	[kHz]

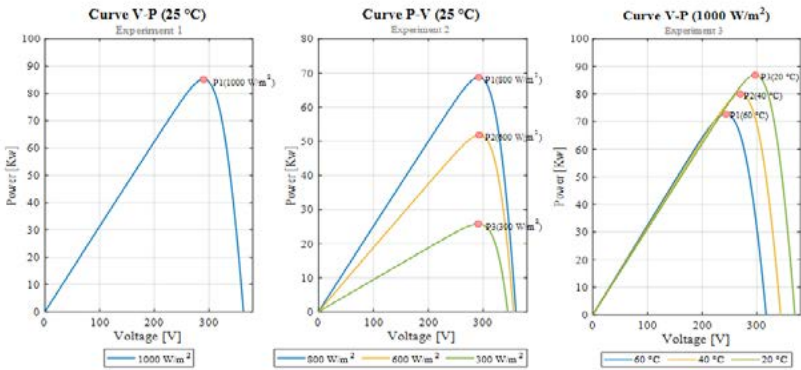
RLC (Filter per phase)	1	[Omhs]
	0.4e-2	[H]
	5,50E-03	[F]
Ideal Inverter Peak Voltage	440	[V]
Load inductive power (Non-linear)	110	[W]

The environmental conditions variation was performed in three different experiments. In the first case, irradiance and temperature are constant, as shown in Fig. 9. The second case, where the current and the irradiance decrease. And the last case, the voltage increases slightly while the irradiance decreases. It is seen that if the temperature decreases; the current would decrease and voltage would increase. Figure 9 shows that if the irradiance decreases; the power would decrease in the same way. However, if the temperature of the array decreases; the power would increase significantly. Table 4 shows the MPP for the three experiments.

**Figure 9**  
V-I curves for each experiment



**Figure 10**  
**V-P curves for each experiment**



**Table 4**  
**Tested experiments MPP**

Experiment	Irradiance [W/m <sup>2</sup> ]	Temperature [°C]	Maximum power point [Kw]
1	1000	25	85.17
2	800	25	68.72
	600		51.8
	300		25.76
3	1000	60	72.7
		40	79.3
		20	86.89

## Results and Analysis

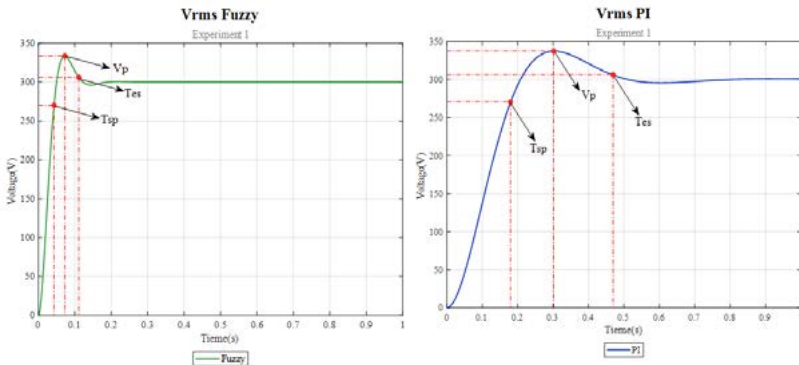
For the analysis of results, it is important to know the performance of a control system that results in a transient response, for which, it is common to take into account the following characteristic terms:

**Table 5**  
**Characteristic control variables**

Term	Description	Unit
Peak voltage	$V_p$	[V]
Over peak time	$T_{sp}$	[s]
Settling time	$T_{es}$	[s]
Steady state error	$e_{es}$	%

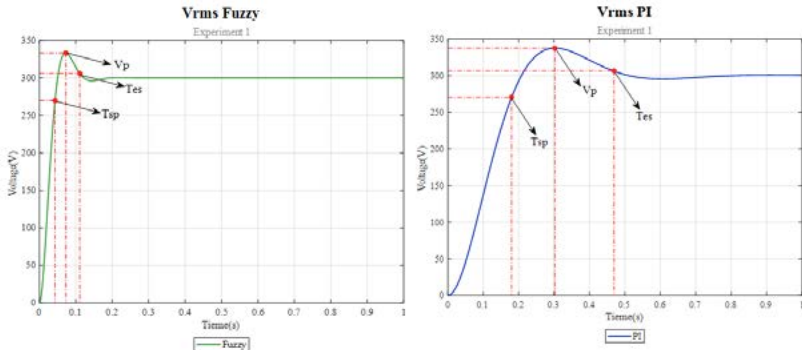
For the evaluation of experiment 1, irradiance and temperature are constant as shown in Table 5. Figure 11 shows the result of the PI vs. diffuse control. The fuzzy logic control stabilizes the voltage at 300 [V] for a time of 0.13 s compared to PI control which takes a longer time 0.478 s.

**Figure 11**  
**Experiment 1, RMS output voltage using**  
**(1) fuzzy and (2) PI controller**



In the second experiment, for the current section, the controllers will be affected by the irradiance changes, while the temperature will be kept constant as described in Table 5. Compared to the previous experiment, the fuzzy control has an unquestionably better settling time of 0.28 s, compared to the PI control which takes 0.46 seconds.

**Figure 12**  
**Experiment 2, RMS output voltage using**  
**(1) fuzzy and (2) PI controller**



For the evaluation of the third and last experimental case, the controllers will be evaluated with temperature changes and constant irradiance as detailed in Table 4. The settling time of the fuzzy controller is 0.181 seconds while the PI is higher at 0.42 seconds.

**Figure 13**  
**Experiment 3, RMS output voltage using**  
**(1) fuzzy and (2) PI controller**

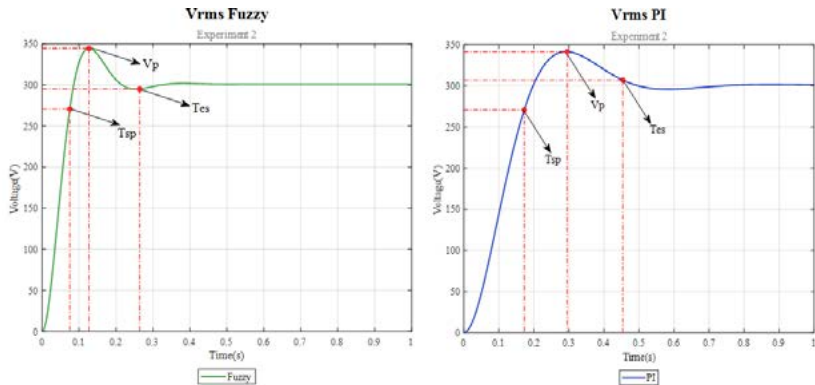




Table 6 shows the comparative results of the controllers, the peak voltage of the Fuzzy control and the PI control are similar so there is little difference, as for the time of over peak the fuzzy control has a better performance compared to the PI control. The settling time of the fuzzy control is less than the PI control because it has a quicker reaction. The percentage of error in steady state of the fuzzy control details better values with 0.096%.

**Table 6**  
**Results of the characteristic variables of the controllers**

Controller	Variable	Experiment	Quantity	Unit
PI	$V_p$	1	337,47	[V]
		2	341,56	
		3	332,06	
	$T_{sp}$	1	0,18	[seg]
		2	0,173	
		3	0,166	
	$T_{es}$	1	0,47	[seg]
		2	0,454	
		3	0,426	
	$e_{es}$	1	0,4667	%
		2	0,812	
		3	0,633	
Fuzzy	$V_p$	1	333,509	[V]
		2	344,41	
		3	331,112	
	$T_{sp}$	1	0,043	[seg]
		2	0,075	
		3	0,0694	
	$T_{es}$	1	0,111	[seg]
		2	0,2641	
		3	0,1774	
	$e_{es}$	1	0,133	%
		2	0,096	
		3	0,0667	

## Conclusions

The fuzzy control strategy successfully fulfills stabilizing the inverter output voltage, in the three experimental cases in its transient and steady state. The system has a quality control with excellent percentages of error in steady stage. For fuzzy control, the error is 0.13%, 0.096% and 0.0667% for each experiment. In comparison, the PI control presents error of 0.4667%, 0.812%, 0.633%. In conclusion, the fuzzy logic controller has a better performance.

The overshoot time for the PI controller is 0.18 seconds. While, the fuzzy controller is faster than PI, since it shows a time below 0.075 seconds as shown in tables 9-14. On the other hand, the PI control peak voltages are 337.47, 341.56 and 332.065 volts. Which are quite similar compared to the fuzzy control, since it describes values of 333.51, 344.41 and 331.113 volts. However, the fuzzy controller has a better robustness over time.

Before the variations of irradiance and temperature, the controllers comply successfully stabilizing the voltage at the desired 300 volts. However, the control strategy that presents better quality at the output inverter is the fuzzy approach. It is known that three-phase inverters are highly nonlinear systems, and the nonlinear fuzzy control responds with high efficiency and quality. In conclusion, fuzzy logic controllers are recommended for medium voltage levels in industrial applications, domestic and also acceptable microgrids integration

## References

- [1] H. M. Foroushani, M. Mohammadrezaei, S. M. Madani, and A. Sayyar, "A single-stage three-phase grid-connected photovoltaic system with fast grid voltage regulation," in *30th Power System Conference, PSC 2015*, Jan. 2017, pp. 305–308, doi: 10.1109/IPSC.2015.7827764.
- [2] I. Sefa, N. Altin, S. Ozdemir, and O. Kaplan, "Fuzzy PI controlled inverter for grid interactive renewable energy systems," *IET Renew. Power Gener.*, vol. 9, no. 7, pp. 729–738, Sep. 2015, doi: 10.1049/iet-rpg.2014.0404.

