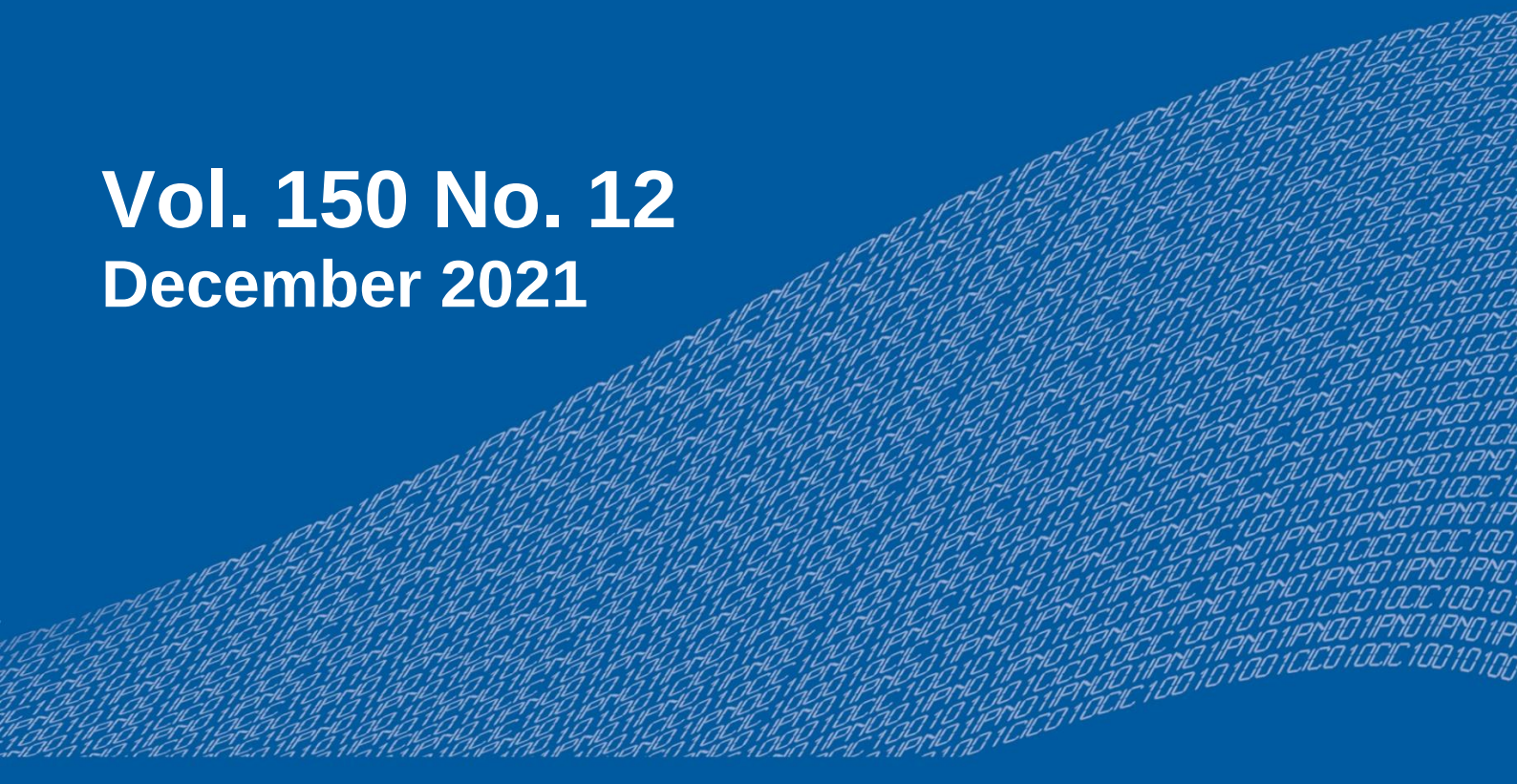




Research in Computing Science

Vol. 150 No. 12
December 2021



Research in Computing Science

Series Editorial Board

Editors-in-Chief:

Grigori Sidorov, *CIC-IPN, Mexico*
Gerhard X. Ritter, *University of Florida, USA*
Jean Serra, *Ecole des Mines de Paris, France*
Ulises Cortés, *UPC, Barcelona, Spain*

Associate Editors:

Jesús Angulo, *Ecole des Mines de Paris, France*
Jihad El-Sana, *Ben-Gurion Univ. of the Negev, Israel*
Alexander Gelbukh, *CIC-IPN, Mexico*
Ioannis Kakadiaris, *University of Houston, USA*
Petros Maragos, *Nat. Tech. Univ. of Athens, Greece*
Julian Padget, *University of Bath, UK*
Mateo Valero, *UPC, Barcelona, Spain*
Olga Kolesnikova, *ESCOM-IPN, Mexico*
Rafael Guzmán, *Univ. of Guanajuato, Mexico*
Juan Manuel Torres Moreno, *U. of Avignon, France*

Editorial Coordination:

Griselda Franco Sánchez

Research in Computing Science, Año 20, Volumen 150, No. 12, diciembre de 2021, es una publicación mensual, editada por el Instituto Politécnico Nacional, a través del Centro de Investigación en Computación. Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othon de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738, Ciudad de México, Tel. 57 29 60 00, ext. 56571. <https://www.rcs.cic.ipn.mx>. Editor responsable: Dr. Grigori Sidorov. Reserva de Derechos al Uso Exclusivo del Título No. 04-2019-082310242100-203. ISSN: en trámite, ambos otorgados por el Instituto Politécnico Nacional de Derecho de Autor. Responsable de la última actualización de este número: el Centro de Investigación en Computación, Dr. Grigori Sidorov, Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othon de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738. Fecha de última modificación 01 de diciembre de 2021.

Las opiniones expresadas por los autores no necesariamente reflejan la postura del editor de la publicación.

Queda estrictamente prohibida la reproducción total o parcial de los contenidos e imágenes de la publicación sin previa autorización del Instituto Politécnico Nacional.

Research in Computing Science, year 20, Volume 150, No. 12, December 2021, is published monthly by the Center for Computing Research of IPN.

The opinions expressed by the authors does not necessarily reflect the editor's posture.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of Centre for Computing Research of the IPN.

Advances in Computing Science and Applications

Abdiel Reyes Vera (ed.)



Instituto Politécnico Nacional
“La Técnica al Servicio de la Patria”



Instituto Politécnico Nacional, Centro de Investigación en Computación
México 2021

ISSN: in process

Copyright © Instituto Politécnico Nacional 2021

Formerly ISSN: 1870-4069, 1665-9899

Instituto Politécnico Nacional (IPN)
Centro de Investigación en Computación (CIC)
Av. Juan de Dios Bátiz s/n esq. M. Othón de Mendizábal
Unidad Profesional “Adolfo López Mateos”, Zacatenco
07738, México D.F., México

<http://www.rcs.cic.ipn.mx>

<http://www.ipn.mx>

<http://www.cic.ipn.mx>

The editors and the publisher of this journal have made their best effort in preparing this special issue, but make no warranty of any kind, expressed or implied, with regard to the information contained in this volume.

All rights reserved. No part of this publication may be reproduced, stored on a retrieval system or transmitted, in any form or by any means, including electronic, mechanical, photocopying, recording, or otherwise, without prior permission of the Instituto Politécnico Nacional, except for personal or classroom use provided that copies bear the full citation notice provided on the first page of each paper.

Indexed in LATINDEX, DBLP and Periodica

Electronic edition

Table of Contents

	Page
Performance Evaluation and Comparison on Lifi, Wifi and Ethernet Networks in Real Environments	5
<i>Gerardo Hernandez-Oregon, Mario E. Rivero-Angeles, Jorge E. Coyac-Torres, Juan C. Chimal-Eguía</i>	
Criptoanálisis y mejora a esquema de cifrado de imágenes hipercaótico	25
<i>Marco T. Ramírez-Torres, César A. Guerra-García, María F. Mendoza Martínez, Jazmín Rangel Saucedo, Carlos Soubervielle Montalvo</i>	
Fusión de sensores mediante filtro de Kalman extendido para seguimiento de trayectorias de un robot móvil tipo rover	37
<i>Jheison Duvier Diaz-Ortega, Ponciano Jorge Escamilla-Ambrosio, Octavio Gutiérrez-Frías, Alberto Luviano-Juárez, José Alejandro Aguirre-Anaya</i>	
Desarrollo de una red neuronal artificial para la predicción de sobre corte en un proceso de manufactura no convencional	49
<i>Gabriela Itzell Rangel Velázquez, José Arturo Pérez Martínez, Everardo Efrén Granda Gutiérrez, Carlos Eduardo Torres Reyes</i>	
Investigación sobre el diseño de software en la última década: Un mapeo sistemático de literatura	63
<i>José Baeza, Omar S. Gómez, Raúl A. Aguilar</i>	
Sistema de detección omnidireccional de gestos corporales para entornos virtuales	75
<i>Iván Cruz García, Héctor Ortiz Cruz, Marco Antonio Solís López</i>	
Análisis y comparación de la clasificación de piezas metálicas regulares mediante redes neuronales convolucionales con y sin análisis de componentes principales	97
<i>Mayra Mateo Jiménez, Carlos Eduardo Torres Reyes, Everardo Efrén Granda Gutiérrez, José Arturo Pérez Martínez</i>	
Análisis de comentarios de turistas sobre áreas naturales protegidas de México utilizando plataformas Google	109
<i>Cynthia Yustis-Cruz, Magdalena Saldana-Perez, Marco Moreno-Ibarra</i>	

Tecnologías emergentes para la reincorporación de personal a sus actividades durante la pandemia de COVID-19	127
<i>Magdalena Saldana-Perez, Fatima Figueroa-Ayala, Carolina Palma Preciado, Marco Moreno-Ibarra</i>	
A First Approach to Asynchronous-Synchronous Tradeoff in 1D Cellular Genetic Algorithms.....	141
<i>José Alejandro Cornejo Acosta¹, Jesús García Díaz</i>	
Aplicación web de secuenciador musical basado en datos de los movimientos oculares (eye-tracking).....	149
<i>Jonathan Manuel Fraire-Lopez, Francisco Lopez-Orozco, Jesus Israel Hernandez-Hernandez, Juan Gabriel Ruiz-Ruiz</i>	
Uso de diccionarios en Python para conocer las tendencias en Twitter sobre el aumento de feminicidios en México	161
<i>Marleni Reyes Monreal, Beatriz Bernábe Loranca, Estefania Sarmiento Barrios, Melissa Isaaly Mendoza Bernábe</i>	
Clasificación de información personal sensible.....	179
<i>Sara De Jesús Sánchez, Jorge Enrique Coyac Torres, Eleazar Aguirre Anaya, Hiram Calvo</i>	
Análisis de sentimientos para el estudio del comportamiento en noticias de COVID-19 en redes sociales como series de tiempo	187
<i>Jesús Hernández Padilla, Hiram Calvo, Ildar Batyrshin</i>	
Estudios sobre el reconocimiento del habla imaginada mediante señales EEG	205
<i>Arturo Iván Sandoval Rodríguez, Francisco Hiram Calvo Castro</i>	
Algoritmo de cifrado de imágenes utilizando autómatas celulares reversibles.....	215
<i>Erendira Corona-Bermúdez, Juan Carlos Chimal-Eguía, Uriel Corona-Bermúdez</i>	
Simulación computacional de nanosistemas para detección de la proteína hemaglutinina	225
<i>J. A. Salazar Cano, J. I. Rodriguez Hernandez, J. C. Chimal Eguía</i>	

Performance Evaluation and Comparison on LiFi, WiFi and Ethernet Networks in Real Environments

Gerardo Hernandez-Oregon, Mario E. Rivero-Angeles
Jorge E. Coyac-Torres, Juan C. Chimal-Eguía

Instituto Politécnico Nacional,
Centro de Investigación en Computación,
Mexico

chimal@cic.ipn.mx

Abstract. Light Fidelity (LiFi) is an emerging technology in the communications and networks fields that uses lighting RF specter as a medium to link data transmissions. In the first instance, this work presents the characterization and operation of two prototypes of LiFi luminaires. Furthermore, this paper shows different approaches, including the creation of a physical layer for LiFi luminaires, demonstrating each of the stages for transmission and acquisition of the signal, the bridges created to achieve the data link, and even the integration of all before phases and factors to result in a whole LiFi communication system ready to be implemented in the real world. Although to guarantee the quality of service in those communication networks, there are many factors to be considered. Therefore, the performance of a LiFi network could be obtained and predicted if an analysis model can consider specific metrics to adjust the network performance to the user's needs. For the specific case of this work, a set of metrics [21], such as packet loss, delay (latency), jitter, and throughput, was experimentally calculated, presenting the graphical results for LiFi metrics performance facing WiFi and Ethernet technologies. Finally, this work shows the importance of data density for LiFi networks by analyzing the results from distance and speed variation.

Keywords: LiFi, VLC, data-link, wireless communication, optical communication, wiFi, ethernet.

1 Introduction

The study and creativity to combine two quotidian human needs, such as lighting and communication in a single solution, have achieved the development of optical devices through researching fields in optoelectronics and computing. Optical technologies, nowadays, represent a solution to current communication problems. Even though these technologies, just a few decades ago, have been studied is surprising that the first experiments with visible light have an antiquity of at least 3000 years [15].

Light-Fidelity(LiFi) emerges as one of the technologies that will shape the future of communication systems. In the last decade, LiFi technology has grown considerably, having as main strengths the range of frequencies in which it works (380 nm to 750

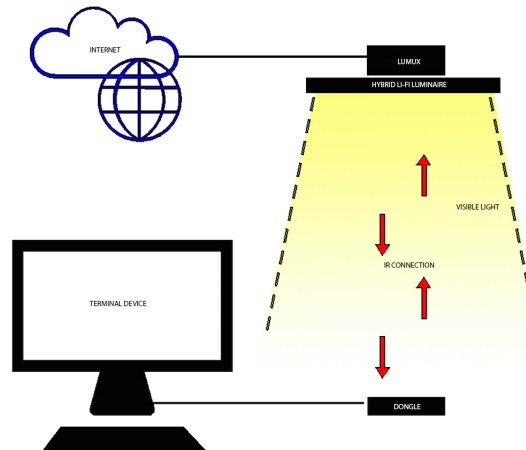


Fig. 1. LiFi connection, through a prototype luminaire, to the Internet.

nm [1]), and the fact of having an infrastructure almost ready for its implementation, making possible the development of new technological perspectives.

Each year the number of users with Internet access is doubled, and the amount of data transmitted each month (estimated at 77 exabytes just for Mobile Data Traffic for the year 2022 [10]) grows in an accelerated way. Therefore, the current networks tend to present congestion and low performance. For this reason, LiFi is presented as a technology not only capable of solving the problems in congestion or performance but as a technological environment capable of interacting and complementing existing technologies such as WiFi or Ethernet.

Human beings have developed lighting systems in any place necessary, and homes were not the exception. For this last reason, the indoor LiFi systems are easy to implement and low cost considering the change of LiFi lamps instead of those currently used. In a typical LiFi network, such as *Figure 1* shows, luminary cells transmit and receive from an access point placed inside a room. Lighting to any point on the receiving surface includes the Line of Sight (LOS) and reflections of objects and walls [16]. LiFi indoors could represent the largest field of action to this technology in the following decades because it represents a potential solution for the Internet of things [12] through experimentation in modulation, color, and power to use this technology in urban environments. In the future, it would be applicable within marine environments or in outer space. Also, it is possible to restructure the lighting in public areas to implement LiFi where necessary.

Although IEEE 802.7.15 and IEEE 802.11 protocols in LiFi sections do not finish yet, this article presents patented prototypes (previously developed by this working group [7, 8]) and finds its main contribution in the methodological, electronic, mechanical, and computational description of LiFi systems. In literature, several models of theoretical operation exist of these systems [5, 19, 20]. So, another essential contribution for this paper is to experimentally demonstrate the performance of current LiFi networks, comparing them concerning similar WiFi and Ethernet environments

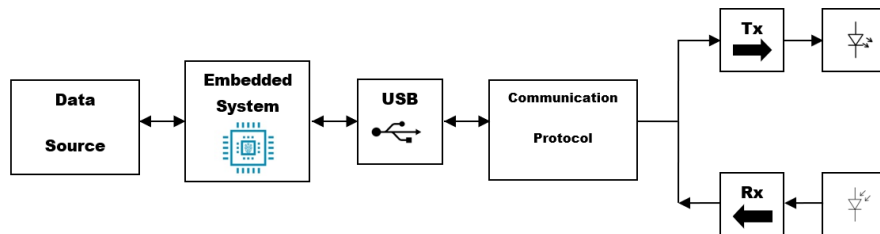


Fig. 2. General LiFi system.

(since these last two technologies are well known, their technical and functional characteristics are not emphasized in this work).

2 LiFi System

Light Fidelity (LiFi) is an emerging technology in the communications and networks fields that uses lighting RF specter as a medium to link data transmissions. Therefore, it uses light radiation, such as emitted light from a Light Emitting Diode (LED), to transfer and receive data and even connect to the Internet. This paper presents the stages of design, implementation, and analysis of two LiFi prototypes. Thus, the activities presented in this work are also thought to give the reader a comparison of this technology against the popular ones that nowadays domain the IT communication technologies like WiFi and Ethernet. The importance of researching in the LiFi field was discussed in Harald Haas and et al. [11] work. They explain the need for a new spectrum for current and future wireless networks. Furthermore, they argue that wireless networks based on light waves are at a very mature state to begin implementing in the industry as an opportunity area. Thus, this work will show the phases and diagrams implemented, during the design and analysis stages, to finally compare these couple of LiFi prototypes against the before mentioned technologies, showing their capability to establish an Internet connection, and giving the user the connectivity experience such as standard communication protocols are.

The diagram shown in *Figure 2* represents the implemented overall system on the LiFi transmitters and receivers. Below is a description of each of the system stages:

- In the first stage, a signal conditioned by the 802.3 protocol is received, which may contain a Manchester or MLT-3 encoding. The MLT-3 coding is specifically used in high-speed communications, due to the fact that MLT-3 has three voltage levels in the signal (-1,0,1) the frequency of the transmitted signal decreases by half, unlike a coding *Manchester no return to zero*. The *Ethernet* signal is isolated by security through a transformer, to be interpreted by a *transceiver*, which will be responsible for encoding and decoding the information provided in *Data Link Layer* to subsequently receive and transmit it to a system embedded.

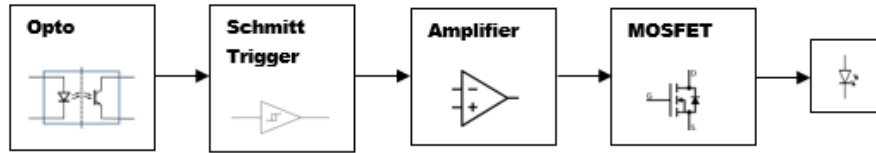


Fig. 3. Transmitter system.

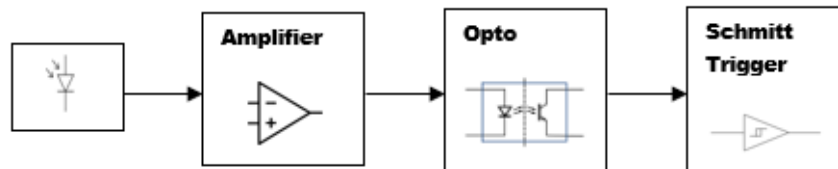


Fig. 4. Receiver system.

- The embedded system is responsible to establishing an information bridge between an *Ethernet port* and a *USB port* through a layer 2 communication protocol.
- Subsequently, the information received in the *USB port* is converted to a communication protocol, which takes data bytes and transmits them sequentially in a certain format. For establish communication, it is necessary to pair the two devices with the same communication parameters.
- In *Figure 3*, the transmitter system is presented. In this part, the signal is optoisolated to the next stage and is processed by a *Schmitt trigger* gate. In order to transmit this signal, it is necessary to condition it with an operational amplifier, which has as a characteristic a higher bandwidth than the bandwidth of the signal to be transmitted (at least ten times greater). At the same time, this amplifier is used as a drive to switch a *MOSFET* device that will turn on and off the *IR* information transmitter.
- In the reception stage, there is a photodiode, which receives the signal to be processed. This current signal passes through a transconductance amplifier and delivers a voltage signal for being amplified to have compatible levels with *TTL* and *CMOS* signals. Once conditioned, the transmission signal is optoisolated to a serial-USB conversion circuit for being reprocessed by the embedded system and returned to the *Ethernet* network. *Figure 4* shows this set of stages.

The sources used in hybrid white LED luminaires systems are Lambertian sources of visible and infrared light [9]. When radiating in all directions, the DC gain through the i -th LOS communications channel(H_{LOS}) [4] corresponds to *Equation 1*:

$$H_{i,LOS} = \begin{cases} \frac{A(m+1)}{2\pi r_i^2} \cos^m \theta_i \cos \psi_i & \psi_i \leq \psi_{max} \\ 0 & otherwise \end{cases}, \quad (1)$$

where,

A : receiver area

m : Lambert index

θ_i : irradiance angle with respect to the transmitter axis

ψ_i : incidence angel with respect to the receiver angle

ψ_{max} : field-of-view (FOV) semiangle of the receiver

r_i : distance between i -th transmitter and i -th optical detector

3 Design of LiFi Luminaires

3.1 General Modifications to Luminaires

The LiFi hybrid(Visible and Non-Visible light) luminaires were designed in two main parts:

The first one is the luminaire (Lumux): it is a case in which the electronic board (required to send and receive the information) is mounted. This information is suitable for transmission by infrared LEDs; even when the visible lights are turned off in the room where this technology is being used, this fact allows to continue with the transmission.

LiFi luminaires used in this work are a hybrid version between visible light and infrared light (*Figure 5*). This luminaire has a variable and adjustable transmission speed, in the direct link, of 1 Mbps, 3 Mbps, 6Mbps, 12Mbps, or 24Mbps. The speed of transmission is due to the adaptation, positioning, and study of the optics of the LEDs. These factors will represent an essential part of the study of the luminaries that transmit data only in visible light.

The second part for a LiFi luminaire is the Dongle(receptor). Unlike Lumux, Dongle is the device that interacts directly with the end-user, so it needs to cover its electronic board completely. Another important difference is that Dongle contains a single receiver and a single transmitter, representing a considerable reduction in its size (compared to the Lumux size). The Dongles presented in this paper, such as in *Figure 6*, are connected to the electronic devices through a male USB type A connector.

3.2 Communication System

Once the signal has been acquired in the electronic devices of the *Physical Layer* and converted into bits, the system uses a protocol of the *Data Link Layer* to encapsulate the frames with IP headers to the *Network Layer*.

The above process will make possible internet connections through the LiFi luminaires. In this work, a *Single-Board Computer* was used in which the necessary

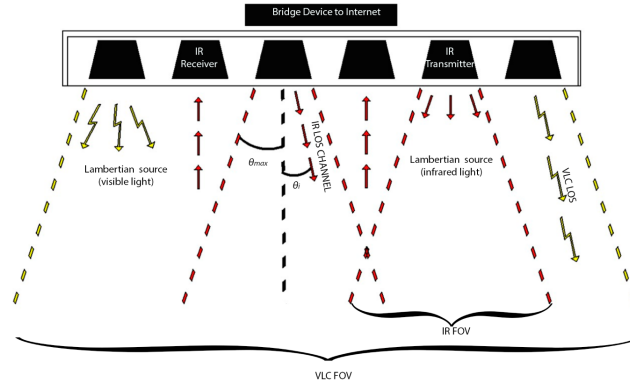


Fig. 5. LiFi Lumux luminaire.

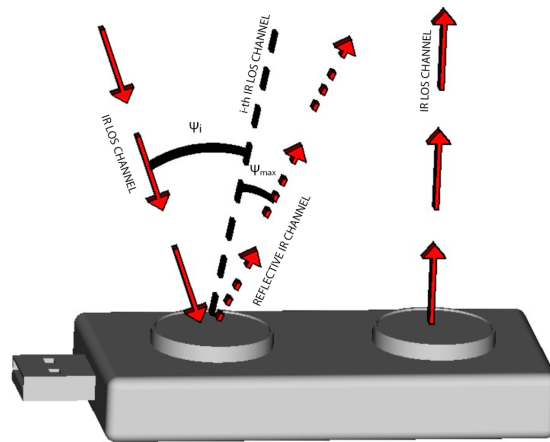


Fig. 6. LiFi receiver scenario, Dongle.

files were created for Lumux and Dongle configurations. The procedure to achieve the communication between Lumux and Dongle is described by *Algorithm 1*.

In addition to the entire functionality process in Lumux and Dongle devices, there are some characteristics in the operation of the LiFi communication prototypes used for this work, these characteristics are described below:

- **Roaming:** As we have seen in several wireless communications technologies, mobility is an essential factor that makes more comfortable and easy communications among multiple users or even a per of them; meanwhile, data exchanges occur for each station.

There are two types of roaming for mobile terminals [3]. The first one is called Intrasytem (intradomain), and another one is called Intersystem (interdomain) roaming. Therefore, we define our proposed roaming as a type of Intrasytem mobility management because it uses network interfaces and protocols similar

Algorithm 1 Internet connection through LiFi luminaires

Input $\leftarrow$ Internet request
Output $\rightarrow$ Internet connection
1: Initialization
2: Settings of communication parameters
3: Settings of proxy connection
4: IP forwarding activation
5: NAT deployment
6: Allowing of Layer two protocols
7: **active_communication** $\leftarrow$ true
8: **while** active_communication **do**
9: Send.Data
10: Receive.Data
11: Check active_communication
12: **end**

among all stations for communicating with each other. Unlike before the roaming type, Intersystems refers to moving between different backbones, protocols, technologies, or service providers. Also, a mechanism that works internally on the OS was created, such as a software service. This mechanism is continuously looking for a connection request from a client station. Once that request is recognized, this service connects from layer two to upper layers' protocols to give the client station an Internet connection.

In the beginning, roaming mechanism activates IP forwarding to apply NAT between upper layers protocols to layer two, converting data into frames ready to be sent to the physical layer of our Li-Fi prototype and vice-versa. Thus, the roaming mechanism needs to define some properties such as physical port, speed, network parameters, and others to initialize this service. Next, the roaming service checks if communication is active; in this case, the bridge starts sending and receiving data and continuously checks if the connection is online. In another case, the bridge gets close to waiting for another connection request, as is shown in *Algorithm 2*.

- **Multiple users per LiFi luminaire:** LiFi communications are characterized by providing dedicated links to the clients that connect to them. However, a single LiFi luminaire can provide connectivity to several terminal devices depending on the dimensions of the Lumux and the scope of its Field of View (FOV) [22]. The LiFi prototypes presented in this work can provide from 4 to 16 users per luminaire, considering the aspects mentioned earlier.
- **Operating systems compatibility:** The integration and compatibility between the technological diversity that exists today is one of the most important challenges of new technologies. The luminaire prototypes used for testing offer compatibility across Windows, Linux, and Mac OS operating systems.

Algorithm 2 Bridge to the Internet connection

```

Input  $\leftarrow$  Internet request
Output  $\rightarrow$  Internet connection
1: Start roaming service
2: r_service  $\leftarrow$  true
3: while r_service do
4:   Check Internet_request
5:   if Internet_request = true then
6:     Initialization
7:     Settings of communication parameters
8:     Settings of proxy connection
9:     IP forwarding activation
10:    NAT deployment
11:    Allowing of Layer two protocols
12:    active_communication  $\leftarrow$  true
13:    while active_communication do
14:      Send_Data
15:      Receive_Data
16:      Check active_communication
17:    end
18:  end
19:  if roaming service is not Alive then
20:    r_service  $\leftarrow$  false
21:  end
22: end

```

- **Data density, bandwidth, and speed:** Two factors define LiFi technology: high data density and bandwidth. Both factors result in the theoretical speed being up to 10Gbps [2], an extremely high speed compared to conventional speeds for WiFi (150 Mbps). It is important to mention that LiFi provides a higher *Data Density* than WiFi since the density of this data is approximately 1000 times greater than in WiFi [18]. This is due to the lower interference that light presents on radio waves; that is why a LiFi system could have the same performance for certain data traffic than a 144.4 Mbps WiFi connection (as will be shown in section III).

3.3 LiFi Prototype Functionality

In order to show the operation of the LiFi prototypes, an Internet connection will be made through a LIFI luminaire, and the Packet Loss metric will be measured. The link will be adequate at a speed of 1 Mbps to make the frequencies easier to observe on the oscilloscope. The hybrid luminaire used has a field of view of $r = 20cm$; this allows mobility in the device's line of sight without losing the connection (LiFi communication was used for 4 hours for this experiment without setbacks). On the other hand, internet connection speed meters were used to measure downlink and uplink. This fact showed that data downlink and uplink speeds were symmetric (an unusual feature in ISP connections since speeds in the downlink are almost always provided lower than the uplink).

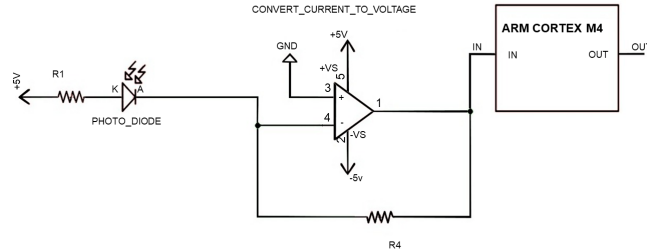


Fig. 7. LiFi receiver circuit.

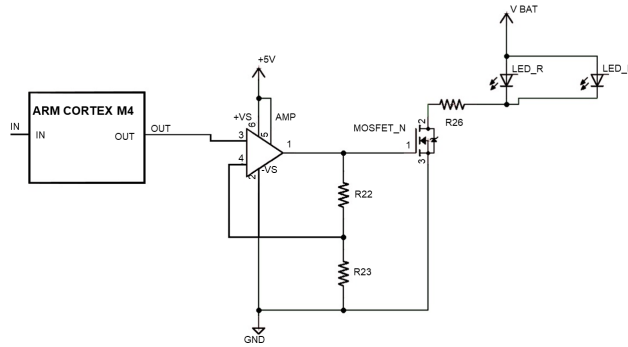


Fig. 8. LiFi transmitter circuit.

The distance between receiver and transmitter was $1.53m$. In these conditions, the noise did not represent a problem. In these LiFi prototypes, noise reduction and speed adjustment are implemented. Another important detail in the measurements is the $1\mu s$ period of the signal, which according to the relationship:

$$T = 1/f, \quad (2)$$

where:

T : period of the signal

f : frequency of the signal

It shows that the frequency of the signal is $1MHz$.

LiFi prototypes implemented in subsequent sections base their operation on the circuits shown in *Figures 7 and 8* [14].

As an additional test of the system, it is proposed to send 10,000 ICMP packets with 64-byte size to a server to the Internet while the connection is also used to explore other

Table 1. ICMP packages' statistics through distance variation in the Li-Fi luminaire.

Distance (m)	Latency (ms)	Packet Loss
1.55	93	0.04%
1.40	121	0.21%
1.25	109	0.24%
1.10	167	0.15%
0.95	112	0.05%
0.80	116	0.05%
0.60	107	0.33%

websites. For this test, the distances between the receiver and the transmitter varied in a $\delta d = 15$ cm. *Table 1* shows the experimental calculation of the *Packet Loss* metric.

4 Performance Comparison (LiFi, WiFi and Ethernet)

This section presents the measurement performance results from *LiFi 1* (3 Mbps), *LiFi 2* (6 Mbps), *WiFi* (144.4 Mbps), and *Ethernet* (1 Gbps) technologies. As is well known, networks generally measure their performance through certain metrics. Below are the used metrics in this document; a summary of each of them is shown next:

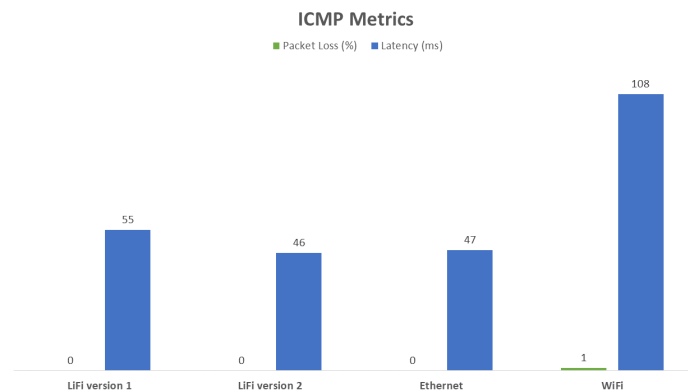
- **Latency:** The time that a packet stays across a network, latency is also related to various delays in packets, transmission queue, propagation, or processes present in a communications network.
- **Throughput:** The sent/received packets by the unit time are linked up with the bandwidth used in a specific unit of time. This parameter is essential to know if a network is in congestion or wasting resources.
- **Jitter:** This metric is the delay variation that emerge in a network, that is, the delay that appears between each packet and is related to variations in traffic load and the number of collisions between packets.
- **Packet Loss:** The packets loss on a certain number of transmissions. This metric encompasses the memory of nodes, congestion, and control policies in a network.

The metrics shown above can be calculated experimentally through network protocols. There are a lot of protocols that can be measured, the most common are those that belong to the layer 4 on the OSI model (TCP, UDP, ICMP). Next, the main characteristics of these protocols are presented:

- **ICMP:** The Internet Control Message Protocol often is used for control and diagnostic purposes, furthermore this protocol is used to find delay and throughput in networking.

Table 2. ICMP metrics.

Technology	Wired	Link Speed (Mbps)	Packet Loss (%)	Latency (ms)
<i>LiFi 1</i>	No	3	0	55
<i>LiFi 2</i>	No	6	0	46
<i>Ethernet</i>	Yes	1000	0	47
<i>WiFi</i>	No	144.4	1	108

**Fig. 9.** ICMP latency and packet loss (LiFi,WiFi, Ethernet).

- **UDP:** The User Datagram Protocol is a protocol that establishes a low-latency link, considered as a connectionless protocol, often used to obtain Jitter in a network.
- **TCP:** The Transmission Control Protocol establishes and maintains a network link. This protocol is impacted by latency, then also, the throughput is impacted by this factor.

Specialized BSD software tools were used to measure each metric in every protocol for all the results shown in this work.

4.1 ICMP Results

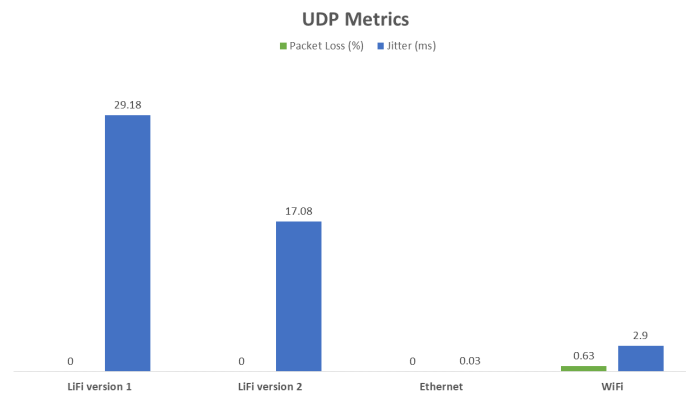
Meanwhile, the round trip was used to get the latency and packet loss, throughput was calculated in terms of the bits transmitted and the latency. The results of developed experiments are shown in *Table 2* and *Figure 9*. The test was done with 100 packets of 32 bytes size through PING software.

4.2 UDP Results

The results obtained for Throughput and Jitter through UDP protocol are shown in *Table 3* and *Figure 10*.

Table 3. UDP metrics (LiFi, WiFi, Ethernet).

Technology	Wired	Link Speed (Mbps)	Packet Loss (%)	Jitter (ms)
<i>LiFi 1</i>	No	3	0	29.18
<i>LiFi 2</i>	No	6	0	17.08
<i>Ethernet</i>	Yes	1000	0	0.03
<i>WiFi</i>	No	144.4	0.63	2.9

**Fig. 10.** UDP jitter and packet loss (LiFi,WiFi, Ethernet).**Table 4.** TCP downlink and uplink throughput (LiFi, WiFi, Ethernet).

Technology	Wired	Link Speed (Mbps)	Download Throughput (Mbps)	Upload Throughput (Mbps)
<i>LiFi 1</i>	No	3	2.50	2.36
<i>LiFi 2</i>	No	6	4.90	4.75
<i>Ethernet</i>	Yes	1000	95.1	94.6
<i>WiFi</i>	No	144.4	21.1	21.0

4.3 TCP Results

These tests were executed to get the Throughput in downlink and uplink through TCP protocol with a window size of 65 KB. Results are shown in *Figure 11*, and *Table 4*.

4.4 Internet Connection Results

These experiments show the link connection reports about bytes received and transmitted, thus allowing to know the data downloaded and uploaded (*Table 5*, *Figure 12*) for a certain period of time (30 minutes).

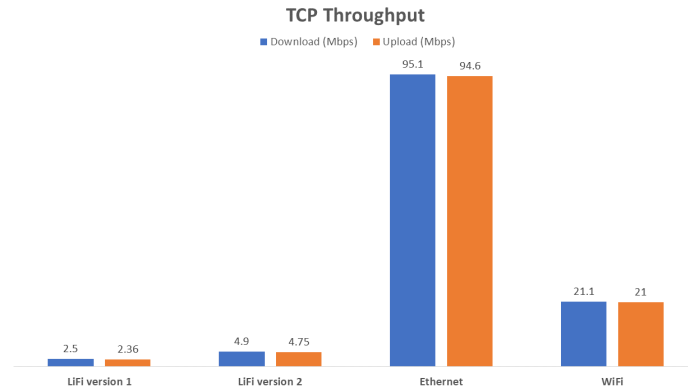


Fig. 11. TCP throughput comparison (LiFi, WiFi, Ethernet).

Table 5. Data transmitted and received (LiFi, WiFi, Ethernet).

Technology	Wired	Link Speed (Mbps)	Download (Bytes)	Upload (Bytes)
<i>LiFi 1</i>	No	3	694,961.18	25,930.75
<i>LiFi 2</i>	No	6	824,726.00	21,483.78
<i>Ethernet</i>	Yes	1000	1,024,507.57	34,504.57
<i>WiFi</i>	No	144.4	951,436.36	63,909.49

5 LiFi Performance through Distance and Speed Variation on LiFi Prototypes

As a final part of this work, the general performance comparison in LiFi prototypes at two different speeds (3 Mbps, 6Mbps) was shown through the methods described in the previous section. These experiments were carried out with the variation of distance between emitter and receiver.

Although distance also affects WiFi and Ethernet technologies, this factor is not considered because its impact is significantly less in short distances (considering 100 meters for Gigabit Ethernet 1000Base-T [13], 100 meters for indoor WiFi, and up to 300 meters outdoor [17]). Such results are demonstrated in Tables 6, 7 and 8, and Figures 13, 14 and 15.

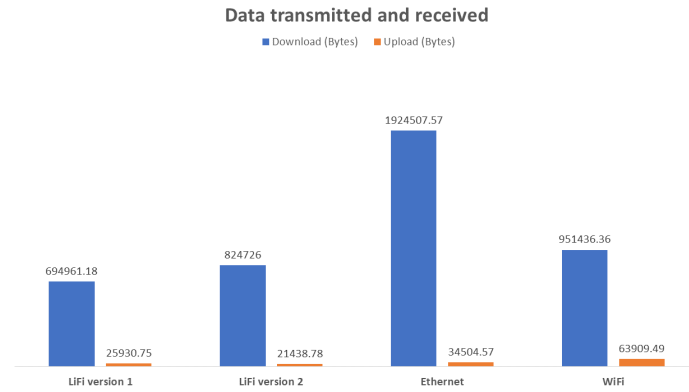


Fig. 12. Data transmitted and received comparison (LiFi,WiFi, Ethernet).

Table 6. LiFi ICMP latency (100 packets, 32 bytes).

Speed Link (Mbps)	Distance (m)	Latency (ms)
3	0.75	49
3	1.0	51
3	1.25	58
3	1.50	64
3	1.75	44
6	0.75	45
6	1.0	52
6	1.25	53
6	1.50	51
6	1.75	46

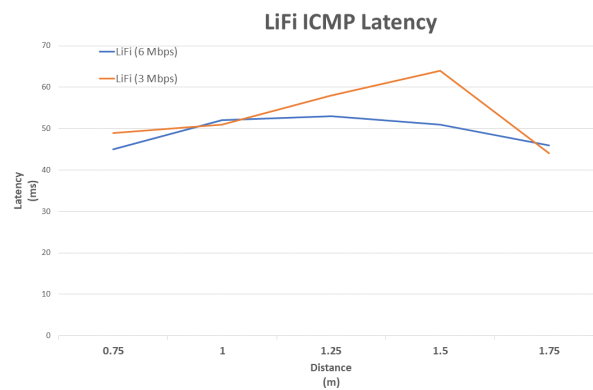
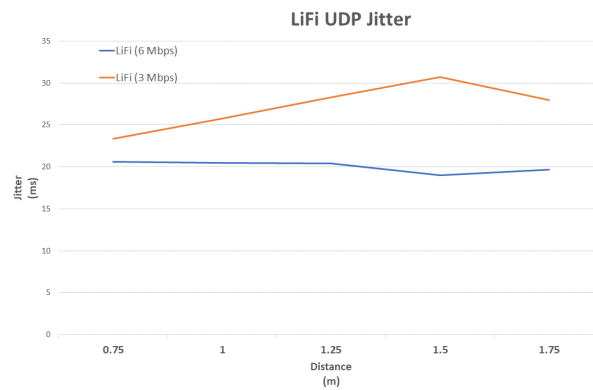


Fig. 13. LiFi ICMP latency at different distances.

Table 7. LiFi UDP Jitter.

Speed Link (Mbps)	Distance (m)	Jitter (ms)
3	0.75	23.38
3	1.0	25.74
3	1.25	28.31
3	1.50	30.69
3	1.75	28.01
6	0.75	20.65
6	1.0	20.48
6	1.25	20.45
6	1.50	19.03
6	1.75	19.69

**Fig. 14.** LiFi UDP Jitter at different distances.

For the last case, it can see in *Figure 15* that the throughput tends to be smaller for long and short distances compared to medium distances. The above is because TCP establishes virtual circuits that require retransmissions, unlike the UDP protocol, which is a connectionless protocol. In addition, for very close distances, the LiFi sensors are easier to align, but sensors could be affected by power saturation; analogously, at very long distances, the sensors may require more energy to amplify the signal, and alignment becomes more complex to find the Line of Sight.

6 Conclusion

The first part of the work was focused on VLC communications and, specifically, on creating LiFi prototypes that can be implemented simply in typical luminaires. The paper presents the electronic blocks for transmission and reception stages, the

Table 8. LiFi TCP (throughput).

Speed Link (Mbps)	Distance (m)	Download Throughput (Mbps)	Upload Throughput (Mbps)
3	0.75	2.50	2.37
3	1.0	2.55	2.41
3	1.25	2.50	2.36
3	1.50	2.55	2.40
3	1.75	1.54	1.41
6	0.75	4.84	4.74
6	1.0	4.89	4.74
6	1.25	4.90	4.75
6	1.50	4.47	4.36
6	1.75	2.1	1.4

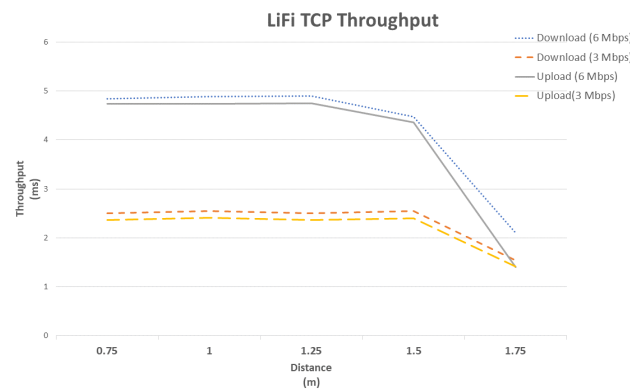


Fig. 15. LiFi TCP throughput at different distances.

mechanical process for LiFi luminaires (Lumux) and LiFi receivers (dongles), and the communications tests between these two devices. It is also shown that the latency times presented in LiFi prototypes are at an acceptable average, and it permits users to surf on the internet in a dedicated link.

The correct implementation of the physical and data-link layers makes it possible to obtain correct communication with the upper layers. In addition to the above, the coverage radii are attached, and the experiment conditions are specified. The manner that humanity communicates is growing accelerated, and it looks for new alternatives to achieve its purpose efficiently. The prototypes presented in this article are an essential starting point for developing LiFi technology and infrastructure that allows an improvement in current communications.

Table 9. QoS metrics.

	Jitter (ms)	Latency (ms)	Packet Loss (%)	Throughput (%)
<i>Excellent / Good</i>	0-20	0-150	0-3	75-100
<i>Medium</i>	21-50	151-400	4-15	25-74
<i>Poor</i>	>50	>400	>15	<25

Another important point of this project was the tests applied to different communication technologies (LiFi, WiFi, Ethernet). In order to compare its performance through certain metrics utilizing specific protocols from the transport layer. After observing the results obtained, we can justify the following premises concerning the proposed experiments and environments:

- Ethernet presented a lower latency and jitter in the ICMP and UDP protocols since cabled technologies have better energy conduction, low loss rate, and no interference. In addition, the PoE in the ICMP packets was 0%. It is important to emphasize that, although the data rates were very similar to the other technologies, Ethernet could support higher data loads through its connection. It is well known that Ethernet is suitable for any protocol of the transport layer, and it was demonstrated in the results of the network performance
- WiFi obtained a Jitter greater than Ethernet but lower than LiFi. Meanwhile, its latency was irregular and therefore obtained the worst values in this test. High latency is a sign of high interference and weak signals, which caused a packet loss of 1%. The data throughput was similar to the other two technologies because the use of the network was similar.
- LiFi presented a latency very close to Ethernet, which indicates that packets within environments that add LiFi present small round trips. With little interference and high data density, the LiFi packet loss was 0%. Meanwhile, the transmitted and received data was very similar to that presented by Ethernet and Wifi. The tests in which LiFi presented the lowest performance were those in which the TCP protocol was used, obtaining a lower data transfer, the previous due to the window sizes, the low bandwidth, and the LiFi network processes.

In general terms, the technologies presented an acceptable performance in Quality of Service (QoS). According to ITU G.114 and TIPHON TR 101 328 standards [6] as is summarized in *Table 9*.

As it can be observed in *Table 10* a denomination EXCELLENT/GOOD was obtained in the three technologies for the case of latency. Jitter metric was categorized as GOOD in all cases, except for LiFi 1 prototype with a MEDIUM score. On the other hand, packet loss was found in a range of EXCELLENT/GOOD range. Although the overall performance was apparently EXCELLENT/GOOD for the throughput test, the test varied by window sizes and packets sent; the experiments showed that although the environments can have a variety of factors that affect a network, in the end, the

Table 10. Summary results.

	Jitter (ms)	Latency (ms)	Packet Loss (%)	Throughput (%)
<i>LiFi 1</i>	29.18	55	0	78.66
<i>LiFi 2</i>	17.08	46	0	79.16
<i>Ethernet</i>	0.03	47	0	94.6
<i>WiFi</i>	2.9	108	1	84

interference in WiFi, the lower LiFi bandwidth and the difficulty for Ethernet cabling are determining factors to find a complement between these technologies.

In the last part of this work, it was showed a comparison of the performance of metrics. These experiments demonstrated that performance was very similar in all cases despite varying distances and speed in LiFi prototypes; this is because of the density of the data found in LiFi transmissions, which allows greater wireless transmission capacity for specific areas.

References

1. Adwani, A., Nagtode, S.: Li-fi: Information transferring through led's. In: Electrical, Electronics, and Optimization Techniques (ICEEOT), International Conference on. pp. 2125–2127. IEEE (2016)
2. Agarwal, C., Kulshrestha, M., Rathore, H., et al.: High speed wireless internet using lifi. In: 2018 IADS International Conference on Computing, Communications & Data Engineering (CCODE). pp. 7–8 (2018)
3. Akyildiz, I. F., Xie, J., Mohanty, S.: A survey of mobility management in next-generation all-ip-based wireless systems. *IEEE Wireless communications*, vol. 11, no. 4, pp. 16–28 (2004)
4. Bao, X., Yu, G., Dai, J., Zhu, X.: Li-fi: Light fidelity-a survey. *Wireless Networks*, vol. 21, no. 6, pp. 1879–1889 (2015)
5. Bober, K., Jungnickel, V., Emmelmann, M., Riegel, M., Tangdiongga, E., Koonen, A., Linnartz, J.-P., Behnke, D., Müller, M., Bök, P.-B., et al.: A flexible system concept for lifi in the internet of things. In: 2020 22nd International Conference on Transparent Optical Networks (ICTON). pp. 1–4. IEEE (2020)
6. Budiman, E., Wicaksono, O.: Measuring quality of service for mobile internet services. In: 2016 2nd International Conference on Science in Information Technology (ICSITech). pp. 300–305. IEEE (2016)
7. Campos, A.: Design and configuration of a system for the bidirectional communication of digital data through the visible light spectrum , in a non - confined medium
8. Campos, A.: Sistema de comunicacioin bidireccional de datos digitales a traves del espectro de luz visible en un medio no confinado.
9. Fath, T., Haas, H.: Performance comparison of mimo techniques for optical wireless communications in indoor environments. *IEEE Transactions on Communications*, vol. 61, no. 2, pp. 733–742 (2012)
10. Forecast, G.: Cisco visual networking index: global mobile data traffic forecast update, 2017–2022. Update, vol. 2017, pp. 2022 (2019)
11. Haas, H., Cogalan, T.: Lifi opportunities and challenges. In: 2019 16th International Symposium on Wireless Communication Systems (ISWCS). pp. 361–366. IEEE (2019)

12. Haas, H., Yin, L., Wang, Y., Chen, C.: What is lifi? *Journal of lightwave technology*, vol. 34, no. 6, pp. 1533–1544 (2015)
13. Hatamian, M., Agazzi, O. E., Creigh, J., Samueli, H., Castellano, A. J., Kruse, D., Madisetti, A., Yousefi, N., Bult, K., Pai, P., et al.: Design considerations for gigabit ethernet 1000base-t twisted pair transceivers. In: *Proceedings of the IEEE 1998 Custom Integrated Circuits Conference (Cat. No. 98CH36143)*. pp. 335–342. IEEE (1998)
14. Hernandez-Oregon, G., Rivero-Angeles, M. E., Chimal-Eguía, J. C., Campos-Fentanes, A., Jimenez-Gallardo, J. G., Estevez-Alva, U. O., Juarez-Gonzalez, O., Rosas-Calderon, P. O., Sandoval-Reyes, S., Menchaca-Mendez, R.: Performance analysis of v2v and v2i lifi communication systems in traffic lights. *Wireless Communications and Mobile Computing*, vol. 2019 (2019)
15. Hutt, D., Snell, K., Belanger, P.: Alexander graham bell’s photophone. *Optics and Photonics News*, vol. 4, no. 6, pp. 20–25 (1993)
16. Kazemi, H., Safari, M., Haas, H.: A wireless backhaul solution using visible light communication for indoor li-fi attocell networks. In: *2017 IEEE International Conference on Communications (ICC)*. pp. 1–7. IEEE (2017)
17. Li, L., Hu, X., Zhang, W.: Design of an arm-based power meter having wifi wireless communication module. In: *2009 4th IEEE Conference on Industrial Electronics and Applications*. pp. 403–407. Ieee (2009)
18. Ramadhani, E., Mahardika, G.: The technology of lifi: A brief introduction. In: *IOP Conference Series: Materials Science and Engineering*. vol. 325, pp. 012013. IOP Publishing (2018)
19. Shao, S., Khreishah, A., Ayyash, M.: Delay analysis of hybrid wifi-lifi system. *arXiv preprint arXiv:1510.00740*, (2015)
20. Soltani, M. D., Wu, X., Safari, M., Haas, H.: Bidirectional user throughput maximization based on feedback reduction in lifi networks. *IEEE Transactions on Communications*, vol. 66, no. 7, pp. 3172–3186 (2018)
21. Sugeng, W., Istiyanto, J. E., Mustofa, K., Ashari, A.: The impact of qos changes towards network performance. *International Journal of Computer Networks and Communications Security*, vol. 3, no. 2, pp. 48–53 (2015)
22. Zeng, Z., Soltani, M. D., Safari, M., Haas, H.: Angle diversity receiver in lifi cellular networks. In: *ICC 2019-2019 IEEE International Conference on Communications (ICC)*. pp. 1–6. IEEE (2019)

Criptografía y mejora a esquema de cifrado de imágenes hipercaótico

Marco T. Ramírez-Torres¹, César A. Guerra-García¹,
María F. Mendoza Martínez¹, Jazmín Rangel Saucedo¹,
Carlos Soubervielle Montalvo²

¹ Universidad Autónoma de San Luis Potosí,
Coordinación Académica Región Altiplano Oeste,
México

² Universidad Autónoma de San Luis Potosí,
Facultad de Ingeniería,
México

{tulio.torres, carlos.soubervielle@uaslp.mx}

Resumen. En los últimos años, debido a la demanda de seguridad los sistemas de cifrado de imágenes basados en sistemas caóticos, han ganado relevancia, debido a su posible aplicación en equipos de cómputo y dispositivos móviles para asegurar la transferencia y almacenamiento de imágenes y videos. En este trabajo, analizamos y probamos un esquema de cifrado de imágenes en escala de grises, basado en un sistema hipercaótico multienroscado. Este esquema consiste en un algoritmo de llave simétrica que utiliza un novedoso Generador de Bits Pseudo-Aleatorios (GBPA), capaz de generar secuencias binarias utilizando los cuatro estados de un sistema hipercaótico multienroscado. Aún cuando esta propuesta se evaluó a través de diversas pruebas de seguridad (análisis de histogramas, entropía, etc.) y el GBPA pasa las pruebas del NIST (National Institute of Standards and Technology), este sistema presenta una debilidad al momento de aplicar un ataque criptoanalítico, conocido como chosen-plain image attack (CPIA). En este trabajo se presenta un criptoanálisis al sistema, donde con dos imágenes arbitrarias es posible recuperar la información original, sin conocer la llave secreta. Además, se propone una mejora para evitar la fuga de información, basada en una función de sustitución de valores utilizando autómatas celulares.

Palabras clave: Criptoanálisis, cifrado imágenes, chosen-plain image attack.

Cryptanalysis and Improvement of a Hyperchaotic Image Encryption Scheme

Abstract. In recent years, due to the demand for security, image encryption systems based on chaotic systems have gained relevance due to their potential application in computers and mobile devices to secure the transfer and storage of images and videos. In this work, we analyze and test a grayscale image encryption scheme based on a multi-threaded hyperchaotic system. This scheme consists of a symmetric key algorithm that uses a novel Pseudo-Random Bit Generator (PRG), capable of generating binary sequences using the four states of

a multi-threaded hyperchaotic system. Although this proposal was evaluated through various security tests (histogram analysis, entropy, etc.) and the PRG passes the NIST (National Institute of Standards and Technology) tests, this system presents a weakness when applying a cryptanalytic attack, known as a chosen-plain image attack (CPIA). This paper presents a cryptanalysis of the system, whereby, with two arbitrary images, it is possible to recover the original information without knowing the secret key. Furthermore, an improvement is proposed to prevent information leakage, based on a value substitution function using cellular automata.

Keywords: Cryptanalysis, image encryption, chosen-plain image attack.

1. Introducción

En la actualidad, el uso generalizado de teléfonos inteligentes y otros tipos de dispositivos ha convertido a las imágenes digitales en uno de los formatos de intercambio de información, más importantes en la vida diaria de las personas.

Asimismo, estas imágenes pueden contener información sensible de diverso tipo, como la seguridad nacional, asuntos legales y/o médicos, entre otros. Según Wu [1], la exposición no planificada de imágenes privadas (fotos) acentúa la importancia de la seguridad de la imagen.

El paradigma del cifrado de cualquier tipo de imágenes (incluidos videos) tiene sus propios requisitos de implementación, principalmente por sus características intrínsecas, por ejemplo, una capacidad de datos masivos y una alta redundancia. A medida que aumenta el ancho de banda en los canales de comunicación, surge un nuevo desafío, un nuevo conjunto de soluciones criptográficas debe mantenerse al día con la gran cantidad de información que debe protegerse. Actualmente, los esquemas de encriptación de imágenes basados en sistemas hipercaóticos han ganado relevancia, siendo una opción que puede brindar seguridad.

En 1989, Matthews publicó el trabajo [2], que se considera una de las primeras propuestas de un esquema de cifrado basado en el caos. En los últimos años, con el deseo de ampliar este enfoque, se han realizado algunas iniciativas y esfuerzos para aplicar diferentes algoritmos de caos a la criptografía [3, 4, 5]. Debido a que existen algunas propiedades que en ambos casos podrían ser similares, por ejemplo, propiedades como la ergodicidad y la sensibilidad de los sistemas caóticos podrían estar relacionadas con las propiedades de confusión y difusión en el cifrado [6, 7]. Entre algunas opciones para el cifrado de imágenes, hay principalmente dos tipos de arquitecturas clásicas consideradas para el cifrado de imágenes: arquitectura de permutación-sustitución y arquitectura de solo permutación.

Fridrich en [8] propuso un cifrado de imagen caótico, que posteriormente se mejoró con otras propuestas como [9, 10, 11, 12], en las que proponen algoritmos enfocados principalmente en nuevas estrategias de permutación y esquemas de difusión. Por ejemplo, en [10] se muestra un algoritmo con permutación caótica con una matriz ergódica. La difusión se lleva a cabo con un sistema dinámico caótico integrado con una operación algebraica S-box, para realizar una sustitución similar a la que se lleva a cabo con las cajas S-box. Y también cuenta con una operación 'XOR plus mod' que se aplica a cada píxel. Asimismo, en [11] los autores propusieron un esquema de cifrado

basado en mapas caóticos bidimensionales. Esta propuesta se basa en la idea de que la función de mapeo baker map bidimensional se amplía para que sea tridimensional y luego se utiliza para componer un esquema de cifrado de imágenes seguro y rápido.

Por otro lado, se ha considerado la arquitectura de solo permutación debido a las ventajas mostradas en propuestas como [13, 14]. En [13] los autores proponen un algoritmo para el cifrado jerárquico de imágenes caóticas, y utilizan una arquitectura VLSI (Very Large Scale Integration). Mientras tanto, en [14] se muestra un algoritmo de esquema de cifrado basado principalmente en las secuencias hipercaóticas mejoradas. Las propuestas antes mencionadas [8-14] tienen un inconveniente en común, son inseguras.

La propuesta presentada en [8], fue analizada por Solak et al. y como mencionan en su trabajo [15], a través del ataque chosen-ciphertext attack (CCA), el algoritmo mostrado en [8] puede ser descifrado.

Por otro lado, las propuestas [9, 10, 11, 12] se centran en la implementación de algoritmos cifrados de arquitectura permutación-sustitución, también se han mostrado inseguras, esto fue expuesto por Li et al. en [16]. Donde se puede revelar la clave equivalente de los esquemas si solo se realiza una ronda de cifrado.

Las propuestas [13, 14], que son algoritmos de solo permutación, han sido infringidas como se muestra en los trabajos [17, 18].

En la misma área de oportunidad de las propuestas de encriptación de imágenes con enfoque caótico, García-Martínez en [19] presentó un esquema de encriptación de imágenes basado en un sistema hipercaótico multienroscado, para generar secuencias pseudoaleatorias. Los autores propusieron un nuevo Generador de Bits Pseudo-Aleatorio (GBPA) capaz de generar secuencias binarias. Para esto implementaron un algoritmo capaz de utilizar los cuatro estados de un sistema hipercaótico multienroscado.

Cabe destacar que esta propuesta fue validada a través de varias pruebas de seguridad, por ejemplo, el análisis de espacios de llave, entropía, análisis de histogramas, entre otros. Dada la originalidad del esquema presentado por García-Martínez et al. y que este esquema puede ser una buena solución para mejorar las tareas de cifrado de imágenes en escala de grises, decidimos estudiar esta propuesta y realizar un criptoanálisis basado en un ataque específico para encontrar una posible vulnerabilidad. Teniendo en cuenta lo anterior, en este trabajo propusimos un criptoanálisis usando el ataque chosen-plain image attack. Con el cual pudimos recuperar las imágenes originales sin conocer la llave secreta, utilizando dos imágenes arbitrarias. Esperamos que, con este trabajo, diseñadores de criptosistemas aprendan sobre los ataques criptoanalíticos y sus consideraciones, y los contemplen en sus propuestas.

Este trabajo se organiza de la siguiente manera, en la sección 2 se describe el esquema de cifrado propuesto por García-Martínez et al. [19]. En el apartado 3 se explica a profundidad nuestra propuesta de criptoanálisis. La sección 4 presenta una mejora propuesta para el esquema de cifrado y una validación. Y finalmente, la sección 5 presenta las conclusiones de este trabajo.

2. Sistema de cifrado

En [19], los autores proponen un esquema de encriptación para imágenes en escala de grises utilizando un nuevo GBPA basado en las series de tiempo obtenidas de los estados del sistema hipercaótico multienroscado, dados por las ecs. 1 con 2 para 2, 4 o 10 enroscados:

$$\dot{X} = AX + B(X), B(X) = \begin{cases} B_1, \text{ if } X \in \mathcal{D}_1; \\ B_2, \text{ if } X \in \mathcal{D}_2; \\ \vdots \\ B_k, \text{ if } X \in \mathcal{D}_k, \end{cases} \quad (1)$$

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1.5 & -1 & -1 & 0 \\ 0 & -1 & 0 & -1 \end{pmatrix}, B = \begin{pmatrix} 0 \\ 0 \\ B_3 \\ B_4 \end{pmatrix}, \quad (2)$$

donde $\mathbf{R}^4 = \cup_{i=1}^k \mathcal{D}_i$, $\mathbf{X} = [x_1, x_2, x_3, x_4]^T \in \mathbf{R}^4$ es el vector de estado. $\mathbf{A} = [a_{ij}] \in \mathbf{R}^{4 \times 4}$, $i, j = 1, 2, 3, 4$ denota una matriz real y $\mathbf{B} = [b_1, b_2, b_3, b_4]^T \in \mathbf{R}^4$ representa un vector real. El proceso para generar las secuencias con el GBPA, consiste en iterar el sistema n veces para obtener una secuencia $\mathbf{X}$ después de 1000 iteraciones del estado transitorio. Cada conjunto de condiciones iniciales $\mathbf{X}_{0p}$ se puede considerar como una llave para el cifrado, donde $p \in \mathbf{Z}^+$, da como resultado p series de tiempo diferentes. El GBPA está definido por la ecuación 3:

$$k_i = \left\lfloor \sum_{j=1}^4 x_j(i) \cdot 10^{14} \right\rfloor \bmod 256. \quad (3)$$

En este caso, las secuencias pseudoaleatorias producidas por GBPA se denominan k_i , donde $k_i \in \{0, 1, 2, \dots, 255\}$, $i = 1, \dots, n$, y n es el número total de píxeles en la imagen de entrada. La operación $\lfloor \cdot \rfloor$ implica la función floor, una función de redondeo de enteros. El factor 10^{14} permite escalar la suma, porque el simulador numérico trabaja con valores de punto flotante. Finalmente, calculando la operación $\bmod 256$, cada valor de k es un entero positivo de 8 bits.

Una vez explicada la generación de las secuencias pseudoaleatorias k , procederemos a explicar las operaciones de cifrado. Las imágenes son cifradas píxel a píxel utilizando las siguientes ecuaciones:

$$\begin{cases} C_1 = P_1 \oplus k_1 \oplus IV \\ C_i = P_i \oplus k_i \oplus C_{i-1} \end{cases} \quad (4)$$

donde C y P representan los píxeles cifrados y en claro respectivamente, con $i = 2 \dots n$. IV representa un vector inicial de 8 bits, k es una secuencia aleatoria de 8 bits obtenida con el GBPA. Y por último, el símbolo $\oplus$ representa la operación XOR. Para obtener el primer píxel cifrado C_1 , se calcula una operación XOR entre el coeficiente del píxel P_1 , la secuencia pseudoaleatoria k_1 y el vector inicial IV . Los demás píxeles cifrados C_i , se

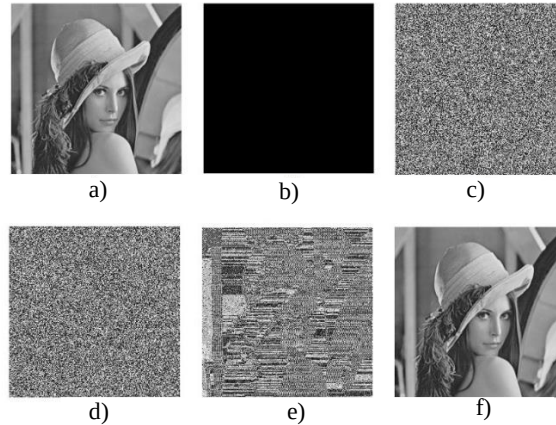


Fig. 1. Chosen-plain image attack. a) Imagen de Lena, b) imagen solida negra, c) imagen de Lena encriptada, d) máscara I_M , e) imagen recuperada con la operación XOR, f) imagen recuperada usando la ec. (3).

obtienen realizando una operación XOR entre el coeficiente del pixel P_i , una nueva secuencia aleatoria k_i y el pixel cifrado previo C_{i-1} . Este proceso se repite hasta cifrar todos los pixeles de la imagen. Las condiciones iniciales del GBPA funcionan como llave secreta.

Como podemos observar el esquema cifra directamente los coeficientes de los pixeles, sin utilizar previamente una operación de sustitución, esto permite a los atacantes introducir datos de manera arbitraria.

3. Criptoanálisis

Para definir lo que es una ruptura o debilidad de un esquema de cifrado, es conveniente mencionar antes, que para esto las reglas de la criptografía académica se relajan considerablemente. Schneier en [20] define que una ruptura en un sistema cifrado es cuando se encuentra una debilidad que se puede explotar con menos complejidad o trabajo que con fuerza-bruta. Recordemos que fuerza-bruta se refiere a los ataques donde se trata de adivinar la llave de cifrado, probando todas las posibles claves para cifrar o descifrar la información. Por lo tanto, se considera que se encontró una debilidad aún cuando en el proceso se requiera una gran cantidad de cifrados o descifrados. La limitante aquí es que sea una cantidad menor que hacerlo por fuerza bruta. Por lo que para demostrar que existe una debilidad, se puede requerir de grandes cantidades de texto plano, texto plano conocido o texto plano elegido, que pueden ser consideradas poco realistas. Es por eso por lo que mencionamos que las reglas de la criptografía se relajan en estos estudios, el propósito se vuelve solo el determinar si el sistema es seguro o no.

En el ataque chosen-plain image attack, el atacante es capaz de seleccionar las imágenes en claro y obtener sus respectivas versiones cifradas [21, 22]. Sin embargo, no posee la llave secreta. Para ilustrar este ataque y todo su proceso, utilizaremos la fig. 1. El ataque comienza seleccionando las dos imágenes en claro, en este caso la imagen de Lena, fig. 1a), y una imagen solida negra, fig. 1b). Ciframos ambas imágenes con el algoritmo de García-Martínez, con las mismas condiciones iniciales. Las imágenes que se obtienen son, fig. 1c), Lena cifrada, y la fig. 1d), la imagen solida cifrada, la cual

llamaremos Mascara I_M . Para recuperar la imagen de Lena sin conocer las condiciones iniciales basta con calcular una operación XOR entre ambas imágenes cifradas, píxel a píxel. El resultado lo podemos ver en la fig. 1e), que como se puede ver esta imagen revela patrones muy significativos de la fig. 1a). Aún así el proceso de recuperación se puede mejorar, y sin conocer las condiciones iniciales la imagen se puede recuperar al 100%, como se puede ver en la fig. 1f). El proceso para obtener esta imagen será explicado a detalle más adelante.

