الجمهـوريـة الجزائريـة الديمقراطيـة الشعبيـة

PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA

University M'Hamed Bougara of Boumerdes

كليـة التكنولوجيـا

Faculty of Technology

The 1st National Workshop on Wireless Network, Cloud Computing and Cryptography (WWN3C'23)

April 26, 2023, Boumerdes

Editors

BAICHE Karim

ACHELI Dalila

AMMAR Mohammed

HARRAR Khaled

DAOUI Abdelhakim

AKROUM Hamza

BELKACEM Samia

AIBECHE Abderezzak

MESSAOUDI Noureddine

MERAIHI Yassine

ISBN: 978-9969-9733-2-7

Dépôt Légal: 9969-2024

Preface

*The M'Hamed Bougara university is very pleased to welcome you in Boumerdes at the first National Workshop on Wireless Network, Cloud Computing and Cryptography (**WWN3C'2023**) held in the Faculty of Technology.*

*The **WWN3C'2023** aims for presenting new advances and research results in the fields of Wireless Networks, Cloud Computing, Cryptography, Telecommunications and Signal Processing.*

The Workshop is expected to provide researchers, lecturers, engineers, and scientists around the world the opportunity to interact and present their latest advanced research.

A wide range of topics is addressed, including Wireless Communications, 5G Communication Networks, Internet of Things, Cloud Computing, Medical Imaging, Pattern Recognition and Machine Learning, Image and Video Processing, Artificial Intelligence and Deep learning.

*We deeply thank the organizing committee and all those who by their help, devotion, competence, and good mood, allowed the good progress of the day of the **WWN3C'2023**.*

We would like to thank the invited speaker for her excellent track record in her own field of research.

We also thank all the participants who insured the animation of the Workshop, and allowed fruitful exchanges and discussions profitable to everybody.

WWN3C'2023

Organizing Committee

Honorary General Chairs

Pr. Mostepha YAHI, Rector of Univ. Boumerdes, Algeria
Pr. Mohamed SAIDI, Dean of Faculty of Technology, Univ. Boumerdes,Algeria

Conference General Chairs

Pr. Dalila ACHELI, Univ. Boumerdes, Algeria
Dr. Karim BAICHE, Univ. Boumerdes, Algeria

Organizing Committee Chair

Dr. Yassine MERAIHI, Univ. Boumerdes, Algeria

Organizing Committee

Miss. Sihem TIRECHE, Univ. Boumerdes, AlgeriaDr.
Youcef TABET, Univ. Boumerdes, Algeria
Dr. Amel BOUSTIL, Univ. Boumerdes, Algeria
Dr. Med Amine RIAHLA, Univ. Boumerdes, Algeria
Dr. Khaled HARRAR, Univ. Boumerdes, Algeria
Dr. Mohammed AMMAR, Univ. Boumerdes, Algeria Dr.
Hamza AKROUM, Univ. Boumerdes, Algeria
Dr. Hadjira BELAIDI, Univ. Boumerdes, Algeria
Miss. Faiza IBAOUNI, Univ. Boumerdes, Algeria
Miss. Selma YAHIA, Univ. Boumerdes, Algeria
Mrs. Sylia MEKHMOUKH, Univ. Boumerdes, AlgeriaMiss.
Amylia AIT SAADI, Univ. Boumerdes, Algeria
Miss. Souad REFAS, Univ. Boumerdes, Algeria
Dr. Abdelkrim AMMAR, Univ. Boumerdes, Algeria
Dr. Fadhila LACHEKHAB, Univ. Boumerdes, Algeria

Scientific Committee

Pr. ACHELI Dalila, Univ. Boumerdes, Algeria
Pr. AZRAR Arab, Univ. Boumerdes, Algeria
Pr. BENTARZI Hamid, Univ. Boumerdes, Algeria
Pr. BOUDEN Toufik, Univ Jijel, Algeria
Pr. BOUDRAH Rabah Univ. Msila, Algeria
Pr. BOUKABOU Abdelkrim, Univ Jijel, Algeria
Pr. BOUKRA Abdelmadjid, Univ. Bab Ezzouar, Algeria
Pr. BOUSHAKI Razika, Univ. Boumerdes, Algeria Pr.
CALLAL Mouloud, Univ. Boumerdes, Algeria
Pr. DAAMOUCHE Abdelhamid, Univ. Boumerdes, Algeria
Pr. FELLAG Sid Ali, Univ. Boumerdes, Algeria
Pr. FERGUENE Farid, Univ. Bab Ezzouar, Algeria Pr.
GHARBI Nawel, Univ. Bab Ezzouar, Algeria
Pr. GROUNI Saïd, Univ. Boumerdes, Algeria
Pr. HAMADOUCHE M'Hamed, Univ. Boumerdes, Algeria

Pr. HARKAT Mohammed Faouzi, Univ. Annaba, Algeria
Pr. HIDOUCI Walid-Khaled, ESI, Oued Smar, Algeria
Pr. HOCINI Abdesselem, Univ. M'sila, Algeria
Pr. KESRAOUI Mohamed, Univ. Boumerdes, Algeria
Pr. KHELDOUN Aissa, Univ. Boumerdes, Algeria Pr.
KHOUAS Abdelhakim, Univ. Boumerdes, Algeria
Pr. KIDOUCHE Madjid, Univ. Boumerdes, Algeria
Pr. KOUADRI Abdelmalek, Univ. Boumerdes, Algeria
Pr. LACHOURI Abderezzak, Univ. Skikda, Algeria
Pr. RAHMOUNE Fayçal, Univ. Boumerdes, Algeria
Pr. RAMDANE-CHERIF Amar, Univ. Versailles, France
Pr. RECIOUI Abdelmadjid, Univ. Boumerdes, Algeria
Pr. ROUABAH Khaled, Univ BBA, Algeria
Dr. AIBECHE Abderezzak, Univ. Boumerdes, Algeria
Dr. AHRICHE Aimad, Univ. Boumerdes, Algeria
Dr. AKLI Isma, CDTA, Algeria
Dr. AKROUM Hamza, Univ. Boumerdes, Algeria
Dr. AMMAR Abdelkarim, Univ. Boumerdes, Algeria Dr.
AMMAR Mohammed, Univ. Boumerdes, Algeria Dr.
AYAD Mouloud, Univ. Bouira, Algeria
Dr. BAICHE Karim, Univ. Boumerdes, Algeria
Dr. BELAIDI Hadjira, Univ. Boumerdes, Algeria
Dr. BELKACEM Samia, Univ. Boumerdes, Algeria
Dr. BENBELKACEM Samir, CDTA, Algeria
Dr. BENHAMOUDA Abdallah, Univ. Constantine 1, Algeria
Dr. BENLATRECHE Mohamed Salah, Univ. Mila, Algeria
Dr. BENMOUSSA Yahia, Univ. Boumerdes, Algeria
Dr. BENZAOUI Amir, Univ. Skikda, Algeria
Dr. BOUDOUDA Aimed, Univ. Boumerdes, Algeria
Dr. BOUSTIL Amel, Univ. Boumerdes, Algeria
Dr. BOUTELLAA El Hocine, CDTA, Algeria
Dr. CHERIFI Dalila, Univ. Boumerdes, Algeria
Dr. DEKHANDJI Fatma Zohra, Univ. Boumerdes, Algeria
Dr. DICHOU Karima, Univ. Boumerdes, Algeria
Dr. DJEKOUNE A. Oualid, CDTA, Algeria
Dr. FLITTI Abdelmadjid, Univ. Mostaghanem, Algeria
Dr. GUERBAI Yasmine, Univ. Boumerdes, Algeria
Dr. HARRAR Khaled, Univ. Boumerdes, Algeria
Dr. HENTOUT Abdelfetah, CDTA, Algeria
Dr. HIMEUR Yacine, CDTA, Algeria
Dr. HOCINE Faiza, Univ. Boumerdes, Algeria
Dr. IDIR Abdelhakim, Univ. M'sila, Algeria
Dr. KERMIA Omar, CDTA, Algeria
Dr. KONATRI Abdelmalek, Univ. Boumerdes, Algeria
Dr. KOUIDER Ahmed, CDTA, Algeria
Dr. MAHDI Ismahane, Univ. Boumerdes, Algeria

Dr. MAHSEUR Mohammed, Univ. Alger1, Algeria
Dr. MAOUCHE Riadh, Univ. Boumerdes, Algeria
Dr. MAOUDJ Abderraouf, CDTA, Algeria
Dr. MANSOUL Ali, CDTA, Algeria
Dr. MERAH MOSTEFA, Univ Mostaghanem, Algeria
Dr. MERAIHI Yassine, Univ. Boumerdes, Algeria
Dr. MESBAH Abdelhak, Univ. Boumerdes, Algeria
Dr. MESSAOUDI Noureddine, Univ. Boumerdes, Algeria
Dr. MEZAI Lamia, Univ. Constantine 2, Algeria
Dr. MOUATSI Abdelmalek, Univ. Boumerdes, Algeria
Dr. MOUHOUCHE Faiza, Univ. Boumerdes, Algeria
Dr. NAFA Fares, Univ. Boumerdes, Algeria
Dr. OMARI Tahar, Univ. Boumerdes, Algeria
Dr. OULEFKI Adel, CDTA, Algeria
Dr. OURARI Samia, Univ. Boumerdes, Algeria
Dr. REZZOUG Abdellah, Univ. Boumerdes, Algeria
Dr. RIAHLA Mohamed Amine, Univ. Boumerdes, Algeria
Dr. SEDJELMACI Ibticeme, Univ. Boumerdes, Algeria
Dr. SMAANI Billel, Univ. Mila, Algeria
Dr. TABET Youcef, Univ. Boumerdes, Algeria
Dr. TAHI Slimane, Univ. Bab Ezzouar, Algeria
Dr. TIBERKAK Allal, Univ. Médéa, Algeria
Dr. ZENATI Nadia, CDTA, Algeria
Dr. ZERAOULIA Khaled, Univ. Bab Ezzouar, Algeria

Invited Speaker

Communication in Smart-grid

Dr. Hadjira BELAIDI
University of Boumerdes, Algeria

Abstract

A smart grid is a modernized electrical grid that uses advanced technology to monitor, control and manage the flow of electricity from all generation sources to meet the changing demand of customers. It uses digital technology to enhance reliability and security and to ensure reliable, and sustainable delivery of electricity. Moreover, it uses sensors, controls, and automation to gather and analyze data on electricity production, distribution, and consumption in real-time, and enables a two-way flow of energy and information between utilities and consumers. In addition, it provides real-time information to consumers, allowing them to make informed decisions about energy use, and improves the overall security and resilience of the grid against cyber and physical attacks. Hence, communication infrastructure is critical to the effective functioning of the expanding smart grid. A robust and pervasive communication infrastructure is critical for both smart grid creation and operation.

In this presentation, the communication technologies used for smart grid data transmission are described and summarized. The three technologies (wireless, Ethernet, and PLC) are compared in terms of speed, stability, ease of connection, and cost …

This presentation focuses on the use of PLC (Power Line Communication) technology which presents an effective means of communication in areas not served by wireless technology. PLC eliminates the need to install new dedicated lines and communication data can be propagated through AC or DC power lines.

Biography

Hadjira BELAIDI received her Doctorate degree from the Institute of Electrical and Electronic Engineering, University M'hamed Bougara of Boumerdes in 2015. She has been an assistant professor at the Institute of Electrical and Electronic Engineering since 2014. Currently, she is the leader of the "Embedded Systems and Robotics" team in the Signals and Systems Laboratory, as well as a lecturer at the Department of Electronics at the Institute of Electrical and Electronic Engineering, University M'hamed Bougara of Boumerdes, Algeria. Her research interests include all-terrain mobile robots, environment modelling and manipulator robots, drone control, and smart-grid. She can be contacted at email: ha.belaidi@univ-boumerdes.dz.

Papers

Contents

Parametric Study Of A Frequency-Selective Surface With MEMS Switches Using The WCIP

Abdelmalik Mekaoussi
Department of Electronics, Faculty of Technology
University of Batna 2 (Mostefa Ben Boulaïd),
Batna 05000, Algeria
abdelmalikmekaoussi@gmail.com

Mohammed Titaouine
Department of Electronics, Faculty of Technology
University of Batna 2 (Mostefa Ben Boulaïd),
Batna 05000, Algeria
mohammedtitaouine@gmail.com

***Abstract*—In our new article we studied a frequency selective surface (FSS), this structure presents as a metal ring with two notches by the method which is called Wave Concept Iterative Procedure (WCIP). This method is based on the fast modal transformation (FMT), two switches of MEMS (microelectromechanical systems) are used to change existing frequencies or create new frequencies. This structure generates frequencies in the X and Y polarization, when the structure is excited along the X direction there are two resonances 7.24 GHz, 12.28 GHz with their bandwidths 2.21 GHz, 0.68 GHz successively, along Y there is a single resonance frequency at 7.6 GHz with a 2.72 GHz bandwidths. The simulation results obtained by the WCIP method and a good agreement should be noted after the comparison with the HFSS results.**

Keywords— MEMS, FSS, WCIP method, FMT, software HFSS.

I. INTRODUCTION

In recent years, the development of space telecommunications has required the construction of increasingly efficient and compact equipment. evolution appears in communication systems. It is accompanied by the design of high frequency circuits with manufacturing precision. The study of these circuits has been the subject of many works in recent years, in particular to reduce the computation time of electromagnetic simulations. Researchers have developed numerical methods to solve complex problems. The use of the wave frequency band in communication systems has stimulated research in the microwave field [1]. wireless communication systems have experienced a remarkable growth in the number of users and the diversity of the means of communication (Antenna, Wi-Fi, radar, tablets, etc.) have begun to integrate wireless communication systems, these means use different communication standards popularized by their commercial designations (bluetooth, WiFi, etc.). [2] Electronically controlled FSS structures are so important for adapting the antenna and compensating manufacturing errors for variations in the operating frequency of wireless communication systems [3]. the iterative procedure WCIP of the concept Wave is able to analyze FSS of rigorously arbitrary form [4], the structure which is proposed to us having two frequencies of resonance, a frequency when the structure is polarized in x and another one when the following polarization there. In order to validate the results of the WCIP method, the HFSS commercial software is used [5]. The use of frequency selective surface antennas, SFS, has attracted the attention of many research groups to minimize unwanted signals, to use the same antenna in at least two configurations, or to improve the performance of the antenna. antenna [6-7], all these results obtained with the WCIP method and are validated by comparison with the software simulation (HFSS).

II. FORMULATION OF WCIP

WCIP is the method introduced by Professor Henri Baudrand in 1995, it is a method of type method of moments adapted to the study of microwave circuits, Studies were therefore conducted on antennas coupled or mounted on ground planes, on coplanar guides and multilayer circuits, frequency selective surfaces SFS [8-9], periodic structures. In order to improve the method, different techniques have been envisaged: a spectral connection technique, a study of the reference impedance and the taking into account of the metallic state, By analogy between the electromagnetic quantities and the circuit-type variables (Equivalent diagrams), the incident "A" and diffracted waves "B" are defined on one (or more) interface (s) of the circuit. In the case where we return to a surface called "Ω" [10], the latter is the interface between two domains "Ω_1" and "Ω_2". The configuration is illustrated in Fig.1. where we show only two domains for the sake of simplicity [11-12].

The figure below represents a plane Ω on a periodic structure and this plane separates the two media (Medium 1, Medium 2), for each one of these media is characterized by the electric permittivity and its magnetic permeability, this is excited by a incident source plane waves $\vec{A}_0$, at the interface Ω we find two types of waves generated: the incident waves $\vec{A}_i$ and the reflected waves $\vec{B}_i$.

These last waves will be reflected by the upper and lower closures of the case to give a new wave that will constitute the incident waves of the next iteration [13].

Incident and reflected waves are defined as follows:

$$\vec{A}_i = \frac{1}{2\sqrt{Z_{0i}}}\left(\vec{E}_{Ti} + Z_{0i}\vec{J}_i\right) \quad (1)$$

$$\vec{B}_i = \frac{1}{2\sqrt{Z_{0i}}}\left(\vec{E}_{Ti} + Z_{0i}\vec{J}_i\right) \quad (2)$$

$$\vec{J}_i = \vec{H}_{Ti} \wedge \vec{n} \quad (3)$$

I. GEOMETRY OF THE FSS

For After the determination of the numerical characterizations of the geometry of the FSS. The lengths lg_1, lg_2, d_1 and d_2 are increased and decreased, the lengths lg_1, lg_2 being openings in the metal part with a notch in the structure.

The width of the ribbons is W, as shown in the figure below, the numerical characteristics is considered as a substrate of electrical permittivity $\varepsilon_r = 4.4$, the thickness h= 1 mm, the unit cell is defined by the dimensions: a = b = 20 mm, dx = 5 mm, dy = 5 mm, W = 2 mm, L_1 = 10 mm.

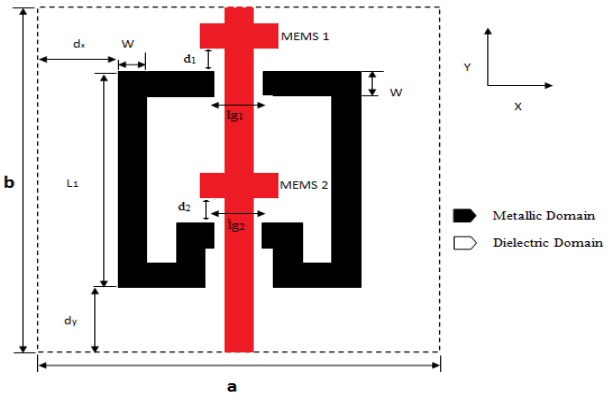


Fig. 1. The section of the structure FSS.

II. RESULTS OF THE FSS

The results that we obtained by the WCIP method and after the simulation by the HFSS when the incident plane wave whose electric field is polarized in X and Y.

A. Results of the simulations

Fig. 2 represents the transmission coefficient as a function of the operating frequency following the X polarization, there are two resonance frequencies observed at 7.24 GHz, 12.28 GHz in our results. the bandwidths in this simulation is 2.21 GHz, 0.68 GHz at -3 dB, Fig. 3 represents the Y polarization, the resonance frequency is observed at 7.6 GHz, the bandwidth is 2.72 GHz at -3 dB.

A good agreement is found in the results of the WCIP and the HFSS.

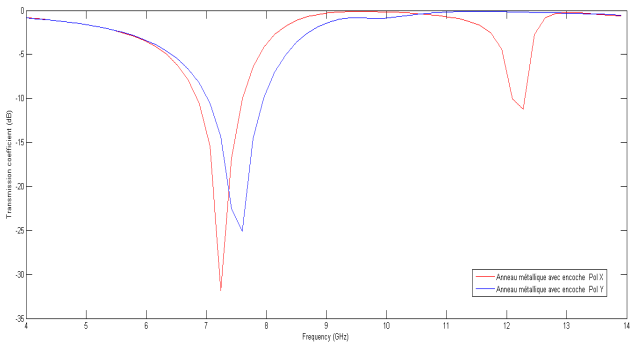


Fig. 2. The variation of the transmission coefficient as a function of the operating frequency, HFSS

Pol X : The X polarization

Pol Y : The Y polarization

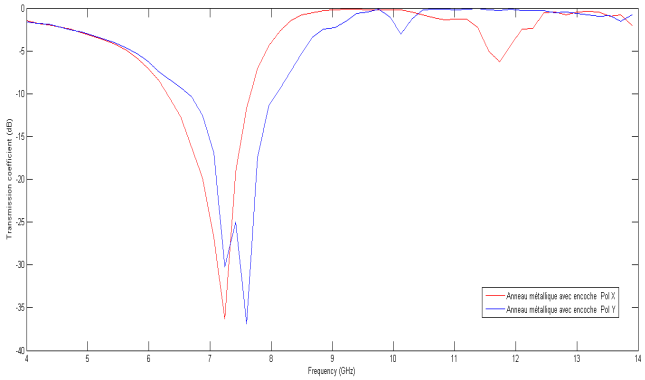


Fig. 3. The variation of the transmission coefficient as a function of the operating frequency, WCIP.

B. The parametric study of the

We make a parametric study of the lengths, the notch and the spaces for the determination of the dimensions of the structure, the variation of the space lg_1 and lg_2 in Fig.4 and Fig.5 we notice that the transmission coefficient of the FSS as a function of the frequency, when the length lg_1 varies from 1 mm to 6 mm the resonance increases but in the length lg_2 when increases from 0.2 mm to 1.6 mm, the resonance remains almost stable at 7.24 GHz and the transmission coefficient decreases by -32.79 GHz to 23.15 GHz.

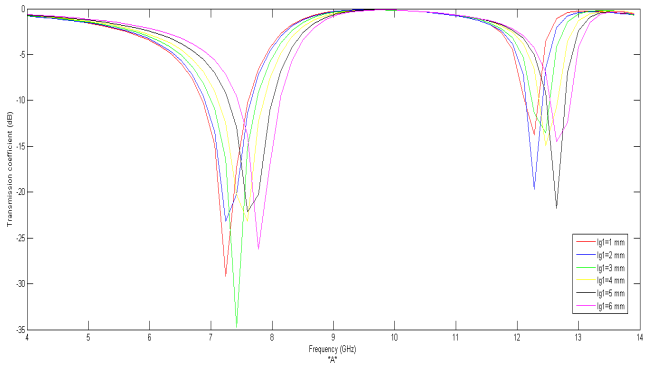


Fig. 4. The transmission coefficient as a function of the operating frequency for different lengths lg_1 .

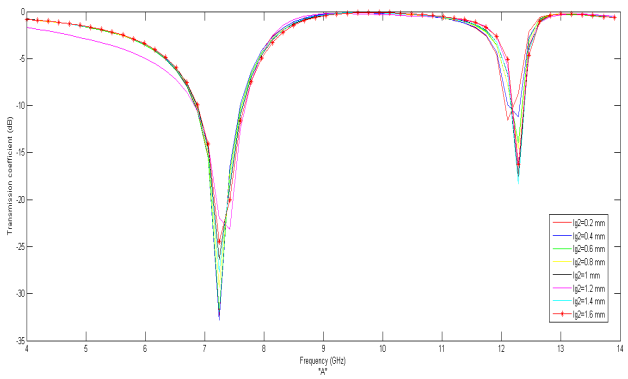


Fig. 5. The transmission coefficient as a function of the operating frequency for different lengths lg_2 .

When the length of lg_1 and lg_2 has been increased at the same time from 0.4 mm to 1.6 mm, the resonance increases.

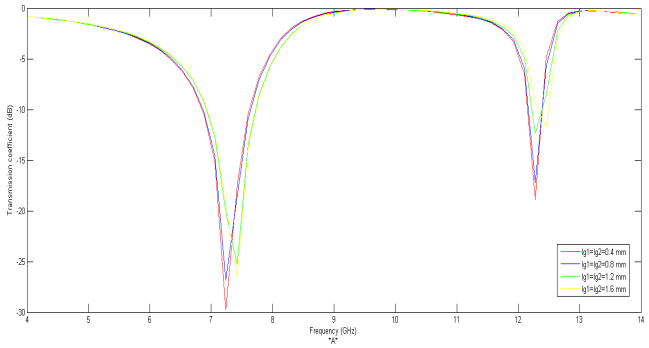


Fig. 6. The transmission coefficient as a function of the operating frequency for different lengths lg_1 and lg_2 .

After we place MEMS 1 in our structure, we study the length d_1 from 0 mm to 2 mm, when we have increased the length of d_1 by a step of 0.1 mm, the first resonance changes its value from 5.26 GHz to 9.94 GHz but the others are almost stable and we can eliminate some resonances as shown Fig.7.

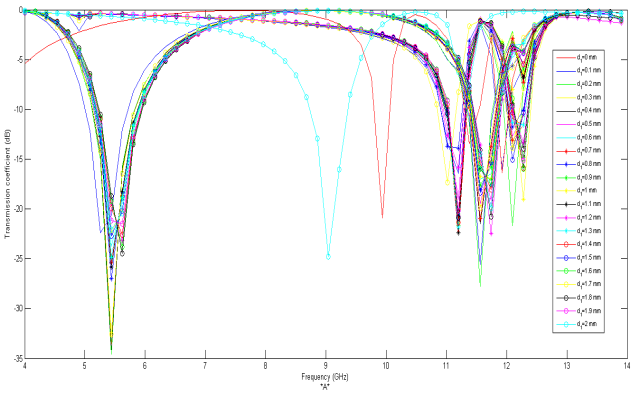


Fig. 7. The transmission coefficient as a function of the operating frequency for different lengths d_1.

In Fig.8 we studied the length d_2 from 0 mm to 2 mm with a step of 0.2 mm after placing the MEMS 2, when we increased the length of d_2, the resonances are almost stable and we can eliminate some resonances every time.

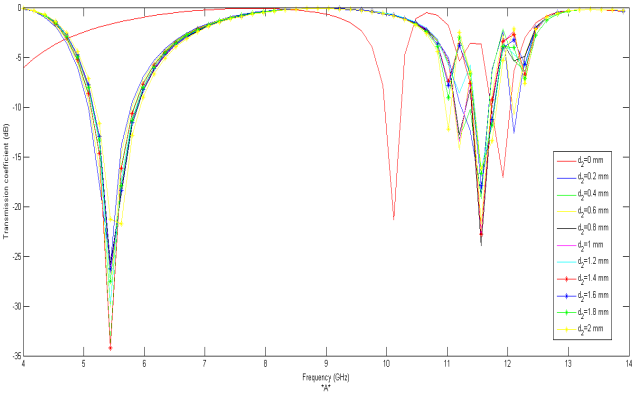


Fig. 8. The transmission coefficient as a function of the operating frequency for different lengths d_2.

In the following curves we studied the closing and the opening of the two MEMS, when we close MEMS 1 we find three frequencies in the X polarization (10.12 GHz, 11.38 GHz and 11.92 GHz). And two frequencies in Y (9.4 GHz and 11.38 GHz).

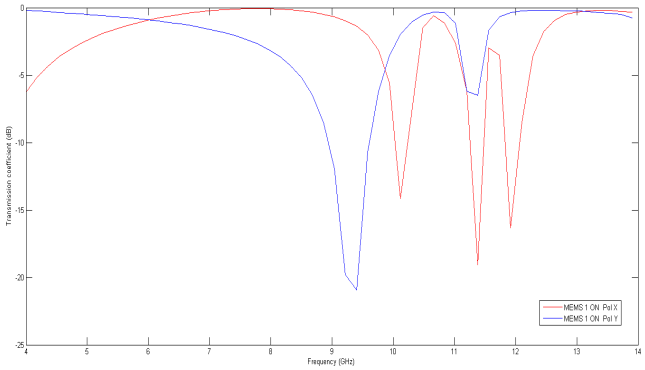


Fig. 9. The transmission coefficient as a function of the operating frequency for the MEMS 1 ON.

We notice in this following curve when we close MEMS 2 we have eliminated a frequency along X so we find two frequencies in the X polarization (9.94 GHz and 11.74 GHz). And two frequencies in Y (9.04 GHz and 12.1 GHz).

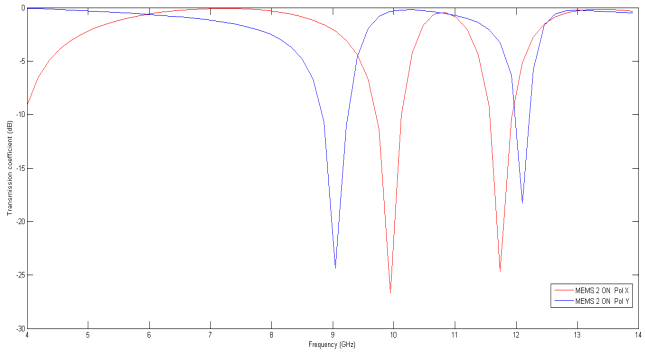


Fig. 10. The transmission coefficient as a function of the operating frequency for the MEMS 2 ON.

This time in the following Fig.11 we open the two MEMS 1 and 2, we found two frequencies in the X polarization (5.44 GHz, 11.56 GHz). And two frequencies in Y (11.2 GHz and 12.1 GHz), but we note that the two frequencies along X are distant and that the others are close.

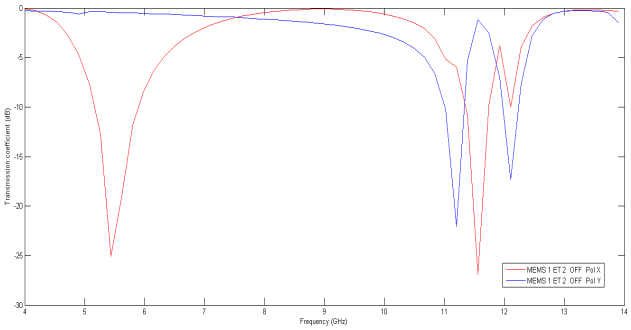


Fig. 11. The transmission coefficient as a function of the operating frequency for the MEMS 1 and 2 OFF.

The following Fig.12 we will close the two MEMS 1 and 2, so we have two frequencies in the X polarization (10.66 GHz, 11.92 GHz) and a frequency in Y (7.78 GHz), in this polarization we eliminate a frequency.

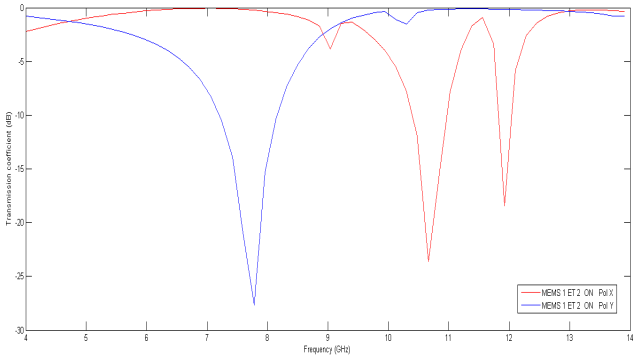


Fig. 12. The transmission coefficient as a function of the operating frequency for the MEMS 1 and 2 ON.

In the following curve we will open MEMS 1 and closed MEMS 2, so we have two frequencies in the X polarization (9.94 GHz, 11.56 GHz) and two frequencies in Y (9.04 GHz and 12.28 GHz), all the frequencies are close each other.

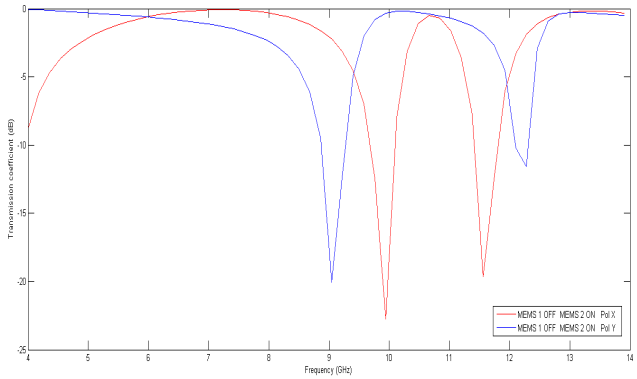


Fig. 13. The transmission coefficient as a function of the operating frequency for the MEMS 1 OFF and MEMS 2 ON.

The following Fig.14 we will closed MEMS 1 and open MEMS 2, we notice that there are two frequencies in the X polarization (10.12 GHz, 11.92 GHz) and a frequency in Y (9.22 GHz).

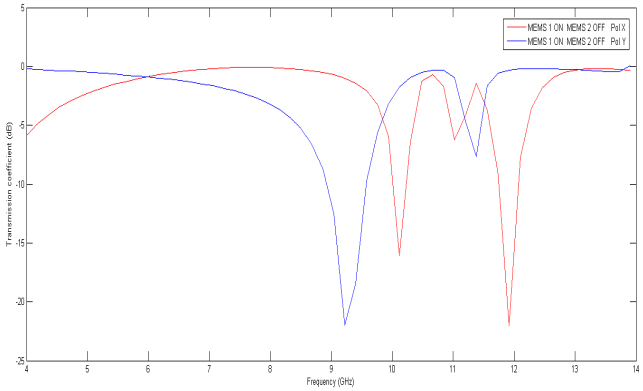


Fig. 14. The transmission coefficient as a function of the operating frequency for the MEMS 1 ON and MEMS 2 OFF.

C. The HFSS model of the FSS geometry

We have presented here the model of the structure in the HFSS software

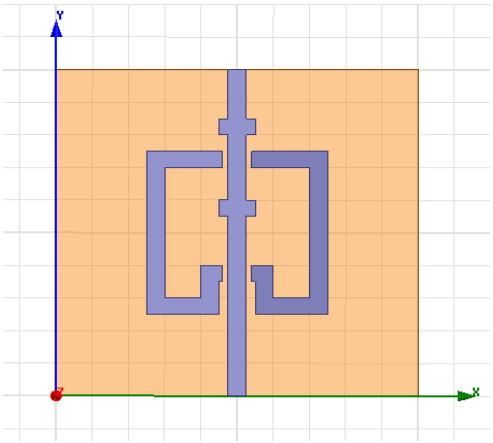


Fig. 15. The section of the structure FSS.

D. Conclusion

We presented in this new article a parametric study of the structure of the FSS, it is shown in this study that there are two resonances 7.24 GHz, 12.28 GHz for the X polarization by a plane wave and 7.6 GHz for the Y polarization, the structure of the FSS is presented for multiband applications, the results of this structure were performed by the WCIP method and validated by the simulation of the HFSS, we found that there is a very good agreement between the results of the method and the simulator.

E. References

[1] I. A. Ahmad, "Contribution to the modeling of SIW and SINRD structures for microwaves and telecommunication applications," PhD Thesis, Toulouse National Institute of Technology (INP Toulouse), Tuesday 25 June 2013.

[2] Alfrêdo Gomes Neto, Jefferson Costa e Silva, Joabson Nogueira de Carvalho, André Nascimento da Silva, Cecília Burle de Aguiar, Deisy Formiga Mamedes, "Analysis of Frequency Selective Surface with U-Shaped Geometry," JMOE, Vol. 14, SI-1, pp.113-122, 2015.

[3] C. Mias, "Tunable C-band Frequency Selective Surface", [on line] http://www.iee.org/OnComms/pn/antennas

[4] M. Titaouine, A. Gomes Neto, H. Baudrand, F. Djahli, "WCIP Method Applied to Active Frequency Selective Surfaces," JMOE, Vol. 6, No.1,pp. 1-16, 2007.

[5] I. Adoui, M. Titaouine, H. Choutri, R. Saidi, T. R. D. Sousa, A. G. Neto, H. Baudrand, "WCIP Method Applied To Modeling an L-Notched Rectangular Metallic Ring FSS for Multiband Applications and its Equivalent Structure," JMOe, Vol. 17, No. 4, pp.457-476, 2018.

[6] A. Edalati, T. A. Denidni, "Frequency selective surfaces for beam-switching applications," IEEE Trans. on Antennas and Propagation, Vol. 61, No. 1, pp. 195–200, 2013.

[7] T.K. Wu, "Frequency-Selective Surface and Grid Array," Wiley, New York, 1995.

[8] C. Girard, "Hybridation de methodes numeriques pour l'etude de la susceptibilité électromagnétique de circuits planaires," PhD Thesis, Toulouse National Institute of Technology (INP Toulouse), jeudi 18 décembre 2014.

[9] M. Azizi, H. Aubert, and H. Baudrand. "A new iterative method for scattering problems," In Microwave Conference, 1995. 25th European, Vol 1, pp 255–258, 1995.

[10] N. Raveu, "Contribution au développement de la formulation en ondes en coordonnées cylindriques : application à l'étude d'antennes placées sur un cylindre parfaitement conducteur," PhD thesis, Toulouse, INPT, 2003.

[11] N. Raveu and H. Baudrand,"Improvement of the WCIP convergence," In Antennas and Propagation Society International Symposium, APSURSI '09. IEEE, pp 1–4, 2009.

[12] N. Fichtner, S. Wane, Damienne Bajon, and P. "Russer, Interfacing the TLM and the TWF method using a diakoptics approach. In Microwave Symposium Digest," IEEE MTT-S International, pp 57–60, 2008.

[13] M. Aouissi, M. Titaouine, T. Rolime De Sousa, I. Adoui, A. G. Neto, H. Baudrand, " Analysis of a joined split-ring FSS structure characterized by three resonant frequencies and a tuned enhanced band using the WCIP method, " Turk J Elec Eng & Comp Sci, pp. 2885-2896, 2017.

Performance analysis of fairness and improved fairness power allocation NOMA Systems

Soulef AYAD
Department of electrical engineering, 20 aout 1955, Skikda University, Algeria
Laboratory of telecommunication, University 8 May 1945 Guelma, Algeria
ayadsoulef @ yahoo.fr

***Abstract*—** Power allocation is an important technique for non orthogonal multiple access (NOMA) system. Power domain (PD) NOMA has attracted great attention of searchers, academic and industrial interest of Mobile communications systems for the reason that it's potential of providing high spectral efficiency and massive connectivity. PD-NOMA is capable to increase the number of users compared to the orthogonal multiple access (OMA). The NOMA technique employs principally two steps which are superimposing coded and successive interference cancelation (SIC). The signal of multiple of users at the transmitter use the superimpose coding simultaneously over the same radio frequency resources and at the receiver the SIC is achieved to minimize the inter-user interference.

In this paper, we compare the results obtained by three methods of power allocation NOMA technique which are fixed, fairness and improved fairness power allocation. Preliminary simulation results have been achieved to confirm that NOMA used the improved methods to obtain the best result of the outage probability compared to classical methods.

Keywords- Non-orthogonal multiple access; Orthogonal multiple access; fairness power allocation, fixed power allocation, outage probability.

1. INTRODUCTION

With the spread of the Covid 19 in the last years, the demands to support wireless connected devices and different advanced applications have increased day by day. For this reason, the searchers have developed a new generation of mobile cellular communications systems, which have come with new standards, the use of advanced multiple access techniques, the change of modulation schemes and features, differentiating it from the previous one. In line with that, the 5G mobile communication systems are supported various advanced services including multimedia applications, Internet-of-Things (IoT)-based applications, and vehicle-to-everything (V2X).

Orthogonal multiple access (OMA) techniques are one of the important modifications to improve the quality of service of wireless communication systems more to more. These techniques have a significant impact on the utilization of the available spectrum, system throughput and latency. Some of orthogonal multiple access techniques which are widely used in the past generations of wireless networks such as: time division multiple access (TDMA), frequency division multiple access (FDMA) and code division multiple access (CDMA). These multiple access techniques characterize by different advantage but it present many disadvantages [1]-[5].

To evade the last problems, the scientific community is more and more oriented to develop a new multiple access technique named Non orthogonal multiple access technique ((NOMA). Therefore, NOMA is an efficient radio access technique, which is a candidate multiple access technique for LTE, 5G for mobile communication system and beyond 5G systems. It is fundamentally different than the classical multiple access techniques because the OMA is provided orthogonal access to the users either in time, frequency, code. In NOMA, each user operates in the same band and at the same time where the signals of the different users are separated by their power levels. NOMA uses superposition coding in the transmitter such that the successive interference cancellation (SIC) in the receiver [3]-[10]. NOMA technique is divided into two types, namely, power-domain and code-domain NOMA [3]-[12]. The users in the NOMA systems which use the code domain multiplexing are separated at the receiver by using redundancy via coding, but in the case of power domain multiplexing is able to perform successive interference cancellation (SIC) for the users with better channel conditions.

This paper is organized as follows: in section 2, we will explain the different steps of the NOMA technique. Section 3 describes the derivation of power allocation coefficients. Then, we explain the different Power allocation methods in section 4. In section 5, we present the simulation comparative results of three methods of the power allocation NOMA system. Finally, a conclusion is given.

2. The different steps of the NOMA technique:

In this part, we interest to study the downlink NOMA phase. Consequently, we consider that x_f and x_n are the signals transmitted from the BS to far and near users (i^{th} users (UES_i)). So, the NOMA system characterize by a high superiority compared to the OMA system because it is based on two mainly steps such as: the superposition coding and Successive Interference Cancellation:

- **Superposition coding:**

The general cases of superposition coded signals transmitted by the BS to the UES_i are given by [3], [5]-[8]:

$$x = \sum_{i=1}^{N} x_i \sqrt{P_i} \qquad (1)$$

Where P_i, i=1,2, is the transmit power for user i and the message signal x_i. The total transmit power of far and near user can then be written as $P_{BS}=P_1+P_2\ldots+P_i$.

The simplification of above expression can be given by:

$$x = \sqrt{P}\left(\sqrt{\alpha_f}.x_f + \sqrt{\alpha_n}.x_n\right) \quad (2)$$

In which, the Users are classified based on their distance from BS.

The transmitted signal when it passes through the channel is affected by the noise which is named the additive White Gaussian Noise (AWGN). This noise is not associated with either fading or any other system parameters. It is just the noise that is added to the signal when it is traveling through the channel. So, high data rate communication over additive white Gaussian noise channel (AWGN) is limited by noise. The mathematical expression in received signal y_i, for the UE_i is represented by following equation [3], [6]-[8], [13]-[16]:

$$y_{i=} h_i.x + n_i \qquad (3)$$

Where, h_i is the Rayleigh fading coefficients of UE_i and the base station, η_i is the additive White Gaussian Noise at the receiver of UE_i.

- **The Successive Interference Cancellation:**

The second step of NOMA systems is the Successive Interference Cancellation (SIC). This stage is used to separate between the different signals of users at the receiver side. Furthermore, we confirm that the use of this technique in the wireless communication system is very important to improve the performance of the receiver. Consequently, the idea of SIC technology is based to eliminate the folding signals by the problem of the interference of the signals which is cause by the multiple of users (as show in fig.2). First, the signals will be sequenced according to the high level of the SNR, then, the signal of the user who had the biggest SNR must be reflected, and eliminate the signal as a disturbance from the receiving signals. Next, it is necessary to reflect the signal of the user which had the second largest SNR [3], [10], [14].

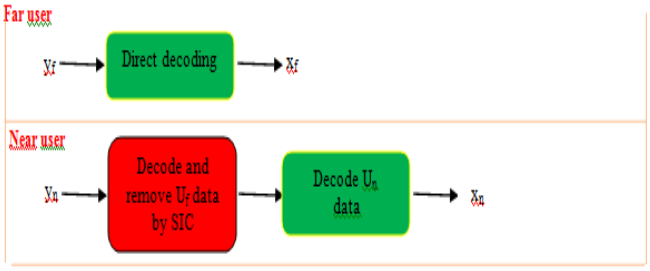


Fig.1. The Successive Interference Cancellation technique.

- **Capacity of the far user:**

The far user is assigned the highest power; it will perform direct decoding from y_f, treating the signals of near user as interference. Thus, the achievable capacity of the far user is [6-7], [10], [12-13], [16]:

$$C_f = \log_2\left(1 + \frac{\alpha_f P|h_f|^2}{\alpha_n P|h_f|^2 + \sigma^2}\right) \qquad (4)$$

From the above equation, we can see an important observation which is the power allocation coefficient $\alpha_n$is present at the denominator, and it is satisfy the following condition $\alpha_f>\alpha_n$. Only then, power of the far user will dominate in the transmit signal, x and in the received signal, y_f.

- **Capacity of the near user:**

Next we will present the formula of the capacity of near user C_{nn} by the equation below. While $\alpha_n<\alpha_f$, the near user must perform successive interference cancellation to remove the data of far user. After removing data of far user by SIC the achievable rate of the near user is given by [6]-[7] [12]-[13]:

$$C_n = \log_2\left(1 + \frac{\alpha_n P|h_n|^2}{\sigma^2}\right) \qquad (5)$$

3. Derivation of power allocation coefficients $\alpha_n$and α_f

To obtain the equations of power allocation coefficients of near user $\boldsymbol{\alpha_n}$ and far user $\boldsymbol{\alpha_f}$, we considers the capacity of the far user C_f equal to C. so, the equation (5) become:

$$C_f = \log_2\left(1 + \frac{\alpha_f P|h_f|^2}{\alpha_n P|h_f|^2 + \sigma^2}\right) = C \qquad (6)$$

To simplified the equation (6), we remove the $\log_2$ by taking 2^x on both sides,

$$1 + \frac{\alpha_f P|h_f|^2}{\alpha_n P|h_f|^2 + \sigma^2} = 2^C \qquad (7)$$

$$\frac{\alpha_f P|h_f|^2}{\alpha_n P|h_f|^2 + \sigma^2} = 2^C - 1 \qquad (8)$$

After that, we use the change of variable of equation (8) by $\mathcal{E}=2^C-1$, while we use the condition of power allocation coefficients which are given by $\alpha_n+\alpha_f=1$, $\alpha_n=1-\alpha_f$

Where $\mathcal{E}$ is the target of the signal-to-interference-plus-noise ratio (SINR) for the far user who has target rate C.

Finaly, we can be given α_n and α_f by:

$$\alpha_f = \frac{\mathcal{E}(P|h_f|^2+\sigma^2)}{P|h_f|^2(1+\mathcal{E})} \quad (9)$$

$$\alpha_n = 1 - \alpha_f \quad (10)$$

4. NOMA technique with different Power allocation methods

In the last generation of Mobile communication systems, NOMA technique which use the power allocation methods become an efficient procedure to improve the result of achievable rate more to more and decreause the outage propabilty of the signal of each user.So, in this study, we explain three methods of power allocation such as: fixed, fairness and improved fairness power allocation [17]-[19].

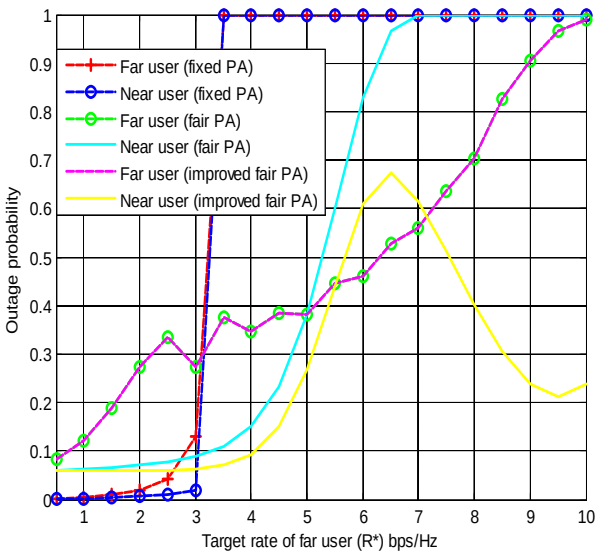
- **Fixed power allocation method:** in this case, the power allocation coefficient are fixed wherever the channel condition for the near and the far user are fixed and equal to α_n = 0.25, α_f=0.75 in the first case (in the second case α_n = 0.9, α_f=0.1).
- **Fairness power allocation method:** the principal point of this step is the use of the fairness power allocation algorithm which is considered as a dynamic scheme. Wherever, the channel of NOMA system changes the values of the power allocation coefficients (α_n and α_f) . Furthermore, the distribution of these values are updated between two successive calls.
- **Improwed fairness power allocation method:** To maximize the sum rate and energy efficiency more to more, we study another method which is the improved fairness power allocation algorithm. Wherever, the power allocation coefficient of the far user target rate is selected in the first case. After that, all the remaining available power is allocated to the near user.

5. DISCUSSION

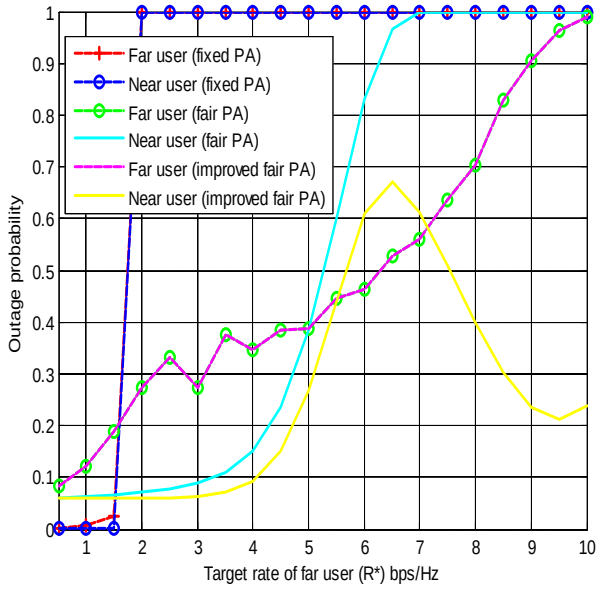
To examine the performance of power domain NOMA system, we compare the sum rate and the outage propability of NOMA system by using three methods such as fixed, Fairness and improved fairness power allocation.

From the Fig. 2., we can show that the outage propabilty result is improved with the varaition of the power allocation methods. So, it can see that a best result is obtained by using fairness power allocation especially improved Fairness method compared to the fixed power allocation. Furthermore, the power allocation coefficients should be selected carefully to reduce outage probability in the fixed power allocation method as shown in the fig.2 (a) and (b).

From the Fig. 2. (a) and (b), we can show that the NOMA system based on fixed power allocation (fixed PA) can be achieve a very poorly resultat wherever its outage probability increase and saturates to 1 when the target rate of the far user R>1.5 bps/Hz because, this method is not used the instantaneous channel state information and it not takes the target rate requirements into account. The fairness power allocation method has lower outage probability because the choose of α_n and $\boldsymbol{\alpha_f}$ are dynamically adjusted based on target rate requirement and the instantaneous channel conditions.

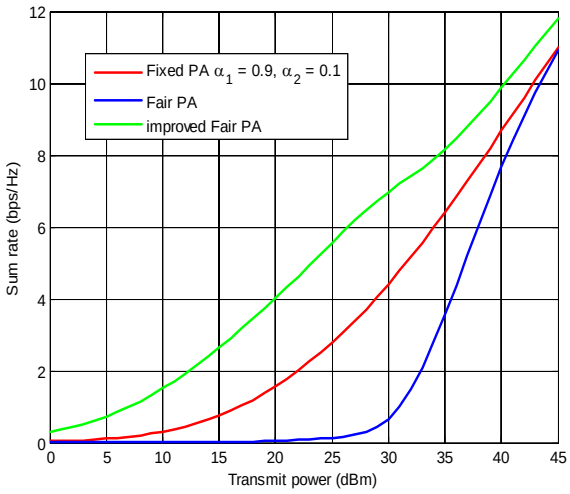


(a): For α_f=0.9, α_n=0.1

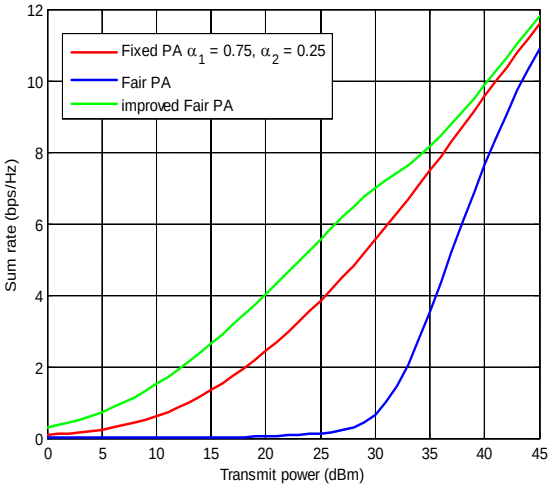


(b): For α_f=0.75, α_n=0.25.

Fig.2. Simulation results of NOMA outage probability using fixed, fairness, improved fairness power allocation methods.