- [3] W. D. P. Vallejos, "Standalone photovoltaic system, using a single stage boost DC/AC power inverter controlled by a double loop control," in *2017 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)*, 2017, pp. 1–6, doi: 10.1109/ISGT-LA.2017.8126705.
- [4] S. Chaudhary and S. N. Singh, "Single Phase Grid Interactive Solar Photovoltaic Inverters: A Review," in *2018 National Power Engineering Conference, NPEC 2018*, Sep. 2018, pp. 1–6, doi: 10.1109/NPEC.2018.8476767.
- [5] D. Sankar and C. A. Babu, "Cascaded H bridge multilevel inverter topologies for PV application: A comparison," in *Proceedings of IEEE International Conference on Circuit, Power and Computing Technologies, ICCPCT 2016*, Aug. 2016, pp. 1–5, doi: 10.1109/ICCPCT.2016.7530140.
- [6] J. M. Carrasco *et al.*, "Power-electronic systems for the grid integration of renewable energy sources: A survey," *IEEE Transactions on Industrial Electronics*, vol. 53, no. 4. pp. 1002–1016, Jun. 2006, doi: 10.1109/TIE.2006.878356.
- [7] M. A. Khan, A. Haque, and V. S. Bharath Kurukuru, "Enhancement of Fault ride through strategy for single-phase grid-connected photovoltaic systems," in *2019 IEEE Industry Applications Society Annual Meeting, IAS 2019*, Sep. 2019, pp. 1–6, doi: 10.1109/IAS.2019.8911895.
- [8] R. Mechouma, B. Azoui, and M. Chaabane, "Three-phase grid connected inverter for photovoltaic systems, a review," in *2012 1st International Conference on Renewable Energies and Vehicular Technology, REVET 2012*, 2012, pp. 37–42, doi: 10.1109/REJET.2012.6195245.
- [9] J. He and X. Zhang, "Comparison of the back-stepping and PID control of the three-phase inverter with fully consideration of implementation cost and performance," *Chinese J. Electr. Eng.*, vol. 4, no. 2, pp. 82–89, 2018, doi: 10.23919/CJEE.2018.8409353.
- [10] D. Baimel, "Implementation of DQ0 control methods in high power electronics devices for renewable energy sources, energy storage and FACTS," *Sustainable Energy, Grids and Networks*, vol. 18. Elsevier Ltd, p. 100218, Jun. 01, 2019, doi: 10.1016/j.segan.2019.100218.
- [11] M. A. Hannan, Z. A. Ghani, M. M. Hoque, P. J. Ker, A. Hussain, and A. Mohamed, "Fuzzy Logic Inverter Controller in Photovoltaic Applications: Issues and Recommendations," *IEEE Access*, vol. 7, pp. 24934–24955, 2019, doi: 10.1109/ACCESS.2019.2899610.
- [12] S. Rakshit and J. Maity, "Three Phase, 10 kVA Dual Conversion type Automatic AC Voltage Regulator-An Approach based on Fuzzy Logic Con-

- trolled ĆUK Converter and PI Controlled Three Phase Inverter,” in *8th IEEE Power India International Conference, PIICON 2018*, Jul. 2018, pp. 1–6, doi: 10.1109/POWERI.2018.8704391.
- [13] D. Pathak, G. Sagar, and P. Gaur, “An Application of Intelligent Non-linear Discrete-PID Controller for MPPT of PV System,” in *Procedia Computer Science*, 2020, vol. 167, pp. 1574–1583, doi: 10.1016/j.procs.2020.03.368.
- [14] M. B. H. Rhouma, A. Gastli, L. Ben Brahim, F. Touati, and M. Benamar, “A simple method for extracting the parameters of the PV cell single-diode model,” *Renew. Energy*, vol. 113, pp. 885–894, Dec. 2017, doi: 10.1016/j.renene.2017.06.064.
- [15] K. B. Adam and H. Miyachi, “Design of PV-DC mini grid using PI-based fuzzy MPPT control considering load demand,” *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, vol. 2017-Decem, pp. 2990–2995, 2017, doi: 10.1109/TENCON.2017.8228374.
- [16] W. J. Praiselin and J. B. Edward, “Voltage Profile Improvement of Solar PV Grid – Connected Inverter with Micro Grid Operation using PI Controller,” *Energy Procedia*, vol. 117, pp. 104–111, 2017, doi: <https://doi.org/10.1016/j.egypro.2017.05.112>.

# FPGA-based Low-Cost Multispectral Camera Prototype for Precision Agriculture Applications

---

Julian Uribe-Rios  
Instituto Tecnológico Metropolitano, Medellín 50034, Colombia  
julianuribe209085@correo.itm.edu.co  
<https://orcid.org/0000-0003-4332-4614>

David Marquez-Viloria  
Instituto Tecnológico Metropolitano, Medellín 50034, Colombia  
davidmarquez@itm.edu.co  
<https://orcid.org/0000-0002-0168-7276>

Luis Castano-Londono  
Instituto Tecnológico Metropolitano, Medellín 50034, Colombia  
luiscastano@itm.edu.co  
<https://orcid.org/0000-0002-6083-3018>

## Abstract

Multispectral systems are used for data acquisition in precision agriculture. These systems need to acquire and process a large amount of data due to the number of spectral bands they handle. If only one processor is used in a multispectral system, the processing time can be high, in many occasions, an attempt is made to solve this problem by using one processor for each spectral band. This method has the disadvantage that the processors must be synchronized to acquire the images simultaneously and increase the size of the design. There is also the alternative of using devices such as GPUs, but this has a higher power consumption. In this work, we propose developing a prototype of a two-band multispectral system that performs the processing using a heterogeneous architecture based on FPGA seeking to achieve hardware acceleration and have a low power consumption to have the possibility of including the prototype in autonomous aerial vehicles. The combination of ARM and FPGA in the heterogeneous architecture achieves 8.47 frames per second (FPS).

## Keywords

Multispectral Camera, FPGA, Precision Agriculture, Distortion Correction, Image Registration, Heterogeneous Architecture.

## Introduction

Multispectral imaging is widely used to identify different properties of nature in different fields such as forensic sciences, environmental studies, art restoration, quality inspection of pharmaceutical production, biochemistry or geology [11]. Multispectral images are some of the data captured in precision agriculture, obtaining several bands of the electromagnetic spectrum that provide information on spectral properties of the soil and plants in their different stages of development such as reflectance, transmittance and absorbance. These properties are closely linked to factors such as photosynthetic activity, amount of water, chlorophyll, among others [3]. The ranges of interest in precision agriculture are mainly in the visible and near-infrared wavelengths of the light spectrum [13].

Multispectral systems are used in remote sensing, often aboard unmanned aerial vehicles (UAVs), to acquire imagery for tasks such as crop classification and mapping, crop status, and detection of disease and nutrient deficiencies. The integration of these systems into UAVs has been made possible by the design of lightweight multispectral cameras, as well as the development and ease of use of UAVs for aerial inspections [14], as they offer longer tracking time compared to satellites, in which positioning must be awaited. In addition, it removes the repetitive and strenuous routines of traditional methods where you must manually walk through the crop taking data with a spectrometer, and provides better spatial resolution which allows for deeper and more accurate data analysis [11].

Different prototypes of multispectral systems for agricultural applications have been presented in the literature. In [11] a multispectral camera with filter wheel is developed, where the authors present a low-cost solution integrated to a UAV to know the vegetation indices in a palm grove. The camera captures images sequentially with a small time delay between the acquisition of each spectral band, making use of a single CMOS image sensor, a rotating wheel that holds the optical band-pass filters, a stepper motor that drives the rotating wheel, an Arduino Nano to control the motor, and an Odroid XU4 as the processing system. In [9] a three-band

multispectral camera system is developed, where two images in the visible spectrum and one NIR image are acquired simultaneously. The system design incorporates three identical monochromatic cameras, three optical filters and an optical system which is composed of a front lens, a cold mirror and a beam splitter. The front lens is responsible for collimating the spectral reflectance of the object, the cold mirror and the beam splitter are responsible for splitting this reflectance to the three cameras. In this system the two visible spectral bands that can be measured range from 400 nm to 700 nm, and the NIR band ranges from 700 nm to 1000 nm. The design is spectrally flexible in spectral band selection by incorporating a switchable filter design. Because of this flexibility the system can be retrofitted for a variety of applications, such as a portable device for food quality inspection and airborne remote sensing using an UAV. A portable 3-band multispectral camera is presented in [10]. The camera is built with a beam-splitter prism array, an XC3S500E FPGA module for system control, and three filters that can be easily changed for the selection of specific spectral bands for an application in the 400 to 1000 nm region. They designed a printed circuit board with firmware programming to achieve concurrent processing, synchronization and independent control of the three CCD sensors to ensure data transfer without significant delay or loss of data. The system can transmit 30 frames per second and an additional electronic design feature includes independent control of each CCD sensor to adjust gain (brightness) and exposure time. They tested the system in the laboratory for defect spots on apples using 740, 820 and 1000 nm filters. In [5] the implementation of a multispectral system incorporated to a UAV to obtain images corresponding to wavelengths in the visible spectrum (blue, green and red) and in the near infrared is presented. In this work they selected the Raspberry PI model B as the processing system obtaining a time of 5 seconds. In [8] the authors developed a modular two-channel multispectral imaging system where a NIR and a SWIR image are acquired in parallel. The system is lightweight and compact for use on a UAV and allows filters to be quickly exchanged in order to adapt the system to a new application with a wide spectral range from the visible to the shortwave infrared spectrum (approx. 400 nm to 1700 nm). The centerpiece of the

developed unit is an Nvidia Jetson TX2. In the work presented in [15] a multispectral camera, which is sensitive to the NIR, red and green bands, is introduced using three KAC-9628 CMOS sensors. In addition, a heterogeneous FPGA-based architecture (NIOS-II provided by Altera) is used to serve as the control unit, the collected information is stored as BMP files on a Compact Flash card. The development of a 6-band multispectral imaging system including the Jetson TK1 embedded system, 6 monochromatic cameras mounted on an UAV with other electronic modules such as GPS and autopilot system for geolocated image acquisition is presented in [12]. However, the image resolution limits the altitude at which the UAV can fly. The authors conducted proof-of-concept tests to validate the proposed system design on various types of crops such as asparagus, grapes, sugar cane and sweet potato. For these tests they used the X8-Skywalker UAV, which integrates the multispectral imaging system composed of an integrated PCduino v2 system connected to two 1.2 MP resolution monochromatic cameras with filters for the red and NIR bands. In [4] the design of a multispectral camera system from two low-cost commercial cameras that can be integrated into a small UAV is presented. The authors used two Raspberry Pi boards to drive each camera, which take aerial photographs in RGB and near-infrared. This system was used to calculate the normalized difference vegetation index (NDVI) with data obtained in a corn field. In [1] they present a low-cost, open-source and rapidly deployable solution by building a multispectral array of cameras for vegetation monitoring in the visible and near-infrared spectrum. The proposed multispectral imaging system consists of 2 to 16 independent cameras linked together as an array. Each of the cameras that make up the multispectral array has an Aptina AR0130 sensor, a band-pass filter, a 1.8mm lens mount and a Raspberry Pi development board (Broadcom BCM2835) to perform image acquisition, processing and control.