Para ilustrar la debilidad observemos la ec. (5), esta ecuación representa el paso del ataque donde se hace la operación XOR a nivel de bit, entre píxeles de ambas imágenes:

$$C_{L1} \oplus C_{M1} = (P_{L1} \oplus k_1 \oplus IV) \oplus (0 \oplus k_1 \oplus IV) = P_{L1}, \quad (5)$$

donde C_{L1} representa el primer píxel de la imagen de Lena cifrada, C_{M1} representa el primer píxel de la Mascara I_M . Y por último, P_{L1} representa el primer píxel de la imagen de Lena. Como se puede ver, al momento de calcular la operación XOR entre ambos píxeles bit a bit, los elementos de la ecuación de cifrado k_1 y IV se reducen aplicando álgebra booleana (la operación XOR de elementos idénticos es igual a cero), y el resultado es el coeficiente de píxel P_{L1} . Para recuperar los demás píxeles se utiliza (6), esta versión es la usada para recuperar la fig. 1f). Para mejorar su desempeño en la recuperación, agregamos una operación XOR entre los píxeles cifrados previos, adicional a la de píxel a píxel, como se muestra a continuación:

$$(C_{Li} \oplus C_{Mi}) \oplus (C_{Li-1} \oplus C_{Mi-1}) = (P_{Li} \oplus k_i \oplus C_{Li-1}) \oplus (0 \oplus k_i \oplus C_{Mi-1}) \oplus (C_{Li-1} \oplus C_{Mi-1}) = P_{Li}, \quad (6)$$

donde C_{Li} y C_{Mi} representan los píxeles cifrados en la imagen de Lena cifrada y de la máscara, posteriores al primero. Los valores C_{Li-1} y C_{Mi-1} , representan el píxel previo cifrado, que es un elemento conocido o disponible para el atacante en la imagen cifrada, que también es usado en las ecuaciones de cifrado, por lo tanto, se puede anular. Como se puede ver, es posible recuperar la imagen original de Lena, sin conocer la llave secreta, solo se requieren dos imágenes.

Este ataque funciona en este esquema, posiblemente por varias razones, pero este trabajo se enfoca en la falta de una función de sustitución. Esto para evitar que el adversario introduzca valores en las ecuaciones a su conveniencia.

4. Mejora propuesta para el esquema

Para mejorar el esquema propuesto por García-Martínez, se sugiere agregar una función de preprocesamiento, capaz de sustituir el texto plano antes de ser cifrado. Sin importar que el texto en claro sea altamente redundante, esta función debe intercambiarlo por diferentes valores de la codificación con igual probabilidad. Si se desea utilizar una caja de sustitución, el esquema debería modificarse aún más, ya que simplemente intercambiar el texto con una S-box antes de cifrar, crearía patrones de la imagen original.

La función de preprocesamiento utilizada en esta mejora se encuentra en la patente [23], que fue presentada en el trabajo [24]. Esta función se basa en la sincronización de autómatas celulares usando la regla local 90. Gracias a su retroalimentación, la función

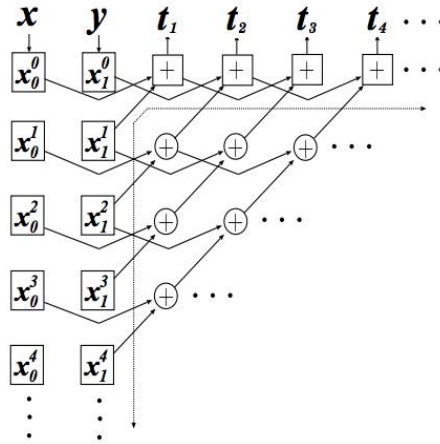


Fig. 2. Generador de secuencias pseudoaleatorias t de n bits, basado en la evolución hacia atrás de la regla 90.

de preprocesamiento puede intercambiar valores idénticos por números diferentes en cada iteración.

Para explicar su funcionamiento, en la fig. 2 se ilustra el generador pseudoaleatorio y el proceso de evolución hacia atrás usando la regla local 90. Se utiliza un vector x y un vector y de n bits y $n + 1$ bits, respectivamente. Se evoluciona hacia atrás, utilizando la operación XOR como indican las flechas, hasta generar el vector t . Esta función es llamada h .

Para aplicar esta operación como función de preprocesamiento, en la fig. 3, podemos observar a la función h como un bloque, y en el lugar del vector x , entra el coeficiente del pixel p , y en el lugar del vector y , un nuevo vector llamado z . Esto para diferenciar su funcionamiento como generador de números pseudoaleatorios y como función de preprocesamiento. El vector de salida es llamado $\hat{p}$, para señalar que es la versión preprocesada de p .

Como se puede ver existe una retroalimentación que actualiza el vector z , este nuevo vector se calcula con el vector de salida $\hat{p}$, concatenando el bit menos significativo del vector z anterior, en la posición de bit más significativo.

La diferencia entre esta función y una caja de sustitución la podemos observar en la fig. 4, donde se intercambian los píxeles de la imagen de Lena, primero con una S-Box y en la segunda imagen con nuestra función.

Se puede observar en la fig. 4, que en el caso de la caja de sustitución se crean patrones, debido a que no existe ningún tipo de dinámica, los coeficientes de la imagen son sustituidos siempre por los mismos valores de la S-box.

En la fig. 5 se muestra el histograma de la imagen de Lena y su versión preprocesada. Como se puede ver el histograma se vuelve uniforme, ocultando la redundancia de la imagen original y previniendo un ataque estadístico.

Al agregar la función de preprocesamiento al algoritmo de García-Martínez et al., antes del cifrado, mejora su desempeño ante el ataque CPIA. Lo anterior se puede confirmar en la fig. 6, donde nuevamente se escogen las mismas dos imágenes de forma arbitraria, y se cifran con las mismas condiciones iniciales.

Como podemos observar el atacante no puede introducir el valor de 0 de forma arbitraria en las ecuaciones de cifrado, porque éste se intercambia. Esto evita que capture

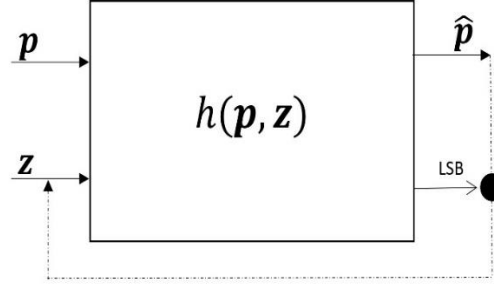


Fig. 3. Función de preprocesamiento. p es el coeficiente del pixel, z un vector aleatorio. A la salida se obtiene el vector preprocesado $\hat{p}$.

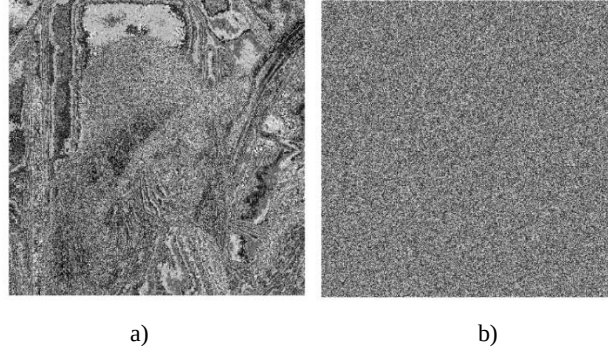


Fig. 4. Imagen de Lena sustituida, a) imagen de Lena sustituida con una S-box del sistema AES (Advanced Encryption Standar), b) imagen de Lena despues de ser preprocesada con la función h .

información en la Máscara I_M . Además, la información no está cifrada de manera directa, el preprocesamiento intercambia los valores en ambas imágenes.

Por lo tanto, ante las condiciones que se dan en este ataque, no queda expuesta la información aún cuando se cifren los píxeles con las mismas secuencias k_i y IV y que con el ataque se anulen. Para ilustrar esto podemos analizar la ecuación (7):

$$C_{L1} \oplus C_{M1} = (\hat{P}_{L1} \oplus k_1 \oplus IV) \oplus (\hat{P}_{IM1} \oplus k_1 \oplus IV) = \hat{P}_{L1} \oplus \hat{P}_{IM1}. \quad (7)$$

5. Conclusiones

Los sistemas caóticos pueden aportar elementos en el diseño de sistemas de cifrado. Sin embargo, su análisis debe ser profundo e interpretar de manera general las pruebas que se aplican a los esquemas de cifrado. Debido a que existe una gran variedad de

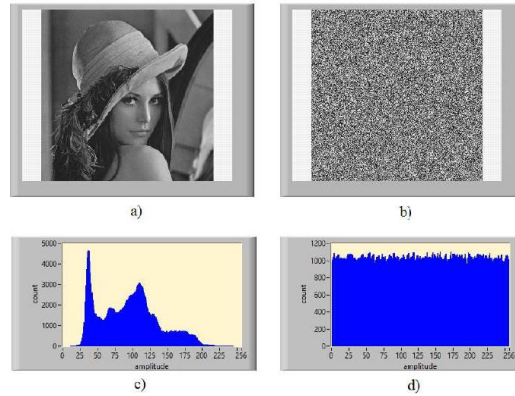


Fig. 5. Análisis de histogramas. a) Imagen de Lena, b) imagen preprocesada de Lena, c) histograma de a) y d) histograma de b).

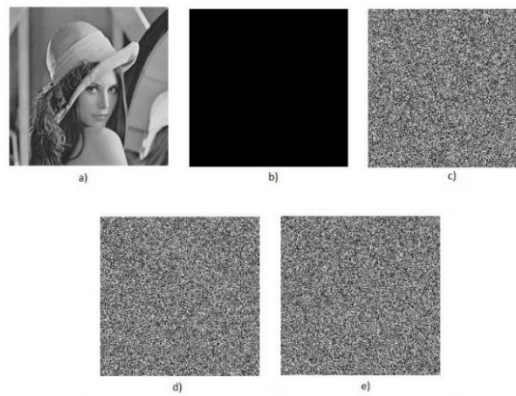


Fig. 6. Chosen-plain image attack. a) Imagen de Lena, b) imagen solida negra, c) imagen de Lena encriptada, d) máscara, e) imagen recuperada con la operación XOR.

ataques con diversos supuestos que permiten concluir cuando un esquema se puede considerar seguro acorde a alguna definición; los diseñadores deben incluir más pruebas en sus propuestas y ser muy explícitos en la forma en que se aplicaron.

En este trabajo se mostró como se puede llevar a cabo una prueba de criptoanálisis. En algún momento se puede pensar que el adversario tiene muchas ventajas, como el hecho de que se utilizan las mismas condiciones iniciales para cifrar ambas imágenes. Estas pruebas con estos mismos supuestos son aprobadas por los esquemas de cifrado que son referente, como lo es el sistema AES [25]. Como se mencionó antes, las reglas de criptografía se relajan cuando se aplica un criptoanálisis, y el objetivo simplemente es definir si el algoritmo analizado es seguro o no.

El cifrado de imágenes sigue presentado áreas de oportunidad y desarrollo, porque cuando un sistema logra solventar las problemáticas de seguridad incrementa su latencia. Por lo que la búsqueda de nuevos métodos se mantiene aún en auge y los sistemas caóticos pueden brindar soluciones a este problema.

Referencias

1. Wu, Y., Noonan, J.P., Yang, G.: Image Encryption Using the Two-Dimensional Logistic Chaotic Map. *Journal of Electronic Imaging*, 21(1), 013014 (2012) doi: 10.1117/1.JEI.21.1.013014.
2. Matthews, R.: On the Derivation of a “Chaotic” Encryption Algorithm. *Cryptologia*, 8(1), pp. 29–41 (1984) doi: 10.5555/67071.67073.
3. Li, Z., Peng, C., Tan, W.: A Novel Chaos-Based Image Encryption Scheme by Using Randomly DNA Encode and Plaintext Related Permutation. *Applied Sciences*, 10(21), 7469 (2020) doi: 10.3390/app10217469.
4. Liao, T.L., Lin, H.R., Wan, P.Y.: Improved Attribute-Based Encryption Using Chaos Synchronization and its Application to MQTT Security. *Applied Sciences*, 9(20), 4454 (2019) doi: 10.3390/app9204454.
5. Datcu, O., Macovei, C., Hobincu, R.: Chaos Based Cryptographic Pseudo-Random Number Generator Template with Dynamic State Change. *Applied Sciences*, 10(2), 451 (2020) doi: 10.451.10.3390/app10020451.
6. Alvarez, G., Montoya, P., Pastor, G.: Chaotic Cryptosystems. In: *Proceedings IEEE 33rd Annual 1999 International Carnahan Conference on Security Technology*, pp. 332–338 (1999) doi: 10.1109/CCST.1999.797933.
7. Kocarev, L., Jakimoski, G., Stojanovski, T.: From Chaotic Maps to Encryption Schemes. In: *ISCAS'98. Proceedings of the 1998 IEEE International Symposium on Circuits and Systems*, 4, pp. 514–517 (1998)
8. Fridrich, J.: Symmetric Ciphers Based on Two-Dimensional Chaotic Maps. *International Journal of Bifurcation and chaos*, 8(06), pp. 1259–1284 (1998) doi: 10.1142/S021812749800098X.
9. Chen, G., Mao, Y., Chui, C.K.: A Symmetric Image Encryption Scheme Based on 3D Chaotic Cat Maps. *Chaos, Solitons & Fractals*, 21(3), pp. 749–761 (2004) doi: 10.1016/j.chaos.2003.12.022.
10. He, X., Zhu, Q., Gu, P.: A New Chaos-Based Encryption Method for Color Image. In: *International Conference on Rough Sets and Knowledge Technology Springer, Berlin, Heidelberg*, pp. 671–678 (2006) doi: 10.1007/11795131_98.
11. Mao, Y., Chen, G., Lian, S.: A Novel Fast Image Encryption Scheme Based on 3D Chaotic Baker Maps. *International Journal of Bifurcation and chaos*, 14(10), pp. 3613–3624 (2004) doi: 10.1142/S021812740401151X.
12. Shen, J., Jin, X., Zhou, C.: A Color Image Encryption Algorithm Based on Magic Cube Transformation and Modular Arithmetic Operation. In: *Pacific-Rim Conference on Multimedia Springer*, pp. 270–280 (2005) doi: 10.1007/11582267_24.
13. Yen, J.C., Guo, J.L.: Efficient Hierarchical Chaotic Image Encryption Algorithm and its VLSI Realisation. In: *IEE Proceedings-Vision, Image and Signal Processing*, 147(2), pp.167–175 (2000) doi: 10.1049/ip-vis:20000208.
14. Zhu, C.: A Novel Image Encryption Scheme Based on Improved Hyperchaotic Sequences. *Optics communications*, 285(1), pp. 29–37 (2012) doi: 10.1016/j.optcom.2011.08.079.
15. Solak, E., Cokal, C., Yildiz, O.T.: Cryptanalysis of Fridrich's Chaotic Image Encryption. *International Journal of Bifurcation and Chaos*, 20(05), pp. 1405–1413 (2010) doi: 10.1142/S0218127410026563.
16. Li, C., Chen, G.: On the Security of a Class of Image Encryption Schemes. In: *2008 IEEE International Symposium on Circuits and Systems IEEE* pp. 3290–3293 (2008) doi: 10.1109/ISCAS.2008.4542161.
17. Li, C.: Cracking a Hierarchical Chaotic Image Encryption Algorithm Based on Permutation. *Signal Processing*, 118, pp. 203–210 (2016) doi: 10.1016/j.sigpro.2015.07.008.

18. Li, C., Liu, Y., Xie, T.: Breaking a Novel Image Encryption Scheme Based on Improved Hyperchaotic Sequences. *Nonlinear Dynamics*, 73(3), pp. 2083–2089 (2013) doi: 10.1007/s11071-013-0924-6.
19. García-Martínez, M., Ontañón-García, L.J., Campos-Cantón, E.: Hyperchaotic Encryption Based on Multi-Scroll Piecewise Linear Systems. *Applied Mathematics and Computation*, 270, pp. 413–424 (2015) doi: 10.1016/j.amc.2015.08.037.
20. Schneier, B.: A Self-Study Course in Block-Cipher Cryptanalysis. *Cryptologia*, 24(1), pp. 18–33 (2000) doi: 10.1080/0161-110091888754.
21. Munir, N., Khan, M., Shah, T.: Cryptanalysis of Nonlinear Confusion Component Based Encryption Algorithm. *Integration*, 79, pp. 41–47 (2021) doi: 10.1016/j.vlsi.2021.03.004.
22. Li, S., Li, C., Chen, G.: Cryptanalysis of the RCES/RSES Image Encryption Scheme. *Journal of Systems and Software*, 81(7), pp. 1130–1143 (2008) doi: 10.1016/j.jss.2007.07.037.
23. Mejía Carlos, M., Murguía Ibarra, J.S., Ramírez Torres, M.T.: Sistema de cifrado de datos digitales IMPI (MX2015008665A) (2015)
24. Ramírez-Torres, M.T., Murguía, J.S., Mejía-Carlos, M.: FPGA Implementation of a Reconfigurable Image Encryption System. In: 2014 International Conference on ReConFigurable Computing and FPGAs (ReConFig14) IEEE, pp. 1–4 (2014) doi: 10.1109/ReConFig.2014.7032524.
25. Heron, S.: Advanced Encryption Standard (AES). *Network Security*, 2009(12), pp. 8–12 (2009) doi: 10.1016/S1353-4858(10)70006-4.

Fusión de sensores mediante filtro de Kalman extendido para seguimiento de trayectorias de un robot móvil tipo rover

Jheison Duvier Diaz-Ortega¹, Ponciano Jorge Escamilla-Ambrosio²,
Octavio Gutiérrez-Frías¹, Alberto Luviano-Juárez¹,
José Alejandro Aguirre-Anaya¹

¹ Instituto Politécnico Nacional,
Unidad Profesional Interdisciplinaria en Ingeniería y Tecnologías Avanzadas,
México

² Instituto Politécnico Nacional,
Centro de Investigación en Computación,
México

{jheisondz, alberto.luviano,
alejandroaguirre1809}@gmail.com, ogutierrezf@ipn.mx,
pescamilla@cic.ipn.mx

Resumen. En este trabajo se presenta la propuesta de un sistema de fusión de sensores aplicado a un robot móvil tipo rover con la finalidad de utilizarse en tareas de seguimiento de trayectorias. El sistema involucra la ley de control, el modelo de transición de estados y tres modelos de medición de los sensores de posición y orientación. Por otro lado, la fusión de sensores se realiza entre los dispositivos que miden la posición en las coordenadas (x, y), y un sensor que mide la orientación, además se presenta el modelo cinemático en dos dimensiones del robot móvil. Finalmente, se muestran los resultados obtenidos mediante simulación numérica a través de los paquetes de Matlab-Simulink, mediante el seguimiento de dos trayectorias deseadas por parte del robot; una en forma circular y otra en forma de lemniscata, observando las ventajas que presenta el empleo de la fusión de sensores, tales como el filtrado del ruido de medición y la disminución del error de posición del robot.

Palabras clave: Fusión de sensores, filtro de Kalman extendido, robot rover, seguimiento de trayectorias, lemniscata.

Sensor Fusion Using an Extended Kalman Filter for Trajectory Tracking of a Rover-Type Mobile Robot

Abstract. This paper presents a proposed sensor fusion system applied to a rover-type mobile robot for path tracking tasks. The system involves the control law, the state transition model, and three measurement models for the position and orientation sensors. Sensor fusion is performed between devices that measure position in the (x, y) coordinates and a sensor that measures orientation. A two-dimensional kinematic model of the mobile robot is also presented. Finally, the

results obtained through numerical simulation using Matlab-Simulink packages are presented, following two desired trajectories by the robot: one circular and the other lemniscate. The advantages of using sensor fusion are observed, such as filtering measurement noise and reducing the robot's position error.

Keywords: Sensor fusion, extended Kalman filter, rover robot, path tracking, lemniscate.

1. Introducción

En los últimos años se ha incrementado el interés por el desarrollo de robots móviles que sean capaces de desplazarse con la mínima o nula intervención del ser humano, sobre terreno poco estructurado con la finalidad de realizar tareas que sean repetitivas, de asistencia o peligrosas.

Dentro de las aplicaciones destacadas de los robots móviles existen las tareas de asistencia al humano, como el apoyo en los quehaceres domésticos o las tareas de guía en un museo, las tareas como búsqueda y rescate en lugares peligrosos, carros autónomos y la exploración planetaria [1, 2].

Para realizar dichas tareas, se requiere de un algoritmo de seguimiento de trayectorias el cual, debe considerar la incertidumbre inherente a los dispositivos de percepción del entorno y la necesidad de explorar ambientes desconocidos de los que no se tiene a priori, mapas confiables y precisos [3]. Con el fin de tomar en cuenta estos factores, la técnica de fusión de datos de sensores es una alternativa, debido a que es el proceso de combinar piezas incompletas e imperfectas de información de sensores mutuamente complementarios, de tal manera que se logra una mejor comprensión de un fenómeno subyacente del mundo real; por lo general, los resultados de fusión exceden lo que se puede producir a partir de la salida de un solo sensor en precisión, confiabilidad o costo. La recopilación, el registro y la alineación adecuados, el filtrado estocástico, el análisis lógico, la integración espaciotemporal, la explotación de redundancias, la evaluación cuantitativa y la visualización adecuada son parte de las técnicas de fusión de sensores [4]. Así entonces, la fusión de sensores es una técnica importante para que múltiples sistemas de sensores realicen detección y monitoreo, y se usa ampliamente en muchos campos, como sistemas inteligentes, sistemas de seguridad o vehículos autónomos [5].

En la literatura, existen algunos trabajos que emplean la técnica de fusión de sensores y que son cercanos a la presente propuesta, como serían: [6] presenta una estrategia de fusión de sensores aplicada a plataformas de tipo UAV (Unmanned Aerial Vehicles), donde se evalúa el rendimiento del sistema en situaciones reales sin considerar las simulaciones. En [7], se muestra la fusión de las lecturas de una cámara y los sensores LIDAR en vehículos autónomos con un método de optimización de redes neuronales llamado "poda knowledge-based pruning". Por otro lado, en [9] se desarrolla un sistema de localización autónomo capaz de entregar estimaciones de posición mucho más precisas que las entregadas por un sistema basado exclusivamente en odómetros.

Si bien el campo de aplicación para los robots rover es extenso, una cantidad importante de los robots desplegados en tareas autónomas dependen de algoritmos basados en sistemas de visión por computadora o sistemas de posicionamiento global (GPS) [10-13]. Este tipo de planteamientos, presentan inconvenientes cuando algún

Tabla 1. Principales características del rover Shrimp.

Características	Descripción
Dimensiones	50.5 * 75 * 60 (cm)
Peso	9.5 (kg)
Sistema de suspensión	Pasivo
Número de ruedas	6
Configuración geométrica de las ruedas	Dos ruedas en cada lado unidas por eslabones en forma de paralelogramo, una rueda adelante móvil unida al chasis mediante un mecanismo de cuatro barras y otra fija, las ruedas se distribuyen de tal manera que la carga sea equitativa.
Sistema de tracción	Cada rueda está actuada.
Sistema de dirección	Dirección independiente en las ruedas frontal y posterior, requiere de 2 actuadores. Capacidad de giro sobre su propio eje y radios de giro pequeños sin deslizamiento de las ruedas.
Capacidad para superar obstáculos	Puede superar obstáculos de alturas iguales al doble del diámetro de sus ruedas.

sistema falla por lo que se requiere de algoritmos de navegación capaces de integrar todos los sensores sobre el robot para la estimación, lectura o adquisición de las variables que le permiten percibir su entorno; adicionalmente, dichos sistemas deben ser lo suficientemente robustos ante eventos que afecten los sensores, como las condiciones de temperatura, iluminación, ruido electromagnético o cambios climáticos.

Con lo anteriormente expuesto, en el presente trabajo, se propone un sistema que utiliza las técnicas de fusión de sensores, de tal manera que involucra la ley de control, el modelo de transición de estados y tres modelos de medición de los sensores de posición y orientación, que son utilizadas para solucionar el problema de seguimiento de trayectorias de un robot móvil tipo rover, para dos diferentes trayectorias como son: la forma circular y la forma de lemniscata. Finalmente, se muestran los resultados obtenidos mediante simulación numérica, obteniendo un filtrado del ruido de medición y la disminución del error de posición del robot.

La estructura de este trabajo se presenta de la siguiente manera: La Sección 2 presenta el sistema robótico tipo rover, el modelado matemático y la estrategia de control utilizada. La Sección 3 describe la técnica de fusión de sensores empleada. En la Sección 4, se muestran los resultados de simulación numérica obtenidos. Y finalmente la última sección se dedica a las conclusiones del trabajo.

2. Sistema robótico tipo rover

El sistema robótico móvil tipo rover empleado como referencia para aplicar la técnica de fusión de sensores se muestra en la fig. 1, fue desarrollado en la Sección de Estudios de Posgrado e Investigación de UPIITA-IPN, el cual tomó como base el robot móvil tipo Shrimp del Instituto Federal Suizo de Tecnología (EPFL) [14]. Las principales características se presentan en la Tabla 1.



Fig. 1. Robot rover Shrimp.

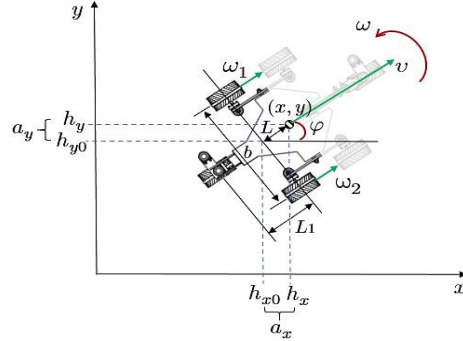


Fig. 2. Diagrama para el modelado cinemático del robot rover Shrimp.

2.1. Modelo cinemático

El modelo cinemático del sistema robótico propuesto se obtiene a partir de una vista superior como se muestra en la fig. 2 considerando un movimiento sobre el plano sin deslizamientos en las ruedas.

La simetría física del robot permite obtener el modelo para la mitad trasera del rover (sección destacada en la fig. 2) y duplicarla para la mitad frontal. El punto de control, es decir, el punto que debe seguir las trayectorias deseadas se sitúa en el centro de gravedad del rover, señalado con el punto de coordenadas (x, y) .

El modelo cinemático del rover Shrimp obtenido en función de las velocidades angulares de las ruedas laterales ω_1 y ω_2 que representan las entradas de control del robot, escrito en representación matricial de la forma $Ax + Bu$, se presenta en la ecuación (1), donde la matriz A que depende de los estados del sistema es nula, pues la cinemática del robot solo depende del vector de entradas de control. Así entonces el modelo presentado solo tiene la matriz B y el vector $u = [\omega_1 \ \omega_2]^T$:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\varphi} \end{bmatrix} = \begin{bmatrix} \frac{R \cos \varphi}{2} + \frac{LR \sin \varphi}{b} & \frac{R \cos \varphi}{2} - \frac{LR \sin \varphi}{b} \\ \frac{R \sin \varphi}{2} - \frac{LR \cos \varphi}{b} & \frac{R \sin \varphi}{2} + \frac{LR \cos \varphi}{b} \\ -\frac{R}{b} & \frac{R}{b} \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}. \quad (1)$$

donde:

- x : Posición del rover en el eje x
- y : Posición del rover en el eje y
- φ : Ángulo de orientación del rover respecto al eje x
- $\dot{x}$: Velocidad del rover en el eje x

- $\dot{y}$: Velocidad del rover en el eje y
- $\dot{\varphi}$: Velocidad de rotación u orientación del rover
- L : Distancia perpendicular desde la mitad del eje al centro de gravedad
- L_1 : Distancia perpendicular desde la mitad del eje a la rueda de dirección
- b : Vía del rover o distancia entre las ruedas laterales
- R : Radio de las ruedas

2.2. Estrategia de control

La estrategia de control utilizada es un controlador PI (Proporcional Integral) el cual se considera suficiente pues se requiere controlar la posición del robot a partir de su modelo cinemático, lo cual incluso es posible únicamente con el control proporcional. El control PI se describe como $\mathbf{u} = \mathbf{K}_p \mathbf{e} + \mathbf{K}_i \int_0^t \mathbf{e}(\tau) d\tau$, donde $\mathbf{e}(\tau)$ es el error entre la señal de salida deseada y la señal de salida real. Así entonces la ley de control planteada tiene la siguiente forma:

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \cos \varphi & -L \sin \varphi \\ \sin \varphi & L \cos \varphi \end{bmatrix}^{-1} \begin{bmatrix} K_{px}(x^*(t) - x(t)) + K_{ix} \int_0^t (x^*(\tau) - x(\tau)) d\tau \\ K_{py}(y^*(t) - y(t)) + K_{iy} \int_0^t (y^*(\tau) - y(\tau)) d\tau \end{bmatrix}, \quad (2)$$

donde K_{px}, K_{ix}, K_{py} y K_{iy} son las constantes para sintonizar el controlador proporcional integral en cada coordenada, mientras que $x^*(t)$ y $y^*(t)$ son las coordenadas (x, y) de la trayectoria deseada.

Por otra parte, el control PI entrega los valores de la velocidad traslacional y la velocidad rotacional del rover, representados por el vector $\mathbf{u} = [u_1 \ u_2]^T$, sin embargo, la cinemática del rover Shrimp depende únicamente de la variación entre las velocidades angulares de las ruedas laterales ω_1 y ω_2 como se muestra en la fig. 2. Así entonces, las velocidades lineal y angular dadas por el controlador se obtienen a través de la siguiente expresión:

$$\begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \begin{bmatrix} \frac{R}{2} & \frac{R}{2} \\ \frac{R}{b} & \frac{R}{b} \end{bmatrix}^{-1} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad (3)$$

en donde u_1 y u_2 son las salidas del controlador.

3. Fusión de sensores

En esta sección se presenta la fusión de sensores utilizando el filtro de Kalman extendido, útil debido a la no linealidad presentada por el modelo cinemático del rover. El esquema general de la propuesta se presenta en la fig. 3; donde las variables

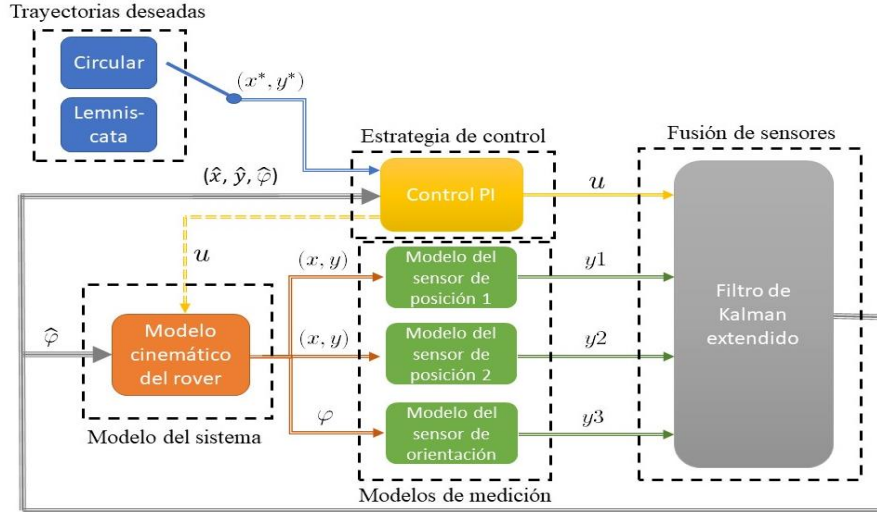


Fig. 3. Esquema general de fusión de sensores.

y_1, y_2, y_3 representan las mediciones de cada sensor; y_1 y y_2 entregan cada una la posición del rover en (x, y) ; mientras que y_3 representa el ángulo de orientación del rover. Las demás variables en el esquema corresponden con las ecuaciones del modelo y la estrategia de control.

Las ecuaciones requeridas por el bloque del filtro de Kalman extendido se presentan en (4) y (5); en las cuales f es la función del modelo discretizado del rover, y g la función de los diferentes modelos de medición implicados en la fusión; w_k y v_k son respectivamente las matrices que representan el ruido del proceso y de medición:

$$x_k = f(x_{k-1}, u_k) + w_k, \quad (4)$$

$$y_k = g(x_k) + v_k. \quad (5)$$

El sistema completo se desarrolla mediante los paquetes de Matlab-Simulink que aportan las librerías para la simulación de distintas configuraciones del filtro de Kalman, además permite agregar los jacobianos de las anteriores ecuaciones, que en el caso del rover Shrimp, son nulos debido a que la matriz A del modelo presentado es cero. La estimación de la posición y orientación del robot ante dos trayectorias deseadas diferentes se realiza utilizando tres sensores, uno mide la orientación angular, y los otros dos la posición respecto de los ejes x, y ; cada uno de ellos tiene una precisión diferente permitiendo que el filtro haga una estimación adecuada para realimentar el algoritmo de control, y conseguir que el robot realice el seguimiento de trayectorias, que en este caso es una forma circular y otra en forma de lemniscata.

La fig. 4 muestra los tres bloques en el entorno de Simulink que contienen las funciones de medición de los 3 sensores, las cuales se encargan de agregar el ruido de medición a cada variable real, como se presenta en la fig. 5.

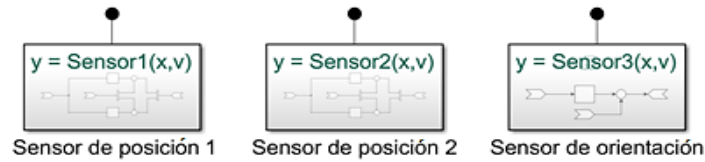


Fig. 4. Funciones de Simulink para los modelos de medición.

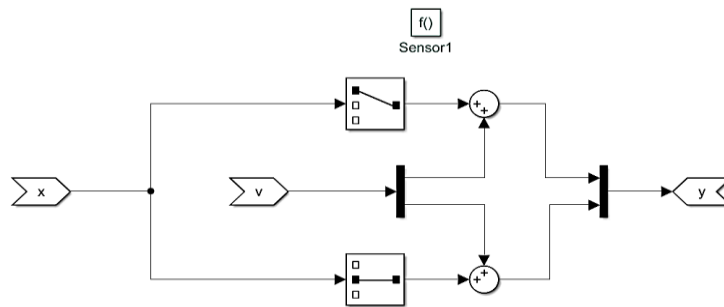


Fig. 5. Bloques del modelo de medición del sensor 1 de posición. x corresponde a la variable real de posición; v es el vector de ruido de medición; la variable y , es la medición de del sensor incluyendo el ruido.

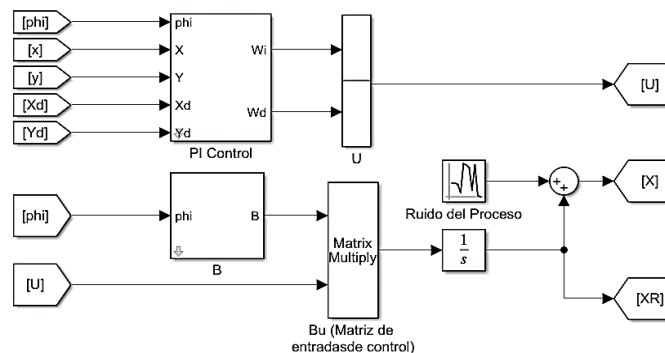


Fig. 6. Bloques del modelo cinemático, representado por la matriz del subsistema “B”, y la ley de control de posición del rover, localizada en el bloque “PI Control”.

Por otra parte, el modelo del sistema y la ley de control son mostrados en la fig. 6, donde las matrices B del modelo cinemático y el subsistema PI contienen las ecuaciones (1 y 2), respectivamente.

Finalmente, el subsistema más relevante dentro del sistema de fusión de sensores se presenta en la fig. 7, denominado “Filtro de Kalman Extendido”, el cual contiene toda la lógica necesaria para simular el filtro propuesto bajo las condiciones dadas. Se puede ver entonces que el filtro utiliza las señales de los tres sensores y de la entrada de control; y entrega una estimación del vector de estados de posición y orientación del robot.

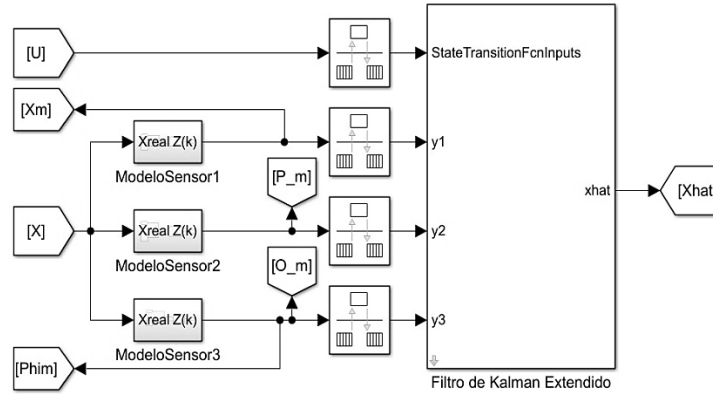


Fig. 7. Fusión de sensores mediante filtro de Kalman extendido.

```
function x = TransicionEstadosRoverShrimp(x,u)

R = 0.062;
L = 0.105;
b = 0.425;

dt = 0.001;
x = x + cumtrapz([ (R/2)*cos(x(3))+(L*R/b)*sin(x(3)) (R/2)*cos(x(3))-(L*R/b)*sin(x(3));
(R/2)*sin(x(3))-(L*R/b)*cos(x(3)) (R/2)*sin(x(3))+(L*R/b)*cos(x(3));
-R/b R/b ]*u)*dt;

end
```

Fig. 8. Función de transición de estados, incluye los valores (R , L , b) reales del robot en metros.

El subsistema del filtro de Kalman extendido requiere de la función de transición de estados, la cual es insertada en el sistema mediante su declaración en un script de Matlab como se muestra en la fig. 8.

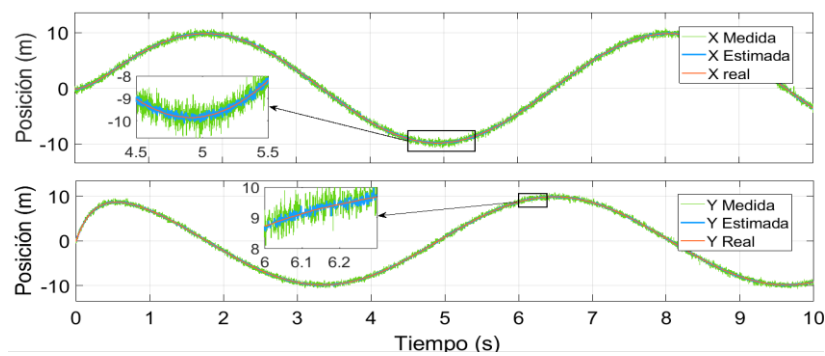
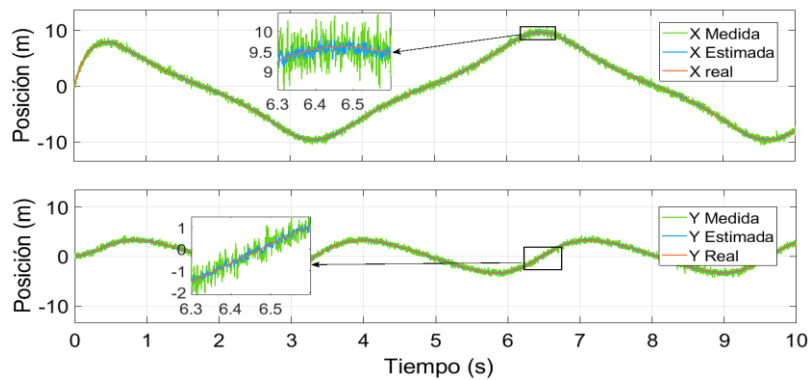
4. Resultados

En esta sección se presentan algunas simulaciones numéricas mediante las cuales se revisa el comportamiento de la estrategia; las simulaciones se realizan mediante Matlab-Simulink con un tiempo de muestreo de 0.01 s; las soluciones se realizan durante 10 s; las trayectorias deseadas propuestas de forma circular y de forma de lemniscata tienen un radio de 10 m; las condiciones iniciales son 0 y la covarianza inicial es 1. Los parámetros del modelo, las constantes del controlador PI y las matrices de covarianza propuestas para la simulación del filtro de Kalman extendido, se presentan en la Tabla 2, cuyas matrices de covarianza representan la precisión propuesta de cada sensor y del modelo cinemático presentado.

Las figs. 9 y 10 muestran la posición en cada eje para las tres señales del sistema (real, medida y estimada) de la trayectoria circular y de lemniscata, respectivamente; de la misma manera en las figs. 11 y 12 se presentan las señales de la orientación del

Tabla 2. Parámetros de simulación.

Parámetros	Valor
R	0.062 m
L	0.105 m
b	0.425
k_{px}, k_{ix}	5, 5
k_{py}, k_{iy}	0.1, 0.1
Covarianza del ruido en el proceso	$[0.012 \ 0 \ 0; 0 \ 0.012 \ 0; 0 \ 0 \ 0.012]$
Covarianza de la medición sensor 1	$[0.1 \ 0; 0 \ 0.1]$
Covarianza de la medición sensor 2	$[0.005 \ 0; 0 \ 0.005]$
Covarianza de la medición sensor 3	0.175

**Fig. 9.** Comparación de la posición en la trayectoria circular.**Fig. 10.** Comparación de la posición en la trayectoria de lemniscata.

robot, obteniendo un error cuadrático medio entre la estimación del filtro y la medición real, de 0.0928 para el círculo y de 0.1152 para la lemniscata.

Finalmente, las trayectorias realizadas por el robot se presentan en la fig. 13; las cuales toman como punto de referencia el centro de gravedad del robot, designado como punto de control en el modelo cinemático del rover para seguir las trayectorias deseadas.

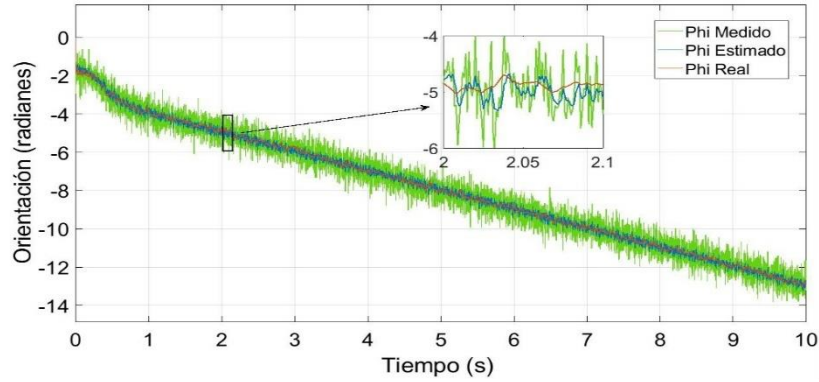


Fig. 11. Comparación de la orientación en la trayectoria circular.

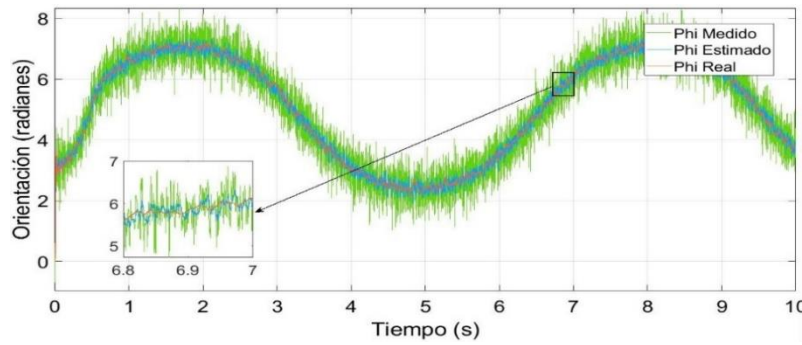


Fig. 12. Comparación de la orientación en la trayectoria de lemniscata.

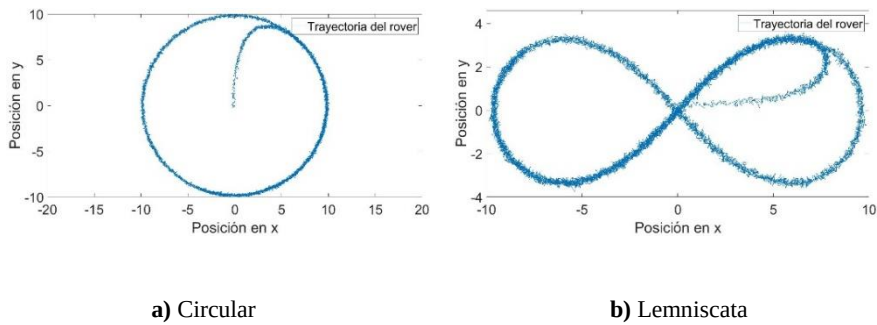


Fig. 13. Trayectoria trazada por el robot rover en el seguimiento de trayectorias.

Los resultados obtenidos exponen rendimientos importantes en cuanto al posicionamiento y el desempeño del controlador. El error cuadrático medio de posición obtenido sin la fusión es de 2.082 para la trayectoria circular, y de 1.616 para la trayectoria de lemniscata considerando las mediciones más precisas.

Por otra parte, si se realimenta el controlador con la señal estimada proveniente de la fusión, se consigue una mejora en el error de posición del 30.31% para el círculo, y del 24.26% para la lemniscata, pues el error cuadrático medio de posición resultante es de 1.451 y 1.224 para cada trayectoria respectivamente; lo cual evidencia un recorrido más preciso al utilizar el esquema de fusión de sensores.

Por otra parte, es importante mencionar que los resultados de la estimación de la orientación del robot dependen en gran parte de la precisión del sensor; esto se debe principalmente a que el sistema planteado utiliza un solo sensor para medir ese estado, adicionándole que el ángulo de orientación es completamente independiente de los otros dos estados y depende únicamente de las entradas de control, consecuencia de que en el modelo presentado la matriz A sea nula.

5. Conclusiones

El sistema de fusión de sensores planteado representa una alternativa para la implementación de algoritmos de seguimiento de trayectorias confiable, y que mejora entre otras cosas la precisión en comparación con algoritmos basados en un solo dispositivo de medición. En este caso se obtiene una mejora en el error cuadrático medio en el vector de estados del robot al usar las señales de estimación, permitiendo tener un error de posicionamiento del robot menor al utilizar la fusión de sensores.

Por otro lado, si bien el modelado del robot desprecia una cantidad importante de aspectos referentes al entorno de trabajo, queda claro que, bajo condiciones controladas, la propuesta presentada aquí es completamente aplicable, e incluso es posible adaptarla para robots rover con otro tipo de configuraciones.

Aunque la precisión del robot no tiene una dependencia directa con la ley de control, el planteamiento futuro de estrategias diferentes al PI; entregarían un punto importante de comparación y una mayor claridad de la relación de los resultados con el controlador utilizado.

El planteamiento de un modelo completo que incluya la cinemática o dinámica de las articulaciones en la suspensión del rover, los deslizamientos de las ruedas y condiciones variables en la inclinación del terreno; conllevaría a una generalización del método presentado aquí y, por ende, a resultados mucho más cercanos a aplicaciones en ambientes diversos, que quedan fuera de los alcances de este trabajo.

Agradecimientos. Este trabajo está financiado por la SIP-IPN, con número de registro 20210268. Jheison Duvier Díaz Ortega agradece a CONACYT; y el apoyo otorgado en la convocatoria de proyectos de desarrollo tecnológico o innovación para estudiantes del IPN 2021.

Referencias

1. Baturone Ollero, A.: Robótica manipuladores y robots móviles, Sevilla. Marcombo (2005)

2. Calle Flores, I.A.: Navegación autónoma de un robot móvil usando técnicas probabilísticas de localización y mapeo basadas en métodos Monte Carlo secuenciales (2014)
3. Acosta, G., Gallardo, J., Pérez, R.: Arquitectura de control reactiva para la navegación autónoma de robots móviles, *Ingeniare. Revista chilena de ingeniería*, 24(1), pp. 173–181 (2016) doi: 10.4067/S0718-33052016000100016.
4. Wolfgang, K.: *Tracking and Sensor Data Fusion: Methodological Framework and Selected Applications*, Berlin. Heidelberg: Springer (2014)
5. Zhu, H., Zou, K., Li, Y.: Robust Sensor Fusion with Heavy-Tailed Noises. *Signal Processing*, 175, 107659 (2020) doi: 10.1016/j.sigpro.2020.107659.
6. García, J., Molina, J.M., Trincado, J.: Real Evaluation for Designing Sensor Fusion in UAV Platforms. *Information Fusion*, 63, pp. 136–152 (2020) doi: 10.1016/j.inffus.2020.06.003.
7. Balemans, D., Casteels, W., Vanneste, S.: Resource Efficient Sensor Fusion by Knowledge-Based Network Pruning. *Internet of Things*, 11, 100231 (2020) doi: 10.1016/j.iot.2020.100231.
8. Bazo, R., Reis, E., Seewald L.A.: Baptizo: A Sensor Fusion Based Model for Tracking the Identity of Human Poses. *Information Fusion*, 62, pp. 1–13 (2020) doi: 10.1016/j.inffus.2020.03.011.
9. Olivares Ávila, M.A., Gallardo Arancibia, J.A.: Sistema de localización autónoma para robots móviles basado en fusión de sensores propioceptivos. *Revista Politécnica*, 11(21), pp. 75–84 (2015)
10. Al Arabi, A., Ul Sakib, H., Sarkar, P.: Autonomous Rover Navigation Using GPS Based Path Planning. In: 2017 Asia Modelling Symposium, pp. 89–94 (2017) doi: 10.1109/AMS.2017.22.
11. Hewitt, R., Ellery, A., Ruiter, A.: Training a Terrain Traversability Classifier for a Planetary Rover Through Simulation. *International Journal of Advanced Robotic Systems*, 14(5) (2017) doi: 10.1177/1729881417735401.
12. Schuster, M.J., Brunner, S.G., Bussmann, K., Büttne, S.: Towards Autonomous Planetary Exploration. *Journal of Intelligent & Robotic Systems*, 93(3) pp. 461–494 (2019)
13. Graven, O., Srisuphab, A., Silapachote, P.: An Autonomous Indoor Exploration Robot Rover and 3D Modeling with Photogrammetry. In: 2018 International ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering, pp. 44–47 (2018) doi: 10.1109/ECTI-NCON.2018.8378279.
14. Aguirre Anaya, J.A.: Diseño y construcción de un móvil tipo Rover de arquitectura abierta. (2016) <https://cloud.upiita.ipn.mx/s/bRRCGpjFXawpznN>. Tesis Maestría en Tecnología Avanzada / file 85.

Desarrollo de una red neuronal artificial para la predicción de sobre corte en un proceso de manufactura no convencional

Gabriela Itzell Rangel Velázquez¹, José Arturo Pérez Martínez²,
Everardo Efrén Granda Gutiérrez¹, Carlos Eduardo Torres Reyes²

¹ Universidad Autónoma del Estado de México,
Centro Universitario UAEM Atlacomulco,
México

² Universidad Autónoma del Estado de México,
Unidad Académica Profesional Tianguistenco,
México

grangelv001@alumno.uaemex.mx,
{eegranda, japerezm, cetorresr}@uaemex.mx

Resumen. El presente trabajo tiene como objetivo el diseño y desarrollo de un modelo de red neuronal artificial de tipo perceptrón multicapa para la predicción del sobre corte en un proceso de maquinado electroquímico por pulsos. Como variables de entrada se consideran el voltaje de polarización, la frecuencia de maquinado y la concentración del electrolito. Se determina un diseño experimental factorial completo 2^3 para la construcción del conjunto de datos, los cuales se segmentan en 70% para entrenamiento y 30% para validación. Como métrica de evaluación de desempeño del modelo, se calcula el error porcentual absoluto medio y, de acuerdo con los resultados obtenidos, se comprueba que el modelo desarrollado es una alternativa factible para utilizar en la predicción de la variable de interés del fenómeno de estudio.

Palabras clave: Maquinado electroquímico por pulsos, perceptrón multicapa, predicción de sobre corte, red neuronal artificial.

Development of an Artificial Neural Network for the Prediction of Overcut in a Non-Conventional Manufacturing Process

Abstract. The present article aims to design and develop a multilayer perceptron-type artificial neural network model to predict the called “overcut” in a pulse electrochemical machining process. The involved input variables are bias voltage, machining frequency and electrolyte concentration. A full factorial experimental design 2^3 is determined for the construction of the data set, which is segmented into 70% for training and 30% for validation. The metric used to evaluate the performance of the model, the mean absolute percentage error is calculated and, according to the obtained results, it is verified that the developed model is a feasible alternative to use in the prediction of the variable of interest of the study phenomenon.

Palabras clave: Artificial neural network, multilayer perceptron, pulse electrochemical machining, overcut prediction.

1. Introducción

La manufactura no convencional de materiales metálicos engloba procesos que permiten trabajar con piezas de geometría compleja, tamaño reducido, alta dureza o con un acabado específico, caracterizándose por evitar la producción de virutas durante su realización. Entre los ejemplos de este tipo de procesos se encuentran el maquinado por chorro de agua, el maquinado por ultrasonido, la electroerosión y el maquinado electroquímico [1]. Este último, es un proceso que permite la remoción de material mediante la generación de un fenómeno denominado disolución anódica, producido por el flujo de electrones que existe desde el cátodo (herramienta de trabajo), hasta el ánodo (pieza de trabajo), a través de un electrolito que actúa como medio conductor de corriente eléctrica [2], con la finalidad de que la pieza de trabajo adquiera la forma del negativo de la figura de la herramienta de trabajo. De acuerdo con McGeough [3], algunos de los factores de importancia que se deben de tomar en consideración para la realización de un proceso de maquinado electroquímico, son la distancia inter-electrodos, el tipo de solución y concentración del electrolito, así como la fuente de alimentación, ya que esta última define si es un proceso de corriente continua o bien, si se trata de un proceso de maquinado electroquímico por pulsos. Además, dichos factores tienen influencia en las variables de salida del proceso, entre las que se encuentran la tasa de remoción de material y el sobre corte, es decir, cuánto material se remueve por unidad de tiempo y la diferencia existente entre el diámetro de la figura hecha en la pieza de trabajo menos el diámetro de la herramienta de trabajo utilizada [2].

Para aportar a la optimización del sector manufacturero, diversos investigadores como [4-8], han estudiado y considerado la implementación de soluciones inteligentes, dando prioridad al uso del *Soft Computing*, que tiene como objetivo obtener y representar información incierta, imprecisa o incompleta para la toma de decisiones de forma automática, reduciendo tiempo e inversión económica, el cual engloba diversas técnicas como la lógica difusa, algoritmos genéticos o las redes neuronales artificiales (RNA), entre otras [9].

Las RNA también conocidas como neuro-redes, son modelos computacionales que tienen el propósito de emular el funcionamiento del cerebro humano, con la finalidad de que un sistema informático sea capaz de tomar decisiones basadas en conocimiento empírico [10]. Estas se conforman por neuronas, las cuales están interconectadas jerárquicamente formando capas para la transmisión de información; además, las RNA son usadas para la resolución de problemas de clasificación mediante clases y etiquetado de metadatos, o para problemas de predicción de valores, también conocidos como problemas de aproximación de funciones, entre los que se encuentran los problemas de regresión y las series temporales [11]. Dentro de las capas que conforman la RNA, se encuentran la capa de entrada, que es aquella que interactúa con el exterior recibiendo la información necesaria para la definición de las variables independientes; la capa oculta, que puede ser desde una hasta las que se consideren necesarias para el entrenamiento de la red y las cuales permiten obtener el aprendizaje deseado;

finalmente, la capa de salida, que contiene la resultante de las variables dependientes del proceso de estudio [10]. Las RNA se han utilizado para contribuir al estudio y mejora de procesos de maquinado electroquímico, tal es el caso de [12], en donde se presenta un modelo para predecir los valores de la tasa de remoción y estudiar el efecto que ejercen sobre ella las variables de entrada, considerando la corriente (50, 100 y 200 A), la concentración del electrolito (10, 25 y 50 g/l) y el gap (0.5, 1 y 1.5 mm).

Emplean un conjunto de 12 datos, de los cuales 9 fueron para entrenamiento y 3 para validación, realizando una comparativa de los valores predichos y los valores reales, de los cuales se obtiene que la predicción correcta es mayor al 70%. Así mismo en [13], proponen un modelo de predicción de la tasa de remoción mediante una RNA perceptrón multicapa con un algoritmo de retropropagación elaborado en MATLAB; consideran el voltaje de polarización (6 y 10 V), la velocidad de avance de la herramienta (0.16 y 0.23 mm/min), la concentración del electrolito (200 y 300 g/l) y el material de la herramienta de trabajo (cobre y latón).

Se emplea un conjunto de datos construido a partir de un factorial completo 2^4 , segmentado en 70% de los datos para entrenamiento y el 30% restante para validación; obtienen que el menor Error Porcentual Absoluto Medio (MAPE, por sus siglas en inglés) resultante fue del 0.0018% correspondiente a 6 V de polarización, 0.23 mm/min de velocidad de avance, 200 g/l de concentración y como material, latón. Por otro lado, en [14] plantean que es necesario mejorar el sistema de control del proceso de maquinado electroquímico para que se puedan aprovechar los beneficios que el uso de este tipo de procesos brinda en la manufactura de metales.

Proponen un modelo de diagnóstico basado en RNA perceptrón multicapa para reducir imprecisiones derivadas de los modelos matemáticos utilizados con fines de control, o posibles fallas provocadas por un diseño incorrecto del proceso, esto con el propósito de determinar ciertas características sobre el maquinado con respecto al equipo de polarización. Realizan pruebas con la variación de la topología de RNA, donde utilizan 4 neuronas en la capa de entrada, 3, 4, 5, 6 y 10 neuronas en la capa oculta y 2 neuronas en la capa de salida; emplean como función de activación la tangente hiperbólica para la capa oculta y función logística para la capa de salida. Obtienen como resultado que el menor error de entrenamiento es del 0.257446% con una topología 4 - 4 - 2; y el menor error de validación es del 0.263862% con una topología 4 - 6 - 2.

El presente artículo tiene como objetivo describir el proceso de diseño y desarrollo de un modelo de RNA perceptrón multicapa para predecir el sobre corte producido en la pieza de trabajo al término de un proceso de maquinado electroquímico por pulsos, con la finalidad de poder ajustar las condiciones de trabajo correspondientes previo a la realización de un experimento real.

Se consideran como variables de entrada la magnitud del voltaje de polarización, la frecuencia de trabajo de los pulsos aplicados durante el proceso y la concentración del electrolito. Además, se construye un conjunto de datos mediante un factorial completo 2^3 , del cual se realiza una segmentación del 70% de datos para entrenamiento y el 30% para validación; finalmente, se evalúa el desempeño del modelo usando como métrica el cálculo del MAPE para determinar si es factible la solución propuesta.

2. Metodología

2.1. Condiciones para la obtención del conjunto de datos

El diseño del experimento (DOE, por sus siglas en inglés) que se determinó para la construcción del conjunto de datos para el entrenamiento y validación del modelo, fue un factorial completo 2^k [15], debido a que el proceso de maquinado electroquímico por pulsos es un fenómeno multivariable. El valor de k es igual a 3, esto significa que se consideran 3 factores: el voltaje de polarización, la frecuencia de maquinado y la concentración del electrolito, con dos niveles determinados para cada uno (nivel alto y nivel bajo). En la Tabla 1 se muestran los niveles de cada factor para la realización de los experimentos correspondientes a la matriz de combinaciones del DOE factorial 2^3 . Los valores de cada nivel sirvieron para poder establecer los parámetros iniciales de cada maquinado y así, obtener un valor del sobre corte por cada barreno realizado.

Una vez definidos los factores de estudio y los niveles de cada factor, se construye la matriz de combinaciones DOE factorial 2^3 , la cual se presenta en la Tabla 2, en donde se pueden observar todas las combinaciones de los factores y sus niveles por cada uno de los experimentos que se deben realizar.

Así mismo, para observar el comportamiento del fenómeno bajo las condiciones de entrada establecidas en la matriz construida, se determinan dos repeticiones por experimento; es decir, se realiza 3 veces cada una de las combinaciones de la variables de entrada, obteniendo un total de 24 respuestas, donde cada tercia de respuestas se promedia para tener un conjunto de datos promedio del valor del sobre corte resultante.

2.2. Prototipo del sistema para el proceso de maquinado electroquímico

Para la experimentación se empleó un prototipo de sistema mecánico de tres ejes coordenados (X, Y, Z) controlado por una Raspberry Pi 3B, el cual se conforma por una bomba que permite el flujo del electrolito, una fuente de alimentación pulsada con una frecuencia de conmutación ajustable entre 20 kHz y 100 kHz, así como un regulador de voltaje de polarización ajustable de 2.5 V a 10 V.

Además, como pieza de trabajo se empleó acero al carbón AISI-1018; como herramienta de trabajo se utilizó una aguja de acero inoxidable de 603 μm de diámetro externo y, como electrolito, se usó una solución de cloruro de sodio ($\text{H}_2\text{O} + \text{NaCl}$). El tiempo de maquinado establecido por cada experimento fue de 60 segundos a lazo abierto con una distancia inicial entre los electrodos (*gap*) de 240 μm . La representación gráfica del esquema general del funcionamiento del prototipo se puede ver en la Figura 1, en donde se muestra cómo fluye el electrolito durante el proceso, así como la corriente real producida en donde se ven involucradas, principalmente, la herramienta y pieza de trabajo.

Por otra parte, el cálculo del sobre corte se realizó considerando la diferencia entre el diámetro del barreno maquinado y el diámetro del electrodo, tal como se muestra en la ecuación (1).

Tabla 1. Niveles alto y bajo de cada factor para experimentos de la matriz factorial 2^3 .

	Voltaje de polarización	Frecuencia	Concentración del electrolito
	[V]	[kHz]	[M]
Nivel bajo	5	60	2.5
Nivel alto	7	80	3

Tabla 2. Matriz de combinaciones del DOE para la obtención del data set.

No. Experimento	Voltaje de polarización [V]	Frecuencia [kHz]	Concentración del electrolito [M]
1	bajo	bajo	bajo
2	alto	bajo	bajo
3	bajo	alto	bajo
4	alto	alto	bajo
5	bajo	bajo	alto
6	alto	bajo	alto
7	bajo	alto	alto
8	alto	alto	alto

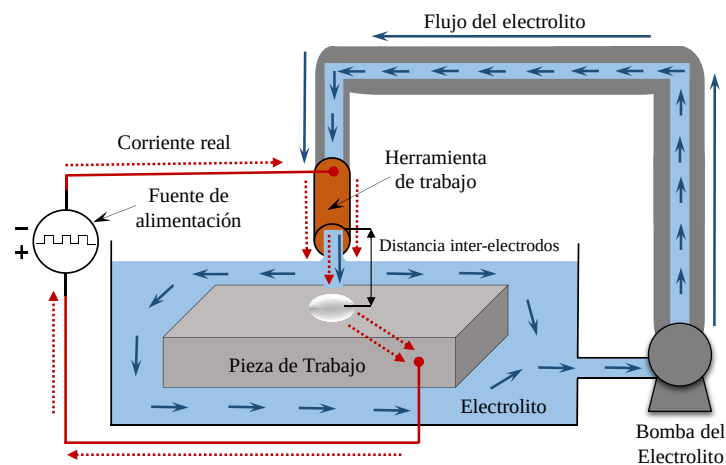


Fig. 1. Representación gráfica del esquema general del prototipo del sistema de maquinado electroquímico utilizado.

$$\text{Sobre Corte} = dP - dH, \quad (1)$$

donde:

dP = diámetro de la pieza de trabajo,

dH = diámetro de la herramienta de trabajo.

De manera ilustrativa, en la Figura 2 se muestra el proceso de generación de sobre corte en un proceso de maquinado electroquímico, donde a) representa el inicio del corte en la pieza, la corriente principal generada a través del flujo del electrolito, la cual sale a través de la herramienta, se coloca en color rojo; posteriormente en b) se muestra

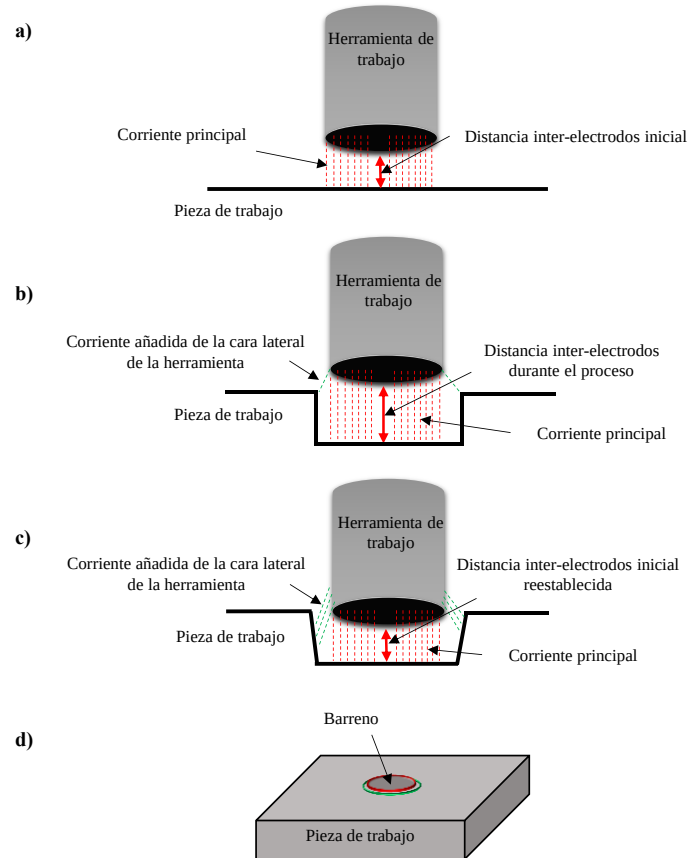


Fig. 2. Generación de sobre corte en un proceso de maquinado electroquímico. a) Inicio del proceso. b) Adición de la corriente de las caras laterales de la herramienta. c) Aumento de la corriente por el descenso de la herramienta. d) Barreno con efecto de sobre corte.

la remoción de material en la pieza de trabajo, el aumento del *gap* y, de color verde, la corriente producida por las caras laterales de la herramienta de trabajo; en c) se representa el descenso de la herramienta con el *gap* reestablecido y la corriente de las caras laterales aumentada. Finalmente, en d) se muestra el maquinado realizado en la pieza de trabajo, el cual es un barreno producido por la corriente principal que tiene el borde en color rojo y, a su vez, un corte adicional producido por la corriente de las caras laterales de la herramienta, el cual tiene el borde en color verde y es conocido como efecto de sobre corte.

En la Figura 3, se muestra un ejemplar de barreno realizado con una concentración de 2.5 M, 5 V de polarización y 80 kHz de frecuencia de maquinado. Se puede observar que el diámetro del barreno es de 836 μm (representado con el borde en color rojo), por lo que, al utilizar la ecuación (1) y restar los 603 μm del diámetro de la herramienta de trabajo (representados con el borde color blanco), se obtiene un sobre corte de 233 μm . Este proceso de obtención de sobre corte se repite en cada una de las pruebas hechas con base en la matriz del DOE factorial completo 2^3 .

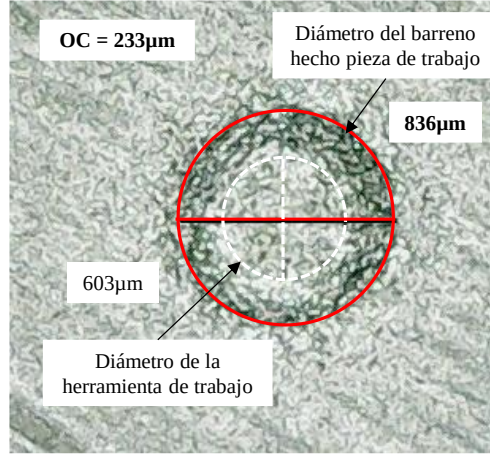


Fig. 3. Barreno realizado a 60 segundos, con una concentración de 2.5 M, 5 V de polarización y 80 kHz de frecuencia de maquinado.

2.3. Modelo de red neuronal para la predicción de sobre corte

El modelo de RNA propuesto para la predicción de sobre corte, es un perceptrón multicapa con una topología definida en la ecuación (2) [16]:

$$F: I^3 \rightarrow \mathbb{R}^1, \quad (2)$$

donde:

F = función continua del modelo de RNA,
 I^3 = matriz de entradas,
 $\mathbb{R}^1$ = matriz de salidas.

La representación gráfica de la topología del modelo propuesto se puede visualizar en la Figura 4, en donde se encuentra la matriz de entradas, que son 3: el voltaje de polarización (v), la frecuencia de maquinado (f) y la concentración del electrolito (ce); también se observa una capa oculta de cuatro neuronas y la matriz de salidas, que en este caso es solo una: el sobre corte (oc).

Cada conexión entre las neuronas que conforman la topología tiene asociado un coeficiente w , conocido como peso sináptico [10]. De modo que, el modelo de RNA propuesto, tiene dos matrices de coeficientes: W^1 , que es una matriz de 4×3 , conformada por las conexiones de las neuronas de la capa de entrada a las neuronas de la capa oculta; y W^2 , que es una matriz de 4×1 , conformada por las conexiones de las neuronas de la capa oculta hacia la neurona de la capa de salida.