(a): For α_f=0.9, α_n=0.1



(b): For α_f=0.75, α_n=0.25

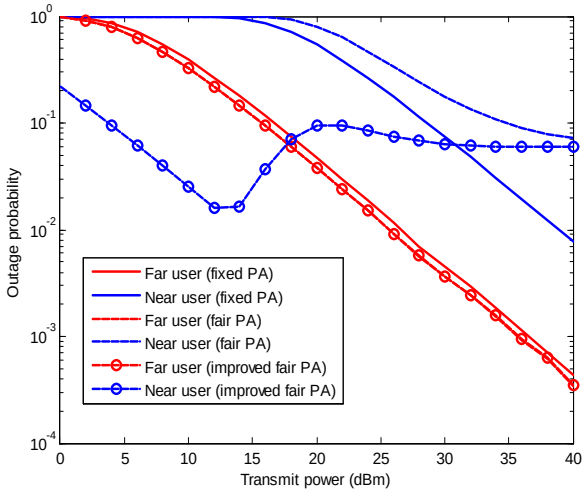
Fig.3. Simulation results of sum rate of NOMA system using fixed, fairness, improved fairness power allocation methods.

Furthermore, the outage probabilty of the far user is increase progressively with the increase of the target rate whether in the fairness and improved fairness power allocation methods. However, the outage of near user shows a sharp transition around 6.5 bps/Hz.

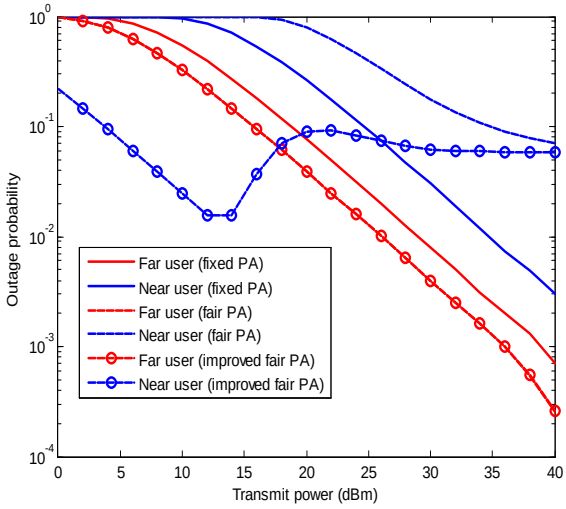
From the Fig.3. (a) and (b), we can find that, the sum rate obtained by using NOMA based on the improved fairness power allocation method is high than the results obtained by NOMA technique used fixed and fairness power allocation methods.

From the Fig.4. (a) and (b), we can see that, the outage propabilty of improved fairness power allocation NOMA system is lower than the results obtained by NOMA technique used fixed and fairness power allocation methods of the near user for the power level inferior to 20 dBm but after that the ouatge probability saturate at the same value.

As a results, we can conclude that the best solution to improve the the sum rate and minimize the outage propability is the improved fairness power allocation method.



(a): For α_f=0.9, α_n=0.1



(b): For α_f=0.75, α_n=0.25

Fig.4. the variation of outage probability of NOMA system with the change of the transmit power
using fixed, fairness, improved fairness power allocation methods.

Furthermore, the outage probability of NOMA system has a strong relationship with the power allocation coefficients and power level of the transmiter.

Conclusion

In this paper, the change of power allocation method consider as a very important idea to improve the efficient of Non-Orthogonal Multiple Access (NOMA) system. Therefore, we show that the outage probability of NOMA has a strong relationship with the values of power allocation coefficients and the power level. The conventional method is the Fixed Power Allocation where the power allocation coefficients are fixed and these coefficients have not related with the instantaneous channel conditions of users. But in the case of the Fairness power allocation the distribution of the power allocation coefficients are dynamically and these coefficients are updated with the change of channel conditions.

As a result, we have shown that the fairness power allocation methods presents many advantages compared to the fixed method, especially the improvement of fairness power allocation method. From the simulation results, we have shown that the improved method provides a minimun of the outage probability compared to the conventional method.

So, we can conclude that the improved fairness power allocation algorithm is able to achieve high sum rate and lower outage probabilty than the fairness and fixed power allocation methods.

References

1. M.A. Ahmed, K. F. Mahmmod, M. M. Azeez "On the performance of non-orthogonal multiple access (NOMA) using FPGA", International Journal of Electrical and Computer Engineering (IJECE), Vol. 10, No. 2, pp. 2151_2163, (April 2020).

2. b. lin, w. ye, x. tang, and z. ghassemlooy, "Experimental demonstration of bidirectional NOMA-OFDMA visible light communications", optics express, Vol. 25, No. 4 , (20 Feb 2017).

3. S.M. Riazul Islam, N. Avazov, O. A. Dobre, and K. Kwak," Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges", National Research Foundation of Korea-Grant funded by the Korean Government , pp. 1-41.

4. M. Vaezi, R. Schober, Z. Ding, and H. V. Poor, " Non Orthogonal Multiple Access: Common Myths and Critical Questions", pp. 1-13, (24 Jun 2019).

5. S. M. R. Islam, M. Zeng, O. A. Dobre, and K. Kwak " Non-Orthogonal Multiple Access (NOMA): How It Meets 5G and Beyond", pp.1-38.

6. K. Higuchi "Non-orthogonal multiple access (NOMA) wit successive interference cancellation for future radio access", IEICE Trans.commun., vol.E98-B,NO.3,pp.403-414, (march 2015).

7. E. S. El-Mokadem, A. M. El-Kassas, T. A.Elgarf, H.El-Hennawy "BER Performance Evaluation for the Downlink NOMA System over Different Fading Channels with Different Modulation Schemes", 5th International Conference on Science and Technology (ICST), Yogyakarta, Indonesia, (2019).

8. S. Soni, M. Jain, D. Rawal, N. Sharma, R. Liyanapathirana" Performance Analysis of DF Cooperative-NOMA System with QPSK-BPSK Scheme in the Presence of SIC Errors", IEEE, (December 11-2021).

9. A. Bal, M. R. H. Khan, M. K. Peyal "Performance Comparison Among Hybrid NOMA Schemes Focusing on Outage Performance and Sum Rate Arrangement", 2nd International Conference for Emerging Technology (INCET), pp.1005-1009, May 21-23, 2021, Belgaum, India.

10. C. A. s Zamora, K. Mezquida, G. Combariza , D. Jaramillo-Ramirez , M. Kountouris "Cooperative Successive Interference Cancellation for NOMA in Downlink Cellular Networks".

11. M. M. S, ahin, H. Arslan, "Waveform-Domain NOMA: The Future of Multiple Access". IEEE, (2020).

12. Z. Sheng, X. Su and X. Zhang , "A Novel Power Allocation Method for Nonorthogonal Multiple Access in Cellular Uplink Network", 13th International Conference on Intelligent Environments, (2017).

13. F. A. Khales, G. A. Hodtani, "An Evaluation of the Coverage Region for Downlink Non-orthogonal Multiple Access (NOMA) Based on Power Allocation Factor ",Iran Workshop on Communication and Information Theory (IWCIT)., (2017).

14. F. Al Rabee, K. Davaslioglu, and R. Gitlin, "The Optimum Received Power Levels of Uplink Non- Orthogonal Multiple Access (NOMA) Signals", IEEE, (2017).

15. A. K. Lamba , R. Kumar , S. Sharma" Power allocation for downlink multiuser hybrid NOMAOMA systems: An auction game approach", John Wiley & Sons, Ltd., pp. 1-9, (8 December 2019).

16. J. L. Q. Wang, "Power Allocation Method of Downlink Non-orthogonal Multiple Access System Based on α Fair Utility Function", J Inf Process Syst, Vol.17, No.2, pp.306~317, (April 2021).

17. A. Muhammad , M. S. Vasanthi ,"A Fair Power Allocation Method For Improving Capacity Of Sc-Noma System", - Elementary Education, Vol 20 (Issue 1), pp.2420-2426, (2021).

18. J. A. Oviedo, H. R. Sadjadpour,"A Fair Power Allocation Approach to NOMA in Multi-user SISO Systems", pp.1-12, (2017).

19. J. A. Oviedo, H. R. Sadjadpour,"A New NOMA Approach for Fair Power Allocation ", IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp.1-5, (2016).

A Comparative Analysis of 2D Image Segmentation Based on the Improved Spatial Fuzzy C-Means Algorithm

Boucif BEDDAD[1], Postaire Jack-Gerard[2] and Frantisek JABLONCIK[3]

[1]LTC Laboratory, University of Saida Dr. TAHAR Moulay, Saida, Algeria
[2]Centre de Recherche en Informatique, Signal et Automatique, Université de Lille, France
[3]Department of Mechatronics and Electronics, University of Zilina, Zilina, Slovakia

***Abstract*— Fuzzy C-means clustering is a distinctive classification technique. To improve this FCM approach in terms of precision, a new Improved Spatial Fuzzy C-Means Algorithm ISFCM is developed by integrating local spatial information. This new technique aims to give good results compared to the other methods discussed by modifying and minimizing the used objective function. In order to obtain clustering results more accurately, a better objective function is required. This developed function is based on adding a new spatial data term with respect to each cluster by introducing a conditional variable based on the modified Nagao filter regions. The cluster count problem and cluster center initialization are well demonstrated to speed up the convergence rate and reduce memory consumption and the number of iterations during the execution processing of the ISFCM algorithm.**

***Index Terms*— Cluster Analysis, Digital Image Segmentation, Fuzzy C-means, Spatial Informations, Objective Function, Euclidean Distance Function.**

I. INTRODUCTION

The traditional Fuzzy C-Means (FCM) approach is one of the most common and widely used methods for unsupervised image segmentation. The FCM technique does not fully utilize spatial information to solve the problem of noise sensitivity and intensity inhomogeneity which significantly affect the performance of medical image segmentation [1]. Based on the various challenges encountered, the current research aims at the comparative study of new image segmentation methods namely SFCM, CsFCM, CsKFCM, GKFCM, RSCFCM, KGFCM, and GKWFLICM [10].

The basic concepts of the fuzzy C-Mean clustering method for medical diagnosis and the relevance of fuzzy theory are also discussed and used to solve the problem of noise sensitivity and intensity inhomogeneity which greatly affect the performance of medical image segmentation. Figure 1 show the flowchart of MRI image segmentation [3]. This paper is structured with literature review in second section, third & four sections for proposed method & result, finally a conclusion.

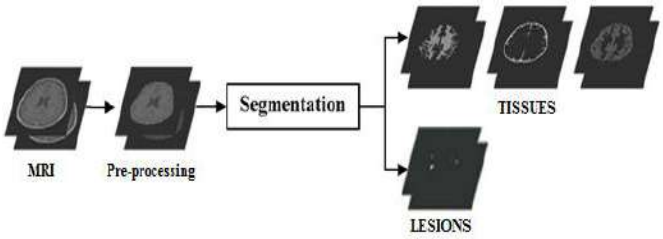


Fig. 1. Brain MRI image segmentation flowchart

II. LITERATURE REVIEW

In this phase, we present a study of new methods of brain MRI segmentation based on the spatial fuzzy c-means algorithm and in the presence of intensity inhomogeneity, significant noise and outliers. The researchers modified and improved the standard FCM method by using the concept of bias field estimation, spatial information, new neighborhood term and additional terms to deal with medical imaging artifacts [2]. Among the new methods presented in the literature, the following algorithms have been studied and discussed: B. Gharnali and S. Alipour [5] presented a new MRI image segmentation approach to solve the problems of the standard FCM algorithm. K.Rehna et al. [6] have created a new method of medical image segmentation where the developed system presents a variant of the c-fuzzy means type algorithm which provides image clustering. Adhikari et al. [4] have developed a conditional and spatial fuzzy C-means clustering algorithm. F. Zhao et al. [7] proposed a new modified version of the fuzzy C-means (FCM) clustering algorithm, the KGFCM approach uses a mean-filtered image and a median-filtered image to define the spatial constraint term. Z.Ji et al. [9] used the negative log posterior as the dissimilarity function, so they introduced a new factor to integrate the bias field estimation model into the fuzzy FCM objective function. K.H. Memon and D. Lee. [8] developed a new variant of FCM clustering that embeds neighborhood information among M-dimensional data, which alleviates the

drawbacks of the standard FCM clustering algorithm (sensitive to noise and outliers, poor performance for different size classes and different densities) and significantly improves clustering performance.

III. PROPOSED METHOD

We propose in this paper, a new spatial and robust FCM framework that incorporates local spatial information and gray level information to segment brain MRI images. The proposed method integrates the average extra filtered image and the median filtered image according to conditions and which can be calculated to replace the neighborhood term by using a mask (5×5) which has a set of 9 sub-windows contained in a neighborhood of (3x3) at each pixel. Figure 2 shows the nine regions of the Nagao filter used in this research.

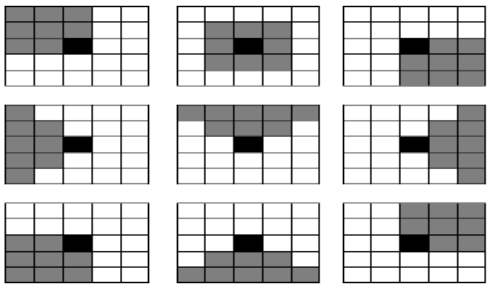


Fig. 2. Representation of the 9 regions of Nagao filter

The main objective is to optimize the segmentation by classification consider the spatial context by minimizing the objective function and integrating its two control parameters $\bar{S}_j$ and $\tilde{S}_j$. The objective function of the ISFCM algorithm using local spatial information and Euclidean distance can be summarized and formulated as follows in eq (1):

$$J_{ISFCM} = \sum_{k=1}^{C}\sum_{j=1}^{N} U_{kj}^{m} \left\| y_j - v_k \right\|^2 + \frac{1}{N_R}\sum_{k=1}^{C}\sum_{j=1}^{N} \bar{S}_j U_{kj}^{m} \sum_{r\in N_j} \left\| y_r - v_k \right\|^2 + \frac{1}{N_R}\sum_{k=1}^{C}\sum_{j=1}^{N} \tilde{S}_j U_{kj}^{m} \sum_{r\in N_j} \left\| y_r - v_k \right\|^2$$

ISFCM Algorithm

Input: the number of clusters C, the image to be segmented converted to gray level, degree of blur m = 2 and the convergent error ε.

1. Read input image and set parameters
2. Choose similarity measure d_{rj} as Euclidean spatial distance
3. Determine the new parameters $\bar{S}_j$et $\tilde{S}_j$
4. Calculate the filtered images mean $\bar{y}_j$ and median $\tilde{y}_j$
5. Initialize cluster centers $v_k^{(0)}$
6. j = 1
7. Repeat

a. j = j + 1

b. Calculate the membership degree matrix $U^{(j)}$ through the cluster centers $v_k^{(j-1)}$

c. Assign each pixel to a specific cluster for which membership is maximal

f. Update the centers $v_k^{(j)}$

8. Until $\left\| v_k^{(j)} - v_k^{(j-1)} \right\| < \varepsilon$

Sortie: Image segmentée à l'aide d'ISFCM

IV. RESULTS AND DISCUSSION

We propose in this paper, a new spatial and robust FCM framework that incorporates local spatial information and gray level information to segment brain MRI images. The proposed method integrates the average extra filtered image and the median filtered image according to conditions and which can be calculated to replace the neighborhood.

TABLE I
SEGMENTATION ACCURACY RESULTS SA(%)
T1 5% NOISE & 20% INH

N°	Methods		T1 5% Noise & 20% INH		
			MB	MG	LCR
1	FCM	MB	86.16	13.85	00.00
		MG	22.31	77.61	00.09
		LCR	02.07	42.93	55.05
2	GFCM	MB	96.94	03.05	00.00
		MG	09.11	90.68	00.22
		LCR	01.58	28.15	70.29
3	GKFCM	MB	97.39	02.61	00.00
		MG	10.36	89.53	00.12
		LCR	01.96	25.46	72.59
4	KGFCM	MB	97.82	02.28	00.00
		MG	09.98	89.96	00.08
		LCR	01.85	25.13	73.02
5	CsKFCM	MB	98.49	01.53	00.00
		MG	07.88	91.94	00.21
		LCR	00.00	10.47	89.54
6	ISFCM	MB	98.83	01.18	00.00
		MG	06.31	93.59	00.13
		LCR	00.00	07.29	92.73

TABLE I I
SEGMENTATION ACCURACY RESULTS SA(%)
T1 7% NOISE & 40% INH

N°	METHODS		T1 7% NOISE & 40% INH		
			MB	MG	LCR
1	FCM	MB	84.65	15.37	00.00
		MG	23.54	76.27	00.21
		LCR	04.61	42.44	52.96
2	GFCM	MB	94.59	05.40	00.00
		MG	13.73	86.08	00.19
		LCR	03.31	30.44	66.28
3	GKFCM	MB	97.41	03.69	00.00
		MG	11.63	88.24	00.14
		LCR	03.01	27.49	69.53

4	KGFCM	MB	96.62	03.39	00.00
		MG	12.44	87.45	00.13
		LCR	2.89	28.38	68.74
5	CsKFCM	MB	96.87	03.14	00.00
		MG	10.08	89.70	00.25
		LCR	00.00	12.22	87.81
6	ISFCM	MB	97.71	02.31	00.00
		MG	08.42	91.47	00.16
		LCR	00.00	09.43	90.59

TABLE I I I
SEGMENTATION ACCURACY RESULTS SA(%)
T1 9% NOISE & 60% INH

N°	Methods		T1 9% Noise & 60% INH		
			MB	MG	LCR
1	FCM	MB	78.76	21.25	00.00
		MG	29.15	70.36	00.51
		LCR	04.35	49.86	45.80
2	GFCM	MB	91.51	08.51	00.00
		MG	14.73	84.77	00.53
		LCR	04.56	29.77	65.68
3	GKFCM	MB	91.25	08.75	00.00
		MG	15.55	84.29	00.15
		LCR	04.66	26.58	68.77
4	KGFCM	MB	90.83	09.19	00.00
		MG	16.02	83.87	00.12
		LCR	4.31	27.36	68.35
5	CsKFCM	MB	93.29	06.73	00.00
		MG	14.35	85.01	00.65
		LCR	00.00	16.19	83.85
6	ISFCM	MB	93.97	06.04	00.00
		MG	10.72	88.80	00.51
		LCR	00.00	13.84	86.17

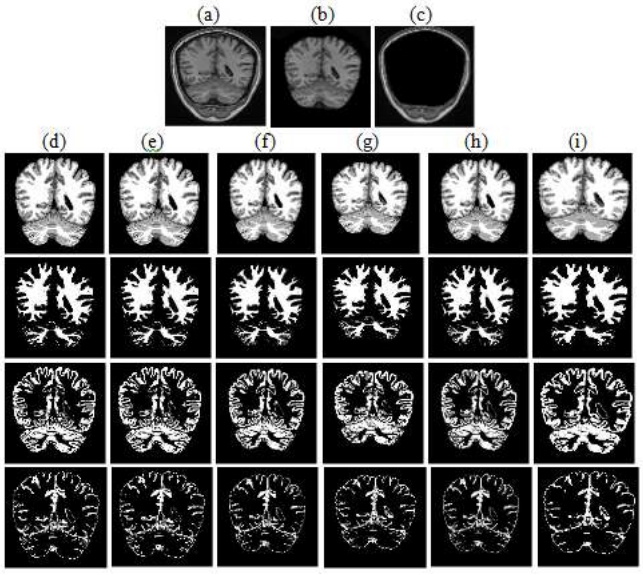


Fig. 3. Comparison of segmentation results on real MRI images: (a) original MRI image, (b) skinning skull, (c) non-brain part, (d) FCM, (e) GFCM, (f) KGFCM, (g) GKFCM (h) CsKFCM, (i) ISFCM

From the comparison of the results obtained in figure 3 and from a clearer point of view, it should be noted that the method developed ISFCM, although still based on an enriched and improved spatial model, nevertheless makes it possible to lead to results which present structural qualities very largely superior to those of the results obtained by the other methods proposed in this thesis, this being explained by the fact that the latter is not based on any preliminary information in particular with regard to the respect for the anatomical properties of the tissues. In the following, we present the results of the SA% segmentation rates by the different algorithms implemented. Some performance metrics like SA are calculated and illustrated in figure 4.

TABLE Iv
SEGMENTATION ACCURACY RESULTS SA(%)

N°	Brain Tissue	# Segmentation Methods					
		FCM	GFCM	GKFCM	KGFCM	CsKFCM	ISFCM
1	MB	85.13	93.17	92.11	94.06	94.30	94.73
2	MG	84.26	86.30	87.14	84.89	90.97	91.19
3	LCR	83.04	85.35	86.48	86.22	89.51	89.74
4	Mean	**84.14**	**88.27**	**88.57**	**88.39**	**91.59**	**91.87**

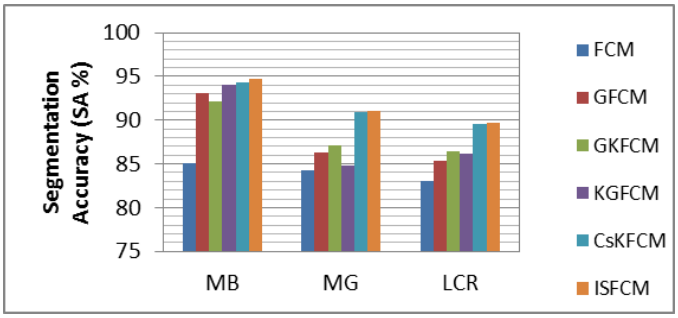


Fig. 4. Graphic representation of Segmentation Accuracy

V. CONCLUSION

The contribution of this research is to provide an improved image segmentation algorithm using the principle of improved fuzzy C-means approach to advance performance with application to medical data analysis. To achieve this objective, an exhaustive study was carried out to analyze the various existing segmentation techniques. The comparative analysis of Enhanced Fuzzy C-Means algorithm with spatial constraints is applied on several medical datasets in terms of optimization. Through a qualitative and quantitative analysis step, the experimental results show that the proposed ISFCM method gives better desired segmentation results, and it is more robust on all medical images tested.

REFERENCES

[1] S. Kamarujjaman, S. Chakraborty, M. Maitra, «An unsupervised modified spatial fuzzy C-mean method for segmentation of brain MR image». Intl. Conference on Advances in Computing, Communications and Informatics, India, pp. 125-131, 2016.

[2] T. Le and K. Gardiner, «A validation method for fuzzy clustering of gene expression data».Intl' Conf. on Bioinformatics and Computational, Las Vegas, Vol. 1, pp. 23-29, 2011.
[3] S. Madhukumar, N. Santhiyakumari, «Evaluation of k-Means and fuzzy C-means segmentation on MR images of brain». Egyptian Society of Radiology and Nuclear Medicine, vol. 46, no. 2, pp. 475-479, 2015.
[4] S. K. Adhikaria, J. K. Singb, D. K. Basub, M. Nasipuriba, «Conditional spatial fuzzy C-means clustering algorithm for segmentation of MRI images». Applied Soft Computing, vol. 34, pp. 758–769, 2015.
[5] B. Gharnali, S. Alipour, «MRI Image Segmentation Using Conditional Spatial FCM Based on Kernel-Induced Distance Measure». Engineering, Technology & Applied Science Research, vol. 8, no. 3, pp. 2985-2990, 2018.
[6] K. Rehna, T. Ciza, M. Abdul Rahiman, «Gaussian Kernel Based Fuzzy C-Means Clustering Algorithm for Image Segmentation». Computer Science &Information Technology, vol. 6, pp. 47–56, 2016.
[7] F. Zhao, L. Jiao, H. Liu, «Kernel generalized fuzzy c-means clustering with spatial information for image segmentation». Digital Signal Processing, vol. 23, pp. 184–199, 2013.
[8] K. H. Memona, D. Lee, «Generalised kernel weighted fuzzy C-means clustering algorithm with local information». Fuzzy Sets and Systems, vol. 340, pp. 91-108, 2018.
[9] Z. Ji et al., «Robust spatially constrained fuzzy c-means algorithm for brain MR image segmentation». Pattern Recognition, vol. 47, No. 7, pp. 2454-2466, 2014.
[10] S. Muhammad, H. Ayyaz, A. J. Muhammad, T. S. Choi, «Fuzzy-based hybrid filter for Rician noise removal». Signal, Image and Video Processing, vol.10, no.2, pp. 215-224, 2016.
[11] Li, C. Gore, J. Davatzikos, C.,« Multiplicative intrinsic component optimization (MICO) for MRI bias field estimation and tissue segmentation». Magn. Reson. Imaging, Vol 32, pp. 913–923, 2014.
[12] A. Madabhushi et J. K. Udupa, «Interplay between intensity standardization and inhomogeneity correction in MR image processing». IEEE Transactions on Medical Imaging, vol. 24, pp. 561-576, 2005.
[13] M. K. Kowar, S. Yadav, «Brain Tumour Detection and Segmentation Using Histogram Thresholding». International Journal of Engineering and Advanced Technology, vol. 1, no. 4, 2012.
[14] S.R. Kannan, S. Ramathilagam, A. Sathya, R. Pandiyarajan, «Effective fuzzy c-means based kernel function in segmenting medical images», Comput. Biol. Med, vol. 40, pp. 572–579, 2010.
[15] S. Krinidis, V. Chatzis, «A robust fuzzy local information C-means clustering algorithm». IEEE Trans. Image Process, vol. 19, pp. 1328–1337, 2010.

Applications' Quality of Experience Enhancement basing on SDN visibility

Hadjira BELAIDI
Signals and Systems Laboratory,
Institute of Electrical and Electronic Engineering, University M'Hamed BOUGARA of Boumerdes
Boumerdes, Algeria
hadjira983@yahoo.fr, ha.belaidi@univ-boumerdes.dz,
ORCID: 0000-0003-2424-626X

Nabil ALIOUAT
Signals and Systems Laboratory,
Institute of Electrical and Electronic Engineering, University M'Hamed BOUGARA of Boumerdes
Boumerdes, Algeria
nabilaliouat2015@gmail.com,
ORCID: 0000-0002-6835-3561

Abdelghani GRICHE
Signals and Systems Laboratory,
Institute of Electrical and Electronic Engineering, University M'Hamed BOUGARA of Boumerdes
Boumerdes, Algeria
gricheabdelghani@gmail.com,
ORCID: 0000-0002-5613-2788

Abstract—**The amount of data traffic has increased dramatically in recent years. This development is inversely-impacts the capacity of traditional Internet Service Providers' (ISPs') networks; leading to their exhaustion and, as a result, a worsening in the Quality of Experience (QoE) observed by consumers.**

This paper attempts to address this issue in the context of video streaming. To accomplish this, the Software-Defined Networking (SDN) controller's high visibility is explored to create a telemetry service that collects network condition information and sends it to the video streaming application, which allows it to adjust its dispersion parameters proportionally. The test results demonstrate the effectiveness of our solution and its impact on the user's QoE.

Keywords— SDN, QoS, QoE, Telemetry, Streaming.

I. INTRODUCTION

Traditional Internet Protocol (IP) based networks are essentially founded on the best-effort approach for transmitting packets to their final destination, with no guarantee of end-to-end quality of service (QoS). This architecture is appropriate for delivering simple data where reliability is the primary need. In multimedia transmission, however, rapid delivery is preferred over reliability. Streaming media applications may have time constraints that cannot be ensured [1]. As a result, it is preferable that the network architecture supports specific methods for providing QoS for multimedia traffic. The Internet Engineering Task Force (IETF) investigated many QoS architectures, including IntServ and Diffserv [2], [3], to this goal. The disadvantage of its solutions is that they are designed on a fully dispersed architecture with little visibility into total network resources. It is harder to manage or integrate new services and technologies in such a distributed, multi-vendor, multi-protocol, non-programmable, human-dependent environment.

The Software Defined Network (SDN) has introduced a paradigm change in network architecture in response to the architectural rigidity of traditional networks [4], [5]. Instead of a distributed control architecture [6], SDN consolidates all control into a single logical entity known as a 'controller,' which serves as the network's brain. Furthermore, the controller has a worldwide perspective of the network, allowing it to collect data on available resources and equipment performance. The considerable visibility and control that SDN provides at the network level can enable new techniques to ensure that applications have the QoS required to give a positive user experience [7], [8]. Hence, it was thus intriguing to study how applications and the network may interact to improve service quality. In other words, using the information offered to the applications, the latter could modify their decisions according to the actual network conditions.

Therefore, a new QoS strategy based on the SDN idea is implemented in this work; where, the right delivery of multimedia streams while protecting them from packet losses is assured. Moreover, our primary concern is boosting their Quality of Experience (QoE).

In order to validate our technique, the example of a video streaming application is tested, which, with greater knowledge of the network's characteristics, it tries to improve its decisions by adjusting its parameters (Bit Rate, Frame per Second (FPS), etc.); then, judge the QoS for the user.

II. CONTROLLER OVERVIEW

A. SDN and OpenFlow

SDN is a comparatively recent network programmable paradigm supported by the ONF (Open Networking Foundation) [9]–[11]. This novel paradigm alters network design and management by introducing an interface that separates the control plane from the data plane [9]. Where, a monitoring program, named a controller, with a global perspective of the overall infrastructure is charged with making decisions while the hardware (router, switch, etc…) is just charged with routing packets to their destinations[12], [13]. OpenFlow is the protocol used for communication between the infrastructure layer and the control layer, it was originally proposed and implemented by Stanford University, and subsequently standardized by the NFB [14], [15]. There have been several controllers built, the most of which are open source and implement the OpenFlow protocol. Table I below covers the most prevalent SDN controllers.

TABLE I. PROMINENT SDN CONTROLLERS

Controller	***Organization***	***Language***	***Functionalities***
NOX [16]	Nicira	C++	The first OpenFlow controller
POX [17]	Nicira	Python	Improve NOX performance
Ryu [18]	NTT, OSRG group	Python	Supports Openstack
Beacon [19]	Stanford	Java	Based on Multithreading

Floodlight [20]	Big Switch	Java	Tested with OpenFlow switches
Opendaylight [21]	Linux Foundation	Java	Supports OSGi and REST API

This work must prove that it is feasible to adjust applications running on network nodes to give a better user experience using data collected from an SDN controller.

B. *Relationship between QoS and QoE*

In the context of a service or network provider, the QoS factors (delay, packet loss rate, etc...) will be used to determi ne compliance with the QoS promises. [22].

Typically, the consumer is unconcerned about how the service or application is implemented. Rather, he wishes to compare the same service provided by other companies. Disparate the QoS factors listed previously, the quality of applications is valued based on other criteria linked to the functioning of the application, such as the image's resolution, the quality of the utilized colors, and so on for the case of multimedia applications. Because these settings are nearer to the end user, where the user quality of experience QoE' is referred rather than network QoS' [22]. Hence, QoE is defined as the users' particular judgements established from their own lived experience, on the quality of the entire application or service provided by vendors or internet service providers. QoE can be measured using even the objective methods or subjective methods [23].

C. *The Proposed Solution Description*

The proposed solution includes two main steps to accomplish our goal: initially the network's QoS metrics must be measured and organized using a telemetry service established for this purpose; then, in the second step the application is permitted to check and infer these metrics to adjust its functioning factors. The flowchart shown in Fig. 1 depicts the operation and the interactions of our platform's components.

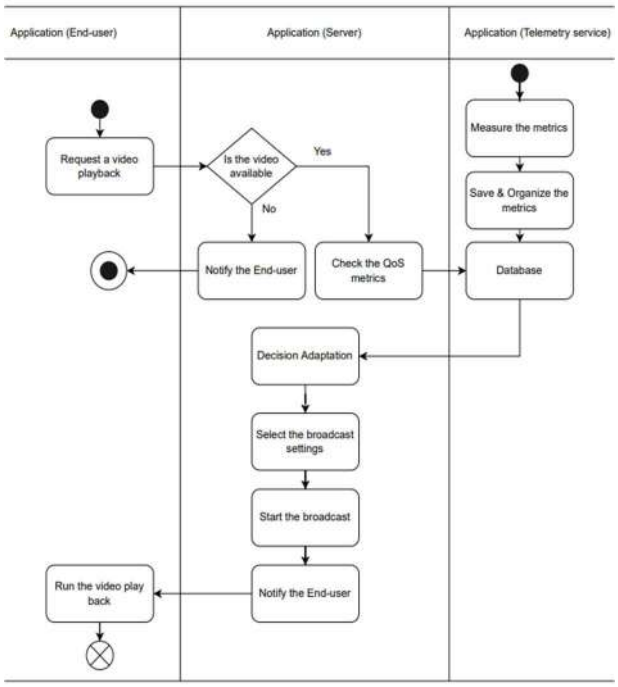


Fig. 1. The interactions between the platform's components.

The study accomplished by [24] is exploited in our work to construct a telemetry service integrated with the controller. This service computes telemetry values like time and packet loss rate; then, organizes and stocks them in a local database. Knowing that the operation of an SDN network is based on the constant exchange of messages between the controller and the switches in order to manage traffic and verify the state of Openflow switches; these messages are employed in our approach to track the delay and packet losses' rate generated in our platform, as illustrated in Fig. 2.

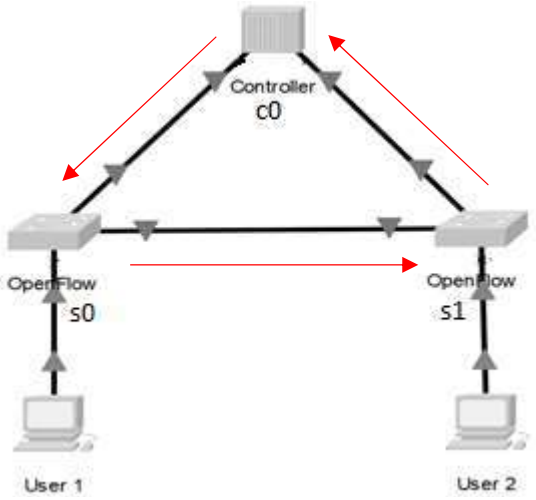


Fig. 2. Test platform.

D. *Delay determination*

Timestamps are placed in specific packages built for this purpose to measure the delay ($T_{latency}$) in our development platform (monitoring packets). Initially, the controller c0 sends a packet with a specified ID and the specified timestamp. This packet will traverse s0 and s1 before returning to the origin point c0. When it is received, the timestamp is retrieved and subtracted from the receipt time to calculate T_{total} the time required to travel the path [c0 →s0→ s1→ c0]. The monitoring packet's journey is depicted in Fig. 2 by red arrows.

However, just the transmission time between s0 and s1 is requested; thus, the transmission times between [c0, s0] and [c0, s1] must be subtracted. These are retrieved owing to the statistical messages that are constantly transmitted between the controller and the switches. Thus, the equation to deduce the latency is defined by Eq. [1]:

$$T_{latency} = T_{total} - (T_{[c0 \rightarrow s0]} + T_{[c0 \rightarrow s1]}) \qquad (1)$$

E. *Packet-loss's rate determination*

Packet-loss's rate is expressed as a proportion of packets lost compared to packets sent. Therefore, the same method used to quantify latency will be used to calculate the rate of lost packets. As a result, the OpenFlow messages sent between the controller and the switches are exploited to calculate the loss rate in the path linking [s0 → s1].

To be more specific, the controller will issue a *Flow-Stats-Request* to s0 and s1. When the controller receives the

reply from s0, it stores the amount of packets transmitted for each stream in *sent_packets*. When the controller receives a reply from s1, it stores the amount in packets received for the same stream in a variable stated as *received_packets*. The difference is the amount of dropped packets. Hence, the loss rate equation can be expressed by Eq. [2] as follow:

$$Lost\_rate = (sent\_packets - received\_packets) * 100 / sent\_packets \quad (2)$$

F. Bandwidth determination

"*iPerf*" , a tool developed by the National Laboratory for Applied Network Research and tailored for client-server systems [25], is used to estimate the bandwidth. To run *iPerf*, the first machine launches it in 'server mode,' and the second in 'client mode', respectively. iPerf is used in our situation to measure the bandwidth of the link between the streaming client and server.

The purpose of the alteration is to improve the application performance in general. In our work, this enables the video streaming application placed on the server to automatically make decisions to alter streaming parameters such as bitrate or frame rate to offer appropriate data streams that match the collected telemetry statistics.

Before starting the broadcast, the application first checks the telemetry values. Then, it determines whether the values retrieved are compatible with the stream's criteria (minimum configuration for its correct operation). Finally, the program modifies its broadcast parameters (as explained in Fig. 1).

III. CONTROLLER CONFIGURATION

A. Controller settings

The controller is initialized via the command "pox.py" which launches the controller and sets it into the up state; it takes as a parameter the "controller.py" script, which allows the controller to determine and organize the telemetry values. After launching the controller, it turns to the idle state until the OpenFlow switches are connected. In addition to, Mininet terminal that is used to initiate the topology that will be connected to the POX controller by calling the "topo.py" script.

Following the launch of the topology, the switches will be connected with the controller, which will begin evaluating the latency and packet loss between s0 and s1 every two seconds; then, displaying and saving the results in a 'csv' format file.

B. Video streaming application

The two users, User 1 and User 2, are used after the test platform is launched. User 1 acts as the streaming server that will broadcast the video, while User 2 acts as the client that will play the broadcast video. The client instantly informs the server that it wishes to play a video after connecting.

Once the server is notified, it consults the "csv" file to retrieve the QoS metrics generated by the controller during the past 20 seconds; it then uses the *iPerf* tool to compute the bandwidth of the link between itself and the client. The application on the server can alter the broadcast parameters (bite rate, resolution and FPS) based on these network factors (delay, bandwidth, and loss rate).

In this paper, the values extracted from [26] and [27] are used to specify the broadcast parameters of our application under different network situations. The values of the QoS metrics (bandwidth and latency) as well as the broadcast parameters (biterate and FPS) necessary for each broadcast quality (resolution) are shown in Table II.

TABLE II. QOS PARAMETERS REQUIRED FOR EACH BROADCAST QUALITY.

Resolution	*FPS*	*Bitrate*	*Bandwidth*	*Delay required*
4K UHD 2160p	60	20,000-51,000Kbps	50mbps	< 150ms
4K UHD 2160P	30	13,000-34,000Kbps	30mbps	
FHD 1080p	60	4,500-9,000Kbps	9mbps	
FHD 1080p	30	3,000-6,000Kbps	6mbps	
HD 720p	60	2,250-6,000Kbps	4mbps	
HD 720p	30	1,500-4,000Kbps	2mbps	
HD 480p	30	500-2,000Kbps	< 2mbps	

Consequently, the diffusion parameters are updated based on the decision retrieved from the execution of the pseudo-code displayed below (see Fig. 3).

```
Broadcast Procedure (integer:BW ,integer:Delay){
Variable:
Integer:BitRate,FPS;
Begin:
if BW<=2 and Delay<=150 then
    FPS<---30;BitRate<---2000;//resolution=480p30
Elseif
    if BW>=2 and BW<=4 and Delay<=150 then
        FPS<---30;BitRate<---4000;//resolution=720p30
    Elseif
        if BW>=4 and BW<=6 and Delay<=150 then
            FPS<---60;BitRate<---6000;//resolution=720p60
        Elseif
            if BW>=6 and BW<=9 and Delay<=150 then
              FPS<---30;BitRate<---6000;//resolution=1080p30
              Elseif
                if BW>=9 and BW<=30 and Delay<=150 then
                    FPS<---60;BitRate<---9000;//resolution=1080p60
                    Elseif
                        if BW>=30 and BW<=50 and Delay<=150 then
                            FPS<---30;BitRate<---34000;//resolution=4Kp30
                            Elseif
                                if BW>=50 and Delay<=150 then
                                    FPS<---60;BitRate<---51000;//resolution=4Kp60
                                End if
                            End if
                End if
            End if
        End if
    End if
End if
//Launch of the broadcast
Excuting the command(vlc -vvv v412:///dev/video0--sout'#transcode{vcodec=mjpg,
                    mux=mpjpeg, dst=10.0.0.1:8080}')
End

Main () {
    Variable:
    Integer:bw,delay;
Begin
        //Retrieve the QoS parameters calculated by the telemetry service
            bw= get_bandwidth(BW_CSV);
            delay=get_delay(Delay_CSV);
        //Call of the procedure and launch of the broadcast
        Broadcast(bw,delay) ;
End
}}
```

Fig. 3. Streaming resolution selection algorithm at the streaming server application level

IV. RESULTS AND DISCUSSION

Live streaming is chosen as a case study to demonstrate the efficiency of our approach. Knowing that network QoS characteristics such as latency and bandwidth have an impact on this latter; hence, Mininet is used to change these parameters in order to mimic different network situations. The application will then be modified in response to these conditions.

Python script telemetry "grph.py" is executed to launch the Matplotlib tool which help us to display the fluctuations of the metrics (delay and packet loss) measured by the controller in real time. Fig. 4 illustrates the display of packet loss rate vs time; the rate changes between 6 and 34 percent.

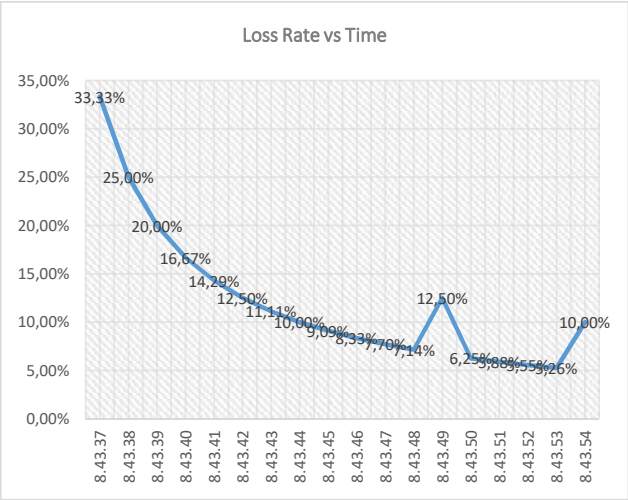


Fig. 4. Display of the metrics fluctuations (delay and packet loss).

In terms of comparison, considering that the application makes no adaptation decisions despite network circumstances in the first scenario. Then, the adaptability feature is triggered in the second case, to test the efficiency of the proposed approach.

A. 1st scenario: without adaptation

In this case, telemetry values are not checked; therefore, no decision is made on the manner to deliver the content. Diffusion is initiated with conventional parameters without specifying the biterate (resolution) or the FPS. The application is tested under various network conditions (good and worse cases).

1) Estimating Good Network conditions:

Mininet is used to simulate optimal network circumstances with short delay (20ms) and adequate bandwidth (20mbps). The results of streaming are presented in Fig. 5.

It can be observed that although the accessibility of resources, the server application provides media with default settings; it does not take advantage of network resource availability to give the greatest possible quality for the user.

2) Estimating Bad Network conditions:

Low settings for bandwidth and delay are selected (delay = 200ms, bandwidth = 1mbps) to mimic a high demand on network resources such as numerous programs requesting bandwidth and heavy traffic; subsequently, streaming starts. Fig. 6 depicts the achieved results.

It is clear that the delivery begins with high settings in comparison to the available resources because the program always uses standard values to deliver the content. This causes instability in streaming video playback and frequent interruptions.

According to these tests, we can infer that the application cannot give a decent quality of experience to the end user because it does not adjust to network conditions. Indeed, it has been deduced that whether there are copious or limited resources, the application always employs standard delivery settings, thus it only gives an adequate service when the conditions are favorable.

Fig. 5. Broadcast under good network conditions.

Fig. 6. Broadcast under poor network conditions.

B. 2nd scenario: Applying our approach

In the second case, our scaling approach is used to address the issue from the prior scenario.

1) Estimating Good Network conditions:

Our approach is tested first by estimating satisfactory network conditions. Mininet is used to emulate optimal network circumstances with short delay (20ms) and adequate capacity (20Mbps). Then our streaming application is started as illustrated in Fig. 7.

After viewing the network status data provided by the telemetry service, it is verified that the streaming application selects optimal streaming settings to deliver the best quality of user experience.

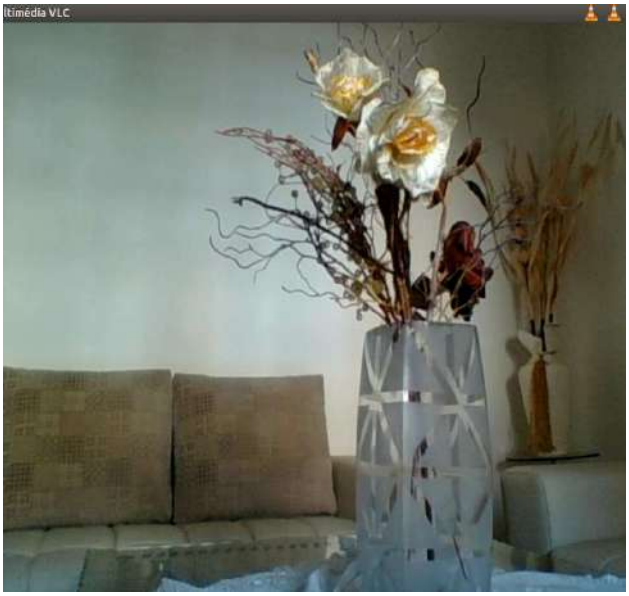


Fig. 7. Broadcast with Adaptation under good conditions

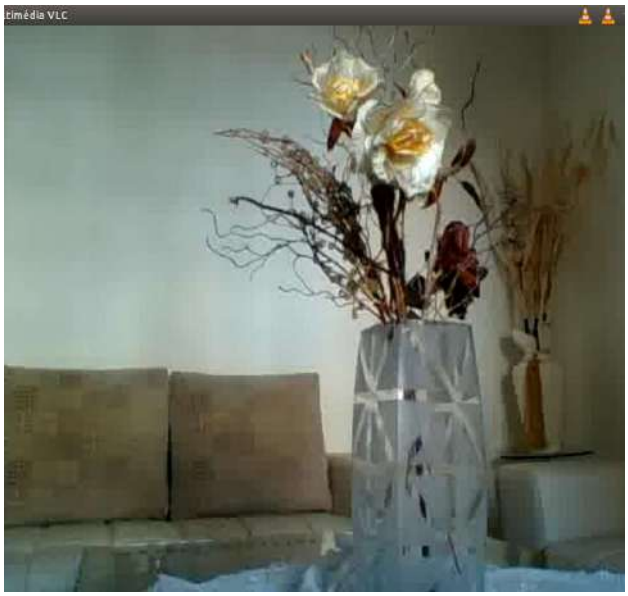


Fig. 8. Braodcast with Adaptation under poor conditions

2) Estimating Bad Network conditions:

A delay of 200ms and 1Mbps bandwidth are used to emulate poor network settings; then, our streaming application is launched as illustrated in Fig. 8.

The application can make decisions based on past knowledge of the network situation to give the greatest possible quality of experience. This highlights the efficiency of our approach for application flexibility based on information transmitted by the telemetry service implemented at the SDN controller level.

The testing of our solution allowed us to prove its effectiveness and highlight the benefits of visibility that the controller offers in an SDN platform to improve the quality of user experience.

V. CONCLUSION

This work stated the way the applications and the networks can cooperate with one another to enhance the service quality for audiovisual flows, using the video streaming application as a case study. Indeed, an enhancement to the video streaming program is recommended by providing it with the ability to adapt to network changes and adjust its broadcast parameters.

To demonstrate the efficiency of our method, tests were developed under various network situations. The obtained results permitted us to conclude that with a better understanding of network conditions, our video streaming application can make better decisions by adapting a more appropriate transcoding and bitrate, resulting in a higher quality of user experience.

The proposed solution can be modified to enable the application to alter its settings during the broadcast, and then the SDN controller can be used to provide QoS processes at the infrastructure layer.

ACKNOWLEDGMENT

We express our acknowledgement for DGRSDT "Direction Générale de la Recherche Scientifique et du Développement Technologique" for its support.

REFERENCES

[1] M. Jalil Piran et al., 'Multimedia communication over cognitive radio networks from QoS/QoE perspective: A comprehensive survey', J. Netw. Comput. Appl., vol. 172, p. 102759, Dec. 2020, doi: 10.1016/j.jnca.2020.102759.

[2] K. Zhigalov, N. A. Skorikova, and I. M. Daudov, 'Interaction of models and methods of providing QoS in networks', J. Phys. Conf. Ser., vol. 1582, no. 1, p. 012095, Jul. 2020, doi: 10.1088/1742-6596/1582/1/012095.

[3] Aprianto Budiman, M. Ficky Duskarnaen, and Hamidillah Ajie, 'ANALISIS QUALITY OF SERVICE (QOS) PADA JARINGAN INTERNET SMK NEGERI 7 JAKARTA', PINTER J. Pendidik. Tek. Inform. Dan Komput., vol. 4, no. 2, pp. 32–36, Dec. 2020, doi: 10.21009/pinter.4.2.6.

[4] S. Scott-Hayward, S. Natarajan, and S. Sezer, 'A Survey of Security in Software Defined Networks', IEEE Commun. Surv. Tutor., vol. 18, no. 1, pp. 623–654, 2016, doi: 10.1109/COMST.2015.2453114.

[5] R. Masoudi and A. Ghaffari, 'Software defined networks: A survey', J. Netw. Comput. Appl., vol. 67, pp. 1–25, May 2016, doi: 10.1016/j.jnca.2016.03.016.

[6] P. Daoutidis, W. Tang, and S. S. Jogwar, 'Decomposing complex plants for distributed control: Perspectives from network theory', Comput. Chem. Eng., vol. 114, pp. 43–51, Jun. 2018, doi: 10.1016/j.compchemeng.2017.10.015.

[7] M. Usman and J. Kim, 'SmartX Multi-View Visibility Framework for unified monitoring of SDN-enabled multisite clouds', Trans. Emerg. Telecommun. Technol., vol. 33, no. 8, Aug. 2022, doi: 10.1002/ett.3819.

[8] M. Alsaeedi, M. M. Mohamad, and A. A. Al-Roubaiey, 'Toward Adaptive and Scalable OpenFlow-SDN Flow Control: A Survey', IEEE Access, vol. 7, pp. 107346–107379, 2019, doi: 10.1109/ACCESS.2019.2932422.

[9] P. Goransson and C. Black, 'Alternative Definitions of SDN', in Software Defined Networks, Elsevier, 2014, pp. 119–144. doi: 10.1016/B978-0-12-416675-2.00006-1.

[10] F. H. P. Fitzek, P. Seeling, T. Höschele, and B. Jacobfeuerborn, 'On the need of computing in future communication networks', in

Computing in Communication Networks, Elsevier, 2020, pp. 3–45. doi: 10.1016/B978-0-12-820488-7.00011-6.

[11] D. Espinel Sarmiento, A. Lebre, L. Nussbaum, and A. Chari, 'Decentralized SDN Control Plane for a Distributed Cloud-Edge Infrastructure: A Survey', IEEE Commun. Surv. Tutor., vol. 23, no. 1, pp. 256–281, 2021, doi: 10.1109/COMST.2021.3050297.

[12] L. Zhu et al., 'SDN Controllers: A Comprehensive Analysis and Performance Evaluation Study', ACM Comput. Surv., vol. 53, no. 6, pp. 1–40, Nov. 2021, doi: 10.1145/3421764.