A multispectral system requires the acquisition, processing and storage of large amounts of data resulting in a high computational demand. For this reason, it is important to properly select the computational system to be used. In a system with several spectral bands, the processing



time is very high if a single processor is used. Some works use one processor for each spectral band [1], but this method increases power consumption and design size, two factors that significantly impact the use of these systems on drones. In addition, all processors must be synchronized to acquire images simultaneously. There is also the alternative of using architectures that execute massively parallel operations such as GPUs and FPGAs, with FPGAs having lower power consumption and more digital input-output interfaces. Implementation using a heterogeneous FPGA-based system can achieve a significant improvement in performance compared to general-purpose processors [6].

This paper presents a system to capture two spectral bands using a heterogeneous architecture based on FPGA, in order to reduce the processing load on the CPU, achieving an acceleration in the system by taking advantage of the parallel processing advantage of FPGA and its low power consumption. The multispectral camera prototype was designed as a low-cost solution affordable for precision agriculture applications in developing countries. This work presents the first stage of development of what is expected to be a future camera with height spectral bands. The implementation accelerates the processing of distortion correction algorithm on a heterogeneous FPGA-based architecture. The results showed an improvement compared to the processing on ARM and an AMD Ryzen 5, and also power measurements are included. The document is organized as follows: Section 2 presents the methodology used to design and implement a multispectral camera prototype. The results of the algorithms and hardware performance measurements are presented in Section 3. Finally, we present the conclusions and future work in Section 4.

## **Methodology**

### *Multispectral camera prototype*

The multispectral system implemented is based on an architecture known as sensor array. The system has image sensors, lenses, and bandpass optical filters to capture each spectral band. The prototype

proposed in this paper is composed of two monochrome cameras of reference ELP-USB130W01MT-B/W and an Ultra96-V2. The Xilinx Zynq UltraScale+ MPSoC ZU3EG A484 is the main chip on the Ultra96-V2. The chip has a heterogeneous architecture based on an ARM known as the processing system (PS) and an FPGA called programmable logic (PL). It has been used in a wide range of applications in recent years, including artificial intelligence, machine learning, IoT connectivity, embedded computing, and robotics. Tables 1 and 2 show the main characteristics of the camera and the Ultra96, respectively.

**Table 1**  
**Monochrome camera specifications**

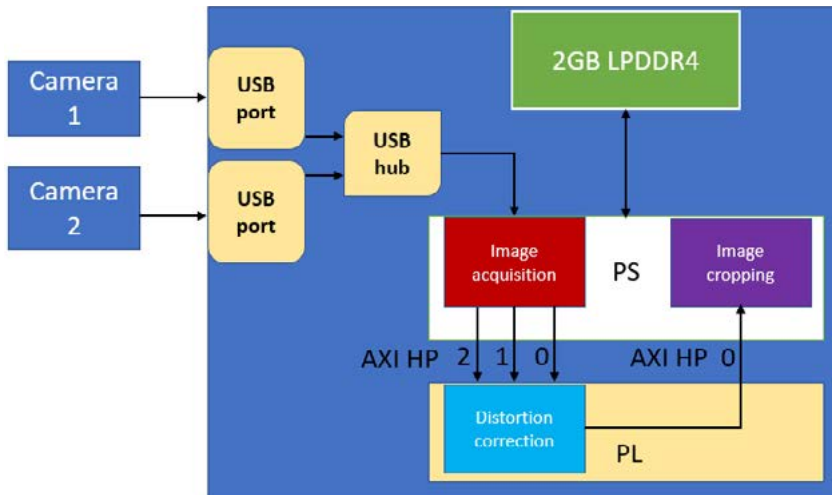
Model	ELP-USB130W01MT-B/W
Sensor	Aptina AR0130
Pixel size	3,75 $\mu\text{m}$
Compression format	MJPEG/YUYV
Resolution	1280x960/1280x720/800x600/640x480/320x240/160x120
Power supply	5V DC

The connection between monochrome cameras and Ultra96 was made through the USB downstream type A ports of the Ultra96. USBs ports are connected internally to the PS of the Zynq chip. In the ultra96, the PYNQ operating system was used, which is an open-source project of Xilinx based on Ubuntu and allows working with the Python language in some of its platforms. OpenCV was used for image capture, acquiring images with a resolution of 720x1280. Fig. 1 shows the block diagram of the architecture used for the multispectral camera prototype. The data flow from camera acquisition to processing in the PS and PL is presented in the figure. The PS handles the acquisition and image cropping process, while the PL performs the distortion correction algorithm. The system has 2GB RAM managed by the PS for temporary data storage.

**Table 2**  
**Ultra96-V2 specifications**

Ultra96 V2
Xilinx Zynq UltraScale+ MPSoC ZU3EG A484
Memory LPDDR4 2GB RAM
16 GB microSD
Wi-Fi 802.11b/g/n
Bluetooth
One USB 3.0 type Micro-B upstream port Two USB 3.0 ports and one USB 2.0 downstream Mini DisplayPort

**Figure 1**  
**Block diagram of the heterogeneous architecture for the multispectral camera system**

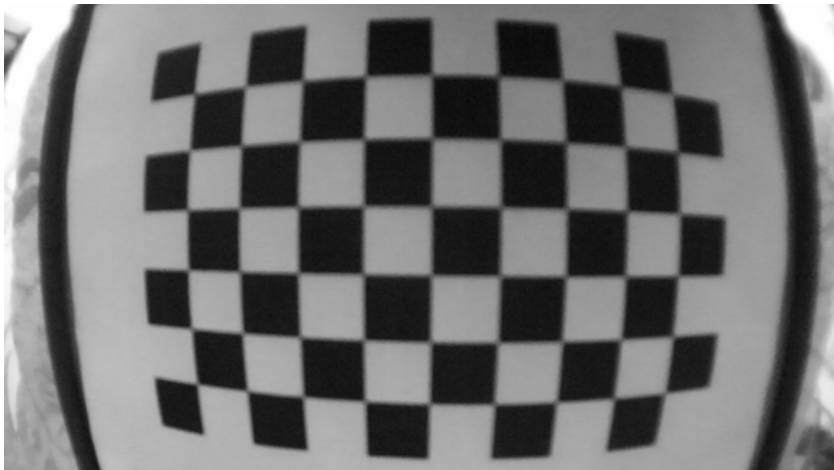


## *Distortion correction implementation on Ultra96-V2*

In the multispectral camera prototype, we have the option of using undistorted cameras that do not require the application of correction algorithms, but these cameras have a reduced field of view (FOV), forcing the image to be captured from a greater height. On the contrary, cameras that allow covering a larger area usually have distortion, and it is necessary to apply a correction algorithm. In this work, the latter option is chosen because it allows capturing images from a lower height.

The acquired images show distortion due to aberrations in the lens of the cameras, see Fig. 2. For this reason, a distortion correction algorithm was applied to the images. The algorithm goes through each pixel in the image, calculating the Euclidean distance from the center. It performs a correction shift radially, finding the appropriate indices to rearrange the pixels in the image to correct the distortion. The algorithm 1 presents the pseudocode that calculates the indices for distortion correction.

**Figure 2**  
**Distortion present in acquired images**



Since the result of the calculation of these indexes will always be the same for given frame size, the results were saved in two text files, one file for the row indexes and the other for the column indexes. The indexes found depend exclusively on the resolution of the images. For this reason, these indexes can be precomputed and stored in text files to avoid calculating them at runtime. In our implementation, two independent text files store the row and column indexes, allowing concurrent reading of both files.

---

**Algorithm 1:** calculation of the distortion correction indices

---

```

Initialization;
row ← 720;
column ← 1280;
halfWidth ← row/2;
halfHeight ← column/2;
correctionRadius ← 0.001566;
for each row do
    for each column do
        newX ← row - halfWidth;
        newY ← column - halfHeight;
        distance ←  $\sqrt{\text{newX}^2 + \text{newY}^2}$ ;
        r ← distance * correctionRadius;
        if r == 0.0 then
            | theta ← 1;
        else
            | theta ← arctan(r)/r;
        end
        correctedX ← round(halfWidth + theta * newX);
        correctedY ← round(halfHeight + theta * newY);
    end
end

```

---

Generating the image without distortion by rearranging the pixels on the ARM takes a lot of processing time, so an intellectual property (IP) block was implemented in the programmable logic using Vivado HLS to perform this procedure.

Zynq chip does not have enough BRAM available to store the 720x1280 images on memory on the chip. Due to this, the processing is performed for sub-images that correspond to one-quarter of the complete image. In this way, a sub-image has a 360x640 size.

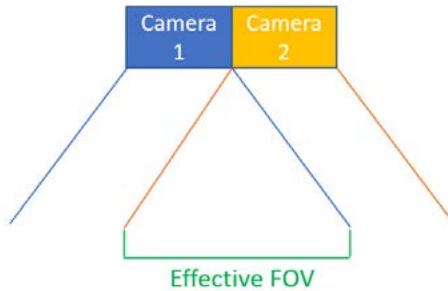
Sub-images and text file values are sent from the PS to PL using the direct memory access (DMA) to transfer the data from RAM to the created IP block via the AXI HP high-speed buses. Three high-speed buses were used, one for the image and the other two for the values of the two text files. After generating the undistorted image in the IP block by rearranging the pixels, the image returns to the RAM using the same bus where this initially arrived in the IP block.

### *Image registration process*

The prototype is a two-band multispectral camera, and each camera has a filter that captures a part of the light spectrum. A possible application in agriculture for this two-band camera is calculating the NDVI using a filter in the near-infrared and another filter in the red band. The two images acquired by the cameras must match pixel by pixel to compose a single multispectral image. However, the optical centers are aligned vertically but not horizontally due to the distance present in both sensors, and this causes a pixel shift. An image registration algorithm is performed in the PS to correct pixel shift when forming a multispectral image using the images acquired with both cameras. Image registration is necessary when the system has a set of sensors because each has a different point of view of the scene, as shown in Fig. 3. The objective is to make a crop in each image so that only the effective FOV that both sensors can see remains.