Considerando la ecuación (3) como la ecuación general para el cálculo de las salidas, y la ecuación (4) para el cálculo de la función de activación rectificada lineal ($relu$) [17], se procede al recorrido del perceptrón de forma unidireccional hacia delante, para obtener el valor de la variable dependiente:

$$Y = wx + u, \quad (3)$$

$$\phi(Y) = \max\{0, Y\}, \text{ siendo } Y \geq 0, \quad (4)$$

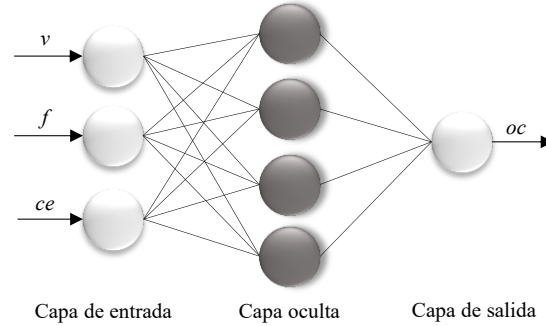


Fig. 4. Topología del modelo de RNA propuesto.

donde:

Y = la salida que se desea obtener,
 w = peso sináptico de la conexión,
 x = valor de entrada,
 u = valor del umbral,
 ϕ = función de activación.

De acuerdo con la estructura de la topología, considerando la ecuación (3) y las matrices de coeficientes W^1 y W^2 , primero se calcula el resultado de la capa de entrada hacia la capa oculta (Y^1) y este se utiliza para calcular el valor de la capa de salida (Y^2), para posteriormente aplicar la función de activación y obtener el valor del sobre corte. Sustituyendo los valores, el recorrido unidireccional hacia delante del perceptrón multicapa, se representa por las ecuaciones (5 - 7), considerando a U^1 y U^2 como las matrices del umbral por cada capa recorrida:

$$Y^1 = \begin{bmatrix} W_{1,1}^1 & W_{1,2}^1 & W_{1,3}^1 \\ W_{2,1}^1 & W_{2,2}^1 & W_{2,3}^1 \\ W_{3,1}^1 & W_{3,2}^1 & W_{3,3}^1 \\ W_{4,1}^1 & W_{4,2}^1 & W_{4,3}^1 \end{bmatrix} * \begin{bmatrix} v \\ f \\ ce \end{bmatrix} + \begin{bmatrix} U_{1,1}^1 \\ U_{2,1}^1 \\ U_{3,1}^1 \\ U_{4,1}^1 \end{bmatrix}, \quad (5)$$

$$Y^2 = \begin{bmatrix} W_{1,1}^2 \\ W_{2,1}^2 \\ W_{3,1}^2 \\ W_{4,1}^2 \end{bmatrix} * \begin{bmatrix} \phi(Y_{1,1}^1) \\ \phi(Y_{2,1}^1) \\ \phi(Y_{3,1}^1) \\ \phi(Y_{4,1}^1) \end{bmatrix} + \begin{bmatrix} U_{1,1}^2 \\ U_{2,1}^2 \\ U_{3,1}^2 \\ U_{4,1}^2 \end{bmatrix}, \quad (6)$$

$$oc = \phi(Y^2). \quad (7)$$

Para llevar a cabo la codificación del modelo se utilizó Python como lenguaje de programación y Google Collaboratory [18] como entorno de desarrollo, fue necesario implementar una normalización de datos para que se establecieran dentro del mismo rango debido a la diferencia existente entre los valores de las entradas y los de la variable dependiente. Dicha normalización se realizó previo al entrenamiento del modelo y para ello se utilizó *preprocessing* de la librería *sklearn* [19]. La función de normalización utilizada fue *QuantileTransformer* [20], que aplica una transformación

no lineal de modo que se hace un mapeo de los datos en una distribución uniforme o gaussiana dentro del rango (0, 1).

Una vez normalizado el conjunto de datos, se procedió al entrenamiento del modelo utilizando una segmentación del 70% de los datos del conjunto. Como métrica para la evaluación del desempeño del modelo se utilizó el MAPE, el cual es el promedio de la diferencia absoluta entre el valor observado y los valores predichos [21], la ecuación que lo define es la ec. (8):

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|p_t - \hat{p}_t|}{|p_t|} * 100, \quad (8)$$

donde:

p_t = valor predicho,

$\hat{p}_t$ = valor real o esperado,

n = número de datos.

3. Resultados y discusión

Los valores de sobre corte obtenidos mediante la aplicación de la matriz del DOE factorial 2^3 , se muestran dentro de la Tabla 3. Los datos observados son el valor promedio de cada terna de pruebas de las 24 en total realizadas, y el cual se consideró como conjunto de datos de entrenamiento (mostrados en color verde) y validación (mostrados en color naranja) del modelo.

Correspondiente a la relación existente entre el sobre corte obtenido y las variables de entrada utilizadas por la combinación de la matriz factorial (ver Figura 5), se puede observar que el menor sobre corte es de 124 μm , el cual se presenta con una concentración de 2.5 M, 60 kHz de frecuencia de maquinado y 5 V de voltaje de polarización; por otro lado, el mayor sobre corte es de 296 μm , obtenido de igual forma con una concentración de 2.5 M y 5 V de voltaje de polarización, pero con una frecuencia de maquinado de 80 kHz. Así mismo, se visualiza que, con una concentración de 3 M, los valores de sobre corte no tienen un cambio significativo entre sí, es decir, los valores de sobre corte obtenidos con esta concentración no varían demasiado entre sí, a pesar de que la frecuencia de maquinado y el voltaje de polarización se alternen en los niveles determinados.

Al realizar la normalización de los datos del conjunto construido con el DOE establecido, se obtuvieron dos matrices: 1) la matriz de datos de entrada normalizados y 2) la matriz de datos de salida normalizados. Para la primera matriz, se tomaron en cuenta los valores del voltaje de polarización, frecuencia de maquinado y concentración del electrolito definidos en las combinaciones del DOE; en el caso de la segunda matriz, se consideraron los valores de la columna de sobre corte mostrados en la Tabla 3, de esta forma, todos los datos utilizados para el entrenamiento estuvieron dentro del rango (0,1). Esto se puede observar en las Figuras 6 y 7.

Tabla 3. Valor promedio de sobre corte obtenido con la matriz de DOE factorial 2³.

No. Experimento	Voltaje de polarización [V]	Frecuencia [kHz]	Concentración del electrolito [M]	Sobre corte [μm]
1	5	60	2.5	124
2	7	60	2.5	248
3	5	80	2.5	296
4	7	80	2.5	217
5	5	60	3	284
6	7	60	3	244
7	5	80	3	269
8	7	80	3	251

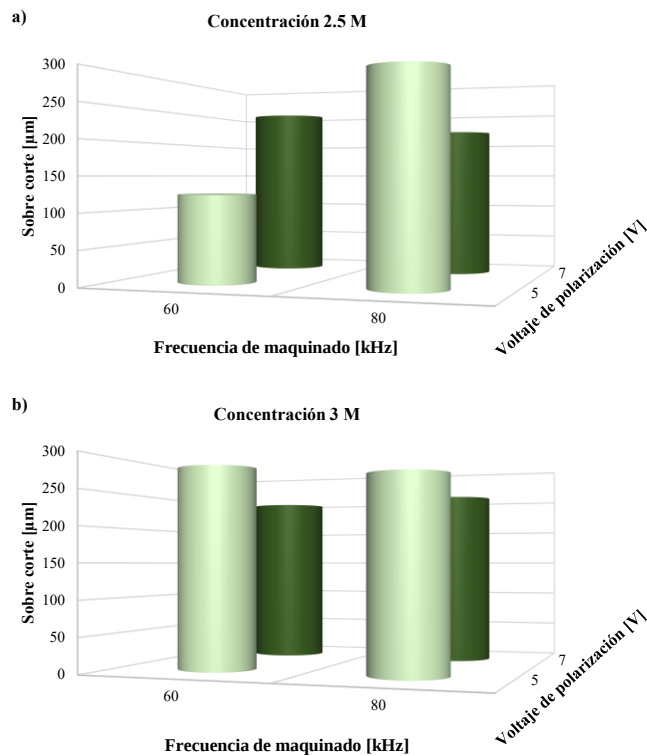
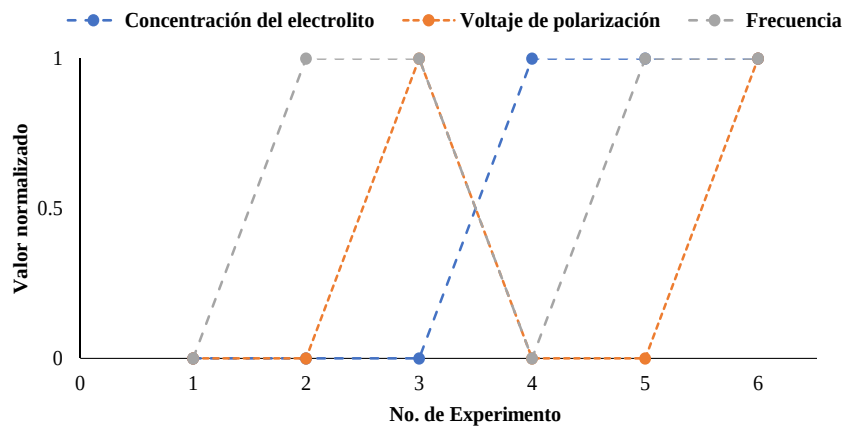


Fig. 5. Relación entre las variables de entrada y la variable de salida para la visualización del comportamiento del fenómeno. a) Concentración de 2.5 M. b) Concentración de 3 M.

Al obtener los datos de la normalización se procedió al entrenamiento del modelo, destacando la importancia de la normalización para la disminución del error en la predicción de los datos. Con la segmentación del 70% de los datos utilizados para el entrenamiento, se obtuvo un MAPE del 7.3%. Se observa que la mayor diferencia de predicción se encuentra en el segundo experimento, con una cantidad de 43.83 μm y la diferencia de predicción más baja, es de la posición cuatro con una cantidad de 2.11 μm (ver Tabla 4). La distribución de los datos predichos por el modelo, en comparación

Tabla 4. Valor del MAPE obtenido en el entrenamiento, considerando los valores reales y los valores predichos por el modelo correspondientes al sobre corte.

No. Experimento	Datos Reales [μm]	Datos Predichos [μm]	MAPE
1	124	146.39	18.05 %
2	296	248.17	16.15 %
3	217	221.62	2.12 %
4	244	241.89	0.86 %
5	269	260.94	2.99 %
6	251	241.89	3.62 %
Total MAPE entrenamiento			7.3 %

**Fig. 6.** Distribución de la matriz de datos de entrada normalizados en un rango de 0 a 1.

con los datos reales del sobre corte promedio obtenido por cada experimento se puede observar en la Figura 8.

De igual forma, con los datos de predicción obtenidos durante el entrenamiento, se pudo observar la relación existente entre las variables de entradas del proceso y el sobre corte obtenido, los cuales se muestran en la Figura 9. Se observa que la predicción de sobre corte más cercana, corresponde a las condiciones de 7 V de polarización, 60 kHz de frecuencia de maquinado y 3 M de concentración del electrolito.

Por otro lado, se realizó una validación con el 30% de los datos no usados para entrenamiento, en donde se obtuvo un MAPE del 16.3%. Los datos predichos en la validación se pueden observar en la Tabla 5, en donde se visualiza que la diferencia más grande de los datos predichos es de 61.11 unidades.

Al obtener los datos de validación, se puede observar que, con relación a las variables de entrada, el valor de predicción más cercano es con 60 kHz de frecuencia de maquinado 5 V de polarización y una concentración del electrolito de 3 M, tal como se muestra en la Figura 10.

Tabla 5. Valor del MAPE obtenido en la validación, considerando los valores reales y los valores predichos por el modelo.

No. Experimento	Datos Reales [μm]	Datos Predichos [μm]	MAPE
1	248	186.89	24.64 %
2	284	260.94	8.11 %
Total MAPE validación			16.3 %

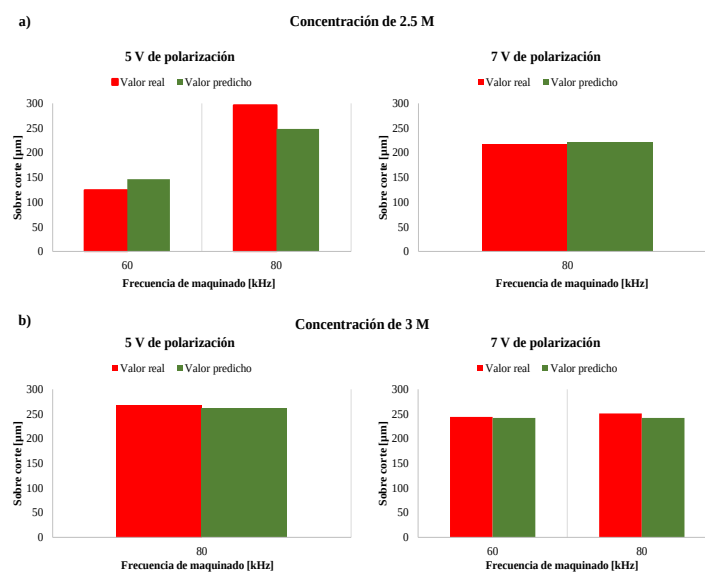


Fig. 9. Relación de las variables de entrada con los valores reales y los valores predichos en el entrenamiento del modelo. a) Concentración de 2.5 M. b) Concentración de 3 M.

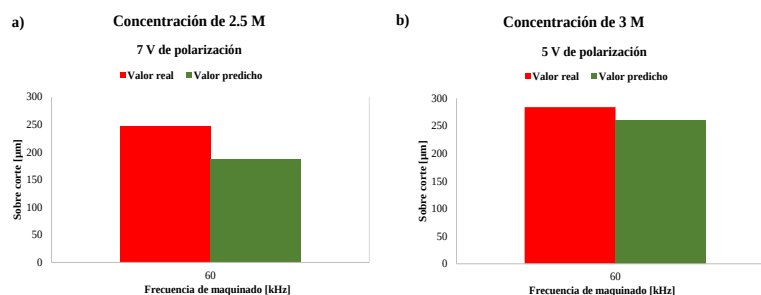


Fig. 10. Relación de las variables de entrada con los valores reales y los valores predichos en la validación del modelo. a) Concentración de 2.5 M. b) Concentración de 3 M.

4. Conclusiones

De acuerdo con los resultados obtenidos del entrenamiento y validación del modelo, se concluye que las redes neuronales artificiales, son una técnica de aprendizaje

supervisado factible para la predicción correcta del sobre corte producido al término de un proceso de maquinado electroquímico por pulsos, coincidiendo con lo expuesto en [12 - 14]. Esto se puede observar con el resultado de la métrica utilizada para la evaluación del desempeño, en donde, con el 70% de datos empleados para el entrenamiento, se obtiene un MAPE del 7.3% y, con el 30% de los datos empleados para la validación, se obtiene un MAPE del 16.3%.

Con respecto al comportamiento del fenómeno, se observa que, al emplear una concentración del electrolito de 3 M, no existe un cambio significativo en el valor del sobre corte, lo cual hace que, como trabajo a futuro, se pueda considerar realizar las pruebas con un tiempo indeterminado para estudiar con mayor profundidad cómo influye el cambio de concentración en el proceso y en el sobre corte.

Así mismo, se propone la construcción de un conjunto con una matriz factorial completo 3^3 , la cual permitirá definir tres valores por factor y ampliar el rango de combinaciones de los experimentos para que, de esta manera, se pueda tener un conjunto de datos de entrenamiento y validación más grande.

Además, se contempla la optimización del algoritmo para la reducción del error obtenido, implementando técnicas de optimización tradicionales tales como un algoritmo de retropropagación o bien, con técnicas híbridas como Algoritmos Genéticos o Lógica Difusa. Finalmente, se propone la variación de los parámetros de la topología propuesta, es decir, se propone la aleatoriedad de las características de la red, tales como el número de capas, el número de épocas o el número de neuronas por capa, continuando con la línea de entrenamiento de una red poco profunda, con el objetivo de verificar el comportamiento y desempeño del modelo bajo estas condiciones.

Referencias

1. Universidad del País Vasco: Métodos no convencionales de mecanizado. Dpto. de Ingeniería Mecánica, Euskadi, España (2015)
2. Schneider, M., Lohrengel, M.: Electrochemical Machining. Springer Hand Book, pp. 941–971 (2017) doi: 10.1007/978-3-662-46657-5_28.
3. Baswaraju, S., Harshvardhan T.: Principles of Electrochemical Machining. Encyclopedia of Chemical Technology, Londres, Inglaterra (2016)
4. Jena, R., Acharjya, D.P., Das, R.: A Fuzzy Graph Recurrent Neural Network Approach for the Prediction of Radial Overcut in Electro Discharge Machining. Advances in Distributed Computing and Machine Learning, 127, pp. 257–270 (2021) doi: 10.1007/978-981-15-4218-3_26.
5. Rouniyar, A.K., Shandilya, P.: Semi-Empirical Modeling and Optimization of Process Parameters on Overcut During MFAPM-EDM of Al6061 Alloy. Part E: Journal of Process Mechanical Engineering, 235, pp. 1784–1796 (2021) doi: 10.1177/09544089211015890.
6. Mirad, M.M., Das, B.: A Critical Review of the State-of-the-Art Literature in the Monitoring of Ultrasonic Machining Process and Tool Failure Prediction. Engineering Failure Analysis, 130 (2021) doi: 10.1016/j.engfailanal.2021.105769.
7. Das, R., Pradhan, M.K.: Artificial Neural Network Training Algorithms in Modeling of Radial Overcut in EDM: A Comparative Study. Research Anthology on Artificial Neural Network Applications, pp. 140–150 (2018) doi: 10.4018/978-1-5225-3035-0.ch006.
8. Sossa, H., Reyes, F.: Inteligencia artificial aplicada a la robótica y automatización. Marcombo-AlfaOmega (2021)

9. Swathi, B., Tiwari, H.: Test Automation Framework using Soft Computing Techniques. International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (2021) doi: 10.1109/ICAECT49130.2021. 9392602.
10. Basogain Olabe, X.: Redes neuronales artificiales y sus aplicaciones. Escuela Superior de Ingeniería de Bilbao (1998)
11. Palacios González F.: Redes neuronales artificiales y métodos de estimación de funciones. Universidad de Granada (2000)
12. Hassan, A., Aghdeab, S.H., Ibrahim, A.F.: Optimization of Electrochemical Machining Process Based on Artificial Neural Network Technique. Eng. & Tech. Journal, 34, pp. 2960–2970 (2016) doi: 10.30684/etj.34.15A.16.
13. Kasdekar, D.K., Parashar, V., Chandan, A.: Artificial Neural Network Models for the Prediction of MRR in Electro-Chemical Machining. Department of Mechanical Engineering, Madhav Institute of Technology & Science, 5, pp. 772–779 (2018) doi: 10.1016/j.matpr.2017.11.146.
14. Masalimov, K.A., Munasypov, R.A.: Neural-Network Diagnostics of Electrochemical Machining. Russian Engineering Research, 7, pp. 817–820 (2017) doi: 10.3103/S1068798X17090179.
15. Gutiérrez Pulido, H., De la Vara Salazar, R.: Análisis y diseño de experimentos. McGraw-Hill (2008)
16. Katsuura, A., Sprecher, D.: Computational Aspects of Kolmogorov Superposition Theorem. Neural Networks, 7, pp. 451– 461 (1994)
17. Vilagran Solsona, A.: Facial Expression Detection using Convolutional Neural Networks. Universidad Politécnica de Cataluña (2018)
18. Google Colab: <https://colab.research.google.com> (2021)
19. Sklearn Preprocessing Data: <https://scikit-learn.org/stable/modules/preprocessing.html> (2021)
20. Sklearn QuantileTransformer: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.QuantileTransformer.html#sklearn.preprocessing.QuantileTransformer> (2021)
21. Faryx: Selección de métricas para aprendizaje automático. https://fayrix.com/machine-learning-metrics_es (2021)

Investigación sobre el diseño de software en la última década: Un mapeo sistemático de literatura

José Baeza¹, Omar S. Gómez², Raúl A. Aguilar¹

¹ Universidad Autónoma de Yucatán,
México

² Escuela Superior Politécnica de Chimborazo,
Ecuador

josedelarosabass@gmail.com, avera@correo.uady.mx,
ogomez@esPOCH.edu.ec

Resumen. El presente estudio tiene como objetivo ofrecer una visión general de la investigación desarrollada en la última década, con base en las preguntas de investigación formuladas, con la intención de identificar áreas de oportunidad para continuar con estudios primarios o secundarios en temáticas específicas. Para lo anterior, se realizó un mapeo sistemático mediante el cual se seleccionaron 60 estudios primarios realizados en la última década, los cuales fueron analizados mediante un enfoque cuantitativo. El estudio plantea el análisis de aspectos como factores de calidad, métodos de representación utilizados, así como el grado de innovación de los métodos, técnicas o metodologías de diseño; adicionalmente, se analizan los métodos de evaluación empírica utilizados y el contexto en que se desarrollan dichos estudios. Los hallazgos indican interés en problemas vinculados con aspectos de calidad, y entre los principales factores considerados en los estudios, se encuentran la fiabilidad y la usabilidad; por otro lado, respecto del grado de innovación, se identificó que poco más de la mitad de los estudios se orienta a la mejora de métodos o técnicas existentes; en relación con el método de validación, el estudio de caso fue el más recurrido, y el contexto industrial duplica en número a los estudios realizados en un contexto académico. El estudio permite concluir que la investigación en el área de diseño software, es aún vigente en el contexto de la Ingeniería de Software.

Palabras clave: Calidad del diseño, diseño software, estudio de mapeo sistemático, métodos de validación empírica.

Research on Software Design in the Last Decade: A Systematic Mapping of Literature

Abstract. This study aims to offer an overview of the research developed in the last decade, based on the research questions formulated, with the intention of identifying areas of opportunity to continue with primary or secondary studies in specific topics. For this, a systematic mapping was carried out through which 60 primary studies carried out in the last decade were selected, which were analyzed using a quantitative approach. The study proposes the analysis of aspects such as quality factors, methods of representation used, as well as the degree of innovation of the methods, techniques or design methodologies. Additionally, the empirical evaluation methods used and the context in which said studies are carried out are analyzed. The findings indicate interest in problems related to quality aspects, and among the main factors considered in the studies are reliability and usability; On the other hand, regarding the degree of innovation, it was identified that just over half of the studies are aimed at improving existing methods or techniques; Regarding the validation method, the case study was the most used, and the industrial context doubles in number to the studies carried out in an academic context. The study allows to conclude that research in the area of software design is still in force in the context of Software Engineering.

Keywords: Design quality, software design, empirical validation methods, systematic mapping study.

1. Introducción

El concepto de diseño se define como “el proceso de definir una arquitectura, componentes, interfaces u otras características de un componente” junto al resultado de este [1]. El diseño representa una fase de vital importancia en las primeras etapas de la construcción del software, en esta se analizan los requisitos del software para poder generar una descripción —generalmente gráfica— de la estructura más adecuada para dar solución a la necesidad planteada por los clientes; como resultado de este proceso de abstracción obtenemos la organización de los diferentes componentes de software así como sus interfaces; de igual forma, este proceso debe describir con alto nivel de detalle todos los componentes, para poder lograr su construcción [2]. Al estar trabajando con en el proceso de diseño, los Ingenieros de Software deben crear diversos modelos —diseños— para dar solución al conjunto de requisitos acordados; en consecuencia, podemos utilizar dichos modelos como punto de partida en la construcción software, de manera que posteriormente se pueda planificar actividades como, codificación, verificación y validación.

De acuerdo con [3] el diseño software consta de dos actividades principales, en primer lugar, se habla de un diseño de alto nivel, en el que se desarrolla la estructura de nivel superior que permite identificar cada uno de los componentes; en segundo lugar, se tiene el diseño detallado, en la cual se especifica cada componente con un

alto nivel de detalle para facilitar su construcción. Por otro lado, existe un área de conocimiento utilizado en las ciencias e ingenierías, para obtener una abstracción de la solución del sistema con cierto nivel de precisión y detalle; esta abstracción del sistema es llamada “Modelo Software” la cual es usada para evaluar la solución antes de codificar, desde diferentes perspectivas o vistas. Este modelado del software, es representado de manera gráfica, usualmente mediante un conjunto de diagramas que forman parte del Lenguaje de Modelado Unificado (UML), los cuales ayudan en el desarrollo del software facilitando la comunicación entre los diferentes actores del proceso. En este contexto, podemos concebir el proceso de modelado del software en dos áreas, el modelado estático, el cual provee una vista del sistema que ilustra la relación entre las clases, y por otro lado, el modelado dinámico, que presenta el comportamiento de los diferentes elementos del sistema [4].

2. Metodología adoptada

Con el propósito de identificar la evidencia disponible publicada a lo largo de los últimos años en el área de diseño del software, y con ello caracterizar la investigación realizada, se optó por utilizar como metodología de estudio, el mapeo sistemático de literatura o estudio de mapeo (EM); dicho tipo de estudio secundario permite la explorar y clasificar la investigación existente, e identificar "lagunas de evidencia" donde se necesitan más estudios primarios [5]. Para el desarrollo del presente estudio se utilizó como referencia la guía propuesta en [6] en la cual se establecen las siguientes tareas:

1. *Formulación de las preguntas de investigación:* El objetivo principal de los EM es proporcionar una descripción general de un área de investigación e identificar la cantidad, tipo de investigación y resultados disponibles.
2. *Búsqueda de estudios primarios:* Los estudios primarios se obtienen utilizando cadenas de búsqueda en Bases de Datos Científicas, o navegando de manera manual en actas de conferencias o publicaciones de revistas relevantes.
3. *Selección de artículos relevantes:* Dicha tarea consiste en aplicar un conjunto de criterios tanto de inclusión, como de exclusión, para seleccionar los artículos primarios más adecuados.
4. *Definición de un esquema de clasificación:* El investigador lee los resúmenes y buscan palabras clave y conceptos que permitan identificar un esquema de clasificación.
5. *Extracción de datos y elaboración del reporte:* En esta tarea, los artículos relevantes se clasifican en un esquema previamente diseñado; es decir, se realiza la extracción de datos y se procede con análisis de los resultados, presentando las frecuencias de las publicaciones para cada categoría, lo anterior hace posible ver qué categorías se han enfatizado en investigaciones anteriores y, por tanto, identificar oportunidades para investigaciones futuras.

3. Planeación del estudio

La revisión del estado del arte plasmada en el SWEBOK [7] y el análisis de materiales considerados como referencias obligadas para el área de diseño software [3, 4], permitieron acumular información suficiente para la formulación de las preguntas de investigación.

3.1 Preguntas de investigación

A continuación, se definieron las siguientes preguntas de investigación en consideración al área de conocimiento de “Diseño Software” de la ingeniería de software.

- *PI01. ¿Cómo se distribuyen los estudios seleccionados en torno al diseño software, en la última década (2010-2020)?*
- *PI02 ¿Cuál es el objeto de estudio que abordan los métodos, técnicas o metodologías abordadas en los estudios primarios seleccionados?*
- *PI03 ¿Qué problemas han sido abordados con los patrones de diseño en los estudios primarios seleccionados?*
- *PI04 ¿Cuáles han sido los métodos de representación gráfica utilizados en los estudios primarios vinculados con el área de Diseño de Software?*
- *PI05 ¿Cuáles son los factores de calidad abordados en los estudios primarios vinculados con el área de Diseño de Software?*
- *PI06 ¿Cuál es el grado de innovación que poseen los métodos, técnicas o metodologías abordadas en los estudios primarios seleccionados?*
- *PI07 ¿Cuáles han sido los métodos de validación utilizados en la última década por los estudios primarios vinculados con el área de Diseño Software?*
- *PI08 ¿En qué contexto han sido desarrollados los estudios primarios vinculados con el área de Diseño Software durante la última década?*

Con la revisión del estado del arte y la formulación de las preguntas de investigación, se realizó un análisis PICOC [8] del área bajo estudio, dicho análisis permitió identificar palabras clave para elaborar la cadena de búsqueda, así como para tener una primera aproximación del instrumento para recolección de la información al momento de la revisión de los estudios que fuesen seleccionados.

- *Población: Diseño Software*
- *Intervención: Estilos arquitectónicos en diseño o Patrones de diseño*
- *Comparación: Factores de calidad para los Modelos de Diseño o Métricas de calidad para los Modelos de Diseño*
- *Resultados: Mejoramiento del Diseño Software o Mejoramiento de los Modelos del Diseño Software.*
- *Contexto: Industrial o Academia.*

Tabla 1. Categorías identificadas para el análisis de los estudios primarios seleccionados.

Aspecto	Categoría
PI02. Objeto de estudio.	Alto nivel, Detallado, Interfaz.
PI04. Modelo UML.	Actividad, estática, interacción, implementación.
PI06. Innovación.	Nuevo, Mejora.
PI07. Método de Validación	Estudios de caso, Experimento, Investigación-acción, observación.
PI08. Contexto del estudio.	Industria, Academia.

Tabla 2. Proceso de selección de estudios primarios.

BD	Fase 1	Fase 2	Fase 3
Google Scholar	276	106	50
IEEE Xplore	524	235	10
		Total	60

Con el análisis PICOC y las preguntas de investigación formuladas, se identificaron posibles categorías que permitiesen al investigador clasificar parte de la información a recopilar durante el análisis de los estudios primarios seleccionados (ver Tabla 1).

3.2 Definición de bases de datos

Para el estudio de Mapeo Sistemático se seleccionaron las siguientes Bases de Datos (BD):

- *Google Scholar*: una BD de artículos académicos (de una amplia gama de fuentes) sin restricciones de idioma, revistas o geográficas, lo cual permite acceder a literatura que no está disponible en otras BD. Si bien dicha base de datos no asegura la disponibilidad de los textos de los artículos, en el caso de un estudio de mapeo, resulta suficiente el acceso a los resúmenes.
- *IEEE Xplore*: una BD de investigación académica en las áreas de Ciencias de la Computación, Ingeniería Eléctrica y Electrónica, en la que los resúmenes son de libre acceso.

3.3 Cadena de Búsqueda

Habiendo identificado los términos claves, se generó la siguiente cadena de búsqueda:

Design AND (“architectural styles” OR patterns OR models OR techniques OR architectures)

Dicha cadena fue evaluada en los repositorios de Google Scholar y de IEEE Xplorer para realizar los ajustes para los formatos que manejan los repositorios teniendo a continuación las siguientes cadenas de búsquedas ajustadas:

- Google Scholar: *allintitle: “Software Design” AND (“architectural styles” OR patterns OR models OR techniques)*
- IEEE Xplore: *(“Document Title”: “Software Design” AND (“architectural styles” OR patterns OR models OR techniques))*

3.4 Criterios de Exclusión e Inclusión

Los criterios acordados para seleccionar los artículos primarios más adecuados, de entre el conjunto que nos arroje la cadena de búsqueda, son los siguientes:

Criterios de Inclusión:

- Publicaciones entre 2010 y 2020.
- Artículos publicados en revistas de investigación.
- Artículos con títulos y resúmenes en idioma inglés.

Criterios de Exclusión:

- Artículos duplicados en los repositorios.
- Artículos que no aborden aspectos vinculados con el proceso software.

4. Ejecución del estudio

El estudio fue desarrollado durante el mes de agosto de 2021, para dicho estudio se filtraron los resultados recuperados de las búsquedas en los repositorios de *Google Scholar* e *IEEE Xplore* se excluyeron aquellos que no estaban alineados con los objetivos del mapeo sistemático, de acuerdo con las siguientes fases:

- Fase 1: Utilizar nuestras cadenas de búsquedas en las BD seleccionados.
- Fase 2: Aplicar los criterios de inclusión al conjunto de estudios obtenidos al final de la fase 1.
- Fase 3: Aplicar los criterios de exclusión al conjunto de artículos obtenidos al final de la fase 2.

La Tabla 2 ilustra el número de estudios obtenidos al final de cada una de las tres fases del proceso de filtrado.

5. Resultados

En este apartado se presentarán los resultados del estudio de acuerdo con las preguntas de investigación planteadas.

PI01. ¿Cómo se distribuyen los estudios seleccionados en torno al diseño software, en la última década (2010-2020)?

En la Figura 1 se presenta un diagrama que ilustra el número de artículos publicados por año, en el periodo del 2010 al 2020. Como se puede observar, durante

Tabla 3. Categorías abordadas en los Patrones de Diseño.

Categoría	f
Factores de Calidad	24
Malas prácticas en el diseño	15
Diseños robustos para aspectos de seguridad	9
Falta de expertiz en los diseñadores	5
Diseños orientados a la automatización	4
No específica	3

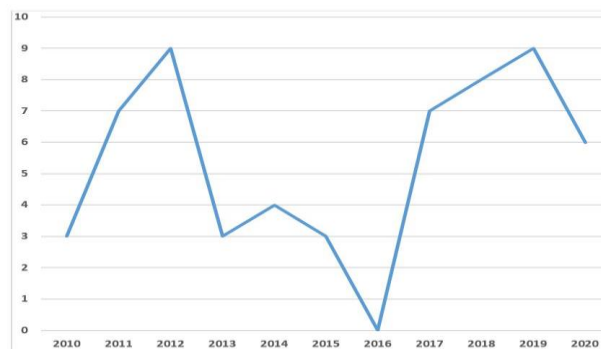


Fig. 1. Frecuencia de estudios seleccionados por año de publicación.

esta década la mayor parte de los años se ha mantenido en un constante número de publicaciones destacando dos años 2012, 2018 y 2019 que acumulan en los tres años el 43% del total de artículos publicados en la década; es de destacar también que en 2016 no se obtuvieron artículos publicados.

PI02 ¿Cuál es el objeto de estudio que abordan los métodos, técnicas o metodologías abordadas en los estudios primarios seleccionados?

En relación con el objeto de estudio, el 28% de los métodos, técnicas o metodologías identificadas, se enfocaron al “diseño a la Interfaz”, el cual representa la manera en la que los usuarios interactuarán con el sistema, junto con ciertas características, prototipos, así como formas de evaluar dichas experiencias. Un segundo grupo de estudios (45%) se enfoca al “diseño detallado”, es decir, técnicas y formas de describir cada uno de los componentes que conforman nuestro sistema; finalmente, el 27% restante habla sobre el “diseño de alto nivel”, por lo que notamos que muy pocos artículos se centran en la forma en la cual se organizan los componentes, o la manera en la que se representan ciertas vistas.

PI03 ¿Qué problemas han sido abordados con los patrones de diseño en los estudios primarios seleccionados?

De acuerdo con el análisis obtenido a partir de los estudios primarios, se identificaron cinco categorías de problemáticas abordadas, la frecuencia de cada una de estas se presenta en Tabla 3.

Tabla 4. Categorías abordadas en los patrones de diseño.

Categoría	f
Estática	24
Implementación	11
Actividad	10
Interacción	8
No específica	7

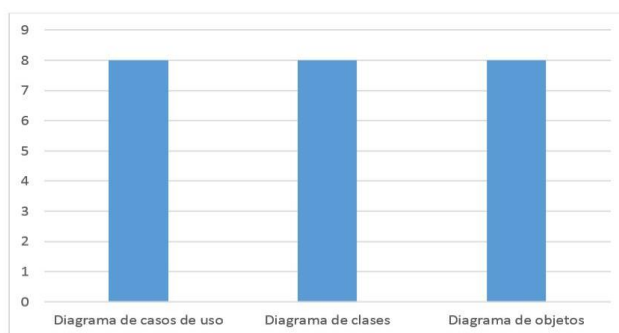


Fig. 3. Diagramas UML más utilizados en la categoría estática.

Podemos identificar que la categoría caracterizada por la moda, tiene que ver con la mejora en algún factor de calidad o con alguna propuesta entre la relación que existe entre crear un buen diseño y obtener un buen producto; por otro lado, la segunda categoría con mayor frecuencia se relaciona con el interés de mejorar los diseños evitando o reduciendo el uso de malas prácticas en dichos diseños, ya que en muchos casos, los diseños son muy complejas y esto genera retrasos o problemas para el desarrollo y entrega de los productos.

PI04 ¿Cuáles han sido los métodos de representación gráfica utilizados en los estudios primarios vinculados con el área de Diseño de Software?

De acuerdo con los tipos de diagramas propuestos por el Lenguaje de Modelado Unificado (UML), en la Tabla 4 se listan los métodos de representación y su frecuencia, identificados en los estudios primarios analizados.

Con base en la categoría que es más utilizada entre todos los estudios primarios, analizamos el tipo de diagrama referido en dicha categoría; la figura 3 ilustra los tres tipos de diagramas que pertenecen a esta categoría —estática— y que coincidentemente se reportaron con la misma frecuencia; cabe resaltar que estos diagramas es común encontrarlos en la documentación de los proyectos software.

PI05 ¿Cuáles son los factores de calidad abordados en los estudios primarios vinculados con el área de Diseño de Software?

Con el análisis de los estudios primarios seleccionados, se observó que en el 23% de ellos se mencionan factores de calidad que no se habían considerado al inicio del estudio; respecto de los aspectos que se habían considerado [9], 21 hablan sobre

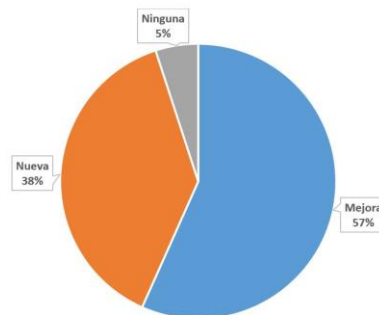


Fig. 2. Innovación en modelos, técnicas o metodologías.

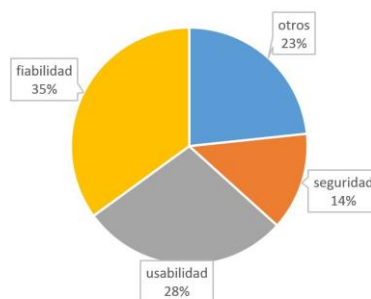


Fig. 4. Factores identificados sobre calidad del diseño en los estudios seleccionados.

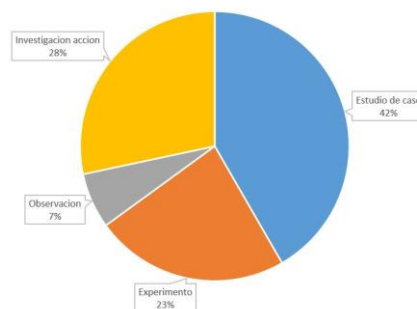


Fig. 5. Utilización de los métodos de validación en los estudios seleccionados.

aspectos de fiabilidad, 17 sobre aspectos de usabilidad, y los 8 restantes se refieren a aspectos vinculados con la seguridad (ver figura 4).

PI06 ¿Cuál es el grado de innovación que poseen los métodos, técnicas o metodologías abordadas en los estudios primarios seleccionados?

En relación con la innovación, se puede observar que se presentaron 23 métodos, técnicas o metodologías que proponen nuevos modelos enfocados en el diseño software con el fin de hacer una evolución a nuevas ideas, por otro lado, 34 artículos

plantearon propuestas de mejora respecto a dichos métodos, técnicas o metodologías existentes; finalmente notamos que 3 publicaciones se enfocaban a la difusión de la investigación, en particular, dar a conocer ciertas técnicas en el ámbito de diseño software.

PI07 ¿Cuáles han sido los métodos de validación utilizados en la última década por los estudios primarios vinculados con el área de Diseño Software?

En cuanto a los métodos de investigación utilizados como mecanismos de validación empírica [10] utilizados en los estudios relacionados a la calidad, se pudo identificar que los Estudios de Caso (25 estudios) son los más recurridos para evaluar y analizar el desempeño de las propuestas en modelos o técnicas existentes al diseño software, ya que gran parte de los artículos se enfocan a la mejora modelos o técnicas y la evaluación de los modelos o técnicas existentes. El segundo método más usado es el de Investigación acción (17 estudios), se explica debido a que gran parte de los artículos buscan comprobar las mejoras que han desarrollado en aspectos de patrones, arquitecturas, etc. En tercer lugar identificamos el método Experimental (14 Estudios) y en cuarto lugar el método de la Observación (4 Estudios). La figura 5 ilustra el uso de los métodos de validación recurridos en los estudios primarios seleccionados.

PI08 ¿En qué contexto han sido desarrollados los estudios primarios vinculados con el área de Diseño Software durante la última década?

Finalmente, en relación con el contexto en el cual se desarrollan los estudios primarios, la categoría identificada con mayor frecuencia, fue la industria en el 58% de los estudios; el contexto académico fue citado en el 25%, y nos llama la atención que el 17% restante no identifica el contexto bajo el cual se desarrolló el estudio. Posiblemente los resultados expliquen el que gran parte de los artículos son propuesta de mejora para realizar cambios en la Industria.

6. Conclusiones

Con el estudio realizado, es posible apreciar, que en la última década se ha mantenido el interés en mejorar los procesos de diseño del software, se pudo identificar que la principal característica al analizar patrones de diseño, radica en el aspecto de la calidad del software, sin embargo, es notorio que gran parte de los estudios se refieren a la tarea de mejorar la creación de los diseños, evitando malas prácticas, es por ello que a los autores les resulta interesante el poder profundizar sobre las características de las estrategias de diseño software para la mejora de aspectos como: complejidad, redundancia, adecuaciones inadecuadas del UML, etc. Por otro lado, también se pudieron identificar ciertos estudios relacionados con los diseños para la automatización, es que aquí donde podemos abordar un nuevo nicho de oportunidades, ya que este término nace alrededor del año 2017 y se tiene muy poco conocimiento sobre el mismo. También se pudo identificar que el Estudio de Caso resulta ser el mecanismo empírico más recurrido para la investigación en el área de Diseño, y que el contexto industrial sigue siendo el entorno de experimentación preferido; ambos resultados coinciden con hallazgos de estudios realizados en

paralelo —por los autores— sobre otras áreas de conocimiento de la Ingeniería de Software [11, 12]. En términos generales, podemos concluir que existen aún problemáticas por abordar en el ámbito del diseño software, lo cual fortalece el interés de los autores por continuar —p.e. con una RSL— profundizando a corto plazo, en alguno de los aspectos identificados en el presente estudio.

Referencias

1. Crosby P.B.: *Quality Is Free*, McGraw-Hill (1979)
2. Humphrey, W.: *Managing the Software Process*, Addison-Wesley (1989)
3. Kan, S.H.: *Metrics and Models in Software Quality Engineering*, Addison Wesley (2002)
4. Gooma, H.: *Software Modeling and Design* (2011)
5. Genero, M. Cruz-Lemus, J., Piattini, M.: *Métodos de investigación en ingeniería de software*, Ed. Ra-Ma (2014)
6. Petersen, K., Feldt, R., Mujtaba, S.: *Systematic Mapping Studies in Software Engineering*. In: *Proceedings 12th International Conference on Evaluation and Assessment in Software Engineering* (2008)
7. Bourque, P., Fairley, R.: *Guide to the Software Engineering Body of Knowledge. SWEBOK V3.0* In: IEEE Computer Society Press (2014)
8. Kitchenham, B., Charters, S.: *Guidelines for Performing Systematic Literature Reviews in Software Engineering*. EBSE (2007)
9. Laporte, C., April, A.: *Software Quality Models, Software Quality Assurance* (2018)
10. Malhotra, R.: *Empirical Research in Software Engineering: Concepts, Analysis, and Applications*. CRC Press (2015)
11. Ávila, D., Aguilar R.: *Research in the Area of Software Quality: A Mapping Study of the Last Decade*. *Proceedings of the 2021 Mexican International Conference on Computer Science (ENC)* (2021) doi: 10.1109/ENC53357.2021.9534818.
12. García, N., Díaz, J., Aguilar R.: *Investigación sobre técnicas y estrategias de prueba del software: Un estudio de mapeo sistemático sobre la última década*. *Research in Computing Science*, 150(11), pp. 245–264 (2021)

Sistema de detección omnidireccional de gestos corporales para entornos virtuales

Iván Cruz García, Héctor Ortiz Cruz, Marco Antonio Solís López

Instituto Politécnico Nacional,
México

Resumen En el presente documento se aborda el diseño detallado de un sistema de detección de gestos para un entorno virtual. Se propone el desarrollo de una plataforma de soporte para caminata omnidireccional en conjunto con un arnés de seguridad aunado a un sistema de clasificación que utiliza una técnica de Time Series Classification, que se vale de un traje de sensores inerciales para identificar los gestos del usuario, las lecturas de los mismos serán tratadas mediante un filtro de Kalman y finalmente se transmiten las señales a un entorno virtual mediante el uso de un dispositivo HID.

Palabras clave: Clasificación, entorno virtual, gestos, inmersión, interfaz, omnidireccional, realidad virtual.

Omnidirectional Body Gesture Recognition System for Virtual Reality Environments

Abstract. This paper presents the detailed design of a gesture detection system for a virtual environment. The proposed system includes the development of a support platform for omnidirectional walking combined with a safety harness, alongside a classification system that uses a Time Series Classification technique. The system leverages an inertial sensor suit to recognize user gestures. The sensor data is processed using a Kalman filter and subsequently transmitted to a virtual environment through a Human Interface Device (HID).

Keywords: Virtual environment, gestures, immersion, interface, omnidirectional, virtual reality.

1. Introducción

Actualmente, la mayoría de las interacciones de un usuario con un entorno virtual están limitadas al uso de un control remoto, lo cual implica el movimiento únicamente de las manos, específicamente de los dedos. Con el desarrollo por parte de consolas comerciales en los últimos años utilizando cámaras y mandos

inteligentes se ha perseguido una nivel cada vez mas avanzado de inmersión, y con los recientes intentos de realidad virtual en conjunto con visores hay un área de oportunidad de desarrollo para esta tecnología. Sin embargo, para lograr una experiencia inmersiva, los gestos motrices y/o movimientos del cuerpo completo deberían verse directamente reflejados en el entorno virtual. En las propuestas tecnológicas antes mencionadas, aunque el usuario puede moverse en un espacio específico, su desplazamiento en realidad está condicionado a que el usuario emule la marcha manteniéndose en su lugar (interpretando el movimiento de las piernas como tal y no como un gesto de caminata natural, en el caso de cámaras), o el desplazamiento obedece a movimientos de las manos (como en los VR Headsets).

Por otro lado, los alcances de la realidad virtual inmersiva en ramas como turismo, cultura y educación son cada vez más prometedores. Los posibles escenarios van desde la visita de museos, hasta apoyo a los modelos educativos actuales con la implementación de tecnología 4.0 que permita al usuario expandir los límites de su experiencia virtual.

Para atender estas circunstancias en esta propuesta de proyecto se plantea diseñar un sistema de detección omnidireccional de movimientos aunado a una plataforma capaz de ofrecerle al usuario mayor libertad para realizar movimientos corporales para su interacción con el entorno virtual, aumentando así el nivel de inmersión y la experiencia de **realidad virtual**. El sistema de detección conlleva una etapa de adquisición de señales de movimientos corporales, una de interpretación y de clasificación de los mismos, finalizando con la transmisión de estos como comandos al entorno virtual. Así mismo, la plataforma de soporte deberá poder restringir el desplazamiento horizontal del usuario sin limitar sus movimientos.

El presente documento comienza ofreciendo un marco histórico de interfaces con entornos virtuales, además, se consideran algunos de los antecedentes con más impacto en éste proyecto. Se plantea el desarrollo e implementación del proyecto así como las técnicas y algoritmos utilizados, además se enlistan los resultados obtenidos. Finalmente se debaten los aportes de este proyecto al estado del arte.

2. Estado del arte

Desde investigaciones en el campo de la Psicología hasta prototipos para entrenamiento militar, las caminadoras omnidireccionales enfocadas a la realidad virtual inmersiva han sido objeto de estudio para diversas compañías y organizaciones. Dentro de las propuestas tecnológicas existen tres principales sistemas(figura 1).

El primero consiste en un arreglo de dos caminadoras lineales independientes, colocadas de manera perpendicular, de modo que permiten el desplazamiento en cualquier dirección en el plano XY. Estas caminadoras están formadas por rodamientos y motorizadas para corregir la posición del usuario manteniéndolo siempre en el centro. Una de las partes más sobresalientes de esta propuesta

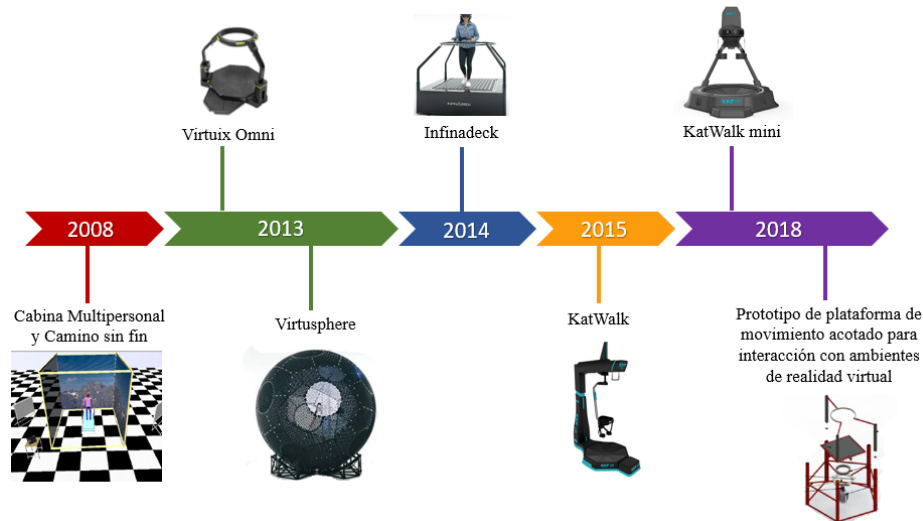


Fig. 1. Línea cronológica del estado del arte.

es el sistema de control, pues para recrear una sensación natural de caminata, la latencia entre el movimiento y la corrección de la posición de la caminadora deben de ser lo más pequeñas posibles, aumentando la complejidad del sistema de control [1].

La segunda propuesta tecnológica consiste en el uso de *wearables*, donde la complejidad del sistema es delegada a sensores colocados en las extremidades del usuario, siendo generalmente giroscopios, acelerómetros y los magnetómetros que detectan los movimientos del usuario para después caracterizarlos, discriminarlos e interpretarlos como comandos. Para mantener al usuario dentro de una zona segura, se utiliza una superficie cóncava donde los pies se ven obligados a regresar al centro por acción gravitatoria combinada con un bajo coeficiente de fricción con la superficie. Dentro de los sistemas comerciales que utilizan este tipo de tecnologías se encuentra el KatWalk VR™. Por último, se han implementado sensores capacitivos dentro de la base que soportará al usuario para medir su posición mientras que los pies se encuentren sobre ella, obteniendo así la posición y velocidad de los pies mientras se encuentren en contacto con la base.

En la tabla 1 se puede observar que la mayoría de estas plataformas están desarrolladas con un enfoque de inmersión en videojuegos o en su defecto en otros países, eso sin mencionar su excesivo costo. Es fundamental en esta propuesta hacer más accesible este tipo de sistemas de inmersión con un propósito más general, con el objetivo de su implementación en un amplio rango de aplicaciones de realidad virtual en nuestro país, además de aportar la detección de las combinaciones de gestos manuales con los del desplazamiento en el entorno.

Tabla 1. Análisis de antecedentes.

Proyecto	Componentes	Contribución	Empresa y año
Cabina Multipersonal y Camino sin fin	Caminadora unidireccional y cabina de espejos con proyector.	Sistema de espejos para proyección de entorno virtual.	CIDETEC, IPN, 2008
Virtuix Omni	Plataforma de soporte para movimiento omnidireccional.	Pionera en movimiento omnidireccional.	Virtuix, 2013
Kat Walk	Plataforma omnidireccional con sistema de detección de caminata.	Plataforma con sistema de detección de movimiento.	Kat VR, 2015
VirtusSphere	Esfera de caminata omnidireccional.	Esfera con sensores inferiores para detectar dirección.	Virtusphere, Inc ,2013
Infinadeck	Caminadora omnidireccional con soporte de usuario.	Caminadora omnidireccional.	Infinadeck, 2014
Prototipo de plataforma de movimiento acotado para interacción con ambientes de realidad virtual	Caminadora unidireccional con motor de redireccionamiento.	Dirección de la mirada y la caminadora alineados por un control PID y un motor.	UPIITA, IPN, 2018

3. Desarrollo

3.1. Plataforma omnidireccional

Se plantea primeramente el uso de una estructura hecha de perfiles unicanal de acero pregalvanizado en conjunto con una superficie cóncava hecha de fibra de vidrio. Se comprobó la funcionalidad de estos materiales para el presente proyecto mediante el uso del software CAD (computer-aided design) *Solid Works*.

En el caso de la plataforma de fibra de vidrio, se necesita que sea capaz de resistir el salto de un usuario y para ello se analizaron los esfuerzos y deformaciones que ocasionaría un impacto sobre la misma. Para ello se realizó un análisis no lineal dinámico. Se utilizó un proyectil de prueba, con un peso igual al del usuario, y se calculó con qué velocidad el objeto tocaría la superficie cóncava. Dicha velocidad se calculó analizando al objeto como un cuerpo en caída libre (ecuación 1), se conocía el peso del objeto (100 kg) y la altura desde la cual caería (20cm ya que fue el salto máximo definido para el proyecto) y un coeficiente de seguridad de tres (figura ??). Con estos datos se puede obtener el

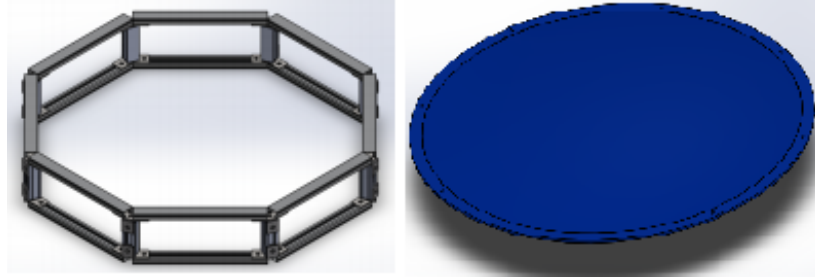


Fig. 2. Estructura de acero pregalvanizado y plataforma de fibra de vidrio.

tiempo de caída (ecuación 2). Por lo tanto, se puede conocer la velocidad con la que se impactaría el cuerpo contra la plataforma (ecuación 3 y ecuación 4).

$$y = y_0 + v_0t + \frac{1}{2}at^2, \quad (1)$$

$$t^2 = \frac{0,2m}{9,81m/s^2}, \quad (2)$$

$$v_f = v_0 + at, \quad (3)$$

$$v_f = 9,81m/s^2(,209). \quad (4)$$

La velocidad es (ecuación 5):

$$v_f = 2m/s. \quad (5)$$

Como se ve en los resultados (figura 3) el impacto del usuario al saltar no sobrepasara el limite elástico de la plataforma, asegurando así que funcionará como superficie de apoyo.

Para la estructura de soporte se necesita de un sistema capaz de resistir el peso del usuario, que no obstruya el movimiento del mismo en ningún momento, pero que también conserve el centro de masa lo más bajo posible y que guarde simetría con los ejes horizontales.

Se propone la construcción de una grúa triple, hecha de perfiles de acero pregalvanizado y ménsulas electrosoldadas. Al analizarla se tomó en cuenta el peso del usuario, el cual actuaría como fuerza vertical en la estructura, la fuerza ejercida por el usuario al generar desplazamiento de manera horizontal y los puntos fijos del sistema, los cuales se encuentran en la parte inferior. Los resultados arrojados por la simulación fueron de $9,32kN/m^2$ (figura 4) actuando en la punta de la grúa, y una deformación máxima de 1.9mm. Con esto se asegura que el material no llegará a su limite de tensión, lo cual asegura que no se romperá, y además no se deformará de una manera considerable, lo cual asegura su utilidad como estructura de soporte.

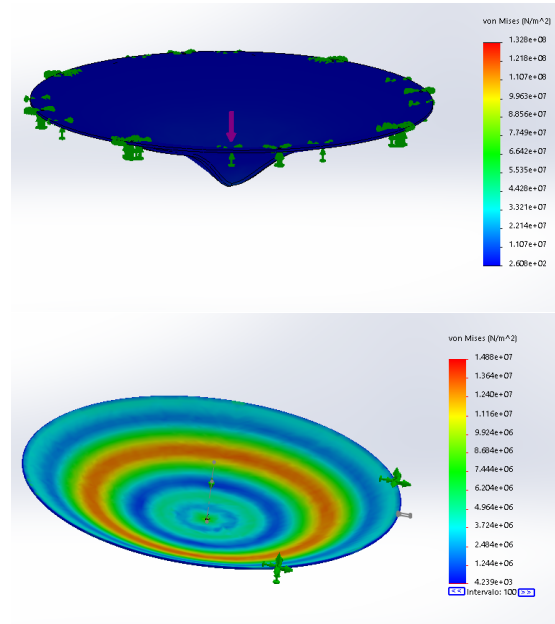


Fig. 3. Deformación máxima debido al impacto del usuario al saltar en la plataforma de vidrio(arriba) y esfuerzo máximo aplicado debido al impacto (abajo).

Finalmente se necesita una estructura que logre limitar el desplazamiento horizontal del usuario, sin limitar su giro y permitiéndole movimiento vertical reducido. Se diseñó un mecanismo que consiste en el uso de un par de tubos cilíndricos doblados, los cuales sujeten al usuario mediante un arnés de seguridad, y se acople a la grúa triple por medio de una plataforma giratoria ubicada en la parte superior de la estructura. Para el diseño del tubo, se tuvo que tener en cuenta las condiciones óptimas para el doblamiento de tubo en frío. Se necesita que el radio del doblado este en un rango de entre 1.5 y 5 veces el diámetro del tubo, ya que se utilizarán dos tubos de 42.1mm y 4.8mm de pared es necesario un radio de doblado entre 63.75mm y 213mm, se eligió un radio de 75mm. Al realizar los análisis de esfuerzos en los tubos se obtuvo un esfuerzo máximo vertical de $9.13 \times 10^7 \text{ N/m}^2$ y un esfuerzo máximo horizontal de $1.1 \times 10^8 \text{ N/m}^2$ (figura 5). Lo cual asegura que los tubos serán capaces de resistir la carga a la que estarán sometidos

3.2. Adquisición de señales de movimientos de los miembros del cuerpo

El proceso de investigación tuvo como objetivo encontrar el tipo de sensores que brindarán la información necesaria para poder medir los movimientos del usuario, buscando opciones que no interfirieran con el movimiento del mismo.

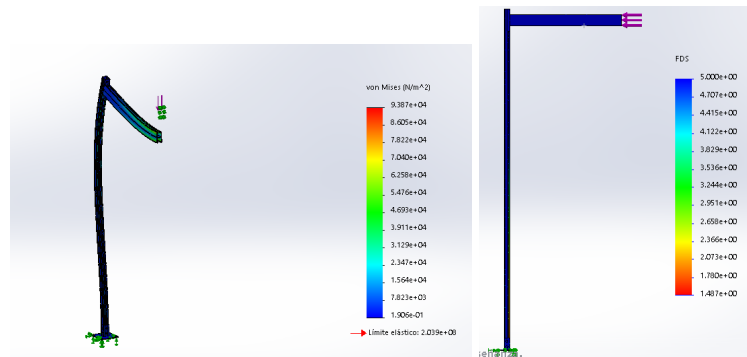


Fig. 4. Análisis de esfuerzo horizontal y vertical en la grúa de soporte de la plataforma.

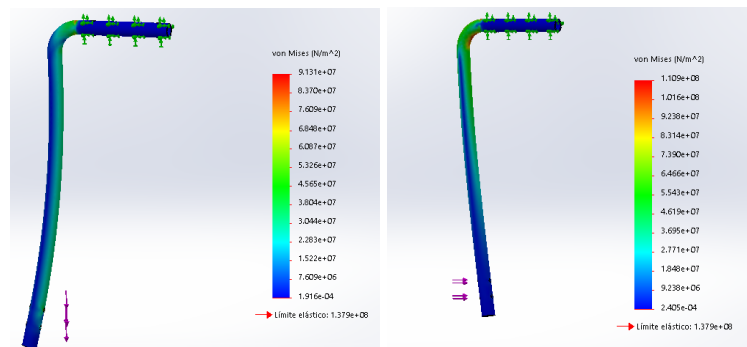


Fig. 5. Análisis de esfuerzo horizontal y vertical en los tubos de soporte para el arnés de la plataforma.

Se plantearon opciones como medición mediante sistemas como exoesqueletos, sistemas basados en visión, y sistemas inerciales. Filippeschi A et al. [4] hacen una reseña de distintos métodos utilizados para la captura del movimiento humano enfocada al miembro superior donde defienden que los sensores inerciales son una opción precisa y barata para la detección de movimientos humanos. Basándose en ésta reseña y otros artículos de investigación [11] [8] que demuestran las ventajas de los IMU sobre otro tipo de sensores, se decidió utilizar sensores inerciales de nueve ejes con acelerómetro, giroscopio y magnetómetro para capturar el movimiento del usuario del sistema. Los mismos envían la información a una unidad central de procesamiento. Para el proyecto se decidió utilizar una tarjeta Raspberry Pi 4, debido a su bajo costo y poder computacional. Se analizaron los puntos de interés para ésta aplicación y se propusieron 14 ubicaciones para la colocación de los sensores (figura 6), creando una red de sensores.

Debido a la naturaleza poco estable de los sensores inerciales, los giroscopios presentan derrapes en sus lecturas y los acelerómetros ruido excesivo, por esto

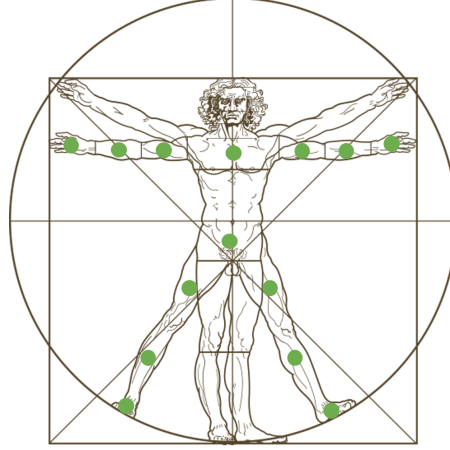


Fig. 6. Puntos de interés para el análisis de movimiento (Verde).

se utilizó un filtro de Kalman para una mejor estimación en las lecturas de los sensores. Las ventajas de este filtro radican en el uso de un modelo matemático, que estime el comportamiento de los sensores, y que al mismo tiempo considere las mediciones actuales [2]. El filtro de Kalman evalúa en cada ciclo el desempeño tanto de la estimación como de las mediciones y decide qué peso darle a cada uno. Se utilizaron las ecuaciones descritas en [9] para la implementación del filtro. La ecuación de estados utilizada (conocida comúnmente como matriz A) es:

$$A = \exp\left(\frac{T}{2}\omega\right), \quad (6)$$

donde ω es la matriz de velocidad angular, la cual puede ser calculada con las mediciones actuales del giroscopio:

$$\omega = \begin{bmatrix} 0 & -w_z^i & w_y^i & 0 & 0 & 0 \\ w_z^i & 0 & -w_x^i & 0 & 0 & 0 \\ -w_y^i & w_x^i & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -w_z^i & w_y^i \\ 0 & 0 & 0 & w_z^i & 0 & -w_x^i \\ 0 & 0 & 0 & -w_y^i & w_x^i & 0 \end{bmatrix}, \quad (7)$$

$w_x^i w_y^i w_z^i$ son respectivamente las mediciones de cada eje de los giroscopios. Para el ruido de procesamiento se utilizó el ruido estimado de los giroscopios, ya que el cálculo de la orientación estimada depende de estos sensores, mientras que el ruido de medición se estima del ruido generado por los acelerómetros y magnetómetros respectivamente. Cabe recalcar que dichos valores, aunque se conocía un estimado gracias a las hojas de datos de los sensores, fueron calculados de manera experimental, aumentando y reduciendo cada uno hasta encontrar los valores óptimos.

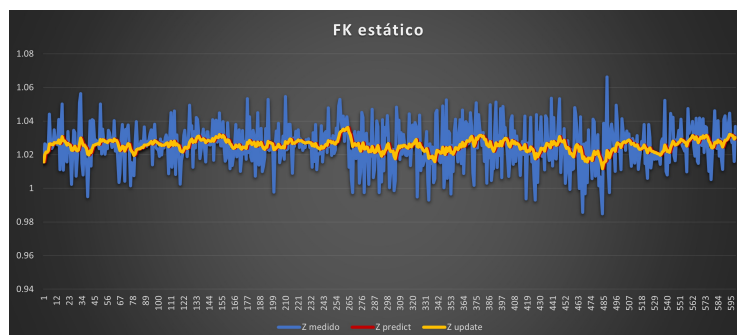


Fig. 7. Mediciones acondicionadas por el filtro de Kalman con el usuario estático.

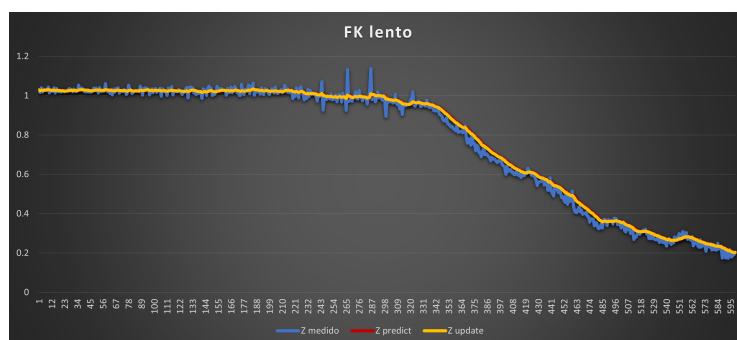


Fig. 8. Mediciones acondicionadas por el filtro de Kalman con el usuario realizando movimientos lentos.

Se probó el desempeño del filtro con mediciones de 3 casos diferentes, el sensor puesto sobre una superficie plana y sin movimiento (figura 7), el sensor siendo movido lentamente (figura 8) y el sensor siendo girado a través del eje x en múltiples ocasiones y de manera rápida (figura 9), siendo éste último el caso óptimo para la aplicación.

El usuario vestirá una red de sensores utilizando un arnés modificado, extendido a las extremidades del cuerpo (piernas y brazos) que se podrá usar sobre ropa ligera, donde se colocarán los IMUs utilizando un sujetador por cada IMU.

3.3. Adquisición de Señales de gestos Manuales

Para la adquisición de señales de gestos manuales se trabajó con la propuesta tecnológica de Shull et al. [7] que consiste un arreglo de sensores barométricos modificados colocados al rededor de la muñeca para la detección de gestos manuales. Un sistema como éste, además de no ser invasivo ya que deja libre el movimiento de los dedos, toma ventaja de la precisión de los barómetros para

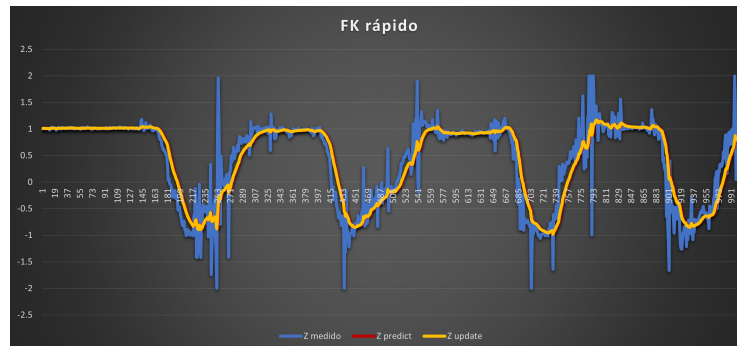


Fig. 9. Mediciones acondicionadas por el filtro de Kalman con el usuario realizando movimientos rápidos (caso ideal).