[13] F. Bannour, S. Souihi, and A. Mellouk, 'Distributed SDN Control: Survey, Taxonomy, and Challenges', IEEE Commun. Surv. Tutor., vol. 20, no. 1, pp. 333–354, 2018, doi: 10.1109/COMST.2017.2782482.

[14] A. Dumka, 'Smart information technology for universal healthcare', in Healthcare Data Analytics and Management, Elsevier, 2019, pp. 211–226. doi: 10.1016/B978-0-12-815368-0.00008-7.

[15] J. Rischke and H. Salah, 'Software-defined networks', in Computing in Communication Networks, Elsevier, 2020, pp. 107–118. doi: 10.1016/B978-0-12-820488-7.00018-9.

[16] S. Shin and G. Gu, 'Attacking software-defined networks: a first feasibility study', in Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking - HotSDN '13, Hong Kong, China, 2013, p. 165. doi: 10.1145/2491185.2491220.

[17] H. M. Noman and M. N. Jasim, 'POX Controller and Open Flow Performance Evaluation in Software Defined Networks (SDN) Using Mininet Emulator', IOP Conf. Ser. Mater. Sci. Eng., vol. 881, no. 1, p. 012102, Jul. 2020, doi: 10.1088/1757-899X/881/1/012102.

[18] Md. T. Islam, N. Islam, and Md. A. Refat, 'Node to Node Performance Evaluation through RYU SDN Controller', Wirel. Pers. Commun., vol. 112, no. 1, pp. 555–570, May 2020, doi: 10.1007/s11277-020-07060-4.

[19] D. Erickson, 'The beacon openflow controller', in Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking - HotSDN '13, Hong Kong, China, 2013, p. 13. doi: 10.1145/2491185.2491189.

[20] P. Göransson, C. Black, and T. Culver, 'SDN Applications', in Software Defined Networks, Elsevier, 2017, pp. 271–301. doi: 10.1016/B978-0-12-804555-8.00012-0.

[21] G. Yao, J. Bi, and P. Xiao, 'Source address validation solution with OpenFlow/NOX architecture', in 2011 19th IEEE International Conference on Network Protocols, Vancouver, AB, Canada, Oct. 2011, pp. 7–12. doi: 10.1109/ICNP.2011.6089085.

[22] E. Gallo, M. Siller, and J. Woods, 'An Ontology for the Quality of Experience framework', in 2007 IEEE International Conference on Systems, Man and Cybernetics, Montreal, QC, Canada, Oct. 2007, pp. 1540–1544. doi: 10.1109/ICSMC.2007.4414109.

[23] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, 'Study of Subjective and Objective Quality Assessment of Video', IEEE Trans. Image Process., vol. 19, no. 6, pp. 1427–1441, Jun. 2010, doi: 10.1109/TIP.2010.2042111.

[24] V. Jara and Y. Shayan, 'Latency Measurement in an SDN Network Using a POX Controller', in 2018 IEEE Canadian Conference on Electrical & Computer Engineering (CCECE), Quebec, QC, Canada, May 2018, pp. 1–5. doi: 10.1109/CCECE.2018.8447647.

[25] Y. Li, H. Qi, G. Lu, F. Jin, Y. Guo, and X. Lu, 'Understanding hot interconnects with an extensive benchmark survey', BenchCouncil Trans. Benchmarks Stand. Eval., vol. 2, no. 3, p. 100074, Jul. 2022, doi: 10.1016/j.tbench.2022.100074.

[26] '"Upload Speeds for 4K Live Streaming". (Mars,2019). https://www.boxcast.com/blog/internet-speeds-for-4k-live-streaming. [cf. p. 63]'.

[27] 'Toni Farley Yan Chen et Nong Ye. "QoS Requirements of Network Applications". Arizona State University, (Jan,2004). [cf. p. 63]'.

Study of reconfigurable Fractal Antennas for Wireless Applications

Benammar Abdelkader
dept. of electrical engienering
Mustapha Stambouli University,Mascara,Algeria
abdelkader.benammar@univ-mascara.dz

Bendaoudi Amina
dept. of electronics
Lepo Lab,UDL-SBA
Sidi belabess, Algeria
aminabendaoudi@yahoo.fr

Mahdjoub Zoubir
dept. of electronics
Lepo Lab,UDL-SBA
Sidi belabess, Algeria
mahdjoubdz@yahoo.fr

***Abstract*— The demand for antennas has exploded due to the rapid expansion of wireless communication systemes, In modern communication links, antennas with enhanced gain, increased bandwidth, small size, and low profile are typically selected.One of methods used to resume as needed is the method based On fractal antennas,In this work different Shapes of Fractal antennas are expected from conventional rectangular patch antenna to T-rectangle Fractal antenna with Sierpinski Carpet in the second iteration and pentagon shapes along the patch surface , with lumped element in the central rectangle in the ON state ,these antennas are designed on Rogers RT 6010 substrate with dielectric constant of 10.2 and thickness of 2.5mm.In addition of the reduction of the antenna size ,fractal antennas designed outperforms the traditional patch in bandwidth and return loss , return loss of -31,44 .Antennas are designed and simulated by CST studio suite 2021**

Keywords- Fractal antennas;Wideband;returnloss;Sierpinski Carpet

I.Introduction

Antennas serve as the wireless communication systemes'eyes and ears' [1], Since antennas are a fundamental component of the entire wireless communication system, modern wireless communication systems place more emphasis on the design of antennas for communication technology advancement. The primary function of an antenna is to act as a transducer between the energy delivered by a transmission line and the energy radiated into space. An antenna is effectively used for electromagnetic wave radiation [2-5]Between an electromagnetic energy source and unoccupied space, it functions as a matching system. Utilizing antennas has only one objective: to enhance this matching. Once the matching is complete, the antenna produces high gain and radiates effectively. In addition to gain, there are numerous other characteristics that influence antenna design. The following is a list of some antenna properties:

"Antenna Pattern," which depicts the strength of the field in various directions, The ratio of power radiated to total power is known as "radiation efficiency." Impedance matching refers to the antenna's input impedance for optimal power transfer. The aforementioned characteristics are essentially constant regardless of the antenna's bandwidth or frequency range,[6-9] The most important requirements nowadays for conformal antennas that may be directly incorporated into a range of microwave circuits are their compactness, light weight, and low profile. They play a significant role in the advancement of printed antennas. They are more commonly usable than the conventional ones since they are inexpensive and simple to fabricate on printed circuit boards (PCB). When developing a microstrip antenna, any geometry shape and size can be taken into consideration. A microstrip antenna has an infinite ground plane on one face and a radiating patch on the other on a dielectric substrate with very little thickness. Any shape, including square, circular, triangular, semi-circular, sectorial, and annular ring shapes, is possible for a patch.

Patch antennas radiate mainly as a result of the fringing fields produced between the patch edge and the ground plane.

Better radiation and greater efficiency are primarily preferred due to the high bandwidth that a low dielectric constant and thick substrate give. Consequently, this design results in a huge antenna size[10], The solution to this issue is to utilize a substrate with a high dielectric constant, which reduces efficiency and narrows bandwidth. As a result, a trade-off between the antenna size and performance must be made[11].

The patch can be constructed in any shape and is typically composed of conductive materials like gold, copper, or silver. The patch region on the substrate and the access conductive substance are typically obtained using the photo-etching process[12]

For an antenna to be effective, it must first and foremost be compact with increased bandwidth. The size of an antenna can be reduced via a variety of methods, including[13];

- utilizing a substrate with high permittivity,
- when reactive or resistive loading is used
- Improved form results in longer electric current
- Notches on certain patches or fractals.

II METHODOLOGY

Begin with a conventional patch antenna, a combination of T-rectangular patch antenna and a Sierpinski Carpet Fractal in the second iteration ,This antenna was designed on Rogers RT Duroid 6010 substrate .Table1 shows specification and properties of this antenna.

TABLE1 Dimensions of conventional patch antenna

Parameters	Values
Frequency	**5.2Ghz**
Substrate length	**27.87 mm**
Substrate Width	**33.3 mm**
Patch length	**8.15mm**

Patch Width	12.19mm
Substrate Thickness	2.5mm
Dielectric constant	10.2
Feed length	2.345 mm
Feed Width	9.192 mm
Inset Distance	2.153 mm
Inset Gap	1.173 mm

The patch antenna was designed by calculating all the required dimensions.the equations used were listed[14]

i)Width of patch

$$Wp = \frac{C}{2f\sqrt{\frac{(\varepsilon_r+1)}{2}}} \quad (1)$$

Where

C:Spead of light

f:center frequency

ε_r :Dielectric constant

ii)Effective dielectric constant for ε_r=2.2,

$$\varepsilon_{reff} = \frac{\varepsilon_r+1}{2} + \frac{\varepsilon_{reff}-1}{2}\left[1 + 12\frac{h}{Wp}\right]^{1/2} \quad (2)$$

Where h: the thickness of the substrate

Wp: the Width of the patch

iii)Length extension

$$\Delta L = 0.412h\frac{(\varepsilon_{reff}+0.3)\left(\frac{Wp}{h}+0.264\right)}{(\varepsilon_{reff}-0.258)\left(\frac{Wp}{h}+0.8\right)} \quad (3)$$

iv) Length of the patch ,

$$Lp = \frac{C}{2f\sqrt{\varepsilon_{reff}}} - 2\Delta L \quad (4)$$

v) Inset-fed depth ,

$$Y = \frac{10^{-4}}{2} Lp\begin{pmatrix} 0.001699\varepsilon_r{}^7 & +0.13761\varepsilon_r{}^6 & -6.1783\varepsilon_r{}^5 \\ +93.187\varepsilon_r{}^4 & 682.69\varepsilon_r{}^3 & +2561.9\varepsilon_r{}^2 - \\ 4043\varepsilon_r{}^1 & +6697 & -0 \end{pmatrix} \quad (5)$$

Figure 1 shows the conventional patch antenna

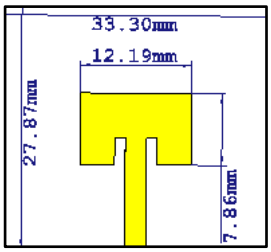


Figure1. Conventional patch antenna

Figures of different models are listed below in Figure2

In the figure2a we subtract a rectangle and add a lumped element.

In Figure 2b we add pentagons in the vicinity of the subtracted rectangle of the last step

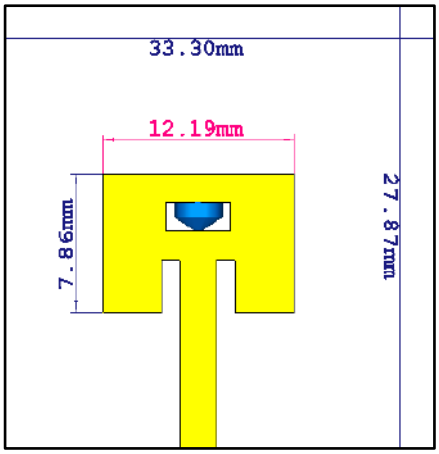


Figure 2a

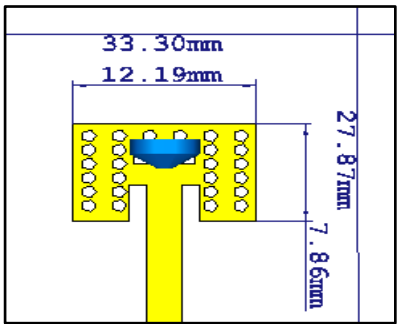


Figure 2b

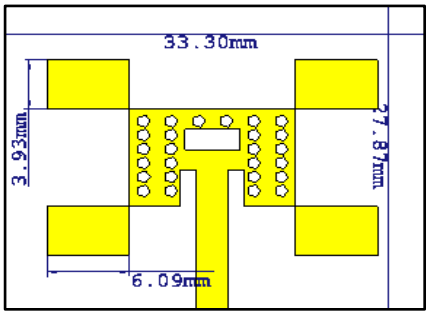


Figure2c

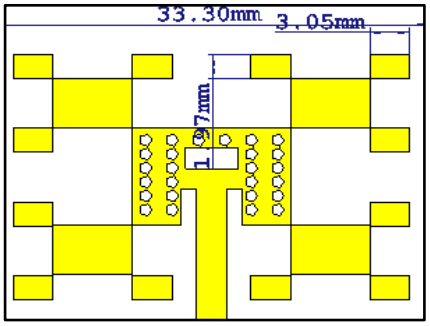


Figure2d

Figure 2a : Sierpinski Carpet in the First iteration

Figure 2b: Sierpinski Carpet with pentagons

Figure 2c: Sierpinski Carpet and T-rectangle First iteration antenna

Figure 2d: Sierpinski Carpet and T-rectangle second iteration antenna

Figure 3 present the return loss of different models

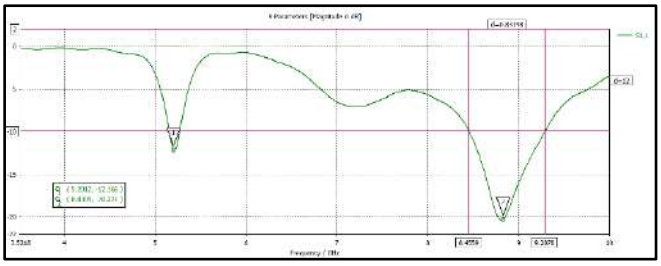


Figure 3a

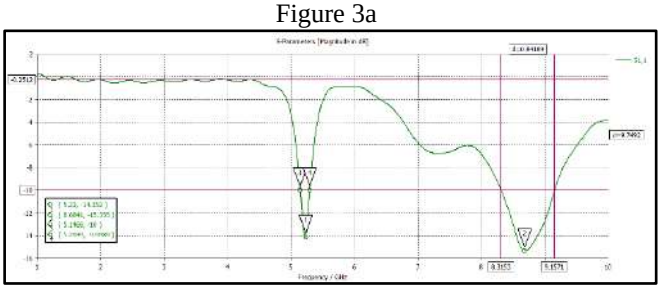


Figure 3b

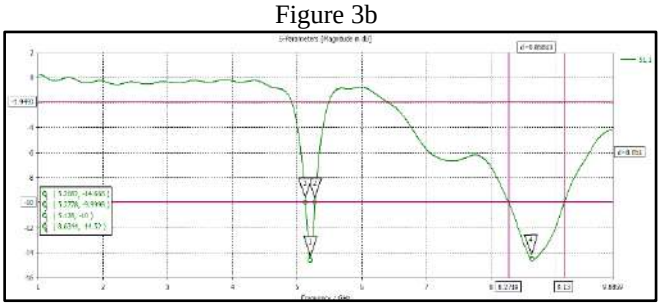


Figure 3c

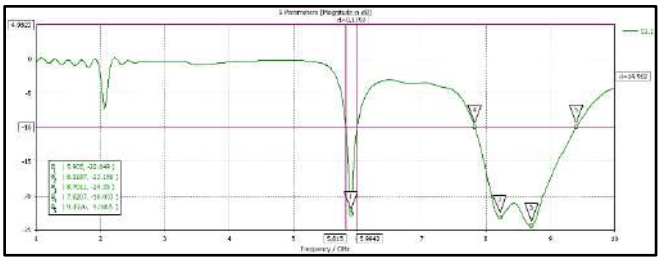


Figure 3d

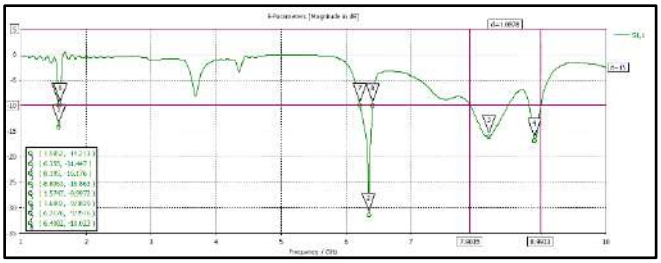


Figure 3e

Figure 2a : return loss plot of cenviontial patch antenna
Figure 2b : return loss plot of Sierpinski Carpet in the First iteration
Figure 2c: return loss plot Sierpinski Carpet with pentagons
Figure 2d: return loss plot Sierpinski Carpet and T-rectangle First iteration antenna
Figure 2e: return loss plot Sierpinski Carpet and T-rectangle second iteration antenna

Conclusion

The proposed Fractal Microstrip Antenna design produced a decent wideband of approximately 1.08 GHz from the 7.90 to 8.99GHz region, offering a bandwidth of about 18% for the many applications,. This investigation led us to the conclusion that as the number of fractal iterations increased, resonant frequency increased, leading to reduced power loss. Providing that the fractal geometry's inherent features are the cause of the broadband frequency response.

REFERENCES

[1] C.Balanus,' Antenna technology :past,present,future,in IEEE international workshop on antenna technology 5-7 March 2012Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) CONFERENCE 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016).

[2] Ankita Tiwari, Dr. Munish Rattan, Isha Gupta, " Review On: Fractal Antenna Design Geometries and Its Applications", International Journal Of Engineering And Computer Science, Volume - 3 Issue -9 September.

[3] R. H. Patel and T. K. Upadhyaya, "Compact planar dual band antenna for wlan application," Progress In Electromagnetics Research Letters, Vol. 70, 89-97, 2017.

[4] Dipika S. Sagne Rahul S. Batra P. L. Zade, "Design of Modified Geometry Sierpinski Carpet Fractal Antenna Array for Wireless Communication", 2013 3rd IEEE International Advance Computing Conference (IACC).

[5] Patel, Riki, and Trushit Upadhyaya. "An electrically small antenna for nearfield biomedical applications." Microwave and Optical Technology Letters 60, no. 3 (2018): 556-561.

[6] M. F. Abd Kadir, A. S. Ja'afar, M. Z. A. Abd Aziz, "Sierpinski Carpet Fractal Antenna", 2007 Asia-Pacific Conference on applied electromagnetic proceedings.

[7] Renu Bansal , P.K.Singhal & P.C.Sharma, "A novel multiband antenna: fractal antenna"

[8] Patel R, Upadhyaya T. An electrically small antenna for nearfield biomedical applications. Microw Opt Technol Lett. 2018;60:556–561. https://doi.org/10.1002/mop.31007

[9] S. Bhatt, P. Mankodi, A. Desai and R. Patel, "Analysis of ultra wideband fractal antenna designs and their applications for wireless communication: A survey," 2017 International Conference on Inventive Systems and Control (ICISC), Coimbatore, 2017,pp. 1-6.

[10] Patel, R. H., Desai, A. and Upadhyaya, T. (2015), A discussion on electrically small antenna property. Microw. Opt. Technol. Lett., 57: 2386–2388.

[11] Upadhyaya, Trushit K., et al. "Novel stacked μ-negative material-loaded

antenna for satellite applications." International Journal of Microwave and Wireless Technologies (2014): 1-7

[12] Payal Prusty, Bani Gandhi, Vinika Yadav, Sanjeev Saxena, "Design and development of microstrip fractal antenna".

[13] Amanpreet Kaur, Gursimranjit singh, "A Review Paper on Fractal Antenna Engineering", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 3, Issue 6, June 2014.

[14] P. Upadhyay and R. Sharma, "Design and Study of Inset Feed Square Microstrip Patch Antenna for S-Band Application," Int. J. Appl. or Innov. Eng. Manag., vol. 2, no. 1, pp. 256–262, 2013.

Randomness evaluation of coupled chaotic maps via NIST tests: A comparative study

Hadjer Bourekouche
LIST Laboratory, Department Engineering of Electrical Systems, Faculty of Technology
University M'Hamed Bougara of Boumerdes
35000 Boumerdes, Algeria
h.bourekouche@univ-boumerdes.dz

Samia Belkacem
LIMOSE Laboratory, Department Engineering of Electrical Systems, Faculty of Technology
University M'Hamed Bougara of Boumerdes
35000 Boumerdes, Algeria
s.belkacem@univ-boumerdes.dz

Noureddine Messaoudi
LIST Laboratory, Department Engineering of Electrical Systems, Faculty of Technology
University M'Hamed Bougara of Boumerdes
35000 Boumerdes, Algeria
n.messaoudi@univ-boumerdes.dz

***Abstract*— A vital requirement for any random number generator based on chaos is to ensure that the generated sequence always benefits of a significant level of randomness. It is critical to examine such sequences by means of Lyapunov exponents, bifur-cation diagrams, or other tests in order to accurately select the parameters of the dynamic system. However, the sequence's randomness quality varies depending on the generator's design and must be examined in different ways. Therefore, we argue to use the National Institute of Standards and Technology (NIST) suite tests to evaluate and compare the randomness properties of two coupled systems found in existing literature: the logistic-sine system (LSS) and the logistic-tent system (LTS). The results reveal that the LSS has much superior statistical features in terms of randomness than the LTS in the range [3.1–4]. This conclusion will substantially affect the selection of the perfect chaotic map to create sequences of keys that match the requirements of cryptography applications.**

Keywords—Random number generator, logistic-tent system, logistic-sine system, NIST, randomness.

I. INTRODUCTION

Random number generators (RNGs) that use chaotic systems (CSs) are fascinating because they provide high throughput data without the requirement for statistical post-processing and may be implemented as electrical circuits using very simple hardware [1]. In reality, using weak random values enables an opponent to break the whole system [2], that's why chaotic sequences have received a lot of interest lately because of their appealing characteristics, including great sensitivity to initial conditions [3]. With the help of this feature, it is possible to create a huge number of sequences with unique correlation properties and unique control pa-rameters, such as beginning conditions and bifurcation parameters [4].

The basic structure and ease of use of typical 1D chaotic maps are highlighted [5], however, they struggle with three major problems, including the limited or dis-continuous range of chaotic behaviors, the sensitivity of low-computation-cost analysis utilizing iteration and correlation functions, and the non-uniform data distribution of chaotic output sequences [6]. Therefore, it is necessary to create new chaotic systems with improved chaotic performance.

For instance, a nonlinear mixture of different 1D chaotic maps, notably the Logistic-Sine System (LSS) and the Logistic-Tent System (LTS), is proposed by Zhou et al. [6] to larger the chaotic ranges and strengthen the chaotic behavior com-pared with their seed maps. These qualities grabbed the attention of many re-searchers, who tried to use them in cryptography applications. Therefore, multiple tests are employed to quantify and evaluate its chaotic behavior, as well as the bi-furcation diagram, Lyapunov exponent, phase portrait, Poincare map, and various entropy metrics.

In [7], the 0-1 test and the three state test are used to provide a thorough investigation of the behavior of the LSS and LTS (3ST). While Zhou et al. [6] have proved that the LSS and LTS exhibit a continuous chaotic behavior in the range $r \in (0,4]$, Muthu et al. [7] portray stronger and weaker regions of chaos, with some regions exhibiting quasiperiodic behavior. In this paper, we aim to demonstrate that the map has the most chaotic nature and is nominated to produce high quality randomness; thus, we analyze and compare the randomness properties of the generated sequences by the LSS and the LTS using the NIST test suite SP 800-22.

The rest of the paper is structured as follows: in Section II, the dynamic behavior of the LSS and LTS via the bifurcation diagram, the Lyapunov exponent, the 0–1 test, and the 3ST are briefed. In section III, we use the NIST tests to compare and discuss the randomness quality of the two chaotic sequences generated by the LSS and LTS, followed by the conclusion of the work in section IV.

II. ANALYZE OF THE COMPORTEMENT BEHAVIOR OF LSS AND LTS FROM THE LITERATURE

The LSS and LTS are a form of chaotic coupled map that are often employed in image encryption due to the benefits of more advanced dynamical behavior. Numerous tests, such as the bifurcation diagram and the Lyapunov exponents in addition to the 0–1 and 3-ST tests, are used to measure its chaotic behaviors. In this section, we cover the theoretical elements of these tests applied to the LSS and the LTS.

A. *Logistic-sine and the logistic-tent systems*

The logistic-sine system (LSS) and logistic-tent system (LTS) proposed by Zhou et al. [6] whose main structure is illustrated in Fig. 1 are a nonlinear mixing of several 1D chaotic maps: the logistic, tent, and sine maps.

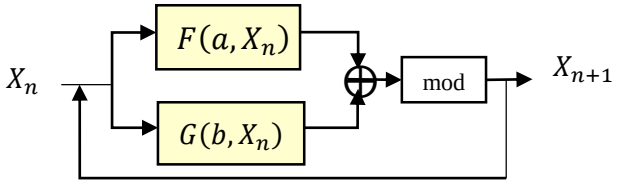


Fig. 1. General structure of 1D combined maps

$$X_{n+1} = \mathcal{F}_{LSS}(r, X_n)$$
$$X_{n+1} = \left[\mathcal{L}(r, X_n) + S\left((4-r), X_n\right)\right] mod 1$$
$$= [rX_n(1-X_n) + (4-r)\sin(\pi X_n)] mod 1 \quad (1)$$
$$r \in (0;\ 4]$$

$$X_{n+1} = G_{LTS}(r, X_n)$$
$$X_{n+1} = \left[\mathcal{L}(r, X_n) + T\left((4-r), X_n\right)\right] mod 1$$
$$= \begin{cases} \left[rX_n(1-X_n) + (4-r)\frac{X_n}{2}\right] \bmod 1 & X_i < 0{,}5 \\ \left[rX_n(1-X_n) + \frac{(4-r)(1-X_n)}{2}\right] \bmod 1 & X_i \geq 0{,}5 \end{cases} \quad (2)$$
$$r \in (0;\ 4]$$

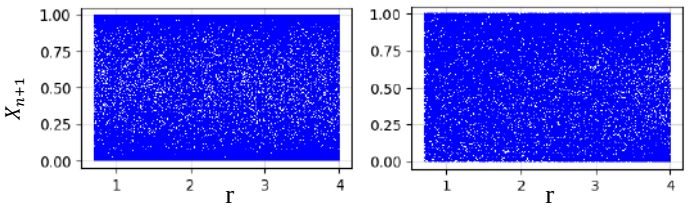
Several tests are available in the literature to emphasize the chaotic behavior of these 1D chaotic systems, including the bifurcation diagram, Lyapunov exponent, 0–1 test, and 3ST.

1) Bifurcation diagram

An essential feature for showing the behavior of chaotic systems is the bifurcation diagram [8] by plotting the output sequences X(n+1) of the chaotic map along with the change of its parameter r [9]. Fig. 2 (a-b) compares the bifurcation diagrams of the LSS and the LTS. From these figures, it is obvious that the chaotic range of the LSS and LTS is inside (0,4], and their bifurcation behavior is evenly dispersed over the full space from 0 to 1. Visually, this finding is not sufficient to compare and determine zones of chaos and regularity; consequently, identifying them using just bifurcation diagrams is difficult. Classification tests are useful allies in dealing with this kind of situation more clearly [10]. In general, the classification by Lyapunov Exponents is the most often used approach in the literature [11].

2) Lyapunov Exponents

When examining the dynamic behavior of chaotic systems, Lyapunov Exponents (LE) as key indicator that examines predictability is used [11]. It is one of the most frequently utilized tests since it is easy to implement when the map f is known explicitly. The Lyapunov Exponent of a discrete time system $X_{n+1} = f(X_n)$ is given by:



(a) (b)

Fig. 2. Bifurcation diagram: (a) logistic-tent system; (b) logistic-sine system

$$\lambda = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} ln|f'(x_i)| \quad (3)$$

Lyapunov Exponents are used to measure chaos. This depends on the sign of Lyapunov exponent λ as follows:

λ > 0, $\{X_n\}$ shows chaotic behavior;

λ < 0, $\{X_n\}$ shows periodic behavior;

λ = 0, a bifurcation occurs.

The Lyapunov Exponent of the LSS and LTS is tested in [6]as shown in Fig. 3. Visually, it is obvious that the LSS and LTS have more complex chaotic qualities as evidenced by their Lyapunov Exponents which are greater than 0 over the whole parameter setting range r, and they consistently behave chaotically in the range r ∈ (0,4] [7].

3) The 0-1 Test

A relatively new method for testing for chaos in deterministic discrete and continuous systems is the 0-1 test [10]. It is used to determine if there is chaos in digital sequences when a mathematical model is not available. Since the test directly applies to time series data and phase space reconstruction is not required, it has been shown to be more favorable than the Lyapunov exponent [7].

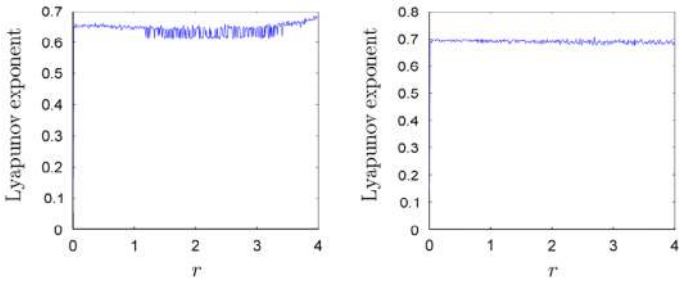
A single real number K and a two-dimensional graph with translation variables p and q make up the test's result. Information about the chaotic sequence may be revealed by the value of K [10]:

$K \approx 0$, Chaotic.

$K \approx 1$, Regular (non-chaotic).

The 0-1 test was experimented by Muthu et al. [7] on the LSS and the LTS with parameters N=5000 and $X_0 = 0.01$.

The K values obtained for the r values are shown in Fig. 4, which demonstrates a slope towards 1 for all values of r in the range [3.1, 4] for the LSS and LTS, demonstrating that these



(a) (b)

Fig. 3. Lyapunov Exponent diagram: (a) logistic-tent system; (b) logistic-sine system [6]

maps do not have a consistently chaotic character over the given range. Furthermore, Muthu et al. [7] demonstrate that the LSS possesses the strongest chaotic nature in most areas of r.

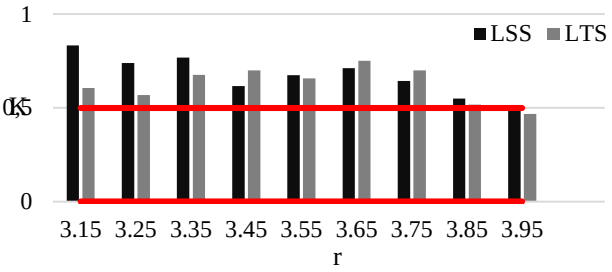


Fig. 4. 0-1 Test results: K values obtained for the r values

TABLE I. Results of behavior comparison of the LSS and LTS using 3ST proved by [7]

r	LSS	LTS
[3,1-3,19]	Chaotic	Quasi-periodic
[3,2-3,29]	Chaotic	Quasi-periodic
[3,3-3,39]	Chaotic	Chaotic
[3,4-3,49]	Quasi-periodic	Chaotic
[3,5-3,59]	Quasi-periodic	Chaotic
[3,6-3,69]	Chaotic	Chaotic
[3,7-3,79]	Quasi-periodic	Chaotic
[3,8-3,89]	Quasi-periodic	Quasi-periodic
[3,9-3,99]	Quasi-periodic	Quasi-periodic

4) 3ST

The 3ST is based on data series pattern analysis. The approach determines whether the dynamics are chaotic or regular by looking at the properties of periodic and quasi-periodic signals. The 3ST looks at how a data series' distribution of system states changes over time [12]. It is aimed at discriminating between the three major dynamics represented by the LE chaotic (> 0), quasi-periodic (<0), and periodic (= 0) dynamics [13].

Muthu et al. [7] performed 3ST on the LSS and LTS in the range r [3.1, 4]. Surprisingly, it clearly differentiates three types of behavior at various r values, periodic, quasi-periodic, and chaotic nature as categories in Table 1, which clearly demonstrate that LTS has a wider chaotic region than other maps. Further, it is demonstrated in [7] that the chaotic behavior of these maps is not uniformly distributed, and certain parts are found to be quasi-periodic in the LSS and LTS. This conclusion refutes what is asserted in [6] that LSS and LTS are chaotic throughout.

III. Experimental Study

In order to further study and compare the random properties of the chaotic sequence formed by the LSS and LTS, the National Institute of Standards and Technology (NIST) is used in this section to identify the areas of randomness and lack of randomness of these two sequences.

A. NIST test of chaotic sequence

In this test, we have expanded our study about the correspondence between the NIST statistical tests for pseudo-random number generators and certain chaotic metrics, including Lyapunov exponents, bifurcation diagrams, the 0-1 Test, and 3ST as a viable technique to verify findings in [6] and [7] .

First, the LSS and LTS systems are utilized to construct the chaotic series X (n+1) with control parameters r in the range (0,4] and X 0= 0.1 using the iterative procedures specified in (1) and (2). The bit length of each sequence n was set to 1000 bits. Then the statistical tests are done using NIST SP 800-22. The 15 sub-tests that make up the NIST test may all be used to assess the randomness of the sequences. By analyzing the sequence's uniformity, the test results largely show the benefits and draw-backs of the pseudo-random sequence [8], in which the probability value (P-value) reflects the regularity of the sequence. The p-value of each subtest is compared to a tester-determined significance threshold (which, for cryptography and in the case of NIST test suite version SP800-22, is commonly set to = 0.01[14]). If the P-value is greater than α, the sequence is random; otherwise, the sequence is not random. The NIST tests may be divided up into four basic categories of testing. These tests include the frequency test from 1–4, repetitive pattern tests from 5–6, pattern matching tests from 7–12, and random walk tests from 13–15 [15].

Since we can't determine which maps exhibit superior randomness simply by doing one test for just one value of r, we have to repeat the test according to r's transitions from 3,15 to 3,95, and then we have calculated the likelihood that a random sequence fails one or more tests for each testing process. Fig. 5 depicts the histogram plot of the uniformity test p-values at three values of r: 3.15, 3.65, and 3.95 for the two maps LSS and LTS. Table 2 shows the results of 15 failed tests at NIST for all r values mentioned.

SS findings exhibit excellent randomness, where all P-values are much over the significance threshold in most r values, expected Binary Matrix Rank Test, Overlap-ping Template Matching Test, and Maurer's Universal Statistical Test. It should be noted that some of these tests are not always appropriate. These tests are run only if the sequence meets certain criteria (for example, passing the frequency test, having more than 500 [16] cycles, and having a sufficient bit-length). However, LSS remains regarded as random even if it fails 3 to 4 tests, according to [16], where data may still be deemed random at the significance level α = 0.01 if they fail fewer than 7 NIST statistical tests.

LTS fails multiple tests when r is in the quasi-periodic range [3.1-3.29] and in the chaotic range [3.4-3.59]. That might be explained by the fact that the randomness of the sequences does not rely only on the chaotic state of the underlying system but also on the post-processing and the generator's design. It is obvious that the randomization qualities of these maps in such a range have exposed major security needs, which make its usage inappropriate for image encryption and demand a solid selection of the chaotic system parameters when employing them.

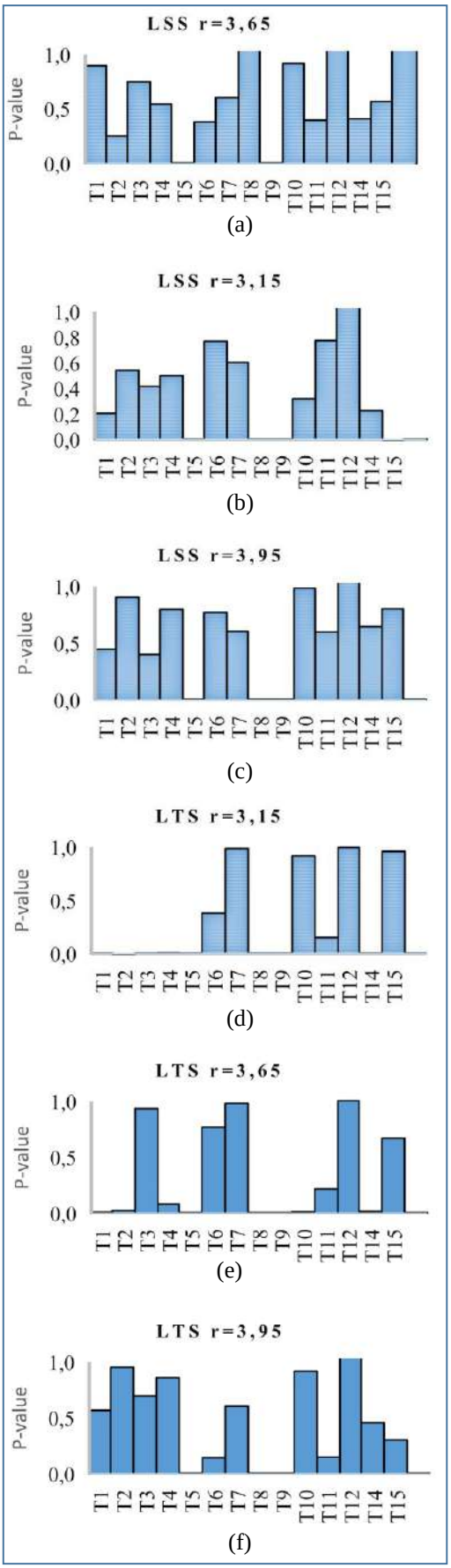


Fig. 5.The P-value of 15 statistical tests, the x axis lists the name of the statistical test in NIST test suit: T1- Frequency, T2- Frequency within a Block,T3- Runs,T4- Longest run of ones, T5- Rank,T6- Spectral,T7- Non-overlapping, T8- Overlapping, T9- Maurer's Universal, T10- Linear complexity, T11- Serial, T12- Approximate Entropy, T13- Cumulative sums, T14- Random Excursions, T15- Random Excursions Variant

TABLE II. Failed tests relatively to 15 tests for r listed values, we have use NR: Not random, R: Random

	r	3,15	3,25	3,35	3,45	3,55	3,65	3,75	3,85	3,95
LSS	Failures	4	3	3	3	3	3	3	3	3
	Conclusion	R	R	R	R	R	R	R	R	R
LTS	Failures	8	7	3	8	7	4	6	3	3
	Conclusion	NR	NR	R	NR	NR	R	R	R	R

Finally, the results reveal a strong relationship between the NIST tests and the chaotic metrics described in Part II. The seeds for which the maps are chaotic, are the seeds that determine a low number of failed NIST tests, which demonstrates the notion that a required criterion for a successful pseudo random number generator is that the development of the underlying system is chaotic. Thus, according to [7], the LSS that processed the strongest chaotic behavior in the range [0-4] is the one with the fewest failed NIST tests and the ability to generate a highly random chaotic sequence.

IV. Conlusion

Coupling chaotic maps is a common way to develop more sophisticated dynamic behavior. This paper gives a comparative examination of two coupled systems, namely the logistic-sine system (LSS) and the logistic-tent system (LTS), considering their randomization qualities to verify which performs the best. Since several test batteries are available to debate which system is best utilized as a random number generator, we pick the NIST test suite, which is regarded as the most appropriate one. We have demonstrated through NIST tests that the LSS presents better properties of randomness than the LTS, while the LTS is discontinuity random in the range of [3.1–3.95], the LSS is random throughout; thus, this will strongly influence the selection of the perfect chaotic map to generate sequences of keys used later for many applications, such as image encryption. We also discovered experimentally that using Lyapunov exponents, bifurcation diagrams, the 0-1 test, 3ST, and NIST tests is required to select the dynamic system characteristics required to develop chaos-based random generators. However, merely making a proper selection of the chaotic system characteristics is not enough; the unpredictability or lack of randomness of such a sequence relies on many aspects, including post-processing and the generator's design, and must be assessed in other ways.

Acknowledgment

The authors would like to thank the anonymous reviewers for their constructive comments on the earlier version of the paper.

References

[1] K. Demir and S. Ergün, "A comparative analysis on chaos-based random number generation methods," *Eur. Phys. J. Plus*, vol. 137, no. 5, Art. no. 5, May 2022, doi: 10.1140/epjp/s13360-022-02793-6.

[2] A. Doğanaksoy, F. Sulak, M. Uğuz, O. Şeker, and Z. Akcengiz, "New Statistical Randomness Tests Based on Length of Runs,"

Mathematical Problems in Engineering, vol. 2015, p. e626408, Apr. 2015, doi: 10.1155/2015/626408.

[3] "Artificial Intelligence and Bioinspired Computational Methods: Proceedings of the 9th Computer Science On-line Conference 2020, Vol. 2 [1st ed.] 9783030519704, 9783030519711," *dokumen.pub*. https://dokumen.pub/artificial-intelligence-and-bioinspired-computational-methods-proceedings-of-the-9th-computer-science-on-line-conference-2020-vol-2-1st-ed-9783030519704-9783030519711.html (accessed Jan. 22, 2023).

[4] H. A. A. Mansour, "Analysis, Study and Optimization of Chaotic Bifurcation Parameters Based on Logistic/Tent Chaotic Maps," in *Artificial Intelligence and Bioinspired Computational Methods*, Cham, 2020, pp. 642–652. doi: 10.1007/978-3-030-51971-1_52.

[5] H. Li, C. Yu, and X. Wang, "A novel 1D chaotic system for image encryption, authentication and compression in cloud," *Multimed Tools Appl*, vol. 80, no. 6, pp. 8721–8758, Mar. 2021, doi: 10.1007/s11042-020-10117-y.

[6] Y. Zhou, L. Bao, and C. L. P. Chen, "A new 1D chaotic system for image encryption," *Signal Processing*, vol. 97, pp. 172–182, Apr. 2014, doi: 10.1016/j.sigpro.2013.10.034.

[7] Joan. S. Muthu, A. J. Paul, and P. Murali, "An Efficient Analyses of the Behavior of One Dimensional Chaotic Maps using 0–1 Test and Three State Test," in *2020 IEEE Recent Advances in Intelligent Computational Systems (RAICS)*, Dec. 2020, pp. 125–130. doi: 10.1109/RAICS51191.2020.9332470.

[8] X. Wang, N. Guan, H. Zhao, S. Wang, and Y. Zhang, "A new image encryption scheme based on coupling map lattices with mixed multi-chaos," *Sci Rep*, vol. 10, no. 1, Art. no. 1, Jun. 2020, doi: 10.1038/s41598-020-66486-9.

[9] N. JITEURTRAGOOL, "Developments of CMOS based Chaotic Oscillator Circuits and Its Application," PhD Thesis, 高知工科大学, 2018. Accessed: Jan. 25, 2023. [Online]. Available: https://ci.nii.ac.jp/naid/500001366092

[10] M. Melosik and W. Marszalek, "On the 0/1 test for chaos in continuous systems," *Bulletin of the Polish Academy of Sciences Technical Sciences*, vol. 64, Sep. 2016, doi: 10.1515/bpasts-2016-0058.

[11] M. Tosin, M. V. Issa, D. Matos, A. D. Nascimento, and A. C. Jr, "Employing 0-1 test for chaos to characterize the chaotic dynamics of a generalized Gauss iterated map," presented at the XIV Conferência Brasileira de Dinâmica, Controle e Aplicações (DINCON 2019), Nov. 2019. Accessed: Jan. 21, 2023. [Online]. Available: https://hal.science/hal-02388470

[12] J. S. A. Eyebe Fouda, J. Y. Effa, M. Kom, and M. Ali, "The three-state test for chaos detection in discrete maps," *Applied Soft Computing*, vol. 13, no. 12, pp. 4731–4737, Dec. 2013, doi: 10.1016/j.asoc.2013.07.020.

[13] J. S. A. Eyebe Fouda and W. Koepf, "Efficient detection of the quasi-periodic route to chaos in discrete maps by the three-state test," *Nonlinear Dyn*, vol. 78, no. 2, pp. 1477–1487, Oct. 2014, doi: 10.1007/s11071-014-1529-4.

[14] R. Stępień and J. Walczak, "Statistical analysis of the LFSR generators in the NIST STS test suite," *Computer Applications in Electrical Engineering*, vol. Vol. 11, 2013, Accessed: Jan. 22, 2023. [Online]. Available: http://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-6b8818c1-d909-4960-8c08-647e878b1531

[15] M. Kaas-Mason, G. Prpić, S. Suriyasuphapong, and N. Bailey, "Comparison of Pseudo, Chaotic and Quantum Random Number Generators and their use in Cyber Security".

[16] M. Sýs, Z. Riha, V. Matyas, K. Marton, and A. Suciu, "On the interpretation of results from the NIST statistical test suite," vol. 18, pp. 18–32, Jan. 2015.

Salt and pepper noise removal method based on anisotropic total variation regularization

1st DIFFELLAH Nacira
dept. of electronics, Faculty of technology
ETA Laboratory, University Mohamed El Bachir El Ibrahimi of Bordj Bou Arreridj
Algeria
diffellahn@gmail.com or 0000-0003-2474-0700

2nd BEKKOUCHE Tewfik
dept. of electromecanics, Faculty of technology
ETA Laboratory, University Mohamed El Bachir El Ibrahimi of Bordj Bou Arreridj
Algeria
bekkou66@hotmail.com or 0000-0002-5405-7382

3rd HAMDINI Rabah
dept. of electronics, Faculty of technology
SET Laboratory,University Saad Dahlab of Blida
Algeria
hamdinirabah@gmail.com or 0000-0003-3127-1367

4th MOKHNACHE Salah
Faculty of technology
University Ferhat Abbas of Setif
Algeria
mokhnachesalah@yahoo.fr or ORCID

***Abstract*—In this paper, an image-denoising technique based on Total Variation (TV) regularization is proposed for Salt-And-Pepper (SAP) noise. We consider two problems: First, in $\ell_1 - norm$ problem, we add squared magnitude as the data term to the penalty function. Second, in $\ell_2 - norm$ problem we add the absolute value of magnitude as a data solution term to the regularization function. The augmented Lagrange Multiplier (ALM) method is used for solving constrained optimization problems. Extensive experimental results demonstrate that ℓ_1 method can get much better performance metrics: Mean Square Error (MSE), Signal Noise–to-Ratio (SNR) Peak Signal-to-Noise Ratio ($PSNR$) and Structural SIMilarity index ($SSIM$), and ℓ_1 method can remove the Salt and pepper noise effectively and keep the image details well in comparison to the ℓ_2 method.**

***Index Terms*—image denoise, total variation, regularization, salt-and-pepper noise, norms, metrics**

I. Introduction

Salt and pepper noise can be introduced during the signal acquisition stage or due to the bit error in the transmission [14]. The image degradation process may be modeled as:

$$f = u + \eta \tag{1}$$

here f is the degraded image, u is an observed image, and η is a salt-and-pepper noise contaminating the measurements, in this noise the corrupted pixels take the maximum value or the minimum value which sets randomly black and white spots on the image [9]. This problem is ill-posed, to overcome this drawback, input data need to be regularized [13]. We choose an anisotropic total variation as a term of regularization, then we solve the following two optimization problems:

$$(P_1): \quad \min_u \frac{\lambda}{2} |u - f|_2^2 + u_{TV} \tag{2}$$

$$(P_2): \quad \min_u \frac{\lambda}{2} |u - f|_1 + u_{TV} \tag{3}$$

The first term is the fidelity and the second one is the regularization. Here $|\cdot|_1$ and $|\cdot|_2$ are defined as $\ell_1 - norm$ and $\ell_2 - norm$, respectively. λ is regularization parameter. u is a grayscale image of size $M \times N$, u_{TV} is anisotropic total variation, it is defined as:

$$\begin{aligned} u_{TV} &= |Du|_1 = |D_x u|_1 + |D_y u|_1 \\ &= |u(x+1,y) - u(x,y)| + |u(x,y+1) - u(x,y)| \end{aligned} \tag{4}$$

where x and y are the row and column indices, respectively, D is the gradient operator ∇ in the discrete setting, D_x and D_y are the horizontal and vertical partial derivative operators, respectively.
The discrete u_{TV} can be defined as:

$$u_{TV\,i,j} = \sum_{i=1}^{M} \sum_{j=1}^{N} \left(|u_{i+1,j} - u_{i,j}| + |u_{i,j+1} - u_{i,j}| \right) \tag{5}$$

We call problem (2) $\ell_1 - \ell_1$ minimization and problem (3) $\ell_2 - \ell_1$ minimization.

To remove SAP noise from the image many algorithms have been proposed median filters [1], [7], fuzzy approach Algorithms based on fuzzy sets [11], [8], Adaptive weighted mean filter [15] iterative nonlocal means filter [12], Adaptive Riesz Mean Filter [5], variational regularisation [3], [12] [2], [6],

The paper is structured as follows: in section I, we present an introduction to SAL noise. Then, we introduce the definition of the total variation of an image as the l1 norm of its gradient field amplitude and the formulation for the discrete TV. In section II, we consider the denoising problems with $\ell_1 - \ell_1$ and $\ell_2 - \ell_1$ approaches to remove SAP noise in image. The section III, describes the evaluation techniques: Mean-Square Error (MSE), Signal-to-noise Ratio (SNR), Peak Signal-to-Noise Ratio ($PSNR$), and Structural SIMilarity index ($SSIM$). Section IV, provides the simulation results. Discussion of experiment results is described in section V,

and section VI, gives the conclusion and perspectives. Finally, the references of this article are listed.

II. Problem formulation

This section considers solving the two problems (2) in subsection II-A and (3) in subsection II-B.

A. First formulation: $\ell_2 - \ell_1$ minimization problem

For solving the problem (2), the following steps must be followed:

- Write problem (2) as:

$$\begin{aligned} &\underset{u,w}{\text{minimize}} \quad \frac{\lambda}{2}|u-f|^2 + |w|_1 \\ &\text{subject to} \quad w = Du \end{aligned} \tag{6}$$

  then set the constraint equal to 0: $w - Du = 0$
- Based on the Augmented Lagrange method, we rewrite the constrained optimization problem (6) to be a saddle-point problem (7):

$$L(u,w,z) = \frac{\lambda}{2}|u-f|^2 + |w|_1 - z^T(w-Du) - \frac{d}{2}|w-Du|^2 \tag{7}$$

  where $L(u,w,v)$ is augmented Lagrangian, d is a regularization parameter and z is the Lagrange multiplier associated.
- Solve sequence of unconstrained minimization of augmented Lagrangian:

$$\begin{cases} u = \arg\min\limits_{u} L(u,w,z) \\ w = \arg\min\limits_{w} L(u,w,z) \end{cases} \tag{8}$$

  and using the following multiplier update [3]:

$$z_{k+1} = z_k - d(w_{k+1} - Du_{k+1}) \tag{9}$$

  Then, apply the method of multipliers:

$$\begin{cases} u_{k+1} = \frac{\lambda}{2}|u-f|^2 - z_k^T(w_k - Du) + \frac{d}{2}|w_k - Du|^2 \\ w_{k+1} = |w|_1 - z_k^T(w - Du_{k+1}) + \frac{d}{2}|w - Du_{k+1}|^2 \end{cases} \tag{10}$$

  Hence (10) has the following solution [3]:

$$\begin{cases} u = TF^{-1}\left\{\dfrac{TF\{\lambda f + dD^T w - D^T z\}}{\lambda + d(TF\{D_x\} + TF\{D_y\})}\right\} \\ w = \max\left\{\left|Du + \frac{1}{d}z\right| - \frac{1}{d}, 0\right\}.sgn\left(Du + \frac{1}{d}z\right) \end{cases} \tag{11}$$

  TF denotes the Fourier transform operator and TF^{-1} is the inverse Fourier transform.