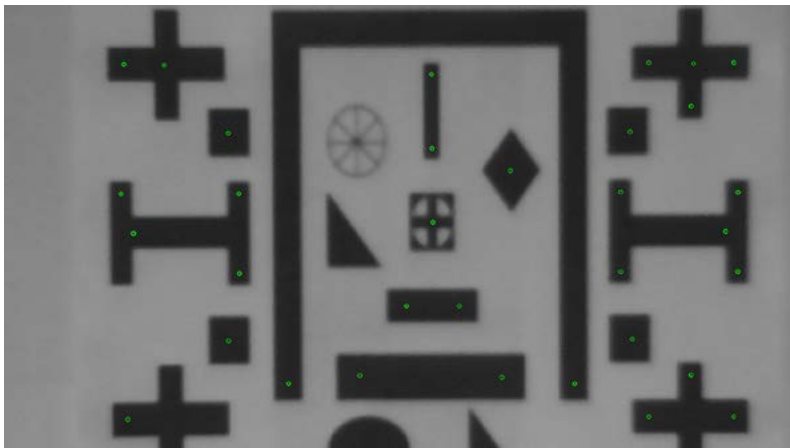
The number of pixels to be cropped in the image is determined by performing feature extraction on the two images acquired by each camera. Scale-invariant feature transform (SIFT) was used for feature extraction.

**Figure 3**  
**Effective FOV in a sensor array**

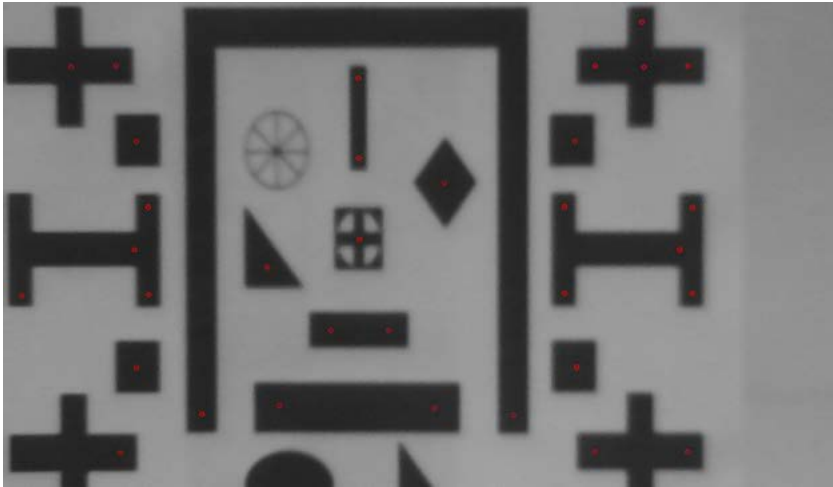


In order to detect corresponding features between the two images, figure patterns were used in the images that facilitate the identification of points of interest. Fig. 4 and Fig. 5 show features extracted from the image captured with camera 1 and camera 2, respectively. Fig. 6 shows the feature correspondences found between both images. OpenCV's BF-Matcher method is used to find the correspondences.

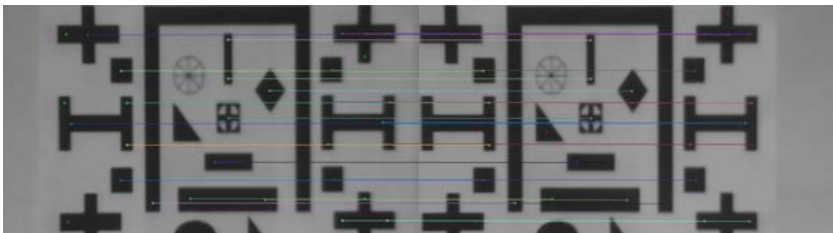
**Figure 4**  
**Characteristics detected in the image taken with camera 1**



**Figure 5**  
**Characteristics detected in the image taken with camera 2**



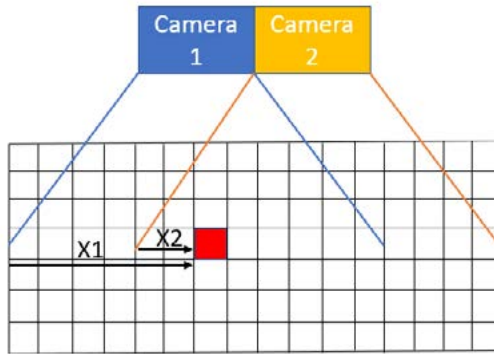
**Figure 6**  
**Corresponding features detected between the two images**



Feature correspondences are used to know the number of pixels to be cropped. For this, the position of a feature in the image acquired with one sensor is obtained. The horizontal displacement in pixels of the corresponding feature in the image acquired with the other sensor is calculated, as shown in Fig. 7. This procedure is performed for each corresponding feature, selecting the modal value of the horizontal offset as the number of pixels to be cropped.



**Figure 7**  
**Number of pixels to crop**



## Results and discussion

Fig. 8 shows the result obtained from the IP block implemented in the PL that performs the distortion correction. The figure shows the chess grid of Fig.2 after passing through the algorithm. The grid is aligned, validating the functionality of the algorithm.

**Figure 8**  
**Image obtained with distortion correction**

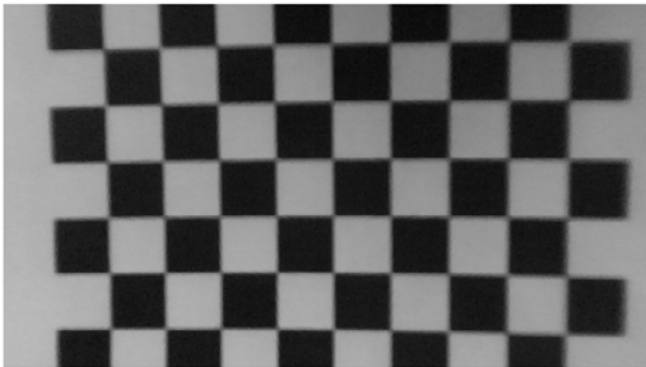


Table 3 shows the resources used by the IP block implemented in the Ultra96, where the highest consumption is in the BRAMs due to the storage of the image pixels that reach the IP.

**Table 3**  
**Resources used on Xilinx Zynq Ultrascale+ MPSoC ZU3EG A484**

	BRAM_18K	DSP48E	FF	LUT
Total	114	0	99	602
Available	432	360	141120	70560
Utilization %	26	0	~0	~0

The execution time of the distortion correction algorithm was measured on quarter images and full images to check the speedup achieved using FPGA+ARM. This was compared with the results of the distortion correction algorithm implemented only on the Ultra96 ARM and an AMD RYZEN 5 processor using the Python language. The execution times on the heterogeneous architecture improve the computation times obtained on the ARM and the AMD Ryzen 5, in 81.4X and 13.5X, respectively. The combination of ARM and FPGA achieve 8.47 frames per second (FPS), as shown in Table 4.

**Table 4**  
**Execution times and FPS of the distortion correction algorithm**

	Ultra96 (ARM)	AMD Ryzen 5	Ultra96 (FPGA+ARM)
Sub-image (360x640)	2,32 s	0,387 s	7,2 ms
Full image (720x1280)	9,37 s	1,57 s	49,2 ms
FPS	0,104 FPS	0,627 FPS	8,47 FPS
Speedup (FPS)	81.44X	13.50X	1X

Fig. 9 shows a 2-band multispectral image when the image registration algorithm is not applied. The displacement that exists between the pixels of both bands can be observed due to the distance between the image sensors. This image has three dimensions: width and height with 1280 pixels and 720 pixels, respectively. The third dimension is the number of spectral bands, two in this case.

**Figure 9**  
**Images captured with both cameras**  
**without performing image registration**

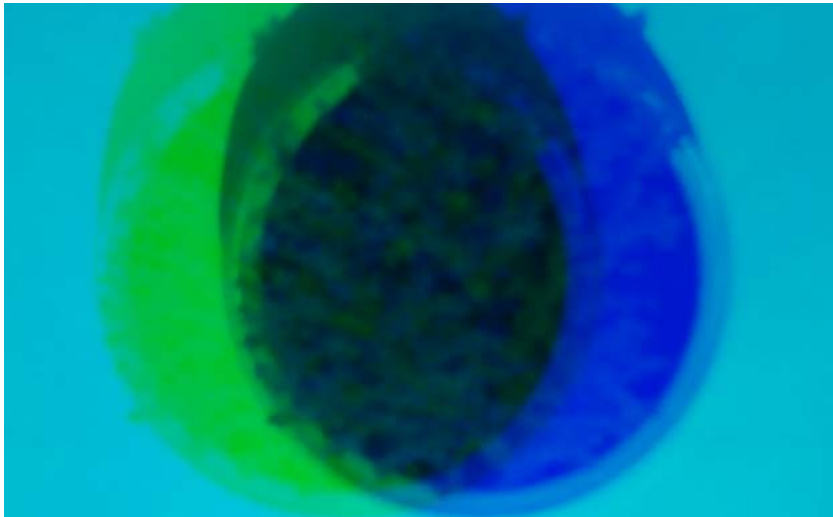


Fig. 10 shows the result of applying for the image registration algorithm aligning that the pixels of the two spectral bands. The pixel value was found to be cropped 164 pixels at the height of approximately 80cm. The first 164 columns of pixels were cropped to the image obtained with the left sensor since they are not within the FOV of the sensor on the right side. Similarly, the last 164 columns of pixels were cropped to the image obtained with the sensor located on the right, resulting in an image of 720x1116x2 containing the FOV of both sensors.