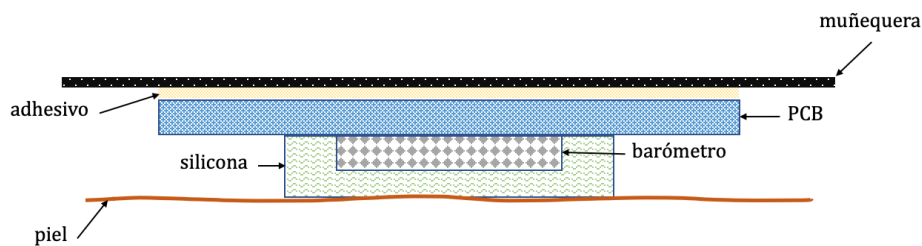


Fig. 10. Corte transversal de un barómetro modificado sujeto a la piel del usuario usando una muñequera.

medir la presión sobre su transductor y la utiliza para medir la tensión en los tendones de la muñeca encargados de articular los movimientos de la mano, alcanzando una precisión de hasta 94 % clasificando 10 gestos distintos.

Al igual que Shull et al. el proceso de la modificación de los barómetros se basa en el trabajo publicado por Tenzer et al. [13]. El objetivo es encapsular el transductor del barómetro en silicona para que en vez de que se detecte la presión atmosférica, al poner el sensor en contacto con la piel sobre los tendones de la muñeca, se detecte la presión ejercida sobre la silicona (figura 10). El diagrama de conexiones eléctricas entre el sistema de adquisición de señales de gestos manuales y la Unidad de procesamiento central incluye también el uso de multiplexores (figura 11). Estos sensores se adherirán en una muñequera ajustada con Velcro® en ambas manos del usuario.

La distribución de los sensores se eligió basándose en las pruebas hechas por Shull et. al, donde se encontró que distribuyendo los sensores en estas posiciones se obtienen resultados de hasta 94 % de precisión (figura 12).

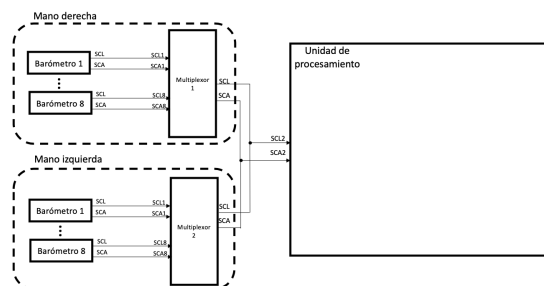


Fig. 11. Diagrama de conexiones eléctricas entre el sistema de adquisición de señales de gestos manuales.

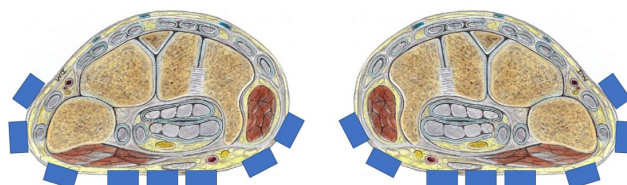


Fig. 12. Ilustración los puntos de contacto de los sensores con la muñeca izquierda y derecha. [3]

3.4. Traducción de datos a movimientos referenciados

Una representación matemática del cuerpo humano permitirá restringir las mediciones de los sensores IMU a ser interpretadas dentro de los límites físicos del cuerpo humano, previniendo así interpretar posiciones que serían imposibles de hacer por un ser humano, que podrían darse debido a errores en la medición.

En el 2004 Rong Zhu et al. [12] propusieron un método para obtener la posición de las partes del cuerpo basándose en el modelo matemático del cuerpo humano de Hanavan [6], el modelo fue modificado para aproximar los segmentos del cuerpo a líneas (figura 13).

Una vez optimizada la señal proveniente de los IMU se necesita traducirlas a las orientaciones relativas de los sensores entre ellos. Para esto se crearán cadenas cinemáticas conformadas por cada segmento del cuerpo como se propuso. El objetivo de esta referenciación es obtener las coordenadas finales de los miembros del cuerpo (manos, pies). Para esto fue fundamental ahondar en los métodos de estimación de orientación existentes para sensores inerciales. El algoritmo de Kalman provee una fusión de sensores para compensar las fallas de los magnetómetros y acelerómetros con las mediciones de los giroscopios, por esto, la proyección de la gravedad y de el norte geomagnético de la tierra son las herramientas a la mano para poder calcular la orientación de cada IMU. Para calcular la orientación de un marco de referencia con respecto a otro puede ser parametrizado utilizando una matriz de rotación entre el sistema coordenado de

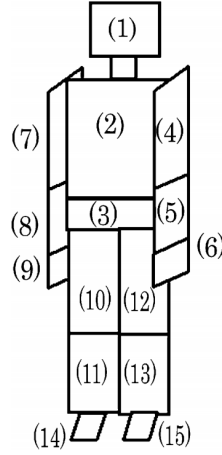


Fig. 13. Cuerpo humano recreado con 15 segmentos para su uso en el modelo matemático.

ejes ortogonales, a esta matriz de rotación se le conoce como matriz de cosenos directores (DCM). Normalmente se requiere conocer la orientación de un marco de referencia del cuerpo (BRF) y un marco de referencia fijo (FRF), en la mayoría de los casos de referenciación de orientación y dirección (AHRS por sus siglas en inglés) el marco de referencia fijo es el ofrecido por la tierra, usando la gravedad y el campo magnético, a este marco de referencia se le conoce como sistema NED (North, East, Down), y con base en este marco se calcula la orientación relativa del cuerpo a observar. Los métodos de estimación más comunes son la estimación con cuaterniones, el método TRIAD y el uso de ángulos euclidianos. Para éste proyecto conviene una parametrización en la que se calcule la orientación relativa entre dos marcos de referencia, ya que la cadena cinemática nos dará finalmente la posición final del miembro del cuerpo. Debido a esto, se eligió la parametrización del método TRIAD. En el método TRIAD la DCM puede determinarse de al menos dos vectores observados [10]. Teniendo dos pares de vectores unitarios escritos como r_1 y r_2 en el marco de referencia fijo y b_1 y b_2 en el marco de referencia del cuerpo y por definición:

$$b_1 = Cr_1, b_2 = Cr_2, \quad (8)$$

donde C es la DCM. Esta ecuación puede ocuparse para determinar C de dos pares de vectores, sin embargo, la solución no es única, por lo que una solución balanceada se puede encontrar en la ecuación 9

$$C_{T3} = b_+ r_+ + b_+ r_+ (b_+ \times b_-)(r_+ \times r_-)^T, \quad (9)$$

donde:

$$r_+ = \frac{r_2 + r_1}{|r_2 + r_1|}, r_- = \frac{r_2 - r_1}{|r_2 - r_1|}, \quad (10)$$

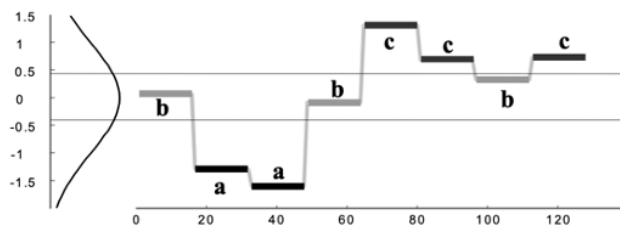


Fig. 14. Discretización realizada por el SAX.

$$b_+ = \frac{b_2 + b_1}{|b_2 + b_1|}, b_- = \frac{b_2 - b_1}{|b_2 - b_1|}. \quad (11)$$

Posteriormente, teniendo los dos vectores optimizados de g y H de los IMU_i e IMU_{i+1} es fácil notar que son correspondientes con el planteamiento del problema de estimación de orientaciones de un marco de referencia relativo entre segmentos del cuerpo humano, por esto se usó esa solución para proponer la matriz de rotación entre esos dos marcos de referencia con el método TRIAD, obteniendo matrices que fueron validadas en Matlab con diferentes pruebas donde se crearon cadenas cinemáticas para verificar los ejes de rotación y los ángulos relativos de los IMU.

Finalmente, con las matrices de rotación relativas entre los segmentos consecutivos de las cadenas cinemáticas del cuerpo humano se propusieron matrices de transformación y se verificó con coordenadas reales del traje de sensores.

3.5. Clasificación de movimientos

En esta sección se abordarán dos Sistemas de clasificación de movimientos, el Sistema de Clasificación de Movimientos Corporales (De ahora en adelante SCMC) y el Sistema de Clasificación de Gestos Manuales (De ahora en adelante SCGM).

Para el SCMC se utilizará un algoritmo de Time Series Classification, el cual constará de una algoritmo SAX (Aproximación de agregación simbólica por sus siglas en ingles), encargado de destemporalizar las medidas obtenidas de la referenciación y un clasificador 1-NN (Vecino mas cercano por sus siglas en inglés) para identificar de qué movimiento se trata. El algoritmo SAX discretiza las señales obtenidas definiendo límites a partir de la media y desviaciones estándar de la curva (figura 14), les asigna un valor simbólico (en este caso letras) y las almacena dentro de un vector cada cierto tiempo. Posteriormente se realiza un barrido de dicho vector para determinar cuántas veces se repitieron secuencias de valores simbólicos iguales y dicha información la almacena en una matriz de frecuencia (para este proyecto de 9x9, (figura 15)). Por otro lado, el clasificador 1-NN toma múltiples muestras de las matrices de frecuencia, creadas por el SAX, y crea una nueva matriz cuyos índices serán el promedio de los

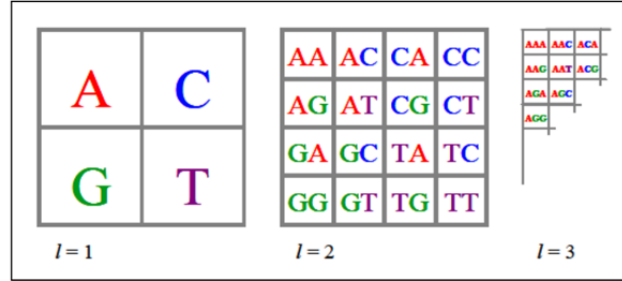


Fig. 15. Creación de matrices de frecuencia.



Fig. 16. Sujeto de pruebas realizando las acciones: agachado, correr y cambio de herramienta. Para el entrenamiento del algoritmo.

índices de todas la muestras obtenidas, a estas matrices se les conoce como centroides y existe un centroide por cada clase a clasificar. A este proceso se le conoce como .entrenamiento” del algoritmo, y para llevarlo a cabo se utilizó el traje de sensores y un sujeto de prueba. El entrenamiento consistió en el sujeto utilizando el traje y realizando continuamente durante 2 minutos uno de los movimientos a clasificar(figura 16), todas las matrices arrojadas por el SAX se almacenaban y procesaban para calcular el centroide correspondiente. Este proceso se repitió con cada uno de los movimientos a clasificar los cuales fueron: agacharse, saltar, correr, picar, usar, colocar, cambio de herramienta y permanecer parado. Finalmente, cada muestra nueva calculada por el SAX se compara con cada uno de los centroides calculados utilizando la distancia euclidiana (ecuación: 12). Aquel centroide al que la muestra es mas cercana es designado como su clase correspondiente (figura 17):

$$D_{ij} = \sqrt{\sum_{k=1}^n (x_{ki} - x_{kj})^2}. \quad (12)$$

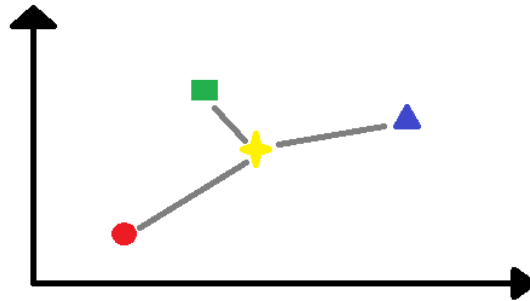


Fig. 17. Distancia a los centroides (modelo de 2 coordenadas).

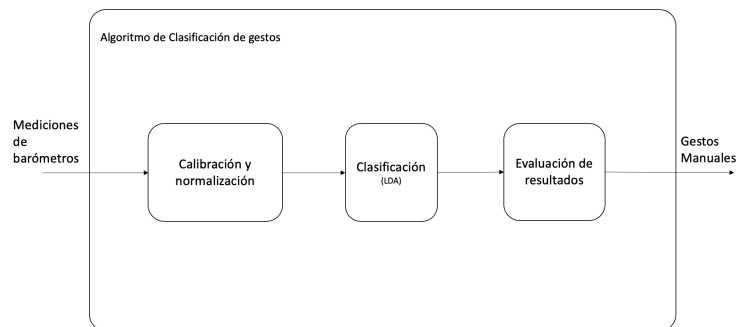


Fig. 18. Algoritmo de clasificación de gestos implementar.

Para el SCGM se desarrolló un algoritmo de clasificación parecido, también 1-NN, pero no se implementó el SAX en este caso pues no se necesitan clasificar movimientos sino solo gestos (figura 18). Al realizar las mediciones, la información de presión de los barómetros de ambas manos es ordenada en dos vectores para su clasificación.

El clasificador utilizado es un 1-NN. El entrenamiento supervisado se realizará midiendo la presión en los barómetros durante 3 minutos cuando la mano del usuario está abierta, cerrada y relajada. Finalmente se calculan los centroides de cada una y se comparan las nuevas mediciones con ellos. La implementación del clasificador seguirá la metodología propuesta por Ghogh B. y Crowley M. [5].

El funcionamiento integral del sistema de traducción de comandos puede representarse con el diagrama a bloques de la figura 19.

3.6. Transmisión de comandos

Teniendo los gestos del usuario identificados se procederá a enviar estos gestos definidos con su correspondiente traducción en el entorno virtual de

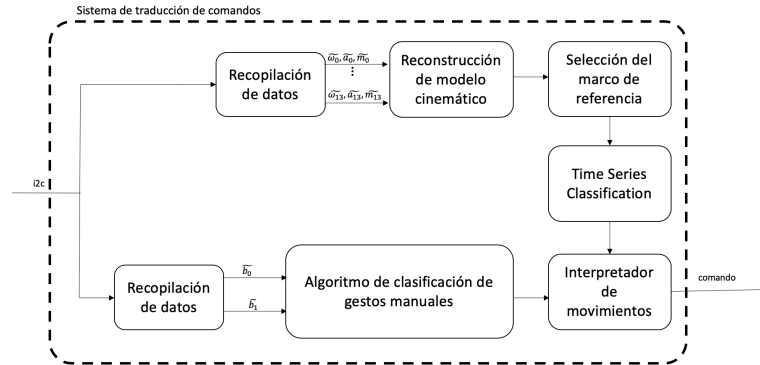


Fig. 19. Diagrama a bloques del sistema de traducción de comandos.

Tabla 2. Equivalencia entre las acciones a identificar, comando HID en el gamepad, comando AT enviado al modulo bluetooth y su boton respectivo en el entorno virtual.

Clase	Comando HID	Comando BLE	X-box
Parado	<i>else</i>	AT+BLEHIDGAMEPAD=0,0,0x00	-
Correr	arrow_up	AT+BLEHIDGAMEPAD=0,-1,0x00	Adelante
Saltar	a_button	AT+BLEHIDGAMEPAD=0,0,0x01	a
Colocar	b_button	AT+BLEHIDGAMEPAD=0,0,0x02	b
Usar	left_trigger	AT+BLEHIDGAMEPAD=0,0,0x08	x
Atacar	right_trigger	AT+BLEHIDGAMEPAD=0,0,0x10	y
Agachar	right_button	AT+BLEHIDGAMEPAD=0,0,0x40	Lb
C. Herr.	right_stick	AT+BLEHIDGAMEPAD=0,0,0x80	Rb

pruebas. Y esto se hace mediante el uso de un módulo BLE (Bluetooth Low Energy) el cual puede interactuar con el entorno utilizando el protocolo de comunicación HID lo que le permite ser identificado como un Gamepad genérico.

El modo de operación que permitirá enviar la información deseada está descrito en el manual de usuario como “AT+BLEHIDGAMEPAD”, el cual envía una carga de datos específica de un HID gamepad por BLE.

Los parámetros de los movimientos naturales traducidos a comandos se almacenan en variables x y y para registrar el movimiento en esa dirección, tiene además, botones con direcciones 0x00-0xFF, lo cual sirve como máscara para hasta 8 botones. Esta información se envía a módulo por medio de comandos AT y a su vez el módulo los comunica al entorno virtual el cual los identifica como activaciones de algún botón.

Las equivalencias de las clases identificadas a comandos AT y sus respectivas acciones en el entorno se presentan en la tabla 2.

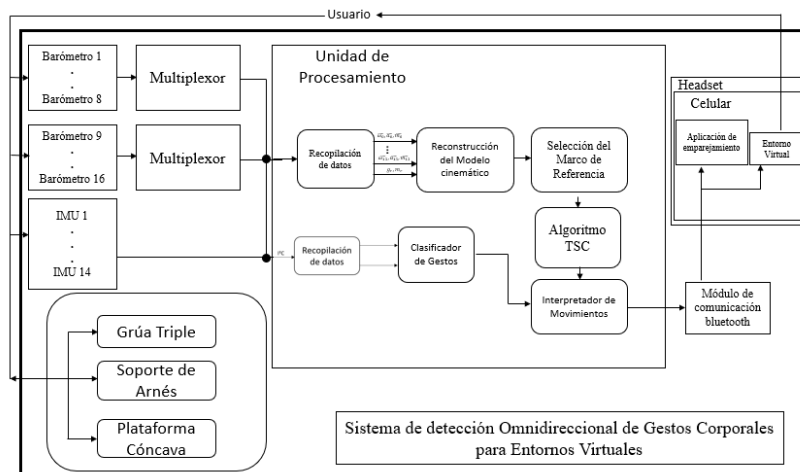


Fig. 20. Diagrama del funcionamiento completo del sistema.

3.7. Sistema completo

En el diagrama de la figura 20 se puede observar la estructura general del proyecto, tanto sus entradas como salidas, y la manera en la que interactúan unos elementos con otros. Se explica desde la manera en la que el usuario influye en el proyecto, también un apartado en el que se menciona la participación de la estructura en el proceso, siguiendo con la adquisición de señales de los sensores, la comunicación a través de I^2C , así como los multiplexores utilizados para la comunicación con los barómetros. Cabe recordar que la unidad de procesamiento mencionada en la cual se integró el proyecto fue una Raspberry Pi 4.

Un diagrama a bloques del funcionamiento y etapas del algoritmo de traducción de movimientos. Posteriormente su interpretación a la equivalencia en comandos para el entorno virtual y su posterior envío al teléfono móvil montado en un headset, para finalmente actuar como la salida por medio de su representación visual para el usuario.

4. Resultados

La integración de los subsistemas diseñados, resulta en un prototipo funcional con las siguientes características:

- La estructura es capaz de resistir las cargas máximas.
- La estructura permite al usuario saltar y agacharse.
- Se logró diseñar un sistema de caminata omnidireccional gracias a la forma de la plataforma elegida que brinda suficiente soporte para el usuario sin limitar el movimiento natural del mismo (figura 21).



Fig. 21. Prototipo de plataforma y sistema de detección integrados.

- Se implementó un algoritmo, basado en la teoría de Kalman, capaz de filtrar de manera confiable las lecturas del traje de sensores y posteriormente interpretarlas como gestos realizados por el usuario.
- El algoritmo de TSC implementado logra identificar 7 de 8 movimientos planteados. Tardando 0.8s en identificar la clase "corriendo", 2.1s "agachando", 1.5s "cambio de herramienta", 0.4s "colocando", "saltando" casi de manera inmediata, 0.8s "picando" 1.1s "parado". Desafortunadamente, durante las pruebas no se identificó la clase "golpear".
- Se logra transmitir los comandos al entorno virtual por medio de protocolo HID (figura 22).

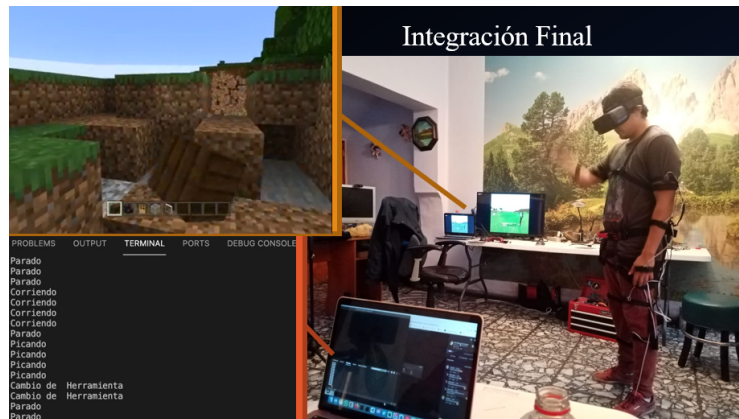


Fig. 22. Sistema integrado y funcional con usuario probando en entorno virtual de Minecraft VR (esquina superior izquierda) y monitoreado a través de la terminal de la Raspberry (esquina inferior izquierda).

5. Abreviaciones y acrónimos

BLE- Bluetooth Low Energy
CAD- Computer Aided Design
HID- Human Interface Device
IMU- Inertial Measurement Unit
I²C- Inter-Integrated Circuit
LDA- Lineal Discrimination Analysis
TSC- Time Series Classification
SCMC- Sistema de Clasificación de Movimientos Corporales
SCGM- Sistema de Clasificación de Gestos Manuales
VR- Virtual Reality

6. Conclusiones

Se desarrolló un sistema de sensores que propone una solución para capturar movimientos de todas las extremidades del cuerpo y los gestos manuales, sin necesidad de que el usuario sostenga con sus manos ningún dispositivo adicional ni del uso de cámaras, esto requirió la indagación en técnicas de clasificación y reveló las ventajas descubiertas en proyectos anteriores sobre el uso de algoritmos de TSC en conjunto con sensores IMU.

Se demostraron las ventajas de estructuras hiper estáticas, como la grúa triple aquí analizada, que brindaran mayor estabilidad debido a su simetría. El sistema de limitación del desplazamiento le permite al usuario simular movimiento omnidireccional sin desplazarse de su lugar. El soporte que se le brinda al usuario podrá resistir su peso total en caso de que éste pierda el equilibrio, por lo que su seguridad no se ve comprometida mientras esté en la plataforma.

Los análisis de esfuerzo realizados en la estructura demostraron la utilidad de los materiales seleccionados, desde la resistencia elástica de la plataforma de fibra de vidrio, así como la seguridad de que los soportes de acero pregalvanizado resistirán las cargas máximas a las que serán sometidos incluso tomando en cuenta un factor de seguridad de tres.

Se implementó un modelo matemático que en conjunto con el filtro de Kalman permite restringir las mediciones de los sensores y crear distintos marcos de referencia para la ubicación de las partes del cuerpo.

Se diseñó un Sistema de Detección de Gestos Manuales usando barómetros modificados que resulta favorable a ésta aplicación, dado que permite detectar los gestos de la mano realizando las mediciones desde la muñeca, dejando así que la mano se mueva de manera libre y natural, además de un sistema interpretador de comandos que permite traducir múltiples movimientos a la vez para enviar los comandos correspondientes a los gestos corporales que el usuario esté realizando. Todos estos algoritmos se ejecutan en una Raspberry Pi 4.

Por último, los comandos son transmitidos al entorno virtual por medio de Bluetooth Low Energy, debido a la estandarización de sus protocolos de envío de datos, mediante el uso de una tarjeta de desarrollo con una modalidad enfocada al envío de comandos compatibles con un HID gamepad, con compatibilidad de emparejamiento con dispositivos móviles.

En lo general, este diseño mecatrónico es una propuesta funcional, que implementa tecnologías multidisciplinarias que trabajan en conjunto. Éste proyecto integra alternativas de solución de diferentes áreas de la ingeniería, ofreciendo al usuario una experiencia de realidad virtual inmersiva. Con ello se comprueba la viabilidad del uso de sensores inerciales como una alternativa eficaz y de bajo costo para la interpretación de gestos corporales. En conjunto con la plataforma omnidireccional y los barómetros modificados, para la interacción entre el usuario y el entorno virtual, demuestran el potencial de esta tecnología en este nicho de desarrollo.

Referencias

1. Caminadora omnidireccional para realidad virtual, productos digitales moviles, <https://www.pdm.com.co/Noticias/2013/Marzo/04/Caminadora-omnidireccional-para-realidad-virtual.html>
2. van Biezen, M.: Lectures in the kalman filter, a
3. D'Antoni, A.: Clinically oriented anatomy, 7th edition, by keith l. moore, arthur f. dalley ii, and anne m. r. agur, baltimore, md: Lippincott williams & wilkins, 2014, 1134 pages, paperback, isbn 978-1-4511-1945-9. Clinical Anatomy, vol. 27 (03 2014) doi: 10.1002/ca.22316
4. Filippeschi, A., Schmitz, N., Miezal, M., Bleser, G., Ruffaldi, E., Stricker, D.: Survey of motion tracking methods based on inertial sensors: A focus on upper limb human motion. Sensors (Basel, Switzerland), vol. 17, no. 6 (June 2017) doi: 10.3390/s17061257
5. Ghogh, B., Crowley, M.: Linear and quadratic discriminant analysis: Tutorial, pp. 17 (06 2019)

6. Hatze, H.: A mathematical model for the computational determination of parameter values of anthropomorphic segments. *Journal of Biomechanics*, vol. 13, no. 10, pp. 833–843 (1980) doi: 10.1016/0021-9290(80)90171-2
7. Jiang, S., Zhu, Y., Zhu, X.: Hand gesture recognition and finger angle estimation via wrist-worn modified barometric pressure sensing. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. PP, pp. 1–1 (03 2019) doi: 10.1109/TNSRE.2019.2905658
8. Jirattigalachote, W., Hunt, M., Cutkosky, M., Delp, S.: Quantified self and human movement: A review on the clinical impact of wearable sensing and feedback for gait analysis and intervention. *Gait & Posture*, vol. 40 (05 2014) doi: 10.1016/j.gaitpost.2014.03.189
9. Manon Kok, J. D. H., Schön, T. B.: Using inertial sensors for position and orientation estimation, <https://www.tecnocurve.es/blog/2019/03/26/curvado-de-tubo-la-importancia-del-radio-de-curvado/>
10. MARKLEY, F.L., N.: Attitude determination using two vector measurements, http://www.archive.org/details/nasa_techdoc_19990052720
11. Patel, S., Park, H.-S., Bonato, P., Chan, L., Rodgers, M.: A review of wearable sensors and systems with application in rehabilitation. *Journal of neuroengineering and rehabilitation*, vol. 9, pp. 21 (04 2012) doi: 10.1186/1743-0003-9-21
12. Rong Zhu, Zhaoying Zhou: A real-time articulated human motion tracking using tri-axis inertial/magnetic sensors package. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 12, no. 2, pp. 295–302 (2004)
13. Tenzer, Y., Jentoft, L., Howe, R.: The feel of mems barometers: Inexpensive and easily customized tactile array sensors. *Robotics & Automation Magazine, IEEE*, vol. 21, pp. 89–95 (09 2014) doi: 10.1109/MRA.2014.2310152

Analysis and Comparison of a Classification of Regular Metal Pieces by Convolutional Neural Networks with and without Principal Component Analysis

Mayra Mateo Jiménez¹, Carlos Eduardo Torres Reyes²,
Everardo Efrén Granda Gutiérrez¹, José Arturo Pérez Martínez²

¹ Universidad Autónoma del Estado de México,
Centro Universitario UAEM Atlacomulco,
Mexico

² Universidad Autónoma del Estado de México,
Unidad Académica Profesional Tianguistenco,
Mexico

`mmateoj481@alumno.uaemex.mx,`
`{cettorresr, eegrandag, japerezm@uaemex.mx}`

Abstract. The classification of images using artificial intelligence techniques (such as neural networks) is fundamentally related to the characteristics of images, such as having a background of a single uniform color. If the illumination scale of the images is different, or the resolution of the image is not sufficient, this may cause false positives when classifying. This work analyzes the importance of classifying images of metal parts (with regular shapes, such as circle, square, rectangle, and flat washer, using a convolutional neural network with and without principal component analysis (PCA), to determine the percentage of classification of images in comparison with the two algorithms by evaluating sensitivity, specificity, accuracy, and precision. The methodology is divided into three steps: obtaining the dataset, training the convolutional neural network, and validating its results. Some representative results of a comparison with a convolutional neural network without PCA were obtained using the RMSProp optimizer with 85.3% precision. PCA achieved the accuracy 98.7%, which indicates that the implementation of PCA improved the classification, with 480 components used.

Keywords: Brightness, convolutional neural network, principal component analysis, classification.

1 Introduction

The regular figures classification is a process where it is necessary to implement pattern recognition algorithms. To carry it out, a wide variety of artificial intelligence techniques are available, and these can be catalogued according to the way in which the data are processed; for example, methods such as deep learning, random forest, and KNN (K-nearest neighbors) [1].

Image processing requires homogeneity, otherwise, loss of important characteristics of the patterns is generated. These problems are caused by certain characteristics of the images, such as their size, width, and length; image noise such as brightness, Gaussian noise, salt and pepper noise, or the resolution. It is difficult to consider all types of noise at the same time in the same job. If the image is not well-defined, the edges or lines of a specific figure cannot be detected. It is important to consider the number of images contained in a dataset. Using many images requires considerable memory storage that is generated due to the size of the dataset. [4].

A pattern detection application focuses on classifying handwritten digits, images, or hyperspectral objects. However, no studies have been done that are related to the classification of metal parts with respect to brightness variation, using algorithms to implement principal component analysis (PCA) and convolutional neural networks [3].

One machine learning application is pattern identification. In manufacturing field, it is necessary in some cases to identify faults in the metal parts produced, such as the manufacture of plugs for lights and the identification of rusted parts to automatically replace them. Using convolutional neural network and vector support machine (SVM) algorithms for the classification of satellite images, a precision value of 83.33% was obtained [6]. In the work of [15], using a convolutional neural network for pattern recognition in images, an accuracy of 92% was obtained.

The PCA algorithm and convolutional neural networks (CNN) belong to the set of deep learning methods because they collect data in a hierarchical way, creating abstract models [5]. However, for most algorithms, inputs for information compression can be data sets with distinct characteristics [2].

In addition to the vector support machine algorithm, PCA can be used for the analysis of data distribution within a Cartesian plane [7, 8].

The extraction of local features is done at high resolution, which minimizes the loss of information. In addition, an alternative implementation is to use a combination of low-resolution images to use more complex functions, if necessary [9, 10].

This type of architecture consists of a variable-length input, which refers to the data entry. In the case of a string or a sentence that has 12 words, the created instances must each be represented by a variable because of its data input of data, in which an internal operation related to the memory of the calculation must be applied. This process is known as a hidden layer [11].

Convolution is the basic component of convolutional neural networks. It consists of an operation between input matrices and filters (kernel), and it results in a feature map. The convolution operation of two $M \times N$ arrays, with an input image $I(x, y) * k(x, y)$, is defined as follows [17, 18]:

$$C[x, y] = I[x, y] * K[x, y] = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I[m, n] K[x - m, y - n] \forall x$$

$$= 1, 2, 3 \dots, \text{Image width}, y = 1, 2, 3 \dots, \text{Image height}, \quad (1)$$

where:

I : input image matrix,

K : kernel,

$\frac{1}{MN}$: normalizing image size,

M, N : matrix of image filters,

x, y : image width and height.

The choice of evaluation method depends on the operation of the algorithm and the results to be obtained. The equations for the metrics are as follows [8]:

$$\text{Precision} = \frac{TP}{TN+FP}, \quad (2)$$

$$\text{Accuracy} = \frac{TP+TN}{TN+TP+FP+FN}, \quad (3)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN}, \quad (4)$$

$$\text{Specificity} = \frac{TN}{TN+FP}, \quad (5)$$

where:

n corresponds to the total number of images.

TP: True Positives,

TN: True Negatives,

FP: False Positives,

FN False Negatives.

Therefore, the point of interest in this work is the percentage obtained from the performance evaluation metrics of the convolutional neural network algorithm with and without PCA.

2 Experimental Set-up

Convolutional neural networks, after being properly trained, can adequately classify classes due to feature extraction. Implementing PCA can improve the disadvantages in processing images. One of them is the extraction of characteristics of the images, avoiding redundancy. This in addition to compressing them, conserving the greatest amount of information. For this reason, we propose comparing the results obtained in the implementation of a convolutional neural network with and without the analysis of the main components in the classification of images of metal parts. Another complication of object detection is that they reflect brightness. We intend to evaluate this issue by pre-treating the image and standardizing the brightness values in the images. The images are captured with an industrial-type 1.3 megapixel camera.

The first stage of this work consists of the construction of the dataset, then the training process of the convolutional neural network, and finally the testing stage. From the batch of images processed in each of the stages, 60% are used for the training stage, 20% for the validation set, and 20% for the test set, for a total dataset of 800 images.

The computer resources utilized for the development of this project are 4GB of RAM, AMD Quad-Core Processor @ 1.00 GHz (up to 1.4 GHz), an industrial-type camera of 1.3 megapixels, camera features (European machine vision association (EMVA), 2006; Model: Basler acA1300-200um). The programming language used was Python 3.8. The stages of data processing are described below, as well as some of the development and performance evaluations and the confusion matrix.

2.1 Dataset

For the training and validation images there are four classes being considered: square, rectangle, circle, and flat washer. Each of the folders contains 120 images, 10 images were taken according to each of the classes (regular pieces) and considering the brightness levels (2000lx, 1500lx, 1000lx, 500lx, 100lx). For the training folder, 10 images of the same images per piece were also obtained. Regarding the distribution of images for each class within the folder, 60% of the total images per piece (equivalent to 30 images per class) were considered for this stage, for a total of 120 images per folder (Table 1).

To obtain the images of the validation folder, 10 images were taken for each of the pieces (Circle, Rectangle, Square, Flat washer). Also, the brightness levels of (2000lx, 1500lx, 1000lx, 500lx, 100lx) were considered; the division is shown in Table 1, according to the two illumination sources: a fluorescent lamp and a light bulb.

This folder contains four classes, which correspond to each of the figures. In total there are 40 images, that is, 10 images of each piece. Another 40 images are contained in the test folder, which is itself divided into two folders in different directories, to corroborate that the classification is correctly done. The number of images corresponds only to one type of brightness source to obtain the same number of samples with respect to the other, thus generating a total of 80 images.

Some images are seen in other rectangular formats. In consideration of this, it is convenient to choose a specific size, which was initially 150 x 150 pixels. When images are being captured with a low intensity of light, such as at scales of 500 and 100 lux, result them too dim, which makes it difficult for the neural network to classify them. Because of this, it is necessary to adjust the brightness.

Brightness adjustment can be done with image pre-processing so that the grayscale images that exceed the threshold calculated according to the percentiles of the images are normalized. This algorithm can also determine which images do not need adjustment. If they do not need it, the algorithm is not applied to them to prevent from increasing in brightness (Figure 1). Some images within the upper levels, such as 2000 lx, 1500 lx, or 1000 lx, were discarded for the application of the brightness adjustment because their percentiles were considered high.

According to this number, a comparison is made with the percentiles of the other images, then the out-of-level percentiles are eliminated. The percentiles obtained are then normalized in a level from 0 to 255. New maximum and minimum percentile values are obtained, and as a result, a new image with the appropriate brightness is generated.

Images that have a suitable brightness scale are discarded from the balance preprocessing. The result of the application of brightness balancing is shown in figure 2. The result of the algorithm is shown in each of the brightness levels, which include high levels such as 2000 lx, 1500 lx, and 1000 lx. For the 500lx and 100lx images, the algorithm is applied according to the percentiles obtained, thus improving the image with respect to its brightness. The edges are also more defined after this configuration.

Table 1. Number of images per piece and brightness levels.

Brightness levels	Square	Number of images per piece		
		Circle	Rectangle	Flat washer
2000 lx	10	10	10	10
1500 lx	10	10	10	10
1000 lx	10	10	10	10
500 lx	10	10	10	10
100 lx	10	10	10	10
Total images	50	50	50	50

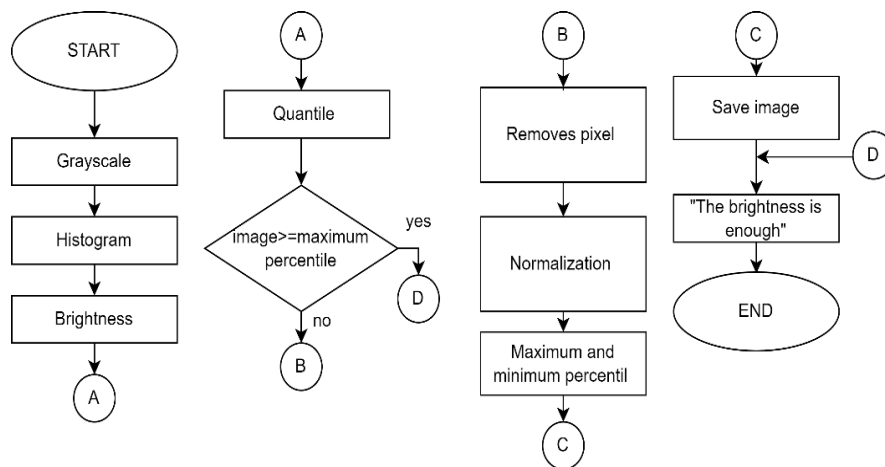


Fig. 1. Brightness balance.

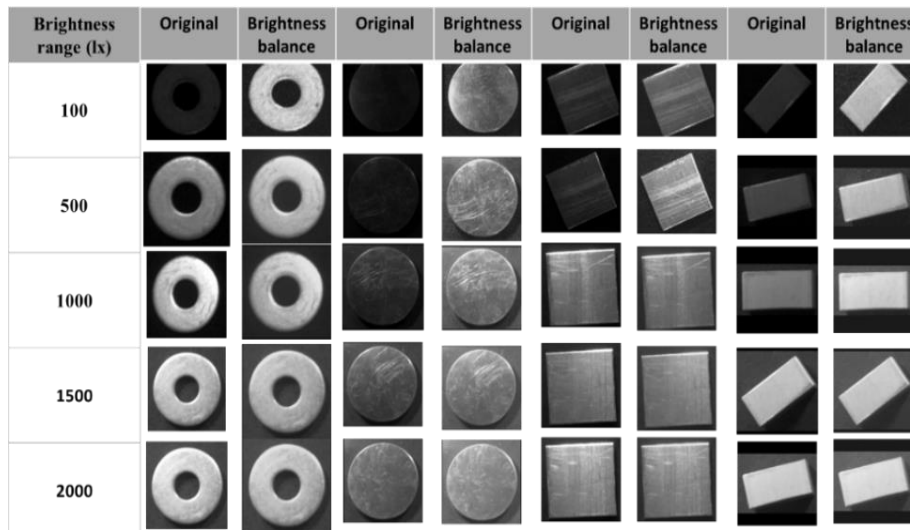


Fig. 2. Brightness balanced samples at each lux level.

2.2 Implementing PCA

The algorithm is divided into three important sections: mean of the data, covariance matrix, and obtaining components. The mean of the data is obtained and subtracted from each of the characteristics (x_1 and x_2), by means of $\frac{x-\bar{x}}{\sigma(x)}$, where x_i is the first extracted feature subtracted from the mean ($\bar{x}$) of the sum of all features in the image. σ refers to the calculation of the covariance, with respect to the characteristics obtained from x . By obtaining a matrix that contains the input images, where from x . Is the set of input images [12, 13].

To acquire an eigenvalue (λ) the equation used is $B \cdot \vec{a} = \lambda \cdot \vec{a}$, where the original matrix (B) has dimension (a square matrix) and an Eigenvector ($\vec{a}$). With the eigenvalues obtained, the total variance of all components is calculated $\sum_{i=1}^p Var(x_i) = \sum_{i=1}^p \lambda_i$, where λ_i they represent the eigenvalues [14, 16].

The eigenvalues and eigenvectors generate the new coordinates corresponding to the components. To reduce the dimensions of the variance obtained, the number of components is calculated from the largest orthogonal variance. The variance consists of the combination of original variables such as the number of pixels, lines, or curves, according to the characteristics of the image.

To calculate a component, the value of the mean of the same variables is subtracted from the variables that make up the image. Thus, an average of zero is achieved in all components. Finally, the covariance of the eigenvectors is calculated. The behavior of the variance distribution can be seen graphically in Figure 3. The calculation of the covariance matrix was performed by using “`cov_matrix = np.matmul(standardized_data.T, standardized_data)`”, which belongs to scikit-learn in the Python language.

The result is a matrix of (784x784) that corresponds to the multiplication of an image size of (28x28) pixels. The result of the calculation of vectors (eigen), by means of “`new_coordinates = np.matmul(vectors, standardized_data.T)`”, is obtained as a result (2, 480), where 2 refers to the x and y coordinates, and 480 is the maximal number of components to use in the implementation of PCA in the convolutional neural network. The optimal number of components levels from 1 to 480. In Figure 4, an approximation with the use of 10 components demonstrate the behavior of the variance.

The result is the generation of a new image with reduced dimensions and non-redundant features, which is stored in a new vector for neural network training. In this case, the convolutional neural network consists of the following features: a convolutional layer with 32 2x2 feature maps, a 2x2 grouping or pooling layer, 64 convolutional layers of 2x2 feature maps, a 2x2 grouping or pooling layer, a 20% dropout, a flattening layer, a connection layer of 784 neurons with a rectifier activation layer, a connection layer of neurons with a *softmax* activation layer, and an-output layer.

The methodology is implemented in Python using Keras and TensorFlow. The metrics used to evaluate the convolutional neural network algorithm with and without PCA are precision (1), accuracy (2), sensitivity (3), and specificity (4). The results obtained are evaluated in a comparison of the accuracy metric, in consideration of the percentage obtained by [15], reaching a percentage greater than or equal to 93%.

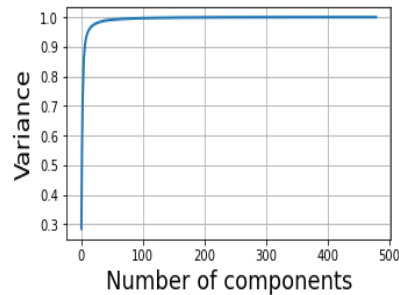


Fig. 3. Number of components in principal component analysis.

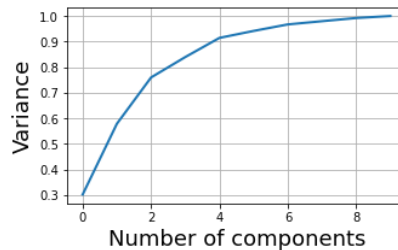


Fig. 4. First 10 components in principal component analysis.

3 Results and Discussion

The following two images are the graphs of the losses obtained for the model without PCA (Figure 6) and with PCA (Figure 7), respectively. In them can see the effect of loss during the training (red line) and validation (green line). In Figure 7, the green line is closer to the red line, that is, the loss is very similar in validation and training.

In the two graphs, it is indicated that the CNN can continue learning and to improve. It can also be observed that the training process stopped prematurely. This effect can be observed at the end of each of the loss curves.

For the execution of the training, the Adam, RMSProp, and Adadelta optimizers were used because they are the most-used optimizers for evaluating neural networks [13]. The results are shown in Table 2. The best percentages are accuracy: 98.8%, sensitivity: 98.8%, accuracy: 98.7%, which were achieved by the implementation of the convolutional neural network with an analysis of main components, using the RMSprop optimizer.

The confusion matrices were obtained using the scikit-learn metrics. The values were obtained automatically in the distribution of the matrix. Results of the classification of the 80 images by class are shown in Figure 8; each one is identified with the name of the algorithm to which it belongs. The labels correspond to 0: circle, 1: square, 2: rectangle, 3: flat washer. The correct classification of true positives is in the white color box in Figure 8, where the highest concentration of classifications

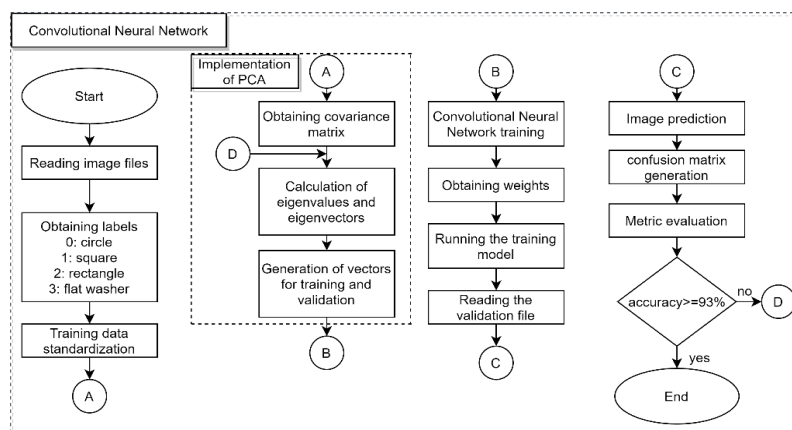


Fig. 5. Implementation of PCA in the neural network.

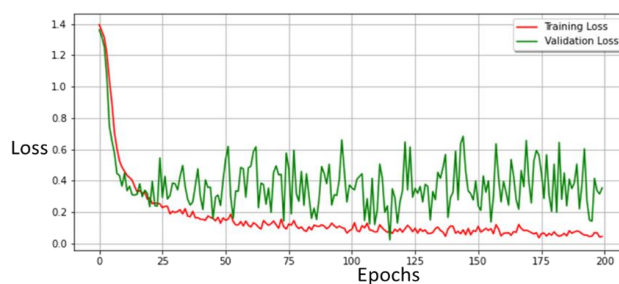


Fig. 6. Training CNN.

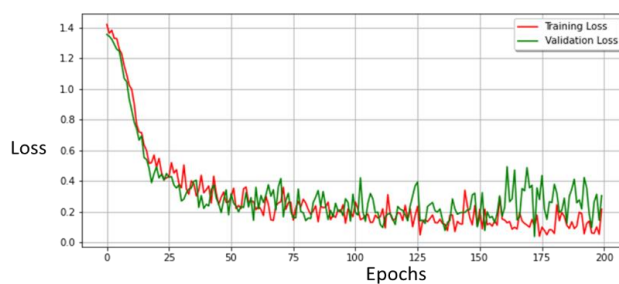


Fig. 7. Training CNN+PCA.

belongs to the confusion matrix of the PCA algorithm together with the convolutional neural network.

The interpretation of the results in the matrix (Figure 8, subparagraph a) indicates that most of the images (true positives) were classified correctly on the main diagonal. However, there are also false positives, which represent images not classified correctly.

Table 2. Comparison of PCA and non-PCA usage, using the Adam optimizer.

ALGORITHMS	Optimizer	METRICS		
		Sensitivity	Specificity	Accuracy
CNN+PCA	Adam	0.969	0.969	0.968
	Adadelta	0.322	0.204	0.321
	RMSprop	0.988	0.987	0.987
CNN	Adam	0.797	0.786	0.796
	Adadelta	0.250	0.1	0.250
	RMSprop	0.853	0.842	0.853

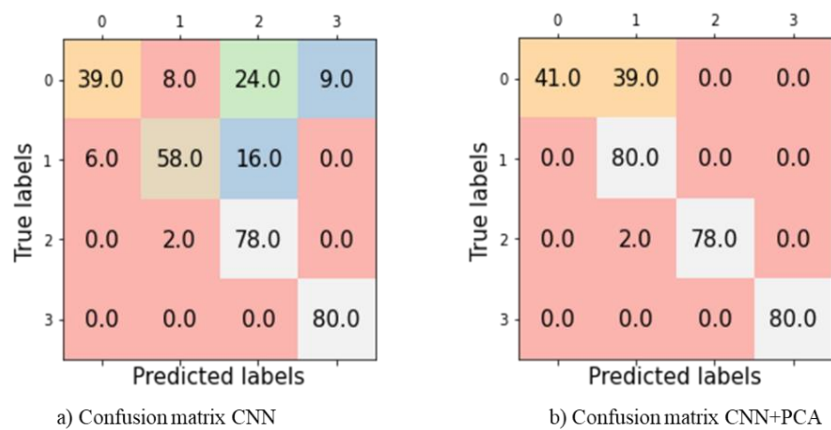


Fig. 8. Confusion matrix results.

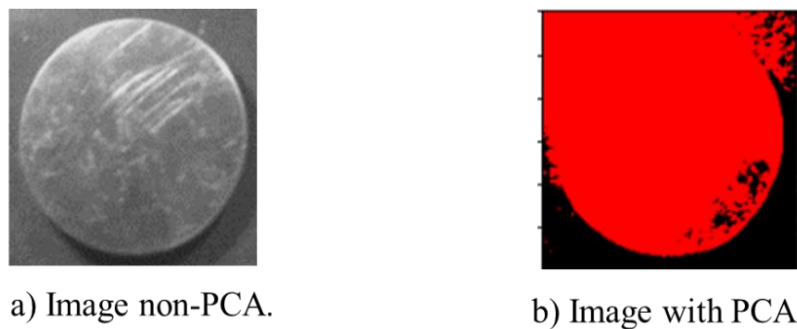


Fig. 9. Results of the application of the PCA algorithm.

For example, in the case of circles, 39 images were considered squares; this is due to the characteristics that the images present, such as the object inclination, its level of brightness or because the image is dark, or because the projections generated by PCA

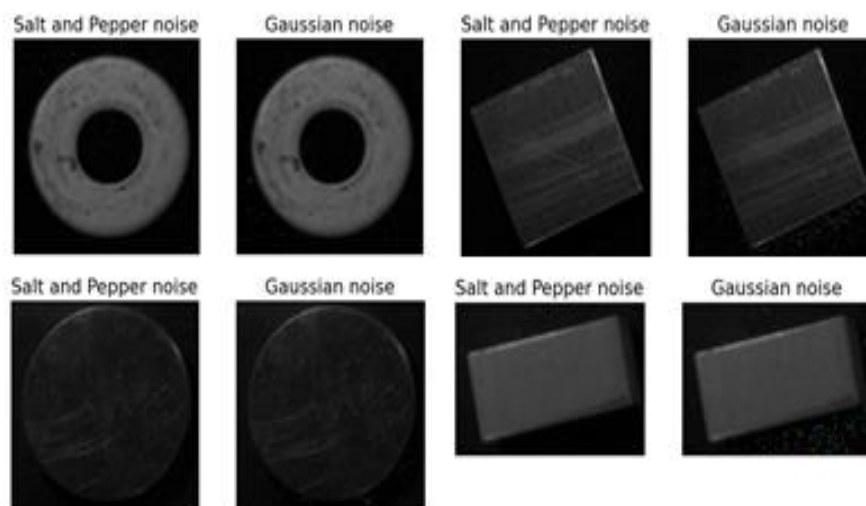


Fig. 10. Types of noise in images.

were not correct and therefore the images were not correctly classified. In the same way, for the images that correspond to the class of the rectangle, two of the images were considered to be squares, when they in fact were not. However, the results are far more accurate than the classification performed without the PCA algorithm, that is, it correctly classifies the largest number of images in each class.

In Figure 8, paragraph b, the only class that prevails as correctly classified is that of the flat washer because the resizing preserves the geometric shape of the image. However, the images are deformed through the convolutions of the CNN algorithm, causing confusion, such that a rectangle, when it is compressed, is classified as a square.

An image without PCA processing is shown in paragraph a) Figure 9. With PCA is shown in paragraph b) Figure 9, some regions of the image are dark, causing the edges to be lost or a shadow to be taken as part of the image to be evaluated. This causes amusing confusion in the convolutional neural network at the time of classification. The loss of characteristics is necessary in the analysis of principal components at the time of calculation, as it reduces the noise that the original image set may contain.

It is important to mention the application of the erode and dilate filter (Figure 10) to the images that contained Gaussian noise, salt and pepper noise, and extra brightness, in order to reduce such noise. The brightness was regularized through the implementation of balancing.

4 Conclusions

The purpose of this study is to demonstrate, through the evaluation of metrics, that the implementation of component analysis in a convolutional neural network improves image classification. Therefore, as assessed by the precision metrics and the RMSProp

optimizer, it shows 98.7%, compared with 85.3% for the convolutional neural network without PCA. Lastly feature extraction reduces feature redundancy in images, resulting in higher classification accuracy.

The position of the metal piece respect to the light source reflects the light in a distributed manner so that, when processed by PCA, sections of the background of the piece of interest end up being considered. This problem causes the component projection to be wrong for regions that should be discarded, as in the case of the circle. Adam (accuracy: 96.8%) and RMSProp (accuracy: 98.7%) resolve decreasing learning rates, compared to training without PCA. In addition, it improves the results obtained compared to the Adadelata optimizer (precision: 32.1% with PCA). Adadelata affects the relationship between the step size and the current gradient. However, the precision remains below 32.1% for each training related to the Adadelata optimizer. Therefore, the classification of true positives is incorrect.

The noise in the images affects the loss of information during the extraction of characteristics by the convolutional neural network with and without PCA. Some types of noise presented in 1.3 MP images are useful: luminance noise, Gaussian noise (which is generated by the camera by itself) or salt and pepper noise.

These types of noise provide complexity in the extraction of main components because the noise is detected as another component of the part, which is unnecessary. Likewise, the PCA algorithm achieves an improvement in the classification. Thus, it is important to consider the number of components that are within the determined level of 1 to 480 components. If fewer are used, the projection is not adequate; therefore, it will generate false positives and false negatives. In the same way, components after the 480th one cannot be picked up, as there are no more combinations to plot on the Cartesian plane. Due to the characteristics of the distribution of the number of images with respect to illumination, no more images were considered.

Therefore, as future work, it is proposed to add more images, with lighting scales of 2000 lx, 1500 lx, 1000 lx, with a lighting balance pretreatment, highlighting the contours of the metal parts, with an opaque black background, to avoid unnecessary light reflection, in order to implement them for training and validation of the neural network according to the proposed balance of 60% training and 40% validation. We also intend to carry out more tests using PCA according to the number of components necessary to project the figure of a circle (which is the piece with greater difficulty to project).

References

1. Arista Jalife, G.C.: Clasificación de imágenes urbanas aéreas: Comparación entre descriptores de bajo nivel y aprendizaje profundo. *Inf. Technol*, 28(3), pp. 209–224 (2020) doi: 10.4067/S0718-07642017000300021.
2. Abiodun, O.I., Aman, J., Omolara, A.E.: State-of-the-Art in Artificial Neural Network Applications: A Survey. *Heliyon*, 4(11) (2018) doi: 10.1016/j.heliyon.2018.e00938.
3. Amir, A.N., Mingui, H.: Convolutional Neural Network with PCA and Batch Normalization for Hyperspectral Image Classification. *School of Electronics and*

- Information Northwestern Polytechnical University, pp. 959–962 (2019) doi: 10.1109/IGARSS.2019.8899329.
4. Berzal, F.: *Redes neuronales y Deep Learning* (2018)
5. Camacho, F.: Redes neuronales convolucionales aplicadas a la traducción del lenguaje verbal español al lenguaje de señas Boliviano. *Revista Ciencia, Tecnología e Innovación*, 12(13), pp. 755–762 (2016)
6. Cruz Roa, A.S.: Classification and Automatic Mapping of Land Covers in Satellite Images Using Convolutional Neural Networks. *Universidad de los Llanos, Villavicencio, Meta. Colombia Suple* (2017)
7. Ruiz, D., Bacca, B.: Hyperspectral Images Classification Based on Inception Network and Kernel PCA. In: *IEEE Latin America Transactions*, 17(12) (2020)
8. Del Carpio Gallegos, J.C.: Mejora de la calidad aplicando la metodología de superficie respuesta y redes neuronales. *Industrial Data*, 9(1), pp. 59–63 (2006)
9. Jaya Prakash-Sahoo, S.A., Kumar-Patra, S.: SVM, Hand Gesture Recognition using PCA based Deep CNN Reduced Features and components. In: *IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS)*, 4 (2019) doi: 10.1109/iSES47678.2019.00056.
10. Lerma, B.C.: Análisis de componentes principales de imágenes multiespectrales en el ámbito del arte rupestre. In: *Conference: 1st Congress in Geomatics Engineering*, pp. 41–47 (2017) doi: 10.4995/CIGeo2017.2017.6597.
11. Ortega, M., González, C.R.: Validación de métodos alternativos para análisis microbiológico de alimentos y aguas. *Revista Cubana de Higiene y Epidemiología*, 51(1), pp. 111–121 (2013)
12. Mardiansyah, A., Contessa-Djamal, E., Nugraha, F.: Multivariate EEG Signal Using PCA and CNN in Post-Stroke Classification. *FORTEI-International Conference on Electrical Engineering (FORTEI-ICEE)*, 6 (2020)
13. Rivas-Asanza, W.: *Redes neuronales artificiales aplicadas al reconocimiento de patrones*. Machala: Universidad Técnica de Machala, 29 (2018)
14. Rodríguez, Y.B.: Integración de la red neuronal convolucional con el algoritmo de función. *Corporación Mexicana de Investigación en Materiales. División de estudios de posgrado*, 87 (2018)
15. Santamaría-Colula, L.Á.: Reconocimiento de patrones en imágenes por medio de redes neuronales convolucionales. Tesis para obtener el título de Maestro en ciencias de la computación. Benemérita Universidad Autónoma de Puebla, pp. 50–76 (2019)
16. Silva-Santos, M., Bernarda-Ludermir, T.: Using Factorial Design to Optimise Neural Networks. *Universidade Federal de Pernambuco*, 2 (2020) doi: 10.1109/IJCNN.1999.831064.
17. Wenqiang, Y., Peng, C.: Recognition Algorithm of Emitter Signals Based on PCA+CNN. In: *IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference*, 5, pp. 2410–2414 (2018)
18. Winarno, E., Husni Al Amin, I., Februriyanti, H.: Attendance System Based on Face Recognition System Using CNN-PCA Method and Real-time Camera. In: *International Seminar on Research of Information Technology and Intelligent Systems*, pp. 301–304 (2019) doi: 10.1109/ISRITI48646.2019.9034596.

Análisis de comentarios de turistas sobre Áreas Naturales Protegidas de México utilizando plataformas Google

Cynthia Yustis-Cruz, Magdalena Saldana-Perez, Marco Moreno-Ibarra

Instituto Politécnico Nacional,
Centro de Investigación en Computación
México

{cyustisc2020, amagdasaldana, marcomoreno}@cic.ipn.mx

Resumen. El turismo es una actividad económica de gran relevancia en México. Las Áreas Naturales Protegidas son áreas resguardadas para la conservación de su biodiversidad, en las que recientemente se ha buscado implementar estrategias que permitan aprovecharlas turísticamente de una forma sustentable. El Procesamiento de Lenguaje Natural es el campo dentro de las ciencias de la computación que conjunta técnicas de inteligencia artificial con lingüística y que, entre otras aplicaciones, permite analizar textos de forma automática. El presente trabajo es una propuesta que analiza 430 comentarios de usuarios de YouTube y Google Maps referentes al Área Natural Protegida Reserva de la Biósfera Tehuacán-Cuicatlán. El estudio consiste en un análisis de sentimientos que otorga una puntuación para identificar los comentarios positivos y negativos. Los datos fueron etiquetados manualmente y automáticamente empleando la *Natural Language API* de Google Cloud. Finalmente, se comparan los resultados obtenidos mediante el etiquetado automático con el etiquetado manual y se presentan las palabras más frecuentes en cada una de las categorías.

Palabras clave: Área natural protegida, procesamiento de lenguaje natural, análisis de sentimientos, natural language API, Google.

Analysis of Tourist Comments on Protected Natural Areas of Mexico Using Google Platforms

Abstract. Tourism is a highly significant economic activity in Mexico. Protected Natural Areas are regions designated for the conservation of biodiversity, where sustainable tourism strategies have recently been promoted. Natural Language Processing (NLP) is a field within computer science that combines artificial intelligence techniques with linguistics and enables automatic text analysis, among other applications. This study proposes the analysis of 430 user comments

from YouTube and Google Maps referring to the Protected Natural Area Tehuacán-Cuicatlán Biosphere Reserve. The analysis involves a sentiment evaluation that assigns a score to identify positive and negative comments. The data was labeled both manually and automatically using Google Cloud's Natural Language API. Finally, the results from automatic and manual labeling are compared, and the most frequent words in each sentiment category are presented.

Palabras claves: Protected natural area, natural language processing, sentiment analysis, natural language API, google.

1. Introducción

El turismo es una actividad económica de gran importancia en México [27], por ello la política sectorial federal ha mostrado interés por diversificar la oferta turística hacia sitios que permitan un mayor contacto con la naturaleza [7], como lo son las Áreas Naturales Protegidas (ANP). Por decreto, las ANP son áreas dedicadas a la protección de la biodiversidad, el cuidado de los paisajes naturales y el mantenimiento de los sistemas ecológicos que la componen, donde se busca gestionar la interacción entre pobladores y visitantes de las mismas para lograr un desarrollo social sostenible. Dichas áreas son gestionadas principalmente por la Comisión Nacional de Áreas Naturales Protegidas (CONANP) [8].

Como parte de la estrategia para el aprovechamiento responsable de las ANP se ha planteado el Marco Estratégico de Turismo Sustentable en Áreas Naturales Protegidas de México [5]. Sin embargo, no se han desarrollado las herramientas suficientes para realizar una gestión turística adecuada que permita tomar decisiones y generar nuevas propuestas basadas en un análisis de los datos que pueden proporcionar los visitantes. Es decir, no se cuenta con la información suficiente para realizar el diagnóstico sobre el comportamiento de los turistas, los sitios que visitan, su percepción de la región y la huella ecológica y digital que representan para las ANP. Es importante mencionar que, pese al avance tecnológico en otras áreas económicas, el turismo, y de manera particular las ANP carecen de las herramientas tecnológicas necesarias para la gestión y promoción eficiente de sus actividades.

Las plataformas digitales son herramientas tecnológicas que la sociedad usa día con día, para comunicarse, expresarse y promoverse. Gracias a estas acciones, los usuarios proporcionan grandes cantidades de información que son usadas por diversas compañías para diferentes objetivos [23]. Entre estos objetivos podemos mencionar la detección de preferencias de usuario, detección de patrones de movimiento en ciudades, marketing y ventas, entre muchas otras.

En el caso que aborda el presente trabajo se analizan datos generados por usuarios de YouTube y Google Maps, en los que muestran opiniones sobre sus experiencias en puntos específicos dentro del Área Natural Protegida (ANP) Reserva de la Biósfera Tehuacán-Cuicatlán [6]. El objetivo es tratar los datos empleando Procesamiento de Lenguaje Natural (PLN) para convertirlos en información que permita comprender mejor la opinión turística acerca del ANP

propuesta. Esto se logra aplicando la herramienta de análisis *Natural Language API* de Google Cloud [15] para realizar un etiquetado automático que permita clasificar los comentarios en positivos, negativos y neutros. Posteriormente, el resultado será comparado con un etiquetado manual. Finalmente, dichos resultados serán separados para entregar las palabras más relevantes de cada categoría.

La sección 2 presenta los antecedentes, que incluyen el estado del arte. Posteriormente, en la sección 3 se presenta la metodología, así como una descripción de las fuentes de datos y herramientas de análisis seleccionadas. Los resultados de los experimentos se describen en la sección 4, para ser discutidos en la sección 5 y el trabajo finaliza con la conclusión en la sección 6.

2. Antecedentes

A continuación, se presentan antecedentes relevantes para el desarrollo y comprensión del artículo. Se incluyen los antecedentes de datos turísticos y aplicaciones tecnológicas empleadas en el sector turístico, así como estado del arte relacionado con el procesamiento de lenguaje natural aplicado al análisis de sentimientos.

2.1. Datos turísticos en México

La CONANP y el gobierno de México cuentan con sitios web [11] que permiten acceder a información relacionada con cada una de las ANP de México. Sin embargo, esta información es muy general y no es suficiente para la gestión turística de dichas áreas. Por un lado, se cuenta con fichas técnicas [4] que permiten conocer información como la categoría que las áreas tienen asignada, la macro ubicación de la misma, extensión territorial, número de habitantes, datos relacionados con el decreto oficial y personal encargado de dirigirla. También se cuenta con un sistema de información geoespacial [9] que proporciona mapas de las delimitaciones territoriales de las ANP, así como mapas disponibles para ser descargados en formatos adecuados para su edición en programas para el procesamiento de datos geoespaciales. Así mismo, existen otras plataformas gubernamentales como es el sitio web de biodiversidad de la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO) [3] que también proporciona información general de ANP.

Por su parte, la Secretaría de Turismo (SECTUR), cuenta con el sistema DataTur [24], una plataforma que proporciona información sobre vuelos, visitas a museos y zonas arqueológicas, indicadores de empleo turístico, encuestas a viajeros internacionales, entre otros datos. De igual forma, proporciona datos geoespaciales referentes a mapas de algunas entidades y rutas turísticas, se puede encontrar una recopilación de dichos datos en [30]. También, cuenta con información sobre atractivos y destinos turísticos dirigida a los visitantes [28]. No obstante, estos datos son dirigidos, en su mayoría, a turistas y no a gestores, además de que hacen referencia al turismo general en México y entre sus estadísticas no se encontró información sobre los turistas que visitan las ANP.

2.2. Tecnologías de la información aplicadas a la gestión del turismo en México

Se denomina Turismo Inteligente al uso de tecnologías de la información, tales como Internet de las Cosas (*IoT*), Cómputo en la Nube (*Cloud Computing*), tecnologías móviles e Inteligencia Artificial (*AI*) para el desarrollo de sistemas inteligentes aplicados al turismo, que a su vez son amigables con los usuarios. En [20] se menciona que el futuro del turismo está basado en el desarrollo tecnológico, la rapidez y la extensión de la aplicación de dichos sistemas a todos los niveles del sector turístico. Se ha estudiado y concluido que se debería tener conocimiento de las áreas que resultan más atractivas para los visitantes y cuáles son sus características principales. La importancia de esta información radica en que estos datos contribuyen como herramientas para un mejor y más inteligente diseño de propuestas, así como para la implementación de estrategias sustentables que llevarán a la optimización de la experiencia turística.

Desde el año 2013, la Sociedad Estatal para la Gestión de la Innovación y las Tecnologías Turísticas (SEGITTUR), una sociedad española dedicada a la gestión e innovación de tecnologías turísticas, junto con la Secretaria de Estado de Turismo de España han impulsado un programa de Destinos Turísticos Inteligentes (DTI) [29]. En 2015, la isla de Cozumel se incorporó a este proyecto, no obstante, el gobierno de México remarcó la necesidad de trabajar en la infraestructura para sistemas de comunicaciones y desarrollo de plataformas tecnológicas, que permita la interacción del gobierno con los prestadores de servicios para establecer nuevos modelos de negocios y operación [26] [25]. Por otro lado, la Zona Dorada de Ciudad Juárez, ubicada en el norte de México, fue tomada como caso de estudio para la búsqueda de áreas de oportunidad hacia la conformación de la zona como un espacio turístico inteligente y sustentable, lo anterior, mediante un estudio de sus componentes e infraestructura tecnológica. Derivado del análisis se concluyó que la zona estudiada no cuenta con el soporte tecnológico para clasificar como DTI, siendo una de las principales limitantes el acceso a redes de internet inalámbricas abiertas. También se mencionó la falta de aplicaciones móviles para facilitar la interacción con los visitantes y un déficit en la recolección de datos que los usuarios pueden estar generando y que podrían ser aprovechados para la creación de un observatorio turístico [12].

Con el desarrollo de este trabajo se pretende estudiar los datos que pueden proporcionar las plataformas digitales y generar información de interés sobre la opinión de los turistas o potenciales turistas del Área Natural Protegida seleccionada. Así como observar el comportamiento de los datos al clasificarlos manualmente y automáticamente. Lo anterior, mediante una metodología propuesta que sea aplicable al estudio de otras Áreas Naturales Protegidas en México.

La investigación se presenta como una propuesta para aprovechar la información generada por usuarios de plataformas digitales en la búsqueda de mejorar la gestión turística en México. Se propone la incorporación de una herramienta de PLN para conocer la opinión de los turistas en un Área Natural Protegida con la visión de proporcionar alternativas que apoyen la

toma de decisiones de administradores del sector turístico. Contar con una herramienta tecnológica que permita generar información a partir del análisis de datos referentes al comportamiento y opiniones de los turistas en ANP representa un avance de impacto en materia de gestión turística a través de la integración de tecnologías de la información como una propuesta para mejorar la administración y promoción de las mismas.

2.3. Procesamiento de lenguaje natural y análisis de sentimientos

Se conoce como Procesamiento de Lenguaje Natural (PLN) al área dentro de las ciencias de la computación que se ocupa de trabajar con aspectos relacionados a la comprensión de la comunicación humana [18]. Esto puede incluir tanto la capacidad de las máquinas para comprender el lenguaje, como la capacidad de generar respuestas basadas en el mismo. Para el desarrollo de algoritmos de PLN se suele tomar un enfoque lingüístico, es decir, se toman elementos semánticos y sintácticos básicos como características [18]. También, se han implementado técnicas de aprendizaje automático para realizar tareas de clasificación dentro del área.