B. Second formulation: $\ell_1 - \ell_1$ minimization

We try to recover the original image by minimizing (3). All steps of resolution are summarized as follows:

- Replace (3) by:

$$\begin{aligned} &\underset{u}{\text{minimize}} \quad \lambda\|u-f\|_1 + \|w\|_1 \\ &\text{subject to} \quad r = u - f \\ &\qquad\qquad\quad w = Du \end{aligned} \tag{12}$$

  then set the constraints equal to 0: $r - u + f = 0$ and $w - Du = 0$
- Converting constrained optimization (12) to unconstrained minimization problem using the Lagragian method:

$$\begin{aligned} L(u,r,w,z_1,z_2) = \lambda|r|_1 + |w|_1 - z_1^T(r-u+f) - z_2^T(w-Du) \\ + \frac{d_1}{2}|r-u+f|^2 + \frac{d_2}{2}|w-Du|^2 \end{aligned} \tag{13}$$

  Here, $L(u,r,w,z_1,z_2)$ is the augmented Lagrangian function, d_1, d_2 are regularization parameters and z_1, z_2 are the Lagrange multipliers associated.
- u, r and w are minimized separately and sequentially as:

$$\begin{cases} u = \arg\min\limits_{u} L(u,r,w,z_1,z_2) \\ r = \arg\min\limits_{r} L(u,r,w,z_1,z_2) \\ w = \arg\min\limits_{w} L(u,r,w,z_1,z_2) \end{cases} \tag{14}$$

  and using the following multipliers update [3]:

$$\begin{cases} z_{1k+1} = z_{1k} - d_1(r_{k+1} - u_{k+1} + f_{k+1}) \\ z_{2k+1} = z_{2k} - d_2(w_{k+1} - Du_{k+1}) \end{cases} \tag{15}$$

  Hence can view the augmented Lagrangian process as:

$$\begin{cases} \min\limits_{u}\left\{z_1^T(u) + z_2^T(Du) + \frac{d_1}{2}|r-u+f|^2 + \frac{d_2}{2}|w-Du|^2\right\} \\ \min\limits_{r}\left\{\lambda|r|_1 - z_1^T(r) + \frac{d_1}{2}|r-u+f|^2\right\} \\ \min\limits_{w}\left\{|w|_1 - z_2^T(w-Du) + \frac{d_2}{2}|w-Du|^2\right\} \end{cases} \tag{16}$$

  The problem (16) can be rewritten as following [3]:

$$\begin{cases} u = TF^{-1}\left\{\dfrac{TF\{d_1 f + d_1 w - z + D^T(d_2 w - z)\}}{d_1 + d_2(TF^2\{D_x\} + TF^2\{D_y\})}\right\} \\ r = \max\left\{\left|u - f + \frac{1}{d_1}z_1\right| - \frac{\lambda}{d_1}, 0\right\}.sgn\left(u - f + \frac{1}{d_1}z_1\right) \\ w = \max\left\{\left|Du + \frac{1}{d_2}z_2\right| - \frac{1}{d_2}, 0\right\}.sgn\left(Du + \frac{1}{d_2}z_2\right) \end{cases} \tag{17}$$

III. Evaluation denoising techniques

We have measured the performance of models by Mean-Square Error (MSE), Signal-to-noise Ratio (SNR), Peak Signal-to-Noise Ratio ($PSNR$),and Structural SIMilarity index ($SSIM$) [4], [10]:

A. Mean Square Error (MSE)

The mean Square Error value denotes the average difference of the pixels all over the image. MSE is defined as:

$$MSE = \frac{1}{M.N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(u\left(i,j\right) - \hat{u}\left(i,j\right)\right)^2 \qquad (18)$$

here, $M \times N$ is the image size, $u\left(i,j\right)$ is the original image, whereas, $\hat{u}\left(i,j\right)$ is the denoised image.

B. Signal Noise–to-Ratio (SNR)

Equation (19) is the formula used to calculate the SNR:

$$SNR = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(u\left(i,j\right)\right)^2}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(u\left(i,j\right) - \hat{u}\left(i,j\right)\right)^2} \qquad (19)$$

C. Peak Signal-to-Noise Ratio ($PSNR$)

For calculating the $PSNR$, MSE is used. $PSNR$ can be calculated as follows:

$$PSNR = 10\log_{10}\left(\frac{d^2}{MSE}\right) \qquad (20)$$

d is the maximum possible pixel value of the image.

D. Structural SIMilarity index ($SSIM$)

The following equations give the expression of $SSIM$:

$$SSIM\left(u,\hat{u}\right) = \frac{\left(2\mu_u\mu_{\hat{u}} + C_1\right)\left(2\sigma_{u\hat{u}} + C_2\right)}{\left(\mu_u^2 + \mu_{\hat{u}}^2 + C_1\right)\left(\sigma_u^2 + \sigma_{\hat{u}}^2 + C_2\right)} \qquad (21)$$

Where $SSIM\left(u,\hat{u}\right)$ is the Structural SIMilarity index between two images, $\mu_u, \mu_{\hat{u}}$ are average of $u, \hat{u}$ respectively, $\sigma_u, \sigma_{\hat{u}}$ are standard deviation of $u, \hat{u}$, $\sigma_{u\hat{u}}$ is standard deviation between u and $\hat{u}$, C_1 ,C_2 are positive constant.

Generally, a smaller value of the MSE and higher values of the SNR, $PSNR$, and $SSIM$ imply a good denoised image.

IV. Results

To compare $\ell_1 - \ell_1$ method with $\ell_2 - \ell_1$ method, seven different types of images are used in our applications such '37073.jpg' , '42049.jpg', '102061.jpg' '03070.jpg', '126007.jpg' of size (321×481) and '172032.jpg', '54082.jpg' ,'376043.jpg' of size (481×321). These images are chosen from the Berkeley Segmentation Data Set ($BSDS300$) available at: https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/. All images were corrupted with salt and pepper noise.

(a) True image u

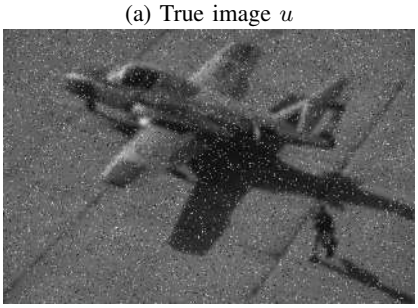
(b) Degraded image f
$MSE = 0.0147$ $SNR = 13.7383dB$
$PSNR = 18.3298dB$ $SSIM = 0.2080$

(c) Restored image with $\ell_1 - \ell_1$
$MSE = 0.0008$ $SNR = 26.4341dB$
$PSNR = 31.1152dB$ $SSIM = 0.8726$

(d) Restored image with $\ell_2 - \ell_1$
$MSE = 0.0017$ $SNR = 22.9407dB$
$PSNR = 27.7000dB$ $SSIM = 0.7839$

Fig. 1
Result of $172032.jpg'$ image

(a) True image u



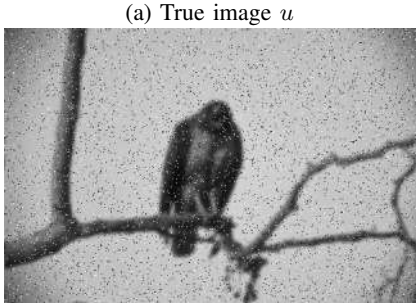
(b) Degraded image f
$MSE = 0.01771$ $SNR = 12.3146dB$
$PSNR = 17.5159dB$ $SSIM = 0.2789$

(c) Restored image with $\ell_1 - \ell_1$
$MSE = 0.0014$ $SNR = 19.8370dB$
$PSNR = 28.5153dB$ $SSIM = 0.7813$

(d) Restored image with $\ell_2 - \ell_1$
$MSE = 0.0026$ $SNR = 17.2429dB$
$PSNR = 25.8312dB$ $SSIM = 0.7020$

Fig. 2
Result of $37073.jpg'$ image

(a) True image u



(b) Degraded image f
$MSE = 0.0188$ $SNR = 13.9288dB$
$PSNR = 17.2662dB$ $SSIM = 0.2344$

(c) Restored image with $\ell_1 - \ell_1$
$MSE = 0.0009$ $SNR = 27.0205dB$
$PSNR = 30.3361dB$ $SSIM = 0.9397$

(d) Restored image with $\ell_2 - \ell_1$
$MSE = 0.0027$ $SNR = 22.2781dB$
$PSNR = 25.7529dB$ $SSIM = 0.8700$

Fig. 3
Result of $42049.jpg'$ image

(a) True image u

(b) Degraded image f
$MSE = 0.0160$ $SNR = 10.6173dB$
$PSNR = 17.9652dB$ $SSIM = 0.2064$

(c) Restored image with $\ell_1 - \ell_1$
$MSE = 0.0019$ $SNR = 19.4110dB$
$PSNR = 27.1386dB$ $SSIM = 0.6964$

(d) Restored image with $\ell_2 - \ell_1$
$MSE = 0.0030$ $SNR = 17.5650dB$
$PSNR = 25.2476dB$ $SSIM = 0.6035$

Fig. 4
Result of $54082.jpg'$ image

(a) True image u

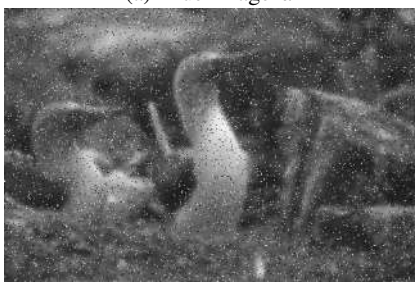
(b) Degraded image f
$MSE = 0.0186$ $SNR = 11.7817dB$
$PSNR = 17.3092dB$ $SSIM = 0.2403$

(c) Restored image with $\ell_1 - \ell_1$
$MSE = 0.0034$ $SNR19.0144dB$
$PSNR = 24.6576dB$ $SSIM = 0.6145$

(d) Restored image with $\ell_2 - \ell_1$
$MSE = 0.0061$ $SNR = 16.4156dB$
$PSNR = 22.1617dB$ $SSIM = 0.3929$

Fig. 5
Result of $376043.jpg'$ image

(a) True image u



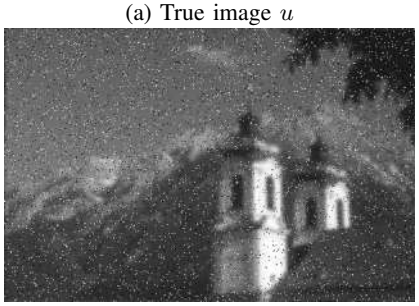
(b) Degraded image f
$MSE = 0.0161$ $SNR = 9.9322dB$
$PSNR = 17.9300dB$ $SSIM = 0.2523$

(c) Restored image with $\ell_1 - \ell_1$
$MSE = 0.0012$ $SNR = 20.7989dB$
$PSNR = 29.2341dB$ $SSIM = 0.8079$

(d) Restored image with $\ell_2 - \ell_1$
$MSE = 0.0025$ $SNR = 17.5945dB$
$PSNR = 25.9885dB$ $SSIM = 0.6800$

Fig. 6
Result of $103070.jpg'$ image

(a) True image u



(b) Degraded image f
$MSE = 0.0170$ $SNR = 9.9395dB$
$PSNR = 17.6930dB$ $SSIM = 0.2229$

(c) Restored image with $\ell_1 - \ell_1$
$MSE = 0.0015$ $SNR = 20.2215dB$
$PSNR = 28.3577dB$ $SSIM = 0.8162$

(d) Restored image with $\ell_2 - \ell_1$
$MSE = 0.0030$ $SNR = 17.1265dB$
$PSNR = 25.2541dB$ $SSIM = 0.7089$

Fig. 7
Result of $126007.jpg'$ image

V. Discussion

Figure 1 to 7 illustrate the original images, images corrupted by salt-and-pepper noise with density 5%, and denoised images by $\ell_1-\ell_1$ and $\ell_2-\ell_1$ minimization techniques. It can be clearly seen that $\ell_1-\ell_1$ method is best than $\ell_2-\ell_1$ method.

The metrics MSE, SNR, $PSNR$, and $SSIM$ are obtained with the two approaches for the various images on the test set can be found in table I. The mean square error values measured between the original image and the denoised one show that those obtained by $\ell_1-\ell_1$ are higher than those measured by $\ell_2-\ell_1$. A lower value of MSE denotes a small difference between the original and the processed images. All the signal-to-noise ratio values are greater than $0\,dB$, which indicates that the image recovery level is greater than the degradation level. The SNR obtained with $\ell_1-\ell_1$ is higher than that obtained with $\ell_2-\ell_1$ for all images . The higher SNR, the better the image quality. The higher the peak-signal-to-noise ratio, the better the degraded image has been denoised to better match the original image. $\ell_1-\ell_1$ presents a better technique of denoised image than$\ell_1-\ell_1$ approach. The Structural Similarity Index is used as a metric to measure the similarity between degraded and denoised grayscale images. If the value tends towards 0, the degraded and reconstructed images are very similar. For all grayscale images, the $SSIM$ values calculated by applying $\ell_1-\ell_1$ are higher than those obtained with $\ell_2-\ell_1$.

TABLE I
Comparison of the denoising results of images

Image		MSE	$SNRdB$	$PSNRdB$	$SSIM$
172032	degraded image	0.0147	13.7383	18.3298	0.2080
	$\ell_1-\ell_1$	0.0008	26.4341	31.1152	0.8726
	$\ell_2-\ell_1$	0.0017	22.9407	27.7000	0.7839
37073	degraded image	0.0177	12.3146	17.5159	0.2789
	$\ell_1-\ell_1$	0.0014	19.8370	28.5153	0.7813
	$\ell_2-\ell_1$	0.0026	17.2429	25.8312	0.7020
42049	degraded image	0.0188	13.9288	17.2662	0.2344
	$\ell_1-\ell_1$	0.0009	27.0205	30.3361	0.9397
	$\ell_2-\ell_1$	0.0027	22.2781	25.7529	0.8700
54082	degraded image	0.0160	10.6173	17.9652	0.2064
	$\ell_1-\ell_1$	0.0019	19.4110	27.1386	0.6964
	$\ell_2-\ell_1$	0.0030	17.5650	25.2476	0.6035
376043	degraded image	0.0186	11.7817	17.3092	0.2403
	$\ell_1-\ell_1$	0.0034	19.0144	24.6576	0.6145
	$\ell_2-\ell_1$	0.0061	16.4156	22.1617	0.3929
103070	degraded image	0.0165	9.8605	17.8348	0.2465
	$\ell_1-\ell_1$	0.0012	20.8131	29.2482	0.8088
	$\ell_2-\ell_1$	0.0025	17.5945	25.9885	0.6800
126007	degraded image	0.0170	9.9395	17.6930	0.2229
	$\ell_1-\ell_1$	0.0015	20.2215	28.3577	0.8162
	$\ell_2-\ell_1$	0.0030	17.1265	25.2541	0.7089

VI. Conclusion

In this paper, we have developed two regularization models based on total variation and ℓ_1 and ℓ_2 norms to remove salt and pepper noise in the image. To compare and evaluate these models, we have used Meam Square Error, Signal Noise–to-Ratio, Peak Signal-to-Noise Ratio, and Structural SIMilarity index as metrics. The experimental results show that $\ell_1-\ell_1$ can get better MSE, SNR, $PSNR$, and $SSIM$ than $\ell_2-\ell_1$. Finally, we want to extend this study to deal with other image processing problems such as image segmentation and image inpainting.

Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions which greatly improved the quality of the paper, the General Directorate for Scientific Research and Technological Development of the Algerian Republic in general and the ETA laboratory of Bordj Bou Arreridj University.

References

[1] J. Astola and P. Kuosmanen, *Fundamentals of nonlinear digital filtering*. CRC press, 1997, vol. 8.

[2] L. Calatroni and K. Papafitsoros, "Analysis and automatic parameter selection of a variational model for mixed gaussian and salt-and-pepper noise removal," *Inverse Problems*, vol. 35, no. 11, p. 114001, 2019.

[3] S. H. Chan, R. Khoshabeh, K. B. Gibson, P. E. Gill, and T. Q. Nguyen, "An augmented lagrangian method for total variation video restoration," *IEEE Transactions on Image Processing*, vol. 20, no. 11, pp. 3097–3111, 2011.

[4] N. Diffellah, R. Hamdini, and T. Bekkouche, "Removal of multiplicative gamma noise from images via srad model amelioration," *Engineering, Technology & Applied Science Research*, vol. 11, no. 6, pp. 7917–7921, 2021.

[5] S. Enginoğlu, U. Erkan, and S. Memiş, "Pixel similarity-based adaptive riesz mean filter for salt-and-pepper noise removal," *Multimedia Tools and Applications*, vol. 78, pp. 35 401–35 418, 2019.

[6] U. Erkan, D. N. Thanh, S. Enginoğlu, and S. Memiş, "Improved adaptive weighted mean filter for salt-and-pepper noise removal," in *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*. IEEE, 2020, pp. 1–5.

[7] A. Farzana FM, H. Arshadh, S. Safreen *et al.*, "Design and analysis for removing salt and pepper noise in image processing," *Indo-Iranian Journal of Scientific Research (IIJSR)*, 2018.

[8] M. Gonzalez-Hidalgo, S. Massanet, A. Mir, and D. Ruiz-Aguilera, "Improving salt and pepper noise removal using a fuzzy mathematical morphology-based filter," *Applied Soft Computing*, vol. 63, pp. 167–180, 2018.

[9] B. Karthik, T. Krishna Kumar, S. Vijayaragavan, and M. Sriram, "Removal of high density salt and pepper noise in color image through modified cascaded filter," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 3901–3908, 2021.

[10] M. Nadipally, "Optimization of methods for image-texture segmentation using ant colony optimization," in *Intelligent data analysis for biomedical applications*. Elsevier, 2019, pp. 21–47.

[11] V. Singh, R. Dev, N. K. Dhar, P. Agrawal, and N. K. Verma, "Adaptive type-2 fuzzy approach for filtering salt and pepper noise in grayscale images," *IEEE transactions on fuzzy systems*, vol. 26, no. 5, pp. 3170–3176, 2018.

[12] X. Wang, S. Shen, G. Shi, Y. Xu, and P. Zhang, "Iterative non-local means filter for salt and pepper noise removal," *Journal of visual communication and image representation*, vol. 38, pp. 440–450, 2016.

[13] T. Wu, X. Gu, Y. Wang, and T. Zeng, "Adaptive total variation based image segmentation with semi-proximal alternating minimization," *Signal Processing*, vol. 183, p. 108017, 2021.

[14] H. Zhang, Y. Zhu, and H. Zheng, "Namf: A non-local adaptive mean filter for salt-and-pepper noise removal," *arXiv preprint arXiv:1910.07787*, 2019.

[15] P. Zhang and F. Li, "A new adaptive weighted mean filter for removing salt-and-pepper noise," *IEEE signal processing letters*, vol. 21, no. 10, pp. 1280–1283, 2014.

A New Medical Image Encryption Using Enhanced Chaotic System.

1st Djamel Herbadji
Electonics Research Laboratory
University 20 August 1955
Skikda 21000, Algeria
herbadjidjamel@gmail.com

2nd Ismail Haddad
Electonics Research Laboratory
University 20 August 1955
Skikda 21000, Algeria
i.haddad@univ-skikda.dz

Abderrahmene Herbadji
Electronic Research Laboratory
Mouhamed Boudiaf University
Msila,Algeria
herbadjiabder@gmail.com

3rd Aissa Belmeguenai
Electonics Research Laboratory
University 20 August 1955
Skikda 21000, Algeria
belmeguenaiaissa@yahoo.fr

***Abstract*—The remarkable progress in image processing and telecommunications has necessitated the protection of the transmitted image from hackers. In this work, we will present a new approach to image encryption based on the Zigzag method, where this zigzag is controlled by enhanced chaotic system . The system is also used in the diffusion process to change the values of image pixels. The method passed all of the tests that were used to evaluate its effectiveness, including entropy, histogram, and correlation coefficient. It has also been shown that the method can respond to a variety of assaults, including statistical and entropy attacks.**

***Index Terms*—Image processing, fractional order, chaotic system, image encryption.**

I. Introduction

In recent years,encryption has become a key instrument for the security of data transported through computer and communication networks . therefore it is regarded as an important approach to safeguard digital images [1]. Image encryption has become a particularly intriguing subject of research [2]–[4].Because the image has distinct features than text, such as vast data. As a result, while certain classic encryption algorithms are excellent for encrypting text, they are ineffective for encrypting pictures.Since chaotic maps have prominent characteristics like highly sensitive dependence on initial conditions and control parameters, ergodicity, unpredictability, and pseudo-randomness, which can satisfy requirements like diffusion and permutation in the direction of image encryption, also chaos-based encryption algorithms are very efficient in terms of speed and security [5]. As a result, chaotic-based encryption systems have received a lot of research attention.The image encryption approaches presented in the literature are implemented on a variety of chaotic maps and employ a variety of encryption algorithms.

So, in this work, we will present a method for encoding images using an enhanced one-dimensional chaotic map. This map offers solutions to several map issues, including finite chaotic fields and non-uniform distribution of random sequences. Which is proven by the bifurcation diagram and Lyapunov exponent.

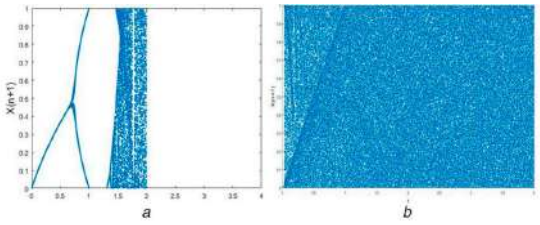


Fig. 1: Bifurcation diagram of (a) Quadratic map, (b) EQM.

II. Enhanced Quadratic Map (EQM) :

In this section, We are employing an enhanced quadratic map as described in [6]. This map overcomes the difficulties of previous maps, such as the restricted chaotic range and the disorganized distribution of chaotic sequences.The EQM's mathematical formula is as follows:

$$X_{n+1} = \mathrm{EQM}\left(r, 2^k\left(X_n\right)\right) = Q\left(r, 2^k\left(Q\left(r, X_n\right)\right) \bmod 1 \right.$$
$$\left. = \left(r - 2^k\left(r - X_n^2\right)^2\right) \bmod 1.\right.$$

the enhanced quadratic map is assessed by the Lyapunov exponent and the bifurcation diagram in Figure 1 and 2.According to Figure 1, the recommended quadratic map exhibits chaotic behavior at any value of r except the extremely tiny range [0, 0.05],furthermore , the chaotic sequences generated by this system have a uniform-distribution in [0.05 ,4].Furthermore, figure 2 demonstrates that the Lyapunov exponent of the enhanced quadratic map is larger than zero for all r values except a tiny range [0, 0.05] .the numerical calculation of largest Lyapunov exponent for the conventional quadratic map and enhanced one is 0.68894 and 5.2713 respectively.Lyapunov?s exponent of enhanced map is larger than that of conventional quadratic maps.It implies that the quadratic map's chaotic properties have been greatly enhanced.

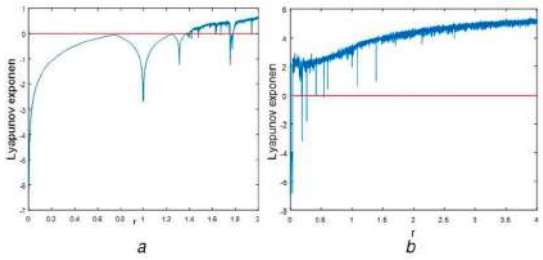


Fig. 2: Lyapunov exponent of (a) Quadratic map, (b) EQM.

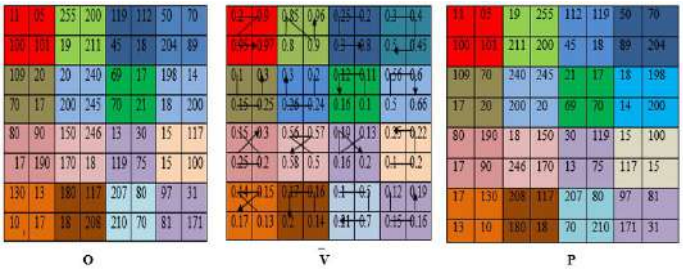


Fig. 3: the mechanism of permutation process.

III. SUGGESTED IMAGE ENCRYPTION SCHEME:

In this part, we will present the proposed technique, which is based on the use of the enhanced quadratic map to change the positions of pixels in the image by a different set of Zigzag types, as it is used in the diffusion process to obtain the encrypted image.

Step 1: We create a random matrix V of size mnusing the enhanced map with the parameters r_1, k_1 and the ini1tial values x_1.

Step 2: We divide the matrix V into blocks of size 22. Then we arrange the elements of each block in ascending order to get different types of Zigzag.

Step 3: We change the locations of the pixels of the original image O with the same types of zigzag of the matrix V to get the permuted image P.Figure 3 shows the mechanism for changing the locations of the pixels of the original image O using the matrix V.

Step 4: We divide the permuted image P into R,G and B components. Then we convert them into one-dimensional vectors R',G' and B'.

Step 5: We diffuse the vectors R', G' and B' using equation 1 and 2 to get the encrypted vectors R", G" and B".

$$\begin{cases} R''_0(i) \leftarrow R'(1) \oplus G'(1). \\ G''_0(i) \leftarrow G'(1) \oplus B'(1). \\ B''_0(i) \leftarrow B'(1) \oplus \text{key } 2(1). \end{cases} \quad (1)$$

key1 and key2 are generated using equation 1 with the parameters r_2,r_3, k_2 and the ini1tial valuesx_2 and x_3.

Step 6: We turn the vectors R?', G?' and B?' into matrices, and then we use these matrices to build the encrypted image.

IV. SIMULATION RESULTS:

In this part, we will give a series of evaluations to assess the efficiency of the recommended approach . Each test was run on an i7 7500u with 8 GB of RAM using Matlab 2019a software.We applied our algorithm to the COVID-19 medical images shown in Figure 5. the encryption results of our approach are shown in figure 6.

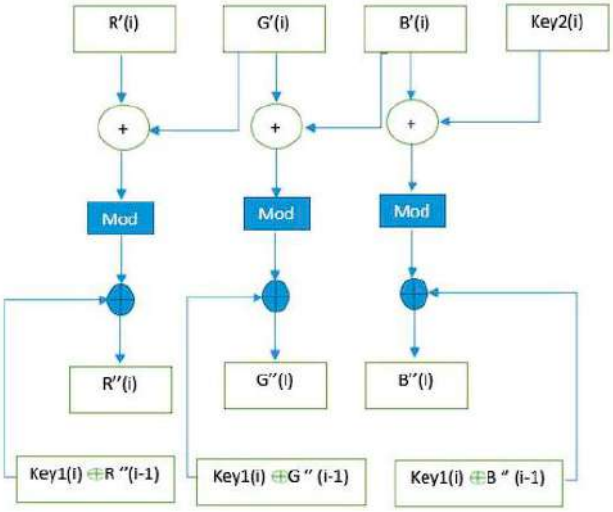


Fig. 4: Diffusion process.

Fig. 5: The original medical images.

$$\begin{cases} R''(i) \leftarrow [(R'(i) + G'(i)) \bmod 256] \oplus [R''(i-1) \oplus \text{key } 1(i)] \\ G''(i) \leftarrow [(B'(i) + G'(i)) \bmod 256] \oplus [G''(i-1) \oplus \text{key } 1(i)] \\ B''(i) \leftarrow [(B'(i) + \text{key } 2(i)) \bmod 256] \oplus [B''(i-1) \oplus \text{keyl}(i)] \end{cases} \quad (2)$$

A. Histogram analysis:

The histogram is a graph that shows the number of pixels in a picture at each intensity value [7].The encrypted images histogram must be uniform so that an adversary cannot extract any information from it.The histogram of the original images and its cipher images is shown in Figure 7.The findings reveal that after encryption, the histograms of the encrypted pictures (Covid1, Covid2, Covid3) are uniform. As a result, the attacker

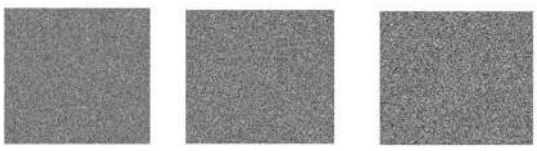


Fig. 6: Encrypted medical images.

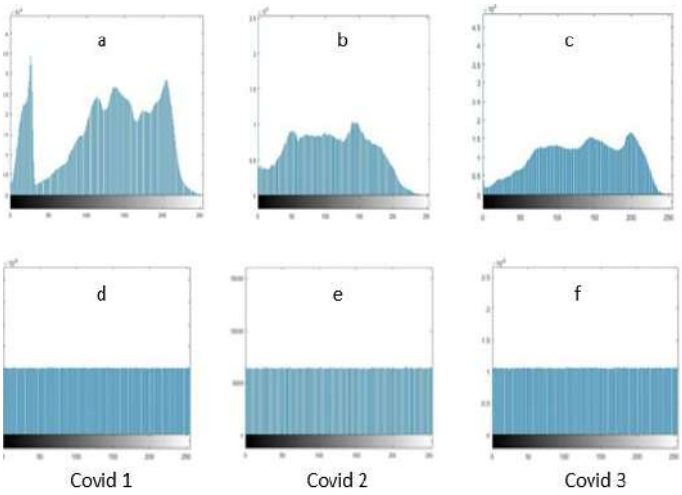


Fig. 7: Histogram analysis results:(a-c) plain images,(d-f) encrypted images.

is unable to extract information from the encrypted image's histogram.

B. Correlation coefficient :

Correlation is a technique for comparing two photos in order to determine pixel changes in one image relative to another reference image. A conventional image's neighboring pixels have a high correlation. To maintain security against statistical analysis, a decent picture encryption technique must eliminate such a correlation.The mathematical formula is as follows [8]:

$$r_{xy} = \frac{cov(x,y)}{\sqrt{D(x)D(y)}} \tag{3}$$

$$cov(x) = E([x - E(x)][y - E(x)]) \tag{4}$$

$$E(x) = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{5}$$

$$D(x) = \frac{1}{N}\sum_{i=1}^{N}[x_i - E(x)]^2 \tag{6}$$

$E(x), E(y)$ denote the means of two adjacent pixel x and y respectively.The results of the correlation coefficient of the ciphered images by our algorithms are shown in Table 1.The outcome demonstrates that after simulating the images, the average value of the correlations. When the encrypted image has a value of 0.001, it is closer to the number 0. This demonstrates that there is no association between neighboring pixels after encryption.

C. Entropy analysis:

The information entropy is a highly important measure of the randomness of information.High entropy values indicate a high degree of unpredictability,and for any message encoded

TABLE I: Correlation coefficient results.

Position	Covid1	Covid2	Covid3
Horizontal	0.00054	0.00012	0.0011
Vertical	-0.0007	0.0004	0.0015
Diagonal	0.0003	-0.0021	-0.0017

TABLE II: Entropy of the ciphered image.

	covid1	covid2	covid3
Entropy values	7,9988	7,9986	7,9990

on M bits, the upper limit of entropy is M.The following formula is used to compute the entropy of a m source:

$$h(m) = -\sum_{i=0}^{2^n-1} p(m_i) log_2(p(m_i)) \tag{7}$$

Where m is the information source. the entropy number for encrypted image must be extremely near to 8.Our entropy results are shown in table 2.The average Entropy of Encrypted Images value is 7.9988, which is near to the value 8. This demonstrates how tough it is to be predictable.

V. Conclusion:

in this study,We concentrated on using a novel image encryption approach via an enhanced quadratic map.This map is utilized to generate the pseudorandom sequence that is employed throughout the zigzag and diffusion procedures. The analytical findings showed that the suggested system has a low correlation coefficient and can withstand various assaults such as statistical attacks and entropy attacks.

References

[1] D. Herbadji, N. Derouiche, A. Belmeguenai, A. Herbadji, and S. Boumerdassi, "A tweakable image encryption algorithm using an improved logistic chaotic map," *Traitement du Signal*, vol. 36, pp. 407–417, 2019.

[2] L. ping Chen, H. Yin, L. guo Yuan, A. M. Lopes, J. A. Machado, and R. chao Wu, "A novel color image encryption algorithm based on a fractional-order discrete chaotic neural network and DNA sequence operations," *Frontiers of Information Technology and Electronic Engineering*, vol. 21, no. 6, pp. 866–879, 2020.

[3] Y. Wan, S. Gu, and B. Du, "A new image encryption algorithm based on composite chaos and hyperchaos combined with dna coding," *Entropy*, vol. 22, 2 2020.

[4] Y. Zhou, L. Bao, and C. L. Chen, "A new 1D chaotic system for image encryption," *Signal Processing*, vol. 97, pp. 172–182, 2014. [Online]. Available: http://dx.doi.org/10.1016/j.sigpro.2013.10.034

[5] X. Zhang, L. Wang, Y. Wang, Y. Niu, and Y. Li, "An Image Encryption Algorithm Based on Hyperchaotic System and Variable-Step Josephus Problem," *International Journal of Optics*, vol. 2020, 2020.

[6] D. Herbadji, A. Belmeguenai, N. Derouiche, and H. Liu, "Colour image encryption scheme based on enhanced quadratic chaotic map," *IET Image Processing*, vol. 14, pp. 40–52, 1 2020.

[7] I. Haddad, A. Belmeguenai, D. Herbadji, and S. Boumerdassi, "Color image encryption based on fractional-order logistic map." Institute of Electrical and Electronics Engineers (IEEE), 6 2022, pp. 1–6.

[8] D. Herbadji, N. Derouiche, A. Belmeguenai, N. Tahat, and S. Boumerdassi, "A new colour image encryption approach using a combination of two 1d chaotic map," *International Journal of Electronic Security and Digital Forensics*, vol. 12, pp. 337–356, 2020.

Packet Delivery Ratio of Wireless UAVs-Ground Base Station Communication in Presence of Interference

Ghania Khraimech, Fatiha Merazka
LISIC Lab. Telecommunications Department,
Electrical Engineering Faculty,
USTHB University,
16111, Algiers, Algeria
{gkhraimech,fmerazka}@usthb.dz

***Abstract*—Unmanned Aerial Vehicles (UAVs) have recently gained significant attention due to their potential applications in various fields, such as agriculture, surveillance, and disaster management. One of the critical challenges for effective UAV operation is establishing reliable communication between the UAVs and ground base stations. In this article, We consider several parameters, such as the distance between the UAVs and the ground base station, the transmission power of the UAVs, and the interference level in the environment. We use the packet delivery ratio (PDR) as a metric to evaluate the performance of the system. Our simulation results show that the proposed communication model can maintain a high packet delivery ratio even in the presence of interference and fading. The proposed system model can be useful in various applications, such as agriculture, surveillance and disaster management, where reliable communication between UAVs and ground base stations is critical.**

***Index Terms*—UAV, ground base stations, interference, fading, PDR,**

I. Introduction

Unmanned aerial vehicles (UAVs) have rapidly evolved in recent years from their initial use in the military and aviation industry to their current, rapidly expanding civilian applications in areas like industrial inspection, scientific research, agricultural practice, security surveillance, emergency rescue, entertainment, and more. The fifth generation (5G) wireless network is currently being prepared for quick implementation, and numerous studies and companies have been looking for various models to accelerate this development and enhance application situations [1]. As it can enable a variety of use cases, including those suggested in the three main principle application scenarios of the International Telecommunication Union (ITU), UAV-aided 5G wireless has attracted a lot of attention [2]. They are specifically massive Machine Type Communications(mMTC), Ultra-reliable low latency communication(URLLC), and enhanced Mobile Broadband (eMBB). For instance, UAV can be extremely helpful in managing other emergency circumstances when uRLLC is needed, improving public safety networks, or offering network service recovery in a disaster-stricken area. In specific, UAV-assisted eMBB can be seen as a crucial addition to the 5G cellular network, which is anticipated to perform 1000 times more successfully than 4G. It has been confirmed that the UAV-enabled 5G communications design problems are more urgent than the conventional 5G communications. This is primarily due to the fact that a UAV-enabled 5G system requires more thorough system designs that take into account several important essential factors, such as wireless communication system [3]. One important aspect of UAV operations is communication, which allows them to transmit data to a ground station or other UAVs. However, UAV communication faces unique challenges, such as limited bandwidth, high altitude, and mobility, which can affect the quality of the communication link [4], [5].

The effectiveness of UAV-enabled communication has previously been examined in a number of studies, including [6], [7]. In particular, a communication system that utilized UAVs as intermediaries between ground endpoints and a BS was explored in [6]. Through the management of the UAV heading angle, a performance optimization method was created in the same article. In [8], the co-channel interference from aerial or ground interfering nodes has been taken into account when analyzing the failure probability (OP) of UAV-based communications for both LoS and non-LoS (NLoS) connections.

Packet Delivery Ratio (PDR) is a commonly used metric to evaluate the performance of communication systems. It measures the proportion of successfully delivered packets out of all the packets sent. In UAV communication, PDR is particularly important as it can affect the reliability and effectiveness of the mission. In this article, we will discuss the use of PDR in evaluating UAV communication [9].

To evaluate UAV communication with PDR, we need to consider several factors that can affect link quality. These include the UAV altitude, speed, distance from the ground station, and the presence of obstacles such as buildings or trees. We also need to consider the type of communication technology used, such as Wi-Fi, cellular, or satellite.

One study conducted by researchers evaluated the performance of Wi-Fi communication between UAVs and a ground station. The researchers measured the PDR for different UAV

altitudes, ranging from 20 to 150 meters. They found that the PDR decreased as the altitude increased, with a sharp drop in performance beyond 100 meters. This was attributed to the weakening of the Wi-Fi signal at higher altitudes in [10].

Another study conducted by researchers evaluated the performance of cellular communication between UAVs and a ground station. The researchers [11] measured the PDR for different UAV speeds, ranging from 0 to 15 m/s. They found that the PDR decreased as the speed increased, with a significant drop in performance beyond 10 m/s. This was attributed to the Doppler effect, which causes a shift in the frequency of the cellular signal as the UAV moves [12].

The following is a description of the paper's structure. The system and channel models examined throughout this work are presented in Section II. Section III describes a performance study of specific scenarios and assumptions made about UAV wireless communication systems. In Section IV, the simulation results are presented and analyzed. Section V, concludes with general conclusions based on the findings of this work.

II. System Model

As depicted in Fig.1, We consider a network of N UAV that are deployed to perform a specific task, such as surveillance or monitoring. Each UAV is equipped with a wireless transceiver that allows it to communicate with other UAVs and a ground base station. The ground base station serves as a control center for the UAVs and receives data from the UAVs. The UAVs and ground base stations communicate using a wireless communication channel. We assume that the communication channel is subject to interference and fading due to various factors, such as atmospheric conditions and obstacles in the environment.

Each UAV is equipped with a wireless transceiver that allows it to communicate with other UAVs and a ground base station. The ground base station serves as a control center for the UAVs and receives data from the UAVs. The UAVs and ground base stations communicate using a wireless communication channel.

During the communication, each UAV sends data to the ground base station by transmitting a packet through the wireless channel. The packet is then received by the ground base station, and the ground base station sends data to the UAVs by transmitting a packet through the wireless channel, which is then received by the UAVs.

The communication channel is subject to interference and fading due to various factors, such as atmospheric conditions and obstacles in the environment. To ensure reliable communication, the system parameters, such as transmission power, packet size, transmission rate, and distance between the UAVs and the ground base station, must be optimized. The optimization aims to improve the packet delivery ratio (PDR), which is the ratio of the number of successfully received packets to the total number of packets transmitted.

A. Path Loss

The path loss is a simplified version of the Friis transmission equation, which calculates the power received by a receiver from a transmitter over a free space propagation path. However, in a real-world environment, the propagation path is not always free space, and there are other factors that affect signal attenuation, such as obstacles, reflection, diffraction, and absorption.

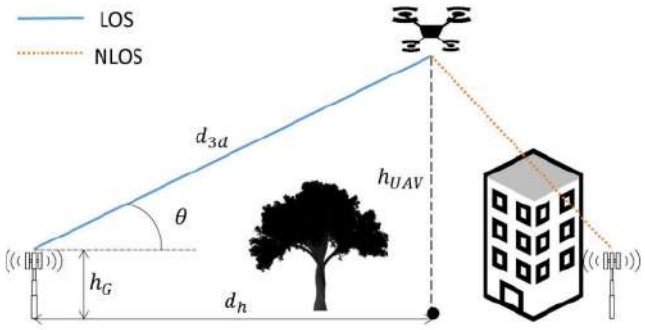


Fig. 1. UAV to ground base station propagation.

To model the path loss accurately, we need to use a more complex path loss model that takes into account these factors. There are several path loss models available in the literature, such as the Okumura-Hata model [13], the [14], and the ITU-R model [15]. The choice of the path loss model depends on the characteristics of the environment, such as the frequency of operation, the type of terrain, and the height of the transmitter and receiver.

Here is an example of the Okumura-Hata path loss model:

$$path\_loss = A + Blog_{10}(distance) + Clog10(frequency) + D + F * log_{10}(h\_receiver). \quad (1)$$

where A, B, C, D, and F are constants that depend on the type of environment, frequency, and height of the transmitter and receiver.

Another example is the COST-231 path loss model:

$$path\_loss = A + Blog_{10}(distance) + Clog_{10}(frequency) + Dlog_{10}(h\_recevier) + Elog_{10}(distance) * log10(frequency) \quad (2)$$

where A, B, C, D, and E are constants that depend on the type of environment, frequency, and height of the transmitter and receiver.

the choice of the path loss equation depends on the specific characteristics of the environment and the frequency of operation.

B. Path Loss and Large-Scale Fading Modeling

In the context of UAV communication, path loss refers to the attenuation of the electromagnetic signal as it travels through space between the transmitter (UAV) and the receiver (ground base station). Large-scale fading, on the other hand, refers to

the variability of the received signal power due to changes in the environment such as obstacles, terrain, and weather.

Path loss and large-scale fading can be modeled using various mathematical equations and models. The most commonly used model for path loss is the Friis transmission equation, which relates the received power to the transmitted power, frequency, distance, and antenna gains. However, this model assumes a free space propagation environment and does not account for obstacles or reflections that can cause signal attenuation.

To account for obstacles and reflections, other models such as the Okumura-Hata model, the COST-231 model, and the Extended Hata model are commonly used. These models take into consideration factors such as the height of the antennas, frequency, and distance between the transmitter and receiver. Large-scale fading can be modeled using statistical models such as the log-normal shadowing model, which assumes that the received signal power follows a log-normal distribution. This model accounts for the variability of the signal power due to environmental factors such as terrain and obstacles.

Path loss and large-scale fading modeling are essential for predicting the performance of wireless communication systems, including UAV communication systems. Accurate modeling of path loss and large-scale fading can help optimize the system design and improve performance

III. Performance Analysis

In order to analyze the performance of UAV communication with PDR, we can vary different system parameters such as the number of UAVs, transmission power, interference level, packet size, and transmission rate, and measure the resulting PDR. By varying the parameters, we can optimize the system for better performance.

- Number of UAVs: The simulation will involve a varying number of UAVs.
- Transmission power of the UAVs: The UAVs will have a fixed transmission power.
- Distance between UAVs and ground base station: The distance between the UAVs and ground base station will be randomly generated within a defined area.
- Interference level in the environment: The interference level in the environment will also be randomly generated.
- Packet size: The packet size will be fixed for all UAVs.
- Transmission rate: The transmission rate will also be fixed for all UAVs.
- Simulation time: The simulation will run for a defined time. UAVs and ground base station
- positions: The positions of the UAVs and ground base station will be randomly generated within a defined area. Packet transmission and reception counters: The counters will be initialized for each UAV. Path loss and fading effects: The path loss and fading effects will be calculated for each UAV and the ground base station.

By optimizing these parameters, we can improve the PDR and enhance the reliability of the communication link between UAVs and the ground base station. We use PDR as a metric to evaluate the performance of the system. Equation (3) shows the calculation of the PDR:

$$PDR = \frac{\sum Total\ number\ of\ packets\ resieved}{\sum Total\ number\ of\ packets\ originated} \quad (3)$$

The PDR is defined as the ratio of the number of successfully received packets to the total number of packets transmitted;

IV. Simulation Results

This section explains the results of average PDR vs. the number of UAVs for different interference levels and transmission powers.

Our simulation, Figs. 2, 3 and 4, results show that the proposed communication model provides a reliable and efficient communication channel for UAVs. We observe that the PDR increases with an increase in the transmission power of the UAVs. We also observe that the PDR decreases with an increase in the distance between the UAVs and the ground base station. Furthermore, we observe that the PDR is affected by the interference level in the environment. However, the system is still able to maintain a high PDR even in the presence of interference. Moreover, the PDR decreases with the increasing number of UAVs, which could be attributed to factors such as interference, limited bandwidth, or increased path loss. The oscillations and non-regular patterns observed could be due to the random nature of the channel or the varying distances between the UAVs and their destination.

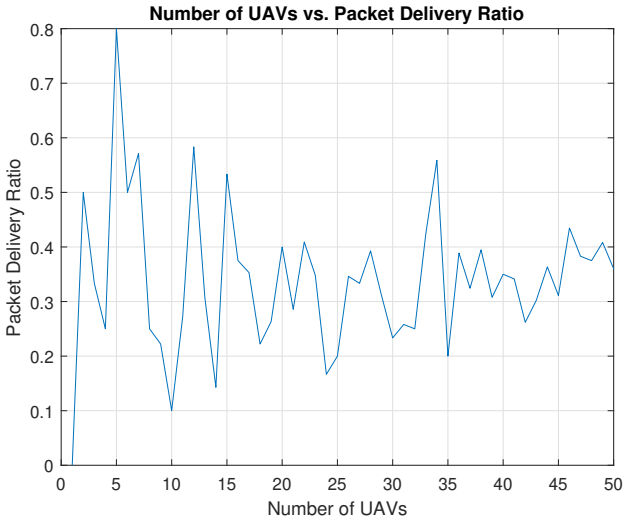


Fig. 2. PDR vs. the number of UAVs N=50 .

V. Conclusion

Overall, these studies demonstrate the importance of evaluating UAV communication with PDR. By measuring PDR, we can gain insights into the performance of the communication link and identify areas for improvement. This can help to increase the reliability and effectiveness of UAV operations in various applications. The recommended system model can be

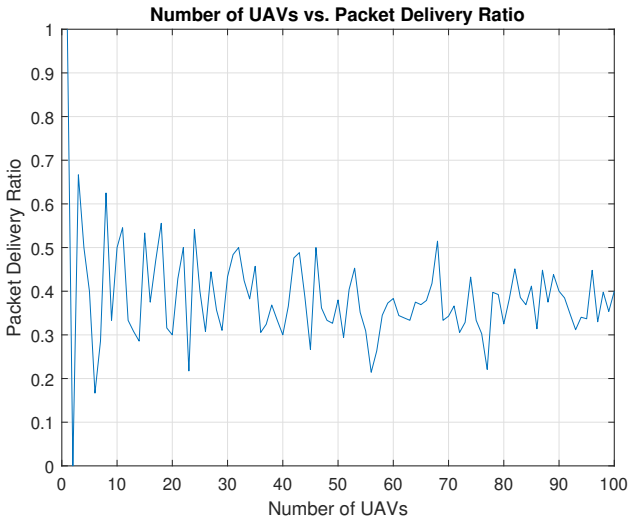


Fig. 3. PDR vs. the number of UAVs N=100.

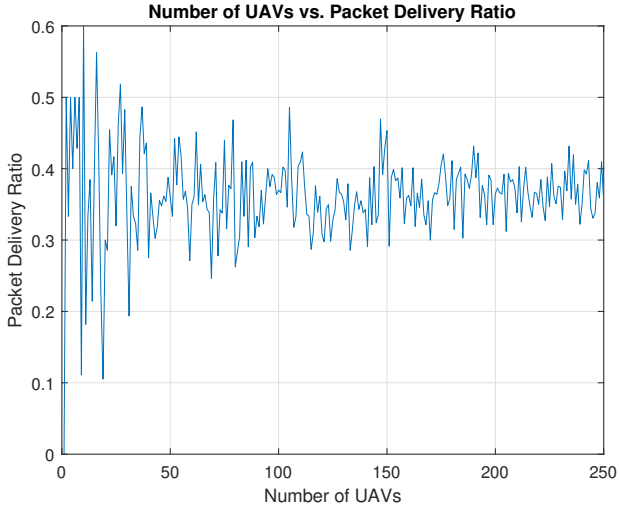


Fig. 4. PDR vs. the number of UAVs N=250

effective in a range of scenarios in which consistent communication between UAVs and ground base stations is essential, including agriculture, monitoring and crisis management.

REFERENCES

[1] Kh., Muhammad Asghar, N. Kumar, S. A. H.n Mohsan, W. U. Khan, M. M. Nasralla, Mohammed H. Alsharif, J. Żywiołek, and I. Ullah, "Swarm of UAVs for network management in 6G: A technical review," IEEE Transactions on Network and Service Management,2022.

[2] M.Mozaffari; W. Saad; M. Bennis; Y.H. Nam; M. Debbah,"A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," IEEE Commun. Surv. Tutor., vol. 21, pp.2334–2360, 2019.

[3] H., Yiming, and al. "Distributed and multi-layer UAV network for the next-generation wireless communication," arXiv preprint arXiv:1805.01534, 2018.

[4] C., Nan, W. Xu, W. Shi, Y. Zhou, N. Lu, H. Zhou, and X. Shen. "Air-ground integrated mobile edge networks: Architecture, challenges, and opportunities," IEEE Communications Magazine 56, no. 8, pp. 26-32, 2018.

[5] A.O. Ahmed, S.A. Busari, and M. Othman, "Physical layer aspects of Terahertz-enabled UAV Communications: Challenges and Opportunities," Vehicular Communications 2022.

[6] P. Zhan, K. Yu, and A. L. Swindlehurst, "Wireless relay communications with unmanned aerial vehicles: Performance and optimization," IEEE Trans. Aerosp. Electron. Syst., vol. 47, no. 3, pp. 2068–2085, Jul. 2011.

[7] B. Galkin, J. Kibida, and L. A. DaSilva, "A stochastic model for UAV networks positioned above demand hotspots in urban environments," IEEE Trans. Veh. Technol., vol. 68, no. 7, pp. 6985–6996, Jul 2019.

[8] M. Kim and J. Lee, "Outage probability of UAV communications in the presence of interference," in IEEE Global Communications Conference, Dec 2018.

[9] U. Saif, K.H. Mohammadani, M. A. Khan Zhi Ren, R. Alkanhel, A. Muthanna, and U. Tariq, "Position-Monitoring-Based Hybrid Routing Protocol for 3D UAV-Based Networks," Drones 6, no. 11, 2022.

[10] V.Evgenii, et al. "Tutorial on UAV: A blue sky view on wireless communication." arXiv preprint arXiv:1901.02306, 2019.

[11] A. Ratnam, N. Vijaya, "Performance Analysis and Enhancement of Unmanned Aerial Vehicle-Based Free Space Optical Communication System," Diss. National Institute of Technology Karnataka, Surathkal, 2022.

[12] T. Wigren and S. Yasini," Passive UAV tracking in wireless networks," IEEE Trans. Aerospace and Electronic Systems, vol. 58, no. 5, pp. 4101-4118, 2022.

[13] A.Nawal, et al." Characterization of Propagation Models in Wireless Communications for 4G Network," International Journal of Electronics and Telecommunications, 68, 2022.

[14] A.H. AH, et al. "Wireless transmissions, propagation and channel modeling for IoT technologies: Applications and challenges," IEEE Access 10, 2022.

[15] O. Zahid and S. Salous. "Long-Term Rain Attenuation Measurement for Short-Range mmWave Fixed Link Using DSD and ITU-R Prediction Models," Radio Science 57.4, 2022.