**Figure 10**  
**Images captured with both cameras performing image registration**

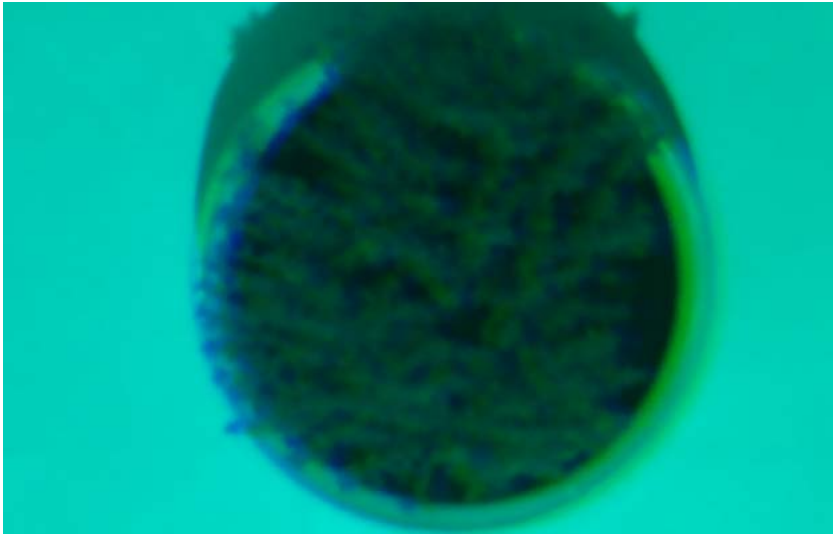
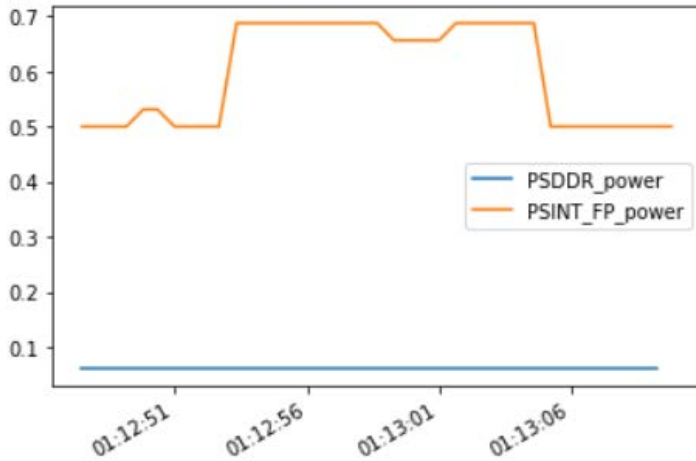


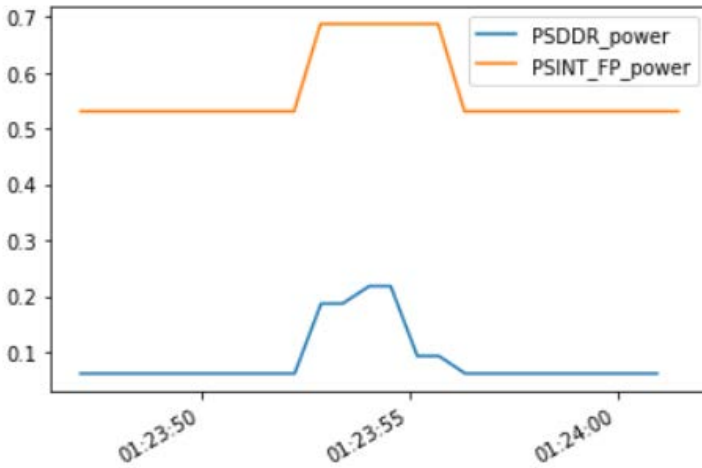
Fig.11 shows the power consumption measured in the Ultra-V2 running the distortion correction and image registration algorithms in the ARM. When the execution of the algorithms starts, the power in the PSINT-FP signal increases up to 0.7 watts. At the end of the execution, it returns to the initial value. The PSDDR power signal remains constant because it measures the power when the DMA is used, which is not the case in this implementation.

The measured power consumptions running distortion correction algorithm in the PL (FPGA) and image registration algorithm in the PS (ARM) are shown in Fig. 12. The same power variation is observed in the PSINT-FP signal compared to the value shown in Fig. 11. During the algorithm's execution, the power rises to almost 0.7 watts. In this implementation, the PSDDR signal increases above 0.25 watts because the PL uses the DMA to read the data from PS.

**Figure 11**  
Power consumption running the algorithms on the PS



**Figure 12**  
Power consumption running the algorithms in the PS and PL



Most of the similar works do not report the number of frames per second they can process, only [5] mention that the multispectral system has an acquisition time of 5 seconds per image. Regarding image registration in [1], they use GSD (Ground Sample Distance) to find the number of pixels to crop. The approach has the disadvantage that the height of the multispectral camera needs to be known to perform the calculations. While using feature extraction, the height value is not necessary. In [2] they use a second-order polynomial transformation to align the images. However, they leave as an alternative the refinement of the algorithm. In [1] the authors propose a Raspberry for each spectral band; this is a drawback for systems to be included in drones as it increases the power consumption and the size of the design, which is not the case in our system. Comparisons of system characteristics are given in Table 5.

**Table 5**  
**Comparison of the system's characteristics with other works**

System	Number of sensors	Processing system	Image registration	FPS
Proposed	2	Ultra96-V2	Yes	8.31
Ref. [1]	2-16	1 Rapsberry Pi per band	Yes	Not reported
Ref. [4]	2	2 Rapsberry Pi	Not reported	Not reported
Ref. [5]	2	1 Rapsberry Pi	Not reported	0.2
Ref. [8]	2	1 Nvidia Jetson TX2	Not reported	Not reported
Ref. [14]	3	Nios-II	Not reported	Not reported
Ref. [12]	2	PCDuino v2	Not reported	Not reported
Ref. [2]	2	2 Rapsberry Pi	Yes	Not reported
Ref. [7]	2	1 Rapsberry Pi	Not reported	Not reported

The design is proposed as a flexible solution for changing filters to suit applications such as measuring water stress in crops. Low-cost commercial solutions such as MicaSense Kit RedEdge-MX have a price of USD \$5500 at the publication of this paper for a camera with five fixed spectral bands. The solution proposed in this paper achieves a price of USD \$646.35 for two bands, and it is expected that the scaling can retain a price well below current commercial solutions. Table 6 shows the price details of the components. The remaining parts needed in a future full version of the prototype, such as 3D printed packaging and screws, are assumed to be negligible in price compared to the components presented in the table.

**Table 6**  
**List of components for the multispectral camera prototype**

Component	Price \$(USD)
1 Ultra96-V2	249
2 Monochrome cameras	87,36
2 Optical filters	310
Total	646,36

## Conclusions

In this work, we presented the prototype of a multispectral system using a heterogeneous FPGA-based architecture for precision agriculture applications. In the first application, we present a prototype of two spectral bands together with distortion correction and image registration algorithms.

One of the characteristics sought in the prototype is to achieve good acquisition and processing times, low power consumption, and low weight of the system to include it in drones. For this reason, we use a heterogeneous architecture based on FPGA to achieve an acceleration in the algorithms using HLS or RTL design methodologies.

The execution times obtained in the distortion correction comparing the results between ARM and PL showed that an acceleration in hardware could be obtained and meet the computational demand that multispectral systems require.

As future work, we propose implementing an 8-band multispectral system and implementing other types of algorithms applied to precision agriculture exploring acceleration techniques.

## Acknowledgments

This work was financed with resources from the *Patrimonio autónomo fondo nacional de financiamiento para ciencia, tecnología e innovación Francisco José de Caldas* through RC 80740-475-2020, supported by the Ministerio de Ciencias, Tecnología e innovación – Minciencias, Instituto Tecnológico Metropolitano – ITM, Corporación Colombiana de Investigación Agropecuaria – AGROSAVIA, and BLACKSQUARE S.A.S.

## References

- [1] Barreiros, J., Magne, N.: Open source multispectral camera array for vegetation analysis. In: 2015 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON). pp. 879-884. IEEE (2015)
- [2] Bokolonga, E., Hauhana, M., Rollings, N., Aitchison, D., Assaf, M.H., Das, S.R., Biswas, S.N., Groza, V., Petriu, E.M.: A compact multispectral image capture unit for deployment on drones. In: 2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings. pp. 1-5. IEEE (2016)
- [3] Candiago, S., Remondino, F., De Giglio, M., Dubbini, M., Gattelli, M.: Evaluating multispectral images and vegetation indices for precision farming applications from uav images. *Remote sensing* 7(4), 4026-4047 (2015)
- [4] Doering, D., Vizzotto, M., Bredemeier, C., da Costa, C., Henriques, R., Pignaton, E., Pereira, C.: Mde-based development of a multispectral camera for precision agriculture. *IFAC-PapersOnLine* 49(30), 24-29 (2016)



- [5] Galindo, A.K.T., Rivera, A.F.G., López, A.F.J.: Development of a multispectral system for precision agriculture applications using embedded devices. *Sistemas & Telemática* 13(33), 27-44 (2015)
- [6] Gulbudak, O., Santi, E.: Fpga-based model predictive controller for direct matrix converter. *IEEE Transactions on Industrial Electronics* 63(7), 4560-4570 (2016)
- [7] Habibullah, M., Mohebian, M.R., Soolanayakanahally, R., Bahar, A.N., Vail, S., Wahid, K.A., Dinh, A.: Low-cost multispectral sensor array for determining leaf nitrogen status. *Nitrogen* 1(1), 67-80 (2020)
- [8] Jenal, A., Bareth, G., Bolten, A., Kneer, C., Weber, I., Bongartz, J.: Development of a vnir/swir multispectral imaging system for vegetation monitoring with unmanned aerial vehicles. *Sensors* 19(24), 5507 (2019)
- [9] Kise, M., Park, B., Heitschmidt, G.W., Lawrence, K.C., Windham, W.R.: Multispectral imaging system with interchangeable filter design. *Computers and Electronics in Agriculture* 72(2), 61-68 (2010)
- [10] Lee, H., Park, S.H., Noh, S.H., Lim, J., Kim, M.S.: Development of a portable 3ccd camera system for multispectral imaging of biological samples. *Sensors* 14(11), 20262-20273 (2014)
- [11] Morales, A., Guerra, R., Horstrand, P., Diaz, M., Jimenez, A., Melian, J., Lopez, S., Lopez, J.F.: A multispectral camera development: From the prototype assembly until its use in a uav system. *Sensors* 20(21), 6129 (2020)
- [12] Paredes, J.A., González, J., Saito, C., Flores, A.: Multispectral imaging system with uav integration capabilities for crop analysis. In: 2017 First IEEE International Symposium of Geoscience and Remote Sensing (GRSS-CHILE). pp. 1-4. IEEE (2017)
- [13] Rivera, A.F.G., Clavijo, F.V.: Agricultura de precisión y sensores multiespectrales aerotransportados. Recuperado de [http://cici.unillanos.edu.co/media2016/memorias/CICI\\_2016\\_paper\\_106.pdf](http://cici.unillanos.edu.co/media2016/memorias/CICI_2016_paper_106.pdf) (2016)
- [14] Sun, B., Zhao, Z., Xie, D., Yuan, N., Yu, Z., Chen, F., Cao, C., de Dravo, V.W.: Sparse spectral signal reconstruction for one proposed nine-band multispectral imaging system. *Mechanical Systems and Signal Processing* 141, 106627 (2020)
- [15] Ze, Z., Minzan, L., Xin, L.: The development of a hand-held multi-spectral camera based on fpga. In: 2010 World Automation Congress. pp. 469-475. IEEE (2010)