El análisis de sentimientos es una de las áreas más activas dentro del PLN. Su tarea es evaluar lenguaje escrito o hablado y determinar si las expresiones evaluadas son negativas, positivas o neutras [18]. Este tipo de análisis es útil en temas relacionados con atención a cliente u opinión de usuarios, medición de estado de ánimo y rastreo de comportamiento humano. Sin embargo, tanto el PLN en general como el análisis de sentimientos son temas complejos que son motivo de libros y artículos totalmente dedicados al tema [1] [2].

Actualmente, existen diversas alternativas para realizar PLN y se pueden ejecutar algoritmos de análisis en distintos lenguajes de programación. Dependiendo de las herramientas que se elijan, el procesamiento puede ir desde la limpieza de los datos hasta la segmentación del texto para el entrenamiento de modelos de clasificación, o bien, existen modelos previamente entrenados que analizan textos en diferentes idiomas. Dichos modelos se pueden encontrar en bibliotecas o módulos previamente desarrollados para distintos lenguajes de programación, o bien, dentro de plataformas de servicios de cómputo en la nube como son Microsoft Azure, Amazon Web Services o Google Cloud. En el caso que se aborda en este trabajo se utilizará la *Natural Language API* de Google Cloud [15].

Un ejemplo del análisis de sentimientos aplicado al turismo puede ser [19], donde se emplea una metodología que emite evaluaciones positivas, negativas y neutrales a datos turísticos tomados de las plataformas TripAdvisor y VirtualTourist. Concluyen que sus aportes pueden ser útiles a los encargados de gestión de los destinos turísticos en su tarea de mejorar la calidad de los servicios. También mencionan la necesidad de profundizar en la generación de datos para la posterior aplicación de ciencia de datos para su análisis.

En 2020, [21] presenta un análisis de sentimientos aplicado a una comunidad de estudiantes, esto se realizó analizando 2866 publicaciones y 3630 comentarios en inglés, obtenidos de la plataforma Google Plus (G+). Para ello hicieron uso de



Fig. 1. Metodología propuesta. (Iconos tomados de [10] y [21])

la *Natural Language API* de Google Cloud. Aunque este trabajo está totalmente dirigido al sector educativo, pudieron implementar de forma satisfactoria la herramienta de PLN de Google Cloud para el análisis, aunque destacaron que, desafortunadamente, uno de los idiomas oficiales de la región de estudio no estaba incluido en la lista de idiomas de la *API* y ello los llevó a limitar su base de datos.

3. Materiales y métodos

El presente trabajo es una propuesta de aplicación de procesamiento de lenguaje natural aplicado a comentarios de usuarios de las plataformas YouTube y Google Maps. Se estudia el caso de la Reserva de la Biosfera Tehuacán-Cuicatlán. En esta sección se describe la metodología establecida, la descripción de las fuentes de datos y la obtención de los mismos, así como la herramienta de análisis automático seleccionada.

3.1. Metodología

La Fig. 1 muestra un diagrama de la metodología establecida para este trabajo, consta de 14 pasos que van desde la selección de las plataformas para la obtención de los datos, hasta la comparación de los resultados obtenidos.

Las plataformas seleccionadas para obtener los datos fueron la red social YouTube y la plataforma Google Maps. Como se muestra en el paso 2, los datos fueron obtenidos a través de las Interfaces de Programación de Aplicaciones

(API) oficiales de cada una de las plataformas y posteriormente se conjuntaron en un solo archivo durante el paso 3. Una vez que se tuvo un primer conjunto de datos se filtraron y etiquetaron manualmente para dejar listo el conjunto de datos que contiene información referente a la reserva natural de interés. Se eliminaron los datos que no estaban relacionados con el área natural seleccionada, dando un total de 430 comentarios relacionados, a estos datos finales se les otorgó una puntuación de -1 a 1, donde los valores cercanos a 1 representan comentarios que se relacionaron con un sentimiento positivo, los cercanos a -1 con sentimientos negativos y los cercanos a 0 se consideran comentarios neutros.

Una vez que se cuenta con el conjunto de datos de interés, y antes de realizar el análisis de sentimientos con la herramienta seleccionada, es necesaria una limpieza básica del texto donde se eliminan símbolos como emoticones, hipervínculos y otros caracteres especiales que pudiera contener el texto. Después del paso 6 se tiene un texto plano en minúsculas, listo para entrar a la herramienta de clasificación, en el paso 7 se implementa la herramienta de procesamiento de lenguaje natural de Google Cloud a través de la *Natural Language API*. Por medio de este último paso se le otorga otra etiqueta a los datos, esta etiqueta también va de -1 a 1, pasando por 0, para comentarios negativos, positivos y neutros respectivamente. Se realiza la diferencia de los valores absolutos del etiquetado manual con el etiquetado automático para saber qué tan distantes son unos valores de otros y se toma esto como el cálculo del error del paso 9.

Una vez que los datos tienen las etiquetas necesarias las frases son segmentadas en palabras durante el paso 10. Una vez en el paso 11, se eliminan las palabras conocidas como palabras vacías o *stopwords*, que incluyen clases de palabras como artículos, pronombres y preposiciones.

A través de las etiquetas generadas durante los pasos 4 y 8, durante el paso 12 se categorizan los datos en tres, los datos que tienen valores entre -1 y -0.2 se catalogan como negativos, los que tienen valor entre 1 y 0.2 como positivos y los restantes como neutros. Esta categorización se realiza tanto para el etiquetado manual como para el automático. De esta forma se tienen las categorías: manual positivo, manual negativo, manual neutro, automático positivo, automático negativo y automático neutro. Finalmente, en el paso 13 se realiza un conteo de las palabras que se repiten más en cada una de las categorías y se comparan los resultados de ambos etiquetados durante el paso 14.

A continuación, se presenta la descripción detallada de algunos de los pasos de esta metodología para su mejor comprensión. Esto incluye la descripción detallada de la obtención de datos de Youtube y el uso de la API para su descarga. De la misma forma para Google Maps y detalles sobre el uso de la API de procesamiento de lenguaje natural de Google Cloud.

3.2. Datos de Youtube

La plataforma YouTube es una red social de Google basada en publicación de videos [31], en los cuales usuarios pueden realizar comentarios. Los datos

de YouTube utilizados para este trabajo consisten en la lista de comentarios provenientes de los primeros 50 videos relacionados a una búsqueda.

Para la realización del presente proyecto, se determinó una lista de palabras para búsqueda y se realizó la obtención de los comentarios mediante la Interfaz de Programación de Aplicaciones (*API*) *YouTube Data API* oficial [17], utilizando el lenguaje de programación Python [22]. La lista de búsquedas incluye nombres de zonas y comunidades de interés dentro del ANP, mismos que se presentan a continuación:

- Tehuacán Cuicatlán
- Salinas Zapotitlán
- Alfareros Mezontla
- Jardín botánico Helia Bravo Hollis Zapotitlán

3.3. Datos de Google Maps

La plataforma de Google Maps es una aplicación de mapas web de Google [14], que permite a los usuarios realizar búsqueda de localizaciones, servicios, trazado de rutas, vistas de mapas desplazables e imágenes satelitales. Google Maps permite dar de alta ubicaciones de servicios, negocios, parques, escuelas, entre otros, y permite a los usuarios y visitantes dejar reseñas y comentarios sobre las ubicaciones.

Para la tarea de recolección de esta propuesta, también se realizó una lista de palabras para búsqueda y se realizó la obtención de los comentarios mediante la *Places API* oficial [16], utilizando el lenguaje de programación Python [22]. A diferencia de YouTube, la obtención de reseñas de lugares en Google Maps tienen la limitante de que solo permite la obtención de hasta cinco reseñas por lugar, a menos de que el solicitante sea el administrador de la ubicación. La lista de búsquedas incluye nombres de zonas y comunidades de interés dentro del ANP, mismos que se presentan a continuación:

- Cañón Alas Verdes
- Tierra Colorada
- Museo Comunitario Paleontológico de San Juan Raya
- Turrítelas
- Huellas de Dinosaurios
- Reserva de la Biósfera Tehuacán-Cuicatlán Zapotitlán Salinas, Puebla
- Zapotitlan Salinas Botanical Garden
- Helia Bravo Hollis Botanical Garden
- Letras Zapotitlán
- Pata de Elefante
- Paleoparque Las ventas
- Mirador ecológico San Antonio Texcala
- El Encinal, Santa Ana Teloxtoc
- Los Reyes Metzontla, Puebla
- Centro Artesanal Comunitario - Alfareros Popolocas de los Reyes Metzontla

- Las Salinas Grandes
- Mirador Tehuacán-Oaxaca
- San Pedro Jaltepetongo
- Reserva de la Biósfera Tehuacán- Cuicatlán Cerro Prieto, Oaxaca
- Temazcal Zapotitlán
- Campestre Zapotitlán

La lista de búsqueda para lugares en Google Maps es mayor a la lista de búsqueda de YouTube debido a que las búsquedas en YouTube regresan videos relacionados y se tendrían muchos datos repetidos al buscar todos los lugares en la última lista presentada, pero Google Maps devuelve los comentarios de lugares específicos por lo que no se tendrán comentarios repetidos, a menos que el usuario haya escrito el mismo comentario en dos ubicaciones diferentes.

3.4. Google Cloud para el procesamiento del lenguaje natural

Google Cloud es un conjunto de servicios de cómputo en la nube ofertados por Google [13]. Como parte de estos servicios, existe la *Natural Language API* que utiliza técnicas de aprendizaje automático para analizar textos no estructurados. Esta *API* cuenta con modelos previamente entrenados que permiten realizar funciones de análisis de opinión, análisis de entidades, análisis de opinión por entidad, clasificación de contenido y análisis de sintaxis. Es importante mencionar que la *Natural Language API* es solamente uno de los tres servicios de procesamiento de lenguaje natural que ofrece Google. Se eligió para probar su capacidad de clasificar la opinión de los textos con los modelos previamente entrenados.

La documentación de Google Cloud para el procesamiento de lenguaje natural describe el análisis de sentimientos como una inspección de textos que identifica la opinión emocional que prevalece dentro del mismo, lo anterior para determinar la actitud positiva, negativa o neutra del autor. Como resultado de la solicitud la *API* entrega una puntuación que va de -1 a 1, donde valores cercanos a -1 corresponden a sentimientos negativos, los cercanos a 1 a sentimientos positivos y los valores alrededor de 0 a comentarios neutros. El método que permite efectuar esta tarea puede analizar un texto general o a nivel de oraciones. Para la metodología propuesta se utilizó la puntuación obtenida a nivel texto.

4. Experimentos y resultados

En esta sección se presentan las gráficas obtenidas siguiendo la metodología establecida, así como los resultados y experimentos adicionales.

Las Fig. 2 y 3 muestran la distribución de los valores obtenidos en el etiquetado manual y el automático. A simple vista se observa que la distribución es muy similar para valores cercanos a 1, es decir, en los valores referentes a la comentarios positivos. Esta distribución va mostrando más cambios para valores menores a 0.5, es decir, para hacia los valores relacionados con los comentarios neutros y negativos.

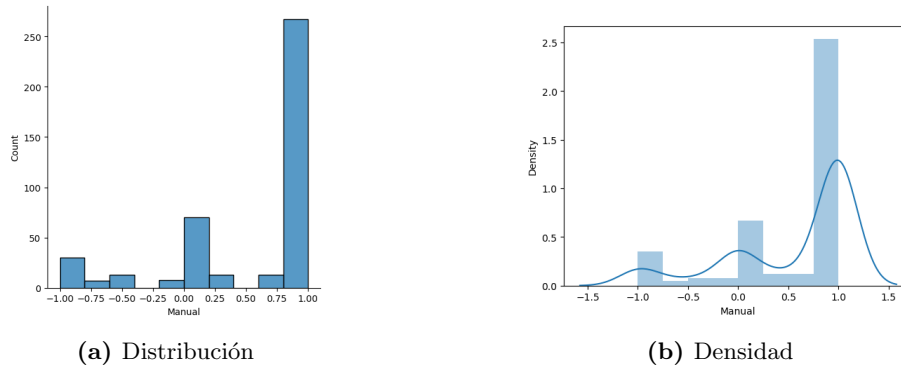


Fig. 2. Distribución de los valores otorgados mediante el etiquetado manual. Se representa la distribución de la puntuación, entre -1 a 1, de cada uno de los 430 comentarios.

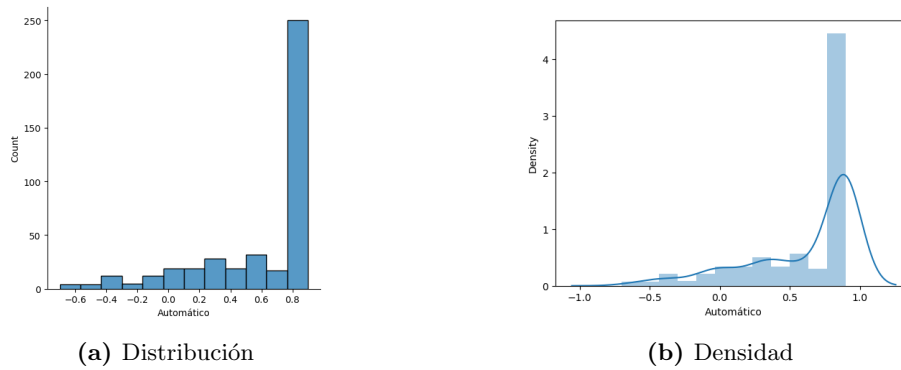


Fig. 3. Distribución de los valores otorgados mediante el etiquetado automático. Se representa la distribución de la puntuación, entre -1 a 1, de cada uno de los 430 comentarios.

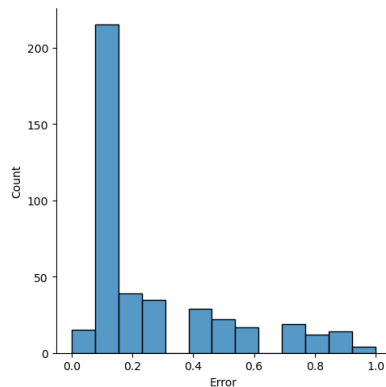


Fig. 4. Distribución del error calculado a través de obtener la diferencia entre la puntuación asignada de manera manual y automática.

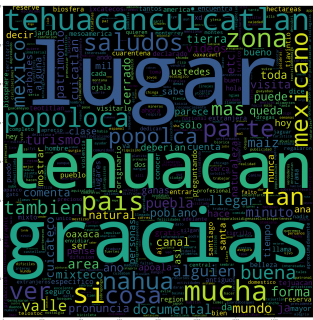


Fig. 5. Nube de palabras de la categoría positiva para el etiquetado manual y automático.

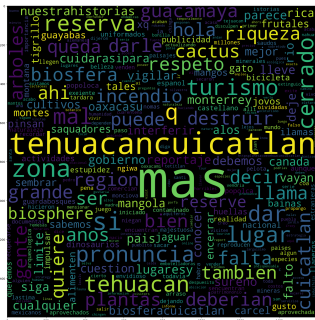
La Fig. 4 muestra la distribución obtenida del error calculado, este error se obtuvo del valor absoluto de la diferencia entre valores absolutos del etiquetado manual y el etiquetado automático. Aunque se muestra que la mayoría de los errores calculados fueron cercanos a 0.1, al obtener el promedio de los valores se tiene que en promedio se tuvo un error de 0.26. Es importante mencionar que la clasificación manual se realizó basada en listas de términos establecidos para diferenciar entre positivos y negativos. Los datos que no tuvieran alguno de ellos se determinaron como neutros. Si se tenían varias frases y cada una expresaba diferentes sentimientos se tomaba el promedio de las puntuaciones. La clasificación manual se logró estableciendo la semántica entre los grupos de

Pos	Manual	Automático
1	('lugar')	('lugar')
2	('tehuacan')	('tehuacan')
3	('mas')	('gracias')
4	('gracias')	('mas')
5	('hermoso')	('hermoso')
6	('excelente')	('excelente')
7	('reserva')	('reserva')
8	('mexico')	('mexico')
9	('bonito')	('puebla')
10	('puebla')	('ver')
11	('tambien')	('tambien')
12	('ver')	('si')
13	('conocer')	('saludos')
14	('gran')	('bonito')
15	('naturaleza')	('tan')
16	('saludos')	('conocer')
17	('san')	('pais')
18	('lugares')	('san')
19	('cuicatlan')	('lugares')
20	('bien')	('zona')
21	('si')	('bien')
22	('zona')	('naturaleza')
23	('ser')	('ser')
24	('oaxaca')	('gran')
25	('mejor')	('oaxaca')
26	('tan')	('cuicatlan')
27	('aqui')	('muchas')
28	('ir')	('parte')
29	('recomiendo')	('documental')
30	('region')	('bien')

Fig. 6. Comparación de las palabras más frecuentes en la categoría positiva para el etiquetado manual y automático.



(a) Manual



(b) Automático

Fig. 7. Nube de palabras de la categoría neutra para el etiquetado manual y automático.

palabras que se referían a aspectos positivos y negativos. Entre estas palabras no se incluyen los nombres propios de los sitios de interés.

Cuando los comentarios contaban con ambas etiquetas fueron separados en subgrupos. Una vez categorizados se utilizó la herramienta nube de palabras para generar una visualización de las palabras más frecuentes en cada categoría. El tipo de gráfico presentado en la Fig. 5 muestra las palabras más frecuentes empleadas en los textos de la categoría positiva, es decir, comentarios con valores



Fig. 8. Nube de palabras de la categoría negativo para el etiquetado manual y automático.

de etiquetado en el intervalo que va de 0.2 a 1, para el método manual en la Fig. 5a y el automático en la Fig. 5b. El tamaño de las palabras mostradas es proporcional a la cantidad de veces que se repiten en la categoría, las palabras con mayor tamaño han sido mencionadas más frecuentemente que las de menor tamaño. Las cercanía de unas palabras con otras también está relacionada con su uso en el texto, es decir, las palabras más cercanas son las que están comúnmente agrupadas en el texto.

En la Fig. 6 se presentan las 30 palabras más frecuentes en la categoría positiva para los comentarios etiquetados manualmente y automáticamente. La primera posición de la lista es la palabra más repetida y descienden conforme desciende su frecuencia de uso. Se usa un semáforo de colores, donde el color verde representa que la palabra se encuentra en ambas listas en la misma posición de lista, el amarillo indica que la palabra se encuentra en ambas listas pero están desfasadas una posición, mientras que el naranja representa que se encuentra en ambas listas pero desfasadas en más de una posición y el rojo representa que la palabra no se encuentran en la otra lista.

Se observa que, de las 30 palabras comparadas solo cinco palabras de cada lista son diferentes y hubo coincidencia en 25 palabras de uso frecuente. Entre las palabras más frecuentes encontramos: lugar, Tehuacán, hermoso, excelente, reserva, México, bonito, Puebla, Naturaleza, conocer y país.

Estos ejercicios se repitieron para la categoría de negativos y neutros. En la Fig. 7 podemos observar la nube de palabras para la categoría de comentarios neutros, es decir, con valores de etiquetado mayores a -0.2 y menores a 0.2. Por su parte, las nubes de palabras para la categoría de comentarios negativos, con un etiquetado entre -0.2 y -1, se pudo observar en la Fig. 8.

La comparación de las palabras más frecuentes para el etiquetado manual y automático de las categorías de comentarios neutros y comentarios negativos se presenta en la Fig. 9. La Fig. 9a muestra las palabras de la categoría de comentarios neutros y la Fig. 9b las palabras de la categoría de comentarios

Pos	Automático	Manual
1	('mas')	('lugar')
2	('tehuacancuicatlan')	('gracias')
3	('si')	('tehuacan')
4	('mal')	('tehuacancuicatlan')
5	('zona')	('si')
6	('q')	('muchas')
7	('turismo')	('pais')
8	('pronuncia')	('tan')
9	('reserva')	('parte')
10	('tehuacan')	('ver')
11	('lugar')	('saludos')
12	('the')	('zona')
13	('dar')	('popolca')
14	('cerrado')	('cosas')
15	('bien')	('tambien')
16	('biosfera')	('mexico')
17	('deberian')	('mas')
18	('arboles')	('alguien')
19	('quiere')	('nahuas')
20	('decir')	('valle')
21	('plantas')	('llegar')
22	('cactus')	('popolocas')
23	('gente')	('puebla')
24	('ser')	('mundo')
25	('falta')	('maiz')
26	('destruir')	('area')
27	('ahi')	('mexicano')
28	('incendio')	('documental')
29	('anos')	('buena')
30	('tambien')	('forma')

Pos	Manual	Automático
1	('mas')	('cerrado')
2	('lugar')	('mas')
3	('cerrado')	('si')
4	('tehuacan')	('tehuacan')
5	('si')	('lastima')
6	('gente')	('mundo')
7	('puebla')	('estan')
8	('mal')	('llegar')
9	('gobierno')	('tan')
10	('estan')	('mal')
11	('mundo')	('popolca')
12	('lastima')	('gobierno')
13	('plantas')	('acabe')
14	('bien')	('dinero')
15	('reserva')	('naturaleza')
16	('region')	('hicieron')
17	('san')	('queremos')
18	('q')	('ver')
19	('hicieron')	('hace')
20	('pena')	('parte')
21	('mexicanos')	('lugar')
22	('pais')	('importante')
23	('cuidar')	('veo')
24	('ser')	('malo')
25	('zapotitan')	('sabe')
26	('venden')	('q')
27	('turistas')	('llueva')
28	('sabemos')	('gente')
29	('parte')	('hora')
30	('juan')	('terraceria')

(a) Categoría de comentarios neutros.

(b) Categoría de comentarios negativos.

Fig. 9. Comparación de las palabras más frecuentes en la categoría neutra y negativa para el etiquetado manual y automático. El color verde representa que la palabra se encuentra en la ambas listas con la misma posición, el amarillo indica que la palabra se encuentra en ambas listas pero desfasadas una posición, el naranja representa que se encuentra en ambas listas pero desfasadas en más de una posición y el rojo representa que las palabras no se encuentran en la otra lista.

negativos. El semáforo de colores, explicado previamente, se mantiene para estas figuras. Con esto podemos visualizar que la cantidad de palabras frecuentes que se repiten es más reducida que en el caso de la categoría de comentarios positivos, siendo la categoría de comentarios neutros la que tiene menor cantidad de palabras repetidas en común entre el etiquetado manual y automático.

Sin embargo, en la Fig. 9 podemos observar que hay palabras repetidas que comparten ambas categorías, es por esto que se propuso un último ejercicio que consiste en unir la categoría de comentarios neutros con negativos para evaluar si existen más palabras repetidas en común. Se creó una categoría de prueba que va de 0.2 a -1 y se realizó el ejercicio de conteo, los resultados de éste pueden ser visualizados en la Fig. 10.

Pos	Manual	Automático
1	('mas')	('lugar')
2	('cerrado')	('mas')
3	('si')	('tehuacan')
4	('tehuacan')	('si')
5	('tehuacancuicatlan')	('cerrado')
6	('mal')	('gracias')
7	('q')	('pais')
8	('lugar')	('puebla')
9	('zona')	('tehuacancuicatlan')
10	('reserva')	('tan')
11	('mundo')	('muchas')
12	('gente')	('mundo')
13	('lastima')	('parte')
14	('plantas')	('ver')
15	('gobierno')	('zona')
16	('bien')	('saludos')
17	('popolca')	('mal')
18	('turismo')	('bien')
19	('naturaleza')	('cosas')
20	('queremos')	('reserva')
21	('pronuncia')	('gente')
22	('hicieron')	('tambien')
23	('anos')	('region')
24	('region')	('estan')
25	('estan')	('llegar')
26	('puede')	('mexicanos')
27	('the')	('mexicano')
28	('dar')	('gobierno')
29	('falta')	('popolca')
30	('llegar')	('ser')

Fig.10. Comparación de las palabras más frecuentes en la categoría neutro-negativo propuesta para el etiquetado manual y automático.

En la Fig. 10 podemos observar que la cantidad de palabras repetidas que coinciden para ambos etiquetados aumentó en comparación con la Fig. 9, teniendo en total 17 palabras de alta frecuencia de repetición para ambos etiquetados. Siendo algunas de ellas: más, cerrado, Tehuacán, tehuacancuicatlan, gobierno, región, zona, llegar, reserva y mundo.

Con este último ejercicio se concluyen los experimentos realizados para este trabajo, en la siguiente sección se discuten los resultados obtenidos.

5. Discusión

En este apartado se comentarán los resultados de la sección anterior, así como algunas propuestas de mejora y trabajo a futuro.

Comenzaremos mencionando que el filtrado de los datos para obtener un cuerpo de comentarios enfocado en el turismo del ANP seleccionada se realizó manualmente ya que aún no se contaba con suficientes datos para realizar un

modelo de clasificación automática. Sin embargo, es una tarea que se plantea para ser desarrollada en futuros trabajos.

En las gráficas de distribución presentadas en las Fig. 2 y 3 se puede observar una distribución similar, donde la mayoría de los comentarios fueron etiquetados con valores correspondientes a sentimientos positivos y los comentarios neutros y negativos representaban la minoría de los datos. Al mostrar el error que se tiene, se encuentra que la variación promedio entre el etiquetado manual y el automático es de aproximadamente 0.26. Para este caso es importante tomar en cuenta que existe un factor de objetividad en el etiquetado manual de los datos, pues la puntuación es ponderada mediante un listado de términos establecidos para identificar contenido positivo, negativo o neutro pero el número específico otorgado es determinado por los investigadores.

Referente a los resultados de la comparación de palabras más frecuentemente repetidas para cada categoría, durante los experimentos quedó claro, mediante la comparación del caso de etiquetado manual y automático, que se tiene mayor coincidencia de palabras positivas. Con lo que se puede concluir que el clasificado de esta categoría es en su mayoría correcto. En el caso de los comentarios neutros y negativos, se observó una diferencia importante entre el etiquetado manual y el automático, una vez que se unificaron las clases neutro y positivo se pudieron encontrar más coincidencias entre los etiquetados, con esto se puede concluir que es más complicado diferenciar entre comentarios neutros y negativo que identificar la clase de comentarios positivo.

En cuanto al desempeño de la *Natural Language API* de Google, se concluye que la clasificación obtenida es comparable con la clasificación que haría un ser humano, resultando una herramienta útil para el análisis de sentimientos, vale la pena resaltar que los comentarios analizados variaban en su longitud, comprobándose que la herramienta de clasificación automática funciona para textos largos y cortos. Como se comentó anteriormente, la función de análisis de sentimientos es sólo una de las tareas que permite realizar la *API*, en el caso que abordó este trabajo se utilizaron comentarios en el idioma español, el cuál es uno de los idiomas compatibles con la funcionalidad de análisis de sentimientos, pero algunos otros métodos solo están habilitados para el idioma inglés. Entre los trabajos a futuro se propone la experimentación mediante los otros métodos disponibles en la *Natural Language API*, así como experimentación con otras herramientas de procesamiento de lenguaje natural de Google Cloud.

6. Conclusiones

El presente trabajo aporta una metodología de análisis de comentarios referentes a opiniones de turistas o potenciales turistas de un Área Natural Protegida en México. Siguiendo dicha metodología es posible identificar palabras frecuentes que usan los usuarios cuando se expresan de forma positiva o negativa de la región. Dichas palabras se pueden consultar en la Fig. 6 y en la Fig. 10.

Entre las palabras que destacan para la categoría positiva se encuentran: lugar, Tehuacán, hermoso, excelente, reserva, México, bonito, Puebla,

Naturaleza, conocer y país, lo que nos da una idea de los entes de los que se expresan de forma positiva.

Por otro lado para el caso de comentarios negativos, las palabras que se destacan son: más, cerrado, Tehuacán, tehuacancuicatlan, gobierno, región, zona, llegar, reserva y mundo, es posible que los usuarios estén haciendo referencia al estado de acceso a las zonas, ya que los últimos meses ha permanecido cerrada por motivos de la pandemia por la COVID-19.

Los resultados obtenidos muestran que el uso de estas herramientas para identificar los sentimientos de los usuarios de las ANP podrían ser provechosas para determinar mejoras en el aprovechamiento de dichas áreas.

También se compara la herramienta de procesamiento de lenguaje natural de Google para el análisis de sentimientos con un etiquetado manual y se concluye que su uso es útil para este tipo de aplicaciones.

Referencias

1. Cambria, E., Das, D., Bandyopadhyay, S., Feraco, A., et al.: A practical guide to sentiment analysis. Springer (2017)
2. Chaudhuri, A.: Visual and Text Sentiment Analysis through Hierarchical Deep Learning Networks. Springer (2019)
3. CONABIO: Áreas protegidas. <https://www.biodiversidad.gob.mx/region/areasprot> (Aug 2020), consultado: 2021-9-20
4. CONANP: Sistema de consulta fichas ANP. https://simec.conanp.gob.mx/consulta_fichas.php, consultado : 2021 – 9 – 20
5. CONANP: Marco estratégico de turismo sustentable en áreas naturales protegidas de México (METS). <https://www.gob.mx/conanp/documentos/marco-estrategico-de-turismo-sustentable-en-areas-naturales-protegidas-de-mexico> (Feb 2019), consultado: 2021-9-20
6. CONANP: Reserva de la biosfera tehuacán - cuicatlán. <https://www.gob.mx/conanp/documentos/reserva-de-la-biosfera-tehuacan-cuicatlan-209465> (Sep 2019), consultado: 2021-9-20
7. CONANP: Turismo y naturaleza en áreas naturales protegidas. <https://www.gob.mx/conanp/prensa/turismo-y-naturaleza-en-areas-naturales-protegidas-211386> (Aug 2019), consultado: 2021-9-20
8. CONANP: Programa nacional de áreas naturales protegidas 2020-2024. <https://www.gob.mx/conanp/documentos/programa-nacional-de-areas-naturales-protegidas-2020-2024> (Sep 2020), consultado: 2021-9-20
9. CONANP: Áreas naturales protegidas de México. <http://sig.conanp.gob.mx/website/pagsig/> (Mar 2021), consultado: 2021-9-20
10. FLATICON: Free vector icons and stickers. <https://www.flaticon.com/> (2010), consultado: 2021-11-18
11. Gobierno de México: Comisión nacional de áreas naturales protegidas. <https://www.gob.mx/conanp>, consultado: 2021-9-20
12. González Herrera, M.R., Miranda de Sá Teles, R., Bauer, R., Marques Gomes, C.: Tecnologias da informação e da comunicação e suas interfaces com o turismo: alguns estudos de caso brasil e México. Departamento de Ciencias Administrativas (2019)

13. GOOGLE: Cloud computing services. <https://cloud.google.com/>, consultado: 2021-9-20
14. GOOGLE: Google maps. <https://www.google.com.mx/maps/>, consultado: 2021-9-20
15. GOOGLE: How to use cloud natural language API. <https://cloud.google.com/natural-language/docs/how-to>, consultado: 2021-9-20
16. GOOGLE: Places API overview. <https://developers.google.com/maps/>, consultado: 2021-9-20
17. GOOGLE: YouTube data API. <https://developers.google.com/youtube/v3/>, consultado: 2021-9-20
18. Kamath, U., Liu, J., Whitaker, J.: Deep learning for NLP and speech recognition, vol. 84. Springer (2019)
19. Molinar, C.M.A., Espinoza, P.M., Llamas, I.O.: Evaluación de destinos turísticos mediante la tecnología de la ciencia de datos. *Estudios y perspectivas en turismo* **26**(2), 286–305 (2017)
20. Muthuraman, S., Al Hazi, M.: Smart tourism destination-new exploration towards sustainable development in sultanate of oman. In: 2019 5th International Conference on Information Management (ICIM). pp. 332–335. IEEE (2019)
21. Pham, T.D., Vo, D., Li, F., Baker, K., Han, B., Lindsay, L., Pashna, M., Rowley, R.: Natural language processing for analysis of student online sentiment in a postgraduate program. *Pacific Journal of Technology Enhanced Learning* **2**(2), 15–30 (2020)
22. PYTHON: Welcome to python. <https://www.python.org/>, consultado: 2021-9-20
23. Sandoval Ortega, A.S., Alcalá De la O, B., Martínez Morales, J.: Marketing digital: Un análisis del consumidor en México. <http://congreso.investiga.fca.unam.mx/docs/xxiii/docs/14.07.pdf> (Oct 2018), consultado: 2021-9-20
24. SECTUR: Datatur3 - mapa del sitio. <https://datatur.sectur.gob.mx>, consultado: 2021-9-20
25. SECTUR: Boletín 81.- SECTUR y SEGITTUR trabajan coordinadamente para desarrollar el primer destino inteligente mexicano. <http://www.sectur.gob.mx/sala-de-prensa/2015/05/06/boletin-81-sectur-y-segittur-trabajan-coordinadamente-para-desarrollar-el-primer-destino-inteligente-mexicano/> (May 2015), consultado: 2021-9-20
26. SECTUR: Cozumel se convertirá en el primer destino turístico inteligente de México. <https://www.gob.mx/sectur/prensa/cozumel-se-convertira-en-el-primer-destino-turistico-inteligente-de-mexico> (Dec 2015), consultado: 2021-9-20
27. SECTUR: El turismo es una de las actividades con más crecimiento en la economía nacional. <https://www.gob.mx/sectur/prensa/el-turismo-es-una-de-las-actividades-con-mas-crecimiento-en-la-economia-nacional> (Jul 2019), consultado: 2021-9-20
28. SECTUR: Atlas turístico de México. <https://www.atlasturistico.sectur.gob.mx> (2020), consultado: 2021-9-20
29. SETUR, SEGITTUR: Red de destinos turísticos inteligentes. <https://www.destinosinteligentes.es/> (2019), consultado: 2021-9-20
30. UNAM: Infraestructura de datos espaciales abiertos. <https://www.gits.igg.unam.mx/idea/descarga> (2017), consultado: 2021-9-20
31. YouTube: YouTube. <https://www.youtube.com/>, consultado: 2021-9-21

Tecnologías emergentes para la reincorporación de personal a sus actividades durante la pandemia de COVID-19

Magdalena Saldana-Perez¹, Fatima Figueroa-Ayala²,
Carolina Palma-Preciado¹, Marco Moreno-Ibarra¹

Instituto Politécnico Nacional,
Centro de Investigación en Computación,
Mexico

{amagdasaldana, cpalmap2020, marcomoreno}@cic.ipn.mx,
ffigueroaa1500@alumno.ipn.mx

Resumen. Es un hecho que el contagio de la COVID-19 es un riesgo que puede disminuirse al aplicar las normas de sanidad adecuadas, también es importante considerar que adaptarnos a dichas medidas sanitarias, y la necesidad de concientización sobre la existencia de virus es totalmente inminente. Las actividades que de alguna forma se vieron trasladadas a un entorno virtual debido al peligro de contagio del virus SARS-Cov-2 que representaban, de manera gradual han comenzado a regresar a su modalidad presencial, permitiendo que las personas comiencen a interactuar de forma física nuevamente, lo que a su vez ha generado la creación de tecnologías emergentes que colaboren a la nueva dinámica que deben llevar las ciudades. Lo anterior genera la necesidad de buscar metodologías y herramientas que permitan mantener a las personas seguras durante su estancia en recintos, y llevar un control de acceso de personal a cada sitio en el que varias personas deban reunirse. La tecnología es un factor importante en esta nueva adaptación de actividades, y debe ser considerada pues las soluciones que genera permiten a las ciudades dirigirse hacia el paradigma de las Ciudades Inteligentes. En el presente documento se detallan algunas soluciones tecnológicas planteadas para adaptar las actividades humanas a la situación de salud mundial actual relacionada con la COVID-19, con la finalidad de que sean conocidas, y de ser posible implementadas con desarrollos tecnológicos producidos en instituciones públicas. También se presenta un análisis sobre la importancia de la tecnología en el avance de las adversidades provocadas por el virus del SARS- CoV-2.

Palabras clave: SARS-CoV-2, COVID-19, tecnologías emergentes, ciudades inteligentes, readaptación.

Emerging Technologies for Returning Staff to Work During the COVID-19 Pandemic

Abstract. It is a fact that COVID-19 infection is a risk that can be reduced by applying appropriate health regulations. It is also important to consider that adapting to these health measures, and the need for awareness about the existence of viruses, is absolutely imminent. Activities that were somehow transferred to a

virtual environment due to the risk of SARS-CoV-2 infection that they represented, have gradually begun to return to their in-person form, allowing people to begin interacting physically again. This, in turn, has led to the creation of emerging technologies that contribute to the new dynamics that cities must navigate. This creates the need to find methodologies and tools that allow people to stay safe during their stay in facilities and to monitor staff access to each location where several people must meet. Technology is an important factor in this new adaptation of activities and must be considered because the solutions it generates allow cities to move toward the Smart Cities paradigm. This document details some proposed technological solutions for adapting human activities to the current global health situation related to COVID-19, with the aim of making them known and, if possible, implemented with technological developments produced in public institutions. It also presents an analysis of the importance of technology in addressing the adversities caused by the SARS-CoV-2 virus.

Keywords: SARS-CoV-2, COVID-19, emerging technologies, smart cities, retrofitting.

1. Introducción

La pandemia ocasionada por el virus SARS-CoV-2, que se convirtió en un problema mundial desde comienzos de 2020, hizo que los gobiernos solicitaran a sus habitantes evitar cuanto fuese posible el dejar sus hogares y aglomeraciones, con la finalidad de mantenerlos fuera del riesgo de contagio. Derivado de la solicitud gubernamental, empresas, escuelas, centros comerciales, cines, parques y múltiples comercios, se vieron en la necesidad de migrar sus actividades a un ambiente completamente virtual, pensando en ello como una solución temporal.

Pese a la adaptación de diversas actividades al entorno virtual, el mantener la situación, en muchos casos se volvió insostenible, pues el avance en diferentes áreas no era equivalente al avance que de forma presencial se habría obtenido, además de que, de acuerdo con el estudio psicológico descrito en [1, 10], las personas requieren del contacto y de la presencia de otras personas, a fin de mantener el equilibrio mental necesario para desarrollar de forma eficiente sus actividades.

Considerando lo anterior, y con la desocupación hospitalaria derivada de los cuidados de los ciudadanos para evitar la propagación del virus SARS-CoV-2, las autoridades comenzaron a permitir la apertura de establecimientos y centros de trabajo respetando ciertas reglas de convivencia como: la sana distancia, el uso de cubrebocas y gel antibacterial, y manteniendo una capacidad máxima del 30% en los recintos. Sin embargo, es necesario considerar la existencia de los super dispensadores del virus que provoca la COVID-19, y las personas asintomáticas, que, de manera indirecta son capaces de contagiar a otras personas el virus SARS-CoV-2, aún sin saberlo [2, 11, 12].

Si bien el virus ha demostrado a lo largo de los últimos meses que puede variar su genoma [3, 13], y que los efectos que causan cada una de sus variantes son diferentes, las reacciones de cada organismo humano también lo son; se sabe que entre los síntomas de contagio iniciales se encuentran el dolor de garganta, dolor de cabeza y fiebre, mismos que van empeorando hasta causar baja de oxigenación del paciente y daño pulmonar [4, 14].

En el presente documento se describen algunas soluciones tecnológicas que han surgido durante el periodo de pandemia con la finalidad de permitir la continuidad de las actividades diarias. Además, se detalla una solución de control de acceso que permite el monitoreo de síntomas de COVID-19, diseñada para escuelas públicas.

El trabajo se conforma de cinco secciones, en la sección 1 la introducción, en la sección 2 se muestran los antecedentes que preceden al trabajo planteado para el control de acceso, en la sección 3 se plantea la metodología propuesta, en la sección 4 se presentan los resultados del proceso; finalmente en la sección 5, se detallan las conclusiones obtenidas de la propuesta.

2. Antecedentes

La pandemia ocasionada por la propagación desmedida del virus que provoca la enfermedad denominada COVID-19 ha traído consigo una crisis de salud pública que ha afectado durante un largo periodo de tiempo, trayendo así la necesidad de adaptación a una nueva normalidad. Es por ello que, los sectores tanto públicos como privados, han decidido enfocar sus recursos al diseño de estrategias seguras para esta nueva etapa en la historia de la humanidad. Un ejemplo es la estrategia de control propuesta por la consultora Gartner Inc. [5]. La cual plantea una propuesta de solución para facilitar el control antes y durante situaciones de crisis. A su vez entre los aportes más destacados, se encuentran los desarrollados por la comunidad científica y tecnológica, ya que a lo largo de este tiempo se ha reconocido la importancia de la telemedicina en conjunto con la inteligencia artificial para el desarrollo de diversos servicios tecnológicos como: aplicaciones web o móviles, chat-bots, robots y dispositivos destinados a realizar mediciones y monitoreo de algunos de los síntomas característicos de la enfermedad.

Un claro ejemplo es el oxímetro electrónico, dispositivo que permite medir el nivel de saturación del oxígeno en la sangre de la persona monitoreada sin necesidad de que la medición sea realizada en el hospital, y el termómetro infrarrojo, que permite medir la temperatura de las personas sin tener contacto directo con ellas, de modo que se pueda monitorear las variaciones de temperatura que tiene una persona a lo largo de los días, y que ha resultado una herramienta útil pero insuficiente en el control de acceso a lugares públicos. Si bien el monitoreo que puede realizarse por este tipo de dispositivos electrónicos no es suficiente para evitar contagios, ha permitido de manera efectiva disminuir la propagación del virus, pues permiten identificar de forma rápida a personas que posiblemente han sido infectadas y se encuentran ya en la fase de contagio, y han ayudado en evitar las aglomeraciones en hospitales.

En el proyecto Detección y seguimiento del contagio mediante tecnologías IoT-Edge: Enfrentando la pandemia COVID-19 [6], desarrollado en Estambul, Turquía, se aprovecharon funcionalidades de diversas tecnologías como lo son el Internet de las cosas (Internet of Things, IOT), el análisis de grandes conjuntos de datos (Big Data) y la inteligencia artificial (Artificial Intelligence, AI). Los autores proponen una estrategia tecnológica para el control de acceso a recintos que consiste en el desarrollo de un mecanismo de comunicación, que mediante el uso de las capas de orden y nube empleando la tecnología IOT y Cómputo de borde (Edge computing), obtienen datos de sensores de temperatura y oxigenación, que son procesados por un microprocesador y transmitidos a base de datos; finalmente, la capa multi edge del modelo realiza el

análisis casi instantáneo de la salud de la persona. El proyecto cuenta con una aplicación móvil que notifica el seguimiento de la persona.

Por otro lado, en Eslovaquia investigadores de la facultad de electromagnetismo y biomédica en la universidad de Ziliana [7], publicaron un artículo donde explican su propuesta para la adaptación de actividades durante la pandemia de la COVID-19. Su propuesta consiste en un software y un hardware destinados a automatizar la toma de temperatura, así como la revisión del uso de cubre bocas. El hardware propuesto fue diseñado de manera que fuera fácil de colocar en cada acceso del recinto de prueba, por medio de un dispositivo electrónico conectado a una cámara, se analiza el rostro de la persona y se muestran los datos obtenidos en una pantalla visible para el usuario.

En el artículo antes mencionado se propone el escaneo y análisis de rostro de la persona, con la finalidad de determinar si el usuario que desea ingresar al recinto tiene la mascarilla puesta, dicho dispositivo también cuenta con un lector de temperatura con el cual realizar la medición de temperatura en tiempo real, en caso de que se detecte que el usuario no cumple con las normas adecuadas para el ingreso, se le niega el acceso al complejo, y se emite una alerta. El funcionamiento del dispositivo depende de una tarjeta controladora raspberry Pi 4, un módulo con cámara v2, un módulo de medición de temperatura tipo MLX90614, un sensor de distancia V153L0X y por último un módulo LCD [7].

En resumen, la información obtenida por el dispositivo de hardware es procesada en tiempo real con ayuda del sistema tipo Laboratorios rápidos de ciencia terapéutica (Rapid Therapeutic Science Laboratory, RTSL), guardando la información en una base de datos MySQL, simultáneamente y con ayuda de los sensores de presencia se toma la información en un perímetro de 4 metros, en el caso de no detectarse nada, el dispositivo se mantiene en un estado de espera hasta que haya un cambio, si el sistema está en pausa se proyecta información general en la pantalla. Adicionalmente al diseño de este dispositivo, se propone el proyectar en la pantalla mediante una barra de progreso el análisis de la temperatura que indica al usuario el tiempo restante de espera para concluir la evaluación.

En cuanto a la detección de la mascarilla, la codificación sigue la estructura de un algoritmo híbrido, el cual implementa el uso de redes neuronales (Neural Network, NN) y un análisis del vector de las características del sistema, toma un enfoque de histograma de gradientes orientados (HGO), este proceso disminuye la resolución de los píxeles con el fin de realizar el reconocimiento facial, para esta tarea se utilizaron modelos faciales disponibles gratuitamente. Finalmente, para la evaluación de la temperatura sólo se usó el dispositivo que realiza la medición, donde en cuanto el sensor detecta presencia se activa.

De forma similar en la universidad australiana Sunshine Coast se realizó una investigación enfocada en la implementación de la inteligencia artificial con el propósito de lograr detectar los casos sospechosos de COVID-19, caracterizándose por recopilar los datos obtenidos de las mediciones de los síntomas en tiempo real [8].

El sistema se compone de sensores infrarrojos destinados a la detección de obstáculos realizando un conteo de las personas que entran y salen, para esta tarea se considera el aforo del recinto donde se desea instalar, es decir, en el momento de ingreso al establecimiento el usuario debe responder un cuestionario referente a los síntomas de COVID-19, al terminar es enviado a un servidor controlado por una Raspberry Pi.

La microcomputadora se encuentra configurada de tal manera que se encarga de actualizar la base de datos con la información recopilada diariamente. Una vez que el usuario termina de llenar el cuestionario avanza un lugar en la fila para colocarse en la posición señalada donde se encuentra el sensor de temperatura IR MLX90614, destinado a medir su temperatura corporal. Así mismo cuando se forma la fila para entrar al recinto se utilizan leds direccionables WS2812B, con el fin de señalar al usuario el lugar en el que se debe ubicar conforme avance la fila, si llegan a la entrada dos personas, el sistema marcará los lugares permitidos para que se sitúen a una sana distancia uno del otro. A medida que ingresan más personas el número de ubicaciones disponibles se actualiza manteniendo una distancia mínima de 1.5 metros entre los lugares ocupados.

Como se observa, muchos son los desarrollos tecnológicos enfocados en el control de acceso a recintos, sin embargo, todos ellos además de útiles resultan relevantes pues muestran distintas soluciones que abordan el mismo problema desde diferentes perspectivas, y que han sido adaptados a las necesidades de los lugares en los que han surgido.

Cabe resaltar que justo en la pandemia, tanto medios de divulgación científicos como repositorios y revistas científicas, así como instituciones de investigación han estado desarrollando desde diferentes frentes, soluciones, estudios y análisis, para comprender de forma conjunta el fenómeno de la pandemia, sus consecuencias y posibles soluciones, por lo que gran número de desarrollos se han centrado en el trascendental tema de la COVID-19.

3. Metodología

En la presente propuesta, se plantea un proyecto para el control de acceso de personas a recintos públicos, que combina componentes de hardware con software para automatizar la autorización de acceso y prevenir el contacto entre personas.

La metodología propuesta se compone de tres factores determinantes referentes al diseño estructural de la arquitectura, estos son: administrador, software y hardware ya que en conjunto se busca obtener un sistema eficiente de monitoreo y realizar el seguimiento de personas que posiblemente estén infectadas con el virus SARS-CoV-2, mediante el monitoreo de su temperatura diaria al momento de acceder al recinto.

La metodología del proyecto se compone de cinco etapas: registro, identificación, recopilación, almacenamiento y decisión, tal como se puede observar en el diagrama de la Fig.1.

Para cumplir con la metodología propuesta se ha diseñado una base de datos en MySQL Workbench, destinada a almacenar la información obtenida tanto de la etapa de registro como de la etapa de recopilación. En la fase de registro se utiliza una aplicación de Internet enriquecida (RIA), por medio de esta aplicación y con la intervención del administrador se llena un formulario previamente diseñado, con la finalidad de recopilar los datos de interés para identificar la identidad de la persona interesada en ingresar al recinto.

La RIA es el medio por el cual el administrador interactúa con la base de datos, reservando un espacio de memoria en la base de datos para cada persona registrada, ese espacio de memoria es ocupado por las mediciones realizadas cada vez que la persona

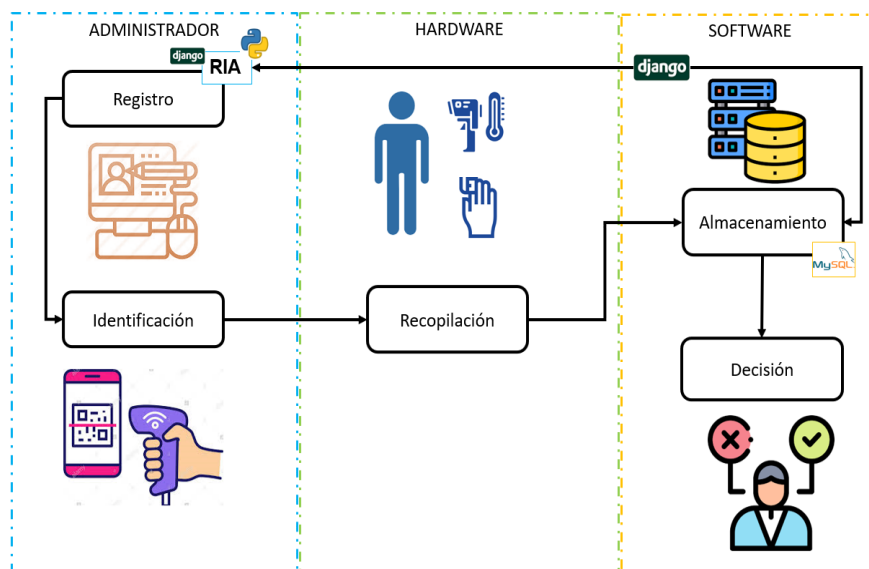


Fig. 1. Diagrama de la metodología propuesta para la solución del problema, en él se muestran las cinco etapas que componen la ejecución del proyecto, desde la toma de temperatura de la persona, hasta la toma de decisión del sistema.

denominada usuario desee ingresar al lugar, generando con esto un historial personalizado de las mediciones obtenidas.

En la base de datos del sistema se cuenta con un registro básico de las personas que regularmente acuden al recinto, a cada persona le es asignado un código de identificación de respuesta rápida (QR, Quick response code), el cual es presentado por el usuario al administrador cada vez que este desee ingresar al recinto (la etapa de registro solo se realiza una vez por persona).

En caso de que la persona que desea acceder sea un visitante, se le asigna un código QR temporal de visita. Para lograr la comunicación de la RIA con la base de datos se hace uso del entorno de desarrollo de Django [9].

Al concluir el registro se inicia la etapa de identificación, el usuario muestra al administrador su código de identificación QR, el administrador lo escanea por medio de una pistola lectora QR y al mismo tiempo el sistema realiza una búsqueda de coincidencias en la base de datos, si se encuentra en la RIA aparece el nombre y los datos de la persona, de otra forma el administrador invitara a la persona a que realice su primer registro.

Posteriormente se llega a la etapa de recopilación, en la cual intervienen dos dispositivos diseñados en este proyecto, uno destinado a realizar la medición de temperatura donde respectiva arquitectura se ilustra en la Fig.2. En la Fig. 3, se muestra la conexión entre uno de los microprocesadores Arduino empleados en el proyecto y el sensor de medición de oxigenación.

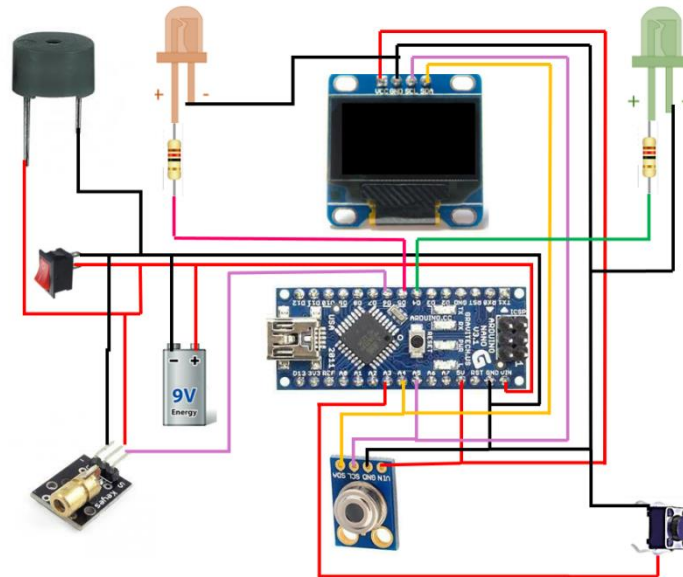


Fig. 2. Esquema de interconexión que existe entre los componentes electrónicos que permiten la medición de temperatura de la persona que desea ingresar al recinto. El cerebro del circuito lo representa el Arduino al que están conectados el sensor infrarrojo de temperatura, la bocina para alertar si la temperatura del visitante es mayor a 37°C, y leds indicadores.

En el proceso para medir la temperatura intervienen diversos componentes electrónicos de los cuáles el Arduino resulta ser el núcleo del dispositivo; el Arduino es un microcontrolador con un ambiente de desarrollo cuyo microprocesador se basa en el microcontrolador ATmega328P, la placa de desarrollo cuenta con 14 terminales digitales y 6 terminales analógicas programables desde el entorno de desarrollo del Arduino.

Para lograr la comunicación entre el Arduino y el sensor de temperatura infrarrojo MLX90614 seleccionado para el proyecto, se realiza una conexión serial SCL (sistema de reloj, system clock) del sensor a la terminal A5, y una conexión SDA del sensor a la terminal A4 del Arduino. Para precisar, la conexión SCL es la línea de pulsos de reloj que sincroniza el sistema, por su parte la conexión SDA (Sistema de datos, system data) permite transportar datos entre los dispositivos. La medición de temperatura obtenida se muestra en una pantalla de diodo orgánico de emisión de luz (OLED, organic light-emitting diode).

El funcionamiento del circuito es el siguiente, primero se activa el mecanismo presionando el interruptor, posteriormente se coloca el usuario frente al sensor de temperatura a una distancia aproximada de 5 cm, cuidando de que la luz láser apunte a su frente ya que esta será la referencia de que se está posicionando de manera correcta.

Una vez que el sensor detecte la presencia del usuario realizará la respectiva medición reflejando el resultado en la pantalla OLED. Internamente el sistema ejecuta

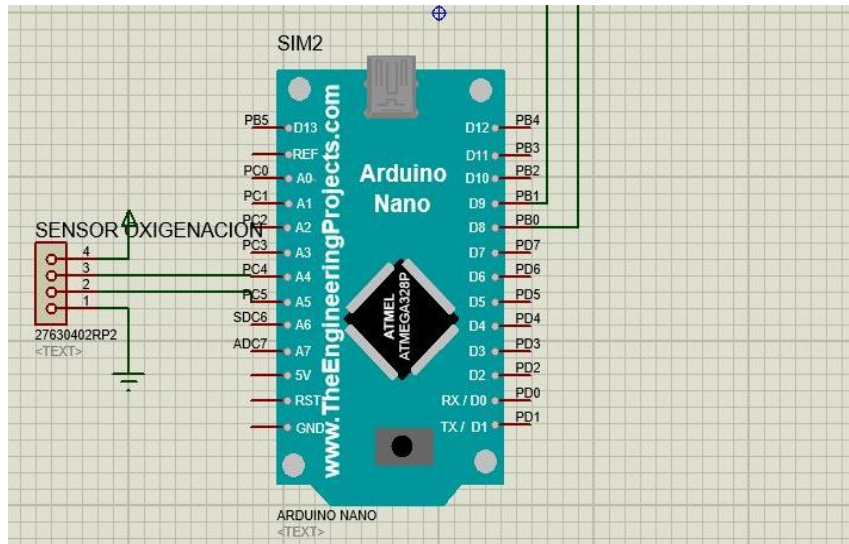


Fig. 3. El esquema de conexión entre el sensor que mide la oxigenación del visitante y el Arduino. El diagrama fue elaborado en Proteus.

una validación de los datos obtenidos, en dicha validación intervienen dos opciones, si el usuario excede la temperatura permitida se activará el LED rojo en señal de acceso denegado y como medida de prevención adicional también se emitirá una alarma emitida por la bocina conectada al circuito, en caso contrario, cuando el usuario presente temperatura normal, se activará el LED verde haciendo referencia a que todo está bien y el usuario puede ingresar al recinto de manera segura.

Para el dispositivo de oxigenación se utiliza otro Arduino, se implementa un Arduino para cada parámetro de medición, para que estos se puedan posicionarse separados de forma estratégica en lugares diferentes del módulo de acceso seguro en el que se utilizarán, de otro modo sería necesario cablear los sensores al Arduino, generando impedancia que podría modificar la calidad de las mediciones.

El dispositivo de oxigenación se compone de un Arduino el cual se encarga de la parte lógica y un sensor de pulso MAX30102 cuyo trabajo es detectar la presencia de la persona, cuando esté es cubierto en su totalidad por el dedo índice del usuario, para así realizar la respectiva medición. El sensor de pulso es un pulsioxímetro y monitor de frecuencia cardíaca, que se caracteriza por contar con un sensor infrarrojo y detectores fotoeléctricos.

El dispositivo Arduino encargado de la medición de oxigenación, está conectado en serie mediante el puerto USB a la computadora designada para el monitoreo del administrador, dicha acción ayuda a entablar la comunicación entre el Arduino y la computadora logrando que se pueda visualizar en la pantalla del equipo el valor de la medición obtenida y registrarla en la base de datos.

La medición de la oxigenación funciona de la siguiente forma, una vez que concluye la medición de temperatura el administrador indica el usuario dónde colocar su dedo. El sistema muestra al administrador por medio de una ventana los resultados obtenidos

de la medición. Una vez concluido el proceso, el administrador informa al usuario su nivel de oxigenación finalizando así el proceso de medición.

Una vez realizada esta acción se recopilan los datos obtenidos y son registrados por el administrador en la RIA, guardándose automáticamente en los espacios de memoria previamente reservados en la primera etapa, acción que cumple la etapa de almacenamiento.

Una vez concluida la etapa de almacenamiento se llega a la etapa de decisión, en esta parte el sistema internamente realiza una comparación de los resultados obtenidos en cada medición, en caso de que el sistema detecte que alguno de los datos registrados por el administrador no cumple con la condición previamente establecida en la codificación del mismo, emite una alerta para avisar al administrador que el usuario no cumple con los requerimientos establecidos para ingresar al sitio. Por otra parte, en caso de que los datos sí sean válidos, no se emitirá ninguna alerta y solo guardará la medición de sus parámetros en la base de datos.

4. Resultados

La relevancia del presente proyecto recae en que su inclusión en los accesos a recintos públicos y privados permitirá tener un mayor control de los posibles síntomas que una persona infectada de con el virus SARS-COV-2 podría presentar, y evitar así que entre en contacto con otras personas, propagando el virus.

El proyecto, además, atiende una de las necesidades más imperiosas que se presentan en la adaptación a esta nueva normalidad, el control de personas en lugares cerrados, y la evasión de contacto entre personas no relacionadas entre sí. Es importante resaltar que la automatización de este tipo de tareas sigue también la idea del desarrollo de ciudades inteligentes.

Se ha demostrado que el virus se propaga principalmente por las gotículas expelidas por la persona al hablar toser o reírse, es decir, si el usuario interesado en ingresar al recinto sólo muestra su código de identificación reduce al máximo la probabilidad de contagio además de agilizar el proceso de entrada al recinto disminuyendo el aforo de personas en la recepción [15].

Como se estableció en la metodología, el proyecto se compone de dos elementos, una de hardware, compuesta por el circuito electrónico que permite llevar a cabo la medición de temperatura y oxigenación, y una de software, que comprende la aplicación web y base de datos del proyecto.

Una vez desarrollados los pasos establecidos en la metodología para el desarrollo del componente electrónico, se logró ensamblar el componente para la medición de temperatura, como se muestra en las Fig. 4a y 4b. Del mismo modo, el dispositivo para la medición de oxigenación se muestra en la Fig. 5. Cabe mencionar que en ambos casos se colocaron los circuitos dentro de cajas de polímero, de forma que pudieran colocarse en el control de acceso, sin sufrir riesgos como desconexión o exposición de los componentes electrónicos.

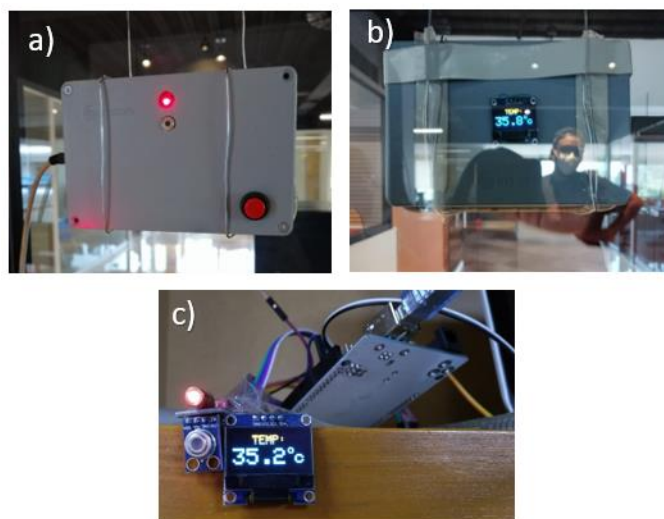


Fig. 4. Componente electrónico para medición de temperatura. La imagen a) muestra el led que debe situarse en la frente de la persona que desea ingresar, para realizar una correcta medición. La figura b) muestra la pantalla OLED que visualiza el visitante mostrando la temperatura medida. La imagen del inciso c) muestra los componentes electrónicos antes de ser protegidos por la carcasa de polímero.

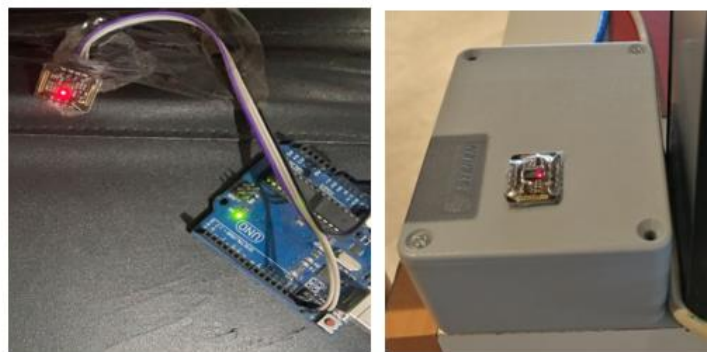


Fig. 5. La imagen de la izquierda muestra los componentes electrónicos conectados para realizar la correcta medición de la oxigenación del visitante, por medio del sensor de contacto. La imagen de la derecha muestra el dispositivo ya ensamblado en una carcasa de polímero para protección de los componentes.

Para el mejor manejo de la herramienta propuesta se diseñó una aplicación web encargada de mostrar una interfaz para el rápido registro de los usuarios, así como para la visualización de los parámetros medidos al momento de su ingreso. Este sistema es



Fig. 6. Instructivo con los pasos que se siguen para realizar la medición de la temperatura y la oxigenación del usuario, en el proyecto planteado.

	A	B	C	D	E	F	G	H	I	J
	NOMBRE	APELLIDOS	NOMBRE DE USUARIO	ULTIMA SESION	BOLETA	TIPO DE USUARIO	BOTON	BOTON	BOTON	BOTON
1				12 de Mayo de 2021 a las 22:26		ESTUDIANTE	EDITAR	AGREGAR	HISTORIAL	ELIMINAR
2				7 de Mayo de 2021 a las 14:35		DOCENTE	INABILITADO	INABILITADO	HISTORIAL	ELIMINAR
3				7 de Mayo de 2021 a las 11:30		DOCENTE	EDITAR	AGREGAR	HISTORIAL	ELIMINAR
4				8 de Mayo de 2021 a las 3:00		ADMINISTRACIÓN	EDITAR	AGREGAR	HISTORIAL	ELIMINAR
5				12 de Mayo de 2021 a las 22:29		ADMINISTRACIÓN	INABILITADO	INABILITADO	HISTORIAL	ELIMINAR
6				12 de Mayo de 2021 a las 22:32		DOCENTE	EDITAR	AGREGAR	HISTORIAL	ELIMINAR

Fig. 7. Muestra de datos capturados en el sistema diseñado para el proyecto.

monitoreado por un administrador. En la Fig. 6, se muestra un instructivo con los pasos ejecutados en el proceso de medición de temperatura y medición de oxigenación.

Hasta el momento, las pruebas realizadas en el sistema se han llevado a cabo en el Centro de Investigación en Computación, por lo cual se generaron tres perfiles para los posibles usuarios: administrador del sistema, personal, y estudiantes. El perfil de administrador es el más completo, pues tiene la capacidad de acceder a los datos de los otros usuarios para verificar que todo esté funcionando correctamente, y en caso de que un usuario sea identificado como posible portador del virus, colocarlo en cuarentena, o

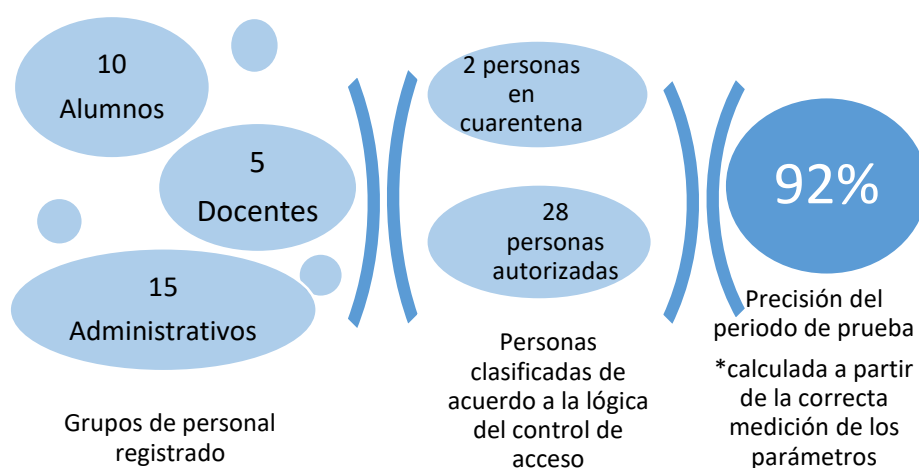


Fig. 8. Diagrama de algunos datos duros obtenidos del uso del dispositivo como control de acceso al área de administración de recursos del Centro de Investigación en Computación.

bien sacarlo de la pila de casos de cuarentena cuando haya pasado su incapacidad y muestre un certificado médico que lo acredite.

La Fig. 7, presenta un extracto de la base de datos del sistema implementado para el proyecto, como se observa, se tiene un campo en el que se monitorea si las mediciones de temperatura y oxigenación de la persona se encuentran dentro de un rango saludable, o si bien parecen ser mediciones de alguien que podría estar contagiado de COVID-19.

Para que un usuario sea colocado en la pila de cuarentena, deberán ocurrir las siguientes condiciones, que el usuario presente una variación considerable en sus registros de temperatura alcanzando mediciones mayores a 37.8 °C, tenga niveles de oxigenación menores al 92%, o bien, presente ambas condiciones. Cuando se cumpla alguna de las tres situaciones sonará una alarma conectada al sistema que alertará a las personas cercanas al acceso, para que la persona sea remitida a atención médica oportuna.

Cabe resaltar que solo el administrador es capaz de mover un registro de la pila de acceso denegado, a la de acceso permitido, y que la revisión del certificado médico que avala dicho movimiento de registros en el sistema se hace de forma manual a través de la terminal colocada en el acceso al recinto. Es también el administrador, quien en caso de que se presente un nuevo usuario no registrado en la base de datos como se ilustra en la Fig. 7, tendrá los permisos para darlo de alta en el sistema y asignar el código.

QR. Por medio de los perfiles de estudiante y personal, solo es posible consultar el historial del usuario generado con cada ingreso del usuario al recinto.

El sistema se probó con un grupo de 30 personas entre las que se cuenta con personal docente, administrativo y alumnos, que acudieron de forma controlada al Centro de Investigación en Computación entre el 1° y el 30 de junio de 2021. Del uso temporal

del sistema en el Centro de Investigación en Computación se obtuvieron los datos duros que se presentan en la Fig. 8.