Segmentation of knee images using U-Net deep learning models

Ibrahim Hamidat
LIST Laboratory,
University M'Hamed Bougara of Boumerdes,
35000, Algeria

Khaled Harrar
LIST Laboratory,
University M'Hamed Bougara of Boumerdes,
35000, Algeria

***Abstract*—Knee bone segmentation is an essential task in medical image analysis, as it can aid in the diagnosis and treatment of knee-related diseases. In recent years, deep learning for semantic segmentation has become a popular tool for this purpose. In this study, we propose using three deep learning models for knee image segmentation: U-Net, Attention U-Net, and Attention Residual U-Net. To compare the performance of these models, we used a dataset of knee MRI images and preprocessed them using normalization and data augmentation techniques. We trained the models using the Adam optimizer and evaluated their performance using Jaccard and Dice coefficient metrics. Our results demonstrate that the Attention Residual U-Net outperforms both the U-Net and Attention U-Net models with an Intersection over Union (IoU) of 91.71% on the test images. This study demonstrates the effectiveness of light deep learning models for image segmentation and highlights the potential of the Attention Residual U-Net model for accurate and efficient segmentation.**

***Index Terms*—Semantic Segmentation, Knee bones, Medical radiography images, U-Net, Attention U-Net, Attention Residual U-Net.**

I. Introduction

Knee-related diseases are a significant cause of pain and disability worldwide. Accurate diagnosis and treatment of such diseases require reliable and efficient analysis of knee images. Automated image segmentation in the medical field is a critical step in the analysis of medical images and it has been widely studied by the community of image analysis as it allows for the identification and localization of regions of interest. This is why accurate and reliable methods are desired to improve the efficiency of clinical work.

Knee image segmentation has been a challenging task due to the complex and variable anatomy of the knee joint. With the advancements of neural networks, such as (CNNs), models can achieve performance levels comparable to those of radiologists. This can greatly reduce the amount of work required in image analysis tasks

Deep learning models have shown great promise in medical image segmentation tasks, including knee image segmentation, which can help in the diagnosis and treatment of knee joint diseases. As a popular tool for this task, deep learning can automatically learn to segment bone structures from medical images. There are several popular deep learning models used for medical image segmentation, such as U-Net, attention U-Net, and attention residual U-Net. The U-Net model uses an encoder-decoder architecture with skip connections to capture both local and global information. Attention U-Net incorporates attention mechanisms to improve the model's focus on relevant features, while attention residual U-Net combines the benefits of residual connections and attention mechanisms for improved performance.

The U-Net architecture is a widely used model for image segmentation, but recent advancements have led to the development of new architectures that aim to improve its performance, as for example; Huang et al. [1] proposed a modified version of attention residual U-Net for slice-based brain tumor segmentation, they used a dataset of 285 brain images only. Caron et al. [2] made a segmentation of trabecular bone microdamage using U-Net. Hwang et al. [3] proposed a segmentation model for a small sample of adherent bone marrow cells. Oktay et al. [4] have proposed an attention U-Net model which has the ability to learn where to look for the pancreas, and Ding et al. [5] also proposed a U-Net architecture multi-scale convolutional network for pediatric hand bone segmentation in X-Ray image. Attention U-Net and attention residual U-Net are two such architectures that incorporate attention mechanisms and residual connections to improve the performance of U-Net. As Alom et al. when Jin et al. [6] made detection early fractures which could be deadly to elderly. Promising results have been published using these models. However, there is still room for improvement, and further research in future.

In this paper, we propose the use of U-Net, Attention U-Net, and Attention Residual U-Net for knee image segmentation and evaluate their performance on a knee dataset with a very low number of parameters yet still remain a stabilized efficiency. We compare the performance of these light modified models and highlight their strengths and limitations. Our study contributes to the growing body of literature on deep learning models for knee image segmentation and provides insights into their potential for clinical applications.

II. Related work

From the vast research papers in the literature, various studies have focused on Knee image segmentation using deep learning models. Do et al. [7] suggested a Multi-Level Seg-U-Net Model with Global and Patch-Based X-ray Images for

Knee Bone Tumor Detection. Liu et al. [8] proposed a multi-level attention U-Net for knee bone and cartilage segmentation. They used a dataset of 216 knee MRI images. Kulseng et al. [9] uses 3D Knee MRI segmentation of a 46 participants and landmark localization to diagnose and treat the fracture risk.

Attention mechanisms and Residual connections have also been used in knee image segmentation to improve the models' performance. Alom et al. [10] (2019) proposed a recurrent residual U-Net with attention mechanisms for knee bone and cartilage segmentation. They used a dataset of 70 knee MRI images and achieved an average Dice coefficient of 0.90. Wang et al. [11] (2020) proposed an attention U-Net with an improved structure for knee bone and cartilage segmentation. They used a dataset of 300 knee MRI images and achieved an average Dice coefficient of 0.91.

Although these studies have shown promising results, there is still a need for further research to explore the potential of deep learning models for knee image segmentation. In this paper, we propose the use of U-Net, Attention U-Net, and Attention Residual U-Net for knee image segmentation and compare their performance on a knee radiography dataset.

III. Methods

A. Data collection and preprocessing

We used a dataset of 209 knee images from the Osteoarthritis Initiative (OAI) dataset [12]. The images were preprocessed and augmented by variating the value of contrast we ended having 450 image data, then we normalized the pixel intensities to have zero mean and unit variance.

B. Model architecture

We implemented three deep learning models for knee image segmentation: U-Net, Attention U-Net, and Attention Residual U-Net using the Keras deep learning framework. We removed the convolutions of 8x8xFN(Filter numbers) to reduce the hyper-parameters from 33 million to less than 8 million and then we made it trainable using low performance machines which has only 4Gb of GPU's VRam (Table I).

TABLE I
NUMBER OF PARAMETERS FOR THE THREE MODELS

Methods	Epoch	Filters	Total Params
U-Net	50	64	7.79 Million
Att U-Net	50	58	7.6 Million
Att Res U-Net	50	50	5.92 Million

The U-Net model had an encoder-decoder architecture with skip connections. The Attention U-Net model incorporated attention mechanisms into the U-Net architecture. The Attention Residual U-Net model combined the benefits of residual connections and attention mechanisms (Fig. 1).

Each model is an improvement from one another, for example:

1) *U-Net:* The U-Net architecture was first introduced in 2015 by Ronneberger et al. [13] for biomedical image segmentation. U-Net is an encoder-decoder architecture that consists of two main parts: the contracting path (Fig.2 (1)) and the expansive path (Fig.2 (2)).

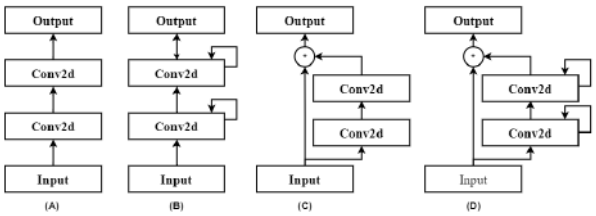


Fig. 1. Different concepts used in the U-Net version

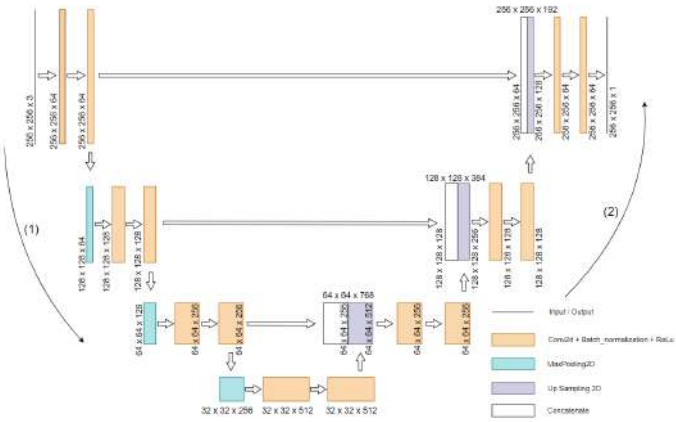


Fig. 2. The U-Net architecture

The contracting path is responsible for capturing the context of the input image, while the expansive path generates the segmentation mask. The contracting path is made up of several convolutional layers, each followed by a max-pooling layer.
The expansive path is composed of a series of up-sampling layers followed by convolutional layers. Skip connections between the contracting and expansive paths help to preserve spatial information and improve the segmentation results.

2) *Attention U-Net:* Attention U-Net was introduced in 2018 by Oktay et al. [4] as a modification of the U-Net architecture (Fig.3). Attention U-Net incorporates attention gates, which allow the network to focus on relevant regions of the input image [10]. Attention gates are added to the skip connections between the contracting and expansive paths, and they use a sigmoid activation function to generate a spatial map that highlights important regions of the input image. This attention mechanism helps to improve the accuracy of the segmentation results by enabling the network to selectively attend to the most informative regions of the input image.

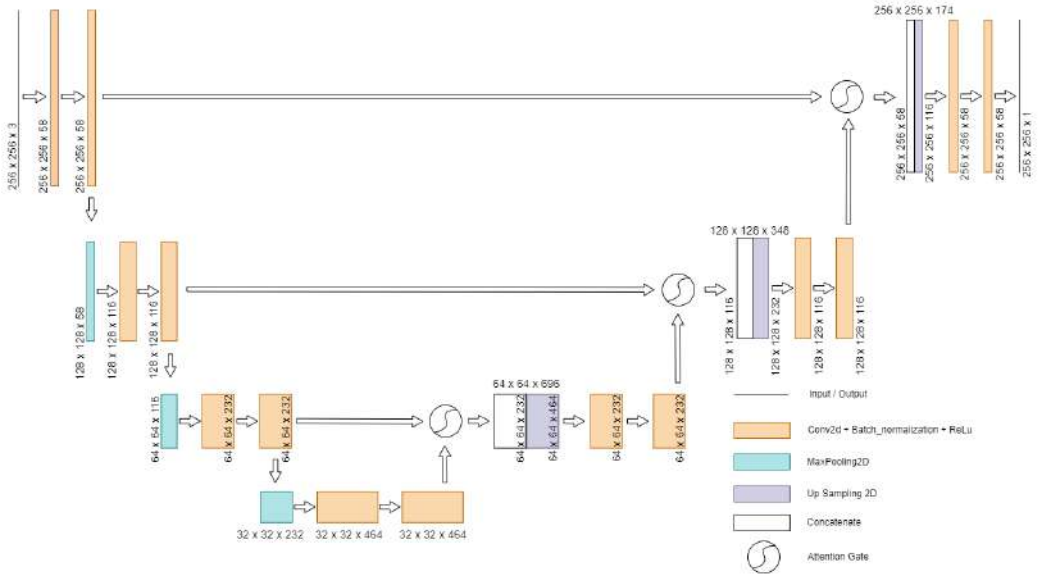


Fig. 3. The attention U-Net architecture

3) *Attention Residual U-Net:* Attention Residual U-Net, introduced in 2018 by Alom et al. [10], is a modification of the U-Net architecture (Fig. 4) that integrates attention gates and residual connections to enhance network performance. Residual connections enable the network to learn residual mappings, which are added to the output of convolutional layers to reduce the gradient vanishing problem and improve the training process. Attention gates are added to the residual connections to selectively attend to relevant regions of the input image. This combination of attention gates and residual connections leads to improved performance compared to both U-Net and Attention U-Net.

C. Training and Evaluation

We randomly splited the dataset into training (70%), validation (20%), and test (10%) sets. We used the training set to train the models using the Adam optimizer with a learning rate of 0.001 and a batch size of 2. We trained each model for 50 epochs and we used binary cross-entropy loss as the objective function. and with different number of parameters for each model (Table I). We evaluated the models' performance on the test set using the Dice coefficient (DSC) and the Jaccard coefficient metrics.

IV. Results and Discussion

The Attention Residual U-Net model outperformed the other models in terms of both DSC (0.9653) and IoU (0.9171) on the test set (Tab. II) unlike during the training and validation process (Tab. III) which gave the lowest values but was more efficient on test images. The U-Net model had a DSC of (0.9631) and an IoU of (0.9138), while the Attention U-Net model had a DSC of (0.9642) and an IoU of (0.9143). The Attention Residual U-Net model showed better segmentation of the knee bone and cartilage structures.

TABLE II
Metric Values for the Testing Results

Methods	DSC	IoU
U-Net	0.9631	0.913878
Att U-Net	0.9642	0.914375
Att Res U-Net	0.9653	0.917172

Our results demonstrate the potential of deep learning models, particularly Attention Residual U-Net, for knee image segmentation. The high performance of these models can help clinicians in accurate diagnosis and treatment of knee-related diseases.

TABLE III
Training and Valitation Metric Results

	Methods	Accuracy	Jaccard	Dice
Train	U-Net	0.971345	0.840333	0.912434
	Att U-Net	0.970738	0.822617	0.901809
	Att Res U-Net	0.971027	0.828688	0.905518
Val	U-Net	0.970023	0.823511	0.902729
	Att U-Net	0.970028	0.820267	0.900836
	Att Res U-Net	0.970526	0.809791	0.894073

A. U-Net

The U-Net model achieved good performance with a Dice coefficient of 0.9631 and an IoU of 0.9138 on the test set. The model was able to segment the knee structures effectively

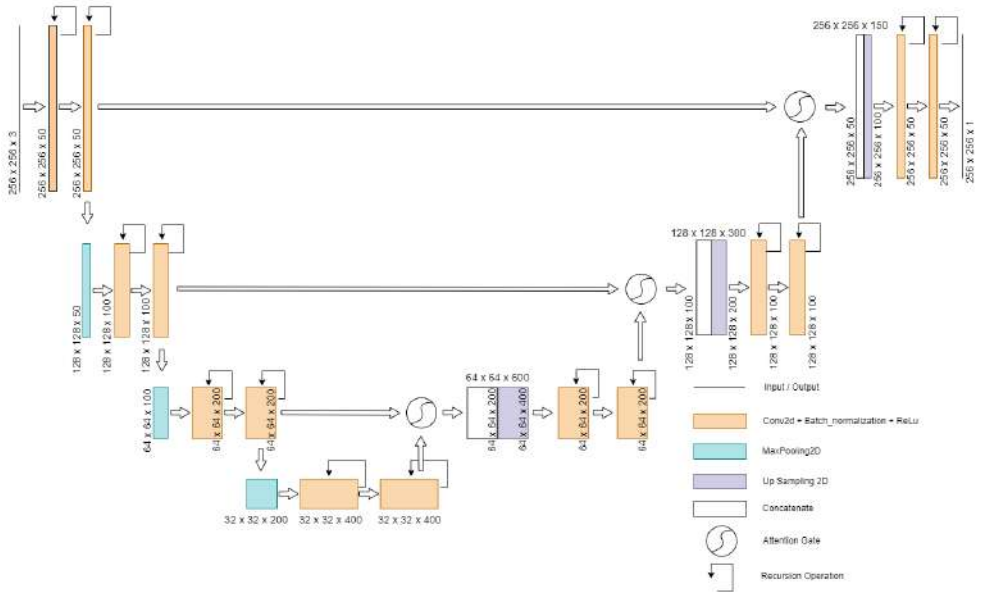


Fig. 4. The Attention residual U-Net architecture

with only 7.79 million hyper-parameters, but it showed some limitations in capturing fine details of the knee bone and cartilage structures. This is likely due to the low training epochs as well as the lack of attention mechanisms and residual connections in the U-Net architecture.

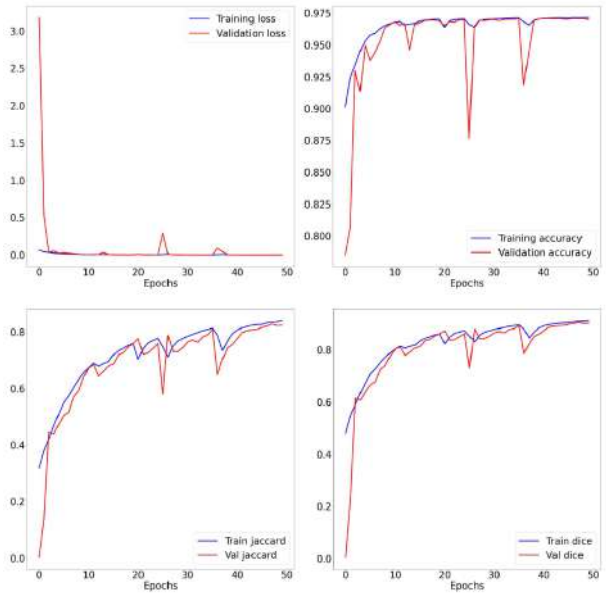


Fig. 5. Graphic results for the loss, accuracy, Jaccard, and dice metrics during the training process for the U-Net model

Figure 5 shows that the model starts with a high loss and low accuracy, but quickly improves within the first few epochs. By the 5th epoch, both the loss and accuracy have stabilized, with the loss fixed at around 0.0029 and the accuracy at around 0.9713.

The Jaccard coefficient and dice coefficient started low but improved quickly, reaching their max values by the 9th epoch. There is a small dips in both metrics, but they quickly recovered and continued to improve, with dice coefficient reaching almost 0.92 by the final epoch.

It can be noticed that the validation metrics follow a similar trend to the training metrics. However, the validation metrics are consistently lower than the training metrics for Dice and Jaccard, indicating that the model may be over-fitting to the training data but accuracy clearly shows a good fit.

B. Attention U-Net

The Attention U-Net model performed better than the U-Net model, achieving a DSC of 0.9642 and an IoU of 0.9143 on the test set. The attention mechanisms in the model improved the segmentation of knee structures by focusing on important regions and suppressing irrelevant information. However, the model still showed some limitations in capturing fine details of the knee bone and cartilage structures, and that's probably due to the low epoch number and the lower number of parameters which are around 7.6 million params (Table I).

Figure 6 shows that loss and accuracy quickly improved within the first few epochs, and they reached their stability values by the 9th epoch for around 0.008 and 0.964. The best value reached by the Loss is at around 0.0028 and the accuracy at around 0.9707. The Jaccard coefficient and dice coefficient started reaching their max values by the 12th epoch with small dives in both metrics, but they quickly recovered and continued to improve, with dice coefficient reaching almost 0.91 by the final epoch.

The graphs also show that the validation metrics follow a similar trend to the training metrics (Table III). However, the validation metrics are consistently lower than the training

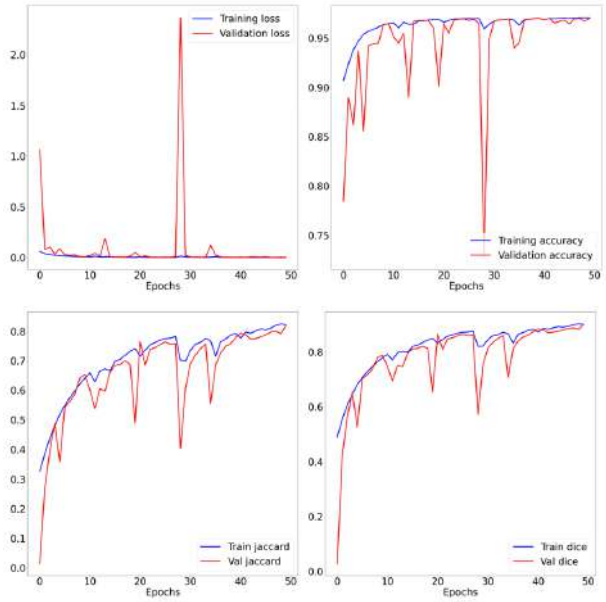


Fig. 6. Graphic results for the loss, accuracy, Jaccard, and dice metrics during the training process for the Attention U-Net model

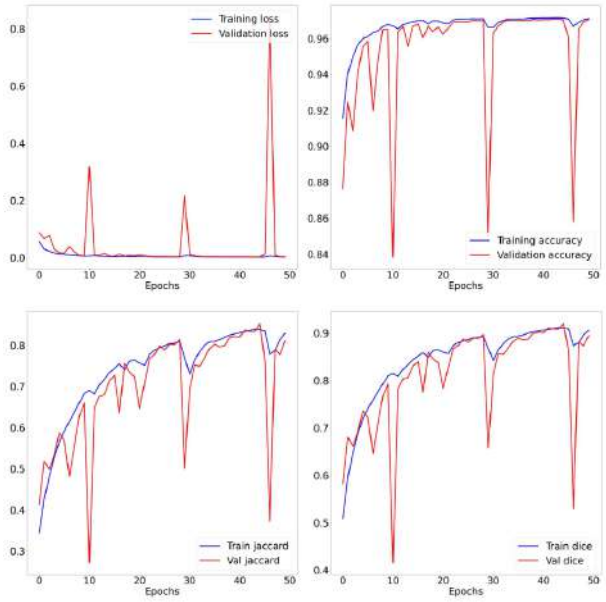


Fig. 7. Graphic results for the loss, accuracy, Jaccard, and dice metrics during the training process for the Attention Residual U-Net model

metrics for Dice and Jaccard, indicating that the model may be over-fitting to the training data, but compared to the number of parameters and filters which been used, the results are satisfying.

C. Attention Residual U-Net

The Attention Residual U-Net model achieved the best performance with a DSC of 0.9653 and an IoU of 0.9171 on the test set (Table II). The model's residual connections and attention mechanisms [10] improved the segmentation of knee structures significantly by enabling the model to capture fine details of the knee bone and cartilage structures. The attention mechanisms in the model helped to highlight important regions, while the residual connections allowed the model to recover finer details of the segmented structures.

Figure 7 shows that loss and accuracy quickly improved within the first few epochs reaching their stability by the 12th epoch for around 0.006 and 0.967. The best value reached by the Loss is around 0.0024 and the accuracy 0.9710. The Jaccard coefficient and dice coefficient started reaching their max values by the 14th epoch with few dives in both metrics, but they quickly recovered and continued to improve, with dice coefficient reaching almost 0.91 by the final epoch.

The graphs also show that the validation metrics follow a similar way to the training metrics. However, the validation metrics are also consistently lower than the training metrics for Dice and Jaccard, indicating that the model may be over-fitting to the training data, but accuracy validation graph shows a good fitting compared to the gap in parameters and filters with U-Net and attention U-Net.

D. Comparison of the models

Our study demonstrates that deep learning models, particularly Attention Residual U-Net, can achieve high performance in knee image segmentation. The use of residual connections and attention mechanisms improved the models' performance compared to the traditional U-Net model. Our results are consistent with previous studies that have reported the benefits of attention mechanisms and residual connections in knee image segmentation (Alom et al. [10]; Chen et al. [16]; Hwang et al. [3]).

Fig. 8. Original Image and its corresponding mask

The high performance of the Attention Residual U-Net model can help clinicians in accurate diagnosis and treatment of knee-related diseases. Accurate segmentation of knee bone and cartilage structures can aid in the detection of abnormalities and assist in surgical planning. The use of deep learning models can also reduce the need for manual segmentation, which is time-consuming and prone to inter-observer variability. However, our study has some limitations. First, we used a relatively small dataset of knee radiography images with its corresponding masks (Fig. 8). Second, we only

evaluated the models' performance on knee bone and cartilage segmentation. Future studies could investigate their potential for segmentation of other knee structures, such as ligaments and tendons. Third, We only reduced the number of parameters to be capable of running on low performance machines with the highest possible efficiency. Further studies will focus on reduce the minimum required performances more with even higher efficient segmentation.

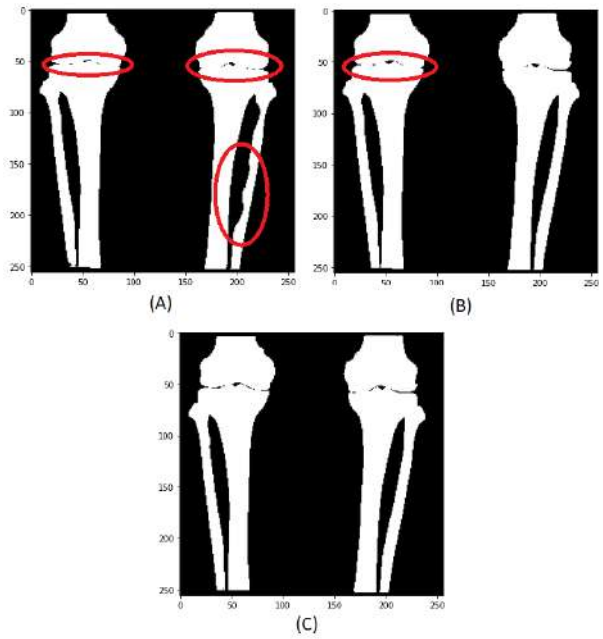


Fig. 9. The predicted masks of the three models, (A) Prediction of the U-Net model, (B) Prediction of the Attention U-Net model, (C) Prediction of the Attention Residual U-Net model.

Overall, the Attention Residual U-Net model outperformed the other models, demonstrating the importance of attention mechanisms and residual connections in knee image segmentation. The U-Net and Attention U-Net models showed good performance, but their limitations in capturing fine details (Fig. 9) of the knee structures suggest that they may not be suitable for more complex segmentation tasks.

V. Conclusion

In summary, U-Net is a popular architecture for image segmentation, but recent advancements have led to the development of new models that aim to improve its performance. Attention U-Net incorporates attention gates to selectively attend to relevant regions of the input image, while Attention Residual U-Net incorporates both attention gates and residual connections to further improve the performance of the network. These newer models have shown improved performance compared to the original U-Net architecture and are therefore worth considering for image segmentation tasks. Our study shows that Attention Residual U-Net is a promising model for knee image segmentation. Further research is needed to explore their potential for clinical applications and to investigate their performance on larger datasets.

References

[1] Huang, Z., Zhao, Y., Liu, Y., & Song, G. (2021). GCAUNet: A group cross-channel attention residual UNet for slice based brain tumor segmentation. Biomedical Signal Processing and Control, 70, 102958.

[2] Caron, R., Londono, I., Seoud, L., & Villemure, I. (2023). Segmentation of trabecular bone microdamage in Xray microCT images using a two-step deep learning method. Journal of the Mechanical Behavior of Biomedical Materials, 137, 105540.

[3] Hwang, Eo-Jin, Sanghee Kim, and Joon-Yong Jung. "Fully automated segmentation of lumbar bone marrow in sagittal, high-resolution T1-weighted magnetic resonance images using 2D U-Net." Computers in Biology and Medicine 140 (2022): 105105.

[4] Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., ... & Rueckert, D. (2018). Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999.

[5] Ding, L., Zhao, K., Zhang, X., Wang, X., & Zhang, J. (2019). A lightweight U-Net architecture multi-scale convolutional network for pediatric hand bone segmentation in X-ray image. IEEE Access, 7, 68436-68445.

[6] Jin, L., Yang, J., Kuang, K., Ni, B., Gao, Y., Sun, Y., ... & Li, M. (2020). Deep-learning-assisted detection and segmentation of rib fractures from CT scans: Development and validation of FracNet. EBioMedicine, 62, 103106.

[7] Do, N. T., Joo, S. D., Yang, H. J., Jung, S. T., & Kim, S. H. (2019, February). Knee bone tumor segmentation from radiographs using Seg-Unet with dice loss. In Proceedings of the 25th International Workshop on Frontiers of Computer Vision (IW-FCV), Gangneung, Korea (Vol. 20).

[8] Liu, J., Hua, C., Zhang, L., Li, P., & Lu, X. (2021). Knee Cartilages Segmentation Based on Multi-scale Cascaded Neural Networks. In Machine Learning in Medical Imaging: 12th International Workshop, MLMI 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings 12 (pp. 20-29). Springer International Publishing.

[9] Kulseng, C. P. S., Nainamalai, V., Grøvik, E., Geitung, J. T., Årøen, A., & Gjesdal, K. I. (2023). Automatic segmentation of human knee anatomy by a convolutional neural network applying a 3D MRI protocol. BMC Musculoskeletal Disorders, 24(1), 1-12.

[10] Alom, M. Z., Hasan, M., Yakopcic, C., Taha, T. M., & Asari, V. K. (2018). Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. arXiv preprint arXiv:1802.06955.

[11] Wang, L., Yang, J., Jin, M., Ding, Y., & Xia, Y. (2020). Attention U-Net with improved structure for knee cartilage segmentation from magnetic resonance images. Journal of Medical Imaging and Health Informatics, 10(3), 576-582. https://doi.org/10.1166/jmihi.2020.2959

[12] G. Lester, "The osteoarthritis initiative: an NIH public-private partnership", HSSJ, vol. 8, pp. 62-63, 2012.

[13] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015.

[14] Chen, Xiaocong, Lina Yao, and Yu Zhang. "Residual attention U-Net for automated multi-class segmentation of covid-19 chest ct images." arXiv preprint arXiv:2004.05645 (2020).

[15] Li, X., Lv, S., Li, M., Zhang, J., Jiang, Y., Qin, Y., ... & Yin, S. (2023). SDMT: Spatial Dependence Multi-Task Transformer Network for 3D Knee MRI Segmentation and Landmark Localization. IEEE Transactions on Medical Imaging.

[16] Chen, H., Qi, X., Yu, L., Dou, Q., Qin, J., & Heng, P. A. (2018). DCAN: Deep contour-aware networks for object instance segmentation from histology images. Medical image analysis, 36, 135-146.

Fast and accurate diagnosing of COVID-19 cases using Convolutional neural network

Djamel Herbadji
Electonics Research Laboratory
University 20 August 1955
Skikda 21000, Algeria

Abderrahmane Herbadji
Department Electonics
University of Setif-1
Setif 19000, Algeria

Ismail Haddad
Electonics Research Laboratory
University 20 August 1955
Skikda 21000, Algeria

Hichem Kahia
LMSE Laboratory
University of biskra
Biskra 7000, BP145, Biskra, Algeria

Aissa Belmeguenai
Electonics Research Laboratory
University 20 August 1955
Skikda 21000, Algeria

***Abstract*—The Coronavirus 2019 (COVID-19), is a viral disease that often causes pneumonia in humans. This virus impacts different parts of body depending on patient's immune system. This infection was first reported in Wuhan city of China in December 2019. After that, it became a global pandemic. Artificial intelligence models have been helpful for successful analyses in the biomedical field. Therefore, this study proposes the detection of COVID-19 from X-ray images using a deep learning model, which is a sub-branch of artificial intelligence. The proposed method is developed to serve as an accurate diagnostics for 2 class classification (i.e. COVID and Normal). The experimental results show high accuracy, thus, the proposed deep learning method could be employed as an initial detection tool to assist the radiologists in fast and accurate diagnosing of COVID-19 cases.**

***Index Terms*—Deep learning, Machine learning, AI, COVID-19, Classification.**

I. Introduction

Coronavirus (COVID-19) is one of the most deadly and dangerous diseases in the world. Recently, the first case of COVID-19 was identified in Wuhan, China [1].The World Health Organization declared COVID-19 an international public health emergency on January 30, 2020 [2].The whole world is facing more than 14.78 million cases. In total, nearly 3,124,726 deaths have been recorded [3].COVID-19 diagnostic systems take time and people spread the virus without even realizing they are truly infected.Viral pneumonia is also caused by a virus, but COVID-19 pneumonia spreads rapidly through the lungs, causing significant damage to lung cells [4-6].Thus, an accurate test method is needed because the inability to detect people with COVID-19 can delay treatment, increasing the risk of spreading COVID-19 infection to others.To date, the real-time reverse transcription polymerase chain reaction (rRT-PCR) test is the best test method for characterizing MERS-CoV [7].RT-PCR can identify CoV-2 RNA from respiratory samples (oral or oral swabs).However, the sensitivity of the RT-PCR test is overshadowed by the limited accessibility of the test kit and the time required for the test result, which usually takes a few hours to one or two days[8].Several researchers have tried to develop an alternative method for rRT-PCR, such as experiments on radiographic images. Their analysis shows that COVID-19 chest x-ray and CT data have unique properties such as the opacity of frosted glass.However, manual scanning of these radiographs takes time. Researchers seek to develop effective COVID-19 detection systems using deep neural networks and machine learning techniques.

978-1-6654-8042-0/22/$31.00

In this paper, a deep-learning COVID-19 detection system that contributes to the automatic detection of coronavirus disease using chest x-ray images is proposed, and a deep CNN model is proposed to extract distinct features and a high level of X-rays.images across a COVID-19 Limited dataset, the model is evaluated using a sufficient number of COVID-19 images. The proposed system can make a significant contribution to the effective diagnosis of COVID-19 in the medical field.

II. MATERIAL AND METHODOLOGY

This section presents the step-by-step methodology needed to develop predictive models to predict future cases of COVID-19. The model requires COVID-19 data such as confirmed cases and recovered cases as input to predict future data corresponding to a specified time period. From the obtained data, a data cleaning and normalization process is performed to remove unwanted fields. The feature is extracted by selecting the dependent and independent variables from the data.The proposed models actively learn real-time data from current COVID-19 observations to predict future outbreaks.

A. Dataset description

As we work with datasets, the machine learning algorithm runs in two steps. We generally divide the data between 20% and 20% between the testing and training phases. As part of supervised learning, we divide the data set into training data and test data in Python ML.

The proposed model was formed and tested on a pooled dataset, which consists of chest x-ray images taken from two

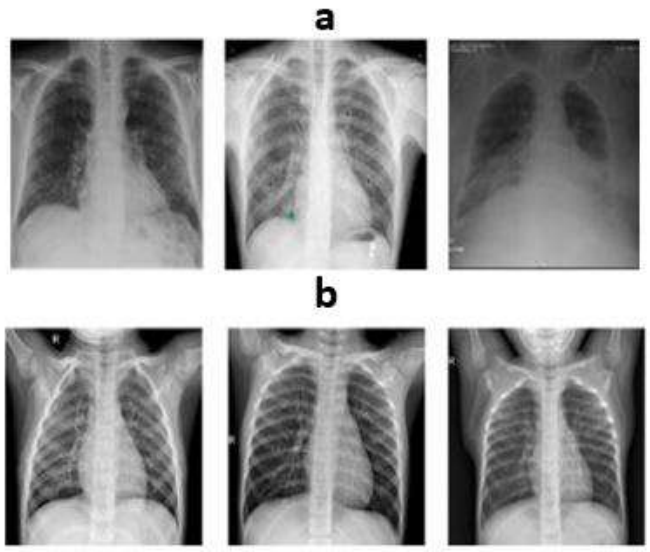


Fig. 1. (a) X-ray images of COVID-19, (b) normal state.

TABLE I
SUMMARY OF PARTITIONING THE DATASET INTO THREE SETS.

Dataset	COVID-19	Normal	Total
Training	920	1108	**2028**
Testing	200	235	**435**
Validation	200	235	**435**
Total	1320	1578	**2898**

different sources to diagnose COVID-19. This combination makes our model more reliable and less sensitive to bias. The source of the dataset is: Chest x-ray images of COVID-19 patients were obtained from the Kaggle COVID-19 X-ray database created by researchers at the University of Dhaka and the University of Qatar, as well as clinicians and collaborators.This database is constantly being updated, so the number of images available in these repositories may change in the future. This dataset contains 1,320 COVID-19 chest x-rays.The second source of chest X-rays was obtained from the Kaggle Repository and contains 1578 chest X-rays. After obtaining a balanced data set.The dataset was mixed and divided into learning sets, test sets and DBAs. Of these, 50% of the images are used for training purposes, 25% of the images are for testing purposes and 25% are validation images.The data was separated into 3 groups: 2028 drive images, 435 check images and 435 test images. Table 1 illustrates this division. Figure 1 shows some typical images of these classes.

B. transfer learning

Recent research has revealed the widespread use of deep CNNs, which provide groundbreaking support for many classification problems.In general, deep CNN models require a large amount of data to achieve good performance. A common challenge associated with the use of such models is the lack of training data.In fact, collecting a large volume of data is a tedious task and no effective solution is available at present. Therefore, the problem of smaller data set is currently solved by using TL technology [9-10],which is very effective in solving the problem of lack of training data.The TL mechanism involves training a CNN model with large amounts of data. In the next step, the model is refined to train on a small set of demand data.Example The student-teacher relationship is an appropriate way to illustrate transformational leadership. The first step is to gather detailed knowledge on the subject[11] Then the teacher introduces a "training course" by imparting information in a "series of lectures" over time. Simply put, the teacher conveys the information to the student.In more detail, the expert (teacher) imparts knowledge (information) to the learner (student). Similarly, the DL network is trained using a huge amount of data, and it also learns biases and weights during the training process.These weights are then transferred to different networks for recycling or testing of a new similar model. So the new model is pre-enabled for strength training rather than requiring training from scratch[12].

C. The architecture of a CNN convolutional neural network

It is a deep learning method that can be used for various tasks, such as object detection, image classification, and other computer vision-related tasks.It showed better performance than traditional machine learning methods. Recently, many researchers have started using deep learning models to diagnose suspected COVID-19 infections.These models show much better results. However, the main drawback is that these models still suffer from processing issues, large number of parameters, usually require a lot of training time and are expensive in real-world applications.The main impetus for developing our shallow CNN architecture is to classify COVID-19 patients based on chest X-ray images while reducing detection time and maintaining high accuracy.The main advantage of this network is its simple and light structure with fewer parameters compared to other models. This greatly reduces computational costs and avoids the possibility of overfitting. As a result, this model can run quickly on low-performance computers.In this section, we propose a six-layer shallow CNN architecture to classify COVID-19. Figure 2 shows the proposed shallow CNN architecture. Our proposed model is based on five components, namely convoluted layer, grouping layer, dense layer, flat layer and activation function.

D. Proposed method architecture

In this study, we used the publicly available dataset of X-ray images of Covid-19. Since these X-ray images were available at different sizes and resolutions, we uniformly resized them to 224×224.We prepared a master model and added pre-learning models such as VGG16, VGG19, RESNET50 and InceptionV3 to get results regarding accuracy, recall, loss and accuracy in the chart. The proposed model is inspired by the work of Şahinbas and Katak . Figure 3 shows the structure of the proposed model and its steps are presented below:

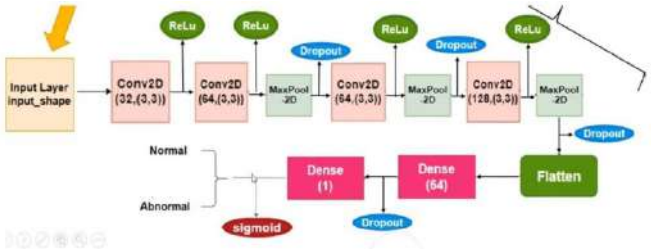


Fig. 2. CNN architecture for the classification of COVID-19 patients.

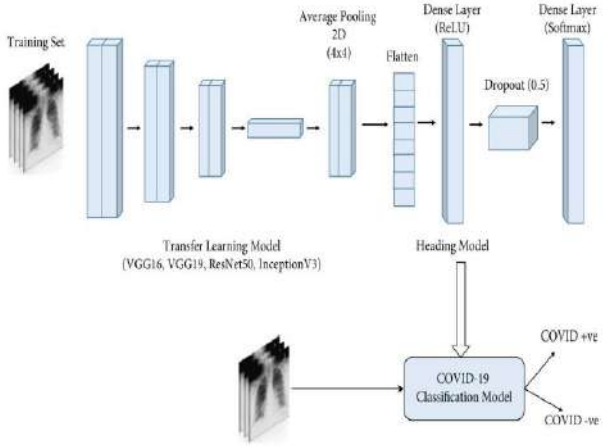


Fig. 3. the architecture of the proposed model.

Step 1: Image Acquisition Initially, x-ray images of Covid-19 patients and non-Covid-19 patients were collected from publicly available sites such as GitHub and Kaggle

Step 2: Refresh data After downloading the dataset, we extracted and collected the labels. All images have been converted from BGR to RGB channels, then resized to 224 x 224. Remember that the purpose of these layers is to filter the image by keeping only distinct information such as atypical geometries.

Step 3; Hot encoding was done on the labels using Label-Binarizer, a Scikit-Learn class that takes input as categorical data and returns an array of Numpy[15].

Step 4; Split and augment the dataset. In this step, the dataset is split into two parts "training" and "testing" with 80% and 20%.

Step 5: Launch of the base model Then we configured the base model with several pre-driven models such as VGG16, VGG19, RESNET50 and ICCECPEV3. However, neither the top of the model nor its head were loaded.

Step 6 Creating the form header. In this step, we have created the basic form header and attached it to the top of the form. (a) The main model starts with the output of the root/base model (b) The average assembly size 4 x 4. is applied. (c) Next, the head model is flattened (d) The dense layer 64 is applied to the head model with the "relu" activation layer (g) A sedimentation layer at a rate of 0.5 is applied to the head model to avoid over processing (e) Finally, a dense layer of size 2 is applied due to binary classification with soft max activation function. When setting up the head model, the base model is placed at the bottom with the head model at high. The complete model was then ready for training.

Step 7:Compilation of the model. Then, the model was compiled using the ADAM optimizer, a combination of the Ada rad and RMSProp algorithms which provides better optimization for noisy data .The initial learning rate selected was 0.001.

Step 8:The model was trained with 25 epochs and 32 lot sizes on 80% of the data.

step 9:Then the model is tested on the remaining 20% of the dataset and gets the desired results for precision, recall, F1 score, specificity, sensitivity, etc.

III. Results and Discussion

In this section, various experiments were performed on the dataset to test the effectiveness of the proposed deep learning models. We performed a comprehensive empirical analysis on X-ray images to predict COVID-19. All models were individually prepared and trained as described above. The proposed work is implemented using Keras package with Python and TensorFlow. The experiment was conducted at the Google Colaboratory using a 16 GB Tesla P100 - PCIe graphics card, a 2.4 GHz Intel Core i5 processor, 16 GB of RAM and a hard drive of 128 GB.

A. Confusion Matrix

Machine learning involves feeding an algorithm with data so that it learns on its own to perform a particular task.In classification problems, it predicts the results that must be compared to reality to measure the degree of its performance.We usually use the confusion matrix, which is a summary of the prediction results for a classification problem.The correct and incorrect predictions are highlighted and divided by class. The results are compared to the actual values. Confusion matrix technology helps measure machine learning classification performance.With this type of model, you can label the model and classify it with known actual values in the test data set and others for validation.This is a method to correctly visualize the number of samples from each label that has been predicted.The beauty of the confusion matrix is that it actually allows us to see where the model has failed and where the model is succeeding, especially when labels are unbalanced.In other words, we can see beyond the precision of the model. This matrix helps to understand how confusing the classification model can be when making predictions.This not only allows us to know what mistakes were made, but especially what kind of mistakes were made.Users can analyze it to determine results that indicate how errors occurred.In addition to machine learning, confusion matrices are also used in statistics, data mining and artificial intelligence. They enable faster analysis of statistical data and facilitate the decoding of results via data visualization. They offer the possibility to analyse errors in statistics, data mining or medical examinations.

B. Classification performance masures

The performance of DL models in determining COVID-related outcomes19 were assessed using various assessment tools and performance measurement is very important to assess the performance of deep learning classification models.Once the training phase was completed, the performance of each model of the test data set was evaluated and compared on the basis of performance measures.It is not enough to have only one result to determine the accuracy. So we used other metrics like. Accuracy, Precision (PPV), Specificity, Sensitivity or Recall, F1 score, these four measurements are normal measurements used in machine learning for classification analysis.The above measures are evaluated using a confusion matrix of four terms: true positive (TP), false positive (FP), true negative (TN) and false negative (FN). All measurements are specified below:

Accuracy:This metric is the most important to evaluate the DL classification.

$$Accuracy = \frac{(PT + TN)}{TP + TN + FP + FN} \tag{1}$$

Precision:It is a measure of precision, calculated by dividing the correct positive predictions (identified as true positives) by the number of all positive predictions.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

Sensitivity, Recall:Perfection is measured, calculated by dividing the number of true positives by the number of true positives

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

Specificity:It is calculated by dividing the number of true negatives by the total number of negatives in the data.

$$Recall = \frac{TN}{TN + FN} \tag{4}$$

F1-score :It is a combination of recall and model accuracy that gives a better measure of misclassified cases

$$F1 = \frac{2(Precision * Sensitivity)}{(Precision + Sensitivity)} \tag{5}$$

PT: This is when the model correctly predicts the positive class. Here, the positive class refers to a patient with COVID-19.

TN: This is when the model correctly predicts the negative class. Here, the negative class refers to a patient NOT suffering from COVID-19.

FP (Type 1 error): when the model incorrectly predicts the positive class. Predicts a patient with COVID-19 but not true.

FN (type 2 error): when the model predicts the negative class incorrectly. Predicts that a patient is NOT suffering from COVID-19 but this is not true

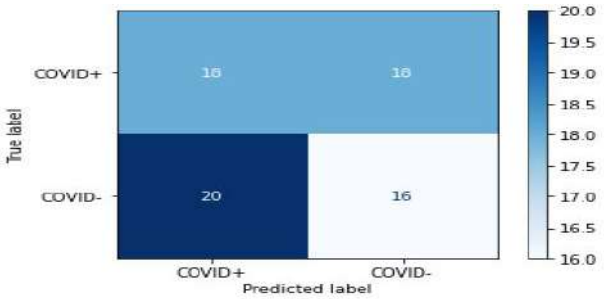


Fig. 4. the confusion matrix.

TABLE II
MODEL-BASED PERFORMANCE EVALUATION

	precision	real	1-score	support
COVID+	0.47	0.50	0.49	36
COVID-	0.47	0.44	0.46	36
Accuracy			0.47	72
Macro avg	0.47	0.47	0.47	72
Weightedavg	0.47	0.47	0.47	72

C. Evaluation and analysis of results

The Confusion Matrix displays 20 of the 38 COVID-19 images included in the normal category and 18 of the 34 normal (non-COVID-19) images classified as COVID-19. Using a confusion matrix, we obtained the other measurements of precision, recall, F1 score, sensitivity and specificity. Avg macro mean For a multi-class classification problem, in addition to class recall, precision and f1 results, we check the macro mean, weight, precision and f1 scores for the entire model. These scores help select the best model for the task at hand. In the above confusion matrix, if we average the precision column, we get 0.47 as shown

D. Differences between micromean and macromean

The calculation and interpretation of these averages varies slightly. Regardless of the scale used, the overall average averages after calculating the scale, regardless of the categories. On the contrary, the partial average will take into account the contributions of each category to calculate the average of the scale. In a multi-category ranking, this approach is often favoured when an imbalance between categories (number, importance, etc.) is suspected.Table 2 shows the model performance score for each of the VOC+ and VOC-classes, the classifier shows an accuracy of 47%, a recall of 50% and a F1 score of 49% For VOC+ cases, it has an accuracy of 47%, an accuracy of 44%, a 44% and 46% F1 grade recall for normal cases (VOC-). The formation results of the proposed models are recorded and presented in the diagram Figure 5.The orange curve is for verification (valdation) and the blue curve is for formation (Accuracy, loss).

Loss : 0.1457 val_{loss}: 0. Accuracy : 0.9622 $val_{accuracy}$: 0.9588 (A) It shows that the amount of training losses decreases rapidly, with an average of about 0.13.During the first

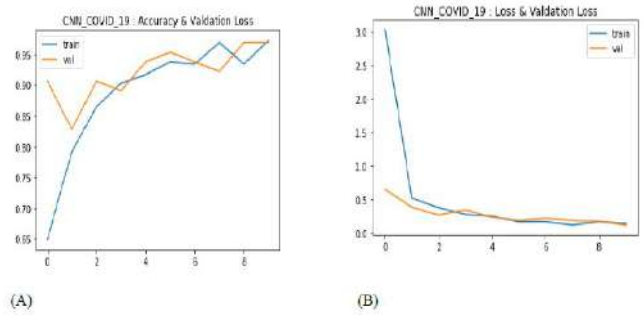


Fig. 5. Graphs of (A) loss and validation loss and (B) accuracy and validation loss.

5 epochs and has continued to decline until I reach almost zero after 8 epochs. As for the rate of losses between tests, its decrease was less severe, which is normal because the data that tested the proposed model were new data. (B) For the precision plot, we have marked a convergence of drive accuracy of 0.96% and a test of 0.95%, it is clear that the proposed model can be generalized, since the plot has a slight difference between drive accuracy and tests, This is a good indication of the effectiveness of the proposed model. The results show that the (accuracy) accuracy reaches 96% and the validation loss (val_{loss}) is reduced to 0.13, which is the culmination. This can be considered a good sign for good classification results, especially in the area of medical diagnosis

IV. CONCLUSION

COVID-19 has caused a serious negative impact on our daily lives ranging from public health systems to the global economy. Individuals infected with COVID-19 are more likely to suffer damage to the lungs which may later lead to death. This study aimed to identify individuals with COVID-19, by employing non-clinical approaches, specifically, artificial intelligence techniques have been adopted for this purpose through X-ray images. The proposed model is capable of providing accurate diagnostics for 2 class classification (i.e. COVID and Normal). In future, we aim to evaluate the proposed deep learning technique by taking into consideration more datasets as well as tuning some existing deep neural network configurations like MobileNet.

REFERENCES

[1] C. Rothe, M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, C. Wallrauch, T. Zimmer, V. Thiel, C. Janke, W. Guggemos, M. Seilmaier, C. Drosten, P. Vollmar, K. Zwirglmaier, S. Zange, R. Wolfel, € M. Hoelscher, Transmission of 2019-nCoV infection from an asymptomatic contact in Germany, N. Engl. J. Med. 382 (2020) 970–971

[2] M.S. Razai, K. Doerholt, S. Ladhani, P. Oakeshott, Coronavirus Disease 2019 (Covid-19): a Guide for UK GPs, vol. 800, 2020, pp. 1–5

[3] X. Peng, X. Xu, Y. Li, L. Cheng, X. Zhou, B. Ren, Transmission routes of 2019-nCoV and controls in dental practice, Int. J. Oral Sci. (2020) 1–6

[4] J.M. Connors, J.H. Levy, COVID-19 and its implications for thrombosis and anticoagulation, Blood J. Am. Soc. Hematol.

[5] I.M. Baltruschat, H. Nickisch, M. Grass, T. Knopp, A. Saalbach, Comparison of deep learning approaches for multi-label chest X-ray classification, Sci. Rep. 9 (2019) 6381

[6] L.T. Duong, N.H. Le, T.B. Tran, V.M. Ngo, P.T. Nguyen, Detection of tuberculosis from chest X-ray images: Boosting the performance with vision transformer and transfer learning, Expert Syst. Appl. 184 (2021) 115519,

[7] CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images

[8] A. Pagano, et al., Lung ultrasound for diagnosis of pneumonia in emergency department,, Intern Emerg Med 10 (7) (2015) 851–854.

[9] Y. Amatya, et al., Diagnostic use of lung ultrasound compared to chest radiograph for suspected pneumonia in a resource limited setting, Int J Emerg Med 11 (1) (2018) 8.

[10] M. Shorfuzzaman, M.S. Hossain, MetaCOVID: A Siamese neural network framework with contrastive loss for n-shot diagnosis of COVID-19 patients, Pattern Recognition, Pattern Recognition 113 (2021) 107700. , pages.

[11] K. Lin, Y. Li, J. Sun, D. Zhou, Q. Zhang, Multi-sensor fusion for body sensor network in medical human–robot interaction scenario, Inf. Fusion 57 (May 2020) 15–26.