# Outcome Prediction of COVID-19 Patients From Clinical Admission Data Using Machine Learning Models

---

Juan Pablo Cartagena  
Instituto Tecnológico Metropolitano, Colombia  
[juancartagena257067@correo.itm.edu.co](mailto:juancartagena257067@correo.itm.edu.co)

Rubén D. Fonnegra  
Institución Universitaria Pascual Bravo, Colombia  
[ruben.fonnegra@pascualbravo.edu.co](mailto:ruben.fonnegra@pascualbravo.edu.co)  
<https://orcid.org/0000-0001-6589-1981>

Fabián Narváez  
Universidad Politécnica Salesiana, Ecuador  
[fnarvaeze@ups.edu.ec](mailto:fnarvaeze@ups.edu.ec)  
<https://orcid.org/0000-0002-2490-6902>

Gloria M. Díaz  
Instituto Tecnológico Metropolitano, Colombia  
[gloriadiaz@itm.edu.co](mailto:gloriadiaz@itm.edu.co)  
<https://orcid.org/0000-0003-102-9111>

## Abstract

The coronavirus pandemic created a collapse in most health systems among other causes, due to the low availability of intensive care units (ICU). As a consequence, some patients ended in a fatal outcome. Some initial studies showed that the presence of underlying conditions could be predictors of the complexity of the disease. However, this is far from conclusive. This article proposes the use of strategies based on artificial intelligence to predict the outcome of the disease in patients diagnosed with COVID-19 from data and results of para-clinical tests obtained at the time of hospital admission. The main goal for the model is developing the ability to identify patients who are going to present complications (admission to the ICU or death), from those who are discharged without having gone through this unit. For this, four characteristics selection techniques are evaluated, which allow establishing the subset of characteristics that provide the most information to the decision, and three prediction models based on supervised clas-

sification techniques: Random Forest, support vector machines (SVM) and multikernel learning (MKL). The assessment was performed on a subset of data from HM-hospitals' COVID Data Save Lives initiative. The highest sensitivity was reported by the combination of a subset of the characteristics selected with the mutual information method, and an SVM classification model; while the highest specificity was obtained by the MKL model (87 %) with a set of characteristics selected by this same model.

## Keywords

COVID-19 · Outcome Prediction · Clinical Data · Admission · Machine Learning · Feature selection.

## 1 Introduction

COVID-19 is a disease caused by SARS Cov-2, a virus of the coronavirus type, identified in December 2019 in Wuhan, China. Despite not being a highly lethal virus, it is characterized by a high degree of infection, which as of December 1, 2020 registers a total of 65 million infections and one and a half million deaths, in the world [22]. This behavior of the disease led the World Health Organization - WHO to declare it a pandemic on March 11, 2020 [11].

COVID-19 presents complications mainly in the elderly or with a notable medical history. SARS-COV 2 is said to accumulate in a higher proportion in the body, resulting in a longer incubation period compared to SARS, as well as being more contagious. It is disclosed that a big problem with this virus is that its infection period begins before the symptoms, this being a strong characteristic that helps its spread [9]. For this reason, hospitals, medical centers and other entities in charge of promoting health were not prepared for the large number of patients. As a result, there has been a high demand for clinical implements, especially in the intensive care unit (ICU), a place where patients in serious condition are admitted [10]. Among the implements that present the greatest demand are: the multipurpose bed, which is electric, has a scale and commands for different functions, allowing constant monitoring of the patient's condition; the mechanical ventilator (which functions as a respirator), being essential for patients who enter with pulmonary insufficiency, among others [13]. In addition to this equipment, the demand for

specialized medical personnel who have the knowledge to operate the technological elements and the direction of all this medical equipment has increased. In this way, not having this type of resources, be they human or technological, implies a collapse of hospital centers in the care of seriously ill patients, due to the health contingency; in addition, a factor that has a high incidence in the final triggering of the patient.

In addition, this virus has particular characteristics in its life cycle. It has been shown that its incubation period begins at the same moment of contagion and up to the fifth day; Unfortunately, his period of infection begins before the first symptoms are observed, I feel this one of the main reasons for spreading. The period of presentation of symptoms occurs immediately after incubation for 10 more days, this being a decisive period for the carrier of the virus, since this is followed by the period of resolution of symptoms, where the final state is determined. of the patient, that is, complications may appear quickly or there may be a recovery. It is vitally important to understand the behavior of the virus and act immediately so that when the resolution period is reached, the patient has already been treated in the best way so as not to have the need to enter the ICU and as a consequence the collapse of the hospitals. For this reason, determining the real state of a patient to establish the level of risk of complication or even predict its outcome is essential to improve treatment, make better use of available resources and help reduce the collapse of the system of Health.

This article presents the evaluation of different machine learning models for the prediction of the final state of a patient diagnosed with COVID-19. These allow defining whether the carrier will have complications or not, based on their clinical and personal data. For the evaluation, a database is taken from which 880 patients with a final state defined as death or domicile are extracted, in addition to sufficient records for the construction of the model. From there, algorithms for the selection of characteristics were implemented in order to identify the impact of each of them with respect to the final outcome of the patients at the time of

diagnosis. This relevance will be an important feature for the evaluation, subsequent analysis and investigation of the disease. Two feature selection techniques were compared, as well as the relevance of the features of the machine learning models used in the early experiments, in order to create the best models. The results obtained lead to identifying the most significant information with which it is possible to know the state that the patient will have and thus have a more assertive care when being treated in a health center. The characteristics obtained with the greatest influence are also presented, since they are the ones that have the greatest relevance in the results of the characteristics extraction techniques. Further; In terms of performance, the evaluated models are made after subsequent analysis and a sensitivity of 83% and a specificity of 87% are reported, these being the best metrics in the experiments carried out.

## 2 Previous Works

Due to the rapid spread of the disease, there were some problems in the health sector, one of these is the collapse of hospitals [26,2,6]. This was due in some cases to a misdiagnosis at the time of the patients' first admission to a health entity. For this reason, most of the studies initially focused on making an early detection of the disease, since this is a decisive stage for carriers of the virus [23]. These works emphasize that this is the main and decisive action against the saturation of hospitals.

Therefore, the most recent works have focused on the prognosis of the outcome of the disease, in order to establish early the needs of intensive care units (ICU), and only give access to people who necessarily require medical equipment and the implements provided by these agencies.

The main source of information analyzed for this type of analysis is clinical data, since they provide the most relevant and accurate information. Different authors use different examinations taken as characteristics to train a machine learning model that is later used to predict when the disease will have complications in a patient, the data most

used in these models are lactate dehydrogenase (LDH), Protein C Reactive (PCR), Neutrophils and Lymphocytes [18,26,23,25]. Additionally, in [18], it is shown that the Urea test may be relevant to this task.

Other works have focused on developing a machine learning model for the prognosis of mortality in infected patients [26,23,25], which, for obvious reasons, is highly related to the prognosis of complications of the disease. However, people who recover are downplayed and this data is very important for our study. XGBoost is a learning model used to predict patient mortality [26,25]. In addition, it is a model based on decision trees, frequently used in clinical decision problems due to its high explain ability, since the model defines the characteristics that are most relevant for each decision.

**Figure 1**  
**General scheme of the methodology used**



In the creation of models based on machine learning, it is chosen to carry out cross validation, which aims to extract the best subset for the training data and as a consequence to obtain a better model. In [26] this process is performed 100 times with a separation of 30 % and 70 % for the test and training subsets respectively. The percentages where the training set is greater than or equal to 80 is not recommended since it can generate an overtraining of the model, which indicates that it may not be as good with new data. With this process in [26] some metrics are obtained such as precision around 90 %, precision for prediction of people who live 96 % and those who are not 95 %.

The analyzed sources have the metrics evaluated in the different models proposed, some with a high percentage in their metrics [26,23,25] due to their small number of samples, with 485, 492 and 404 samples respectively, it may not cover all the sufficient amount of information in order to give a general prediction. This problem demands a large number of samples due to its high number of variables that are involved for its prediction, so the more data a better and more general model will be created. One of the advantages of this study is that it has a very large dataset, so we would not have that problem.

## Methodology

Figure 1 illustrates the methodology used in this work for the construction of the learning models that were evaluated. First, the data set provided by the COVID Data Save Lives initiative was refined to extract only those patients who had final outcome information, be it hospital discharge (without admission to the ICU), or who were admitted to the ICU or who died. These patients were grouped into two classes, with complications (admission to the ICU or death) and without complications (hospital discharge).

The clinical data that were going to be used for the construction of the model were extracted and completed from the included patients, which were passed through three models of selection of characteristics. Finally, a subset of the best characteristics was used to train three classification models, which had to decide whether or not the patient would have complications.

## *Dataset and Experimental Setup*

The main input was a set of data from the COVID Data Save Lives initiative of HM hospitals<sup>1</sup>, which includes a total of 2307 patients who

---

1 Full details and information of HM hospitals and COVID Data Save Lives can be found in <https://www.hmhospital.es/coronavirus/covid-data-save-lives/englishversion>

were admitted to the hospital between December 26, 2019 and April 20, 2020, due to COVID-19. This dataset was also used in [17]. The data set includes information on the general data of each patient, the vital signs record (temperature, blood pressure, oxygen content, heart rate and blood glucose), medication supplied to the patient, results of tests performed, hospital follow-up record, emergency register with all procedures performed according to the international disease classification code, 10th edition (ICD-10) and diagnostic images (X-ray and Tomography).

For this work, only patients with a defined initial and final state registered (maximum two days after hospital admission) of para-clinical examinations were considered. Resulting in a data set of 881 patients, 640 of them who were discharged without having registered admission to the ICU, and 241 who had at least one admission to the ICU or even died. Outcome prediction was performed considering these two conditions.

Regarding the data used for the evaluation of the models, only general data, vital signs, and results of the para-clinical exams were considered. Table 1 lists the data and results of para-clinical examinations that were employed. As can be seen, the data set consisted of values of 35 clinical features that were used to estimate final outcome.