En la figura anterior se observa la representación visual de los diferentes grupos de estudio, estos se dividen en 3 tipos: Alumnos, docentes y administrativos sumando un total de 30 pruebas. Seguido de eso se observa la clasificación de personas de acuerdo con la lógica del sistema, dando como resultado 28 personas autorizadas y 2 personas en cuarentena.

Donde, las personas que cumplan con las condiciones de acceso previamente establecidas en la lógica del sistema, se consideran personas autorizadas y las personas que no cumplan con las condiciones se consideran personas en cuarentena.

Cabe destacar que los 2 casos clasificados como personas en cuarentena son canalizadas de inmediato a la instancia correspondiente para realizarle una prueba rápida, el resultado de la misma debe ser notificado al administrador quien dará seguimiento al caso en cuestión, protegiendo la integridad y anonimato del caso positivo.

5. Conclusiones

En este trabajo se presenta una metodología que mediante el uso de software y hardware libre propone un sistema de monitoreo y seguimiento de personas para controlar su acceso a espacios interiores promoviendo el uso exitoso de instrumentos de medición como oxímetros y termómetros para detectar a posibles individuos que estén infectados de COVID-19.

La metodología compuesta de cinco etapas: registro, identificación, recopilación, almacenamiento y decisión, permite facilitar el proceso de control de interacciones entre personas en espacios cerrados. Debido a los criterios de registro, las personas proporcionan datos valiosos que permiten identificar a personas que probablemente estén infectadas de SARS-CoV-2 y canalizarlas adecuadamente.

El uso de una página web como método de registro, y la asignación de código de identificación QR, disminuyen la probabilidad de contagio entre el personal de recepción del edificio y el usuario interesado en entrar. Además, la automatización de esta tarea ayuda a agilizar el flujo de personas en la recepción, evitando así las aglomeraciones.

La técnica de almacenamiento utilizada en el manejo de la base de datos, facilitan la seguridad y administración de la información, aunado a la ayuda que ofrece el rol del administrador o personal de recepción designados, se logra un monitoreo confiable de los síntomas de las personas en el edificio.

Para finalizar el proyecto desarrollado resuelve el problema de monitoreo de personas en instalaciones, el cual puede ser replicado para usar en diferentes edificios de la institución, permite generar un seguimiento de los posibles casos y casos confirmados de COVID-19, e implementa varios pasos de verificación para impedir el paso a los casos que no cumplen con los requisitos mínimos necesarios para ingresar a un área.

Referencias

1. Valdés, M.A.P., Morales, N.E.Á., Cárdenas, A.E.R.: Repercusión psicológica y social de la pandemia COVID-19. *Medimay*, 27(2), pp. 252–261 (2020)
2. Víctor, R.Z., Eder, W.P., Rocío, L.D.: COVID-19: El enigma de los caminantes asintomáticos. *Rev Mex Patol Clin*, 67(1), pp.59–60 (2020)
3. Avendaño, L.F.: ¿Es controlable la COVID-19?. *Neumología Pediátrica*, 16(3), pp. 95–96 (2021) doi: 10.51451/np.v16i3.438.
4. Marín, J.E.O.: Sintomatología no respiratoria de COVID-19. *Alerta, Revista científica del Instituto Nacional de Salud*, 4(2), pp. 61–68 (2021) doi: 10.5377/alerta.v4i2.9923.
5. Guía de mercado para soluciones de gestión de crisis/emergencias y regreso seguro al trabajo COVID-19. <https://www.gartner.com/guest/purchase/registration?resId=3989277>.
6. Ashraf, MU, Hannan, A., Cheema, SM.: Detección y seguimiento del contagio mediante tecnologías IoT-Edge: Enfrentando la pandemia COVID-19. In: 2020 International Conference on Electrical, Communication and Computer Engineering, IEEE, pp. 1–6 (2020) doi: 10.1109/ICECCE49384.2020.9179284.
7. Bashir, A., Izhar, U., Jones, C.: IoT-Based COVID-19 SOP Compliance and Monitoring System for Businesses and Public Offices. In *Engineering Proceedings, Multidisciplinary Digital Publishing Institute*, 2(1) (2020) doi: 10.3390/ecsa-7-08267.
8. Barabas, J., Zalman, R., Kochlan, M.: Automated Evaluation of COVID-19 Risk Factors Coupled with Real-Time, Indoor, Personal Localization Data for Potential Disease Identification, Prevention and Smart Quarantining. In: *IEEE 43rd International Conference on Telecommunications and Signal Processing*, pp. 645–648 (2020) doi: 10.1109/TSP49548.2020.9163461.
9. Django sitio oficial: <https://www.djangoproject.com> (2020)
10. Saladino, V., Algeri, D., Auriemma, V.: The Psychological and Social Impact of Covid-19: New Perspectives of Well-Being. *Frontiers in Psychology*, 11 (2020) doi: 10.3389/fpsyg.2020.577684.
11. Stein, R.A.: Superspreaders in Infectious Diseases. *International Journal of Infectious Diseases*, 15(8), pp. 510–513 (2011) doi: 10.1016/j.ijid.2010.06.020.
12. Beldomenico, P.M.: Do Superspreaders Generate New Superspreaders? A Hypothesis to Explain the Propagation Pattern of COVID-19. *International Journal of Infectious Diseases*, 96, pp. 461–463 (2020) doi: 10.1016/j.ijid.2020.05.025.
13. Liu, Y., Yan, L.-M., Wan, L.: Viral Dynamics in Mild and Severe Cases of COVID-19. *The Lancet Infectious Diseases*, 20(6), pp. 656–657 (2020) doi: 10.1016/S1473-3099(20)30232-2.
14. World Health Organization: World Health Organization. Obtenido de WHO Coronavirus (COVID-19) (2021) https://covid19.who.int/?gclid=CjwKCAjwoZWHBhBgEiwAiMN66SxMOoXtbNyo2oFAMR16sVHhxMVz2GogHfk19wa27xBy567kyS0_FB0CTRQQAyD BwE.
15. Coello Barzola, K.J., Naranjo Segovia, A.S.: Percepciones de cuidadores familiares frente al cuidado de personas adultas con discapacidad mental e intelectual severa ante el confinamiento por la crisis sanitaria COVID-19 entre marzo y mayo de 2020 en la ciudad de Guayaquil (2020)

A First Approach to Asynchronous-Synchronous Tradeoff in 1D Cellular Genetic Algorithms

José Alejandro Cornejo Acosta¹, Jesús García Díaz^{1,2}

¹ Instituto Nacional de Astrofísica, Óptica y Electrónica,
Coordinación de Ciencias Computacionales,
Mexico

² Consejo Nacional de Ciencia y Tecnología,
Mexico

`alexcornejo@inaoep.mx, jesus.garcia@conacyt.mx`

Abstract. This paper explores the effect of the asynchronous-synchronous tradeoff in Cellular Genetic Algorithms (cGA). We perform this exploration by introducing an update policy called *k*-Fixed Line Sweep (*k*FLS). We tested the *k*FLS in a cGA over different uni-modal and multi-modal continuous optimization functions. The results show that for some cases, the cGA can converge more quickly to an optimal solution if the value of *k* is appropriately adjusted.

Keywords: Cellular genetic algorithms, asynchronous update, fixed line sweep.

1 Introduction

Cellular Genetic Algorithms (cGA) are a class of Evolutionary Algorithms (EA) that evolve a population of individuals, with overlapped neighborhoods, by applying genetic operators over a *small* neighborhood of each individual. cGAs have been used for finding near-optimal solutions to different kinds of optimization problems. This includes combinatorial [4] and multi-objective optimization [5, 6] problems. One of the advantages of cGAs is that parallelism mechanisms can be directly used. Usually, the individuals of a cGA are distributed in a regular 1D, 2D, or 3D grid [1]. Nevertheless, this paper only considers 1D grids with closed boundary conditions (a.k.a. rings).

There are many parameters that can influence a cGA's performance. These include the neighborhood, topology, and update policy. This work focuses on the latter. An update policy indicates the order in which individuals are explored and how they affect each other. Depending on the used policy, an individual's codification can spread differently across the grid. An update policy can be asynchronous or synchronous. Synchronous (SYN) update consists in independently updating all the cells and then replacing the whole population. In an asynchronous update, the cells are sequentially updated, immediately replacing the individuals. The main asynchronous update policies reported in the literature are the following [2]:

- *Uniform Choice* (UC). The cell to be updated is selected uniformly at random, with replacement.
- *Fixed Line Sweep* (FLS). The cells are updated row by row. The cells in the sequence are adjacent in the ring.
- *Fixed Random Sweep* (FRS). It is similar to FLS. However, the sequence of cells is generated at random and, once defined, it never changes.
- *New Random Sweep* (NRS). It is similar to FRS. However, every time a sequence is fully explored, a new random sequence is sampled.

Independently of the update policy, a *time step* is usually defined as updating all the cells in the grid. However, in this paper, we define a *time step* as the action of synchronously updating a subset of cells, i.e., not all of the cells are necessarily explored. This way, we set the basis for empirically investigating the tradeoff between asynchrony and synchrony. In more detail, this work introduces a generalization of the FLS update policy. We refer to this proposal as *k*-Fixed Line Sweep (*k*FLS). In fact, *k*FLS with $k = 1$ is equivalent to FLS and is equivalent to SYN when $k = n$, where n is the length of the ring. The following section introduces *k*FLS in detail.

The remaining part of the paper is structured as follows. Section 2 introduces the *k*FLS update policy. Then, Section 3 presents the experimental setup and results. Finally, Section 4 presents the concluding remarks and possible future work.

2 *k*-Fixed Line Sweep

This paper explores an update policy named *k*-Fixed Line Sweep (*k*FLS). This policy is straightforward. It is similar to FLS; however, it synchronously explores k adjacent cells instead of exploring cells one by one. Thus, *k*FLS with $k = 1$ (1FLS) is equivalent to FLS, and *k*FLS with $k = n$ is equivalent to SYN, where n is the number of cells. Since we are working with a 1D grid, the number of cells and the length of the ring are the same. Figure 1 shows an example of how FLS works. In these diagrams, each row represents a time step. Namely, a step where all the black cells are synchronously updated. In these figures, the length of the ring is 11, and the number of time steps is 13. Observe that the number of explored cells in Figure 1(a) is 13, and it is 65 in Figure 1(b).

As mentioned before, FLS and SYN are extreme cases of *k*FLS. In order to explore the effect of different values of k , we applied *k*FLS over a set of continuous optimization problems. The following section presents the experimental setup and results.

3 Experimental Results

This section presents the results of using the *k*FLS update policy over a cGA for solving a set of continuous single-objective optimization benchmark functions. These functions are separated into uni-modal and multi-modal (Tables 1 and

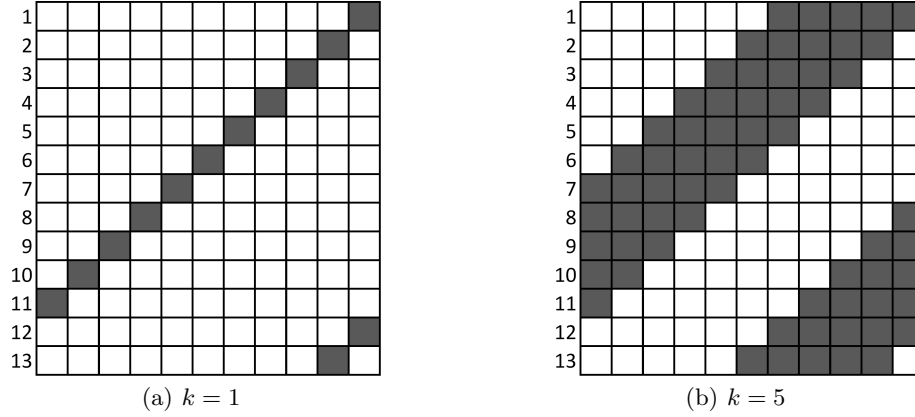

Fig. 1. k FLS.

Table 1. Benchmark uni-modal functions.

Name	Function	Domain
Sphere	$\sum_{i=1}^d x_i^2$	$[-5.12, 5.12]$
Sum Squares	$\sum_{i=1}^d ix_i^2$	$[-10, 10]$
Rotated Hyper-Ellipsoid	$\sum_{i=1}^d \sum_{j=1}^i x_j^2$	$[-65.536, 65.536]$
Zakharov	$\sum_{i=1}^d x_i^2 + \left(\sum_{i=1}^d 0.5ix_i \right)^2 + \left(\sum_{i=1}^d 0.5ix_i \right)^4$	$[-5, 10]$

2). Table 3 shows a simple configuration used for all experiments. The ring has closed boundary conditions, and its length is 100. So, we test k FLS with values $k \in \{1, 2, \dots, 100\}$. The selected neighborhood is one of the simplest possible, EAST-WEST (left and right cells). Concerning recombination, the mother is the incumbent individual, and the father is randomly selected from its neighborhood. The crossover operator is the blending method, which generates the offspring by applying Eq. 1, where β is a random number in $[0, 1]$ and x_n (y_n) is the allele of the n^{th} gene in the chromosome of the mother (father) [3]. If the resulting offspring is better than the incumbent individual, it replaces it. All functions from Tables 1 and 2 have a fitness value of zero for every $d \in \mathbb{Z}^+$. All functions were solved with $d = 1$ and $d = 2$. For practical purposes, we established as “optimal” any solution of size less than 10^{-4} (10^{-2}) for $d = 1$ ($d = 2$).

$$z_n = \beta x_n + (1 - \beta)y_n, \quad (1)$$

Figures 2 and 3 show the results obtained by executing a cGA, using k FLS as update policy, over the set of uni-modal and multi-modal functions from Tables

Table 2. Benchmark multi-modal functions.

Name	Function	Domain
Ackley	$-20 \exp \left(-0.2 \sqrt{\frac{1}{d} \sum_{i=1}^d x_i^2} \right) \exp \left(\frac{1}{d} \sum_{i=1}^d \cos(2\pi x_i) \right) + 20 + \exp(1)$	$[-5.12, 5.12]$
Rastrigin	$10d + \sum_{i=1}^d [x_i^2 - 10 \cos(2\pi x_i)]$	$[-5.12, 5.12]$
Schwefel	$418.9829d - \sum_{i=1}^d x_i \sin(\sqrt{ x_i })$	$[-500, 500]$
Griewank	$\sum_{i=1}^d \frac{x_i^2}{4000} - \prod_{i=1}^d \cos \left(\frac{x_i}{\sqrt{i}} \right) + 1$	$[-600, 600]$

Table 3. cGA's configuration.

Ring length	100
Boundary conditions	Closed
Neighborhood	East-West
Crossover operator	Blending method
Mutation operator	Uniform random
Mutation probability	0.1
Replacement criterion	Improvement

1 and 2. The reported values are the average number of crossovers performed by the cGA. For each value of k , the cGA was executed 300 times; the stop condition being to find an optimal solution. The black points represent the average of 30 different executions. So, for each value of k , 10 black points are reported. The red point represents the average of these 10 black points. From Figure 2, observe that all uni-modal functions with $d = 1$ tend to be solved faster for values of k around 10 and 20. Intuitively, this makes sense because uni-modal functions are relatively easy to solve. Thus, having more exploitation than exploration elements should improve the convergence time to optimal solutions. So, k FLS with small values of k has more exploitation elements because it takes advantage of the previously explored cells more quickly. Now, for $d = 2$, the value of k that improves the convergence time to optimal solutions is slightly greater. This makes sense too, because for $d = 2$ the functions are more difficult to solve. Thus, they should require less exploitation. Nevertheless, notice that the results from Figure 2(h) suggest that there might be functions such that there is no difference regarding the value of k .

Regarding multi-modal functions in Figure 3, the observations are similar to those obtained for the uni-modal functions. Namely, for $d = 1$, a greater value of k , around 20 and 40, helps improve the cGA's running time. Now, for $d = 2$, it is not easy to infer any improvement using intermediate values of k . From

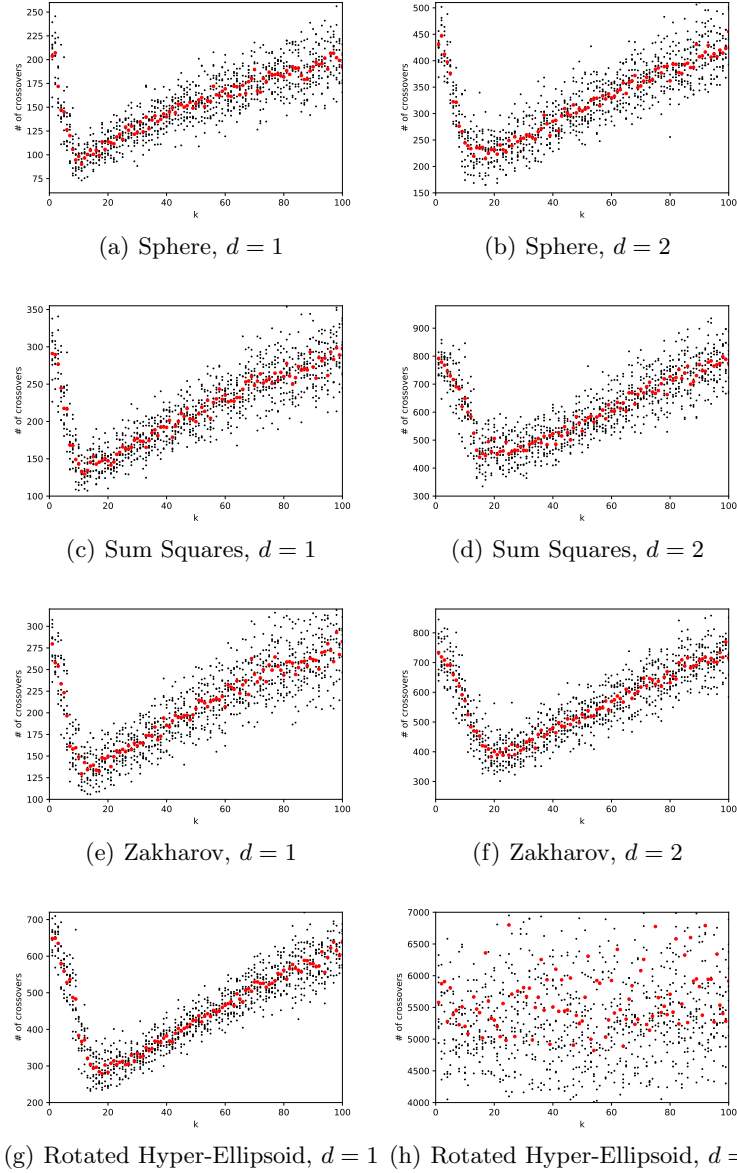


Fig. 2. Average running time of the cGA over some uni-modal functions.

these last figures, we cannot make any strong conjecture. Thus, we left a more rigorous analysis as future work.

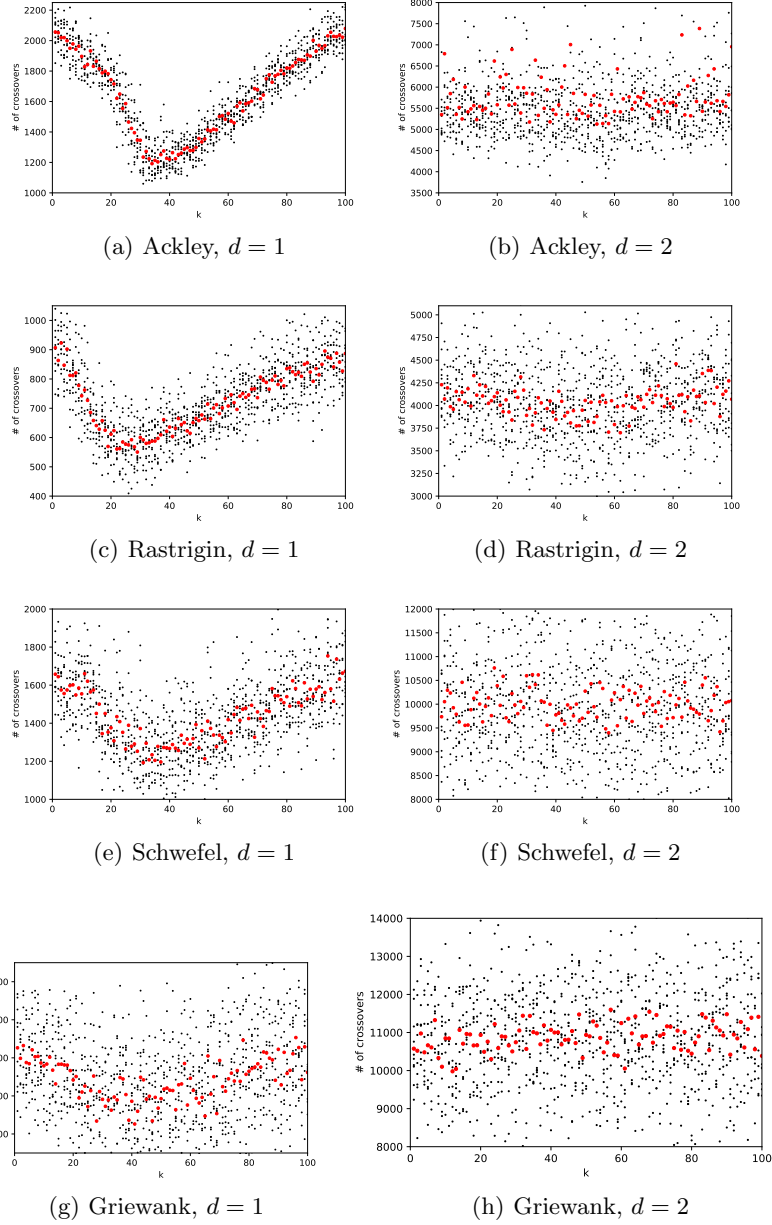


Fig. 3. Average running time of the cGA over some multi-modal functions.

4 Conclusions and Future Work

In this work, we attempted to explore the tradeoff between asynchronous and synchronous update policies in cGAs. To do so, we introduce an update policy,

named k -Fixed Line Sweep (k FLS). This policy was applied over a cGA for solving different single-objective benchmark continuous optimization functions. The results show that, for some cases, k FLS can accelerate convergence to near-optimal solutions if the value of k is appropriately adjusted. As a first approach, the performed empirical analysis may not be rigorous enough. Thus, in the future, we would like to perform a more robust and rigorous statistical analysis. Besides, we would like to explore the effect of different neighborhoods, topologies, and strategies, such as Cellular Memetic Algorithms (cMA). Finally, we are especially interested in performing an analysis of k FLS for $\mathcal{NP}$ -hard combinatorial optimization problems.

References

1. Alba, E., Dorronsoro, B.: Cellular Genetic Algorithms, vol. 42. Springer Science & Business Media (2009)
2. Alba, E., Giacobini, M., Tomassini, M., Romero, S.: Comparing synchronous and asynchronous cellular genetic algorithms. In: International Conference on Parallel Problem Solving from Nature. pp. 601–610. Springer (2002)
3. Haupt, R.L., Haupt, S.E.: Practical Genetic Algorithms. John Wiley & Sons, 2 edn. (2004)
4. Luque, G., Alba, E., Dorronsoro, B.: An asynchronous parallel implementation of a cellular genetic algorithm for combinatorial optimization. In: Proceedings of the 11th Annual Conference on Genetic and Evolutionary Computation. p. 1395–1402. GECCO '09, Association for Computing Machinery, New York, NY, USA (2009), <https://doi.org/10.1145/1569901.1570088>
5. Nebro, A.J., Durillo, J.J., Luna, F., Dorronsoro, B., Alba, E.: Mocell: A cellular genetic algorithm for multiobjective optimization. International Journal of Intelligent Systems 24(7), 726–746 (2009), <https://onlinelibrary.wiley.com/doi/abs/10.1002/int.20358>
6. Zhang, B., Xu, L., Zhang, J.: A multi-objective cellular genetic algorithm for energy-oriented balancing and sequencing problem of mixed-model assembly line. Journal of Cleaner Production 244, 118845 (2020), <https://www.sciencedirect.com/science/article/pii/S0959652619337151>

Aplicación Web de secuenciador musical basado en datos de los movimientos oculares (*eye-tracking*)

Jonathan Manuel Fraire-López¹, Francisco López-Orozco^{1,2},
Jesús Israel Hernández-Hernández^{1,2}, Juan Gabriel Ruiz-Ruiz³

¹ Universidad Autónoma de Ciudad Juárez,
División Multidisciplinaria en Ciudad Universitaria,
México

² Universidad Autónoma de Ciudad Juárez,
Maestría en Cómputo Aplicado,
Instituto de Ingeniería y Tecnología,
México

³ Universidad de la Sierra Juárez,
México

{al159943}@alumnos.uacj.mx
{francisco.orozco,israel.hernandez}@uacj.mx
{jugaruiz}@unsij.edu.mx

Resumen Este artículo presenta el desarrollo de una aplicación web de un secuenciador de instrumentos musicales que se utiliza con datos del movimiento de los ojos. Una ventaja de la aplicación es la habilidad para reproducir y grabar los instrumentos usando una webcam convencional. El proyecto está basado en *WebGazer* una biblioteca de software libre escrita en javascript. A partir de ésta se obtienen los datos de entrada de la vista por parte de los usuarios y se estima la posición de la mirada sobre la pantalla de la computadora. Las fases de prueba y validación fueron llevadas a cabo con 30 músicos experimentados. Los resultados muestran que la propuesta es correcta y aceptada por estos usuarios. Sin embargo, nuestro desarrollo puede ser mejorado en varios aspectos como la inclusión de más instrumentos musicales entre otros.

Palabras clave: Interacción basada en la vista, eye tracking, secuenciación de instrumentos musicales, interacción por cámara web.

Web Application of a Music Sequencer based on Eye Movement Data (Eye-Tracking)

Abstract. This paper presents the development of a web-based application for a musical instrument sequencer controlled through eye movement data. A key advantage of the application is its ability to play and record instruments using a standard webcam. The project is based

on WebGazer, an open-source JavaScript library. Using this library, gaze data from users is captured, and the system estimates the position of the user's gaze on the computer screen. Testing and validation phases were conducted with 30 experienced musicians. The results show that the proposed system is functional and well received by users. However, the application can be improved in several aspects, such as the inclusion of more musical instruments, among others.

Keywords: Gaze-based interaction, eye tracking, musical instrument sequencing, webcam interaction.

1. Introducción

La música ha evolucionado a través del tiempo así como los instrumentos que se utilizan para su ejecución. Algunas herramientas de software usadas en esta área han también evolucionado y se han convertido de mucha importancia para los músicos hoy en día. Todo esto ha sido ocasionado entre otras razones por el incremento de la digitalización musical en nuestros días. De la misma forma, algunas tecnologías también han cobrado importancia como las del seguimiento ocular o *eye-tracking* [11]. El *eye-tracking* ha sido previamente usada en el ámbito musical aunque no ha alcanzado un suficiente grado de popularidad entre los músicos todavía. En este proyecto se propone una solución a uno de los problemas o dificultades actuales del uso de la tecnología *eye-tracking* en la música. El *eye-tracking* ha fallado en su intento por captar la atención de usuarios dentro del campo musical. Esta dificultad refiere al hecho de que las aplicaciones existentes para reproducir o registrar instrumentos, requieren el uso de un eye-tracker especializado, al que usualmente no tienen acceso la mayoría de los usuarios [12]. Por esta razón, este desarrollo es una propuesta de software para la reproducción de música a partir de una cámara web convencional sin la necesidad de un eye-tracker.

2. Trabajos previos

La interpretación de la música es normalmente asociada a las extremidades del cuerpo o la voz, pero hoy en día hay diferentes alternativas a estas extremidades. Por ejemplo, hay instrumentos musicales virtuales que pueden tocarse sin la necesidad de las extremidades, usando sensores o la vista [1]. De esta forma se ofrecen nuevas alternativas a los músicos, teniendo la oportunidad de contar con uno o más elementos para desarrollar cierta independencia musical. Una ventaja podría ser usar la vista para tocar un instrumento y una extremidad para tocar otro al mismo tiempo. Adicionalmente, esto abre la posibilidad para que personas con capacidades diferentes y a las que se les imposibilita tocar un instrumento musical físico, puedan aventurarse en el campo de la música. La necesidad de adaptar y simplificar la forma en la que los instrumentos musicales virtuales son tocados o ejecutados ha permitido la experimentación

con secuenciadores virtuales, para grabar diferentes pistas que se reproducen repetidamente y así con más posibilidades de reproducir otros instrumentos sobre las pistas cuando se están ejecutando [2].

Diferentes trabajos y proyectos han sido desarrollados en este campo haciendo énfasis en instrumentos musicales y secuenciadores en la Web. Un secuenciador con características interesantes es *Soundation* [3], una herramienta de pago dedicada a la producción musical y edición de pistas. Los instrumentos son ejecutados solo con controladores físicos como con *Musical Instrument Digital Interfaces (midi)*. Un secuenciador muy práctico es Online Sequencer [4], cuya interfaz basada en una cuadrícula que segmenta los ritmos en el cual las notas se dibujan en tiempos específicos para todos los compases. Otro ejemplo es *midi.city* [8], una herramienta gratuita que permite secuenciar un número de instrumentos musicales. También encontramos *EyeHarp* [5], una aplicación de escritorio no gratuita que permite la ejecución de instrumentos con la vista. Requiere de un *eye tracker* para ejecutar los instrumentos y navegar dentro de la aplicación con la vista.

Por otro lado, se pueden encontrar dos aplicaciones de Microsoft para la secuenciación musical: *Microsoft Hands-Free Sound Machine* [6] y *Microsoft Hands-Free Sound Jam* [10]. Ambas son aplicaciones gratuitas que requieren un *eye tracker* para funcionar. El primero es un secuenciador que combina sonidos multimedia previamente creados con instrumentos musicales de la aplicación y el segundo está más orientado a la composición musical.

En el proyecto que se describe en el presente artículo, el objetivo fue desarrollar una alternativa de uso del *eye tracking*. La propuesta es una aplicación web para la ejecución y secuenciación de instrumentos musicales mediante la vista usando una cámara web convencional. Se plantearon dos preguntas importantes: ¿cuál sería el rendimiento de la aplicación web y cuáles serían las ventajas de usar este secuenciador comparado con algún controlador para un instrumento musical?

3. Metodología e implementación

La metodología de cascada fue usada para el desarrollo de este proyecto con las siguientes etapas: análisis de requisitos, diseño, codificación, pruebas y mantenimiento. Por las características de este modelo secuencial, cada etapa comienza y termina antes de empezar la siguiente [7].

En la primera etapa, se llevó a cabo la elicitación de requerimientos de la aplicación, definiendo los funcionales como la *Reproducción y Grabación de instrumentos* y no funcionales como la *Calibración de la biblioteca* o la *Webcam necesaria*, además de la elaboración de diagramas de casos de uso. Los casos de uso describen la manera en la que el usuario puede interactuar con el sistema y de las funciones disponibles para él.

Durante la etapa de diseño se recolectaron las muestras o *samples*, se diseñaron los instrumentos, el menú y el diseño general de la aplicación.

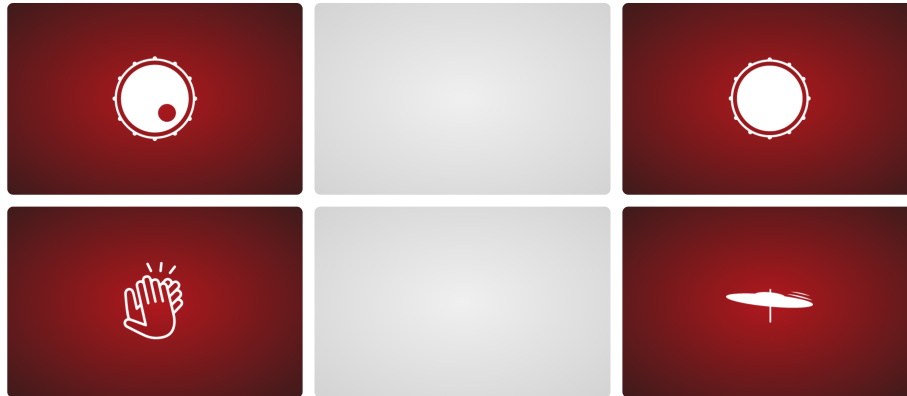


Fig. 1. Diseño de instrumentos.



Fig. 2. Controles del secuenciador.

Cada nota o elemento de un instrumento musical corresponde a un *sample* en un archivo de audio MP3. Para el caso de los instrumentos piano, guitarra y bajo se utilizó la aplicación Logic Pro X, la cual cuenta con una gran variedad de instrumentos virtuales que pueden ser ejecutados vía MIDI. En el caso de la batería se usaron *samples* gratis obtenidos de la página de SamplesWrap [9]. Una vez recolectadas las muestras fue necesario implementar una interfaz de cada instrumento en forma de *pad* rectangular. En la Figura 1 se muestran los cuatro *pads* a los costados y en el centro dos *pads* de descanso donde los usuarios pueden mirar sin que la aplicación lance sonido alguno. Esto último fue basado en las sugerencias aportadas en [14]. Posteriormente se diseñó un menú estático de forma horizontal cuyo apariencia y botones pueden apreciarse en la Figura 2.

Finalmente en la Figura 3 se presenta la interfaz principal del secuenciador dividido en 3 partes. En la sección 1 aparece el visualizador de la cámara web del usuario en tiempo real de la biblioteca *WebGazer*. En 2, corresponde al menú de opciones y en la 3 aparece el instrumento seleccionado. La codificación del secuenciador fue dividido en dos partes principales: el secuenciador en sí y por otro lado la parte de la biblioteca *WebGazer* que se utilizó para la ejecución con la vista. Para el desarrollo del secuenciador se usaron diferentes bibliotecas. En la Figura 4 se muestra el orden del flujo de datos y conexión de las bibliotecas utilizadas para la función del secuenciador. La primera biblioteca para la programación del secuenciador fue *Howler JS*, usada para cargar los *samples*. Al seleccionar cada nota o elemento del instrumento se lanza el sonido de éste. Para evitar retrasos y lograr que la interpretación sea en tiempo real

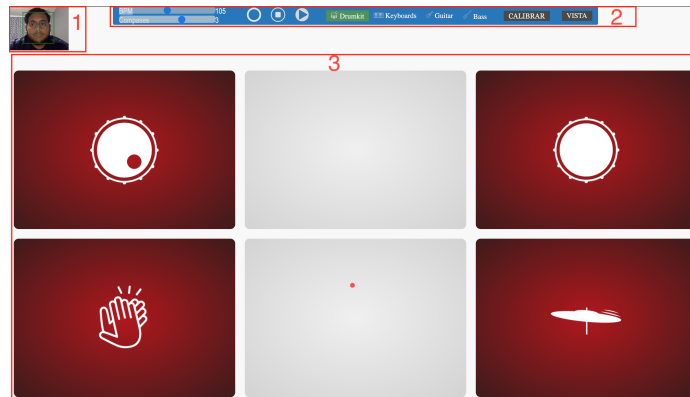


Fig. 3. Interfaz general de la aplicación.

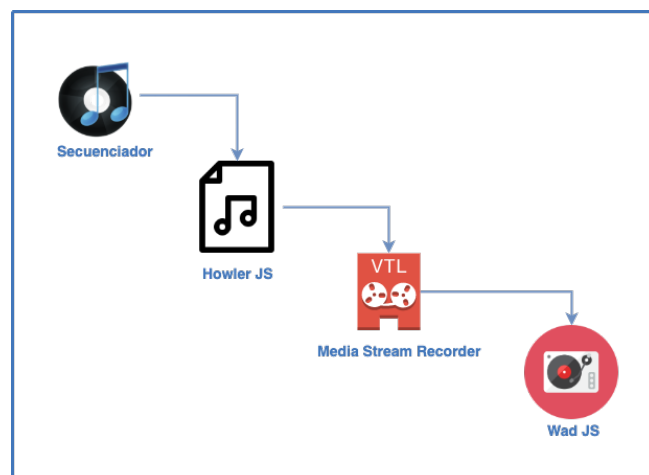


Fig. 4. Bibliotecas utilizadas.

se usó *Howler JS*, la cual provee la función de crear *sprites*. Estos consisten en juntar y unir cada archivo de audio MP3 o un archivo WebM en uno solo y así crear un archivo tipo JSON, donde reconoce el segmento del archivo correspondiente y lo separa en segundos, para así llamar siempre al mismo archivo desde diferentes partes.

Una vez definidos los *sprites* de cada instrumento del secuenciador, se pasa la información de la biblioteca *Howler JS* a la biblioteca *Media Stream Recorder*. Esta es la encargada de grabar el audio. Básicamente esta biblioteca y *Howler JS* funcionan mediante entrada y salida. Se tomó un objeto de salida de la biblioteca de *Howler JS* y se pasó a la entrada de *Media Stream Recorder*. Una vez con los

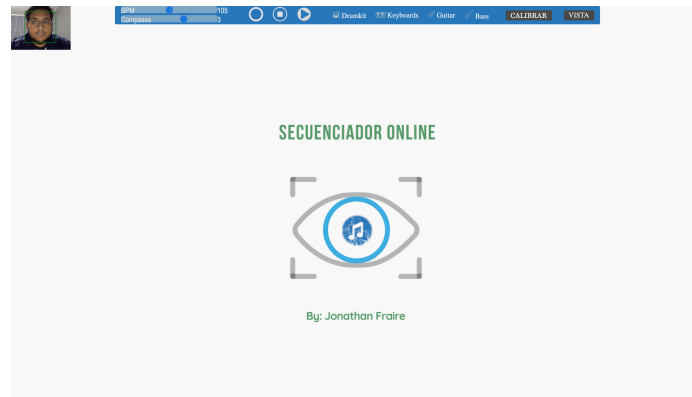


Fig. 5. Pantalla de inicio de la aplicación.

datos dentro de esta biblioteca, estos están listos para ser procesados y enviados a la salida. Asimismo, la salida de *Media Stream Recorder* se conectó a la entrada de la biblioteca *Wad JS*. Esto fue posible porque se creó un objeto tipo *Wad*, el cual recibe como fuente de entrada el objeto tipo *blob* que genera la salida de *Media Stream Recorder*.

Para iniciar la grabación se llama a la función *.start()* de la biblioteca *Media Stream Recorder*. Después de la grabación que el usuario hace, se lanza un evento de la biblioteca *Media Stream Recorder*, y justo cuando la grabación se termina se llama a este evento llamado *.onStop()*, en el cual se genera el archivo tipo *blob* que se menciona anteriormente. Dentro de este mismo evento cada grabación se va guardando en un arreglo donde cada posición contiene un archivo tipo *Wad* que representa la grabación hecha.

Después de la programación de las funciones del secuenciador, se realizó la conexión con la biblioteca *WebGazer*. Esta biblioteca es la encargada de estimar la posición de dónde el usuario está mirando en la pantalla y permite conectarla con el secuenciador, de tal forma que los sonidos puedan ejecutarse al momento que el usuario fije la vista en un elemento o nota de un instrumento. Cabe hacer mención que la etapa de pruebas y evaluación se apoyó de la aplicación de un *After-Scenario Questionnaire (ASQ)*, que se detalla más adelante en la sección 4.

3.1. Secuencia de uso de la aplicación

Al acceder a la aplicación, lo primero que se encuentra es la pantalla principal de inicio como se aprecia en la Figura 5. La pantalla principal cuenta con el menú de opciones, dentro del cual se encuentra el botón **CALIBRAR** que se encarga de dirigir a la pantalla del sistema de calibración para la biblioteca *WebGazer*. Una vez que se accede a la pantalla de calibración, es necesario completarla para poder regresar a la pantalla principal. El proceso de calibración consta de seguir con la mirada el cursor y dar 5 clics en cada uno de los puntos grises y así estos

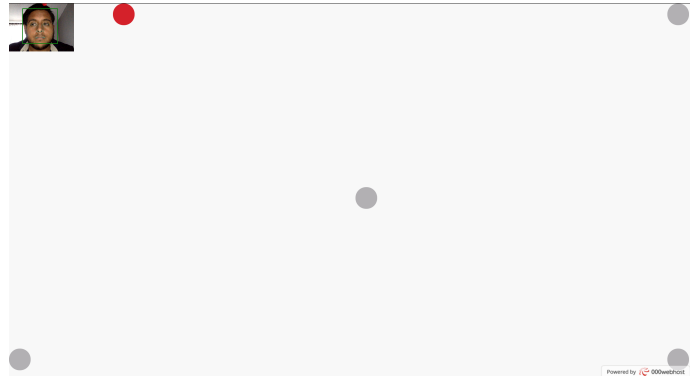


Fig. 6. Pantalla de calibración.

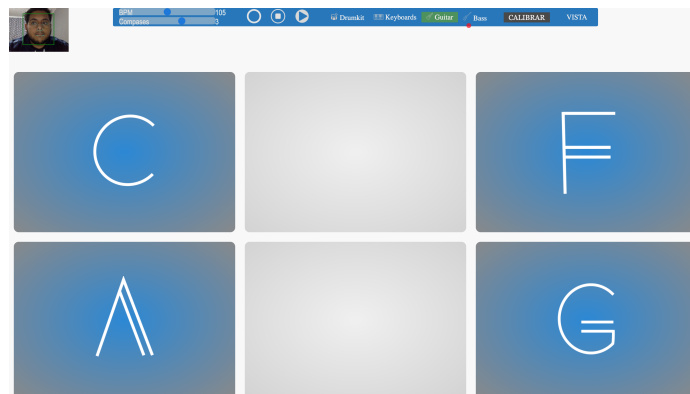


Fig. 7. Pantalla de instrumento seleccionado.

cambiarán a color rojo como se muestra en la Figura 6. Terminada la calibración, la aplicación regresará a la pantalla de inicio como se puede ver en la Figura 7, donde se puede seleccionar cualquier instrumento para ser ejecutado o grabado.

4. Resultados y discusión

En esta sección se presentan los resultados de las pruebas realizadas con 30 músicos que pudieron manifestar su aceptación o rechazo de la aplicación web. De estos, 28 fueron hombres y 2 mujeres entre 17 a 38 años. Todos, experimentados en al menos uno de los cuatro instrumentos que tiene la aplicación: bajo, batería, teclado y guitarra. A cada uno de los músicos se les pidió realizar dos tareas y la evaluación se llevó a cabo tanto de forma cualitativa como cuantitativa. En el

Tabla 1. Promedio de respuestas en evaluación cualitativa para Tarea 1 y Tarea 2.

Pregunta	Tarea 1 ($\bar{x}$ σ)	Tarea 2 ($\bar{x}$ σ)
1	5.8 1.09	5.36 1.27
2	5.73 1.14	5.63 1.06
3	5.76 1.30	5.66 1.09

primer caso, mediante un cuestionario tipo *After-Scenario Questionnaire (ASQ)* que se aplicó después de cada tarea completada, para que cada usuario reflejara su grado de satisfacción con la aplicación web. La evaluación cuantitativa fue basada en dos métricas: porcentaje de tareas completadas y tiempo que le llevó realizar cada tarea. La aplicación desarrollada es accesible a partir de un *host* gratuito del proveedor *000WebHost*[13] desde la siguiente URL 1. El código fuente del proyecto es accesible en URL 2.

4.1. Tarea 1

En esta tarea se les pidió a los usuarios que grabaran cualquier progresión de acordes a su gusto, mediante clics del ratón. La primera pregunta para los usuarios después de la tarea fue: **En general, ¿Qué tan satisfecho estoy con la facilidad de completar la tarea en este escenario?** La escala de respuestas para esta pregunta fue de 1 (totalmente en desacuerdo) hasta 7 (totalmente de acuerdo). El promedio de las respuestas fue 5.8 ($\sigma = 1.09$), que se inclina hacia el contexto totalmente de acuerdo.

Las respuestas más frecuentes fueron 7, 6 y 5, con 10, 8 y 9 usuarios respectivamente, lo que significa que la mayoría de usuarios están conformes con la facilidad de usar la aplicación al realizar la tarea. La segunda pregunta fue: **En general, ¿Qué tan satisfecho estoy con la cantidad de tiempo que tomó completar la tarea en este escenario?** En este caso, nuevamente los resultados reflejan que la mayoría de usuarios están satisfechos con el tiempo que les tomó realizar la tarea, dado que las respuestas fueron 7, 6 y 5, con 10, 7 y 9 usuarios respectivamente. El promedio de respuestas de los usuarios fue 5.73 ($\sigma = 1.14$), con una tendencia hacia el lado totalmente de acuerdo. Para el caso de la pregunta tres, el promedio fue 5.76 ($\sigma = 1.30$), la cual cuestiona: **¿Qué tan satisfecho estoy con la información de soporte (ayuda en línea, mensajes) al completar la tarea?** La respuesta más frecuente fue el número 7 con 11 votos, seguida del número 6 con 8 y el número 5 con 7 votos de usuarios.

4.2. Tarea 2

En la segunda tarea se graban los instrumentos con la vista. Sin embargo, al final el mismo cuestionario fue aplicado a los usuarios. Acerca de la primera pregunta de la facilidad para completar la tarea, el promedio de respuestas fue de 5.36 ($\sigma = 1.27$), el cual a diferencia de la Tarea 1 (5.8, $\sigma = 1.09$), fue menor,

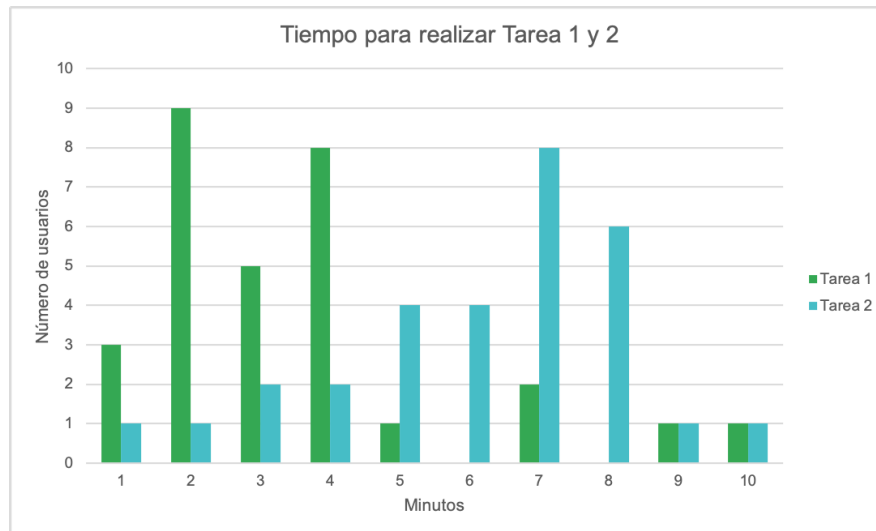


Fig. 8. Comparativa de tiempo para realizar ambas tareas.

pero se mantiene inclinado hacia el contexto totalmente de acuerdo, mostrando como resultado más frecuente el número 6. Para la segunda pregunta se tiene también como respuesta más frecuente el número 6 al igual que en la pregunta anterior, además de un promedio de 5.63 ($\sigma = 1.06$). Para la última pregunta, nuevamente el resultado más frecuente es 6, con promedio 5.66 ($\sigma = 1.09$), lo cual, al igual que en las otras 2 preguntas, es un resultado un poco más bajo a comparación con la Tarea 1, lo cual refleja menos satisfacción para realizar la tarea pero no llegando a ser un mal resultado sobrepasando la mitad de la escala, encontrándose inclinado hacia el contexto totalmente de acuerdo. La Tabla 1 resume estos resultados.

Acerca de la evaluación cuantitativa, como ya se había mencionado anteriormente, los resultados se distribuyeron a través de 2 variables: el porcentaje de tareas completadas por el usuario y el tiempo que tomó el usuario para realizar la tarea. Ambas tareas se completaron al 100 % por los 30 usuarios. Respecto a la segunda variable, en la Figura 8 se observa que la mayoría de usuarios tardaron alrededor de 2 y 4 minutos en realizar la Tarea 1, con un tiempo promedio de 3.53 min ($\sigma = 2.20$). La segunda tarea les llevo más tiempo a los usuarios en realizarla, en promedio 6.13 min ($\sigma = 2.09$). En la Figura 8 se observa que les tomó entre 7 y 8 minutos. Además, se puede observar una comparativa de ambas tareas, reflejando el número de usuarios y el tiempo que les tomó realizar cada tarea.

La mayoría de usuarios tardaron poco tiempo en realizar la Tarea 1 y que pocos tardaron mucho tiempo en completarla. En la Tarea 2 se presenta lo contrario, muchos usuarios tardaron mucho tiempo y pocos usuarios pudieron

realizarla en un tiempo corto. Este comportamiento se pudo deber al grado de familiarización de los músicos con aplicaciones relacionadas con el seguimiento ocular, la mayoría de usuarios presentaban un grado bajo o nulo de experiencia con la función de ejecutar los instrumentos con la vista la cual era el diferenciador de la Tarea 2 con la Tarea 1. Debido a esto, al no haber usado antes una aplicación como la que describe este proyecto, los usuarios tendían a tomar más tiempo para familiarizarse primero con la función además de corregir factores externos como puede ser la iluminación del entorno para que la biblioteca *WebGazer* pudiera captar sin problema el rostro, postura del cuerpo y rostro frente a la *webcam*, distracciones que daban como resultado el perder la calibración, entre otros. Una vez que ya se encontraban familiarizados procedían a realizar la grabación.

5. Conclusiones

En el presente artículo se presentó el desarrollo de un secuenciador musical web, compuesto de cuatro instrumentos musicales que pueden ser ejecutados y grabados mediante la vista a través de una *webcam* con la biblioteca *WebGazer* que se encarga de estimar el seguimiento ocular del usuario. En este desarrollo no se requiere de un dispositivo *eye-tracker* como en el caso de trabajos reportados en el área y citados en este artículo. El objetivo de este proyecto fue desarrollar una aplicación web para secuenciar instrumentos musicales mediante la vista a través de una *webcam* convencional.

Además se plantearon los siguientes objetivos específicos:

- diseñar un secuenciador web con los instrumentos musicales convencionales,
- determinar el tamaño y separación adecuada de los instrumentos musicales virtuales y
- establecer el protocolo de comunicación entre los instrumentos musicales y la *webcam* para el manejo del secuenciador.

Todos fueron cumplidos gracias a la biblioteca *WebGazer*. Después de la fase de evaluación del proyecto con los usuarios profesionales de la música, se concluye que el desempeño del secuenciador es bueno, dado que la mayoría de los usuarios manifestaron estar totalmente de acuerdo con la aplicación.

También se puede destacar que la desventaja principal es que los usuarios no están familiarizados con la función para ejecutar los instrumentos con la vista, ya que existen controladores para ejecución de instrumentos más comunes, además de otras desventajas, como lo son factores externos que se mencionan y se describen anteriormente en este artículo.

Sin embargo, a pesar de que el proyecto final es un producto totalmente funcional, aún existen mejoras que se pueden hacer para enriquecer futuros proyectos o productos similares. Algunas de estas mejoras podrían ser el desarrollo de funciones extras de edición del audio grabado por el usuario, la creación de instrucciones o tutoriales más ilustrativos, que puedan captar la atención del usuario y sean de ayuda para el correcto uso de la aplicación, así como la adición de funciones MIDI para uso de controladores externos.

Referencias

1. Z. Vamvakousis, R. Ramirez: The EyeHarp: An eye-tracking-based musical instrument. In: Proceedings of the 8th Sound and Music Computing Conference, SMC 2011 (2011)
2. J. Palazón-Herrera: Creación musical en Educación Secundaria a través del uso desecunciadores multipista (2016)
3. Soundation — Make music online. (2021) [Online]. Available: <https://soundation.com/>.
4. Make music online. (2021) [Online]. Available: <https://onlinesequencer.net/>.
5. Inicio - The EyeHarp. (2021) [Online]. Available: <http://theeyeharp.org/es/inicio/>.
6. Microsoft Hands-Free Sound Machine. (2021) [Online]. Available: <https://www.microsoft.com/en-us/research/project/microsoft-hands-free-sound-machine/>.
7. W. Van Casteren: The Waterfall Model and the Agile Methodologies: A comparison by project characteristics. (2017)
8. midi.city - free online synth instrument. (2021) [Online]. Available: <https://midi.city/>.
9. Free audio samples, drum loops & kits, vocals, royalty free music. (2021) [Online]. Available: <https://sampleswap.org/>.
10. Microsoft Hands-Free Sound Jam - Microsoft Research. (2021) [Online]. Available: <https://www.microsoft.com/en-us/research/project/microsoft-hands-free-sound-jam/>.
11. A. Poole, L. Ball: Eye Tracking in Human-Computer Interaction and Usability Research : Current Status and Future Prospects (2004)
12. L. González and J. D. Velásquez, “Una aplicación de herramientas de eye-tracking para analizar las preferencias de contenido de los usuarios de sitios web, ”Revista de ingeniería e sistemas, vol. 26, no. 1, pp. 95–118 (2012)
13. Mejor Hosting Gratis de Mexico 2020 — Hosting Gratuito. (2020) [Online]. Available: <https://mex.000webhost.com/>.
14. C. A. Gainer, R. W. Obermayer: Pilot Eye Fixations While Flying Selected Maneuvers Using Two Instrument Panels. Human Factors, vol. 6, no. 5, pp. 485–501, SAGE Publications Inc. (1964) Available:<https://doi.org/10.1177/001872086400600507>.

Uso de diccionarios en Python para conocer las tendencias en Twitter sobre el aumento de feminicidios en México

Marleni Reyes Monreal¹, Beatriz Bernábe Loranca²,
Estefania Sarmiento Barrios², Melissa Isaaly Mendoza Bernábe³

¹ Benemérita Universidad Autónoma de Puebla,
Escuela de Artes Plásticas y Audiovisuales,
México

² Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Computación,
México

³ Universidad Iberoamericana Puebla,
México

`marleni.reyes@correo.buap.mx`

Resumen. Debido a las noticias alarmantes en las redes y diversos medios de difusión, se han establecido distintos supuestos sobre un aumento de violencia en contra de la mujer, en particular en el asunto de feminicidios. Esta situación, ha dado lugar a la necesidad de identificar, las tendencias de opinión y forma en que se difunden y viralizan los contenidos relacionados con el tema de feminicidios en México. En este punto, la técnica de creación de diccionarios aplicados al estudio de contenidos en Twitter mostrará las relaciones entre palabras usadas para referirse al tema y el sentido general en el que se comprende el tema. Considerando que el desarrollo de una investigación centrada en un problema del tamaño de feminicidios implica un elevado costo económico, de tiempo, de recursos humanos, entre otros. Identificar y establecer justicia o solución al este problema conduce a un trabajo mayúsculo con múltiples variables, con áreas interminables de analizar y representa un elevado precio en distintas direcciones. Es en este punto donde se inserta el artículo que presenta una propuesta de creación de diccionarios respecto a comentarios en Twitter con el propósito de organizar las palabras representativas sobre feminicidios de acuerdo con las opiniones de los usuarios de Twitter, una de las redes sociales más utilizadas para compartir noticias. De esta manera, es posible proponer actividades de reflexión, difusión, prevención e incluso diagnóstico sobre la magnitud e incidencia del problema sumada a las investigaciones de otras áreas que tratan el tema y sin el elevado costo que tiene una investigación tradicional.

Palabras clave: Redes sociales, análisis de datos, diccionarios, feminicidios.

Using Dictionaries in Python to Understand Twitter Trends about the Rise in Femicides in Mexico

Abstract. Due to alarming news stories on social media and in various media outlets, various assumptions have arisen about an increase in violence against women, particularly in the area of femicide. This situation has led to the need to identify opinion trends and the way in which content related to femicide in Mexico is disseminated and goes viral. At this point, the dictionary creation technique applied to the study of content on Twitter will reveal the relationships between words used to refer to the topic and the general understanding of the topic. Considering that conducting research focused on a problem the size of femicide entails a high cost in terms of finances, time, human resources, and other aspects. Identifying and establishing justice or a solution to this problem requires a massive undertaking with multiple variables, endless areas to analyze, and represents a high price in different directions. This is where the article presents a proposal for the creation of dictionaries of Twitter comments. This would organize representative words about femicides based on the opinions of Twitter users, one of the most widely used social networks for sharing news. This would allow for activities for reflection, dissemination, prevention, and even diagnosis of the magnitude and impact of the problem, combined with research from other areas that address the topic and without the high cost of traditional research.

Keywords: Social networks, data analysis, dictionaries, femicides.

1. Introducción

El concepto de feminicidio fue propuesto por la norteamericana Diana E.H Russell, quien lo considera como “el asesinato de mujeres a manos de hombres debido a que son mujeres” [1]. En México, los casos de asesinatos de mujeres en Ciudad Juárez Chihuahua se comenzaron a denunciar en 1993 dando lugar a que el tema se posiciona como un problema de relevancia nacional e internacional. A partir de ese momento, surgieron distintas protestas para exigir justicia y seguridad para la ciudadanía mexicana, particularmente para las víctimas; sin embargo, tuvieron que pasar casi 20 años para que el país incluyera en su código penal el delito de feminicidio [2].

El aumento considerable de la violencia en el país que se traduce en crímenes, intimidaciones, abusos, injusticias y arbitrariedades contra las mujeres, no ha sido suficiente para que las autoridades establezcan un orden y medidas adecuadas para detener la violencia de género. Una gran mayoría de los casos de feminicidio continúan sin ser formalmente registrados, investigados, juzgados o sancionados. Se asume que la ausencia de grupos especializados que trabajen de manera eficiente y transparente limita las investigaciones necesarias. Por otra parte, la falta de datos estadísticos y

sistemas de información verídicos o confiables, que contribuyan a los procesos de las investigaciones, impiden la valoración de la gravedad del problema.

De acuerdo con lo anterior, se propone un estudio adicional sobre problema que analice lo referente a su difusión y las opiniones que genera. Por tanto, el objetivo de este trabajo es realizar un estudio social y computacional de los comentarios en Twitter que permita generar datos e información para conocer el impacto en la visión del público del problema de feminicidios en México en los últimos años. Consecuentemente, al caracterizar el problema dentro de una red social, es posible estructurar estadísticamente los resultados obtenidos y proponer un programa de difusión que incida en los valores de la gente que transita en las redes sociales.

Adicionalmente, se busca identificar la presencia del tema de feminicidios respecto a opiniones, relatos, lugares y maneras que se asocian a este problema, con el fin de estructurar la información en diccionarios que contienen las palabras más representativas relacionadas con feminicidio.

El trabajo se encuentra estructurado de la siguiente manera, A) presentar algunos trabajos relacionados, B) exponer el primer acercamiento a la solución del problema usando PLN, C) crear los diccionarios, D) presentar conclusiones y trabajo futuro.

2. Preliminares y trabajos relacionados

En este apartado se incluyen algunos artículos que abordan el tema del análisis sobre feminicidios usando Procesamiento de Lenguaje Natural (PLN), además de trabajos que permiten conocer los conceptos básicos como feminicidio, violencia, violencia de género y dignidad. El orden de aparición no dictamina la importancia.

2.1 Trabajos relacionados

Redes en Twitter y la defensa de la mujer peruana (ante el acoso y el feminicidio)

En este trabajo del 2020 se analiza el tema de feminicidios en Perú y el uso de Twitter como forma de protección para la mujer. Para hacerlo, se recolectan Tweets, usando etiquetas/hashtags y palabras clave, se aplican técnicas de PLN para automatizar el análisis de los mensajes y se usan métricas de centralidad durante el análisis de redes para la identificación de los perfiles más relevantes. Finalmente, el objetivo principal del texto ha sido diseñar, modelar y analizar los datos de Twitter. El resultado del este trabajo consistió en identificar las personas influenciadoras y las tendencias que las personas en la red discuten sobre la defensa de los derechos de las mujeres y que deberían tener principal atención por parte de las autoridades [3].

Tecnologías de la lengua para análisis de opiniones en redes sociales: Procesamiento del Lenguaje Natural

Este trabajo trata sobre el análisis y monitorización del flujo de datos, utilizando métodos inteligentes de acceso a la información, así como el análisis de sentimientos o minería de opiniones. Mediante el desarrollo de este proyecto se trata de afrontar los desafíos como el ruido de los textos analizados con el fin de desarrollar un sistema

efectivo de minería de opiniones sobre micro textos escritos en español y gallego. El resultado del estudio es el desarrollo de relevantes en analizadores de dependencias basados en grafos, diseño e implementación de sistemas de minería de opiniones multilingües no supervisados [4].

Análisis de Flujos de Información de la Red Social Twitter Caso de Estudio: Impacto de las Tendencias en Redes Sociales para las Operaciones del Mercado de Acciones de la Empresa ECOPETROL

El desarrollo de la investigación sobre flujos de información tiene el objetivo de analizar datos provenientes de Twitter utilizando una API de captura de flujos de información y procesando la misma con el fin de realizar categorizaciones cuantificables de la información para mejorar la toma de decisiones para la operación de valores en el mercado del grupo de acciones de la empresa Ecopetrol S.A [5]. Este estudio no aborda al feminicidio, pero aporta a la investigación de tendencias en Twitter y su uso en la toma de decisiones; sus métodos también pueden ser aplicados a la toma de decisiones con respecto al tema de feminicidios.

2.2 Feminicidio: Terminología

a) Motivación: Sobre la violencia, dignidad y derecho de la mujer

El significado de violencia siempre llevará a una connotación negativa y se refiere a la manifestación de actos de fuerza física desmedida, siendo este concepto el más difundido en los organismos jurídicos. Cabe aclarar que este no es el único tipo de violencia documentado.

El termino de violencia de genero ha tomado presencia en México debido aumento de casos y denuncias en los últimos años. En 1993, la declaración sobre la eliminación de la violencia contra la mujer de la Organización de las Naciones Unidad ONU defina a la violencia contra la mujer como “Todo acto de violencia basado en el género que tiene como resultado posible o real un daño físico, sexual o psicológico, incluidas las amenazas, la coerción o la privación arbitraria de la libertad, ya sea que ocurra en la vida pública o en la vida privada” [6].

En el convenio del Consejo de Europa sobre prevención y lucha contra la agresión generalizada a las mujeres, el término violencia de género se refiere a la violencia que se ejerce directamente contra la mujer porque el hecho de ser mujer [7].

El concepto de violencia de género y la lucha contra su erradicación ha cobrado importancia recientemente, por múltiples factores como el aumento de su difusión en redes. Aunque se estima que es alrededor de los 90s cuando las iniciativas para la eliminación de la violencia contra la mujer se consolidan.

Al estudiar los documentos sobre el tema destaca la palabra dignidad, la cual se encuentra relacionada directamente al tema, valor necesario en los derechos a la mujer [9]. Por el lado de la ética, se dice que la dignidad se sustenta en la “moral” que da forma a la conciencia de la responsabilidad que se orienta a una regencia sobre las acciones propias, por tanto, es el resultado de la manera de comportarse “íntegramente”, por encima de todo precio [10]. En este aspecto, es posible afirmar que la dignidad, no es aislada del respeto, consideración y valoración a la mujer, por lo cual, decir que la

mujer se dignifica, no es un asunto de mera capacidad reproductiva y por tanto es necesario aplicar esfuerzos para contrarrestar actitudes discriminatorias [11].

b) Sobre herramientas y técnicas de Procesamiento en Lenguaje Natural.

El Lenguaje Natural (LN) “es el medio que se utiliza de manera cotidiana para establecer comunicación entre personas” [12]. En otras fuentes, se afirma que el LN ha evolucionado para establecer la comunicación entre seres humanos [13]. Dos definiciones sencillas que sugieren que, a pesar de que el medio de comunicación puede resultar simple, este se ha vuelto un proceso que involucra distintas herramientas que van desde las intuitivas hasta las sistemáticas. El LN ha ido creciendo debido a las exigencias de las plataformas para interpretar y analizar comunicación en distintos medios.

Procesamiento del lenguaje natural

El Procesamiento de Lenguaje Natural (PLN) es la rama de las ciencias computacionales encargada del diseño e implementación de los elementos de software y hardware necesarios para el tratamiento computacional del LN. Su objetivo es la comprensión del LN por parte de la computadora [14]. Este campo ha evolucionado los últimos años debido al gran volumen de datos disponible y a la capacidad de utilizar algoritmos nuevos para su procesamiento. Por otra parte, el PLN se incluye dentro de ciencias de la computación e Inteligencia Artificial (IA), del mismo modo se incorpora en Ingeniería del conocimiento, que se entiende como el “conjunto de conocimientos y técnicas que permiten aplicar el saber científico a la utilización del conocimiento (entendimiento, inteligencia o razón natural)” [15].

Twitter

Las redes sociales permiten la relación entre usuarios, quienes comparten, difunden y transforman información, cabe resaltar que son servicios basados en web [16]. Debido a esta nueva tecnología se ha generado una gran cantidad de información que es compartida por miles de usuarios, generando a su vez más información. La gran cantidad de información y usuarios en esta red impacta en la construcción de códigos culturales, lo que hace necesario entender que el procesamiento de este volumen de información ayuda en el análisis de diversas situaciones en distintos contextos que son compartidos por la red. La red social de interés en este documento es Twitter, entendido como un servicio que permite compartir pensamientos, enlaces, información, etc. en formato de microblog, usando tweets que son pequeñas publicaciones basadas en texto de no más de 280 caracteres. Sobre la interacción, Twitter permite la interacción con cualquier usuario web, que pertenezca a la red social, así como comunicarse de forma privada o pública [3]. Twitter es una red en tiempo real que posibilita la interacción con información que está sucediendo en el momento, instantáneamente.

Twitter posee una serie de características que permiten su aplicación dentro del proceso de análisis y extracción de información con el fin de obtener conocimiento. Los mensajes cortos con base en texto de Twitter la hacen excelente red social para ser estudiada PLN.

Python

Para este estudio se empleó el lenguaje de programación interpretado y multipropósito Python. Este lenguaje de *scripting*, dinámico, funcional y de alta precisión numérica. Debido a su popularidad, se puede aplicar y encontrar en diversas áreas que requieren desarrollo, “se trata de un lenguaje potente, flexible, con una sintaxis muy limpia que favorece un código legible” [17].

Feminicidio en PLN

Los feminicidios son asesinatos de mujeres por razones de género, debido a las consecuencias de estructuras de poder en un sistema patriarcal [3]. La palabra feminicidio se suele adaptar a contextos culturales y sociales, por lo que si bien, existe una idea general del concepto, también existen muchas formas de expresarla, entenderla y legislarla. Este problema ha sido analizado con diversas herramientas computacionales incluyendo PLN.

3. Primera etapa: Descubrimiento de información y primer acercamiento a la solución del problema

Twitter permite exponer noticias y opiniones de un tema en tiempo real que pueden ser medidas en un tiempo determinado. Los datos de Twitter son únicos y se pueden extraer a partir de la información compartida en otras plataformas sociales

3.1 Arquitectura del problema de feminicidios en PLN

La arquitectura del problema que tratamos en este artículo induce a señalar una serie de pasos a seguir para tener una implementación deseable. En términos amplios, se exhiben las indicaciones en secuencia iniciando con el descubrimiento de la información y terminando con la conclusión de los resultados (ver Fig. 1).

3.2 Selección de palabras clave y hashtags

El paso previo para la extracción de datos es plantear las palabras clave que describen al problema para facilitar las descargas de los Tweets y mantener una relación constante del tema que se aborda. La manera más útil consiste en buscar las cuentas de aquellas personas influenciadoras relacionadas con el tema de feminicidios. Una vez encontradas las cuentas asociadas, se identifican las palabras más representativas del problema y se realizan búsquedas con esas palabras a través de hashtags. Con este procedimiento, las palabras con más énfasis en cuanto a la frecuencia son las siguientes: #NiU-naMas, #NiUnaMenos, #Feminicidio, #VivasNosQueremos, Acoso sexual, Movimiento feminista, Asesinato machista, Feminicida.

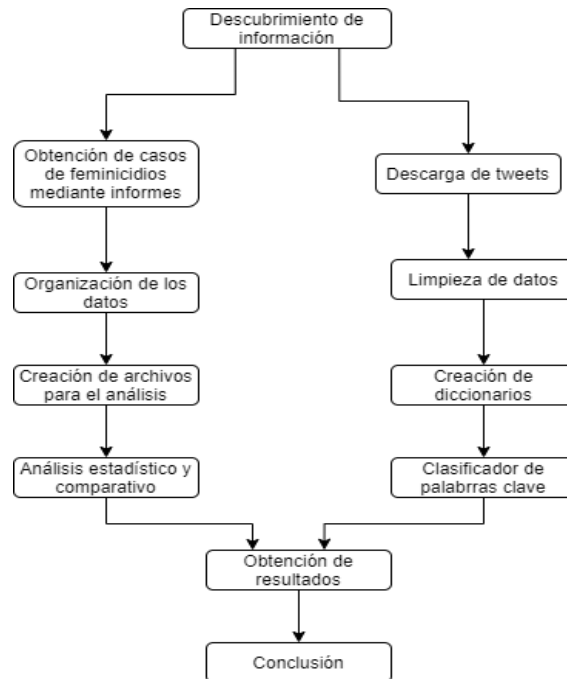


Fig. 1. Arquitectura del problema de feminicidios.