[12] S.U. Amin, et al., Deep Learning for EEG motor imagery classification based on multi-layer CNNs feature fusion, Future Generation Computer Systems 101 (2019) 542–554.

An Efficient Adaptive Filtering Algorithm with Double Talk Detection in Car cabin

Islam Hassani
Laboratoire de Développement Durable et Informatique (LDDI), University of Adrar
Adrar, Algeria
islamhassani65@yahoo.com

Rédha Bendoumia
Detection, Information and Communication Laboratory (DIC),
University of Blida1, Algeria
bendoumia_redha@univ-blida.dz

Elhachemi Kouddad
Telecommunication and Digital Signal Processing Laboratory, Department of Telecommunication,
University of Sidi Bel Abbes
hachemi.kouddad@dl.univ-sba.dz

***Abstract*—This paper addresses the problem of Acoustic Echo Cancelation (AEC) by adaptive filtering algorithms. In this paper, we propose to integrate a Double Talk Detection (DTD) based on Normalized Cross Correlation (NCC) into Fast Normalized Least Mean Square (FNLMS) algorithm. The proposed algorithm shows an improvement of the convergence speed. Simulation results have confirmed the good behavior of the proposed structure to cancel acoustic echo in case of a real impulse response which is car cabin.**

Keywords—Acoustic Echo Cancellation, adaptive filtering, Fast-NLMS, Double Talk Detection.

I. INTRODUCTION

Acoustic echo cancellers are used to suppress acoustic echoes [1] in modern communication systems. An acoustic echo canceller (AEC) is generally implemented by an adaptive finite impulse response (FIR) filter as illustrated in Fig. 1. The presence of the near-end speech makes the adaptation of the echo cancellation filter problematic. A strong near-end speech acts as a large disturbance to the adaptive filtering algorithm and may cause the echo cancellation filter to diverge [2]. It is for this reason that the update of the echo cancellation filter should be turned off when near-end speech is detected. An algorithm that detects the presence of near-end speech is called a double talk detector (DTD).

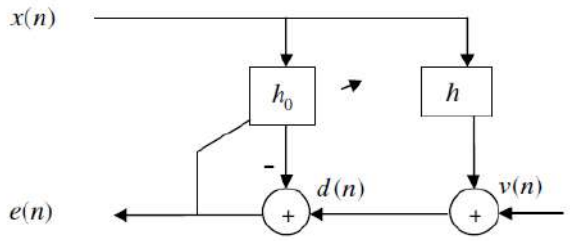


Fig.1. Basic AEC model

Several methods of double-talk detection have been proposed in the literature. Such as, a level energy based DTD: the well-known Geigel algorithm, presented in [3],uses an amplitude comparison between near-end and far-end speech signals, and a method based on the signal envelope is presented in [4], Holder inequality based method is presented in [5]. DTD using coherence is developed in [6], also, methods based on cross-correlation (CC) and normalized cross correlation (NCC) algorithms are proposed in [7]-[8].

II. ACOUSTIC ECHO CANCELLATION SYSTEM

The principle of an echo cancellation is shown in Fig.2. the echo signal received on the microphone is obtained by the convolution between the room impulse response h(n) and the input signal x(n), then the echo signal is:

$$y(n) = \mathbf{h}^{\mathrm{T}}\mathbf{x}(n) \quad (1)$$

Where $\mathbf{h} = [h_0\ h_1\ , \dots\ , h_{L-1}]$

L is the length of the echo path, the superscript denotes transpose of a vector

$\mathbf{x}(n) = [x(n)\ x(n-1), \dots \dots, x(n-L+1)]^T$ is the vector contains the last L samples of the far-end speech signal $x(n)$. The desired signal is:

$$d(n) = y(n) + s(n) + v(n) \quad (2)$$

where $s(n)$ is the near-end speech signal and $v(n)$ is the background noise.

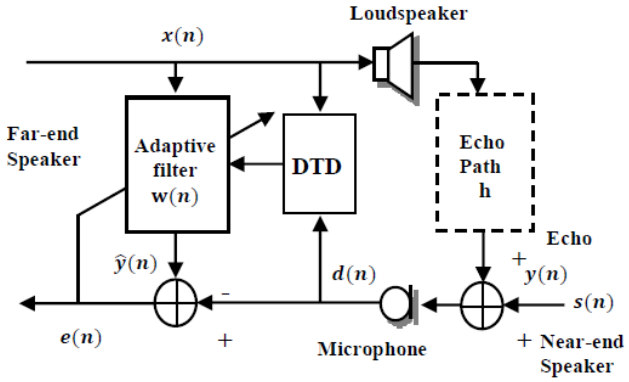


Fig.2. Principle of echo cancellation

$\mathbf{w}(n)$ is a finite impulse response (FIR) filter. The FIR filter coefficient $\mathbf{w}(n)$ should be identical with the impulse response $\mathbf{h}(n)$.

A. *Adaptive filtering*

The estimated echo $\hat{y}(n)$ is created by the convolution of the coefficient vector of adaptive filter $\mathbf{w}(n) = [w_0(n)\ w_1(n)\ , \dots \dots , w_{L-1}(n)]$, with the received input signal $\mathbf{x}(n)$. The estimated echo is:

$$\hat{y}(n) = \mathbf{w}^T(n-1)\mathbf{x}(n) \tag{3}$$

The adaptive algorithms enable filter $h(n)$ to be estimated by $\mathbf{w}(n)$ using a priori estimation error, is written, for each n:

$$e(n) = d(n) - \hat{y}(n) \tag{4}$$

The filter is updated at each instant by feedback of the estimation error proportional to the adaptation gain, denoted as $\mathbf{g}(n)$, and according to:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mathbf{g}(n)e(n) \tag{5}$$

The different adaptive algorithms are differentiated by the adaptation gain calculation.

The LMS algorithms derived from the gradient [9], based on the minimization of the mean-square error. For the Normalized LMS (NLMS) algorithm, the adaptation gain is given by:

$$\mathbf{g}(n) = \frac{\mu}{L\pi_x(n)+c_0}\mathbf{x}(n) \tag{6}$$

where μ is the adaptation step and c_0 is a small positive constant used to avoid division by zero in absence of the input signal and $\pi_x(n)$ is the power of input signal.

The RLS algorithm, for which minimizes a deterministic sum of squared errors [9].

Fast versions of these algorithms are derived from the RLS by using forward and backward linear prediction analysis over the signal $\mathbf{x}(n)$, the adaptation gain is given by :

$$\mathbf{g}(n) = \gamma(n)\tilde{\mathbf{c}}(n) \tag{7}$$

The variables $\gamma(n)$ and $\tilde{\mathbf{c}}(n)$ indicate respectively the likelihood variable and normalized Kalman vector.

Recently, a new adaptive algorithm with fast convergence and low complexity is proposed [10]. This algorithm FNLMS derived from the FRLS algorithm where the adaptation gain is obtained by discarding completely the forward and backward predictors. Thus, in this algorithm, we proposed a simplified adaptation gain:

$$\begin{bmatrix}\tilde{\mathbf{c}}(n)\\ c(n)\end{bmatrix} = \begin{bmatrix}-\frac{\varepsilon(n)}{\lambda\alpha(n-1)+c_0}\\ \tilde{\mathbf{c}}(n-1)\end{bmatrix} \tag{8}$$

where $\varepsilon(n)$ is the prediction error and $\alpha(n)$ is the forward prediction error variance.

$$\varepsilon(n) = x(n) - a(n)x(n-1) \tag{9}$$

where $a(n)$ is the prediction parameter and it is estimated by the following equation :

$$a(n) = \frac{r_{1x}(n)}{r_{0x}(n)+c_a} \tag{10}$$

where $r_{1x}(n)$ is an estimate of the first lag correlation function of $x(n)$, $r_{0x}(n)$ an estimate of the input signal power and c_a is a small positive constant used to avoid division by zero. The forward prediction error variance is now evaluated by:

$$\alpha(n) = \lambda\alpha(n-1) + \varepsilon^2(n) \tag{11}$$

B. *Double talk-detection*

In Acoustic Echo Cancellation, the most difficult problem is to handle with the situation of Double-talk presence. Double-talk occurs when far-end and near-end talk at the same time, as a result, the far-end speech signal is corrupted by near-end signal. To solve this problem, one introduces the Double-talk Detector. The task of DTD is freezes the adaptation step during filtering algorithm in case of near-end speech present to avoid the divergence of adaptive algorithm. Without DTD, when the near-end talking would make the system estimation process fail and produce extremely erroneous results. Typically, the DTD calculates a variable decision $\xi(n)$, and double-talk is declared if is lower than a given threshold value T [11]. The optimum variable decision $\xi(n)$ for double-talk detection will behave as follows:

- If $s(n) = 0$ (double talk is not present), $\xi(n) \geq T$.
- If $s(n) \neq 0$ (double talk is present), $\xi(n) < T$.

The control of the adaptive filter by DTD is defined as:

$$decision = \begin{cases}\xi(n) \geq T\ , DTD = 0\ , update\ filter\ coeff\\ \xi(n) < T\ , DTD = 1\ , freeze\ adaptation\end{cases}$$

Geigel is one of the most popular algorithms of DTD, which compares the magnitude of the microphone signal with the recent history of the far-end signal, where the variable decision of this latter is defined as:

$$\xi_G(n) = \frac{\max\{|x(n)|,\dots,|x(n-L+1)|\}}{|d(n)|} \tag{12}$$

A method of DTD based on a Cross-Correlation (CC) between the far-end and error signals is proposed in [12]. Moreover, approximate versions, such as a Normalized Cross-Correlation method (NCC) are developed in [13] [14]. This method calculates a variable decision $\xi_{NCC}(n)$ which is defined by:

$$\xi_{NCC}(\mathrm{n}) = 1 - \frac{\hat{r}_{ed}}{\hat{\sigma}_d^2} \tag{13}$$

where $r_{ed} = E\{e(n)d(n)\}$ is the cross-correlation between $e(n)$ and $d(n)$. $E\{.\}$ denotes the mathematical expectation and σ_d^2 is the variance of $d(n)$.

This variable decision is based on the estimates $\hat{r}_{ed}$ and $\hat{\sigma}_d^2$ where are found using the recursive form [15]:

$$\hat{r}_{ed}(n) = \lambda\hat{r}_{ed}(n-1) + (1-\lambda)e(n)d(n) \tag{14}$$

$$\hat{\sigma}_d^2(n) = \lambda\hat{\sigma}_d^2(n-1) + (1-\lambda)d(n)d(n) \tag{15}$$

where

λ is the exponential weighting factor $\lambda < 1\ and\ \lambda \approx 1)$

III. THE STUDIED DTD FOR AEC

In this section, we propose to use the Fast-NLMS type adaptive algorithm with DTD, the method used in this article is known as Normalized Cross-correlation technique. It is based on the

calculation of the estimates using the exponential recursive weighting algorithm to obtain the values of the cross-correlation (r_{em}) between error signal and microphone signal and the variance (σ_m^2) of the microphone signal, then to achieve the decision statistic (ξ_{DTD}) from these values. Finally, we compare this value to the threshold (T) to make the decision of Double-talk Detector in case of double talk.

Because of the convergence time of the adaptive filter, we must setup the first time (DTDbegin) when the Double-talk Detector start working.

The proposed algorithm is demonstrated in the table below:

TABLE I:
Fast-NLMS with DTD based on NCC

Initialisation:

L=length(h);

$\mathbf{w}(0) = \tilde{\mathbf{c}}(0) = 0, \gamma(0) = 1, r_1(0) = 0,$

$\alpha(0) = r_0(0) = E_0$ where E_0 is an initialization constant.

T=0.92; Set the threshold

$\lambda_{DTD} = 0.95$; to calculate decision statistic

DTDbegin=20000; time to activate DTD

$d(n) = y(n) + s(n) + v(n)$;

For each instant of time n=1,2…

Prediction error:

$$r_1(n) = \lambda_a r_1(n-1) + x(n)x(n-1)$$
$$r_0(n) = \lambda_a r_0(n-1) + x^2(n)$$
$$a(n) = \frac{r_1(n)}{r_0(n) + c_a}$$
$$\varepsilon(n) = x(n) - a(n)x(n-1)$$
$$\alpha(n) = \lambda\alpha(n-1) + \varepsilon^2(n)$$

Adaptation gain :

$$\begin{bmatrix}\tilde{\mathbf{c}}(n)\\ c(n)\end{bmatrix} = \begin{bmatrix}-\frac{\varepsilon(n)}{\lambda\alpha(n-1) + c_0}\\ \tilde{\mathbf{c}}(n-1)\end{bmatrix}$$
$$\delta(n) = c(n)x(n-L) + \frac{x(n)\varepsilon(n)}{\lambda\alpha(n-1) + c_0}$$
$$\gamma(n) = \frac{\gamma(n-1)}{1 + \gamma(n-1)\delta(n)}$$

Filtering Part:

$$e(n) = d(n) - w^T(n-1)x(n)$$

---------- NCC for DTD----------

threshold(n)=T;

if *DTDbegin*>=n then start updating filter coefficients

else compute

$$\hat{r}_{ed}(n) = \lambda\hat{r}_{ed}(n-1) + (1-\lambda)e(n)d(n)$$
$$\hat{\sigma}_d^2(n) = \lambda\hat{\sigma}_d^2(n-1) + (1-\lambda)d(n)d(n)$$

$$\xi_{NCC} = 1 - \frac{\hat{r}_{ed}}{\hat{\sigma}_d^2}$$

if ξ_{NCC}> threshold

$\mu = 1$, update filter coefficients.

end

$$\mathbf{w}(n) = \mathbf{w}(n-1) - \mu e(n)\gamma(n)\tilde{\mathbf{c}}(n)$$

From Table 1, we can see that the proposed algorithm updates the adaptive filter when double talk is not detected (i.e ξ_{NCC}> threshold) and stop the adaptation when the double talk is detected to prevent the adaptive filter from divergence.

Mean Square Error (MSE) :

The purpose of the adaptive filter is minimizing the Mean Square Error MSE:

$$MSE(dB) = 10log_{10}(E\{|e(n)|^2\})$$

Therefore, the values and graph of this quantity will be essential to evaluate the performance of the adaptive filter. If the adaptive algorithm works well, after convergence time, the value of MSE should be reduced gradually to zero (for the case of no near-end signal).

IV. SIMULATION RESULTS

We used in our simulation two different types of input signals. The first one is a stationary correlated noise, with a spectrum equivalent to the average spectrum of speech. It is usually called USASI (USA Standards Institute, now ANSI) noise in the field of acoustic echo cancellation. Since in real situation the input signal is nonstationary, we also ran the algorithm on a second signal, which is a typical speech signal. These two signals, sampled at 16 kHz, are filtered by two impulse responses to obtain the desired signals. The first impulse response, represents a low-pass filter of size 32, and the second represents a real impulse response measured in a car and truncated to N = 256 samples.

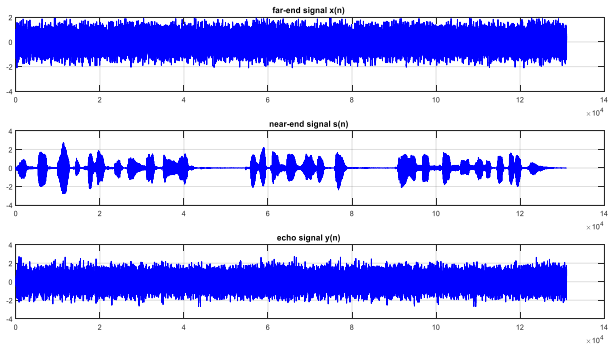


Fig.3.the used signals x(n):USASI noise , s(n): speech signal used for Double-Talk and y(n):measured echo with USASI noise in car.

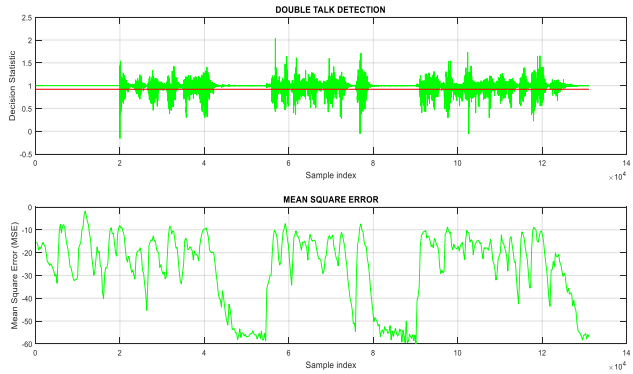


Fig.4. MSE with the data of **fig.3** ., $\lambda_{DTD} = 0.95$, DTDbegin=20000.

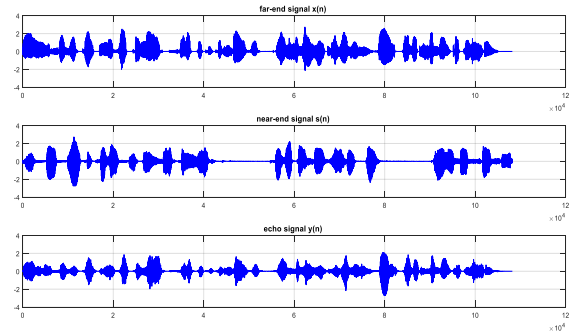


Fig.5.the used signals x(n):speech input signal , s(n): speech signal used for Double-Talk and y(n):measured echo with speech signal in audioconference room ACN.

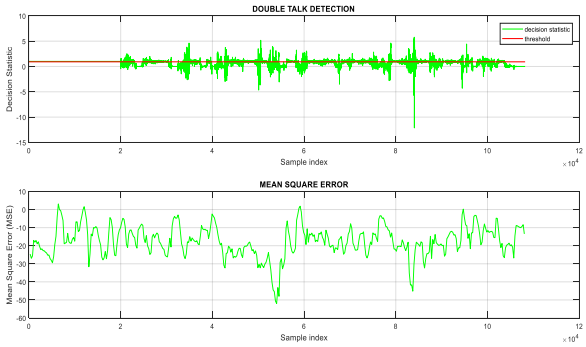


Fig.6. MSE with the data of **fig.5** . $\lambda_{DTD} = 0.95$, DTDbegin=20000.

From figures 4 and 6 we can well see that the proposed algorithm converges rapidly in the silent periods, which means that the proposed structure has rapidly detected the double talk, and this is noticed in the convergence rate of the MSE criterion. In addition, these experiments are done in a real car cabin and using two different input signals noise and real speech.

V. CONCLUSION

In this paper, we have presented a DTD based on NCC, where the main purpose of this DTD is to halt the update of filter coefficients during the double-talk periods. DT periods are determined from decision statistic depending on the estimated variance of near-end signal and cross-correlation between error signal and microphone signal. The experimental results demonstrate a good performance with FNLMS-type algorithm with different input signals under a real acoustic impulse response.

REFERENCES

[1] Gänsler, Tomas, Jacob Benesty, and Steven L. Gay. "Double-talk detection schemes for acoustic echo cancellation." Acoustic signal processing for telecommunication. Springer, Boston, MA, 2000. 81-97.

[2] Wang, Qingyun, et al. "A frequency-domain nonlinear echo processing algorithm for high quality hands-free voice communication devices." Multimedia Tools and Applications 80 (2021): 10777-10796.

[3] Duttweiler, D. "A twelve-channel digital echo canceler." IEEE Transactions on Communications 26.5 (1978): 647-653.

[4] Szwoch, Grzegorz, Andrzej Czyżewski, and Maciej Kulesza. "A low complexity double-talk detector based on the signal envelope." Signal Processing 88.11 (2008): 2856-2862.

[5] Paleologu, Constantin, et al. "Class of double-talk detectors based on the holder inequality." Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on. IEEE, 2011.

[6] Gansler, Tomas, et al. "A double-talk detector based on coherence." IEEE Transactions on Communications 44.11 (1996): 1421-1427.

[7] Benesty, Jacob, Dennis R. Morgan, and Jun H. Cho. "A new class of doubletalk detectors based on cross-correlation." IEEE Transactions on Speech and Audio Processing 8.2 (2000): 168-172.

[8] Gänsler, Tomas, and Jacob Benesty. "The fast normalized cross-correlation double-talk detector." Signal Processing 86.6 (2006): 1124-1139.

[9] Haykin, Simon O. Adaptive filter theory. Pearson Higher Ed, 2013.

Gay, Steven L., and Jacob Benesty, eds. Acoustic signal processing for telecommunication. Vol. 551. Springer Science & Business Media, 2012.

[10] A. Benallal, M. Arezki, "A fast convergence normalized least-mean-square type Algorithm for Adaptive Filtering", *International Journal Adaptive Control and Signal Processing*, 2013 John Wiley & Sons, DOI: 10.1002/acs.2423.

[11] Hassani, I., Benallal, A., & Bendoumia, R. (2022). Double-Talk Robust Fast Converging and Low Complexity Algorithm for Acoustic Echo Cancellation in Teleconferencing System. In Artificial Intelligence and Heuristics for Smart Energy Efficiency in Smart Cities: Case Study: Tipasa, Algeria (pp. 409-420). Springer International Publishing.

[12] J. Benesty, D. R. Morgan, J. H. Cho, "A new class of doubletalk detectors based on cross-correlation," IEEE Trans on speech and audio processing, vol. 8, no. 2, pp. 168–172, 2000.

[13] T. Gänsler and J. Benesty, "The fast normalized cross-correlation double-talk detector," Signal Processing, vol. 86, no. 6, pp. 1124–1139, 2006.

[14] J. Benesty and T. Gänsler, "A multichannel acoustic echo canceler double-talk detector based on a normalized cross-correlation matrix," European Tran on telecommunications, vol. 13, no. 2, pp. 95–101, 2002.

[15] Hassani, Islam, Madjid Arezki, and Ahmed Benallal. "The fast-NLMS algorithm for acoustic echo cancellation with double-talk detection." Proceedings of the 4th International Conference on Electrical Engineering and Control Applications: ICEECA 2019, 17–19 December 2019, Constantine, Algeria. Springer Singapore, 2021.

Contactless fingerprint and Ear based multi-biometric recognition system using HOG descriptor

Abderrahmane Herbadji
Department of Electronics
University of Setif
Setif, Algeria
herbadjiabder@gmail.com

Djamel Herbadji
Department of Electronics
University of Skikda
Skikda, Algeria

Hichem Kahia
LMSE laboratory
University of Biskra
Biskra 7000, BP145, Biskra, Algeria

***Abstract*—In the high-security scenarios, biometrics require stringent accuracy and performance criteria. Towards this goal , multi-biometric systems that combine information from multiple sources of biometric have exhibited to diminish the error rates and alleviate inherent frailties of single biometric systems. To this aim, in this paper, the authors present a multibiometric system using fingerprint and ear based on HOG descriptor, in order to extract the salient characteristics of these traits . The proposed system is evaluated on two publicly available databases, namely IIT-Delhi-2 ear and polyU contactless fingerprint database. Experimental analyzes confirmed that the proposed multi-biometric system is able to decrease the error rates compared with with ones produced by individual systems, achieving an error rate of 0.28%.**

***Index Terms*—Fingerprint, Ear, Person recognition, multibiometric, Security**

I. Introduction

Robust identity management systems have become an international requirement in various fields [1]. In fact, security and privacy has become a sensitive issue for citizens, governments, and businesses due to extensive information and data theft [2]. Identification systems based on biometrics prove to be effective to respond the requirement of security. Biometrics identifies individuals based on their biological characteristics like fingerprint, handwritten signature, face, voice, palmprint, DNA, and iris [3], [4]. Biometrics is currently used in many areas such as social networking applications, healthcare systems, entertainment technologies, cloud computing, and homeland security. In addition, individuals unlock their phones using either the fingerprint or face modalities [5].

Fingerprint is considered the most common biometric trait widely employed by national ID programs around the world, the fingerprint recognition systems could capture the fingerprint images using capacitive or optical sensors, among others [6]. Also, contact-based fingerprint sensors should deal with other concerns like deformation, sensor surface noise, and hygienic concerns. In order to address these issues, contactless fingerprint has been recently investigated for more security and better hygienic purposes [7], [8]. contactless Imaging of fingerprint helps to collect patterns without any deformation, and thereby, better matching performance is expected.

Just like a fingerprint or a face, the ear can be easily used for subject identification owing to its unique structure. In addition, the ear features are stable over time compared to the face, and are not affected by external factors like expression. Also, ear recognition does not require subject cooperation compared with iris recognition. It should be noted that main ear recognition drawback is that the ear may be occluded (partial or full) by hair or some other head-ware. However, ear recognition has an additional benefit compared to face recognition, which is that the degree of privacy concerns is lower when the ear image is collected and stored in a database rather than the face image [9].

In the present study, we propose a multimodal fingerprint-ear recognition system. Many reasons for our choice of this multimodal system. First, contactless fingerprint and ear data acquisition demands no cooperation from the user, besides, it can be captured using inexpensive sensor (e.g. conventional cameras). Second, the human ear and fingerprint patterns offer rich and stable information. Third, the proposed multibiometric recognition system could work faster as well as efficiently even with the images having low resolution. Fourth, fingerprint and ear are very suitable in identification systems due to their broad public acceptance. Fifth, we observed that suggested biometric system using HOG descriptor [10] is able to achieve high person recognition performance. We structure the rest of this research paper as follows. A general idea of the proposed contactless fingerprint and ear framework is depicted in the section II. Section III discusses the experimental protocol, sectionIV provide the experimental results, and finally a conclusion and future directions are drawn in Section V.

II. Proposed multi-biometric system

A biometric system is basically a pattern recognition/matching system. An automatic fingerprint and ear recognition systems are a essentially processing chain that is split into two stages: enrollment and recognition phases During enrollment, the fingerprint and ear biometric trait of the user are captured and processed. Once the fingerprint and ear image

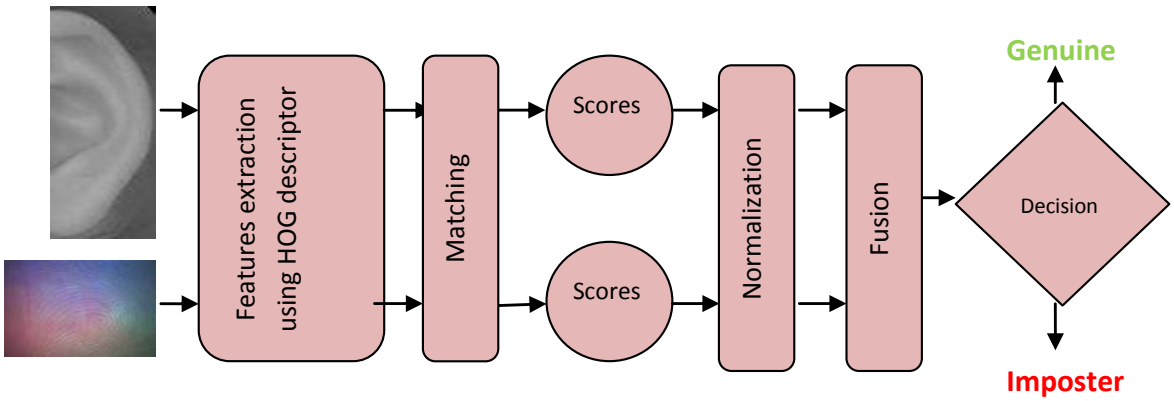


Fig. 1 *Proposed Fingerprint-ear biometric system based on HOG descriptor*

are collected, the HOG descriptor is applied to informative features. Then these features are saved in a database as a reference model along with the subject identity. During the authentication phase, the same biometric trait (here, the ear) of the subject is collected, processed, and the features are extracted using the same descriptor. This time, HOG features are compared with those stored in the database using chi-square distance to calculate the matching score. If the matching score is greater than a threshold, the subjected is rejected as imposter, otherwise as genuine.
Fig. 1 shows an architecture illustrating the overall procedure of a multi-biometric person recognition framework that integrates information from FKP and iris biometric sources.

III. Experiments

In this section, we present an experimental analysis of the proposed fingerprint and ear based multi-biometric person verification system on two publicly available databases.

A. Databases

The proposed contactless multibiometric system that fuse information from fingerprint and ear traits is evaluated on IIT-Delhi-2 ear [11] database and PolyU Contactless to Contact-based Fingerprint database [7]. The IIT Delhi ear database was collected at IIT Delhi campus, India. This database consists of ear images collected from a distance in an indoor environment. The number of subjects of this database is 221 subjects (i.e. students and staff of university), with at least three ear images for each subject. The subjects are in the age range of 14 to 58 years. The resolution of the images is 272 × 204 pixels. Besides, cropped and normalised ear images are available with resolution of 50 ×180 pixels as shown in Fig. 2.

In the PolyU Contactless to Contact-based Fingerprint database [7], fingerprint images (i.e., contact based and contactless) acquired from 336 volunteers are available (see fig. 3). This database provides six fingerprint images for each user besides to their corresponding contact-based images.

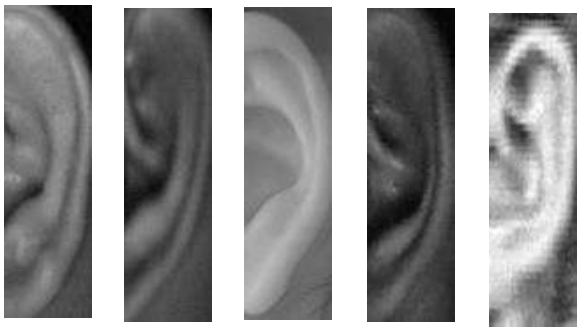


Fig. 2 *Some normalised ear images from IIT Delhi-2 database*

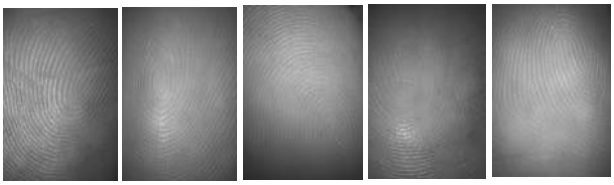


Fig. 3 *Some normalised fingerprint images from polyU contactless fingerprint database*

B. Experimental protocol

We selected 221 users from each database, we choose two samples to generate the registered model and one samples as a test set. Here, each image of the test set is associated with

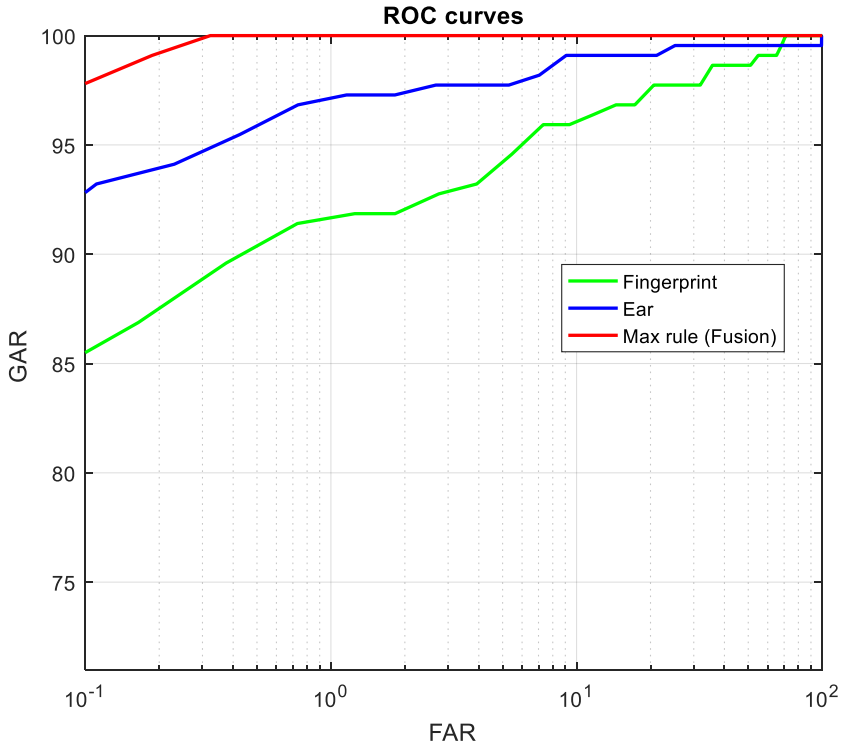


Fig. 4 *ROCs of uni-biometrics fingerprint and ear along with and their fusion using both feature and score fusion*

the two images of the registered model for each person. We therefore have 221 authentic scores and 48620 (221 × 220) impostor scores for each of the modalities. The performances of proposed fusion scheme are analysed utilising receiver operating characteristic (ROC) curve. An ROC curve is a plot of FAR versus GAR. GAR, FAR and FRR are GAR such as GAR = 1–FRR (i.e. ratio of legitimate users accepted as genuine), FAR (i.e. ratio of impostors accepted as legitimate) and false rejection rate (FRR) (i.e. ratio of legitimate users rejected as impostors), respectively. Equal error rate (EER), where FRR and FAR are equal has been also used to evaluate performance.

IV. Results

In the proposed multimodal fingerprint-ear biometric system, features have been extracted from two modalities using HOG descriptor. We combine these two traits using numerous fusion rules such as weighted sum, weighted product, max rule, min rule, product, and sum rule in order to show how can these rules could be usfeaul for the proposed multibiometric system. Three evaluation metrics [12], wich are of equal error rate (EER), decidability index (d'), and receiver operator characteristic curve (ROC) are used to evaluate verification accuracy in our experiments.

A. Performance of multimodal fingerprint and ear system using EER

We present the EERs of the fingerprint and ear based unimodal systems and their multibiometric system in table I. From this Table, we can observe that the proposed multibiometric fusion based on max rule achieves lower EER than the corresponding fingerprint and ear unimodal systems. For example, using only the fingerprint and ear traits, the EER reaches 5.44% and 2.40%, while fusing these modalities via max rule results in 0.28% EER.

TABLE I *Comparison of ERRs obtained for the uni-biometrics fingerprint and ear systems and their fusion via different methods*

Biometric System	EER
Fingerprint	5.44
Ear	2.40
Product	1.53
Sum rule	1.55
Max rule	**0.28**
Min rule	3.26
Weighted sum	0.67
Weighted product	1.08

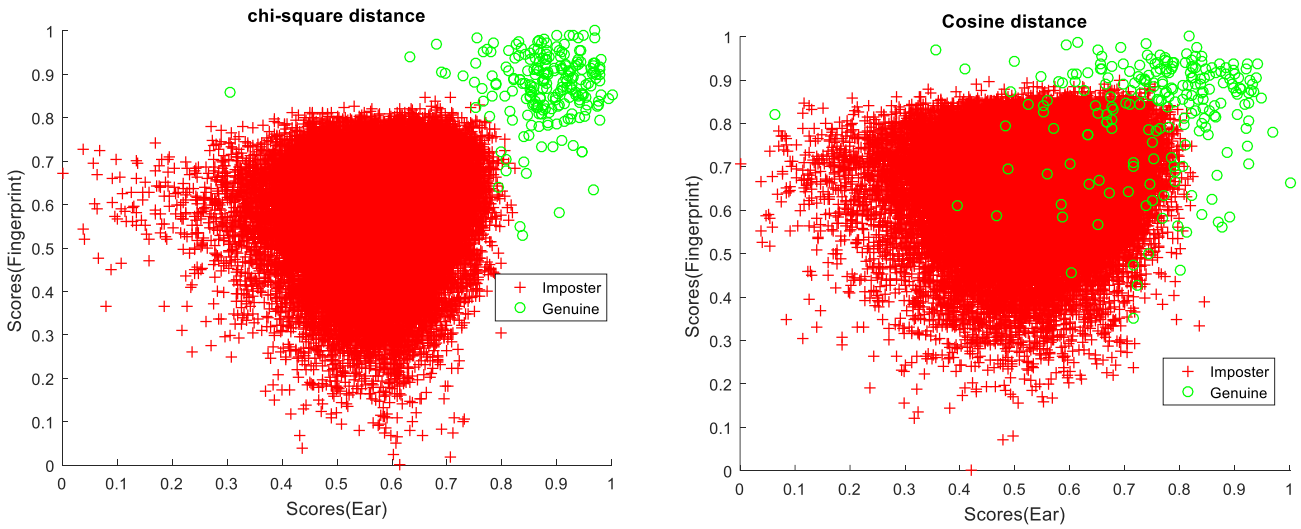


Fig. 5 *Distributions of normalized biometric scores from two databases, where x-axis is ear scores and y-axis is fingerprint scores, which were calculated using two distances: (a) chi-square distance; (b) cosine distance.*

B. Performance of multimodal fingerprint and ear system using d'

Table II gives the d' values of the fused modalities on virtual multimodal database using different fusion methods such as sum, min, and max rules. From this table, it is evident that product and max rules perform better than other fusion techniques as d' values corresponding to former are higher than the latter.

TABLE II *Comparison of decidability index obtained for the uni-biometrics fingerprint and ear systems and their fusion via different methods*

Biometric System	d'
Fingerprint	3.10
Ear	3.87
Product	4.88
Sum rule	4.74
Max rule	**4.83**
Min rule	3.63
Weighted sum	4.81
Weighted product	4.63

C. Performance Evaluation Using Verification Rate (VR)

Fig. 4 illustrates ROC's uni-biometrics systems and and corrosponding multi-biometric recognition system by invoking the score level fusion using max rule on PolyU contactless fingerprint and IID-Delhi-2 ear databases. At FAR of 0.1%, the GARs of fingerprint and ear are 85.80% and 93% respectively. But with max rule, GAR of 100% is obtained with the same FAR. Other fusion rules such as weighted sum, product, sum rule are also used for the score level fusion and the results are almost identical to those of max rule.

Overall, the max rule fusion with the chi-square distance gives the best results in comparison with other matching distances such cosine distance, where a GAR = 97.90% at FAR = 0.1% of the operating point is achieved. The possible reason could be that the genuine scores and the impostor scores of the fingerprint and ear calculated using the chi-square distance are well separated as shown figure 5.

TABLE III *Comparison of VRs obtained for the uni-biometrics fingerprint and ear systems and their fusion via different methods*

Biometric System	VR@0.1%FAR	VR@1%FAR
Fingerprint	85.80	91.90
Ear	93.00	97.25
Product	96.75	98.00
Sum rule	97.30	98.20
Max rule	97.90	**100**
Min rule	90.50	94.70
Weighted sum	97.90	99.90
Weighted product	97.75	98.95

V. Conclusion

In this work, we have presented an approach for multi-modal contactless fingerprint and ear biometric recognition using HOG texture descriptor at score-level fusion. Extensive experimental analysis on a multimodal database using different fusion rules such as weighted sum showed a significant improvement in verification rates in comparison with unimodal systems. In future work, we aim to compare the study of the score-level with feature and decision level fusion methods.

References

[1] B. Bhanu and V. Govindaraju, *Multibiometrics for human identification*. Cambridge University Press, 2011.

[2] A. Jain, R. Bolle, and S. Pankanti, "Introduction to biometrics," in *Biometrics*. Springer, 1996, pp. 1–41.

[3] P. J. Phillips, P. J. Flynn, J. R. Beveridge, W. T. Scruggs, A. J. O'toole, D. Bolme, K. W. Bowyer, B. A. Draper, G. H. Givens, Y. M. Lui *et al.*, "Overview of the multiple biometrics grand challenge," in *International conference on biometrics*. Springer, 2009, pp. 705–714.

[4] W. Kabir, M. O. Ahmad, and M. Swamy, "Score reliability based weighting technique for score-level fusion in multi-biometric systems," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2016, pp. 1–7.

[5] Z. Akhtar, A. Buriro, B. Crispo, and T. H. Falk, "Multimodal smartphone user authentication using touchstroke, phone-movement and face patterns," in *2017 IEEE global conference on signal and information processing (GlobalSIP)*. IEEE, 2017, pp. 1368–1372.

[6] M. Bellaaj, R. Boukhris, A. Damak, and D. Sellami, "Possibilistic modeling palmprint and fingerprint based multimodal biometric recognition system," in *2016 International Image Processing, Applications and Systems (IPAS)*. IEEE, 2016, pp. 1–8.

[7] C. Lin and A. Kumar, "Matching contactless and contact-based conventional fingerprint images for biometrics identification," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 2008–2021, 2018.

[8] A. Herbadji, N. Guermat, L. Ziet, Z. Akhtar, M. Cheniti, and D. Herbadji, "Contactless multi-biometric system using fingerprint and palmprint selfies." *Traitement du Signal*, vol. 37, no. 6, 2020.

[9] N. Hezil and A. Boukrouche, "Multimodal biometric recognition using human ear and palmprint," *IET Biometrics*, vol. 6, no. 5, pp. 351–359, 2017.

[10] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, vol. 1. Ieee, 2005, pp. 886–893.

[11] A. Kumar and C. Wu, "Automated human identification using ear imaging," *Pattern Recognition*, vol. 45, no. 3, pp. 956–968, 2012.

[12] M. Hanmandlu, J. Grover, A. Gureja, and H. M. Gupta, "Score level fusion of multimodal biometrics using triangular norms," *Pattern recognition letters*, vol. 32, no. 14, pp. 1843–1850, 2011.

AIoT-based Quality Control for Gear Production: A Complete Autonomous Environment

Mohamed Yanis Hiou
Dept. of Electrical Systems Engineering
Faculty of Technology, University of Boumerdes
Boumerdes, Algeria
m.hiou@univ-boumerdes.dz

Hamza Akroum
Systems and Telecommunications Engineering Laboratory
Faculty of Technology, University of Boumerdes
Boumerdes, Algeria
akroum@univ-boumerdes.dz

***Abstract*—The fourth industrial revolution is characterized by the convergence of advanced technologies such as Artificial Intelligence (AI) and the Internet of Things (IoT), with edge computing playing a crucial role in enabling efficient data processing and communication in close proximity to data sources. Despite its benefits, edge computing presents challenges such as security risks, data-intensive services, incomplete data, and high investment and maintenance costs, which can be mitigated using cloud computing technology.**

This paper proposes an AIoT-based quality control production line that utilizes a combination of AI and IoT technologies to enhance production management. The production line focuses on classifying gears and detecting defects, with data being collected and processed on a server for analytics to be displayed on Android, web, and desktop applications. By leveraging AIoT, the proposed solution aims to create a complete autonomous environment and improve overall efficiency in gear production.

***Index Terms*—Cloud Computing, Artificial Intelligence, Internet of Things, Artificial Intelligence of Things, Industry 4.0**

I. Introduction

Nowadays, industrial developments in various countries around the world are all pursuing Industry 4.0. The primary goals are to digitalize machining data in the manufacturing process, and import intelligence techniques [1]. One of the techniques is Artificial Intelligence of Things technology. Artificial Intelligence of Things (AIoT) is the natural evolution for both Artificial Intelligence (AI) and Internet of Things (IoT) because they are mutually beneficial [2]. Existing research indicated that advanced manufacturing technologies such as intelligent robotic systems, smart manufacturing systems, industrial automation, and intelligent monitoring systems have shown a great dependency on AI and IoT-based technology [3]. IoT is a sort of "universal global neural network" in the cloud which connects various devices. The IoT is an intelligently connected devices and systems which be made up of smart machines interacting and communicating with other machines [4]. And When we combine the industrial IoT with cloud computing, we'll have the CloudIoT Technology [5].

In our view, industrialized countries such as China, the United States, and Germany are investing heavily to ensure the adoption of Industry 4.0 in their industries in order to keep up with the global digital transformation wave. Therefore, Algerian manufacturers must also embrace this digital transformation and adopt Industry 4.0 initiatives in order to remain competitive. This challenge requires a commitment to investing in the necessary expertise, organizational and professional skills, and the planning, execution, and improvement of transition projects towards Industry 4.0. The main contributions of this work include training the model with different architectures, implementing the model with the highest accuracy on the server, building the production line and connecting it to the cloud, and displaying the data on different platforms. The main contributions of this work can be summarized below:

1) Training the model with different architectures;
2) Implementing the model with the highest accuracy on the server;
3) Building the production line and connect it to the cloud;
4) Display the data on different platforms.

II. Materials And Methods

A. Convolutional Neural Network

The term Deep Learning or Deep Neural Network refers to Artificial Neural Networks (ANN) with multi layers. Over the last few decades, it has been considered to be one of the most powerful tools, and has become very popular in the literature as it is able to handle a huge amount of data. The interest in having deeper hidden layers has recently begun to surpass classical methods performance in different fields; especially in pattern recognition. One of the most popular deep neural networks is the Convolutional Neural Network (CNN). It take this name from mathematical linear operation between matrices called convolution. CNN have multiple layers; including convolutional layer, non-linearity layer, pooling layer and fully-connected layer. The convolutional and fully-connected layers have parameters but pooling and non-linearity layers don't have parameters. The CNN has an excellent performance in machine learning problems. Specially the applications that deal with image data, such as largest image classification data set (Image Net), computer vision, and in natural language processing (NLP) and the results achieved were very amazing. In this paper we will explain and define all the elements and important issues related to CNN, and how these elements work. In addition, we will also state the parameters that effect CNN efficiency. This paper assumes that the readers have adequate knowledge about both machine learning and artificial neural network [6].

In this paper, we propose several Convolutional Neural Network (CNN) architectures for training our model, each with unique properties aimed at improving performance and decreasing processing time. While these architectures differ in their design, they all share the same objective. Some architectures have shown to be particularly effective across a wide range of applications, such as VGG16, ResNet, AlexNet, and Xception, which are among the most commonly used architectures for industrial image pattern recognition.

For the image classification of the gears, we utilized three distinct models: VGG16, ResNet, and AlexNet.

1) VGG Net: VGG Net is a pre-trained Convolutional Neural Network architecture invented by Si-monyan and Zisserman from the Visual Geometry Group (VGG) [7]. It was trained to extract features that can identify objects and classify unseen items to improve classification accuracy by increasing the depth of CNNs. VGG 16 architecture, having 16 hidden layers, was used for image classification. It accepts input images of dimensions (224,224,3) and passes them through a series of convolutional layers with a 3x3 filter. There exist five max-pooling layers between convolutional layers to subsample the input images. Convolutional layers are followed by three fully connected layers with the softmax being the last layer.
2) ResNet: The ResNet [8], [9] is a modern deep neural network architecture. The core idea is to use an identity shortcut connection to skip some layers to deal with vanishing gradients' problem. It allows us to stack convolutional layers and obtain better recognition quality than shallower models. So, ResNet is a scalable and accurate model that founds wide application in practice.
3) AlexNet: AlexNet architecture was developed by Alex krishevesky et al [10], which won the ILVRC (ImageNet Large Scale Visual Recognition Challenge) in 2012. AlexNet architecture consists of five convolutional (conv) layer, three pooling layer (Pool) which is followed by three full connected (FC) layer. To reduce overfitting problem these fully connected layers are used with dropout layer. Convolution layer uses number of filters to convolve the image, and generating feature maps. Rectified linear unit (ReLU) layer is used along with convolutional layer as it performs non-linear operation and converts all negative value to zero. The task of pooling layer is to reduce the spatial dimension (feature map) which is derived from previous layer.

B. Hardware Used

To train the model, we utilized a Jetson Nano with the OpenVino Intel Movidius Vision Processing Unit (VPU) Kit, while a Raspberry Pi served as the server. In addition, an STM32 Microcontroller was employed to control the production line, and the image was sent using the ESP32-Cam.

1) Jetson Nano: NVIDIA Jetson Nano, the entry-level board of the NVIDIA Jetson ecosystem, is a small, powerful single- board computer that allows parallel operation of multiple neural networks for applications such as image classification, object detection, segmentation, and speech processing. It has a comprehensive development environment (JetPack SDK) and libraries developed for embedded applications, deep learning, IoT, computer vision, graphics, multimedia and more. Using Jetson Nano with a GeForce-enabled graphics processor (GPU) using the same CUDA cores creates a very powerful development environment for applications [11]. In addition, Jetson Nano has a CPU-GPU heterogeneous architecture in which the operating system can be booted by the CPU and can be programmed to speed up complex machine learning tasks of the CUDA capable GPU [12].
2) Vision Processing Unit: The VPU is an ultra-low power application-specific integrated circuit (ASIC) that is designed for image processing and computer vision workloads, in particular, CNNs. The ultra-low power consumption makes it ideal for mobile edge devices where real-time processing and low latency are vital [13] [14].
3) Raspberry Pi: Raspberry Pi is a powerful computer in a size of a credit card. It was created by the Raspberry Pi Foundation charity, whose primary goal was the reintroduction of computer skills learning among students. The Raspbian operating system is based on Linux and offers an excellent working environment for students, as it includes software solutions designed to expand students' knowledge of computer science in an interesting way. Besides computer science knowledge, it is also suitable for acquiring basic knowledge in electronics. It is also very important to have the possibility of interconnecting such acquired knowledge [15].
4) STM32: STMicroelectronics' STM32 microcontrollers represent a full line of 32-bit products designed to perform real-time low-voltage digital signal processing. Within this range, we find in particular the Cortex-M7F, Cortex-M4F, Cortex-M3, Cortex-M0+ and Cortex-M0. They are accompanied by a wide choice of tools and software, making this family of products the platform recognized as excellence and all the STM32 MCUs are based on ARM architecture.
5) ESP32: The ESP32 is a low-cost, low-power system on a chip series of microcontrollers with Wi-Fi and Bluetooth capabilities and a highly integrated structure powered by a dual-core Tensilica Xtensa LX6 microprocessor [16].

C. Software Used

For displaying the data, we employed various platforms, including HTML, CSS, JS, Node.JS, and MySQL, to create a web-based dashboard. To develop the mobile application, we utilized the Flutter framework, and the desktop app was created using the Processing3 programming language.

1) Flutter: Flutter is a free and open-source mobile User Interface (UI) framework created by Google and released

in May 2017. In a few words, it allows you to create a native mobile application with only one codebase [17].

2) Processing3: The Processing Development Environment (PDE) makes it easy to write Processing programs. Programs are written in the Text Editor and started by pressing the Run button. In Processing, a computer program is called a sketch. Sketches are stored in the Sketchbook, which is a folder on your computer [18].

III. Experiments And Results

A. Training's Process And Result

In this section, we will describe the process of training and evaluating three models using the architectures presented earlier, and measuring their performance using different metrics. The dataset was divided into training (80%) and validation (20%) sets, and the training data was augmented using the ImageDataGenerator class from the Keras library. Transfer learning was applied to the models, starting with loading the corresponding architecture from Keras library, followed by batch normalization, a dense layer, a dropout of 0.5 to prevent overfitting, a flatten layer, and a dense layer with 38 neurons and softmax activation function for the output layer. The models were trained with 1500 pictures for each class (defective gears and healthy gears). After trial-and-error, we determined the hyperparameters for training our models as listed in "Table. I":

TABLE I
HYPERPARAMETER FOR MODELS' TRAINING

Parameter	AlexNet	VGG16	ResNet
Input size	(224,244,3)	(224,244,3)	(224,244,3)
Optimizer	Adamax	SGD	ADAM
Loss function	Binary cross entropy	Binary cross entropy	Binary cross entropy
Classifier	Softmax	Softmax	Softmax
Epochs	15	15	15
Batch size	128	128	128
Dropout rate	0.5	0.8	0.5

Upon completion of the training process, we observed that the AlexNet architecture achieved the highest accuracy among the other architectures utilized in this study. The training process resulted in an accuracy rate of (83%), which we consider satisfactory.

The "Fig. 1", depicts the accuracy and loss of the model.

B. Hardware Development

Having achieved a satisfactory level of accuracy, we proceeded to implement the model on the Raspberry Pi, which serves as our server in this project. The process of the server is illustrated in "Fig. 2".