The data were categorized, using numerical data, in order to ease processing. Thus, for example: Sex, for male 0 and for female 1; final state of the patient, death 0, Address 1, Transfer to hospital 2, Transfer to a associated-health center 3, Voluntary discharge 4 and no data 5. Additionally, a unification process was carried out in the vital signs records, due to the fact that on the same day there were several takes of these data. This unification consisted of calculating the average per day. Once this was done, only the data of the first days in the records were taken per patient, considering that the decisive period to define whether they will have complications with the disease occurs in the first 10 days after infection. In addition, for the processing and centralization of the information, the data was loaded into a relational database in order to favor the extraction of information [21]. For the training and test sets, a



bootstrapping validation scheme with 80% to train and 20% to test was used to avoid high computational cost for implementing the models. Additionally, all algorithms were implemented using a workstation with an Ubuntu system under a CPU Intel Xeon with 32 cores @ 3.8GHz with 32GB of RAM. Besides, models were trained using Python and essential machine learning libraries such as Numpy, Pandas, Matplotlib, Scipy and Scikit-learn.

Finally, considering that all patients did not undergo the same tests, and that some of these results were included as part of the characteristics of the model, a process was carried out to complete the missing data. This process was carried out by implementing *k*-Nearest Neighbors imputation (*k*- NNimpute), which generates the required value, such as the average of the *k* closest neighbors [19].

**Table 1**  
**Clinical variables used in the training of the models**

Data Source	Feature
2*Basic information	Gender Age
27*Medical exams	Alanine aminotransferase GPT (ALT) Aspartate aminotransferase GOT (AST) Anisocytosis coefficient Creatinine Eosinophil Glucose Red blood cells Hematocrit Hemoglobine Mean Corpuscular Hemoglobin Lactate dehydrogenase LDH Leukocytes Lymphocytes Monocytes Neutrophils Percentage of basophils Eosinophil percentage

	Lymphocyte percentage Monocyte percentage Neutrophil percentage Potassium C-reactive protein Platelet count Urea Medium corpuscular volume Medium platelet volume Sodium
6*Vital signs	Temperature Maximum blood pressure Minimum blood pressure Oxygen saturation Blood glucose Heart rate

### *Feature selection*

In order to find out those variables in the data set that provided the greatest relevance or information to distinguish between patients who presented complications and those who did not, it was decided to evaluate the following characteristics selection techniques:

- **Minimum Redundancy - Maximum Relevance (mRMR):** It uses a selection criterion that seeks to establish, from a feature space  $M$ - dimensional, a subspace of  $m$  features in which ( $m < M$ ) with the greatest relevance to identify the desired class and the least redundancy between them [15].
- **Mutual Information (MI).** Select the features that maximize the mutual information between two random variables that model each feature and the classification labels [20].
- **Random forest (RF).** Although in a strict sense this is not a method for selecting characteristics, this classification method assigns a relevance value to each characteristic, according to the weight they acquire in the decision tree that allows the classification to be carried out [5].

- **Multiple Kernel learning (MKL).** Like Random Forest, MKL is also a classification model. However, in this study the strategy proposed by Areiza et al. [1], who implement a kernel for each characteristic or variable to evaluate, and then take as relevance, the weight of the coefficients of each kernel in the process of combining these.

## *Prediction Models*

Three binary classification models were evaluated, which were trained to distinguish between patients who will require admission to the ICU from those who do not.

- **Random Forest.** It is a widely used model in clinical decision problems, due to its high interpretability. This is an assembly learning method, in which several decision trees are trained, taking subsets of randomly selected characteristics, and the final decision results from averaging the output of these [5,17].
- **Support Vector Machines (SVM).** It is one of the most widely used binary classification models in the literature for offering a performance that surpasses most of the classic artificial learning methods. The method finds a hyperplane that optimally separates the training instances in a feature space, so that a new instance is classified according to its location relative to that hyperplane [8].
- **Multiple Kernel Learning (MKL).** This method generates a linear combination of multiple kernels, which are trained to optimally separate classes based on a set of characteristics [7]. In this case, each kernel was trained to distinguish the two classes based on each attribute, that is, each subset was composed of a single characteristic.

## *Performance metrics*

In order to make an objective comparison of the results, the following performance metrics are extracted. For this, the values are assu-

med as with complications (1) and without them (0). The confusion matrix is calculated and the values of the true positives (TP), true negatives (VN), false positives (FP) and false negatives (FN) are taken. With these values the metrics described below are estimated.

- **Accuracy and Precision.** Accuracy is defined as the number of positive predictions made by the model regarding the actual complication of the patient. On the other hand, precision is defined as the percentage of cases of complication or non-complication correct by the model. Both metrics are defined mathematically from the equations 1 y 2.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

- **Sensitivity and specificity.** Sensitivity and specificity are metrics that make it possible to identify the ability of the model to discriminate those cases that present complications from those that do not. Sensitivity points to the portion of cases that will be complicated, while specificity points to the fraction of cases that will not be complicated. These rules are expressed mathematically in the equations 3 y 4.

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

$$specificity = \frac{TN}{TN + FP} \quad (4)$$

- **F1-Score and Geometric mean.** The objective of estimating the F1-Score and the geometric mean lies mainly in that both estimate the harmonic mean between the precision and sensitivity

values previously calculated. This harmonic mean allows identifying the balance between the detection of cases that present medical complications and those that do not. Both are calculated mathematically according to the equations 5 y 6

$$F1 - Score = 2 \times \frac{Precision \times Sensitivity}{Precision + Sensitivity} \quad (5)$$

$$Geom Mean = \sqrt{Sensitivity \times} \quad (6)$$

## Results

### Feature selection

Table 2 presents the 15 attributes or characteristics that were reported as the most relevant for each of the selection methods implemented. In bold, those data that were considered relevant in the four methods are highlighted, while those that were presented in three of them are indicated with underlined text. As can be seen, three para-clinical results were considered relevant by the four selection models, these are: LDH, Creatinine and sodium level. Regarding the vital signs on admission and general data, it is observed that, although some models consider the values of temperature (mRMR) and oxygen saturation (MI) relevant, they do not seem to be really relevant for the decision. On the other hand, age is identified as a variable that contributes to the decision in three of the four models.

In order to determine how much the results obtained by each model were similar, the Euclidean distance between them is calculated, taking the position of each attribute given by each model as an attribute in a 35-dimensional space. Table 3 presents these results. Note that the selection method of maximum relevance-minimum redundancy is the one with the greatest distance from the other three. This can also be seen in Table 2, since this is the method in which the least selected attributes

match those of the other methods. Likewise, the mutual information and Random Forest methods show the highest agreement.

### *Outcome prediction*

To determine whether or not a patient was going to present complications, and consequently if they would require admission to the ICU, the performance of the three classification models described above was evaluated. Each model was trained with different subsets of characteristics (10-20 and 35). The subset of 20 features being the one that offered the best performance. Tables 4, 5 and 6 present the performance with Random Forest, SVM and MKL respectively. The number of decision trees was set to 100 for the Random Forest model, due to the fact that it is widely used [27,12]. The SVM was trained with a linear kernel, in addition, an additional process is carried out in which the data is scaled, normalized and the regularization value is set at 1 [16]. The Random Forest and SVM models were trained with 70 % of the data and with the remaining 30 % an evaluation was made and thus the metrics were generated. In the case of MKL, radial base kernels were used, whose parameters were optimized with respect to the F1-Score, giving greater weight to sensitivity.

The method that reports the highest accuracy (85.66 %) is the combination of the features selected by the mutual information method and SVM. This method also offers the best sensitivity (83.02%) with good specificity (86.32 %), which affects the other performance metrics, which depend on these such F1-Score and geometric mean. The best specificity was obtained with the combination of the attributes selected and classified with MKL (87.54 %), which also presents an acceptable sensitivity for this task (78.39 %). Note that the precision in general is low (65 %), which may be precisely due to the imbalance between the classes presented by the dataset.

**Table 2**  
**15 most relevant features resulting from each selection method**

<b>Random Forest</b>	<b>mRMR</b>
<b>Lactate dehydrogenase LDH</b>	<b>Lactate dehydrogenase LDH</b>
<b>Sodium</b>	Gender
Neutrophil percentage	<u>Lymphocytes</u>
<u>Urea</u>	Potassium
Lymphocyte percentage	Medium platelet volume
Age	Red blood cells
Creatinine	Temperature
<i>C-reactive protein</i>	Mean Corpuscular Hemoglobin
Glucose	Percentage of basophils
Platelet count	Monocytes
<u>Lymphocytes</u>	Eosinophil percentage
<u>Medium corpuscular volume</u>	<b>Creatinine</b>
Neutrophils	<u>Leukocytes</u>
Leukocytes	<b>Sodium</b>
Monocyte percentage	Hemoglobin
<b>Mutual Information</b>	<b>MKL</b>
Neutrophil percentage	<b>Sodium</b>
<b>Lactate dehydrogenase LDH</b>	Age
Lymphocyte percentage	Aspartate aminotransferase GOT (AST)
Neutrophils	<u>Lymphocyte percentage</u>
<u>Leukocytes</u>	Medium platelet volume
<b>Creatinine</b>	<u>Urea</u>

<u>Age</u>	Platelet count
Sodium	<u>Medium corpuscular volume</u>
Glucose	Alanine aminotransferase GPT (ALT)
<u>Urea</u>	<b>Lactate dehydrogenase LDH</b>
<u>Lymphocytes</u>	<b>Creatinine</b>
Monocyte percentage	Anisocytosis coefficient
<u>Medium corpuscular volume</u>	Hematocrit
Oxygen saturation	Mean Corpuscular Hemoglobin
Anisocytosis coefficient	C-reactive protein

## Discussion

This work proposes the use of supervised machine learning models to determine if a patient is going to present complications and in this way contribute to better planning regarding the use of Intensive Care Units, using basic clinical data that are acquired on admission, hospital with a patient interview and routine clinical tests.