3.3 Extracción de datos con Vicinitas

El procedimiento general consistió en descargar Tweets iniciando con 10 días consecutivos para continuar con descargas semanales. Las palabras clave se identificaron como #NiUnaMas, # NiUnaMenos, #Feminicidio, #VivasNosQueremos, Acoso sexual, Movimiento feminista, Asesinato machista y Feminicida. Estas palabras permitieron la extracción de los Tweets que son de interés. La limpieza de datos es un proceso necesario en la información extraída y que debe atenderse de manera implícita en las descargas debido a que muchas palabras son conectores o abreviaturas y además diversos Tweets no son relevantes para el problema por su bajo contenido significativo en feminicidios.

En este escenario, al principio del descubrimiento de información en Twitter, se centró la atención en las descargas iniciales de los Tweets por 10 días consecutivos para analizar la información y continuar con descargas semanales. Para este propósito se utilizó una aplicación en la web llamada Vicinitas dado que proporciona datos básicos de los tweets del problema a explorar [18]. En un segundo paso, dos usuarios independientes analizaron los tweets principales (se excluyeron los re-tweets y las respuestas) para un análisis de contenido cualitativo y el desacuerdo se resolvió mediante la consulta de un tercer revisor independiente para señalar los tweets relevantes.

En este trabajo, la primera parte de descargas se realizó con Vicinitas para experimentar su utilidad, sin embargo, genera repetición de tweets por semana, esta situación dio lugar a controlar las descargas por día utilizando la API de Python.

Para utilizar el API de Twitter se necesita la librería tweepy de Python. También se requiere una cuenta en Twitter como desarrollador para dar de alta una aplicación y conseguir las credenciales de acceso/Tokens a las API de Twitter. Este acceso generalmente es desde un procedimiento en Python que permite las descargas de datos que se asocian a la búsqueda de palabras claves del tópico a tratar. El procedimiento posterior consiste en usar filtros de limpieza de datos y crear los archivos Excel de tipo .csv con los datos resultantes. El pseudocódigo del programa utilizado se presenta a continuación.

```
1.      Inicio
2.      Asignar credenciales
3.      Escribir palabras a buscar
4.      Leer palabra a buscar
5.      Escribir número de tweets a capturar
6.      Leer número de tweets
7.      Crear archivo de salida
8.      Obtener fecha y hora
9.      Asignar nombre de archivo con palabra, fecha y hora
10.     Inicializar tweets descargados como 0
11.     Mientras tweets descargados < tweets a capturar hacer
        Obtener tweet
        Si tweet contiene 'áéíóú' entonces
            Cambiar valor por 'aeiou'
            Continuar
        Sino
            Continuar
        Fin si
        Si tweet contiene url entonces
            Sustraer url
            Continuar
        Sino
            Continuar
        Fin si
        Si tweet tiene caracteres especiales entonces
            Sustraer caracteres especiales
            Continuar
        Sino
            Continuar
        Fin si
        Escribir tweet en el archivo creado
        Aumentar el valor de tweet descargo
        Repetir
12.     Fin
```

3.4 Creación de diccionarios y justificación de la exclusión de análisis de sentimientos

Para la creación de diccionarios se desarrolló una aplicación en Python que contabiliza las palabras en los archivos csv generados por la aplicación Vicinitas de Python. El programa realiza el conteo de la frecuencia de cada palabra que contengan

Tabla 1. Cifras de casos de presunto feminicidio dadas por el INEGI [18]

ESTADO	EN	FE	MA	AB	MA	JU	JU	AG	SE	OC	NO	DI	TOTA
	E	B	R	R	Y	N	L	O	P	T	V	C	L
AGUASCALIENTES	6	0	1	0	0	0	0	4	1	0	5	0	17
BAJA CALIFORNIA	12	0	0	3	14	19	23	7	3	0	2	10	93
BAJA CALIFORNIA SUR	1	0	1	4	2	0	0	1	0	1	0	0	10
CAMPECHE	0	0	0	0	1	0	1	1	0	0	8	3	14
COAHUILA	1	4	3	3	5	10	2	1	1	1	6	2	39
COLIMA	0	0	0	0	4	4	4	7	0	0	0	2	21
CHIAPAS	0	0	0	1	5	2	3	1	1	1	1	4	19
CHIHUAHUA	14	5	9	9	18	24	28	16	9	9	8	20	169
DISTRITO FEDERAL	4	0	1	6	18	10	14	4	3	1	1	13	75
DURANGO	2	0	3	0	3	3	2	1	0	1	1	0	16
GUANAJUATO	4	6	1	8	24	17	31	28	29	30	32	28	238
GUERRERO	6	5	0	0	9	18	20	3	0	1	9	13	84
HIDALGO	4	0	0	7	2	5	6	1	0	1	0	3	29
JALISCO	6	0	1	5	10	19	10	4	1	2	3	11	72
MÉXICO	13	13	14	22	18	13	21	10	12	9	18	19	182
MICHOACÁN	2	0	0	2	10	13	15	6	2	0	2	7	59
MORELOS	2	0	2	2	7	8	7	1	2	1	2	3	37
NAYARIT	0	0	0	0	3	2	1	1	0	0	1	0	8
NUEVO LEÓN	3	3	2	7	5	4	14	11	4	6	1	8	68
OAXACA	2	0	1	5	12	17	3	1	0	0	2	9	52
PUEBLA	3	3	1	2	9	12	12	2	3	8	5	7	67
QUERÉTARO	0	0	0	0	3	4	1	2	0	0	2	0	12
QUINTANA ROO	2	0	1	2	3	3	5	2	2	0	4	5	29
SAN LUIS POTOSÍ	1	0	2	9	3	6	1	3	0	1	2	11	39
SINALOA	6	2	0	1	2	5	10	4	6	1	2	3	42
SONORA	5	3	5	3	6	10	3	5	7	2	11	6	66
TABASCO	0	1	2	4	5	9	4	1	1	0	2	3	32
TAMAULIPAS	2	0	0	2	6	1	2	2	1	2	1	5	24
TLAXCALA	1	0	0	0	2	7	1	1	0	0	5	1	18
VERACRUZ	5	3	0	12	10	22	17	5	0	0	1	9	84
YUCATÁN	0	0	0	1	2	0	0	0	0	0	0	1	4
ZACATECAS	4	1	1	7	6	3	9	2	6	3	2	3	47
TOTALES	111	49	51	127	227	27	27	138	94	81	139	20	1766
						0	0					9	

más de 5 letras, las palabras con menos de 5 caracteres se excluyen porque no son significativas para el estudio debido a que son preposiciones o bien palabras fuera del contexto. No se incluyen siglas ni acrónimos en el estudio. En general las palabras relacionadas con feminicidios son mayor o igual a 5 letras. Posteriormente se genera un archivo Excel para el tratamiento de los registros de las palabras y la frecuencia asociada.

Los datos se analizaron por semana para efectos de comparación sobre la intensidad de las palabras, frecuencia, ausencia o aparición de una nueva palabra, la primera semana del 15 al 21 de marzo y la segunda semana del 22 al 27 de marzo del año en curso.

El Análisis de Sentimientos AS también conocido como minería de opinión, un término muy discutido, pero no del todo comprendido, es muy útil en la monitorización de las redes sociales porque permite proporcionar una idea de la opinión pública sobre temas de interés. Para el trabajo que se aborda en este documento, se desarrolló AS, sin embargo, no fue posible dividir el conjunto de palabras en sentimientos positivos o

Tabla 2. Palabras con más frecuencia dentro de los diccionarios

PALABRA	FRECUENCIA	PALABRA	FRECUENCIA	PALABRA	FRECUENCIA
FEMINICIDIO	10089	DENUNCIA	3874	JUSTICIA	2816
FEMINICIDA	9283	CHICA	3836	REPORTAR	2813
VICTIMAS	9129	ASEINADA	3679	LUCHA	2715
MUJER	7538	PUEBLA	3383	911	2707
DEMASIADAS	5628	MACHISTA	3237	FORZÁNDOLA	2707
CRISIS	5393	CASOS	3114	ASESINO	2620
VIDA	5147	TODAS	3049	DENUNCIADO	2601
HOMBRES	4902	MÉXICO	2972	DENUNCIAS	2585
VICTIMA	4598	FEMINISTA	2881	CUARENTENA	2570
SOFIA	4242	HISTORIA	2828	FEMINICIDIOS	2493

Tabla 3. Diccionario extendido

PALABRA	FRECUENCIA	PALABRA	FRECUENCIA	PALABRA	FRECUENCIA
MARZO	2328	#FEMINICIDIOS	1918	AGRESOR	1711
ASESINATO	2278	VIOLACION	1901	FEMINISTA	1707
CADAVER	2252	#VIOLENCIAMACHISTA	1872	ABUSO	1607
#FEMINICIDIO	2213	IVANA	1776	MATO	1603
#VIOLENCIADEGENERO	2176	FAMILIA	1773	DERECHOS	1598
ASESINADAS	2171	MUCHAS	1772	MIEDO	1589
DENUNCIAR	2139	ESTADO	1765	HERMANA	1547
MENOR	2112	DETENIDO	1762	MADRE	1528
NIÑAS	2099	#FEMINICIDIOINFANTIL	1756	AGRESORES	1512
#TODASDUELEN	2024	NICOLE	1724	ULTIMA	1466



Fig. 2. Wordcloud, palabras más frecuentes en el diccionario.

negativos dado que predominó en las pruebas la neutralidad de las palabras en el análisis. La tendencia es constante en otros estudios y se presume que se debe a la cantidad de caracteres permitidos por la red social, aunque se ha observado que el uso de las palabras emotivas aumenta la viralización.

4. Estudio de estadísticas de casos de feminicidios y creación de diccionarios

Con la finalidad de entender el problema base y la posterior correlación con lo expuesto en la red social se estudiaron las estadísticas de los casos de feminicidios en México a través de los datos del Instituto Nacional de Estadística, Geografía e Informática INEGI. El mayor número de muertes contabilizadas extraoficialmente son etiquetadas como “homicidios” con la etiqueta del género de la persona que falleció.

El repositorio del INEGI concentra de manera oficial entre 1990 a 2018, cifras de asesinatos en una base de datos, pero no especifica como feminicidios las muertes a mujeres cuando el crimen es por cuestión de género (dada la dificultad de señalar un feminicidio). En general, las bases de datos oficiales no presentan información concluyente a las muertes registradas de mujeres, ni como asesinatos o “feminicidios”, por tanto, no es posible categorizar “la clase de muertes de feminicidios”, por ello la importancia de acudir a otras herramientas, tanto computacionales como manuales, para clasificar esos casos. En este punto, es posible analizar los homicidios ocurridos a mujeres en distintos rangos y características con diccionarios a través de descargas en Twitter.

Para este trabajo y para efectos de comparación con los diccionarios PLN, fueron considerados los “homicidios” de mujeres del INEGI por mes en cada uno de los estados de la República. La Tabla 1 muestra el concentrado de datos según cifras del INEGI.

Herramientas para descargas en Twitter y descripción de diccionarios: Tweepy y API Search

4.1 Creación de Diccionarios

Un diccionario es una colección no ordenada de objetos cuya estructura es empleada para almacenar información de toda clase. Para este artículo, los diccionarios tienen una importancia sustancial para el análisis estadístico descriptivo y también para identificar las palabras de mayor presencia en el tema de feminicidios.

El procedimiento aquí consiste en abrir diferentes archivos para enumerar la frecuencia de cada una de las palabras almacenadas en los archivos y consecuentemente crear un archivo .csv que contenga la palabra vinculada al número de frecuencia.

Por otra parte, se recurrió a la generación de una Wordcloud, para la creación de una nube de palabras desarrollada en Python. Este programa hace uso de distintas librerías como Wordcloud, de útil conveniencia para la creación de ilustraciones, donde destaca el análisis textual de datos. El resultado de este script fue una imagen con las palabras que más representan el tema, de esta forma se puede tener una idea aproximada del contexto que abarca el tema “feminicidios”. Las palabras se muestran la Fig. 2.

Cada una de las palabras resaltadas en la Figura 2 han sido consideradas por las bondades que suministran al programa de diccionarios. En la siguiente Tabla 2 se muestran las 10 palabras más notificadas en todos los datos recolectados, así como su respectiva frecuencia.



Fig. 3. Gráfico asociado a la tabla 2.

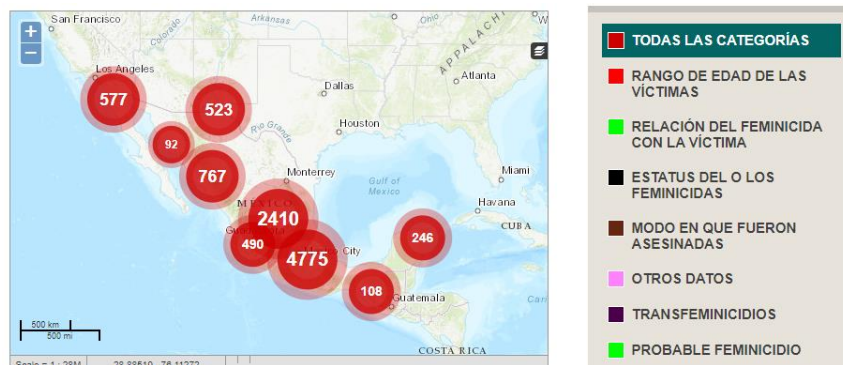


Fig. 4. Mapa de feminicidios de María Salguero.

La periodicidad de los hashtags responde al conteo de 30 días para crear el diccionario final como se observa en las tablas 2 y 3. En las tablas las palabras se mantienen sin acentos como se encontraron en Twitter. En dichas tablas, todas las columnas son las palabras mas representativas que conforman el diccionario de feminicidios. En la siguiente figura se señalan las palabras más insistentes:

Es interesante resaltar que el tiempo real en las redes posiciona algunos casos en particular en el interés del público como como Nicole, Ivana y Sofía, tres nombres de mujeres víctimas de feminicidio del periodo estudiado. Estos nombres desaparecen y aparecerán otros en las redes. Adicionalmente, las palabras más usadas con relación al feminicidio son: víctimas, mujer y demasiadas, y los hashtags NiUnaMas, NiUnaMenos.

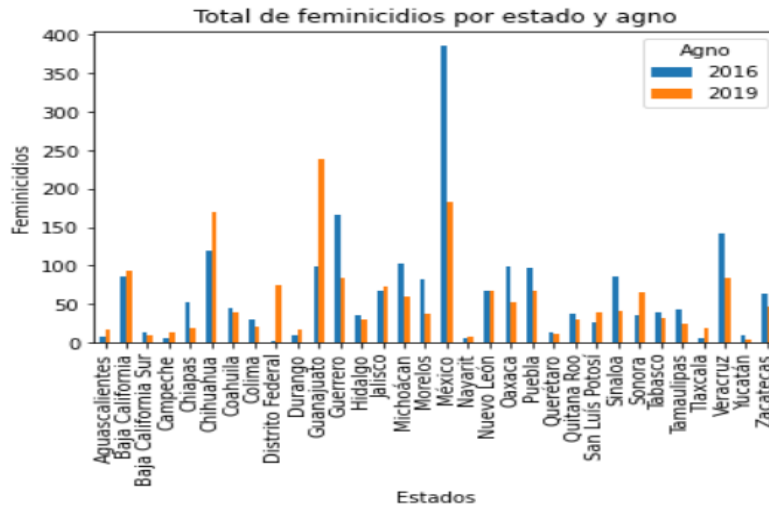


Fig. 5. Casos de feminicidios por estado en los años 2016 y 2019.

Obtención de casos de feminicidios

Una vez creados los diccionarios y sumados a la información obtenida del INEGI, se compararon los resultados con el mapa publicado por María Salguero, en el cual se documenta mediante informes aprobados los casos de feminicidios por año en todo el país [19]. La figura 4 revela gráficamente diversas categorías, pero no existe una manera exacta de saber el número de casos de feminicidios por estado. En este punto, se contabilizaron los casos registrados por estado entre 2016 y 2019.

Resultados: Casos de feminicidios

Se utilizó la librería “pandas” de Python para ordenar y ajustar la información del INEGI que se presentó en la tabla 1. Como consecuencia se generó la información contenida en la tabla 4 donde se revelan los resultados agrupados por año y por estado, lo cual pone en evidencia de forma tabular el aumento de feminicidios.

Se observa una inconsistencia en los feminicidios registrados con baja presencia, y la respuesta no es tan simple. Los feminicidios realmente son cifras bajas y cuando sucede uno de ellos no se califica inmediatamente como feminicidio. Esta situación se rodea de la investigación profunda para señalar como feminicidio un descenso de una mujer, generalmente se registran con otro nombre.

La Tabla 4 refleja el número de feminicidios registrados en cada estado en 2016 y en 2019. El estado con mayor número de feminicidios en 2016 fue el estado de México y los menores fueron Campeche y Tlaxcala con 5. Del mismo modo para 2019 (en la tabla aparece el distrito federal por separado, pero se sumarán sus datos al estado de México dando un total de 388 casos). Del mismo modo para 2019, el estado con mayor número de feminicidios fue Guanajuato con 238 y el menor fue Nayarit.

Tabla 3. Medidas de tendencia central.

	FEMINICIDIOS EN 2016	FEMINICIDIOS EN 2019
CONTEO	32.0000	32.0000
PROMEDIO	65.0625	55.1875
DESV. STD.	72.7248	53.2640
MINIMO	2.0000	4.0000
25%	14.0000	18.7500
50%	44.0000	39.0000
75%	88.7500	69.0000
MÁXIMO	386.0000	238.0000

Tabla 4. Feminicidios por estado en los años 2016 y 2019.

ESTADO	2016	2019	ESTADO	2016	2019
AGUASCALIENTES	8	17	MICHOACÁN	102	59
BAJA CALIFORNIA	86	93	MORELOS	82	37
BAJA CALIFORNIA SUR	14	10	NAYARIT	6	8
CAMPECHE	5	14	NUEVO LEÓN	68	68
COAHUILA	45	39	OAXACA	99	52
COLIMA	30	21	PUEBLA	97	67
CHIAPAS	52	19	QUERÉTARO	14	12
CHIHUAHUA	119	169	QUINTANA ROO	37	29
DISTRITO FEDERAL	2	75	SAN LUÍS POTOSÍ	27	39
DURANGO	10	16	SINALOA	85	42
GUANAJUATO	99	238	SONORA	36	66
GUERRERO	166	84	TABASCO	39	32
HIDALGO	36	29	TAMAULIPAS	43	24
JALISCO	68	72	TLAXCALA	5	18
MÉXICO	386	182	VERACRUZ	142	84

En la Figura 5, se puede notar que, en 2016 los estados de México, Veracruz y Guerrero registraron más feminicidios.

Sin embargo, para el año 2019 el grafico muestra que los estados en donde se registraron más casos de feminicidio son Guanajuato, el estado de México y Chihuahua, y también se distingue una disminución de los casos registrados.

Con esta información podemos concluir que en el año 2019 menos feminicidios se denunciaron y aprobaron con ese nombre. Estos resultados hacen suponer que descenso obedece a que se popularizo la denuncia formal y acusación popular contra la violencia de género gracias a la capacidad expresiva de las redes sociales, lo que provocó una alerta en que las autoridades para atender casos de violencia y evitar un feminicidio.

Por otra parte, se considera que tanto en campañas de sensibilización como en la conciencia en las medidas de castigo fueron diversificadas y extrapoladas, lo cual fomentó la prevención de feminicidios.

La información de la Tabla 4 afirma que en el año 2016 ocurrieron por estado un promedio de 65 feminicidios, mientras que en el año 2019 se redujo a 55. También se observa que en 2016 el mínimo número de casos registrados es de 2 y el máximo de 386, mientras que en el año 2019 el mínimo de casos registrados fue de 4 y el máximo de 238.

5. Conclusiones

Podemos concluir que el estudio de los datos orientados a un aspecto social es muy importante porque con un orden y análisis consecuente, es posible encontrar patrones, realizar investigaciones, dar respuesta a preguntas, entre otros. Lo importante es saber obtener esta información para después transformarla en conocimiento.

Para el problema de feminicidios en particular, realizar un estudio de lo sucedido en México se considera de sustancial importancia porque se consiguió reconocer que no se le da el debido seguimiento a este delito, se habla de casos y noticias particulares y se difunden las noticias y comentarios sobre el tema masivamente, aunque no de fuentes confiables.

Con ayuda del conteo de los casos de feminicidios obtenidos del INEGI, se identificó que la violencia de género era muy común, entonces, el número de mujeres muertas no era considerado tan alarmante. Sin pérdida de generalidad, con el tiempo, estas muertes tomaron más importancia porque el número fue creciendo considerablemente. A partir de ello fue de interés el uso de las redes sociales que proporcionan generosa información pública sobre opiniones. Décadas atrás, cuando estos medios no eran tan utilizados, se había establecido la conjetura de que los feminicidios aumentaron debido a que la información no se extrapolaba y por tanto no se conocían ni la cantidad ni la magnitud del problema. El PLN ha contribuido a encontrar información y tendencias en la gran cantidad de mensajes de twitter y otras redes, para entender problemas sociales como el que se ha abordado en este trabajo. Se desea que los resultados sean difundidos a las autoridades competentes para la toma de decisiones.

Cabe subrayar que la información extraída por las redes sociales, permiten establecer el problema central con la alternativa de subrayar las palabras que lo definen, y consecuentemente sea posible crear diccionarios con las palabras de mayor ocurrencia. Estas palabras se identifican con el descubrimiento de información utilizando descargas de la red social como es el caso de este trabajo, donde diariamente se re-copilaron Tweets asociados al tópico de feminicidios y así, se crearon los respectivos diccionarios con palabras útiles al tema. Con este procedimiento ha sido posible contar con información sobre denuncias y delitos diversos sobre feminicidios, de tal modo que dependencias legales tendrán a su disposición información calibrada y estructurada que obedece a un procesamiento computacional.

6. Trabajo futuro

A partir de las conclusiones ya mencionadas y observando los resultados alcanzados, es posible identificar que este tipo de temas requiere de un seguimiento e investigación

extensa y constante para obtener resultados mucho más precisos, por ello se han considerado algunos cambios en la creación de diccionarios para poder desarrollar un análisis de sentimientos consistente.

Principalmente es importante la recolección de datos con la API REST empresarial de Twitter para facilitar la extracción de la información en distintos periodos de tiempo y sin restricciones.

Para los datos estadísticos de los casos registrados, es necesario considerar más variables que ayuden a describir los factores más comunes presentes en este tipo de delito para que ayuden a una mejor interpretación. Por otra parte, se plantea considerarse la combinación de los programas de creación de diccionarios y de Wordcloud, para poder realizar el proceso de manera simultánea.

Siendo optimistas, el uso de estas redes teóricamente permite empatía y unión, lo que presume una mayor difusión del problema, y al mismo tiempo establece una reflexión en las personas.

Referencias

1. Russell, D.; Harmes, R.; Lagarde, M.: Definición de feminicidio y conceptos relacionados. Feminicidio: una perspectiva global. Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades, Universidad Nacional Autónoma de México (CEIICH - UNAM), pp. 73–96 (2006)
2. Denis, D., Rodríguez, A.: Especial: Feminicidios en México, una epidemia en aumento. El país. (2020) <https://elpais.com/especiales/2017/feminicidios-en-mexico/#>.
3. Reyes Silva, P.G.: Redes en Twitter y la defensa de la mujer peruana (ante el acoso y el feminicidio). Universitat Oberta de Catalunya, repositorio institucional (2020) <http://hdl.handle.net/10609/104066>.
4. Vilares, M., Trigo, E.S., Gómez Rodríguez, C., Alonso, M. A.: Campus de Elvina, A.: Tecnologías de la lengua para análisis de opiniones en redes sociales. Procesamiento del Lenguaje Natural (2020) <http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/5502>.
5. Sarmiento Pacanchique, E. L., Silva Téllez, D. C.: Análisis de Flujos de Información de la Red Social Twitter Caso de Estudio: Impacto de las Tendencias en Redes Sociales para las Operaciones del Mercado de Acciones de la Empresa ECOPETROL. Repositorio Institucional Universidad Distrital – RIUD. <http://repository.udistrital.edu.co/bitstream/11349/5911/1/TesisAnalisisDeFlujosDeInformacionDeLaRedSocialTwitter.pdf>.
6. ONU: Declaración sobre la eliminación de la violencia contra la mujer. Naciones Unidas Derechos Humanos, oficina del alto comisionado (1993) <https://www.ohchr.org/sp/professionalinterest/pages/violenceagainstwomen.aspx>.
7. COE: Convenio del Consejo de Europa sobre prevención y lucha contra la violencia contra las mujeres y la violencia doméstica, Serie de Tratados del Consejo de Europa – nº 210. Concil of Europe (2011) <https://rm.coe.int/1680462543>.
8. Abreu Maqueda, M.L.: La violencia de género: Entre el concepto jurídico y la realidad social. Revista electrónica de ciencia penal y criminología (2006) <http://criminnet.ugr.es/recpc/08/recpc08-02.pdf>.
9. Martínez Bullé-Goyri, V.M.: Reflexiones sobre la dignidad en la actualidad. Boletín Mexicano de Derecho Comparado. (2013) <https://www.elsevier.es/es-revista-boletin>

- mexicano-derecho-comparado-77-articulo-reflexiones-sobre-dignidad-humana-actualidad-S0041863313711219.
10. Sepulveda López, M.: La dignidad humana como un valor ético jurídico implicado en la bioética y el bioderecho. *Misión Jurídica: Revista de derecho y ciencias sociales* (2009) <https://www.revistamisionjuridica.com/la-dignidad-humana-como-valor-etico-juridico-implicado-en-la-bioetica-y-el-bioderecho/>.
 11. Gil Millán, M.: La violencia sexual como un atentado contra la dignidad de la mujer. *Revista de Derecho UNED* (2015) http://e-spacio.uned.es/fez/eserv/tesisuned:Derecho-Mgil/GIL_MILLAN_Maximiliana_Tesis.pdf.
 12. López Takeyas, B.: *Lenguaje Natural*. (2017) [http://www.itnuevolaredo.edu.mx/takeyas/Apuntes/Inteligencia%20Artificial/Apuntes/tareas_alumnos/PLN/PLN\(2005-II\)](http://www.itnuevolaredo.edu.mx/takeyas/Apuntes/Inteligencia%20Artificial/Apuntes/tareas_alumnos/PLN/PLN(2005-II)).
 13. Brookshear, J.G.: *Teoría de la Computación: lenguajes formales, autómatas y complejidad*. México: Pearson Educación (1993)
 14. Martín Vide, C.; Mitrana, V.; Paun, G.: *Formal Languages and Applications*. Springer (2004)
 15. Palma, J.T.; Paniagua-Aris, E.: *Ingeniería del Conocimiento. De la Extracción al Modelado de Conocimiento. Inteligencia Artificial* (2000)
 16. Orihuela, J.: *Internet: la hora de las redes sociales*. Nueva Revista de política, cultura y arte. (2008). http://mccd.udc.es/orihuela/documentos/nueva_revista_08.pdf.
 17. González Duque, R.: *Introducción - ¿Qué es Python? Python para todos*, pp.7 (2010)
 18. *Defunciones por homicidios*. INEGI (2020) <https://www.inegi.org.mx/sistemas/olap/proyectos/bd/continuas/mortalidad/defuncioneshom.asp?s=est>.
 19. Salguero, M.: *Los Feminicidios en México* (2020) <https://feminicidiosmx.crowdmap.com/Feminicidios-Mx>.

Clasificación de información personal sensible

Sara De Jesús Sánchez, Jorge Enrique Coyac Torres,
Eleazar Aguirre Anaya, Hiram Calvo

Instituto Politécnico Nacional,
Centro de Investigación en Computación,
México

sdejesuss2100@alumno.ipn.mx, jcoyact1900@alumno.ipn.mx,
eaguirrea@ipn.mx, fcalvo@ipn.mx

Resumen. La exposición de la información personal sensible en medios públicos representa grandes riesgos tanto para los individuos, quienes son los titulares de sus propios datos personales, como para las empresas que requieren tratar los datos personales de sus clientes, proveedores y empleados, siendo las mismas empresas responsables de mantener la privacidad de dichos datos, más aún, tratándose de datos personales sensibles. Para disminuir los riesgos de exposición de la información personal sensible, es preciso clasificarla, así los responsables podrán tomar las medidas correspondientes para mantener la confidencialidad de los datos. Este trabajo muestra la factibilidad de utilizar algoritmos de aprendizaje automático en la identificación y clasificación de textos que contienen datos personales sensibles. Los resultados de los experimentos fueron satisfactorios, empleando los algoritmos naive Bayes, regresión logística y máquina de vectores de soporte y algunas técnicas de procesamiento de lenguaje natural, como la normalización y eliminación de palabras auxiliares.

Palabras clave: Ciberseguridad, información sensible, aprendizaje automático, clasificación, procesamiento de lenguaje natural.

Classification of Sensitive Personal Information

Abstract. The exposure of sensitive personal information in public media represents big risks both for individuals, who are holders of their own personal data, and for companies that must process the personal data of their customers, suppliers and employees. These companies are responsible for maintaining the privacy of such data, even more so, in the case of sensitive personal data. In order to reduce the risks of exposure of sensitive personal information, it must be classified, so those responsible take the corresponding measures to maintain the data confidentiality. This work shows the feasibility of using machine learning algorithms to identify and classify texts that contain sensitive personal data. The results of the experiments were satisfactory, using

naive Bayes, logistic regression and support vector machine algorithms, in addition to some natural language processing techniques, such as lemmatization and stop words elimination.

Keywords: Cybersecurity, sensitive information, machine learning, classification, natural language processing.

1. Introducción

La protección de las personas con respecto al tratamiento automatizado de los datos personales tiene origen en el Convenio 108 del Consejo de Europa en 1981. Además de los países miembros de la Unión Europea, en América, países como Argentina, Uruguay y México han adoptado este convenio como una herramienta global para el intercambio efectivo y seguro de información. Uno de los puntos principales de este convenio es garantizar la confidencialidad de los datos personales sensibles por parte de las organizaciones responsables de ellos [2].

Los datos personales, en general, son los concernientes a una persona física identificada o identificable cuya manifestación puede ser textual, gráfica, acústica o fotográfica [13]. Pueden ser de diferentes tipos: identificativos, laborales, académicos, de salud, de patrimonio.

Un subconjunto de los datos personales son los llamados datos personales sensibles, definidos por la Comisión Europea como los datos que, por su naturaleza, puedan atentar contra las libertades fundamentales o la intimidad, y están sujetos a condiciones de tratamiento específicas [14].

Los datos personales se consideran sensibles en caso de que revelen: origen racial o étnico, opiniones políticas, creencias religiosas, filosóficas y morales, afiliación sindical, datos genéticos, datos biométricos, datos relativos a la salud, datos relativos a la vida sexual o a la orientación sexual de una persona [3].

Las organizaciones deben garantizar la seguridad de la información personal de la cual son responsables, de acuerdo con el nivel de sensibilidad, valor y criticidad de ésta [10]. Por tal motivo, clasificarla es la base para disminuir los riesgos en su seguridad. Estos riesgos pueden llegar a ser enormes tanto para las personas, como para las organizaciones y los gobiernos.

En México, por ejemplo, la Ley Federal de Protección de Datos Personales en Posesión de Particulares (LFPDPPP) establece que las multas a los responsables de los datos personales van desde los 100 hasta 320,000 días de salario mínimo vigente en la CDMX y que, en el caso de datos personales sensibles, las sanciones podrán incrementarse hasta por dos veces los montos establecidos [4].

Pero las repercusiones no son únicamente económicas, sino también en la credibilidad y confianza en las organizaciones responsables de los datos personales y en las tecnologías que se utilizan para el tratamiento (obtención, almacenamiento, modificación, copia, eliminación, procesamiento, etc.) de los datos.

Ahora bien, como se mencionó anteriormente, la clasificación de la información es la base para garantizar su seguridad, es decir, para garantizar la confidencialidad, integridad y disponibilidad de esa información. Este tipo de clasificación es desafiante, debido a diversos factores, por ejemplo: no existe consenso sobre los datos que componen la categoría de datos personales sensibles [11]; se especifican de acuerdo con el riesgo que implican, por lo tanto, la información sensible es incierta; es compleja, pues algunas frases pueden ser sensibles en un contexto y no serlo en otro; el conjunto de datos de entrenamiento debe mantenerse actualizado, es cambiante; se requiere un alto grado de precisión al asignar el nivel de seguridad para que las organizaciones responsables puedan tomar las medidas correspondientes. La clasificación de seguridad de la información puede hacerse en dos niveles (sensible o no) o en más (no clasificada, confidencial, secreta y alto secreto, por ejemplo). La clasificación también podría hacerse sobre todo un documento o en las partes que lo componen.

El tipo de información a clasificar en este proyecto es textual, contenida en documentos o textos que pudieran ser expuestos en algún medio electrónico. La clasificación textual de información personal sensible se aborda como un problema de procesamiento de lenguaje natural o NLP (por sus siglas en inglés), que usa el lenguaje dentro de un documento para clasificarlo en una categoría particular [1].

La clasificación automática de seguridad es una tecnología que permite predecir el nivel de seguridad de la información de un documento, se abordó como un problema de investigación en 2005 [7] y desde entonces se han propuesto numerosos métodos para resolverlo, tales como el aprendizaje automático y las redes neuronales, principalmente. Se han empleado distintos tipos de documentos en ámbitos específicos, como el médico, el organizacional y el militar, en lenguas china, coreana e inglesa.

El objetivo de este proyecto es identificar y clasificar textos con información personal sensible, mediante técnicas de aprendizaje automático, para prevenir su exposición en espacios públicos. Este proyecto está enfocado en la clasificación de textos en español para saber si contienen datos personales sensibles o no. Para ello se aplicaron tres algoritmos de aprendizaje automático junto con técnicas de NLP, con el fin de encontrar aquéllos que ofrezcan la mayor precisión en la clasificación.

2. Estado del arte

En 2005 Hassan Mathkour utilizó árboles binarios para la clasificación de seguridad basada en estructuras retóricas en idioma Árabe [7], como se mencionó anteriormente, esta publicación planteó el problema de la clasificación automática de seguridad.

Graham McDonald et al. emplearon, en 2015 y 2017, n-gramas, análisis gramatical (Part of Speech POS) y máquinas de vectores de soporte (SVM) obteniendo una efectividad del 90 % con patrones fijos para la clasificación de

sensibilidad de textos en inglés [9]. El uso de n-gramas y análisis gramatical para el preprocesamiento es una referencia para identificar secuencias de textos sensibles.

En 2019, Yan Liang et al. utilizaron aprendizaje incremental y comparación de similitud (ILSC) para la clasificación de textos sensibles en Chino con una efectividad del 86 %, incremental support vector machine (ISVM) con 87 %, online random forest (ORF) con 84 % y naive Bayes (NB) con 82 % [6]. Se compararon tres algoritmos de aprendizaje automático con buenos resultados, siendo ISVM el más efectivo.

En ese mismo año, Gousheng Xu et al. emplearon redes neuronales de convolución (CNN) con 95 % de efectividad y las recurrent neural networks (RNN) con 94 % para detectar información sensible en textos no estructurados en chino [15]. Se compararon las dos principales redes neuronales con una efectividad alta.

Huimin Jiang et al. también en 2019 diseñaron un clasificador de datos médicos sensibles en chino, obteniendo un 90 % de precisión con el algoritmo SVM, 85 % con naive Bayes y 80 % con KNN [5]. Nuevamente vemos que el algoritmo SVM alcanzó mayor precisión.

En 2020, Srdjan Matic et al. diseñaron un clasificador de URLs sensibles con un 88 % de efectividad [8]. No mencionan el algoritmo utilizado, sin embargo la efectividad resultó similar a los otros trabajos revisados.

Ji-Sung Park et al., en 2020 desarrollaron un sistema de prevención de pérdida de datos (DLP) para clasificar palabras en coreano, en categorías de datos personales sensibles utilizando reconocimiento de entidades nombradas (NER) [12]. Se puede considerar el uso de NER para identificar entidades como nombres, lugares, organizaciones que podrían clasificarse como información sensible.

Como podemos ver, los métodos que actualmente han logrado mayor precisión en la clasificación de información sensible, son los basados en aprendizaje automático y los basados en redes neuronales. Los tipos de documentos y los ámbitos sobre los que se han aplicado son muy variados, así como los idiomas. Se han aplicado técnicas de PLN como n-gramas, POS y NER.

No se encontraron referencias sobre sensibilidad de datos personales en idioma español, tampoco sobre el algoritmo de regresión logística, ni la normalización o eliminación de palabras auxiliares.

En el presente artículo se presenta una clasificación de textos en español, para identificar si contienen información personal sensible, utilizando algoritmos de aprendizaje automático y técnicas de PLN. En esta propuesta la clasificación es binaria, las dos clases de objetos son: sensible (1) y no sensible (0).

3. Desarrollo de la solución

Para esta solución inicial se propone emplear dos técnicas de PLN y tres de los algoritmos más representativos del aprendizaje automático sobre un conjunto de datos de prueba con textos cortos en español para clasificarlos según contengan o no información sensible.

Tabla 1. Exactitud obtenida

RL	SVM	nB
0.8401	0.8525	0.8395

El conjunto de datos consiste en 60 textos, el 60 % de ellos contiene información personal sensible, el 40 % no. El conjunto de datos es almacenado en un archivo de texto separado por tabuladores, donde en cada renglón se tiene un texto y su etiqueta asignada con los valores de 1, si contiene información personal sensible, y de 0 en el caso contrario. El conjunto de datos está conformado por textos cortos en español, obtenidos de publicaciones en redes sociales y etiquetados manualmente.

El modelo de espacio vectorial está basado en el modelo bolsa de palabras (Bag of Words BOW), donde las palabras del vocabulario son las características de los objetos. Los vectores, en este caso, contienen los valores de frecuencia de término (Tf), es decir qué tanto se repite cada palabra en un texto.

Con el fin de reducir el tamaño de los vectores y con ello facilitar la clasificación de los textos, se hace un preprocesamiento a los objetos antes de entrenar y probar los algoritmos. Este preprocesamiento consiste en aplicar técnicas de PLN a los textos del conjunto de datos.

Las técnicas de PLN empleadas son la normalización y la eliminación de palabras auxiliares. Normalizar un texto es sustituir cada palabra flexionada o derivada (es decir, en plural, conjugada, etc.) por su forma normal o lema. Las palabras auxiliares o stop words son aquellas que, aunque son muy comunes, son poco relevantes (como los artículos, conjunciones, preposiciones, etc.), por tal motivo se filtran. El preprocesamiento se lleva a cabo mediante funciones de la biblioteca Stanza de Python.

Posteriormente el conjunto de datos es dividido de la siguiente manera: se toma el 75 % de los objetos para el entrenamiento y el 25 % para las pruebas. Haciendo uso de la biblioteca sklearn de Python, se aplican tres de los principales algoritmos clasificadores de aprendizaje automático: Regresión Logística, naive Bayes y Máquina de Vectores de Soporte.

4. Experimentos y resultados

Con una muestra de 20 ejecuciones, los resultados obtenidos en la exactitud (Accuracy) de los algoritmos de Regresión Logística, Máquina de Vectores de Soporte y naive Bayes, se muestran en el Cuadro 1.

Se observan exactitudes similares en los tres algoritmos, ligeramente superior con SVM. Los resultados son cercanos a los reportados con algoritmos de aprendizaje automático encontrados en el estado del arte.

5. Conclusiones y trabajo futuro

En el presente artículo se ha mostrado la factibilidad de emplear algoritmos de aprendizaje automático en la clasificación textual de la información personal sensible, para coadyuvar en su confidencialidad y, por lo tanto, en su seguridad. Se utilizaron tres algoritmos de aprendizaje automático para la clasificación de textos con información personal sensible, obteniendo resultados similares a los vistos en el estado del arte.

Es necesario considerar que el conjunto de datos empleado es pequeño y es conveniente incrementarlo para los siguientes estudios.

Se debe analizar si es necesario agregar otras técnicas de PLN en el preprocesamiento, tales como el análisis de expresiones regulares (Regular Expressions RE), el reconocimiento de entidades nombradas (Named Entity Recognition NER), la desambiguación del sentido de las palabras (Word Sense Disambiguation WSD), o el uso de n-gramas, así como formar los vectores utilizando TfIdf como valores de las características de los objetos. Un trabajo posterior es emplear algoritmos de clasificación basados en redes neuronales, que probablemente requieran de algún modelo de espacio vectorial distinto a BOW.

El trabajo por hacer en este proyecto, además de la utilización de distintos algoritmos de clasificación, es el diseño de una herramienta que utilice los algoritmos que funcionan mejor en la clasificación de información personal sensible, para prevenir su exposición en espacios públicos.

Agradecimientos. Este trabajo fue apoyado por CONACyT, COFAA, IPN, IPN-EDI, IPN-SIP, OEA, Cisco y la Fundación Citi, gracias a los proyectos SIP 20210189, SIP 20211758 y al proyecto Plataforma de Identificación, Clasificación y Monitoreo de Información sensible para entidades de Gobierno Federal, ganador del Fondo de Innovación en Ciberseguridad de Latinoamérica en 2021.

Referencias

- [1] Bassens, A., Beyleveld, G., Krohn, J.: Deep Learning Illustrated (2019)
- [2] Diario Oficial de la Federación: Decreto promulgatorio del protocolo adicional al convenio para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal (2018), https://www.dof.gob.mx/nota_detalle.php?codigo=5539474&fecha=28/09/2018
- [3] Europea, C.: ¿Que datos personales se consideran sensibles? (2021), https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_es
- [4] INAI: Ley federal de protección de datos personales en posesión de particulares. En línea (2021), https://home.inai.org.mx/?page_id=1870I&mat=p

- [5] Jiang, H., Chen, C., Wu, S., Guo, Y.: Classification of medical sensitive data based on text classification. IEEE (may 2019) doi: 10.1109/icce-tw46550.2019.8991726
- [6] Liang, Y., Wen, Z., Tao, Y., Li, G., Guo, B.: Automatic security classification based on incremental learning and similarity comparison. IEEE (may 2019) doi: 10.1109/itaic.2019.8785798
- [7] Mathkour, H., Touir, A., Al-Sanie, W.: Automatic information classifier using rhetorical structure theory. In: Klopotek, M. A., Wierzchon, S. T., Trojanowski, K. (eds) Intelligent Information Processing and Web Mining, Proceedings of the International IIS: IIPWM'05 Conference held in Gdansk, Poland, June 13-16, 2005. Advances in Soft Computing, vol. 31, pp. 229–236. Springer (2005)
- [8] Matic, S., Iordanou, C., Smaragdakis, G., Laoutaris, N.: Identifying sensitive urls at web-scale. In: Proceedings of the ACM Internet Measurement Conference. pp. 619–633. IMC '20, Association for Computing Machinery, New York, NY, USA (10 2020) doi: 10.1145/3419394.3423653 , <https://doi.org/10.1145/3419394.3423653>
- [9] McDonald, G., Macdonald, C., Ounis, I.: Using part-of-speech n-grams for sensitive-text classification. In: Proceedings of the 2015 International Conference on The Theory of Information Retrieval. pp. 381–384. ICTIR '15, Association for Computing Machinery, New York, NY, USA (9 2015) doi: 10.1145/2808194.2809496 , <https://doi.org/10.1145/2808194.2809496>
- [10] OEA: Clasificación de datos. En línea (2019), <https://www.oas.org/es/sms/cicte/docs/ESP-Clasificacion-de-Datos.pdf>
- [11] OEA: Glosario-Protección de Datos Personales (2021), http://www.oas.org/es/sla/ddi/proteccion_datos_personales_glosario.asp
- [12] sung Park, J., woo Kim, G., ho Lee, D.: Sensitive data identification in structured data through genner model based on text generation and ner. In: Proceedings of the 2020 International Conference on Computing, Networks and Internet of Things. pp. 36–40. CNIOT2020, Association for Computing Machinery, New York, NY, USA (4 2020) doi: 10.1145/3398329.3398335 , <https://doi.org/10.1145/3398329.3398335>
- [13] Union Europea: Grupo de trabajo del artículo 29l dictamen 4, 2007 sobre el concepto de datos personales. ¿que son los datos personales? (2007), <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136.es.pdf>
- [14] Union Europea: Reglamento ue 2019/679 del parlamento europeo y del consejo del 27 de abril de 2016, relativo a la proteccion de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulacion de estos datos (2016), <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32016R0679#d1e1547-1-1>.
- [15] Xu, G., Qi, C., Yu, H., Xu, S., Zhao, C., Yuan, J.: Detecting sensitive information of unstructured text using convolutional neural network. IEEE (oct 2019) doi: 10.1109/cyberc.2019.00087

Análisis de sentimientos para el estudio del comportamiento en noticias de COVID-19 en redes sociales como series de tiempo

Jesús Hernández Padilla, Hiram Calvo, Ildar Batyrshin

Instituto Politécnico Nacional,
Centro de Investigación en Computación,
Laboratorio de Ciencias Cognitivas Computacionales,
México

jhernandezp2019@alumno.ipn.mx, hcalvo@ipn.mx,
batyr1@cic.ipn.mx

Resumen. En la última década, los usuarios en redes sociales se han incrementado exponencialmente. Las redes sociales cada vez están más inmersas en la vida humana y en consecuencia cada vez tienen mayor influencia en la información de los usuarios: qué les gusta, qué les disgusta, que los motiva, etc. Twitter en la actualidad es una excelente fuente para generar un corpus robusto de las emociones de las personas y en consecuencia es una fuente muy importante para experimentar y comprobar si la información textual fortalece los modelos de series de tiempo en medida que una figura pública influye en las aplicaciones de redes sociales en mayor o menor medida. Un ejemplo reciente es el comportamiento del bitcoin a partir del interés que mostró Elon Musk como una divisa para adquirir productos tesla, otro que es el se analizara en este ejercicio es el comportamiento de las personas frente al primer bimestre cuando la Organización Mundial de la Salud declaró el inicio de pandemia el 11 de marzo de 2020. Se ha encontrado que si se representan las emociones o sentimientos reflejados en Twitter. Se puede analizar una tendencia en los sentimientos textuales en redes sociales y grupos o clusters de tipos de usuarios agrupados por emociones o sentimientos. En este trabajo se analizaron *tweets* recopilados durante el inicio de la pandemia, en el bimestre marzo y abril de año 2020. Se evaluaron los modelos de clasificación de regresión logística, *random forest*, *modelos bayesianos* y máquinas de soporte vectorial. Adicionalmente se buscó entender y observar la tendencia del comportamiento de la fluctuación de los tweets dentro de la red a partir de una serie temporal.

Palabras clave: Análisis de sentimientos, COVID-19, twitter, series de tiempo, redes sociales, análisis de comportamiento, agrupamiento por emociones.

Sentiment Analysis for the Study of Behavior in COVID-19 News on Social Media as Time Series

Abstract. In the last decade, the number of social media users has grown exponentially. These platforms are increasingly embedded in daily human life and, as a result, have a significant influence on users' information consumption — what they like, dislike, what motivates them, and more. Today, Twitter is an excellent source for generating a robust corpus of human emotions, making it a valuable resource for experimenting with and validating whether textual information can enhance time series models — especially when public figures influence how content spreads across social networks. A recent example is the behavior of Bitcoin following Elon Musk's expressed interest in accepting it as payment for Tesla products. Another example, analyzed in this work, is the public's reaction during the first two months after the World Health Organization declared the COVID-19 pandemic on March 11, 2020. This study demonstrates that emotions or sentiments expressed on Twitter can be modeled to analyze trends over time and cluster user types by sentiment. Tweets collected during March and April 2020 were analyzed. Sentiment classification models — including logistic regression, random forest, Bayesian models, and support vector machines — were evaluated. Additionally, temporal analysis was conducted to understand the behavior and fluctuation of tweets over time.

Keywords: COVID-19, twitter, time series, social media, behavior analysis, emotion clustering.

1. Introducción

Algunos trabajos de los investigadores sobre el comportamiento de las finanzas, han arrojado que el mercado puede ser influenciado por las reacciones emocionales de los agentes económicos encargados elementalmente del consumo.

Esto brinda una oportunidad estratégica a los agentes destinados exclusivamente a la producción de bienes y servicios. Pues de algunos años a nuestra actualidad, en algunas investigaciones se ha encontrado que las aplicaciones de las redes sociales son bastante rentables para predecir las acciones o la demanda de mercado, específico o general como necesidades médicas, insumos de primera necesidad . Esto dependiendo de la temporalidad, estacionalidad y tendencia del comportamiento de los demandantes de servicios y de los bienes. A partir de la percepción de las personas (*influencers*) y productos inmersos en el mercado.

Por otro lado, no solo un producto nuevo o innovador, o una inclinación preferencial de alguna figura influyente en los grupos sociales genera una tendencia en las aplicaciones de redes sociales y en el mundo real. Pues también los problemas políticos y de salud crean nuevas tendencias sociales solo que

con mayor correlación en el sector farmacéutico y de seguridad social. Es por ello que la situación por la que atravesamos frente a la pandemia declarada por la Organización Mundial de la Salud el 11 de marzo de 2020 a causa del virus SARS-Cov-2, es que en este trabajo se busca analizar algunos tipos de *tweets* recopilados durante tiempos de pandemia. Específicamente de marzo y abril, 2020.

La humanidad siempre se ha preocupado por intentar predecir el futuro. Predecir el clima ha sido muy importante una vez nos volvimos sedentarios, ya que nos ayuda a prepararnos para las cosechas en temporadas de sequía o en tiempos de abundante lluvia. La incertidumbre en la vida humana siempre ha sido un tema muy importante, tal vez tenga que ver con que somos seres con miedo y la mayoría del tiempo queremos saber que nos espera en el futuro.

Tenemos miedo de perder una economía estable, del comportamiento de fenómenos meteorológicos y/o sus consecuencias como la sequía extrema, el calentamiento global, ciclones tropicales, de la propagación de un virus nuevo y mortal como el SARS-CoV-2, etc. es decir existen muchas posibilidades de perder nuestras pertenencias incluyendo la vida por una u otra cosa. Estudiar el futuro nos ayuda interpretar que errores podemos evitar y tener un mejor futuro para la humanidad.

En la última década, la actividad en redes sociales se ha incrementado exponencialmente. A tal nivel que hoy día han superado y desbancado a la comunicación telefónica tradicional y a los medios de comunicación. Posiblemente después de la emergencia sanitaria por la que pasa el mundo en la actualidad, los viajes de negocios podrían ser sustituidos por video conferencias de manera impresionante.

Las redes sociales cada vez abarcan más áreas de interacción humana, se han creado extensiones de compra y venta de artículos, extensiones adicionales para conocer usuarios nuevos, canales de entretenimiento, con el objetivo que el usuario este mas tiempo inmerso en la red y genere nuevos datos en beneficio de los propietarios de las redes sociales. Lo anterior es la razón por la que las redes sociales poseen una gran influencia en las ciencias económicas, especialmente en el área del mercado bursátil [1], así como en otras áreas.

Para el análisis de series temporales se utilizan métodos que tienen en común el supuesto de que múltiples variables muestran dependencia mutua a lo largo del tiempo. Es decir, una variable x_1 en el momento t puede predecirse hasta cierto punto por esa misma variable en el momento $t-1$, o por otra variable x_2 en el momento $t-1$. Por ejemplo, el indicador NASDAQ-100 puede ser predictivo de sí mismo en el momento $t-1$ y el análisis de emociones junto con la ansiedad social futura en el momento t de los consumidores de Netflix, Facebook y Microsoft. A su vez, también el análisis de emociones junto con la ansiedad social futura puede afectarse a sí mismas en $t-1$ y el indicador NASDAQ-100 en el momento.

En este caso aprovecharemos de la información generada por los usuarios de Twitter. A mediados de la década pasada [4] Jiang ZHU publicaba en un artículo que las redes sociales son una excelente fuente para generar un corpus de las emociones de las personas debido al crecimiento desmesurado de usuarios

en la red. Expresaba que una nueva manera de agrupar las sociedades que se estaban formando en las redes sociales como avatar era a través del estudio de las emociones de los usuarios [2].

2. Estado del arte

Presentaremos el estado del arte considerando primero los trabajos relacionados con análisis de sentimientos, y posteriormente nos enfocaremos en las series de tiempo.

2.1. Análisis de sentimientos

El análisis de sentimientos es el proceso de analizar texto con herramientas como el aprendizaje automático para identificar y clasificar los textos de un conjunto de datos. Se utilizan diferentes fuentes de texto o corpus para realizar análisis de sentimientos. Un artículo que investiga el futuro de Twitter como herramienta para extraer metadatos a partir de la información textual publicada en Twitter, menciona que es una herramienta auxiliar considerable para pronósticos en las series de tiempo [3][4][5].

Por otro lado, Nisar también menciona que Twitter será una de las fuentes dominantes y más utilizadas para realizar análisis de sentimientos [6]. En la actualidad Twitter se considera un pilar clave de las redes sociales ya que se desempeña como un podio para que celebridades, deportistas, políticos, científicos, etc. expresen sus opiniones sobre un tema e influyan en las sociedades virtuales.

Nalini [7] analizó los sentimientos primarios que se han generado durante la pandemia. Se centró en la reacción de las emociones de usuarios de Twitter al encontrarse con noticias falsas. El análisis textual se realizó mediante la herramienta de análisis textual *Bidirectional Encoder Representations from Transformers* (BERT). En este artículo se compararon los resultados de BERT con los de otros modelos tales como regresión logística (LR), máquinas de vectores de soporte (SVM), y redes de memoria a corto y largo plazo (LSTM). Para nuestro caso de estudio se aprovecharán estos resultados para compararlos con otros modelos del estado del arte.

2.2. Series de tiempo

Una serie de tiempo es un conjunto de valores observados cronológicamente durante un intervalo temporal. Las series de tiempo nos ayudan a realizar pronósticos con el fin de orientar las decisiones en muchas áreas del mundo como los mercados, el transporte, la identificación de fallas, el clima, entre otros. Para pronosticar una variable se debe construir un modelo y estimar sus parámetros usando datos históricos, es decir, un pronóstico se logra a través de una caracterización estadística de los enlaces entre el presente y el pasado.

En nuestro trabajo de estudio se pretende buscar la existencia de correlación entre el comportamiento de indicadores bursátiles y análisis de emociones primarias de usuarios relacionados con algunas de las empresas con mayor influencia en los distintos indicadores bursátiles extraídas de los comentarios en Twitter, debido al cierto determinista del comportamiento humano. Sin embargo, el alcance mostrado en este trabajo solo converge hasta el análisis de sentimientos y experimentación de la tendencia de la fluctuación de *tweets* en la red con etiquetas relacionadas con la pandemia o el virus SARS-Cov-2.

El análisis de emociones es afectado constantemente por las preferencias de consumidores, noticias falsas en las redes sociales, experiencia del cliente y comentarios de influencias sobre algún producto o servicio, por lo que el problema de predicción de indicadores bursátiles se considerara como un modelo con variables exógenas. El análisis de series de tiempo requiere ciertas técnicas que nos permitan distinguir cuánto de esa variabilidad en el tiempo depende de determinados factores y cuál es la correlación entre variables de tal manera que podamos explicarlo dentro de nuestros modelos.

Se efectuó un análisis de las emociones representadas en series de tiempo donde se describía el comportamiento emocional del usuario a través del tiempo. El resultado de este trabajo aporó que al analizar los sentimientos textuales en redes sociales las emociones fueron cambiando al paso del tiempo [3]. . Tomando este artículo como referencia inicial, en este trabajo se pretende analizar las emociones de los usuarios de Twitter durante la pandemia generada a causa del virus SARS-Cov-2.

2.3. Diferencia de esta propuesta con respecto al estado del arte presentado anteriormente

Se ha visto en el estado del arte que el análisis de textos ha tenido buenos resultados al ser tratados con *transformers* empleando modelos BERT y modelos T5. Comparado con los resultados obtenidos en los modelos de regresión logística, *random forest*, *modelos bayesianos* y máquinas de soporte vectorial aplicados a la clasificación binaria y clasificación multiclase muestran sin duda una un área de mejoría.

En este trabajo, se realizó parte del preprocesamiento de datos como la tokenización y stemming con librerías como NLTK. Como trabajo futuro se pretende comparar los resultados del preprocesamiento de datos obtenidos con la librería Spacy y Stanford.

3. Desarrollo de la solución

3.1. Descripción de la solución al problema planteado

Inicialmente los datos se normalizan y estandarizan para reducir la complejidad de los datos y lograr un modelo más general.

Posteriormente se realizó un análisis del procesamiento del lenguaje natural, el primer ejercicio en esta etapa de experimentación fue utilizar los modelos de

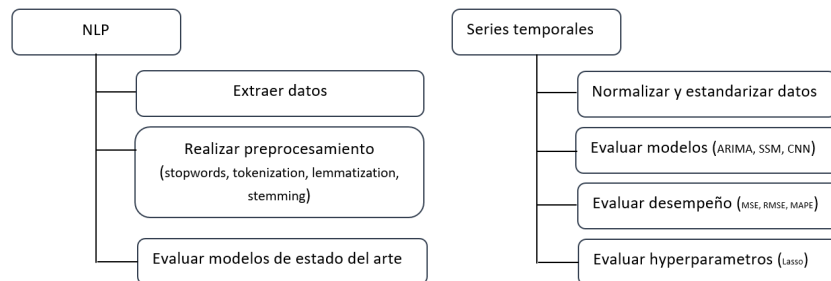


Fig. 1. Diagrama de flujo del proceso para clasificar sentimientos.

clasificación de regresión logística, *random forest*, *modelos bayesianos* y máquinas de soporte vectorial para clasificación multiclase, para realizar una clasificación binaria y experimentar con una clasificación multiclase, con esta misma familia de modelos de *machine learning*. Posteriormente se utilizó el mismo conjunto de datos para trabajar en un modelo ARIMA para predecir la tendencia de la fluctuación de *tweets* relacionados con la pandemia generada a causa de virus SARS-Cov-2.

Las métricas de desempeño utilizadas se dividirían en dos rubros. Para los modelos de procesamiento de lenguaje natural usaremos las métricas *recall*, *accuracy*, *precision* y *f1-score*. Y para el análisis de series temporales la métrica *MSE*.

En un trabajo futuro, se pretende cambiar la familia de modelos de *machine learning* por modelos relacionados con *transformers* como los modelos BERT y modelos T5. Así mismo, se pretende generar una serie temporal de la fluctuación de *tweets* a partir de etiquetas relacionadas con la pandemia y el virus SARS-Cov-2 pero con un intervalo cercano a un año. Posteriormente, se pretende buscar correlación con otras series de tiempo relacionadas con el sector financiero.

3.2. Novedad científica

Aprovechando la riqueza brindada en el trabajo del estado del arte en el que se encuentran trabajando muchos de nuestro colegas y mentores. Y en consecuencia gracias a la evaluación y aplicación de diferentes tipos de vista de AA para series de tiempo y análisis de textos nos da ya mejor percepción de nuestro estudio.

La novedad de este trabajo se centra en la búsqueda y análisis de la relación de las emociones de los usuarios de twitter durante la pandemia generada por el SARS-Cov-2. Y su correlación en entre las emociones como la tristeza, alegría, temor e ira. En la figura 1 se muestra el diagrama de flujo del proceso para clasificar sentimientos.

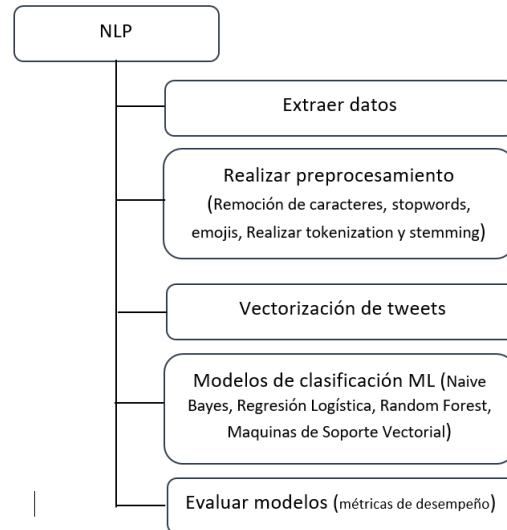


Fig. 2. Diagrama de flujo del proceso para clasificar sentimientos.

3.3. Evaluación

En este trabajo se evaluaron los experimentos realizados para la clasificación de emociones y para la predicción de series temporales. Primeramente, se evaluó el desempeño de los algoritmos implementados para la clasificación binaria y clasificación multiclase. Posteriormente, se evaluó un experimento de series temporales creado a partir del flujo de los *tweets* en la red registrados en el mismo conjunto de datos empleado para el análisis de sentimientos.

Las medidas de desempeño que fueron empleadas para los modelos de procesamiento de lenguaje natural fueron las métricas *recall*, *accuracy*, *precision* y *f1-score*. Y la métrica empleada para las series de tiempo fue *MSE*, en trabajos futuros, se pretende realizar un análisis de series temporales profundo, en el cual se considera emplear adicionalmente las medidas de desempeño *MSPE* y *MAPE*.

4. Experimentos y resultados

En este trabajo se analizaron *tweets* recopilados durante el inicio de la pandemia. Específicamente del bimestre marzo y abril de año 2020. Como bien se mencionaba con anterioridad, en esta experimentación se trabajó con modelos de clasificación para predecir el sentimiento de los *tweets* relacionados con el virus SARS-Cov-2.