In addition, we designed the production line using SolidWorks and manufactured the parts using a 3D printer. The 3D design of the production line prior to printing the parts is presented in "Fig. 3".

The STM32 microcontroller controls the production line while the ESP32-CAM captures and sends the image. An ESP-01, which is based on the ESP8266 chip, is connected to all the microcontrollers and acts as a master device that collects

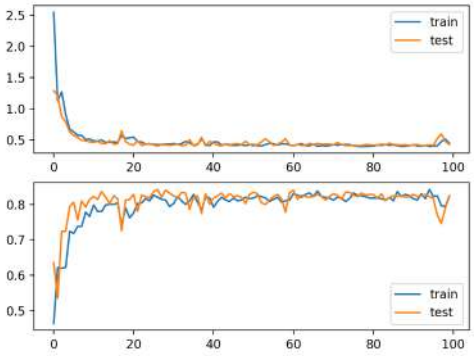


Fig. 1. Training's loss and accuracy.

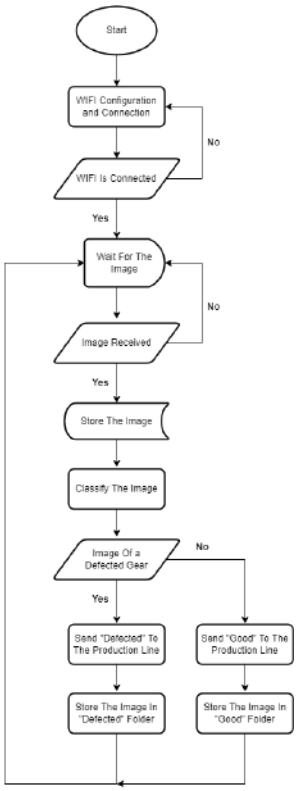


Fig. 2. Server block diagram.

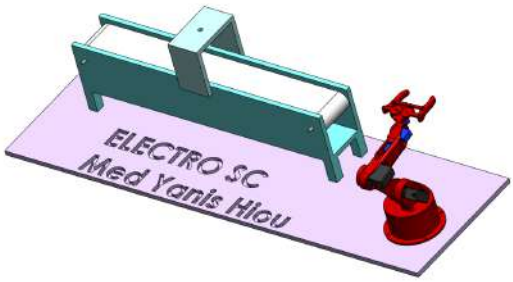


Fig. 3. 3D model of the production line.

data from all production lines and sends it to the cloud. The project flowchart is depicted in "Fig. 4".

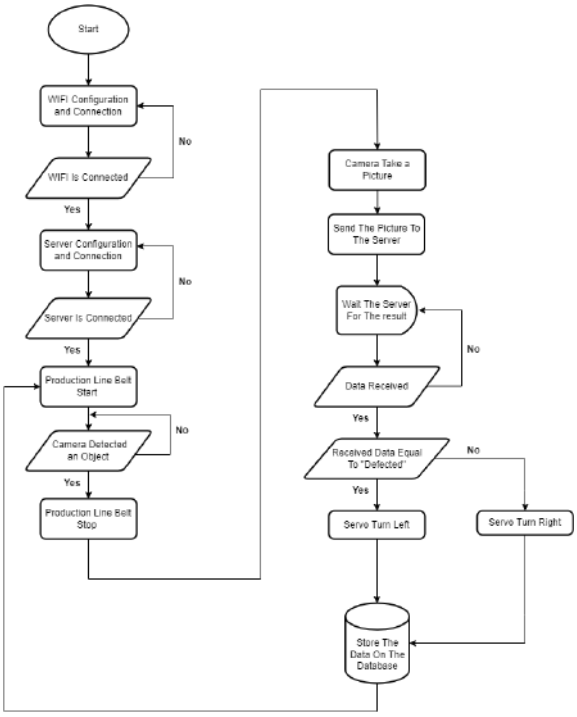


Fig. 4. General block diagram.

The server receives the image from the ESP32-CAM, applies the trained model to classify the gear as defective or not, and sends the result back. The communication between the devices in our production line is illustrated "Fig. 5".

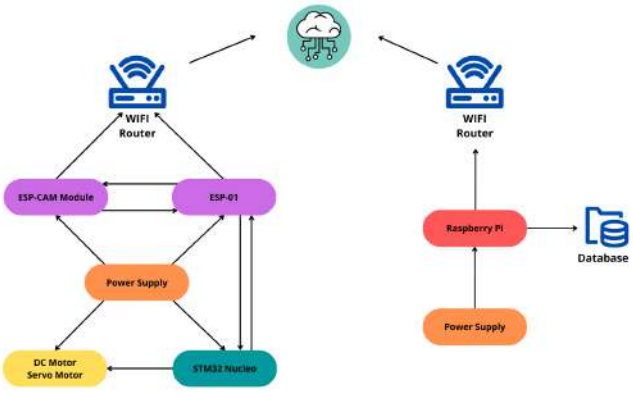


Fig. 5. Networking block diagram.

C. Software Development

As we said in "Section 2.3", we created dashboards to present the data obtained from the production line, such as power consumption, the number of healthy and defected gears, camera's frames per second (FPS), server's performance, and more. Additionally, we can manage the production line wirelessly using the dashboards. The "Fig. 6" is a screenshot of the web-based platform home page.

Fig. 6. Web-based real time dashboard.

D. Printing And Assembly

The process of the production line was simulated using SolidWorks, and the parts were manufactured using a 3D printer. The "Fig. 7" is a picture of our final production line prototype.

Fig. 7. Final prototype of the production line.

IV. Problems Faced

The project encountered issues with the Raspberry Pi overheating during data processing, causing it to crash and freeze. To address this, a cooling system was installed to reduce the server's temperature.

Another challenge was the model occasionally misclassifying objects. To improve accuracy, additional pictures were added to the data-set and the model was retrained.

Slow internet speed also posed difficulties in data processing. To mitigate this, an image compression algorithm was implemented in the ESP-CAM to reduce image size.

V. Discussion

The aim of this project was to gain knowledge in building a server and implementing AI on it, while also exploring the transmission of data between the production line and server. Our approach involved the use of artificial intelligence and cloud computing to detect and identify defective gears. The models required over 12 hours of training due to the large dataset used, which was trained on the NVIDIA Jetson Nano kit with the Intel Movidius VPU. AlexNet architecture

provided the highest accuracy at (83%), and the model was implemented on a Raspberry Pi server.

There are several areas for improvement in this project, including:

1) Upgrading the server to a more powerful computer;
2) Connecting the server and production line with a 5G network;
3) Encrypting the transmitted data for added security;
4) Implementing real-time and remote monitoring;
5) Integrating predictive maintenance to prevent potential issues.

VI. Conclusion And Future Work

This project aims to create a smart production line that utilizes AI and IoT technologies to detect and sort defected gears. The production line includes a camera that captures an image of each gear, which is then sent to a server where an AI model is applied for image classification. The server analyzes the image and determines whether the gear is healthy or defected. Based on the classification result, the production line sorts the gear into its corresponding package.

In addition to the core functionalities of the production line, the system also collects and stores data on power consumption, FPS, server performance, and the number of healthy and defected gears. This data is transmitted to a centralized database and displayed in real-time on a variety of dashboards, including web-based, android app, and desktop app interfaces.

Although the project achieved satisfactory results, there were some challenges that needed to be addressed. For instance, the server experienced overheating issues during processing, which required the addition of a cooling system. Furthermore, the AI model occasionally misclassified gears, highlighting the need for a larger and more diverse data-set.

To improve the system, some future enhancements can be considered, such as replacing the current server with a more powerful computer, implementing a 5G network for better connectivity, encrypting data transmissions for security, implementing real-time and remote monitoring, and adding predictive maintenance capabilities.

In our upcoming project, we plan to build a high-performance computer by utilizing FPGAs and creating an FPGA Cluster. Our goal is to implement some computing accelerator algorithms on this system. The FPGAs will act as clients and will be connected to a central server. This setup is expected to significantly decrease the processing time and increase productivity and quality. By leveraging the capabilities of FPGAs, we can accelerate the execution of complex algorithms and significantly reduce the time required for computation. This project will require extensive knowledge of hardware design, FPGA programming, and High Performance Computing algorithms. However, if successful, it has the potential to revolutionize the field of artificial intelligence and greatly improve the speed and efficiency of various computations.

Acknowledgment

First and Foremost, praise is to ALLAH, the Almighty, the greatest of all, on whom ultimately, we depend for sustenance and guidance. Last but not least, my warmest gratitude goes to Electro Scientific Club which always provided the most suitable working environment without which this work would not be in its present form.

References

[1] E. -S. Cheng, J. -Y. Yang, J. -D. Lee, K. -Y. Chen, N. -Z. Hu and L. -Y. Chen, "AIoT module development for automated production," 2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), 2021, pp. 1-2, doi: 10.1109/ICCE-TW52618.2021.9603135.

[2] M. Karner, J. Hillebrand, M. Klocker and R. Sámano-Robles, "Going to the Edge - Bringing Internet of Things and Artificial Intelligence Together," 2021 24th Euromicro Conference on Digital System Design (DSD), 2021, pp. 295-302, doi: 10.1109/DSD53832.2021.00052.

[3] S. Teerasoponpong and P. Sugunnasil, "Review on Artificial Intelligence Applications in Manufacturing Industrial Supply Chain – Industry 4.0's Perspective," 2022 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT and NCON), 2022, pp. 406-411, doi: 10.1109/ECTIDAMT-NCON53731.2022.9720417.

[4] P. V. Dudhe, N. V. Kadam, R. M. Hushangabade and M. S. Deshmukh, "Internet of Things (IOT): An overview and its applications," 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), 2017, pp. 2650-2653, doi: 10.1109/ICECDS.2017.8389935.

[5] V. C. Jadala, S. K. Pasupuletti, S. H. Raju, S. Kavitha, C. M. H. Sai Bhaba and B. Sreedhar, "Need of Intenet of Things, Industrial IoT, Industry 4.0 and Integration of Cloud for Industrial Revolution," 2021 Innovations in Power and Advanced Computing Technologies (i-PACT), 2021, pp. 1-5, doi: 10.1109/i-PACT52855.2021.9696696.

[6] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.

[7] U. Atila, M. Uc ar, K Akyoul, et al., "Plant leaf disease classification using EfficientNet deep learning model", Ecological Informatics, 2019.

[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), 2016, pp. 770–778.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in European conference on computer vision (ECCV). Springer, 2016, pp. 630–645.

[10] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton "ImageNet Classification with Deep Convolutional Neural Networks", International Conference on Neural Information Processing Systems - Volume 1, June 2016

[11] Saeedi, B., "The Jetson Artificial Intelligence Tool Chain (JAI-TC)", Master of Science (M.S.), Oregon State University,2019.

[12] Mittal, S., "A Survey on optimized implementation of deep learning models on the NVIDIA Jetson platform", Journal of Systems Architecture, pp. 428-442, 2019. doi.org/10.1016/j.sysarc.2019.01.011.

[13] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and¡ 0.5 mb model size," arXiv preprint arXiv:1602.07360, 2016.

[14] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 6848–6856.

[15] B. Balon and M. Simić, "Using Raspberry Pi Computers in Education," 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2019, pp. 671-676, doi: 10.23919/MIPRO.2019.8756967.

[16] A. Maier, A. Sharp and Y. Vagapov, "Comparative analysis and practical implementation of the ESP32 microcontroller module for the internet of things," 2017 Internet Technologies and Applications (ITA), 2017, pp. 143-148, doi: 10.1109/ITECHA.2017.8101926.

[17] G. Thomas, "What is Flutter and Why You Should Learn it in 2020." https://www.freecodecamp.org/. https://www.freecodecamp.org/news/what-is-flutter-and-why-you-should-learn-it-in-2020 (accessed Dec. 20, 2022).

[18] "Processing3, Integrated Development Environment Overviews," https://processing.org. https://processing.org/environment/ (accessed Dec. 19, 2022).

[19] J. Mesquita, D. Guimarães, C. Pereira, F. Santos and L. Almeida, "Assessing the ESP8266 WiFi module for the Internet of Things," 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), 2018, pp. 784-791, doi: 10.1109/ETFA.2018.8502562.

An improved BSIF descriptor based on GFIPML model and deep architecture for Arabic handwriting recognition

Aicha KORICHI
LINFI Laboratory
University of Mohammed Khider
Biskra, Algeria
aicha.korichi@univ-biskra.dz

***Abstract*— The main aim of this paper is to present a robust offline Arabic handwriting recognition system based on an improved version of BSIF descriptor. The main idea behind the proposed improved descriptor is to learn the filters using small patches from the same dataset as the context of the training. Thus, unlike the traditional BSIF descriptor that uses filters learned from natural images, filters of the improved BSIF includes characters which are learned and convolved with handwriting word images. Moreover, a simple lightweight deep architecture is integrated where the filters are learned via layers. In order to improve the efficiency of the proposed architecture, Generic Feature Independent Pyramid Multi-Level GFIPML model is used to extract the features. Extensive experiments on the public AHDB dataset were conducted. The obtained results have proven the efficiency of the proposed system.**

Keywords*— *BSIF, Deep learning, Offline recognition, Arabic handwriting

I. INTRODUCTION

Recently, a great deal of interest has been devoted to the recognition of offline Arabic handwriting text, which refers to the task of automatically recognizing and interpreting the word/text already existing in the image [1]. Despite its importance in many real applications such as bank check processing and automatic postal mail sorting form data entry, recognizing Arabic handwriting poses unique challenges due to the complexity and variability of Arabic script and considerable efforts remain to be done.

According to the data acquisition method, two branches of handwriting recognition systems are existing offline and online branches [1]. In the offline mode, only static data representing the pixel values (0 or 1) is provided because the input data is taken from a text scan image. However, the input data for the online mode is obtained through a tactile digital screen, and both static and dynamic information on the handwriting trajectory, such as the trajectory coordinates, temporal order, speed, and acceleration, are provided [2]. Moreover, because of the lack of dynamic information, offline handwriting recognition is much more difficult than online handwriting recognition.

Arabic literal amounts recognition is a common problem in Arabic handwriting recognition, as checks are widely used in financial transactions in many countries. Despite human skills to perceive the literal amount of checks, the error rate is related to the number of checks in hand. Therefore, reliance on human agents in such a critical scenario may result in multiple errors in check processing.

From one point of view, existing work can be categorized into four categories. Some existing works [3–6] deal with this issue considered the structural features, whereas, certain others [7, 8] distinguished literal amounts based on statistical features. However, Recent advancements in deep learning techniques, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have significantly improved the accuracy of Arabic handwriting recognition systems.

In this paper, an improved version of a deep BSIF descriptor for extracting features from Arabic literal words integrated with the GFIPML model is proposed. Unlike traditional BSIF, which uses natural images to learn the filters, the filters of the proposed improved version are learned using Arabic words that maximize the quality of the proposed system. Moreover, in order to improve the accuracy of the proposed system, a deep architecture with the GFIPML model is proposed.

The remainder of the paper is structured as follows: Section 2 reviews some relative works on Arabic handwritten amount recognition. Section 3 describes the newly proposed improved BSIF descriptor for extracting features from Arabic handwriting images. Section 4 contains information on the database used and the experimental evaluation outcomes. Finally, we conclude the paper by summarizing some findings and future work.

II. RELATED WORK

Despite recent technological advances that have facilitated financial transactions by automating data processing, over one hundred billion checks are handled globally each year, with the majority of them being manually processed by human agents [9, 10].

Several works have been proposed during the last two decades to address the problem of automatic Arabic handwriting literal check amount recognition. The presented methods can be classified into four types based on the nature of the features used: statistical, structural, hybrid, and deep learning based features.

Recently, several researchers have proposed utilizing statistical features to define Arabic handwriting amounts. In [7], the authors developed an automatic holistic method based on Gabor filters and Bag of Features (BoF) for the recognition of Arabic handwritten literal amounts. Images were filtered by a set of Gabor filters with varying sizes and orientations to extract local characteristics. Gabor filter responses are sorted and then given to BoF frameworks. Their proposed method was evaluated on the CENPARMI database, which contains Arabic handwritten checks, and an identification rate of

86.44% was obtained. For detecting handwritten literal amounts, [11] developed a holistic technique based on SIFT descriptor and PCA for dimensionality reduction. SVM properly identified 58.55% of the words in the AHDB database's 12 classifications. The authors of [12] examined the performance of many CNN networks against some statistical descriptors based on the PML model. The best recognition rate achieved by the LPQ descriptor was 91.52%.

On the other hand, some researchers have proposed that Arabic handwritten literal check amounts be defined by a collection of structural traits. For example, Farah et al [3] proposed using some structural elements (ascenders, descenders, loops, etc.) recovered from 4800 word images representing Arabic handwritten literal amounts. In order to improve system performance further, the scientists combined the results of three classifiers: multilayer neural network (MLP), k-nearest neighbor (KNN), and fuzzy k-nearest neighbor (FKNN), with a recognition rate of 94%. The authors of [4] presented two combination techniques. The first scheme is based on feature combination, and it takes into account four features: angle, distance, horizontal and vertical span. The second strategy use the majority vote technique to join three ELM classifiers. The proposed technique was evaluated using the same database (AHDB), and the recognition rate was 64.63%. The same authors introduced an improved quadratic feature model in [5], with an attained recognition score of 83.06%. Authors of [13] proposed a new model based on integrating N-grams. This system is divided into two parts: the first is concerned with word recognition, while the second incorporates an N-gram model to improve the accuracy acquired in the first stage.

Other studies have used a combination of features (structural, statistical, and transformation-based features) to describe Arabic handwritten amounts. Menasria et al. [14] proposed combining statistical features extracted from the entire image word, such as local chain code histograms (CCH), zoning, Zernike moment invariants (ZMI), and density profile histograms (DPH), with structural features extracted from different image parts to describe Arabic handwriting literal amounts. (i.e., PAWs). SVM correctly classified 95.91% of the 61 classes in AHDB. Images were initially preprocessed in [15], and then transformation-based and statistical features such as Discrete Cosine Transform (DCT) and Histogram of Oriented Gradient (HOG) were recovered from them. The Neural Nets classifier correctly detected 95% of the words in the AHDB database.

III. Proposed system based on an Improved BSIF Descriptor

Preprocessing, feature extraction, and classification are some of the main procedures involved in any Arabic handwriting recognition system. The reprocessing phase is the first step in the development of any recognition systems. It aims to remove irrelevant information, which may harm the overall recognition [16], in order to enhance the image. This stage employs techniques such as noise reduction, thresholding, and normalization. Then, during the feature extraction stage, pertinent features are extracted from each letter or word to reflect its shape, stroke direction, and other attributes. The gathered features are finally classified into appropriate Arabic words during the classification stage using machine-learning methods.

As previously stated, feature extraction approaches can be roughly categorized into three types: structural, statistical, and spatial transformation-based features [17]. The image's structural properties comprise both topological and geometrical characteristics [18]. The total number of dots and their position relative to the baseline, the end of points, and loops are examples of structural elements [19][20]. Hough Transform, Wavelets, Gabor Transform, Fourier Transform, Karhunen-Loeve expression, and moments are examples of space transformation-based features. Statistical characteristics, on the other hand, are numerical measures derived from pixel distribution.

Statistical features are known to have low computational complexity and high speed [21], for this reason, we considered in this work taking advantage of these kinds of features. More precisely, we opt for using Binarized Statistical Image Feature (BSIF) [22]. Indeed, this descriptor has been shown to be effective for a variety of image analysis tasks, including face recognition, texture classification, and object detection. It is particularly useful for images with complex textures and lighting conditions, as it is robust to illumination changes and noise. However, the filters used for BSIF are learned using natural images. We propose an improved version of BSIF descriptor where the filters are learned using the same dataset as the training. Moreover, a simple deep architecture is designed for BSIF in order to extract well the features.

A. Improved BSIF descriptor

Binarized Statistical Image Feature (BSIF) is a texture descriptor that is commonly used in image processing and computer vision. It was introduced in 2012 by Kannala and Rahtu [22] as a method for describing local texture patterns in grayscale images.

The BSIF descriptor works by analyzing the local texture patterns in an image by dividing the image into small overlapping patches. For each patch, a binary code is generated that represents the presence or absence of specific texture patterns. This binary code is computed by comparing the pixel intensities in the patch to a set of predefined binary patterns. Each pattern consists of a center pixel and its neighboring pixels, which are assigned a binary value based on their relative intensities.

To extract the BSIF descriptor from an image, the image is first convolved with a set of linear filters. In contrast to other methodologies, the filters used by BSIf are learned from statistics of a small set of natural image patches by maximizing the statistical independence of the filter responses (ICA) [23]. The resulting codes are concatenated to form a feature vector that represents the texture characteristics of the image.

As we have mentioned BSIF uses natural images to learn the filters, our idea in this paper is to learn the filters using the same dataset as the training. Indeed, to extract features from Arabic handwriting word, filters containing Arabic characters should be used; which motivate us to change the manner of filters used.

B. Proposed system

A handwriting words recognition system is an automatic system that simulates human reading abilities. The proposed

system is divided into two major stages: feature extraction and classification.

For extracting features, we used the proposed improved BSIF descriptor with GFIPML model (Figure 1).

Algorithm 1 Handwriting recognition

Background: a set of images from AHDB database
Input: Handwriting descriptors : HV = HV1 , HV2 , ...HVn, Images , Number of levels : M, Combinaison scheme models
Output: Digital format of the image content : Dg
for all image ∈ AHDB test do
  RV(image)← []
  for i=1 to M do
    [X,Y]← size(image)
    for j=1 to i do
      H ← floor(X/j)
      W ← floor(Y/j)
      HL ← mod(X/H)
      WL ← mod(Y/W)
      for xx=1 to X-HL step=H do
        for yy=1 to Y-WL step=W do
          block ← image(xx:xx+H-1,yy:yy+W-1)
          RV(image) ← RV(image)+HV(block)
  image ← resize(image,X/2,Y/2)
First decisions ← fed each RV for the same classifier
Second decisions ← apply combinaison schemes(First decisions)
Final decisions Dg ← Majority voting of (Second decisions)

Figure 1: GFIPML model Algorithm [25].

The procedure of learning the filters is passed through a deep architecture in order to learn useful filters. As for the classification, we use SVM classifier to recognize words. The general scheme of the proposed system is illustrated in Figure2 :

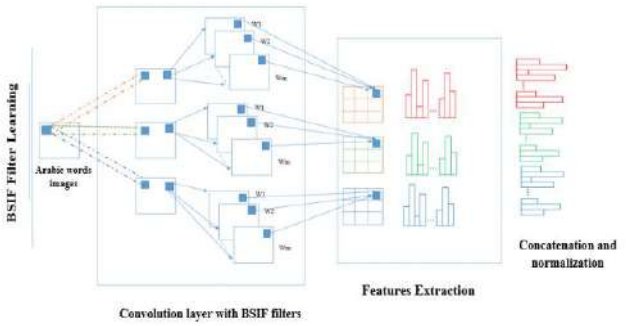


Figure 2: The general scheme of the proposed system.

As it is shown in Figure 2, the proposed system contains three main stages: learning BSIF filters, which are used as detectors for the input Arabic handwriting images, features extraction for computing the features, and finally a normalization stage for histogram normalization where we use a robust Tied Rank normalization technique (see Figure 3).

Algorithm 1 Histogram normalization based on Tied Rank (TR) principal

Input: Extracted block-wise histograms of an image : H, Number of channels D, Number of blocks B
Output: TR normalized histogram : v
V=[]
for d=1:D do
  for b=1:B do
    Compute $\overline{H}_b^d$ the the tied rank of H_b^d
    $V_b^d = \sqrt{\overline{H}_b^d}$
    $\hat{V}_b^d = L_2$ normalization of V_b^d
    $V=V+\hat{V}_b^d$

Figure 3: TR nomalization

IV. Experimental results

In this section, we provide additional details regarding the proposed system for Arabic handwriting literal amount recognition, in which a thorough comparison was performed between several models, and the related stateof-the-art methods. To evaluate the performance of the proposed model, we conduct experiments on the challenging AHDB database [24]. The AHDB database contains the words that appear the most frequently in checks. It is composed of 70 classes representing various words, with each word written by 105 different writers. We divide the 6615-image dataset into three folds, each of which contains 2189, 2183, and 2243 images, respectively. Two folds are used for training and the remaining folds are used for testing in each cross-validation iteration. Table 1 displays a sample of each class from the database.

Table 1: Arabic words used to express amounts on checks extracted from AHDB database

N°	Arabic name	N°	Arabic name	N°	Arabic name	N°	Arabic name	N°	Arabic name
1	أحد	14	خمسة	27	تسعة	40	ستة	53	ثلاثمئة
2	احدى	15	خمسمئة	28	تسعمئة	41	ستمئة	54	ثلاثمائة
3	واحد	16	خمسمائة	29	تسعمائة	42	ستمائة	55	عشرون
4	ثمان	17	أربع	30	تسعون	43	ستون	56	عشرين
5	ثمانية	18	أربعة	31	تسعين	44	ستين	57	اثنان
6	ثمانمئة	19	أربعمئة	32	لا	45	عشر	58	اثنين
7	ثمانمائة	20	أربعمائة	33	سبع	46	عشرة	59	مئتين
8	ثمانون	21	أربعون	34	سبعة	47	ثلاثون	60	مائتين
9	ثمانين	22	أربعين	35	سبعمئة	48	ثلاثين	61	ألفان
10	اثنى	23	مئة	36	سبعمائة	49	ألف	62	ألفين
11	خمسون	24	مائة	37	سبعون	50	آلاف	63	غير
12	خمسين	25	مليون	38	سبعين	51	ثلاث	/	/
13	خمس	26	تسع	39	ست	52	ثلاثة	/	/

All the experiments done in this section were obtained using: MATLAB installed on a computer with a Core i7 "7th generation" processor, 16 GB of RAM memory, and AMDA Radeon graphical card. As an evaluation metric, we have used the accuracy metric, which is the quotient of the total number of correctly classified words corrected over the total number of words.

By remembering our aim that is proposing an improved version of BSIF descriptor for extracting features from Arabic handwriting images. Thus, extensive experiments including the impact of the improved version of BSIF, the impact of TR normalization and GFIPML model are conducted. All results are summarized in Figure 4.

According to the results illustrated in Figure4, the improved version of BSIF gives very high accuracy compared to traditional BSIF which prove our hypothesis that filters learned via small patches of words gives better results than filters trained via natural images for the context of Arabic handwriting recognition. Moreover, including TR

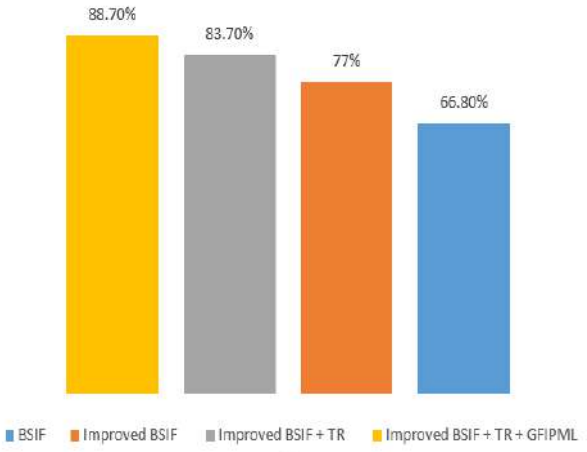


Figure 4: Experimental results.

normalization improves well the results. Our idea of integrating TR with GFIPML model gives the best and the highest results.

V. Conclusion

Nowadays, deep learning techniques have proven their efficiency for several recognitions tasks. However, they are known also to be data hungry. In this paper, we proposed a novel improved version of BSIF descriptor that uses filters including characters learned from the same dataset as the training. A simple deep architecture is proposed then to learn useful filters .Moreover, an efficient GFIPML model is used to extract the features. The experimental results conducted on the public challenging dataset proved the efficiency of the proposed system.

References

1. LORIGO, Liana M. et GOVINDARAJU, Venugopal. Offline Arabic handwriting recognition: a survey. IEEE transactions on pattern analysis and machine intelligence, 2006, vol. 28, no 5, p. 712-724.
2. ZOUARI, Ramzi, BOUBAKER, Houcine, et KHERALLAH, Monji. A time delay neural network for online arabic handwriting recognition. In : Intelligent Systems Design and Applications: 16th International Conference on Intelligent Systems Design and Applications (ISDA 2016) held in Porto, Portugal, December 16-18, 2016. Springer International Publishing, 2017. p. 1005-1014.
3. AL-NUZAILI, Qais, HAMDI, Ali, HASHIM, Siti Z. Mohd, et al. An enhanced quadratic angular feature extraction model for arabic handwritten literal amount recognition. In : Recent Trends in Information and Communication Technology: Proceedings of the 2nd International Conference of Reliable Information and Communication Technology (IRICT 2017). Springer International Publishing, 2018. p. 369-377.
4. AL-NUZAILI, Qais, HASHIM, Siti Z. Mohd, SAEED, Faisal, et al. Enhanced structural perceptual feature extraction model for Arabic literal amount recognition. International Journal of Intelligent Systems Technologies and Applications, 2016, vol. 15, no 3, p. 240-254.
5. AL-NUZAILI, Qais Ali, HASHIM, Siti Z. Mohd, SAEED, Faisal, et al. Pixel distribution-based features for offline Arabic handwritten word recognition. International Journal of Computational Vision and Robotics, 2017, vol. 7, no 1-2, p. 99-122.
6. FARAH, Nadir, SOUICI, Labiba, et SELLAMI, Mokhtar. Classifiers combination and syntax analysis for Arabic literal amount recognition. Engineering Applications of Artificial Intelligence, 2006, vol. 19, no 1, p. 29-39.
7. ASSAYONY, Mohammed O. et MAHMOUD, Sabri A. Recognition of Arabic handwritten words using Gabor-based bag-of-features framework. International Journal of Computing and Digital Systems, 2018, vol. 7, no 01, p. 35-42.
8. HASSEN, Hanadi et AL-MAADEED, Somaya. Arabic handwriting recognition using sequential minimal optimization. In : 2017 1st international workshop on arabic script analysis and recognition (ASAR). IEEE, 2017. p. 79-84.
9. AHMAD, Irfan et MAHMOUD, Sabri A. Arabic bank check analysis and zone extraction. In : Image Analysis and Recognition: 9th International Conference, ICIAR 2012, Aveiro, Portugal, June 25-27, 2012. Proceedings, Part I 9. Springer Berlin Heidelberg, 2012. p. 141-148.
10. AHMAD, Irfan et MAHMOUD, Sabri A. Arabic bank check processing: state of the art. Journal of Computer Science and Technology, 2013, vol. 28, no 2, p. 285-299.
11. QAWASMEH, Yasser, AWWAD, Sari, OTOOM, Ahmed Fawzi, et al. Local patterns for offline Arabic handwritten recognition. International Journal of Advanced Intelligence Paradigms, 2020, vol. 16, no 2, p. 203-215.
12. KORICHI, Aicha, SLATNIA, Sihem, AIADI, Oussama, et al. Arabic handwriting recognition: Between handcrafted methods and deep learning techniques. In : 2020 21st International Arab Conference on Information Technology (ACIT). IEEE, 2020. p. 1-6.
13. LAMSAF, Asmae, AITKERROUM, Mounir, BOULAKNADEL, Siham, et al. Recognition of arabic handwritten text by integrating n-gram model. In : Innovations in Smart Cities Applications Volume 4: The Proceedings of the 5th International Conference on Smart City Applications. Springer International Publishing, 2021. p. 1490-1502.
14. MENASRIA, Azzeddine, BENNIA, Abdelhak, NEMISSI, Mohamed, et al. Multiclassifiers system for handwritten Arabic literal amounts recognition based on enhanced feature extraction model. Journal of Electronic Imaging, 2018, vol. 27, no 3, p. 033024-033024.
15. HASSAN, Alia Karim Abdul et KADHM, Mustafa S. Handwriting word recognition based on neural networks. International Journal of Applied Engineering Research, 2015, vol. 10, no 22, p. 43120-43124.
16. HUANG, Bing Quan, ZHANG, Y. B., et KECHADI, Mohand Tahar. Preprocessing techniques for online handwriting recognition. In : Seventh International Conference on Intelligent Systems Design and Applications (ISDA 2007). IEEE, 2007. p. 793-800.
17. EL QACIMY, Bouchra, HAMMOUCH, Ahmed, et KERROUM, Mounir Ait. A review of feature extraction techniques for handwritten Arabic text recognition. In : 2015 International Conference on Electrical and Information Technologies (ICEIT). IEEE, 2015. p. 241-245.
18. CHERIET, Mohamed, KHARMA, Nawwaf, SUEN, Ching, et al. Character recognition systems: a guide for students and practitioners. John Wiley & Sons, 2007.
19. ARICA, Nafiz et YARMAN-VURAL, Fatos T. An overview of character recognition focused on off-line handwriting. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2001, vol. 31, no 2, p. 216-233.
20. ZAIZ, Faouzi, BABAHENINI, Mohamed Chaouki, et DJEFFAL, Abdelhamid. Puzzle based system for improving Arabic handwriting recognition. Engineering Applications of Artificial Intelligence, 2016, vol. 56, p. 222-229.
21. KHORSHEED, Mohammad S. Off-line Arabic character recognition–a review. Pattern analysis & applications, 2002, vol. 5, p. 31-45.
22. KANNALA, Juho et RAHTU, Esa. Bsif: Binarized statistical image features. In : Proceedings of the 21st international conference on pattern recognition (ICPR2012). IEEE, 2012. p. 1363-1366.
23. HYVÄRINEN, Aapo et OJA, Erkki. Independent component analysis: algorithms and applications. Neural networks, 2000, vol. 13, no 4-5, p. 411-430.

24. AL-MA'ADEED, Somaya, ELLIMAN, Dave, et HIGGINS, Colin A. A data base for Arabic handwritten text recognition research. In : Proceedings eighth international workshop on frontiers in handwriting recognition. IEEE, 2002. p. 485-489.
25. KORICHI, Aicha, SLATNIA, Sihem, AIADI, Oussama, et al. A generic feature-independent pyramid multilevel model for Arabic handwriting recognition. Multimedia Tools and Applications, 2022, vol. 81, no 15, p. 20719-20739.

Proposal for a four-channel all-optical filter based on linear photonic crystals

1st Kouddad Elhachemi
Telecommunication and Digital Signal Processing Laboratory, Faculty of Electrical Engineering, Department of Telecommunications, University Djillali Liabes, Sidi-Bel-Abbes22000, Algeria
Kouddad20@hotmail.fr

2nd Dekkiche Leila
Telecommunication and Digital Signal Processing Laboratory, Faculty of Electrical Engineering, Department of Telecommunications, University Djillali Liabes, Sidi-Bel-Abbes22000, Algeria
ldekkiche_dz@yahoo.com

3rd Hassani Islam
Laboratoire de développent durable et informatique (LDDI), University of Adrar, Algeria
islamhassani65@yahoo.com

4rd Yaichi Ibrahim
Laboratoire de développent durable et informatique (LDDI), University of Adrar, Algeria
i.yaichi@univ-adrar.edu.dz

***Abstract*— In our work, we have designed a new four-channel all-optical filter based on two-dimensional photonic crystals. This filter can extract four totally different wavelengths at each output from some techniques used like ring resonators, scattering rods, resonant microcavities. The two well-known methods like PWE and FDTD are often applied to calculate the band gap diagram and to study the behavior of the light inside our filter. The obtained results show that the proposed filter has a high transmission power is 99.1%, a very high-quality factor is 15 628, with a very low crosstalk is -41.6. In addition, the operating parameters show that the high-performance results have been confirmed with other recently published work. The maximum response time of our proposed filter is ultra-fast with an ultra-compact size, which makes this filter easy to use in integrated photonic circuits.**

Keywords*— *All-optical filter, Integrated optical circuits, Quality factor, Resonant cavity, Ring resonator, Photonic crystal

I. INTRODUCTION

In recent years, researchers have tried to use all-optical circuits in several fields due to the development of photonics [1]. Important fields such as computer science, chemistry, medicine, astronomy, pharmacy, and telecommunications, are exploiting photonics to improve the infrastructure that has enabled the development of each field [2]. Enabled applications that have based on Photonic Crystals (PhC) in design and realization such as filters [3-6] logic gates [7-10],
encoders[11], PhC optical fibers [12], sensors [13], de-multiplexers [14]. All-optical add/drop filters are commonly used which consist of a single input channel and three output channels. To evaluate the performance in all types of all-optical filters parameters such as: Quality factor, data rate, channel spacing, and crosstalk are used. Ghorbanpour and Makouei [15] designed a four-output all-optical filter characterized by quality factor, transmission power, crosstalk, and channel spacing of 5967, 90%, -16.5 dB, 3 nm, respectively. In another work, Alipour-Banaei et al. [16] presented a design of a four-channel all-optical filter characterized by quality factor, transmission power, crosstalk, and channel spacing of 387, 77%, -13.5 dB, 4.4 nm, respectively. The four-output all-optical filter was designed by Delphi et al. [17] characterized by quality factor, transmission power, crosstalk, and channel spacing of 5443, 93.5%, -14.9 dB, 2.6 nm, respectively.

In this work, we proposed a linear all-optical filter based on photonic crystals which is designed by four channels using resonant cavities coupled with two ring resonators placed in cascade. The characteristics of this proposed design are adequate to havea very high-quality factor, very high bit rate operate, very low crosstalk, and ultra-compact size of this filter.

II. PROPOSED ALL-OPTICAL FILTER

The proposed filter (figure (1)) consists of two cascaded ring resonators, with several resonant cavities and eight scattering rods placed at the corner of each resonator. A disconnected square array of 28 × 43 rods has a refractive index of 3.8 (GaAs) suspended in air. The area of the fundamental filter structure is equal to 357.86 μm^2. The distance between the two rod centers is assumed to be "a", and the radius of dielectric rods in red is R= 0.2 × a and all rods radii forming the resonant cavities of each output are shown in table 1.

The figure (2) shows the band diagram of the proposed filter using the PWE (Plane Wave Expansion) method, in which appears two PBGs in TE mode (Transverse Electric): $0.26< a/\lambda < 0.40$ and $0.69< a/\lambda < 0.71$.The first PBG covers the third telecommunication window. For analyze the emission, accuracy, and speed of the light inside ourfilter we have used the FDTD (Finite Difference time Domain) method. To determine the mesh size and to ensure the stability of the system, the two following equations (1) and (2) have been considered:

$$\Delta x = \Delta z = \frac{a}{2} \qquad (1)$$

$$\Delta t \leq \frac{1}{c\sqrt{\frac{1}{\Delta x^2}\frac{1}{\Delta z^2}}} \qquad (2)$$

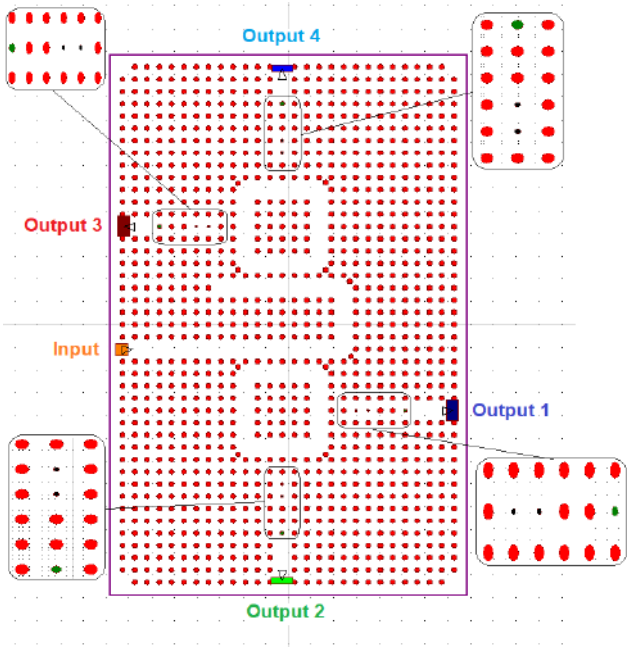


Fig. 1. The proposed all-optical filter.

TABLE 1: Rods radii of each channel of the proposed all-optical filter shownby the figure (1)

Rods	Output 1	Output 2	Output 3	Output 4
Black	R*0.409	R*0.42	R*0.44	R*0.46
Red	R	R	R	R
Green	R*0.61	R*0.71	R*0.81	R*0.91

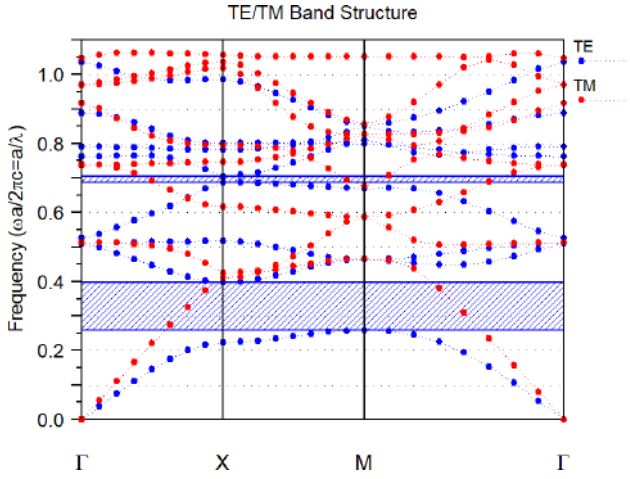


Fig. 2. The band diagram for our all-optical filter based on a square lattice.

III. THE NUMERICAL RESULTS OF THE ALL-OPTICAL FILTER

To calculate the response of our proposed all-optical filter,we excite with an optical Gaussian pulse at the input, as shown in figure (3a). The designed structure allows certain frequenciesto pass from the input to the output. An optical signal centered at 1522.6 nm comes out of output 1 of the filter, with a transmission efficiency of 92.06%, a quality factor of 7613, and an optical signal centered at 1539.6 nm comes out of output 2 with a transmission efficiency of 99.1 %, a quality factor of 15 396. The output 3 can guide the optical signal centered at 1562.8 nm with a transmission efficiency of 54.43 %, a quality factor of 15 628, and output 4 carries the optical signal at 1592.8 nm with a transmission efficiency of 56 % and a quality factor of Q = 7964. The figure (3b) shows the all-optical filter responses in dB to evaluate the crosstalk of the filter for the four outputs. Figure 4 illustrates the behavior of the optical field inside the filter and Table.II shows the results obtained, which are better than other recent published works.

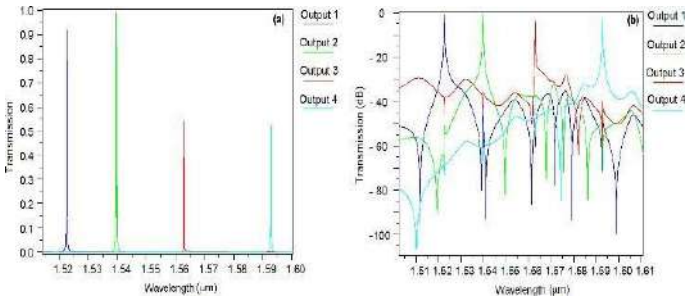


Fig. 3. Illustration of the transmission (a) percentage (%), (b) in decibel (dB).

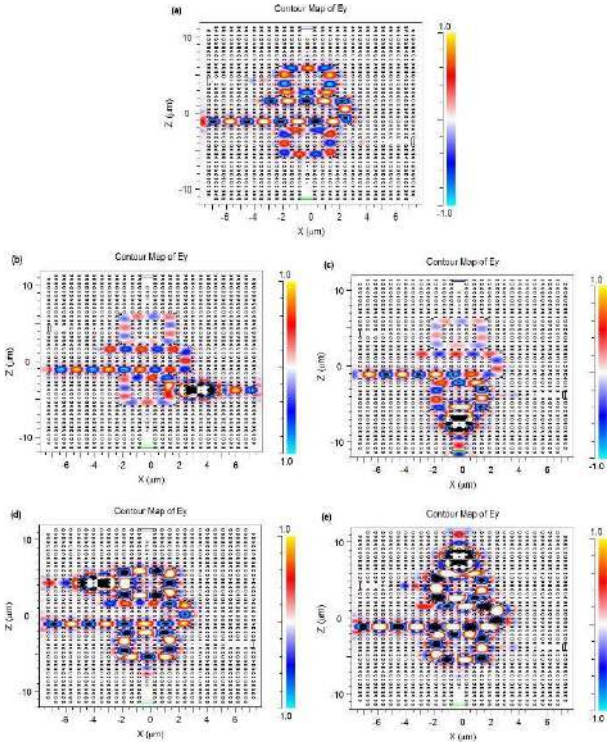


Fig. 4. The optical field propagation within the proposed all-optical filter for input optical pulses centered at (a) 1.5226 µm, (b) 1.5396 µm, (c) 1.5628 µm, (d) 1.5928 µm.

TABLE II: The obtained results compared with previously published works

Previously published works	Number of output channels	Transmission Power (%)	Quality Factor (Q)	Crosstalk in (dB)
Our work	4	99.1	15 628	-41.6
[18]	2	90	5967	-16.5
[19]	4	98	3805	-41
[20]	2	93.5	5443	-14.19
[21]	2	77.5	387	-13.5

CONCLUSION

The proposed work presents a contribution to the design of a four-channel all-optical filter based on two resonators ring placed in cascade with photonic crystal resonant cavities. The designed filter is characterized by very important advantages such as low crosstalk, high quality factor for each channel, very high bit rate, narrow inter-channel spacing, high spectral efficiency, with a compact size, which shows that this filter is suitable for optical integrated circuits, and has a better performance compared to other recent filters already published.

ACKNOWLEDGMENT

This work was supported by the Directorate General for Scientific Research and Technological Development (DGRSDT).

REFERENCES

[1] Wabnitz, S., & Eggleton, B. J. (2015). All-optical signal processing. Springer Series in Optical Sciences, 194.

[2] Razavi, B. (2012). Design of integrated circuits for optical communications. John Wiley & Sons.

[3] Butt, H., Dai, Q., Farah, P., Butler, T., Wilkinson, T. D., Baumberg, J. J., & Amaratunga, G. A. (2010). Metamaterial high pass filter based on periodic wire arrays of multiwalled carbon nanotubes. Applied physics letters, 97(16), 163102.

[4] McCrindle, I. J., Grant, J., Drysdale, T. D., & Cumming, D. R. (2014). Multi – spectral materials: hybridisation of optical plasmonic filters and a terahertz metamaterial absorber. Advanced Optical Materials, 2(2), 149-153.

[5] Hosseinzadeh Sani, M., Ghanbari, A., & Saghaei, H. (2020). An ultra-narrowband all-optical filter based on the resonant cavities in rod-based photonic crystal microstructure. Optical and Quantum Electronics, 52(6), 1-15.

[6] Naghizade, S., & Saghaei, H. (2020). Tunable graphene-on-insulator band-stop filter at the mid-infrared region. Optical and Quantum Electronics, 52(4), 1-13.

[7] Kouddad, E., & Naoum, R. (2020). Optimization of an all-optical photonic crystal NOT logic gate using switch based on nonlinear Kerr effect and ring resonator. Sensor Letters, 18(2), 89-94.

[8] Kouddad, E., Naoum, R., Vigneswaran, D., & Maheswar, R. (2021). Performance evaluation of all-optical NOT, XOR, NOR, and XNOR logic gates based on 2D nonlinear resonant cavity photonic crystals. Optical and Quantum Electronics, 53(12), 1-15.

[9] Kouddad, E., & Rafah, N. (2020). A novel proposal based on 2D linear resonant cavity photonic crystals for all-optical NOT, XOR and XNOR logic gates. Journal of Optical Communications.

[10] Kouddad, E., Vigneswaran, D., Rafah, N., Koundal, D., & Leila, D. (2022). All optical logic gates function by ring resonator properties aiding photonic crystal. Physica Scripta, 97(10), 105502.

[11] Ouahab, I., & Naoum, R. (2016). A novel all optical 4× 2 encoder switch based on photonic crystal ring resonators. Optik, 127(19), 7835-7841.

[12] Saghaei, H. (2018). Dispersion-engineered microstructured optical fiber for mid-infrared supercontinuum generation. Applied optics, 57(20), 5591-5598.

[13] Kouddad, E, Rafah, N., & Leila, D. (2021). High Sensitivity And Ultra-High Quality Factor For An All-Optical Temperature Sensor Based On Photonic Crystal Technology.

[14] Talebzadeh, R., Soroosh, M., Kavian, Y. S., & Mehdizadeh, F. (2017). Eight-channel all-optical demultiplexer based on photonic crystal resonant cavities. Optik, 140, 331-337.

[15] Ghorbanpour, H., & Makouei, S. (2013). 2-channel all optical demultiplexer based on photonic crystal ring resonator. Frontiers of Optoelectronics, 6(2), 224-227.

[16] Alipour-Banaei, H., Serajmohammadi, S., & Mehdizadeh, F. (2015). Effect of scattering rods in the frequency response of photonic crystal demultiplexers. J. Optoelectron. Adv. Mater, 17(3-4), 259-263.

[17] Delphi, G., Olyaee, S., Seifouri, M., & Mohebzadeh-Bahabady, A. (2019). Design of an add filter and a 2-channel optical demultiplexer with high-quality factor based on nano-ring resonator. Journal of Computational Electronics, 18(4), 1372-1378.

[18] Ghorbanpour, H., & Makouei, S. (2013). 2-channel all optical demultiplexer based on photonic crystal ring resonator. Frontiers of Optoelectronics, 6(2), 224-227.

[19] Foroughifar, A., Saghaei, H., & Veisi, E. (2021). Design and analysis of a novel four – channel optical filter using ring resonators and line defects in photonic crystal microstructure. Optical and Quantum Electronics, 53(2), 1-12.

[20] Delphi, G., Olyaee, S., Seifouri, M., & Mohebzadeh-Bahabady, A. (2019). Design of an add filter and a 2-channel optical demultiplexer with high-quality factor based on nano-ring resonator. Journal of Computational Electronics, 18(4), 1372-1378.

[21] Alipour-Banaei, H., Serajmohammadi, S., & Mehdizadeh, F. (2015). Optical wavelength demultiplexer based on photonic crystal ring resonators. Photonic Network Communications, 29(2), 146-150.

Deep Learning for Atherosclerosis Detection: a study of Explainable AI Algorithms

Amel Laidi
Faculty of Engineering, LIMOSE Laboratory
M'Hamed Bougara University
Boumerdes, Algeria
a.laidi@univ-boumerdes.dz

Mohammed Ammar
Engineering Systems and Telecommunication Laboratory
University M'Hamed Bougara
Boumerdes, Algeria
m.ammar@univ-boumerdes.dz

***Abstract*— Atherosclerosis is a cardiovascular disease that can be diagnosed through non-invasive imaging techniques such as Coronary CT Angiography (CCTA). Deep learning algorithms have shown great potential in assisting the classification of CCTA images. However, the lack of transparency in the decision-making process of black-box AI models is a major challenge in the medical field, where healthcare professionals need to understand the reasoning behind AI systems' recommendations. In this study we explore the potential benefits of Explainable AI (XAI) algorithms for a previously developed deep learning model for atherosclerosis classification from CCTA images. The study showed promising results not only in explaining the model's decisions, but in recognizing falsely classified cases and eventually improving the model's performance.**

Keywords*— *deep learning, atherosclerosis, coronary CT angiography, Explainable AI, GradCam.