**Table 3**  
**Distance between the position of relevance established**  
**by each selected method and its counterparts**

Selection Technique	RF	MKL	mRMR	MI
RF	0.00	47.56	50.97	17.97
MKL		0.00	50.24	52.53
mRMR			0.00	49.53
MI				0.00



**Table 4**  
**Classification performance of the Random Forest algorithm**  
**using the 20 most relevant features of each feature selector**

Metrics	RF	MKL	mRMR	MI
Accuracy	<b>83.77 %</b>	82.26 %	81.89 %	<b>83.77 %</b>
Sensitivity	75.00 %	74.07 %	71.93 %	<b>77.78 %</b>
Precision	<b>61.64 %</b>	54.79 %	56.16 %	57.53 %
Specificity	<b>86.34 %</b>	84.36 %	84.62 %	85.31 %
F1-Score	<b>67.67 %</b>	62.99 %	68.08 %	66.14 %
Geo-mean	80.47 %	79.05 %	78.02 %	<b>81.46 %</b>

**Table 5**  
**Classification performance of the SVM algorithm**  
**using the 20 most relevant features of each feature selector**

Metrics	RF	MKL	mRMR	MI
Accuracy	85.28 %	84.15 %	84.91 %	<b>85.66 %</b>
Sensitivity	81.48 %	76.27 %	80.00 %	<b>83.02 %</b>
Precision	60.27 %	<b>61.64 %</b>	60.27 %	60.27 %
Specificity	86.26 %	<b>86.41 %</b>	86.19 %	86.32 %
F1-Score	69.29 %	68.18 %	68.75 %	<b>69.84 %</b>
Geo-mean	83.83 %	81.18 %	83.04 %	<b>84.65 %</b>

**Table 6**  
**Classification performance of the MKL algorithm**  
**using the 20 most relevant features of each feature selector**

Metrics	RF	MKL	mRMR	MI
Accuracy	80.70 %	85.47 %	81.95 %	81.84 %
Sensitivity	65.92 %	78.39 %	68.30 %	67.69 %
Precision	61.00 %	64.73 %	63.49 %	64.32 %
Specificity	85.71 %	87.54 %	86.61 %	86.81 %
F1-Score	63.36 %	70.91 %	65.81 %	65.96 %
Geo-mean	75.17 %	82.84 %	76.91 %	76.65 %

In the process of selection of characteristics, it is observed that the four models used coincide in including among the 15 most relevant characteristics to three results of paraclinical tests, these are: the levels of lactate dehydrogenase (LDH), creatinine and sodium. Other characteristics that are also relevant and appear in at least 3 of the 4 selection techniques are: Urea, percentage of lymphocytes, lymphocytes, mean corpuscular volume, leukocytes and age; which is consistent with previous work, for example [26,23] identified that Lactate dehydrogenase (LDH), Lymphocytes have a high relevance when determining the future state of a patient.

In the analysis of the metrics, it can be seen that in the SVM model, with a data set given by mutual information, the highest sensitivity is found, so we can say that it has the greatest capacity to determine the people who will be complicated in the data resolution period. However, it is worth noting that the most specific is found in the MKL model using its own data set after the selection process, this indicates that this model is adequate to determine those patients who will not have complications in the outcome of the disease.

This work had some limitations that prevent reaching fully generalizable conclusions. First, the data used for the analysis have been extracted from a specific population group, which does not represent the world population. In this sense, specific variables associated with regional conditions, or the variation in knowledge of the SARS-Cov-2 virus acquired by this population could be the determinants of complications in patients [14,4]. In this sense, an experimental validation with additional data belonging to different populations and / or at different times of the pandemic is required in order to determine the applicability of the results. Additionally, there is no information on the initial state of admission of the patient, that is, if he had an advanced or initial state of the disease. This is critical to determine the true early predictive ability of these methods.

On the other hand, the data set did not contain information on medical history or habits, which have shown a high correlation with possible

complications of the disease [24,3]. In addition, the authors decided not to include variables associated with the initial symptoms, due to the fact that a high number of incomplete data were presented. These parameters could be considered in a larger study. Furthermore, the use of other information, such as medical images, in combination with the paraclinical results, could have a positive effect on the performance of the methods.

Finally, due to time constraints and access to computing resources, this study did not carry out an intensive or optimal search for the configuration parameters of the classification models. In this sense, an exhaustive evaluation of the different parameters could report a better performance of some of the methods.

## Conclusions

In this work, the use of computational learning models has been proposed for the prediction of the outcome of patients diagnosed with COVID-19. The results obtained are promising, since they allow obtaining a good balance between sensitivity and specificity, with data that can be obtained easily, even in low complexity medical centers. Although the prediction model does not allow to be totally conclusive, regarding the possible outcome of the disease in a specific patient, these can be used in combination with other clinical parameters, to help the doctor in making decisions.

Although this is an initial work, as future work it is intended to include the use of different types of medical images in the learning models, since the databases used have this information but these have not been analyzed or used in this work. Including multimodal data fusion models, it is intended to increase performance rates to determine more reliably the complication status of the disease.

## Acknowledgements

This work was supported by the Agencia de Educación Superior de Medellín – Sapiencia, the Universidad Politécnica Salesiana, Ecu-

dor, the Institución Universitaria Pascual Bravo, Colombia and the Instituto Tecnológico Metropolitano, Colombia. We also want to give special thanks to the HM Hospitals, Spain for providing access to the COVID-19 data employed to develop this work.

## References

- [1] Areiza-Laverde, H.J., Castro-Ospina, A.E., Hernández, M.L., D'íaz, G.M.: A novel method for objective selection of information sources using multi-kernel svm and local scaling. *Sensors* 20(14), 3919 (2020)
- [2] Assaf, D., Gutman, Y., Neuman, Y., Segal, G., Amit, S., Gefen-Halevi, S., Shilo, N., Epstein, A., Mor-Cohen, R., Biber, A., et al.: Utilization of machine-learning models to accurately predict the risk for critical covid-19. *Internal and emergency medicine* 15(8), 1435–1443 (2020)
- [3] Bajgain, K.T., Badal, S., Bajgain, B.B., Santana, M.J.: Prevalence of comorbidities among individuals with covid-19: A rapid review of current literature. *American journal of infection control* (2020)
- [4] Brantley, H.L., Yoo, R.M., Jones, G.I., Stock, M.A., Park, P.J., Sheils, N.E., Kohane, I.S.: Variation across population subgroups of covid-19 antibody testing performance. *medRxiv* (2020)
- [5] Breiman, L.: Random forests. *Machine learning* 45(1), 5–32 (2001)
- [6] Condes, E., Arribas, J.R., et al.: Impact of covid-19 on madrid hospital system. *Enfermedades Infecciosas Y Microbiología Clínica* (2020)
- [7] Gönen, M., Alpaydın, E.: Multiple kernel learning algorithms. *The Journal of Machine Learning Research* 12, 2211–2268 (2011)
- [8] Hearst, M.A., Dumais, S.T., Osuna, E., Platt, J., Scholkopf, B.: Support vector machines. *IEEE Intelligent Systems and their applications* 13(4), 18–28 (1998)
- [9] Lauer, S.A., Grantz, K.H., Bi, Q., Jones, F.K., Zheng, Q., Meredith, H.R., Azman, A.S., Reich, N.G., Lessler, J.: The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine* 172(9), 577–582 (2020)
- [10] Leisman, D.E., Deutschman, C.S., Legrand, M.: Facing covid-19 in the icu: vascular dysfunction, thrombosis, and dysregulated inflammation. *Intensive Care Medicine* pp. 1–4 (2020)
- [11] Mahase, E.: China coronavirus: Who declares international emergency as death toll exceeds 200. *BMJ: British Medical Journal (Online)* 368 (2020)

- [12] Meng, C.z., Liu, B.s., Zhou, L.: The practice study of consumer credit risk based on random forest. In: 2019 International Conference on Modeling, Analysis, Simulation Technologies and Applications (MASTA 2019). Atlantis Press (2019)
- [13] Murillo-Cabezas, F., Vigil-Martín, E., Raimondi, N., Pérez-Fernández, J.: Covid19 pandemic and digital transformation in critical care units. *Medicina intensiva* (2020)
- [14] Nepomuceno, M.R., Acosta, E., Alburez-Gutierrez, D., Aburto, J.M., Gagnon, A., Turra, C.M.: Besides population age structure, health and other demographic factors can contribute to understanding the covid-19 burden. *Proceedings of the National Academy of Sciences* 117(25), 13881–13883 (2020)
- [15] Radovic, M., Ghalwash, M., Filipovic, N., Obradovic, Z.: Minimum redundancy maximum relevance feature selection approach for temporal gene expression data. *BMC bioinformatics* 18(1), 1–14 (2017)
- [16] Raihan-Al-Masud, M., Mondal, M.R.H.: Data-driven diagnosis of spinal abnormalities using feature selection and machine learning algorithms. *Plos one* 15(2), e0228422 (2020)
- [17] Sánchez-Montan˜es, M., Rodríguez-Belenguer, P., Serrano-Lo´pez, A.J., SoriaOlivas, E., Alakhdar-Mohmara, Y.: Machine learning for mortality analysis in patients with covid-19. *International journal of environmental research and public health* 17(22), 8386 (2020)
- [18] Siordia Jr, J.A.: Epidemiology and clinical features of covid-19: A review of current literature. *Journal of Clinical Virology* p. 104357 (2020)
- [19] Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., Altman, R.B.: Missing value estimation methods for DNA microarrays. *Bioinformatics* 17(6), 520–525 (06 2001). <https://doi.org/10.1093/bioinformatics/17.6.520>
- [20] Vergara, J.R., Estévez, P.A.: A review of feature selection methods based on mutual information. *Neural computing and applications* 24(1), 175–186 (2014)
- [21] Widenius, M., Axmark, D., Arno, K.: MySQL reference manual: documentation from the source. ” O’Reilly Media, Inc.” (2002)
- [22] World Health Organization (WHO): COVID-19 Weekly Epidemiological Update. Tech. Rep. November, World Health Organization (2020), <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20201012-weekly-epi-update-9.pdf>

- [23] Wu, Q., Wang, S., Li, L., Wu, Q., Qian, W., Hu, Y., Li, L., Zhou, X., Ma, H., Li, H., et al.: Radiomics analysis of computed tomography helps predict poor prognostic outcome in covid-19. *Theranostics* 10(16), 7231 (2020)
- [24] Wynants, L., Van Calster, B., Collins, G.S., Riley, R.D., Heinze, G., Schuit, E., Bonten, M.M., Dahly, D.L., Damen, J.A., Debray, T.P., et al.: Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *bmj* 369 (2020)
- [25] Yan, L., Zhang, H.T., Goncalves, J., Xiao, Y., Wang, M., Guo, Y., Sun, C., Tang, X., Jin, L., Zhang, M., et al.: A machine learning-based model for survival prediction in patients with severe covid-19 infection. *MedRxiv* (2020)
- [26] Yan, L., Zhang, H.T., Goncalves, J., Xiao, Y., Wang, M., Guo, Y., Sun, C., Tang, X., Jing, L., Zhang, M., et al.: An interpretable mortality prediction model for covid-19 patients. *Nature Machine Intelligence* pp. 1–6 (2020)
- [27] Zhang, H., Bi, Y., Jiang, W., Luo, C., Cao, S., Guo, P., Zhang, J.: Application of random forest classifier in loan default forecast. In: *International Conference on Artificial Intelligence and Security*. pp. 410–420. Springer (2020)