Como se mencionaba en la sección introductoria, el número de casos acelerado a nivel mundial se convirtió en pánico, el miedo y la ansiedad entre la sociedad.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 41157 entries, 0 to 41156
Data columns (total 6 columns):
#   Column          Non-Null Count  Dtype
---  -
0   UserName         41157 non-null  int64
1   ScreenName       41157 non-null  int64
2   Location         32567 non-null  object
3   TweetAt         41157 non-null  object
4   OriginalTweet    41157 non-null  object
5   Sentiment        41157 non-null  object
dtypes: int64(2), object(4)
memory usage: 1.9+ MB
```

Fig. 3. Información de conjunto de datos tomada de Python.

Y uno de los medios donde se ha reglado es en las herramientas digitales de conexión social como Twitter.

Los datos utilizados para esta parte experimental se refieren a un conjunto de datos estructurado a partir de 5 variables, una etiqueta de clasificación de sentimientos y 41157 registros.

De este conjunto de datos podemos ver que la mayor participación porcentual de *tweets*, emergen de ciudades grandes como Londres, Nueva York y Los Angeles. Para esta etapa de pruebas se ha recopilado únicamente *tweets* en idioma inglés, es por ello que no aparece aún alguna ciudad de México. La recopilación de información en español, así como la información de un mayor número de meses está considerada para la siguiente etapa de esta investigación.

Para esta etapa experimental, se analizaron cinco clases de sentimientos a partir del texto contenido en los *tweets* reflejados en la etiqueta *Sentiment*. Esta etiqueta tiene cinco tipos de sentimientos como se mencionaba previamente y son los siguientes, extremadamente negativos, negativos, neutrales, positivos y extremadamente positivos.

A continuación, podemos encontrar en la siguiente gráfica de anillo la participación porcentual las ciudades a partir de la agrupación de los *tweets* a nivel global.

En la Figura 5, podemos ver la estructura del conjunto de datos utilizado en esta etapa de experimentos y resultados. Inicialmente en la columna *OriginalTweet*, se encuentra el contenido textual tal y como se recopiló en la *Tweeter*. En la parte del procesamiento de la información se normalizará el texto. Es importante mencionar que para esta experimentación los caracteres especiales y emojis fueron removidos. Sin embargo, se tiene presente que estos caracteres deben de evitar ser removidos durante la limpieza de texto para darle mayor importancia al análisis de texto en el campo de procesamiento del lenguaje natural. Debido a que no dejan de ser un lenguaje y la consecuencia de ser removidos es que con ellos se remueve información valiosa.

Distribucion de Tweets (COVID-19) por ciudades

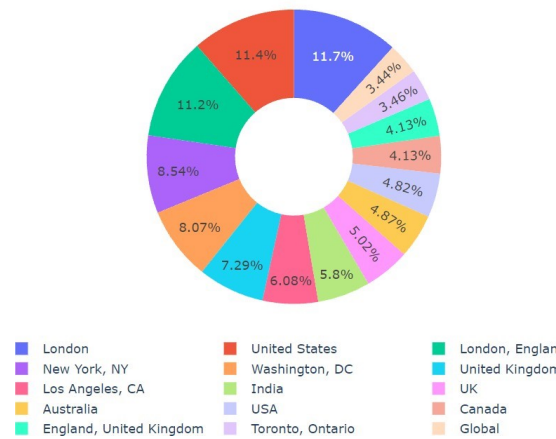


Fig. 4. Participación porcentual de tweets por ciudades.

	UserName	ScreenName	Location	TweetAt	OriginalTweet	Sentiment
0	3799	48751	London	16-03-2020	@MeNyrbie @Phil_Gahan @Chrisitv https://t.co/IFz9FAn2Pa and https://t.co/XX6ghGFzCC and https://t.co/I2NlzdXNo8	Neutral
1	3800	48752	UK	16-03-2020	advice Talk to your neighbours family to exchange phone numbers create contact list with phone numbers of neighbours schools employer chemist GP set up online shopping accounts if poss adequate su...	Positive
2	3801	48753	Vagabonds	16-03-2020	Coronavirus Australia: Woolworths to give elderly, disabled dedicated shopping hours amid COVID-19 outbreak https://t.co/blnCA9Vp8P	Positive
3	3802	48754	NaN	16-03-2020	My food stock is not the only one which is empty... PLEASE, don't panic, THERE WILL BE ENOUGH FOOD FOR EVERYONE if you do not take more than you need. Stay calm, stay safe.	Positive
4	3803	48755	NaN	16-03-2020	Me, ready to go at supermarket during the #COVID19 outbreak. Not because I'm paranoid, but because my food stock is literally empty. The #coronavirus is a serious thing, but please, don...	Extremely Negative

Fig. 5. Estructura del conjunto de datos.

4.1. Limpieza de texto

Como podemos ver, hasta ahora se ha comenzado a trabajar con una análisis y exploración de datos. A continuación, se hablara sobre la parte del preprocesamiento de los datos. En esta sección, lo primero que se realizo fue la remoción de usuarios en los *tweets*. Es decir, se removieron los @usuario del tweet.

Posteriormente se removieron las direcciones HTTP y URLs. Para este paso se utilizó la librería 're' de Python, por sus siglas en ingles de expresión regular.

4.2. Stopwords

Los siguientes pasos que se llevaron en la experimentación fue continuar con la limpieza de datos, removiendo la puntuación, los números, caracteres especiales



Fig. 7. Nube de palabras de conjunto positivo.



Fig. 8. Nube de palabras de conjunto extremadamente negativo.

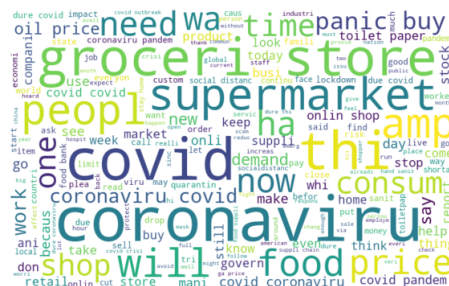


Fig. 9. Nube de palabras de conjunto negativo.

4.5. Conjunto de entrenamiento y conjunto de prueba

Posteriormente a este pequeño análisis exploratorio, se generó un conjunto de entrenamiento y un conjunto de prueba con la librería de *machine learning* scikit-learn. Para el ejercicio de clasificación multiclase. En esta parte de experimentos y resultados se evaluaron los modelos de clasificación de regresión logística, *random forest*, *modelos bayesianos* y máquinas de soporte vectorial.

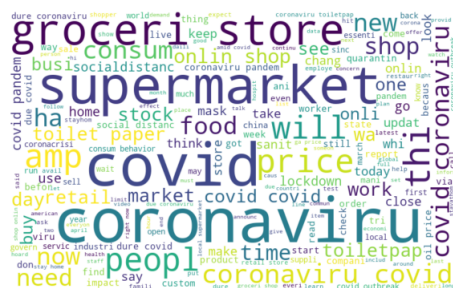


Fig. 10. Nube de palabras de conjunto neutro.

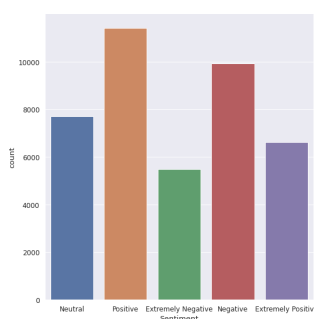


Fig. 11. Distribución de sentimientos en tweets.

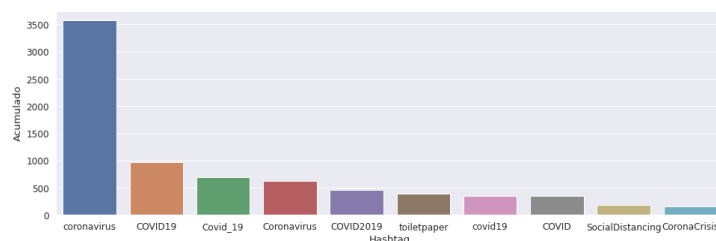


Fig. 12. Top 10 de hastag con mayor frecuencia en tweets positivos.

4.6. CountVectorizer

Posteriormente se le aplico la función de *CountVectorizer* de *scikit-learn*. La función *CountVectorizer* es un proceso de extracción de características o vectorización y es empleado para convertir documentos de texto en un vector o matriz de recuentos de tokens de cada palabra que aparece en todo el texto.

Este proceso de *CountVectorizer*, nos funcionara también para los *transformers* empleados en la siguiente etapa de la investigación. El trabajo con modelos BERT, T5 text-to-text-transfer-transformer.

A continuación, se muestra un ejemplo ilustrativo.

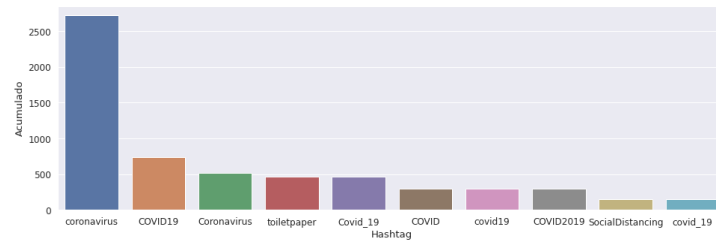


Fig. 13. Top 10 de hashtag con mayor frecuencia en tweets neutros.

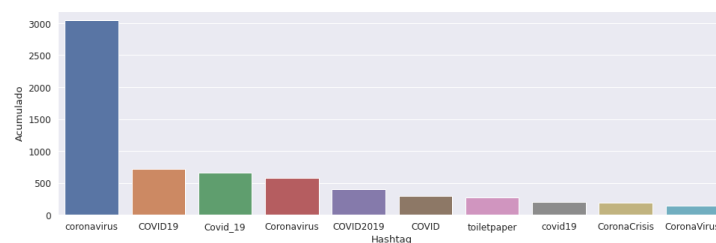


Fig. 14. Top 10 de hashtag con mayor frecuencia en tweets negativos.

	at	each	four	help	helps	many	one
document [0]	0	0	0	0	1	0	0
document [1]	0	0	1	1	0	0	0
document [2]	1	1	0	0	1	1	0

Fig. 15. Ejemplo de proceso de CountVectorizer.

El primer ejercicio en esta etapa de experimentación fue utilizar los modelos de clasificación de regresión logística, *random forest*, *modelos bayesianos* y máquinas de soporte vectorial para clasificación multiclase debido a las cinco etiquetas de nuestro conjunto de datos (extremadamente negativos, negativos, neutrales, positivos y extremadamente positivos).

Los resultados del desempeño de los indicadores *precision*, *recall* y *f1-score* en los clasificadores empleados son los siguientes:

En cuanto a los resultados obtenidos, se puede ver que indudablemente el modelo de regresión logística es el que tiene un mejor desempeño, aun cuando su desempeño es sin duda mejorable, utilizando algún método más sofisticado o un método del estado del arte.

Para este grupo de medidas de desempeño, el siguiente modelo que podemos ver que ofrece un segundo mejor resultado es el modelo de máquinas de soporte vectorial.

Por otro lado, si nos guiáramos únicamente por el desempeño calculado por *accuracy*, encontraríamos que nuevamente el modelo con mejor desempeño en

Maquinas de Soporte Vectorial				
	precision	recall	f1-score	
Extremely Negative	0.48	0.71	0.57	
Extremely Positive	0.53	0.78	0.63	
Negative	0.58	0.55	0.56	
Neutral	0.71	0.64	0.67	
Positive	0.67	0.55	0.61	

Regresión Logística				
	precision	recall	f1-score	
Extremely Negative	0.62	0.68	0.65	
Extremely Positive	0.62	0.71	0.66	
Negative	0.54	0.56	0.55	
Neutral	0.72	0.64	0.68	
Positive	0.62	0.58	0.59	

Fig. 16. Resultados de precision, recall y f1-score.

Modelo	Accuracy
Regresión Logística	0.859451
Random Forest	0.828231
Naive Bayes	0.791667
Maquinas de Soporte Vectorial	0.607264

Fig. 17. Resultados de accuracy.

este ejercicio es el modelo de regresión logística, solo que en este caso es seguido del modelo *random forest*.

Otro experimento que se realizó en este trabajo, fue emplear la misma familia de modelos de *machine learning* pero para una clasificación binaria extremadamente negativos + negativos= negativos, extremadamente positivos + positivos =positivos.

Y los resultados son los siguientes:

En estos resultados, se pudo observar que nuevamente el modelo de regresión logística fue la que tuvo el mejor desempeño, sin embargo, no estuvo muy alejado del resto de los modelos. Por otro lado, podemos observar que el desempeño mejoro ligeramente en los clasificadores binarios contra los clasificadores multiclase.

El orden de los resultados en el *accuracy* para los clasificadores binarios fue similar al de los clasificadores multiclase.

Naive Bayes

	precision	recall	f1-score
0	0.69	0.74	0.71
1	0.85	0.82	0.84

Regresión Logística

	precision	recall	f1-score
0	0.77	0.84	0.80
1	0.92	0.87	0.89

Fig. 18. Resultados de precision, recall y f1-score.

Random Forest

	precision	recall	f1-score
0	0.70	0.82	0.75
1	0.91	0.83	0.87

Maquinas de Soporte Vectorial

	precision	recall	f1-score
0	0.69	0.87	0.77
1	0.94	0.84	0.88

Fig. 19. Resultados de precision, recall y f1-score.

	Modelo	Accuracy
	Regresión Logística	0.859451
	Maquinas de Soporte Vectorial	0.845603
	Random Forest	0.828474
	Naive Bayes	0.791667

Fig. 20. Resultados de Accuracy.

Si bien aún no se han realizado experimentos con datasets relacionados directamente con series temporales como el dataset HistoricalData, el cual contiene información del indicador bursátil NASDAQ-100.

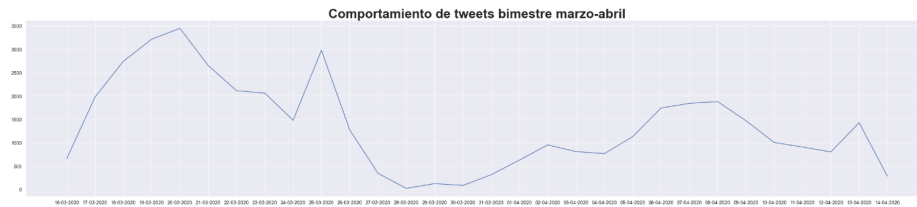


Fig. 21. Comportamiento de tweets marzo-abril, 2020.

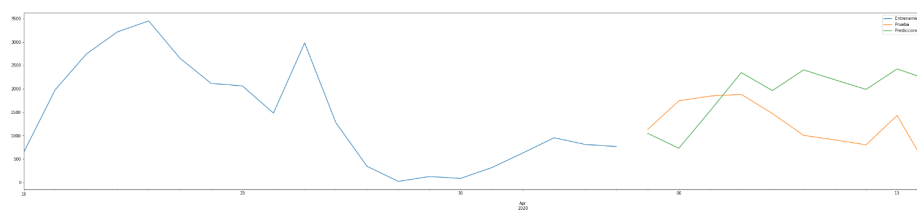


Fig. 22. Predicción de tweets marzo-abril, 2020.



Fig. 23. Comportamiento de tweets positivos marzo-abril, 2020.

En este trabajo se realizaron experimentos iniciales, en el cual se buscó entender y observar la tendencia del comportamiento de la fluctuación de los *tweets* dentro de la red. Para la predicción del modelo ARIMA el error fue grande MSE 1121060.09. En trabajos futuros se pretende buscar mejorar este error a partir de enriquecer el conjunto de datos y probar con modelos *CNN*, como se mencionaba con anterioridad.

Inicialmente, se representó gráficamente el comportamiento general del tráfico de *tweets* en la red. Es decir, se consideraron las cinco etiquetas como una sola, para tener un agrupamiento global de la presencia de *tweets*. Distinguiéndolos a partir de la fecha de emisión en un formato de año-mes-día.

A continuación, se muestra la tendencia de la fluctuación de *tweets* en este bimestre marzo-abril, 2020. Se pudo observar que en consecuencia de que los *tweets* fueran mayoritariamente positivos y negativos. La tendencias global se vio marcada con un comportamiento muy similar.

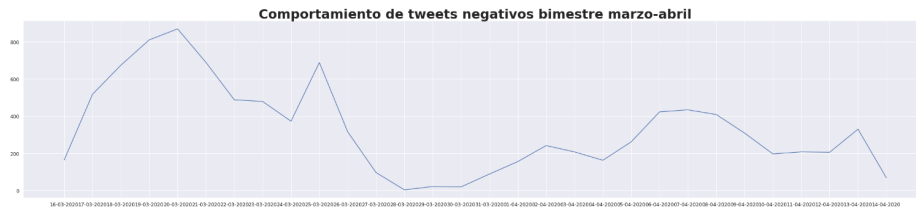


Fig. 24. Comportamiento de tweets negativos marzo-abril, 2020.

5. Conclusiones y trabajo futuro

En este trabajo se ha propuesto una solución posible para identificar sentimientos en *tweets* relacionados con la enfermedad de la pandemia causada por el virus SARS-CoV-2. Se ha revisado el estado del arte, y se han identificado diversas soluciones para resolver esta tarea. Presentamos resultados utilizando modelos bayesianos, regresión logística, *random forest*, y máquinas de soporte vectorial, así como una propuesta de utilización para el estudio de comportamiento en noticias en redes sociales como series de tiempo. El desempeño es similar al obtenido en el estado del arte, aunque existe todavía margen de mejora. Como trabajo futuro, hemos propuesto utilizar BERT y modelos T5 para obtener un mejor desempeño, para la parte de procesamiento de lenguaje natural. Para la parte de las series temporales, se pretende utilizar modelos CNN y comparar los resultados con otros modelos como ARIMA y SSM. En el experimento de tratar los datos de la fluctuación de *tweets* en la red se pudo encontrar que los datos empleados para la etapa experimental, deben ser más grande. Lo que nos llevara a abordar otro reto asociado con el *big data*.

Agradecimientos. Este trabajo fue apoyado por el CONACyT, COFAA y el Instituto Politécnico Nacional, IPN-SIP gracias a los proyectos SIP 2083 y SIP 20210189. J.H.P.: Agradezco la atención y el apoyo de mis asesores, por esta oportunidad. Así mismo les agradezco por incentivar la curiosidad y el desarrollo científico constante que me han brindado.

Referencias

1. Chen, C. Y.-H., Härdle, W. K., Klochov, Y.: SONIC: SOcial Network analysis with Influencers and Communities. *Journal of Econometrics*, (2021)
2. Li, J., Yang, G.: Network embedding enhanced intelligent recommendation for online social networks. *Future Generation Computer Systems*, vol. 119 (2021)
3. Sul, H. K., Dennis, A. R., Yuan, L.: Trading on twitter: Using social media sentiment to predict stock returns., (2017)
4. Zhu, J., Wang, B., Wu, B.: Social network users clustering based on multivariate time series of emotional behavior. *The Journal of China Universities of Posts and Telecommunications*, vol. 2014 (2014)

Estudios sobre el reconocimiento del habla imaginada mediante señales EEG

Arturo Iván Sandoval Rodríguez, Francisco Hiram Calvo Castro

Instituto Politécnico Nacional,
Centro de Investigación en Computación,
Laboratorio de Ciencias Cognitivas Computacionales,
México

asandovalr2021@cic.ipn.mx, fcalvo@ipn.mx

Resumen. El habla imaginada es un área con una gran cantidad de posibilidades para mejorar la calidad de vida de las personas que han sufrido de accidentes que les impidan comunicarse mediante el habla o que no puedan realizar actividad motriz. Es por eso que este tema de estudio resulta interesante, mediante el estudio del cerebro, más específicamente, señales electroencefalográficas (EEG) posibilitar una nueva forma de comunicación mediante una interfaz cerebro-computadora. En el presente trabajo se realiza un estudio del estado del arte, así como una comparativa de trabajos recientes. Finalmente se concluye con una propuesta de que es lo que se puede innovar teniendo en cuenta el estudio del estado del arte realizado.

Palabras clave: Habla imaginada, señales EEG, interfaz cerebro-computadora, comunicación asistida, estado del arte.

Studies on Imagined Speech Recognition Using EEG Signals

Abstract. Imagined speech is a field with great potential to improve the quality of life for individuals who have suffered accidents that prevent them from communicating through speech or performing motor activities. For this reason, the topic is of particular interest, as it explores the brain — more specifically, electroencephalographic (EEG) signals — to enable a new form of communication through a brain-computer interface (BCI). This work presents a study of the state of the art in imagined speech recognition, along with a comparison of recent research. Finally, it concludes with a proposal on potential innovations based on the insights gained from the state-of-the-art analysis.

Keywords: Imagined speech, EEG signals, brain-computer interface, assistive communication, state of the art.

1. Introducción

Para comenzar es necesario dar una definición de lo que es el habla imaginada, en palabras de D'zmura et al., (2009) corresponde a imaginar la pronunciación del habla en ausencia de salida articulatoria y acústica. Ahora bien, dentro del mundo de las ciencias cognitivas, el cerebro es el principal objeto de estudio, aunque debido a la complejidad de su entendimiento, a lo largo de los años se han ido desarrollado herramientas para tener un mejor entendimiento de este. Una de las más utilizadas son las Interfaces Cerebro-Computadora o BCI (por sus siglas en inglés, Brain Computer Interface), que en palabras de Valerdi et al., (2019) son sistemas que miden la actividad del Sistema Nervioso Central y la convierten en salidas que reemplazan, restauran, aumentan, suplementan o mejoran las salidas naturales de dicho sistema y, por lo tanto, cambian las interacciones en curso entre el ser humano y su ambiente interno o externo.

En otros estudios se define como un sistema informático que adquiere señales cerebrales, las analiza y las traduce en comandos que se transmiten a un dispositivo de salida para llevar a cabo una acción deseada [2].

Estas interfaces o bien sistemas están basadas en señales electroencefalográficas (EEG), las cuales resultan ser una excelente opción de estudio pues no son invasivas, son relativamente simples, económicas e insensibles a ambientes con grandes cantidades de ruido audible [3].

El EEG estándar es una exploración indolora, no invasiva, de bajo coste, que puede ser de gran utilidad en la práctica clínica. Se realiza colocando electrodos de superficie adheridos al cuero cabelludo por un gel conductor. Se posicionan de acuerdo con el sistema internacional 10-20 [4, 5].

Cada canal de registro mide la diferencia de voltaje entre dos electrodos (uno es el activo y otro el de referencia). Lo habitual es que se usen de 16 a 24 derivaciones en cada montaje. Los distintos pares de electrodos se combinan constituyendo los montajes. Hay dos tipos básicos de montajes: bipolar (transversal y longitudinal) y monopolar (o referencial). [7] El bipolar registra la diferencia de voltaje entre dos electrodos colocados en áreas de actividad cerebral, mientras que el monopolar registra la diferencia de potencial entre un electrodo ubicado en una zona cerebral activa y otro colocado sobre un área sin actividad o neutra [7].

2. Estado del arte

Los trabajos que se encontraron donde se hace hincapié en el habla imaginada se detallan a continuación:

2.1. Decoding imagined, heard, and spoken speech: classification and regression of EEG using a 14 channel dry-contact mobile headset

(Clayton et al., 2020) [8]: En este trabajo los autores hacen énfasis en que los dispositivos de tipo científico de recolección de datos EEG resultan ineficientes

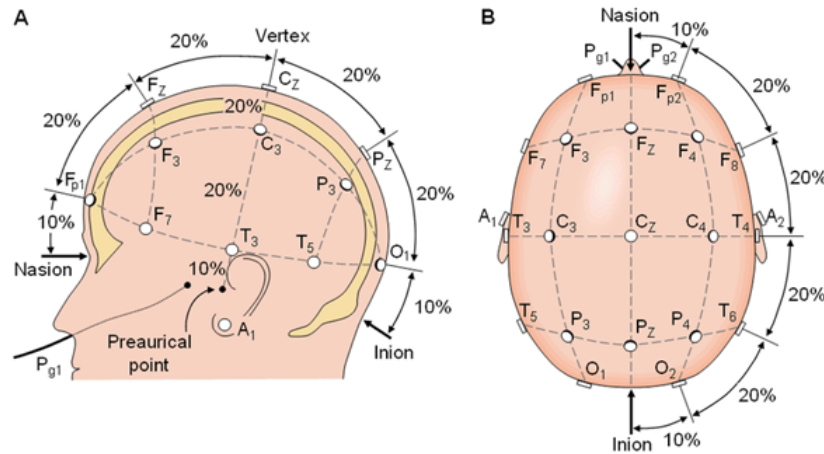


Fig. 1. Sistema Internacional 10-20 para la colocación de los electrodos extracraneales. Las letras señalan el área (Fp, prefrontal; F, frontal; C, central; P, parietal; T, temporal y O, occipital), mientras que los números designan el hemisferio (pares del derecho, nones del izquierdo) y los electrodos de la línea media se señalan con una "z"; por lo que Fz se encuentra frontalmente en la línea media [6].

en los escenarios cotidianos, por esa razón optaron por usar un dispositivo de tipo comercial que, si bien ofrece una menor fidelidad con menos electrodos, resulta ser más práctico para el día a día.

Sus objetivos primarios son:

- Evaluar la efectividad del dispositivo "ligero" para la decodificación del habla, al comparar el rendimiento con el de dispositivos de grado investigativo.
- Investigar diferentes modelos de machine learning para la clasificación del habla y tareas de regresión con EEG.
- Presentar un nuevo corpus o bien base de datos, con licencia de grabaciones de EEG evocadas por voz además de resultados y código.

La base de datos presentada fue llamada FEIS (por sus siglas en inglés Fourteen-channel EEG for Imagined Speech) y contiene grabaciones de 21 participantes de habla inglesa, estas grabaciones se realizaron con un dispositivo portátil, de 14 canales con electrodos secos (el Emotiv EPOC+) [9].

Las grabaciones EEG constan en 160 ensayos, los cuales comprenden 6 fonemas en 10 repeticiones, aleatorias para mantener la atención del participante. Cada prueba tiene cuatro "épocas" de cinco segundos. Este procedimiento se ilustra en la figura 2.

Primero se realiza una época de "descanso", en la cual a los participantes se les muestra en el monitor la palabra "REST" y tratan de despejar su mente, esto también reduce la carga cognitiva. Después viene una época de "estímulo" en la cual a los participantes se les reproduce su propia grabación del fonema

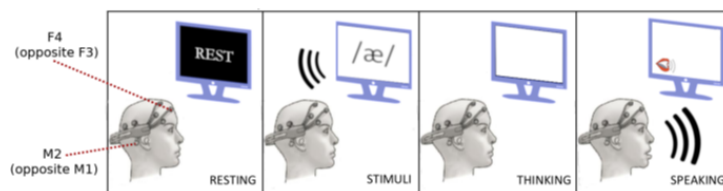


Fig. 2. Los participantes escuchan cinco veces repetidas un fonema (el cual es su propia voz grabada previamente), luego se imaginan hablando este fonema cinco veces (con el mismo ritmo), y después hablan abiertamente el fonema cinco veces [8].

Tabla 1. Resultados de los modelos empleados.

Modelo	Porcentaje de precisión (Desviación estandar)		
	Escuchar	Pensar	Hablar
SVM	64.0(16.5)	69.0(13.2)	63.7(21.2)
CNN	51.2(5.7)	49.0(6.1)	49.4(6.3)

en bucle cinco veces, y se les muestra en el monitor una representación IPA del fonema. Lo siguiente es una época de “pensamiento”, la cual consiste en mostrar a los participantes el monitor en pantalla negra y que se imaginen repitiendo el fonema, pero sin realizar ningún movimiento. Finalmente, la época del “habla”, donde a los participantes se les muestra una imagen de una boca para que posteriormente vocalicen el fonema.

Para el modelado se realizó una división de los datos del 80/10/10, para el entrenamiento, prueba y validación respectivamente. Para la máquina de vectores de Soporte (por sus siglas en inglés SVM, Support Vector Machine) se usó un 80/20, entrenamiento/prueba con validación cruzada de cinco iteraciones. Para cada tipo de habla, escuchada, imaginada y hablada, los datos fueron divididos en épocas de cinco segundos. Las características son extraídas de estas divisiones y son usadas para el entrenamiento del clasificador usando las etiquetas correspondientes al tipo de fonema, y un modelo de regresión usando las características extraídas del estímulo de audio que fue escuchado, imaginado y hablado. Para la clasificación se probaron dos tipos de modelo, máquina de vectores de soporte(SVM) y las redes neuronales convolucionales (por sus siglas en inglés, Convolutional Neural Networks).

Los resultados más destacables de este trabajo se muestran en la tabla 1. En esta se puede observar que el modelo SVM tuvo mejor rendimiento, sin embargo, los mismos autores mencionan que es posible obtener mejores resultados con las CNN aunque esto requeriría de mayor investigación.

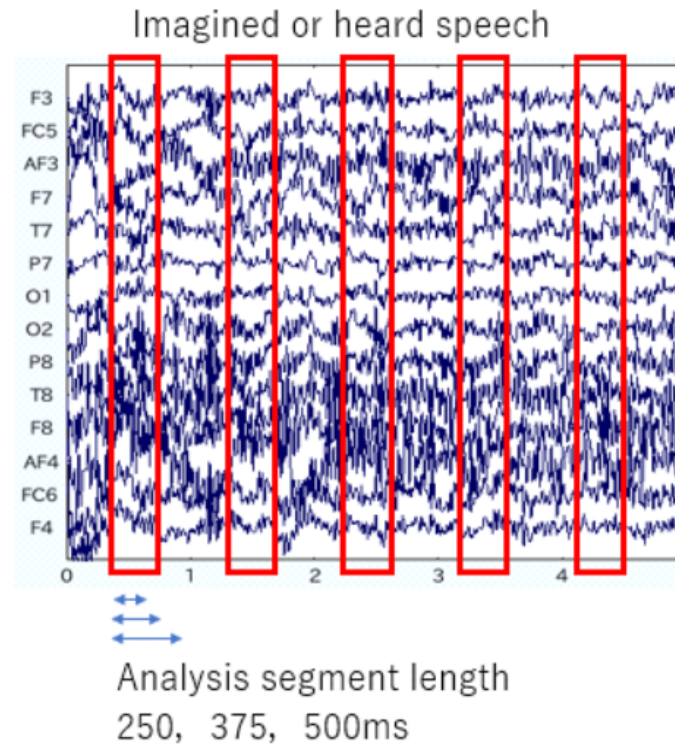


Fig. 3. Extracción de características de señal EEG.

2.2. Classification of Imagined and Heard Speech Using Amplitude Spectrum and Relative Phase of EEG

(Sakai et al., 2021) [10]: En esta investigación precisamente se retoma el trabajo de Clayton et al., pues principalmente para su estudio se hace uso de la base de datos FEIS, sin embargo, este trabajo se enfoca en la amplitud del espectro y la fase relativa de las señales EEG. Para aumentar el número de muestras, dividen la información en cinco partes (de un segundo cada uno) y es tratada cada parte como información separada. Con el fin de investigar en qué periodo de tiempo las características aparecen, los experimentos fueron llevados a cabo con tres longitudes de segmento de análisis de 250, 375 y 500 ms desde el inicio de cada estímulo o pensamiento repetido, esto es ilustrado en la figura 3.

Como clasificador, se usó una máquina de vectores de soporte (SVM) y redes neuronales (NN). Por una parte, la SVM realiza una validación cruzada de cinco iteraciones en la información preprocesada de cada participante con un kernel gaussiano. Por otro lado, la red neuronal usada tiene una capa oculta con el mismo número de nodos que de dimensiones de entrada. Se aplicó ReLU como función de activación, la función sigmoide se aplicó en la capa final, entropía cruzada como la función de pérdida, y Adam como función de optimización. Se

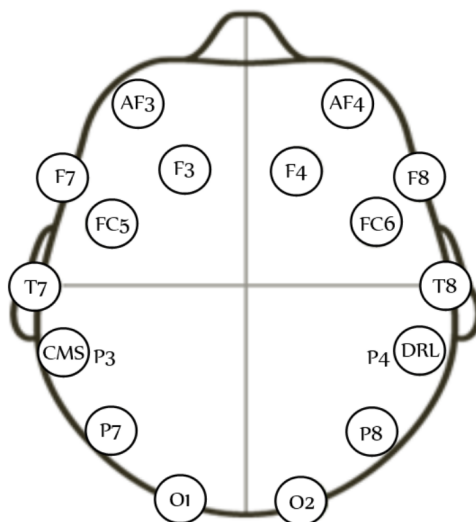


Fig. 4. Localización de los electrodos en el kit Emotiv [11].

dividió la información de la siguiente manera, 80 % para entrenamiento, 10 % para validación y 10 % para pruebas. Finalmente se obtuvo una mejor precisión usando tanto el espectro de amplitud como la fase relativa, lo que indica la efectividad de ambas en el reconocimiento del habla imaginada. Respecto a los modelos, las redes neuronales presentan una mejor clasificación, por otro lado, los experimentos con SVM no mejoraron los resultados de trabajos previos.

2.3. Análisis de señales electroencefalográficas para la clasificación de habla imaginada

(Torres et al., 2012)[3] Este trabajo se enfoca en la identificación de cinco palabras a través del habla imaginada. Para la metodología que se presenta está dividida en cinco etapas, las cuales son: Adquisición de la actividad cerebral, Preprocesamiento, Extracción de características, Selección de características, y Clasificación. Para la parte de adquisición de datos se usó el kit EMOTIV, el cual tiene catorce canales de alta resolución. Para la adquisición de datos se tiene un mayor interés a los canales F7, FC5, T7, P7(ver fig. 4) puesto a que están más cercanas a las áreas del modelo Geschwind-Wernicke, lugar donde las señales EEG están más relacionadas a la producción del habla.

Además, se usó un pequeño protocolo para delimitar el inicio y fin de una palabra imaginada. Un conjunto de muestras entre ambos etiquetados (inicio y fin) es llamado ventana o instancia. El preprocesamiento consiste en filtrar la información obtenida de los canales previamente mencionados mediante un filtro FIR pasa bandas en el rango de 4 a 25 Hz. Cabe mencionar que la duración de cada ventana del habla imaginada es variable, tanto para un sujeto como para diferentes sujetos, por lo que resulta innecesario establecer un tamaño

específico para todas las ventanas. En este punto, las ventanas con 256 muestras y una frecuencia entre 4 a 25 Hz se mantienen para el uso de creación de la información experimental. Las ventanas con un menor número de muestras a 256 son completadas con ceros y aquellas que sobrepasen las 256 muestras son descartadas.

Para la extracción de características se usa la Transformada Wavelet Discreta, su ecuación está definida por [12]:

$$W(j, k) = \sum_j \sum_k f(x) 2^{-j/2} \psi(2^{-j}x - k). \quad (1)$$

La DWT (por sus siglas en inglés, Discrete Wavelet Transform) provee una representación wavelet altamente eficiente mediante la restricción de la variación en la traslación y la escala, usualmente a potencias de dos. Por lo que se aplica la DWT con seis niveles de descomposición, usando como wavelet madre a una Daubechies de segundo orden(db2). Se obtiene un vector con 269 coeficientes wavelet por cada ventana en cada canal de interés. Por lo tanto, los coeficientes en el mismo intervalo de tiempo que pertenecen a los cuatro canales son concatenados en el siguiente orden F7-FC5-T7-P7. En este punto se obtiene un vector con 1076 características y su etiqueta de clase es obtenida.

Para la selección de características, los subconjuntos que sean mayores a 25 Hz son descartados, después el subconjunto seleccionado consiste en los coeficientes D2 al D6 y a la aproximación A6(estos coeficientes obtenidos previamente de la DWT) lo cual reduce el tamaño del vector de características y al mismo tiempo reduce el impacto del problema de la dimensionalidad. Con esto, cada ventana de cada canal es representado con 140 coeficientes wavelet. Después, los coeficientes DWT de ventanas en el mismo intervalo de tiempo son concatenadas como parte de la etapa de extracción de características.

Finalmente tres tipos de clasificadores fueron entrenados y probados: Maquinas de vectores de soporte (SVM), Random Forest (RF) y Naive Bayes(NB). La precisión de la clasificación es obtenida a través de una validación cruzada de 10 iteraciones, los resultados preliminares(realizados a los datos de tres participantes) se muestran a en la tabla 2.

Tabla 2. Porcentaje de precisión de cada uno de los tres clasificadores utilizando los vectores de características reducido.

Participante	NB	RF	SVM
S1	23.35	24.08	23.35
S2	17.09	31.63	24.78
S3	35.75	41.21	18.18

Con esto se procedió a usar a los clasificadores para entrenamiento y prueba en los 21 participantes, y como en los resultados preliminares se obtuvo mayor

precisión con RF, se experimentó con está teniendo aún mejores resultados bajo el nombre de "Bagging-RF".

3. Conclusiones y trabajo futuro

A pesar de que hay trabajos previos sobre el habla imaginada, esta sigue siendo un mundo poco explorado, por lo que siguen existiendo muchísimas posibilidades por crear e innovar. Cabe mencionar que los trabajos previos tienen su alta relevancia y podrían funcionar como punto de inicio para futuras investigaciones. La selección de los trabajos para esta investigación se debe a que son estudios recientes y que tienen su importancia en este campo, la identificación del habla imaginada, además se cuenta con su base de datos por lo que se plantea en primera instancia replicar su trabajo para posteriormente proponer un mayor desarrollo. Con esta premisa se propone realizar un estudio donde con un conjunto de sílabas, separadas por pocos segundos, se empiecen a identificar palabras sencillas, claro esta con su respectiva limitación a cierto número de sílabas y las posibles combinaciones que resulten para las palabras a identificar.

Agradecimientos. Este trabajo fue apoyado por el CONACyT, COFAA y el Instituto Politécnico Nacional, IPN-EDI, IPN-SIP gracias a los proyectos SIP 2083 y 20210189.

Referencias

1. D'Zmura, M., Deng, S., Lappas, T., Thorpe, S., Srinivasan, R.: Toward EEG sensing of imagined speech. In: Human-Computer Interaction. New Trends, ed J. A. Jacko (Berlin; Heidelberg: Springer Berlin Heidelberg), 40–48 (2009)
2. Alonso-Valerdi, L. M., Arreola-Villarruel, M. A., Argüello-García, J.: Interfaces Cerebro-Computadora: Conceptualización, Retos de Rediseño e Impacto Social. Mexican Journal of Biomedical Engineering, 40(3), 1–18 (2019)
3. Jerry J. Shih, Dean J. Krusienski, Jonathan R. Wolpaw: Brain-Computer Interfaces in Medicine. Mayo Clinic Proceedings (2012)
4. Torres-García, A., Reyes-García, C. A., Villaseñor-Pineda, L.: Toward a silent speech interface based on unspoken speech. In: BIOSIGNALS 2012 - Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing (2012)
5. Homan, R.W., Herman, J., Purdy, P.: Cerebral location of international 10-20 system electrode placement. Electroencephalogr Clin Neurophysiol, 66, 376–382 (1987)
6. Myslobodsky, M.S., Coppola, R., Bar-Ziv, J., Weinberger, D.R.: Adequacy of the International 10-20 electrode system for computed neurophysiologic topography. J Clin Neurophysiol 7, 507–518 (1990)
7. Novo-Olivas, C., Guitiérrez, L., Bribiesca, J.: Mapeo Electroencefalográfico y Neurofeedback. (2010)

8. F. Ramos-Argüelles, G. Morales, S. Egozcue, R.M. Pabón, M.T. Alonso: Técnicas básicas de electroencefalografía: principios y aplicaciones clínicas, *An. Sist. Sanit. Navar.* 32 (Supl. 3): 69–82 (2009)
9. Clayton, J., Wellington, S., Valentini-Botinhao, C., Watts, O.: Decoding imagined, heard, and spoken speech: Classification and regression of EEG using a 14-channel dry-contact mobile headset. In: *Proceedings Interspeech 2020* (Vol. 2020-October, pp. 4886-4890). International Speech Communication Association (2020) <https://doi.org/10.21437/Interspeech.2020-2745>
10. Emotiv EPOC+. [Online]. Available: <https://www.emotiv.com/epoc/>
11. R. Sakai, A. Kai. S. Nakagawa: Classification of Imagined and Heard Speech Using Amplitude Spectrum and Relative Phase of EEG. In: *IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech)*, pp. 373-375 (2021) doi: 10.1109/LifeTech52111.2021.9391883.
12. Torres-García, A.A., Reyes-García, C.A., Villaseñor-Pineda. L., et al.: Análisis de Señales Electroencefalográficas para la Clasificación de Habla Imaginada. *Rev Mex Ing Biomed.* 34(1):23–39 (2013)
13. Priestley M.: Wavelets and timedependent spectral analysis. *Journal of Time Series Analysis*, vol. 17, no. 1, pp. 85–103 (2008)

Algoritmo de cifrado de imágenes utilizando autómatas celulares reversibles

Erendira Corona-Bermúdez, Juan Carlos Chimal-Eguía,
Uriel Corona-Bermúdez

Instituto Politécnico Nacional,
Centro de Investigación en Computación,
México

ecoronab2020@cic.ipn.mx

Resumen. Durante años se han desarrollado esquemas de cifrado y algoritmos, que han ayudado al ser humano a preservar la información compartida; además, durante la pandemia el uso del internet como medio de comunicación ha incrementado significativamente. Por ese motivo, y otros, la criptografía juega un rol importante en la actualidad. En este artículo se propone un esquema de cifrado y descifrado de imágenes, el cual se desarrolla con autómatas celulares reversibles como salida se obtienen imágenes, tales que, visualmente no revelan información acerca de la imagen original. Todo el proceso se realiza durante un tiempo de ejecución de segundos.

Palabras clave: Autómatas celulares reversibles (AC), cifrado de imágenes, criptografía, esquema criptográfico.

Image Encryption Algorithm Using Reversible Cellular Automata

Abstract. For years, encryption schemes and algorithms have been developed to help preserve shared information. Moreover, during the pandemic, the use of the internet as a means of communication increased significantly. For this reason, among others, cryptography plays an important role today. This article proposes an image encryption and decryption scheme developed using reversible cellular automata. As output, the process generates images that, visually, do not reveal any information about the original image. The entire process is executed within a matter of seconds.

Keywords: Reversible cellular automata (RCA), image encryption, cryptography, cryptographic scheme.

1. Introducción

La seguridad de la información es necesaria para los individuos. La mayoría de los datos son compartidos de forma digital, además, son almacenados y

procesados en la nube, por esa razón la criptografía juega un papel importante en la actualidad. Existen métodos y herramientas diferentes para desarrollar criptografía, entre ellos se encuentran los autómatas celulares que son un tipo de sistema dinámico, los cuales se han utilizado de manera eficaz para la construcción de criptosistemas robustos aprovechando sus propiedades [1].

Jun [2] presentó un esquema de cifrado / descifrado de imágenes. Se investigó el comportamiento de varios autómatas celulares elementales, como resultado obtuvieron buenas propiedades de confusión y difusión en su esquema. Del mismo modo en [3] se propuso una nueva estructura para el cifrado de imágenes utilizando Ácido Desoxirribonucleico (ADN) y autómatas celulares recursivos (RCA). El cifrado de imágenes se realizó en dos fases independientes las cuales son permutación y difusión.

Dos años después, [4] propuso un esquema de cifrado de imágenes basado en autómatas celulares unidimensionales reversibles, el cuál fue completamente paralelizable ya que las tareas de cifrado / descifrado se ejecutaban utilizando múltiples procesos independientes para la misma imagen única.

Ping, P. et al., [5] en 2016 presentaron un cifrador de imágenes basado en el caos y los autómatas celulares (AC) similares a las reglas de *life*. El cifrado de imagen propuesto está integrado de dos subprocesos: permutación y sustitución. Similarmente Abdo, A. et al., [6] propusieron un algoritmo con AC elementales, periódicos, y con atractores unitarios; el número de estados atractores de autómatas celulares cambiaba con respecto a la imagen cifrada, y se ocuparon diferentes claves para cifrar diferentes imágenes simples.

Debido a que la transmisión de datos de sensores a través del canal de comunicación inalámbrica juega un papel importante, Satyabrata et. al. [7] presentaron una técnica de criptografía de clave simétrica de cifrado de bloques utilizando reglas de autómatas celulares aplicadas a datos de sensores en WSN. De forma semejante los mismos autores [8], mostraron una nueva técnica de cifrado de imágenes segura y eficiente, usando un autómata celular de dos dimensiones usando la vecindad de Moore, una ventaja de tal técnica es que no tiene pérdidas y es fácil de implementar en hardware.

Finalmente, en [9] se introdujo un esquema de cifrado de imágenes de alta seguridad para la comunicación y el almacenamiento de estas. La imagen permutada se cifra en función de un único número aleatorio generado por un *tent map* sesgado.

2. Metodología

En esta sección se describen conceptos que nos ayudarán en el desarrollo del esquema criptográfico, además, se describirá el modelo propuesto.

2.1. Preliminares

Autómatas Celulares Los autómatas celulares (AC) son una clase de sistemas matemáticos espacialmente y temporalmente discretos caracterizados

por la interacción local y una forma de evolución paralela [10].

Los modelos de AC utilizados usualmente poseen 5 características:

- **Teselado discreto de células:** El sistema consiste en un teselado de células de una, dos, o hasta d-dimensiones.
- **Homogeneidad:** Todas las células son equivalentes.
- **Estados discretos:** Cada célula toma uno de un número finitos de posibles estados discretos.
- **Interacción Local:** Cada célula interactúa solo con células que están en su vecindario local.
- **Dinámica discreta:** A cada unidad discreta de tiempo, cada célula actualiza su estado presente de acuerdo a una regla de transición, tomando en cuenta los estados de las células en su vecindario.

Autómatas celulares reversibles Para describir los autómatas celulares reversibles (ACR) también llamados automatas celulares invertibles, primero se introduce el concepto de reversibilidad y la relación con estos.

Un AC es inyectivo si cada configuración tiene como máximo un elemento del dominio, y es sobreyectiva si cada configuración tiene al menos un elemento del dominio, si el mapa es tanto inyectivo como sobreyectivo, es biyectivo (un mapa biyectivo tiene una función inversa). Llamamos reversible a un AC biyectivo si también la función inversa es un autómata celular [11,12].

Cifradores

Los cifradores simétricos se pueden dividir en dos: cifradores de flujo y cifradores de bloque [13].

Cifradores de flujo: Cifran bits individualmente. Esto se logra agregando un bit de una secuencia de claves a un bit de texto plano.

Cifradores de bloque: Cifra un bloque completo de bits de texto plano a la vez con la misma clave. Esto significa que el cifrado de cualquier bit de texto plano en un bloque determinado depende de cada otro bit de texto plano en el mismo bloque.

Confusión y difusión Existen operaciones primitivas que se pueden aplicar para lograr un cifrado seguro según Claude Shannon [13]. Son dos operaciones primitivas con las que se pueden construir algoritmos de cifrado fuertes:

1. **Confusión:** Es una operación de cifrado donde la relación entre la clave y el texto cifrado se oscurece.
2. **Difusión:** Es una operación de cifrado donde la influencia de un símbolo de texto sin formato se extiende sobre muchos símbolos de texto cifrado con el objetivo de ocultar las propiedades estadísticas del texto sin formato.

Regla 232								
0	0	0	0	0	0	0	0	Valor del bit en bloque de imagen
0	0	0	0	1	1	1	1	Valor del bit en bloque de llave (i-1)
0	0	1	1	0	0	1	1	Valor del bit en bloque de llave (i)
0	1	0	1	0	1	0	1	Valor del bit en bloque de llave (i+1)
1	1	1	0	1	0	0	0	Valor del bit en bloque de salida

Regla 23								
1	1	1	1	1	1	1	1	Valor del bit en bloque de imagen
0	0	0	0	1	1	1	1	Valor del bit en bloque de llave (i-1)
0	0	1	1	0	0	1	1	Valor del bit en bloque de llave (i)
0	1	0	1	0	1	0	1	Valor del bit en bloque de llave (i+1)
0	0	0	1	0	1	1	1	Valor del bit en bloque de salida

Fig. 1. Reglas de los AC y modo en que evolucionan.

2.2. Motivación y objetivos

Con el tiempo, se han creado algunos protocolos, esquemas y mecanismos criptográficos para preservar nuestra información; pero a medida que avanza la criptografía, también lo hace el criptoanálisis. Dado que el uso del internet y la comunicación por el ha incrementado, es fundamental cuidar nuestros datos, con los diferentes servicios criptográficos. Este trabajo tiene como objetivo diseñar un esquema de cifrado de imágenes, como entrada se tendrá una imagen y como salida se busca obtener otra imagen que no muestre información visual relacionada con la imagen original todo este proceso usando autómatas celulares reversibles.

2.3. Modelo de cifrado-descifrado usando autómatas celulares reversibles

Se propone un modelo de cifrado - descifrado de imágenes utilizando dos AC uno-dimensionales. Para las reglas de evolución, en el primer AC, se toma en cuenta el valor de los vecinos en radio uno de la llave; para el segundo se toma en cuenta el valor que tiene, en la misma posición que la célula que se está evolucionando, la imagen. A partir de ambos datos se toma una decisión del nuevo valor que tendrá el bit de la imagen cifrada. Lo anterior es explicado y colocado como una tabla de verdad en la Figura 1.

La salida que se obtiene del modelo es otra imagen, tal que, no tiene relación visual con la imagen de entrada, o su relación sea mínima. El flujo del modelo es presentado en un diagrama de bloques en la Figura 2.

Los AC unidimensionales ocupados en este algoritmo están definidos como una 4-tupla (S, N, f, d) , donde:
 $S : \{0, 1\}$

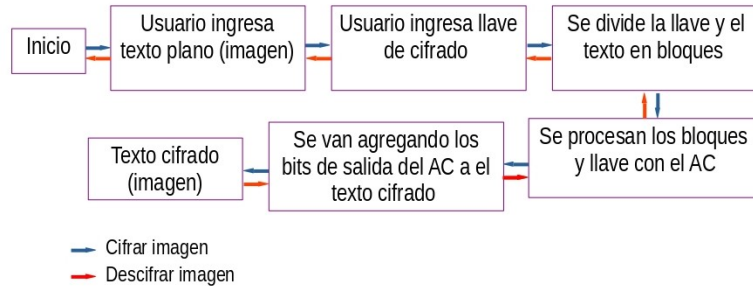


Fig. 2. Diagrama de bloque para explicar el flujo del modelo.

N : Es considerado radio 1.

$f: S^n \rightarrow S$: Regla 232 y 23.

$d \in Z^+$: Uno dimensional.

Dentro de las 256 reglas de Wolfram para los autómatas celulares unidimensionales, solo existen seis que son reversibles; se eligen las reglas 232 y 23 de estas porque tienen la propiedad de ser AC reversibles, además la probabilidad de ser la salida 0 o 1 es la misma. Otra propiedad que tiene es que la función que cifra el archivo es la misma que lo descifra. En el algoritmo 1 se muestra el flujo que se ocupó para desarrollar el modelo.

Algorithm 1 Algoritmo del modelo propuesto para cifrar/descifrar imágenes

InputEntrada OutputSalida

Imagen en claro D , llave K Imagen cifrada E

Inicialización El usuario ingresa la imagen a cifrar D y la llave K Se realiza una transformación a la llave para obtener tamaño de bloque fijo k Se parte la imagen en j bloques de tamaño k existan bloques de la imagen valor del bloque j_i de la imagen en la posición $[i] = 0$ Se toma la regla 232 de Wolfram para realizar cifrado Se ve el valor de los bits de la llave en la posición $i - 1, i, i + 1$ Se regresa el valor del bit de salida, ya sea 0 o 1

Se toma la regla 23 de Wolfram para realizar cifrado Se ve el valor de los bits de la llave en la posición $i - 1, i, i + 1$ Se regresa el valor del bit de salida, ya sea 0 o 1

Se agrega a la nueva imagen cifrada E el valor de bit regresado. Se guarda la imagen E .

3. Resultados

El modelo propuesto fue simulado en el lenguaje de programación *Python*, con un procesador intel core i5 y 4 GB de RAM. El procesamiento se hizo a nivel de bits, esto es, se realizaron corrimientos y operaciones lógicas para este proceso, esto hizo que la ejecución del algoritmo fuera rápida. Además, se

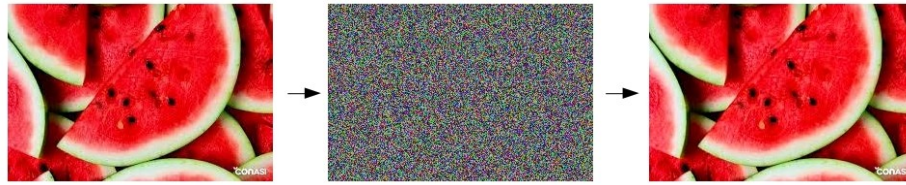


Fig. 3. Figura de sandías en plano, imagen cifrada con el algoritmo propuesto y salida obtenida.



Fig. 4. Figura de frutas en plano, e imagen cifrada con el algoritmo propuesto y salida obtenida.

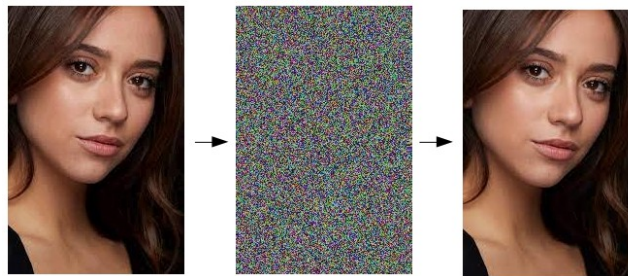


Fig. 5. Figura de rostro en plano, e imagen cifrada con el algoritmo propuesto y salida obtenida.



Fig. 6. Figura de gama de colores en plano, e imagen cifrada con el algoritmo propuesto y salida obtenida.

obtuvieron como salida imágenes, las cuales no tenían relación visual con las imágenes de entrada.

Tabla 1. Valores de la imagen en claro, y de la imagen cifrada obtenida con el modelo propuesto

Figura	Entropía texto cifrado	Entropía texto plano
Sandías	7.6552590369485	7.4757733701720
Frutas	7.6626427475365	7.3809175636137
Rostro	7.6595972685455	7.4639996896011
Gama de colores	7.65532224872648	7.3518466899058

En las Figuras 3, 4, 5 y 6 se observa la imagen original a cifrar, la cual llamamos texto plano o imagen en plano; seguida a esta, se muestra lo que es llamado imagen cifrada, que es la que se obtiene al procesar la imagen con una llave k . Finalmente, se muestra la imagen que obtenemos al hacer el proceso inverso a la imagen cifrada; se regresa a la imagen original teniendo la misma llave k , esto es, de no poner la llave correcta no se obtendría la misma imagen.

3.1. Entropía

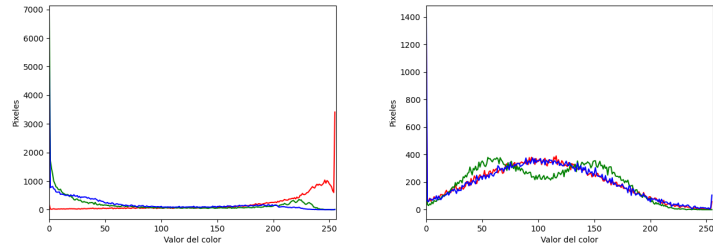
La entropía de la información es una medida estadística de aleatoriedad, por lo tanto, se puede utilizar para caracterizar la textura de una imagen de entrada. Si la entropía de la imagen cifrada es menor que la entropía de la imagen en plano, entonces la imagen puede predecirse y amenaza su seguridad [14]. En la tabla 1 se muestran los valores obtenidos de la entropía de las imágenes 3, 4, 5 y 6.

La entropía $H(m)$ de un mensaje m , con $p(m_i)$ representa la probabilidad del símbolo m_i y N es el número total de píxeles en la imagen y la entropía se expresa en bits se puede calcular como:

$$H(m) = \sum_{i=1}^{N-1} p(m_i) \log_2 \frac{1}{p(m_i)} \text{ bits.} \quad (1)$$

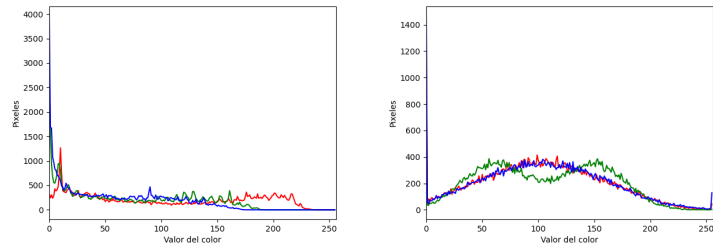
3.2. Histogramas

En las Figuras 7, 8 y 9 se muestran los histogramas de las figuras presentadas previamente. Con los histogramas se muestra que, a pesar de la distribución que tenga la imagen original, todas las imagen cifradas tienden a poseer distribuciones similares. Eso es, debido a las operaciones que se realizan en el cifrador, además de que la llave que se ocupó fue la misma para todas las imágenes; los histogramas no muestran patrones existentes.



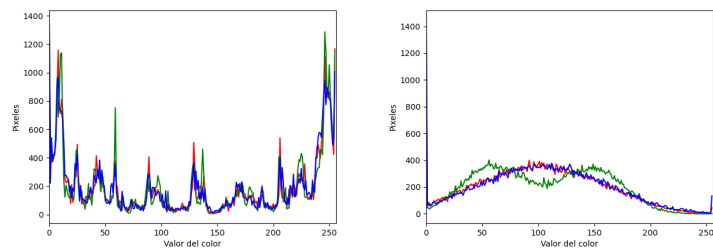
(a) Histograma de imagen en plano de Sandía. (b) Histograma de imagen cifrado de Sandía.

Fig. 7. Histogramas de la imagen 3.



(a) Histograma de imagen en plano de rostro. (b) Histograma de imagen cifrado de rostro.

Fig. 8. Histogramas de la imagen 5.



(a) Histograma de imagen en plano de gama de colores. (b) Histograma de imagen cifrado de gama de colores.

Fig. 9. Histogramas de la imagen 6.

4. Conclusión

En este trabajo un novedoso esquema de cifrado de imágenes fue propuesto. Además, cuenta con alta entropía y salida de imágenes sin relación visual, su

tiempo de ejecución es de segundos debido a que todas las operaciones fueron realizadas a niveles de bits. En el proceso de cifrado y descifrado no se cuenta con pérdida de información. Por lo descrito anteriormente, se dice que el sistema mantiene la confidencialidad en las imágenes. El algoritmo se puede aplicar para cualquier tipo de imagen representada por una profundidad de 24 bits así como a diferentes formatos de imagen (.JPEG, .JPG, .BMP, .PNG).

Referencias

1. Hameed, Y., Nada, M.: Enhanced RC5 key schedule using one-dimensional cellular automata for audio file encryption. *Iraqi Journal of Science*, 60, 388–401 (2019)
2. Jun, J.: An image encryption based on elementary cellular automata. *Optics and Lasers in Engineering* 20(12), 1836–1843 (2012)
3. Abdorreza, B., Homayun, M., Rasul, E.: A new permutation-diffusion-based image encryption technique using cellular automata and DNA sequence. *Optik*, 203 (2020)
4. Faraoun, K.: A parallel block-based encryption schema for digital images using reversible cellular automata. *Engineering Science and Technology, an International* 17(2), 85–94 (2014)
5. Ping, P., Jinjie, W., Yingchi, M., Feng, X., Jinyang, F.: Design of image cipher using life-like cellular automata and chaotic map. *Signal Processing* 150, 233–247 (2018)
6. Abdo, A., Shiguo, L., Ismail, A., Amin, M., Diab, H.: A cryptosystem based on elementary cellular automata. *Communications in Nonlinear Science and Numerical Simulation*, 18(1), 136–147 (2013)
7. Satyabrata, R., Jyotirmoy K., Rawat, U., Dayama, P., Nilanjan, D.: Symmetric Key Encryption Technique: A Cellular Automata based Approach in Wireless Sensor Networks. *Procedia Computer Science* 78, 408–414 (2016)
8. Satyabrata, R., Manu, S., Umashankar, R., Chirag, V., Sanjeet, K.: IESCA: An efficient image encryption scheme using 2-D cellular automata. *Journal of Information Security and Applications* 61, 2126–2214 (2021)
9. Bhaskar, M., Shrey, S., Prabhakar, K.: A secure image encryption scheme based on cellular automata and chaotic skew tent map. *Journal of Information Security and Applications*, 45, 117–130 (2019)
10. Ilachinski, A.: *Cellular automata a discrete universe*. 2nd edn. World Scientific (2001)
11. Hedlund, G.: Endomorphisms and automorphisms of the shift dynamical systems. *Mathematical Systems Theory* 3(4), 320–375 (1969)
12. Kari, J.: *Reversible Cellular Automata: From Fundamental Classical Results to Recent Developments*. *New Generation Computing* 36, (2018)
13. Christof, P., Jan, P.: *Understanding cryptography*. Springer, New York (2009)
14. Kaur, T., Sharma, R.: Image cryptography by TJ-SCA: Supplementary cryptographic algorithm for color images. *International Journal of Scientific & Engineering Research (IJSER)*, 4(7) (2013)

Simulación computacional de nanosistemas paradetección de la proteína hemaglutinina

J.A. Salazar Cano¹, J. I. Rodriguez Hernandez², J.C. Chimal Eguía¹

¹ Instituto Politécnico Nacional,
Centro de Investigación en Computación,
México

² Instituto Politécnico Nacional,
Escuela Superior de Física y Matemáticas,
México

`jsalazarc2019@cic.ipn.mx`

Resumen. La presente investigación se basa en un estudio de acoplamiento molecular de la hemaglutinina (HA), del virus de influenza AH1N1, con ocho diferentes ligandos predichos por la mecánica cuántica: i) ácido siálico aislado (Nc) y, siete nanoestructuras estables formadas por ácido siálico con estructuras metálicas complementarias: ii) cluster con cuatro átomos de oro con simetría romboédrica (Au4- rhombo1- Nc); iii) cluster de veinte átomos de cobre con arreglo simétrico tetragonal (Cu20-tetra-Nc); iv) cluster con cuatro átomos de oro con simetría piramidal (Au4-pyra-Nc); v) cluster con veinte átomos de oro con simetría tetragonal (Au20-tetra-Nc); vi) cluster de oro con cuatro átomos y simetría romboédrica (Au4-rhomboVar-Nc); vii) cúmulo de plata con veinte átomos y simetría tetragonal (Ag20- tetra-Nc) y; viii) cúmulo de cobre con simetría icosaédrica (Cu20- ico-Nc). Al intentar evaluar las estructuras anteriores en su acoplamiento con HA, se pretendió que sobre todo la parte de ácido siálico estuviera ligada incluso cuando el software no contenía las subrutinas (campo de fuerza) para átomos metálicos. La estructura de HA se obtuvo de la base de datos Protein Data Bank (PDB) con el código 3LZG, y a la cual se aplicó un método de acoplamiento para determinar los mejores sitios de unión ligando-proteína. Los resultados de la simulación muestran que la mejor energía de interacción corresponde al mínimo global asociado con Cu20- ico-Nc.

Palabras clave: Virus de la Influenza H1N1, hemaglutinina, ácido siálico, nanoestructuras.

Computational Simulation of Nanosystems for Detection of Hemagglutinin Protein

Abstract. The present research is based on a study of molecular coupling of hemagglutinin (HA), of the influenza virus AH1N1, with eight different ligands predicted by quantum mechanics: i) isolated sialic acid (Nc) and, seven stable nanostructures formed by sialic acid with complementary metal structures: ii) cluster with four gold atoms with rhombohedral symmetry (Au4-rhombo1- Nc);

iii) cluster of twenty copper atoms with a tetragonal symmetric arrangement (Cu₂₀-tetra-Nc); iv) cluster with four gold atoms with pyramidal symmetry (Au₄-pyra-Nc); v) cluster with twenty gold atoms with tetragonal symmetry (Au₂₀-tetra-Nc); vi) cluster of gold with four atoms and rhombohedral symmetry (Au₄-rhomboVar-Nc); vii) cluster of silver with twenty atoms and tetragonal symmetry (Ag₂₀- tetra-Nc) and; viii) cluster of copper with icosahedral symmetry (Cu₂₀- ico-Nc). When trying to evaluate the above structures in their coupling with HA, it was intended that above all the sialic acid part was ligated even when the software did not contain the subroutines (force field) for metal atoms. The structure of HA was obtained from the Protein Data Bank (PDB) database with code 3LZG, and to which a docking method was applied in order to determine the best ligand-protein binding sites. The simulation results show that the best interaction energy corresponds to the global minimum associated with Cu₂₀- ico-Nc.

Keywords: H1N1 influenza virus, hemagglutinin, sialic acid, nanostructures.

1. Introducción

A lo largo de la historia las pandemias han representado una alarmante preocupación para la humanidad, debido a las pérdidas que representan cada una de estas. Actualmente, el virus de la influenza aun representa una seria amenaza año con año, debido a que es la principal responsable de la epidemia estacional, atacando tracto respiratorio, garganta, bronquios y la gravedad de esta variara dependiendo de diversos factores causando efectos leves o incluso la muerte.

Desde su origen, esta nueva cepa afectó a miles de personas en el país, extendiéndose alrededor del mundo debido a que las personas infectadas desconocían la existencia del virus, confundiéndolo con una simple gripa o neumonía. El virus de la influenza A circula en diferentes tipos de huéspedes, tales como humanos, aves, caballos, puercos, entre otros, por lo que es latente la presencia de nuevas cepas, generando así una nueva repercusión en aspectos relacionados en el área de la salud y economía.

Actualmente las herramientas computacionales han permitido comprender las estructuras y sus interacciones a un nivel sorprendente como es el caso del docking molecular, en donde es posible identificar sitios de unión de un ligando y una proteína viral. La técnica de docking molecular implica dos procesos: La predicción de la conformación del ligando, posición y orientación dentro de los sitios y la evaluación de la afinidad de unión. Esto se puede lograr mediante, el uso de algoritmos de muestreo tales como algoritmos genéticos o de búsqueda estocástica por mencionar algunos, mientras que la función de puntuación se basa en campos de fuerza, empíricas y basadas en conocimiento.

La influenza H1N1, pertenece a los Orthomyxoviridae los cuales son virus de 150 – 200 nm de diámetro, el cual se encuentra envuelto en una capa lipídica sobre la cual se sitúan dos proteínas principales (figura 1): la hemagglutina (HA) que facilitan la entrada del virus, y neuraminidasa (NA), que permiten la liberación viral [1]. Se han identificado dieciséis subtipos diferentes de HA y nueve de NA (dieciocho HA y once NA, si se incluyen virus similares a la influenza A en murciélagos) entre las diferentes cepas del virus de la influenza A. [2]

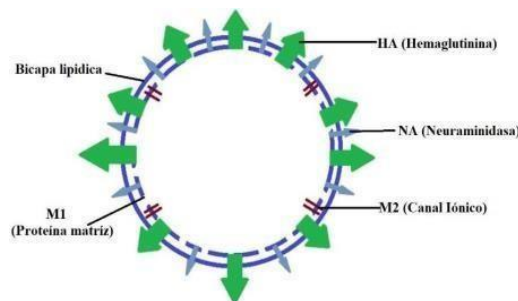


Fig. 1. Representación esquemática del virus de la influenza.

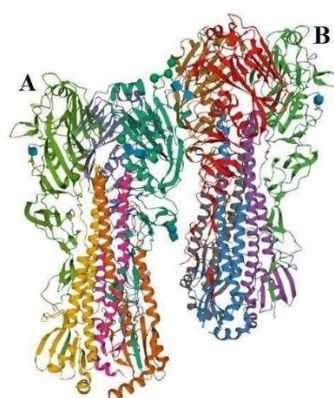


Fig. 2. Hemaglutinina del virus de la influenza H1N1 2009 (3LZG PDB). A) Subunidad HA1, B) Subunidad HA2

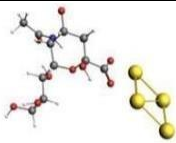
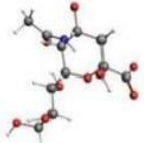
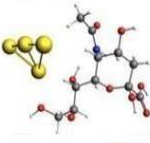
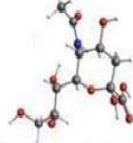
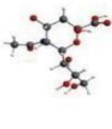
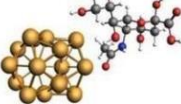
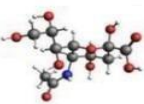
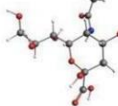
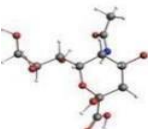
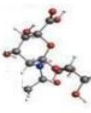
La creación de nuevas vacunas junto con la generación de biosensores desempeñan un papel fundamental en el control y prevención del virus de la influenza. Actualmente existen técnicas para la detección y el diagnóstico del virus de la influenza necesitan equipo especializado, personal capacitado, reactivos costosos por lo que actualmente se han venido desarrollando diversos biosensores portátiles con una alta sensibilidad, tales como los sensores basados en microfluidos, interferometría, nanotubos de carbono, cristal de cuarzo, electroquímicos de superficie de diamante dopado con boro, entre otros. [3-8].

2. Materiales y métodos

La hemaglutinina es una de las principales proteínas del virus de la influenza, la cual juega un papel fundamental en la unión y penetración en la célula huésped.[2], tiene una forma cilíndrica con dimensiones aproximadas de 135 Å de longitud y de 35 a 70 Å de radio formando picos que se proyectan externamente [9]. Esta proteína está constituida de dos partes HA1 y HA2 [10].

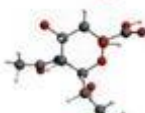
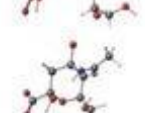
La proteína HA seleccionada para esta investigación (figura3) fue la 3LZG de acuerdo con el Protein Data Bank (PDB), la cual está constituida por dos subunidades HA1 y HA2 (figura 4 y 5).

Tabla 1. En la primera columna se muestran las estructuras formadas por un nanoclúster metálico y una molécula de ácido siálico obtenidas mediante simulaciones computacionales basadas en DFT, en la segunda columna se encuentra el ácido siálico conservando la primera columna y en la tercera columna se muestran las etiquetas para cada sistema.

Imagen tridimensional	Ácido Siálico	Etiqueta
		Au₄-rombo1-Nc
		Au₄-pira-Nc
		Au₂₀-tetra-Nc
		Cu₂₀-tetra-Nc
		Cu₂₀-ico-Nc
		Nc
		Ag₂₀-tetra-Nc

Los ligandos candidatos para llevar a cabo la interacción con la Hemaglutinina fueron los siguientes. Sin embargo, debido a que el software no posee el campo de fuerza adecuado para tratarla parte metálica, se trabajó únicamente con la molécula del ácido siálico a una primera aproximación. Para llevar a cabo la interacción de la hemaglutinina (HA) con cada uno de los ligandos mencionados en la tabla 4, fue necesario que los

Tabla 2. Se muestra que el ácido siálico y la proteína HA pueden formar complejos o moléculas estables. En la tercera columna se muestran los mínimos globales asociados a un score.

Ácido Siálico	Descripción	Energía de docking (kcal/mol)
	Cu₂₀-ico-Nc	-7.4
	Nc	-6.8
	Au₂₀-tetra-Nc	-6.7
	Cu₂₀-tetra-Nc	-6.6
	Ag₂₀-tetra-Nc	-6.6
	Au₄-pira-Nc	-6.4
	Au₄-rombo1-Nc	-6.2

formatos XYZ de cada uno de ellos fueran transformados al formato PDB, esto con el fin de que el software UCSF Chimera pueda leer la información topológica de cada uno de los mismos.

3. Resultados

Los resultados de la primera etapa del acoplamiento molecular de los ocho ligado candidatos y la HA (hemaglutinina).

4. Conclusiones

La presente investigación muestra los resultados de la técnica de docking molecular entre ocho ligandos (ácido siálico) y la hemagglutina del virus de la influenza H1N1, por medio del uso del software UCSF Chimera, encontrando que todos se ligan a la proteína. El análisis de la exploración del espacio conformacional nos indica que los posibles sitios de unión se encuentran distribuidos tanto en la subunidad HA1 y HA2, así como también entre ambas subunidades. Posterior a dicha exploración del espacio conformacional se encontró que el mejor candidato para ser utilizado en el biosensor fue el ácido siálico asociado al nanocúmulo de cobre icosaedro con un score asociado de $|-7.4| = 7.4$ kcal/mol, debido a que presenta la mejor energía de afinidad.

Referencias

1. Lamb, R.A.: Influenza, Editor(s): Brian W.J. Mahy, Marc H.V. Van Regenmortel, Encyclopedia of Virology, Academic Press, pp. 95–104 (2008)
2. Fouchier, R.A.: Characterization of a new hemagglutinin subtype of influenza A (H16) virus obtained from laughing gulls. *J. Virol.*, 79, pp. 2814–2822 (2005)
3. Singh, R., Hong, S.: Label-free Detection of Influenza Viruses using a Reduced Graphene Oxide- based Electrochemical Immunosensor Integrated with a Microfluidic Platform. *J. Sci Rep.*, 7, 42771 (2017)
4. Farris, L.R., Wu, N., Wang, W., Clarizia, L.J., Wang, X., McDonald, M.J.: Immuno-interferometric sensor for the detection of influenza A nucleoprotein. *Anal Bioanal Chem*, 396(2), pp. 667–674 (2010)
5. Cai, Q., Xinsheng, T., She, C., Jing-Hua, Y.: Detection of avian influenza virus subtype H5 using a biosensor based on imaging ellipsometry. *Biosensors and Bioelectronics*, 25(6), pp. 1530 (2009) doi:10.1016/j.bios.2009.10.030.
6. Qi, C., Tian, X.S., Chen, S., Yan, J.H., Cao, Z., Tian, K.G., Gao, G.F., Jin, G.: *Biosens Bioelectron*, 25(6), pp. 1530 (2010)
7. Lee, D., Chander, Y., Goyal, S.M., Cui, T.: Carbon nanotube electric immunoassay for the detection of swine influenza virus H1N1. *Biosens Bioelectron*, 26(8), pp. 3482 (2011)
8. Su, C., Chang, C.M., Tseng, Y.L., Chang, Y.F., Li, Y.C., Chang, Y.S., Chou, C.: Rapid and highly sensitive method for influenza A (H1N1) virus detection. *Anal Chem.*, 84(9), pp. 3914–3920 (2012)
9. Peduru Hewa, T.M., Tannock, G.A., Mainwaring, D.E., Harrison, S., Fecondo, J.V.: The detection of influenza A and B viruses in clinical specimens using a quartzcrystal microbalance. *J VirolMethods*, 162(1-2), pp. 14–21 (2009)
10. Palese, P., Tobita, K., Ueda, M., Compans, R.W.: Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. *Virology*, 61, pp. 397–410 (1974)
11. Mitnaul, L.J., Matrosovich, M.N., Castrucci, M.R., Tuzikov, A.B., Bovin, N.V., Kobasa, D., Kawaoka, Y.: Balanced hemagglutinin and neuraminidase activities

Simulación computacional de nanosistemas para detección de la proteína hemaglutinina

are critical for efficient replication of influenza. A virus. J Virol , 74, pp. 6015–6020 (2000)

Electronic edition
Available online: <http://www.rcs.cic.ipn.mx>



<http://rsc.cic.ipn.mx>



Centro de Investigación
en Computación