I. INTRODUCTION (*HEADING 1*)

Explainable AI (XAI) is a relatively new concept that refers to the ability of an artificial intelligence system to provide understandable and transparent explanations of its decision-making process. This is particularly important in healthcare due to the high risks associated with the decision-making process, but it can be as important in other fields such as finance, security, and criminal justice.

When a model provides explanations of its decisions, it allows for a better understanding of how the system reached a specific decision or recommendation, which in turn helps build trust in the system and enables the user to improve accountability and to identify and correct any biases or errors in the system. Additionally, explainable AI can help comply with regulatory requirements and ethical standards, as well as enhance the overall effectiveness and fairness of the AI system. [1]

In medical imaging, AI systems are used to assist in the diagnosis of diseases, by analyzing medical images and providing recommendations to healthcare professionals, such decisions are directly responsible for the medical treatment plan of critical patients. In the case of atherosclerosis screening from CCTA images, there can be serious consequences if an AI system fails to detect the presence of the disease. This could result in a missed diagnosis, delayed treatment, and potentially life-threatening complications. If an AI system makes a recommendation without providing an explanation of how it arrived at that decision, healthcare professionals may not fully trust the system and may be hesitant to act on its recommendations.

In this paper we study the overall explainability of a Deep Learning model created for atherosclerosis screening from Coronary CT Angiography images. The model we are trying to explain is a previously developed model, trained and tuned for the specific task of atherosclerosis screening.

II. BACKGROUND

A. *What is AI ? and why deep learning?*

artificial intelligence is a wide term that refers to all types of computer systems trained to perform tasks that are normally associated with human intelligence or abilities, such as perception, reasoning, learning, and decision-making. AI systems are designed to analyze large amounts of data, recognize patterns, and make predictions or decisions based on that data.

There are different types of AI systems, including:

- **Rule-based systems:** AI systems that use pre-defined rules to make decisions or perform tasks. They are limited to the rules that have been programmed into them and cannot adapt to new situations.
- **Machine learning systems** : AI systems that can learn from data and improve their performance over time. There are different types of machine learning, including supervised learning, unsupervised learning, and reinforcement learning.
- **Deep learning systems** : a subset of machine learning systems that use artificial neural networks to analyze large amounts of data and learn from it. Deep learning systems are particularly effective in tasks that involve image and speech recognition, natural language processing, and other complex tasks.

Deep learning involves the use of artificial neural networks that consist of multiple layers of interconnected nodes, each of which performs a specific function in the processing of data. The input data is fed into the network, and the network gradually learns to recognize patterns in the data by adjusting the weights of the connections between nodes.

Deep learning has shown tremendous promise in computer vision and healthcare applications. It can analyze and interpret complex images, such as X-rays or MRI scans, with greater accuracy than traditional image analysis techniques. This ability has enabled medical professionals to diagnose diseases and conditions at an early stage, leading to better treatment outcomes. Additionally, deep learning has been used in

healthcare to improve patient outcomes through personalized treatment plans. By analyzing vast amounts of patient data, deep learning algorithms can identify patterns and correlations that would be impossible for human doctors to detect. This has the potential to lead to more accurate diagnoses, better treatment plans, and ultimately, improved patient outcomes. Overall, the combination of deep learning, computer vision, and healthcare has the potential to revolutionize the way we diagnose and treat diseases, ultimately leading to better healthcare for everyone. [2]

B. Deep learning for atherosclerosis detection from Coronary CT Angiography

Atherosclerosis is a common cardiovascular disease that is characterized by the accumulation of fatty deposits in the walls of arteries. Coronary CT Angiography (CCTA) is the key technique for atherosclerosis screening, it is a non-invasive imaging technique that can be used to visualize the coronary arteries and detect the presence of atherosclerotic plaques [3]. However, interpreting CCTA images can be challenging and time-consuming, particularly when there are multiple plaques or complex plaque morphology.

Deep learning algorithms have been developed to assist in the classification of CCTA images by automatically identifying and characterizing atherosclerotic plaques based on their morphology, composition, and location within the coronary arteries. These algorithms can analyze large amounts of data and learn to recognize patterns that are indicative of different types of plaques, such as calcified or non-calcified plaques.

In a recent study [4], we created a deep learning model and trained it on a publicly available dataset.

The dataset is an open-source collection of Coronary CT Angiography images for screening atherosclerosis. It consists of Mosaic Projection View (MPV) images of 18 views of straightened coronary arteries from 500 patients, created by combining unique ray-traced projections. The dataset was partitioned into 300 training images, 100 testing images, and 100 validation images. To balance the dataset, the training images were augmented 6-fold. The validation dataset contains one randomly selected artery per normal case and one diseased case. [5]

The model we used was a pretrained Residual network [6] with 101 layers. It scored 95.21% of accuracy, 90.48% positive predictive value, and 95.6% negative predictive value.

C. Challenges and limitations of black-box AI models

Black-box AI models, which refer to AI systems that are difficult or impossible to interpret, particularly Deep learning models, have become increasingly popular in the medical field for tasks such as atherosclerosis screening. However, there are several challenges and limitations associated with these models that need to be addressed to ensure their safe and effective use in clinical settings.

Not understanding the full process happening inside the model leads to a lack of transparency in the decision-making process. This can be particularly problematic in the medical field, if healthcare professionals cannot understand how an AI system arrived at a particular diagnosis or recommendation, they may be hesitant to act on it.

To address these challenges and limitations, recent research worked on developing AI models that are more transparent and interpretable, and to ensure that they are trained on unbiased and representative data. Additionally, there needs to be ongoing monitoring and evaluation of AI systems to ensure that they remain accurate and up to date with changes in the data and patient population. By addressing these challenges and limitations, the potential of black-box AI models to improve atherosclerosis screening and other medical tasks can be fully realized.

D. Explainable AI Algorithms

Explainable AI (XAI) algorithms are a type of AI system that is designed to be transparent and interpretable, allowing humans to understand how the AI system arrived at its decision or recommendation. XAI algorithms are intended to address the limitations of black-box AI models, which are often difficult or impossible to interpret.

There are several popular explainable AI algorithms that are used to provide transparency and interpretability to AI models. In this paper we took interest in three algorithms:

LIME (Local Interpretable Model-agnostic Explanations): LIME is a model-agnostic method for explaining the predictions of any black-box model. LIME works by creating a simpler, interpretable model that is locally faithful to the black-box model around the instance being explained. The simpler model can then be used to provide an explanation for the black-box model's prediction. [7]

LIME works by generating perturbations around the instance being explained and measuring the impact of these perturbations on the model's output. The algorithm then uses these perturbations to create a simpler, interpretable model that is locally faithful to the black-box model around the instance being explained.

LIME can provide a way to explain the prediction of any black-box model, without requiring knowledge of its internal workings. LIME is also a model-agnostic algorithm, meaning that it can be used with any type of model, including neural networks, decision trees, and support vector machines.

Grad-CAM (Gradient-weighted Class Activation Mapping): Grad-CAM is an algorithm for generating heatmaps that visualize the importance of different regions of an image for a given classification decision. Grad-CAM works by computing the gradients of the output class score with respect to the feature maps in the final convolutional layer of a neural network. The resulting gradient-weighted maps are then used to generate a final heatmap that visualizes the importance of different regions of an image for a given classification decision. [8]

The advantage of Grad-CAM over black-box models is that it provides a way to visualize the decision-making process of the neural network, which can be difficult to interpret using traditional methods. Grad-CAM also provides a way to identify which regions of an image are most important for the model's decision, which can be particularly useful in medical imaging.

***Occlusion Sensitivity**:* Occlusion sensitivity is a technique for visualizing the importance of different regions of an image for a given classification decision by systematically occluding different parts of the image and measuring the resulting change in the model's output. By comparing the model's output for the original image and the occluded images, it is possible to identify the most important regions of the image for the model's decision. [9]

Occlusion sensitivity is also a simple and easy-to-implement algorithm, making it a useful tool for interpreting the decisions of machine learning models.

E. XAI algorithms in atherosclerosis classification.

The use of explainable AI algorithms in atherosclerosis classification offers the typical advantages such as increased transparency and improved decision making, but it also adds another layer to the task of the classification. The main advantage of classifying the images directly, without first segmenting them, is the rapidity of the task. When using XAI algorithms, we can add a layer on the explanation, where the highlighted areas are most likely areas of high activity, in our case it could be the location of buildup. This not only allows for a quick detection of the disease, but it also guides the doctor to which region to investigate for higher risk of plaque.

III. RESULTS AND DISCUSSION

For a full analysis of the model, we ran the three algorithms on a number of images from the validation dataset. We tested true positives , true negatives, and false negatives.

False negatives were not part of the scope if this study, because they do not pose an urgent risk. A healthy patient that has been falsely classified as sick would be kept under medical care and eventually identified as healthy with further tests. While sick patients that were classified as healthy can be discharged without medical care, leading to further complications of their health.

A. True positive

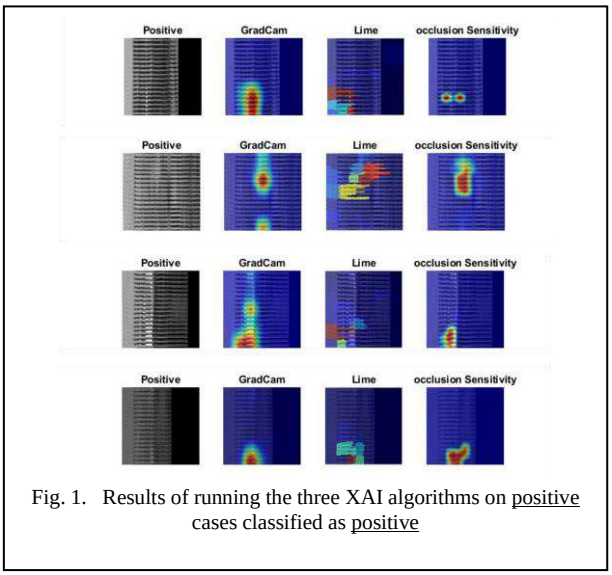


Fig. 1. Results of running the three XAI algorithms on positive cases classified as positive

True positive (TP) is a term used to describe a correctly predicted positive case. In other words, it is the case where the model correctly identifies a positive example as positive. In our case, the model has 90.48% positive predictive value, so it should be good at classifying positive cases.

At first sight, all three algorithms show similar activity at the same regions of the image (Fig. 1). This means that the deep learning model we used relied heavily on this region for its final decision.

The focus is on one small region, being highlighted in red, with no significant activity anywhere else.
The explainable models offer an extra layer of performance, it shows the region with higher probability of plaque and residue, which is useful for the rest of the medical care process.

B. True Negative (TN)

True negative (TN) is a term used to describe a correctly predicted negative case. In other words, true negative occurs when the actual class of an image is negative, and the model correctly predicts it as negative.
Our model has a 95.6% negative predictive value; and although it is considered as an impressive performance, it would help more to understand the reason behind the classification.
Both GradCam and LIME show a more spread-out activity (Fig. 2), which is quite different from positive cases. However, occlusion sensitivity for TN does not look too different from the one of TP, so it cannot be used for comparison.

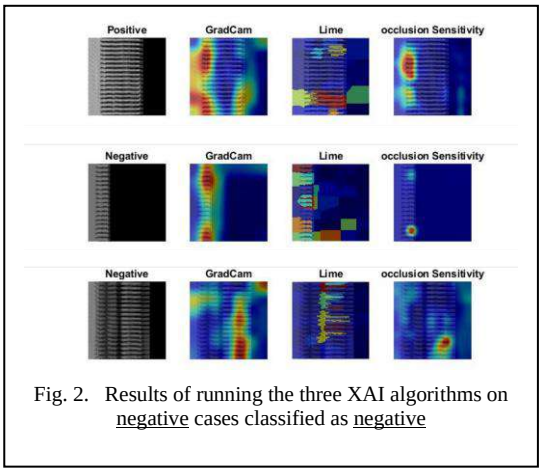


Fig. 2. Results of running the three XAI algorithms on negative cases classified as negative

C. False Negative

A false negative (FN) is an error in binary classification where a negative outcome is predicted when the actual outcome is positive. In other words, a false negative occurs when a model fails to recognize a positive image or when it incorrectly classifies a positive example as negative. False negatives are particularly important in medical diagnoses as they can lead to the dismissal of sick people when their case ca be critical or even fatal.

All three algorithms show an activity that is less focused than true positives (Fig. 3), which explains the error in classification. It still is different from true negatives.
The images have small areas of high activity (red regions), but also significantly large areas with lesser activity (orange and yellow regions).

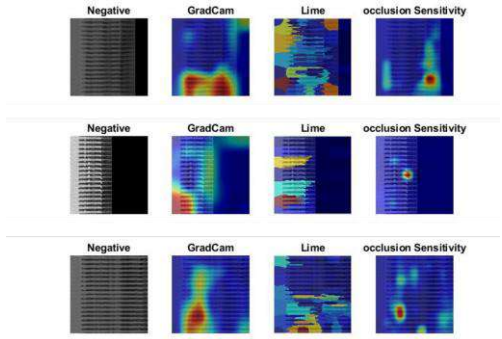


Fig. 3. Results of running the three XAI algorithms on positive cases classified as negative

D. *Comparaison between FN and TN*

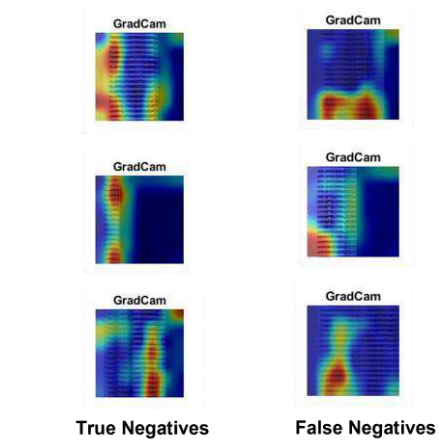


Fig. 4. A comparison between generated heatmaps of negative case classified as negative (left), and positive case classified as negative (right)

As mentioned earlier, false negatives are of critical importance in medical applications, because they refer to patients who are positive but were classified by the model as negative. Their dismissal can lead to fatal complications.

In this experiment, we compared the GradCam analysis of true negative cases and false negative cases. The choice of GradCam was due to it being the only algorithm that showed a significant variation in behavior .

Fig. 4. shows that true negative cases generate heatmaps with a larger surface for high activity (red area) that is well spread vertically. This suggests that a larger region of the image was considered for the final decision of the model. Unlike true positive cases, where only a small region was considered, probably the location of plaque.

As for false negatives, the heatmaps have a smaller localized surface for high activity (red area), with other region with lesser activity (orange and yellow areas). Having the lesser activity can explain WHY the images were falsely classified as negative, a larger area was considered for the final decision.

In conclusion, GradCam shows a tangible difference between false negatives and true negatives, showing significant activity in small areas for false negatives, indicating not only that they were wrongfully classified , but the region for the physician to focus on before making their final decision. Eventually improving the performance of the model.

IV. CONCLUSION

Artificial intelligence techniques have revolutionized healthcare applications, offering human-like performance on classification tasks in a fraction of the time. However, they still can't be trusted by most healthcare professionals due to the ambiguity behind the decision making process.

Explainability algorithms allow us to unlock the full potential of AI, by offering an extra layer of trustworthiness in addition to the human level performance.

In this paper we revisited a deep learning model used for atherosclerosis screening from Coronary CT Angiography. In an attempt to understand the reasoning of the model, we applied three of the most prominent algorithms used in explainable AI: Gradient-weighted Class Activation Mapping (GradCam), Local Interpretable Model-agnostic Explanations (LIME), and Occlusion sensitivity. The results we obtained offered not only more information about the classification, but also somehow improved the model's performance.

Positive cases correctly classified as positive (TP) showed high focused activity in one small area, which would suggest it is the location of plaque.

Positive images that have been wrongfully classified as negative (FN) have the same focused activity in a small region, which is an indication for the doctor to take a second look on the regions to make the final decision of the existence of plaque or buildup, and eventually the decision of the presence of atherosclerosis.

Negative cases correctly classified as negative (TN) have a more spread-out heatmap, meaning that the model did not use one region for its final decision, but considered a wider area of the image. GradCam was the best algorithm to spot the difference between FN and TN.

The main contribution of this study is proving that the deep learning model relies on detecting focalized high activity (most probably plaque position) for its final decision, this explanation makes the model more trustworthy. In addition, the study offers an extra layer of performance, as it shows the probable location of plaque in positive cases, and most importantly offers a way for telling true negative cases from false negative cases, which improves the overall performance of the model, and reduces the risk of early dismissal of sick patients.

The study was limited to a number of samples , and should be better generalized in future research, other models with different performance should also be tested for comparison.

REFERENCES

[1] C. Molnar, Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models, Lulu. com, 2020.

[2] A. Rajkomar, J. Dean and I. Kohane, "Machine learning in medicine," *New England Journal of Medicine*, vol. 380, no. 14, pp. 1347--1358, 2019.

[3] M. Portegies, P. Koudstaal and M. Ikram, "Cerebrovascular disease," in *Handbook of Clinical Neurology*, F. B. D. F. S. Michael J. Aminoff, Ed., Elsevier, 2016, pp. 239-261.

[4] A. Laidi, M. Ammar, M. E. H. Daho and S. Mahmoudi, "Deep Learning Models for Coronary Atherosclerosis Detection in Coronary CT Angiography," *Current Medical Imaging*, vol. 19, 2023.

[5] V. Gupta, M. Bigelow, B. Erdal, L. Prevedello and R. White, "Image dataset for a CNN algorithm development to detect coronary atherosclerosis in coronary CT angiography," *Mendeley Data*, vol. 1, 2019.

[6] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[7] M. T. Ribeiro, S. Singh and C. Guestrin, "" Why should i trust you?" Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016.

[8] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization.," *arXiv preprint arXiv:1610.02391*.

[9] "visualizing and understanding convolutional networks.," in *European conference on computer vision*, 2014.

Simulated Annealing for the UAVs placement problem

Sylia Mekhmoukh Taleb
LIST Laboratory
University of M'Hamed Bougara
Boumerdes, Algeria
s.mekhmoukh@univ-boumerdes.dz

Yassine Meraihi
LIST Laboratory
University of M'Hamed Bougara
Boumerdes, Algeria
y.meraihi@univ-boumerdes.dz

Selma Yahia
LIST Laboratory
University of M'Hamed Bougara
Boumerdes, Algeria
s.yahia@univ-boumerdes.dz

Amylia Ait Saadi
LIST Laboratory
University of M'Hamed Bougara
Boumerdes, Algeria
a.aitsaadi@univ-boumerdes.dz

Asma Benmessaoud Gabis
LMCS Laboratory
Ecole Nationale Supérieure d'Informatique
Algiers, Algeria
a_benmessaoud@esi.dz

Amar Ramdane-Cherif
LISV Laboratory
University of Paris-Saclay
Velizy, France
rca@lisv.uvsq.fr

***Abstract*—Unmanned Aerial Vehicles (UAVs) have emerged as a promising solution for providing wireless coverage in areas that are difficult to access, such as disaster zones and remote areas. The placement of UAVs in the network is a challenging optimization problem. In fact, the performance of the network is mainly depends on the locations of UAVs. This paper suggest the application of Simulated Annealing (SA) for solving the UAVs placement problem. The main objective is to determine the best locations of UAVs in order to maximize the number of covered users. SA is validated in terms of user coverage using six instances with different numbers of UAVs. The simulation results demonstrated that SA gives better results in comparison Particle Swarm Optimization (PSO).**

***Index Terms*—Simulated Annealing, UAVs placement problem, optimization**

I. Introduction

Unmanned Aerial vehicles (UAVs) have received a lot of attention as a prospective wireless connectivity solution. UAVs can be rapidly deployed in emergency scenarios, such as terrestrial BSs failure, to meet the sudden demand for wireless devices [1, 2, 3]. Clearly, UAVs should be deployed at locations that maximize the number of covered users.

The placement of UAVs in a network is a complex optimization problem that belongs to the family of NP-Hard problems. Meta-heuristics have been proposed to solve this problem with reasonable time execution, such as Genetic Algorithm (GA) [5], Particle Swarm Optimization (PSO) [6], Artificial Bee Colony (ABC) [7], Grey Wolf Optimizer [8], and Brain Storm Optimization (BSO) [9].

In this study, we suggest the use of SA for solving the UAVs placement problem. SA was validated using six instances with various numbers of UAVs.

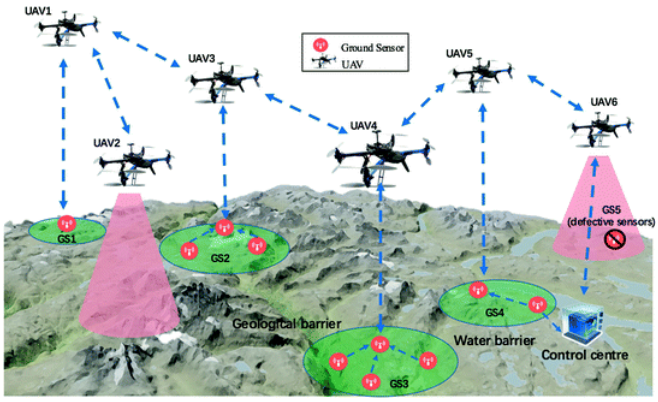


Fig. 1: UAV-based cooperative wireless network for disaster situation [4]

The remaining sections of the paper are structured as follows. The formulation of the UAVs placement problem is presented in Section 2. SA is described in Section 3. Results of simulation are shown in section 4. Section 5 contains the conclusion and upcoming works.

II. The UAV placement problem formulation

The system is modeled in a 3-dimensional deployment area with two types of nodes such as users and UAVs. Therefore, let:

- UV is the set of m UAVs: $UV = \{UV_1, UV_2, ..., UV_m\}$. Each UAV is equipped with one radio interface to communicate with users and other UAVs. The position of each UAV is defined by its coordinates (x,y,z). The height (z coordinate) of the drone is defined based on the coverage radius and the visibility angle.
- US is the set of n users $US = \{US_1, US_2, ..., U_n\}$, we assume that users are randomly distributed in a

$2D$ rectangle area of dimension WxH. The user US_i is covered by a UAV UV_j if it is within its coverage radius R. Each user can be associated with at most one UAV. It can be within a coverage radius of various UAVs but it is associated with the closest UAV.

The main objective of this work is to determine the best placement of a given number of UAVs in order to maximize the coverage metric, depending on the locations of users. The coverage metric for solution G is described below:

$$Cov(G) = \frac{\sum_{i=1}^{n}(max_{j\in\{1,...m\}}\sigma_{ij})}{n} \quad (1)$$

where the coverage variable σ_{ij} is defined as follows:

$$\sigma_{ij} = \begin{cases} 1 & if\ US_i\ is\ covered\ by\ UV_j \\ 0 & Otherwise. \end{cases} \quad (2)$$

III. Simulated Annealing (SA)

In 1983, Kirkpatrick et al. [10] developed the single-based meta-heuristic known as Simulated Annealing (SA). The main concept of SA is based on the annealing theory which simulates the cooling process of metal atoms. Numerous optimization problems, such as the issue of node placement [11, 12, 13], have been addressed using SA.

SA begins with an initial solution X and Temperature Tmp. For each iteration t in $[1tmax]$, SA searches for X' the neighbor of the current solution X. X' solution is accepted only in two cases: Firstly, if $\delta \leq 0$, where $\delta = f(X') - f(X)$, $f(X')$ and $f(X)$ are fitness values of the neighbor and current solution, respectively. Secondly, if the Boltzmann probability $P = e^{\Delta/Tmp}$ is greater then a random value r and $\delta > 0$. At the end of the iteration, the temperature Tmp decreases with a cooling factor Cf. This process is repeated until the maximum number of iterations is reached. Algorithm 1 illustrates the SA algorithm's pseudo-code.

IV. Simulations

The performance of the SA in contrast to PSO will be assessed in this section. The evaluation is done in terms of coverage metric. The parameters used in simulation are described in Table I.

Figures 2.a and 2.b report examples of a planned network using SA and PSO, respectively. The planned network is a solution of network instance with 12 UAVs and 100 users. Green points represent users randomly distributed in the deployment area of $200m \times 200m \times 25m$. Each installed UAV is represented by orange diamond. These figures show that there is 15 and 25 uncovered users by SA and PSO, respectively. Thus, SA gives betters results in terms of coverage when compared with PSO.

As illustrated in Figure 3 and Table II, the performance of SA is evaluated in terms of coverage under various numbers of UAVs. The presented results demonstrated that as the number of UAVs increases, the coverage increases. In fact, the uncovered users have more chance to be covered by the added UAVs. Again, the results proved the effectiveness of SA in comparison with PSO.

Algorithm 1 The pseudo-code of Simulated Annealing algorithm

```
Initialize SA parameters: Initial Temperature Tmp_0, cooling factor Cf, and number of the neighborhoods ns in the search space.
Generate initial solution X
while (t < tmax) do
    while (it < n) do
        Generate X' using equation
        Calculate Δ = f(X') − f(X)
        Generate a random uniform variable r
        if (Δ < 0) then
            Determine X' the neighbor of X
        else
            if (exp^{−Δ/Tmp} > r) then
                X = X'
            end if
        end if
        it = it + 1
    end while
    t = t + 1
end while
return The best solution X_best
```

TABLE I: Parameters values considered in our simulations

Parameter	Value
N° of users n	100
N° of UAVs m	[2 12]
Coverage radius R	25 m
Width W	200 m
Length L	200 m
Height H	25 m
Total number of iterations	1000
N° of runs	10

TABLE II: Coverage under various number of UAVs

m	**2**	**4**	**6**	**8**	**10**	**12**
SA	23.3%	**41.5%**	**56.9%**	**69.2%**	**78.3%**	**86.4%**
PSO	**25%**	37.9%	49.7%	62%	68.5%	73.3%

Figures 4a, 4b, 4c, 4d, 4e, and 4f depict the convergence of the proposed algorithms under various numbers of UAVs (2, 4, 6, 8, 10, and 12, respectively). The presented results showed that SA needs more iterations to achieve the global optimum in comparison with PSO. However, the coverage reached by SA is the best for most of cases.

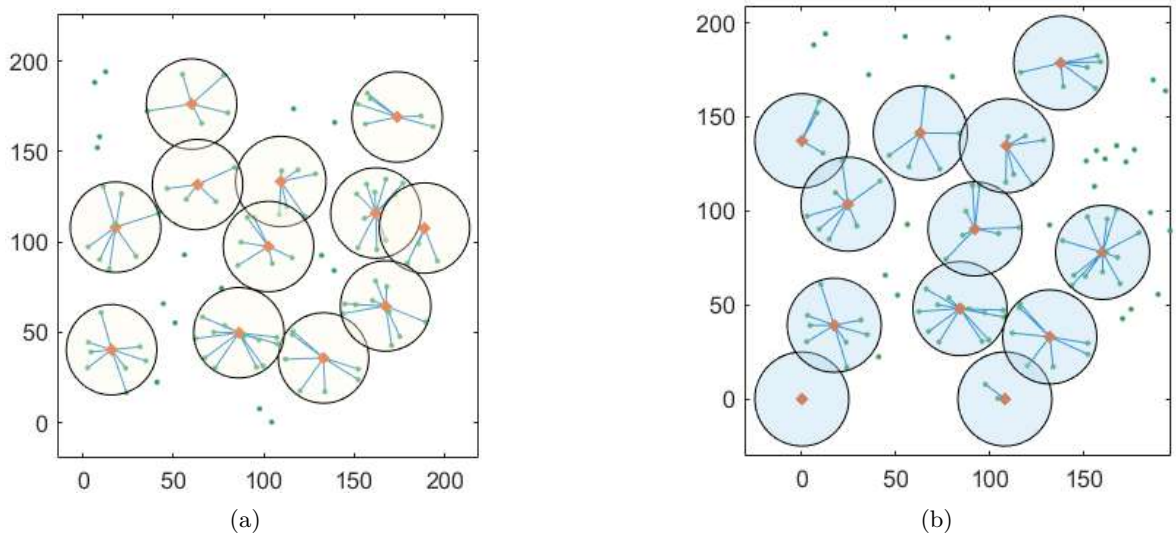


Fig. 2: obtained placement using: (a) SA (b) PSO

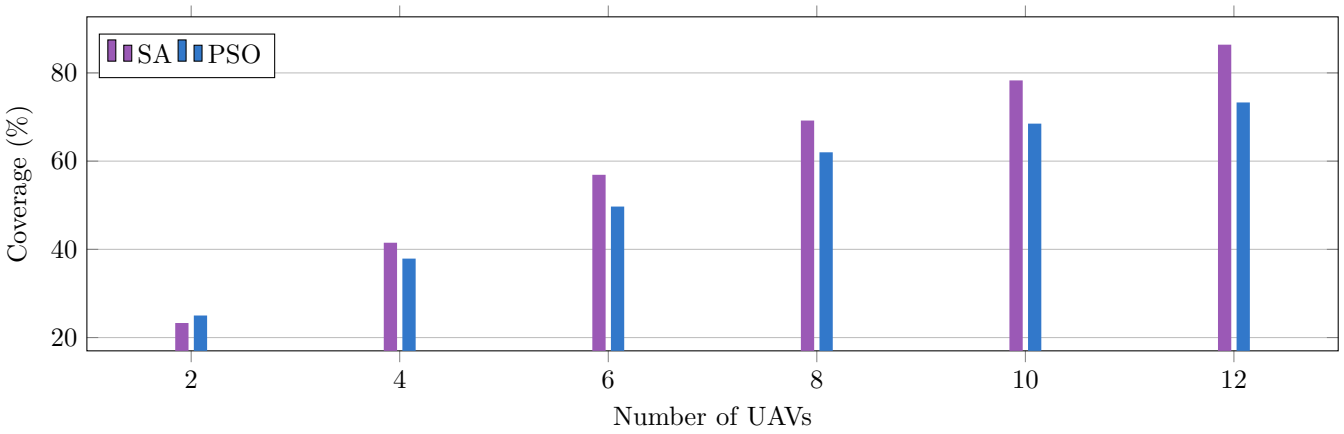


Fig. 3: Coverage under various numbers of UAVs

V. Conclusion

This paper presents the optimization of UAVs placement for wireless coverage, which is a critical challenge for providing connectivity in areas that are difficult to access. The study proposes the use of SA as a meta-heuristic optimization technique to solve the UAVs placement problem. The results of the simulation experiments demonstrate that SA outperforms PSO in terms of coverage metric. For future work, we attempt to combine the global search capability of PSO with the local search capability of SA for solving the UAVs placement problem

References

[1] Yong Zeng, Rui Zhang, and Teng Joon Lim. Wireless communications with unmanned aerial vehicles: Opportunities and challenges. *IEEE Communications Magazine*, 54(5):36–42, 2016.

[2] Elham Kalantari, Muhammad Zeeshan Shakir, Halim Yanikomeroglu, and Abbas Yongacoglu. Backhaul-aware robust 3d drone placement in 5g+ wireless networks. In *2017 IEEE international conference on communications workshops (ICC workshops)*, pages 109–114. IEEE, 2017.

[3] Irem Bor-Yaliniz and Halim Yanikomeroglu. The new frontier in ran heterogeneity: Multi-tier drone-cells. *IEEE Communications Magazine*, 54(11):48–55, 2016.

[4] Chunbo Luo, Wang Miao, Hanif Ullah, Sally McClean, Gerard Parr, and Geyong Min. Unmanned aerial vehicles for disaster management. In *Geological Disaster Monitoring Based on Sensor Networks*, pages 83–107. Springer, 2019.

[5] Xukai Zhong, Yiming Huo, Xiaodai Dong, and Zhonghua Liang. Qos-compliant 3-d deployment optimization strategy for uav base stations. *IEEE Systems Journal*, 15(2):1795–1803, 2020.

[6] Hazim Shakhatreh, Abdallah Khreishah, Ayoub Alsarhan, Issa Khalil, Ahmad Sawalmeh, and

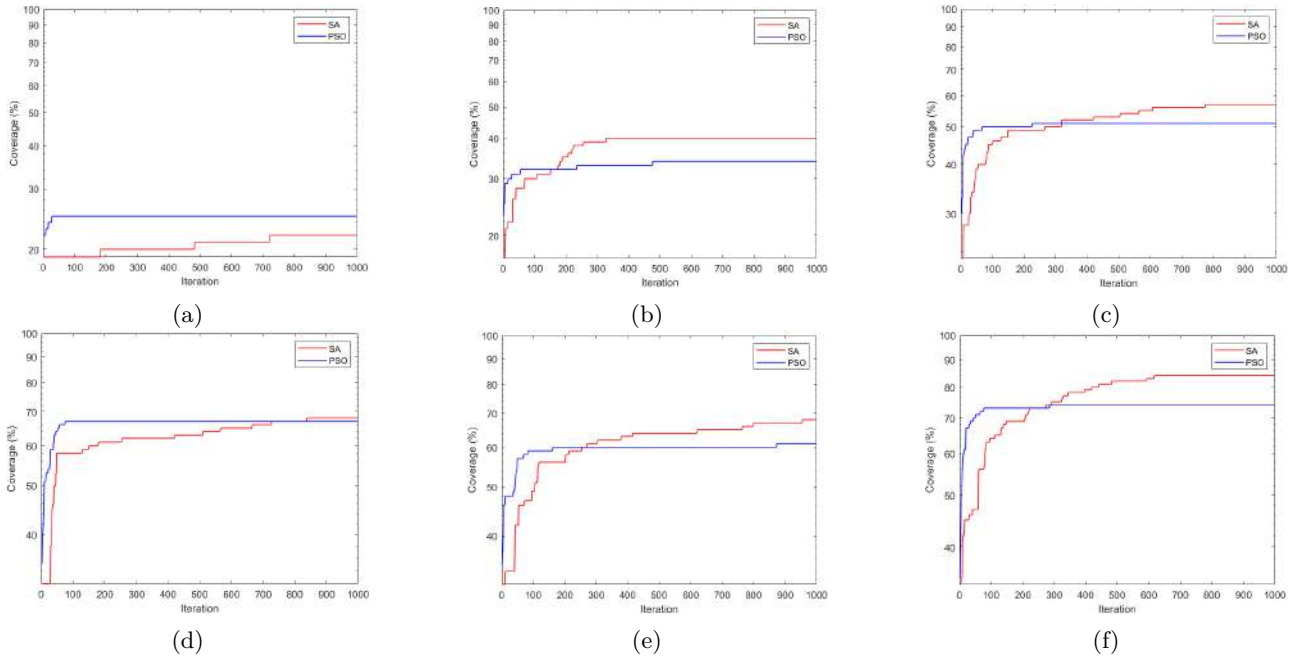


Fig. 4: Example of algorithms convergence using: (a) 2 UAVs (b) 4 UAVs (c) 6 UAVs (d) 8 UAVs (e) 10 UAVs (f) 12 UAVs

Noor Shamsiah Othman. Efficient 3d placement of a uav using particle swarm optimization. In *2017 8th International Conference on Information and Communication Systems (ICICS)*, pages 258–263. IEEE, 2017.

[7] Selcuk Aslan and Sercan Demirci. Solving uav localization problem with artificial bee colony (abc) algorithm. In *2019 4th International Conference on Computer Science and Engineering (UBMK)*, pages 735–738. IEEE, 2019.

[8] Mohamed Amine Ouamri, Marius-Emil Oteşteanu, Gordana Barb, and Cedric Gueguen. Coverage analysis and efficient placement of drone-bss in 5g networks. *Engineering Proceedings*, 14(1):18, 2022.

[9] Eva Tuba, Romana Capor-Hrosik, Adis Alihodzic, and Milan Tuba. Drone placement for optimal coverage by brain storm optimization algorithm. In *International Conference on Hybrid Intelligent Systems*, pages 167–176. Springer, 2017.

[10] Scott Kirkpatrick, C Daniel Gelatt Jr, and Mario P Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.

[11] Nicholas Hao Zheng Lim, Ying Loong Lee, Mau Luen Tham, Yoong Choon Chang, Allyson Gek Hong Sim, and Donghong Qin. Coverage optimization for uav base stations using simulated annealing. In *2021 IEEE 15th Malaysia International Conference on Communication (MICC)*, pages 43–48. IEEE, 2021.

[12] Majid Raissi-Dehkordi, Karthikeyan Chandrashekar, and John S Baras. Uav placement for enhanced connectivity in wireless ad-hoc networks. Technical report, 2004.

[13] Fatos Xhafa, Admir Barolli, Christian Sánchez, and Leonard Barolli. A simulated annealing algorithm for router nodes placement problem in wireless mesh networks. *Simulation Modelling Practice and Theory*, 19(10):2276–2284, 2011.

Hybrid classification system using GoogLeNet and support vector regression for the diagnosis of knee osteoarthritis

Khadidja Messaoudene
LIMOSE Laboratory
University M'Hamed Bougara
Boumerdes, Algeria
k.messaoudene@univ-boumerdes.dz

Khaled Harrar
LIST Laboratory
University M'Hamed Bougara
Boumerdes, Algeria
khaled.harrar@univ-boumerdes.dz

Dehia Abdiche
LIST Laboratory
University M'Hamed Bougara
Boumerdes, Algeria
d.abdiche@univ-boumerdes.dz

***Abstract*—Osteoarthritis (OA) of the knee is a prevalent and chronic degenerative joint disease, affecting a substantial portion of the global population. Diagnosis of OA can be challenging due to the need for extensive analysis of medical imaging data, such as X-rays and Magnetic Resonance Imaging (MRI). In this study, we present a hybrid system that combines deep feature-based and Machine Learning (ML) approaches for detecting knee OA. The proposed method employs Gabor filter-based preprocessing, data augmentation with translation and rotation, feature extraction using the GoogLeNet model, feature selection via F-Score, and classification using the Support Vector Regression (SVR) model. Our experimental results demonstrate that the proposed approach outperforms existing state-of-the-art methods, achieving an accuracy rate of 83.6%. These findings suggest that the hybrid system has the potential to improve the accuracy and efficiency of diagnosing knee OA.**

***Index Terms*—KOA, Gabor filter, GoogLeNet, F-Sore, SVR.**

I. Introduction

Osteoarthritis of the knee is a degenerative joint disease that is characterized by the loss of cartilage and the deterioration of joint function [1]. It is a major cause of pain, disability, and reduced quality of life for millions of people worldwide. The use of imaging techniques, such as X-rays and Magnetic Resonance Imaging (MRI) [2], has long been a valuable tool in the diagnosis of knee OA. However, the interpretation of these images can be subjective and can vary between radiologists. To address this challenge, researchers have turned to Artificial Intelligence (AI) and Deep Learning (DL) techniques to assist in the diagnosis of knee OA. Machine Learning (ML) techniques have the potential to automate this analysis, enabling faster and more accurate diagnosis of knee OA. However, the performance of such techniques depends heavily on the preprocessing and feature extraction methods used. Preprocessing techniques such as edge detection and image enhancement can improve the quality of the images and facilitate the identification of key features. Features extraction techniques can capture the most informative features in the images, enabling accurate classification of knee OA. Several studies have explored the use of Convolutional Neural Networks (CNNs) for the detection of knee OA from radiographs. For example, Jakaite et al. [3] used different texture features and ML methods to analyze high-resolution X-ray images. The Haralick features and Zernike moments were used to optimize the performances of the ML techniques, including Random Forest (RF), Support Vector Machines (SVM), Artificial Neural Network (ANN), and the proposed GMDH-type network. Zernike moments with DMDH provided better accuracy (77.5%) than Haralick features. Wang et al. [4] proposed an automated approach for knee OA classification using Deep Neural Network (DNN). The approach involves preprocessing the knee X-ray images using frequency-domain filtering and histogram normalization, and a two-step classification strategy is used to extract the joint center and classify OA grades. The method reported a classification accuracy of 81.41%. The proposed techniques appear effective, but further validation is needed on larger and more diverse datasets before implementing them in the clinical setting. Cueva et al. [5] proposed a semi-automatic CADx model based on Deep Siamese CNN and a fine-tuned ResNet-34 for detecting OA lesions in the two knees according to the KL scale. The model was trained using a public dataset, and the validation was performed using a private dataset. The imbalanced dataset problem was addressed using transfer learning. The model achieved an average multi-class accuracy of 61%, with better performance for classes KL0, KL3, and KL4 than KL1 and KL2. The model results were compared and validated using the classification of experienced radiologists. Lim et al. [6] presented a DL approach to predict the presence of OA in subjects aged 50 years and older using statistical data. The study uses a DNN and Principal Component Analysis (PCA) to automatically generate features from the data and identify risk factors for OA prevalence. They achieved an accuracy of 71.97%.

In this study, we propose a novel ML-based approach for the detection of knee OA that includes several innovative techniques, including Gabor filter-based preprocessing, data augmentation with translation and rotation, feature extraction with GoogLeNet model, feature selection with F-Score, and classification with Support Vector Regression (SVR) machine.

The proposed approach aims to improve the accuracy and interpretability of knee OA detection, making it more suitable for clinical use.

The remainder of this paper is organized as follows: Section II presents materials and methods where the dataset and the methods used are described. In Section III the obtained results are reported and discussed. A conclusion is provided in Section IV.

II. Materials & Methods

A. Dataset

The dataset used in our experiment was obtained from the publicly accessible Osteoarthritis Initiative (OAI) database [7]. The data contains 688 radiographs of the knee that have been categorized using the Kellgren and Lawrence (KL) rating system (KL0, KL2). We compared grade (no OA) to disease overall grade KL2 (mild OA). In our study, we worked with the medial region of the knee radiographic image. Fig. 1 shows the Region Of Interest (ROI) used in our work. Table I demonstrates the data distribution.

TABLE I
Data distribution

Data	Training data	Test data	Validation data
KL0	262	58	24
KL2	262	58	24

Fig. 1. Knee radiographic image

B. Methods

Our approach to classifying KOA involves several techniques, which we will describe in this section. The method we propose consists of four main steps: firstly, we perform artifact removal through preprocessing, followed by deep feature extraction. Then, we apply feature selection using F-Score and finally classify the results into KL0 and KL2 categories. Fig. 2 provides a visual representation of the flowchart of our proposed approach.

1) Preprocessing: The Gabor descriptor is widely used for textured images and has been shown to be robust to changes in illumination and scale in several works in the literature. This descriptor allows the extraction of microstructure and macrostructure information that describes the texture in an image. Linear Gabor filters are used for various applications such as iris recognition, object detection, and medical imaging [8]–[10].

The Gabor filter is a complex wavelet filter created by multiplying a modulated sine and cosine wave by a two-dimensional Gaussian window. The Gaussian function is used to define the spatial extent and the bandwidth of the filter, while the sinusoidal wave is used to define the frequency and orientation of the filter. By varying the parameters of the Gaussian and sinusoidal functions, Gabor filters can be designed to selectively highlight different features of an image, such as edges, textures, or blobs. The 2D Gabor filter is described mathematically in equations (1) and (2):

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{1}{2}\frac{(\bar{x}^2+\bar{y}^2)}{\sigma_x^2+\sigma_y^2}+2\pi jW\bar{x}\right) \quad (1)$$

$$\bar{x} = x\cos\theta + y\sin\theta \bar{y} = -x\sin\theta + y\cos\theta \quad (2)$$

where : σ_x and y indicate the spread of the current pixel in the neighborhood. W is the center frequency of the complex sinusoid. $\theta \in [0,\pi]$ denotes the orientation of the bands. The number of oscillations and the angle of the Gabor filter are determined by the frequency and orientation. These two main parameters allow a representation of multi-scale and multi-orientation textures, the adjustment of which is given by the following formulas (3) and (4):

$$\frac{dy}{dx} = f(x,y) \quad (3)$$

$$y(x_0) = y_0 \quad (4)$$

In summary, in order to extract the Gabor descriptor, the image is convoluted with a bank of Gabor filters at different frequencies and orientations, and the resulting histograms are used as feature vectors for texture classification. Fig. 3 shows the result of the preprocessing stepp.

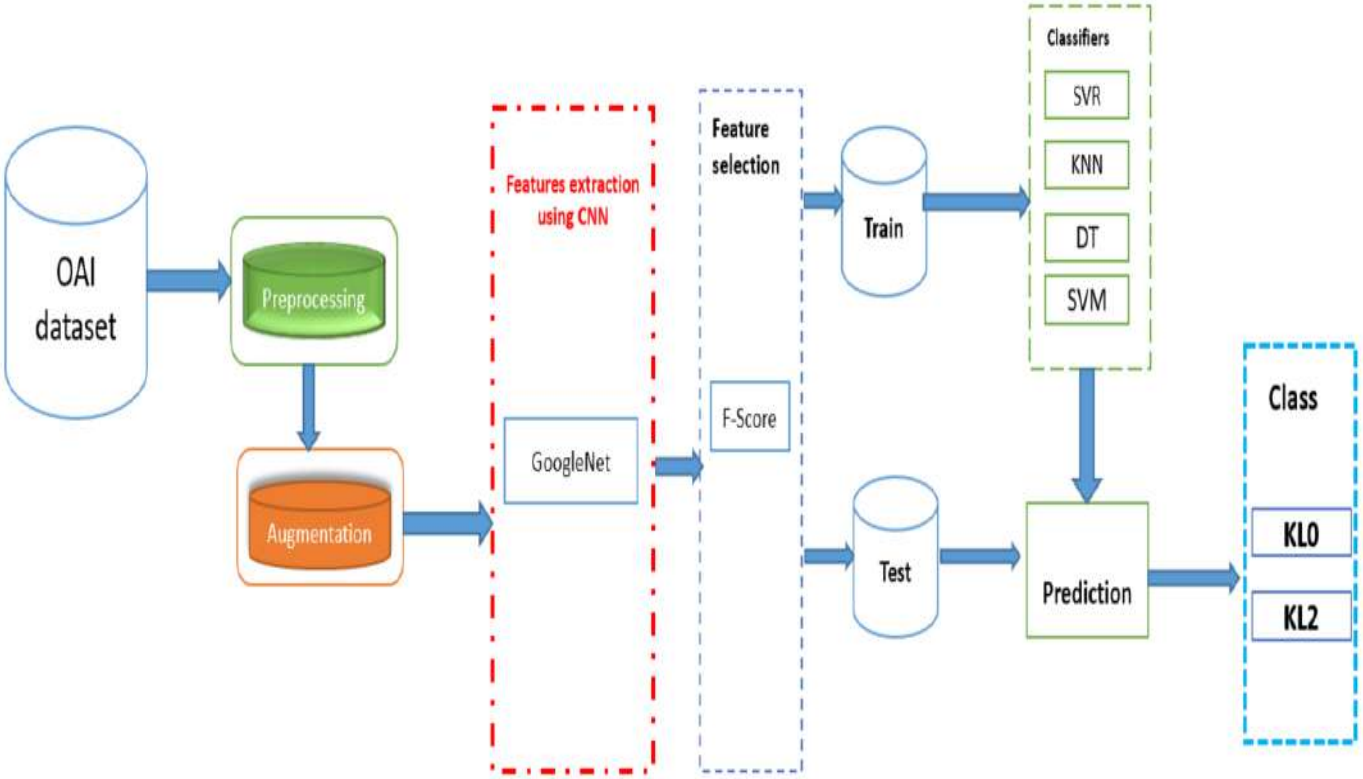


Fig. 2. Proposed method

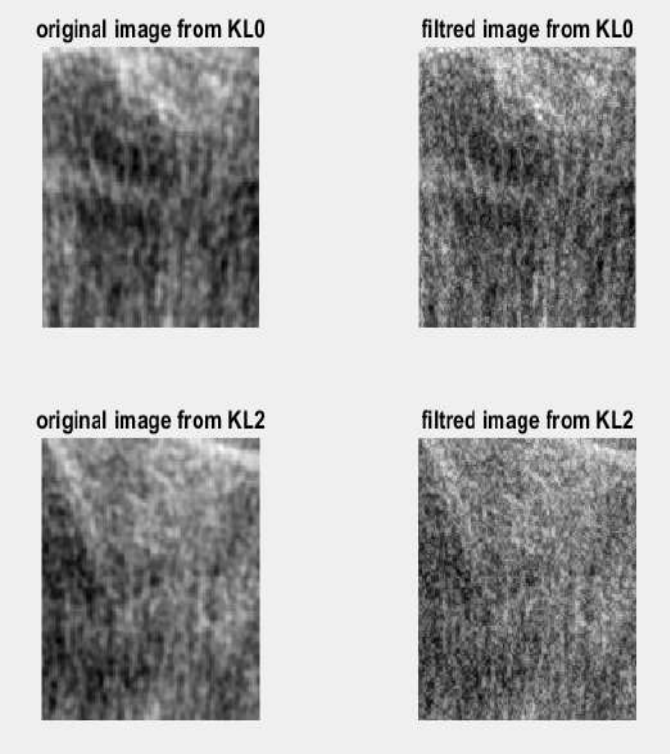


Fig. 3. The result of preprocessing

2) Data augmentation: To increase the size and diversity of the training dataset, we applied data augmentation techniques, including translation and rotation. The translation is performed by shifting the image horizontally and vertically, while rotation is performed by rotating the image around its center. This approach creates new training examples that are similar to the original images but with slightly different perspectives. This helps to improve the generalization ability of the ML model and reduce overfitting.

3) Features extraction: DL is a subset of ML that uses ANN to solve tasks in various application domains such as object detection, speech recognition, and image classification. Over time, several variants of ANN have been developed, such as Recurrent Neural Network (RNN), Long-Short-Term Memory (LSTM), and CNN. The latter is the subject of our study for a texture classification task in medical images revealing AO.

CNNs use a mathematical operation called convolution to extract relevant and deep features from images, then apply pooling to reduce the dimensionality of these features. These two techniques are organized into a layer system. Different arrangements and settings of these parameters and layers have given rise to several CNN variants. The different CNN variants that have won the ImageNet Large-Scale Visual Recognition

Challenge (ILSVRC) each in a specific year are AlexNet, ResNet, VGG, GoogLeNet, etc. GoogLeNet [11], which we will discuss in more detail, is the CNN variant we used in this article.

GoogLeNet [11], also known as Inception v1, is a deep convolutional neural network architecture developed by Google researchers in 2014. It was designed to improve the accuracy and efficiency of image classification tasks by reducing the number of parameters in the network while increasing its depth [12]. The main innovation of the GoogleNet architecture is the use of a module called the Inception module, which combines filters of different sizes (1x1, 3x3, and 5x5) in parallel to capture different levels of spatial information within an image. By using filters of different sizes, the Inception module can capture both fine-grained and coarse-grained features of an image. Another key feature of the GoogleNet architecture is the use of the auxiliary classifier, which is designed to mitigate the vanishing gradient problem that occurs in deep neural networks. The auxiliary classifier consists of a small classifier attached to an intermediate layer of the network, which encourages the network to learn more robust features. Overall, the GoogleNet architecture achieved state-of-the-art performance on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2014, and has since become a popular choice for image classification tasks. Its success has inspired further development of the Inception family of models, including Inception v2, Inception v3, and Inception-ResNet. Fig. 4, illustrates the architecture of the Inception model with several parallel convolutional layers of different filter sizes (1x1, 3x3, 5x5), allowing it to detect features at different scales. In addition, it also uses pooling convolutions to reduce the dimensionality of the data and thus reduce computational costs. The outputs of each branch are concatenated and passed to the next layer [13].

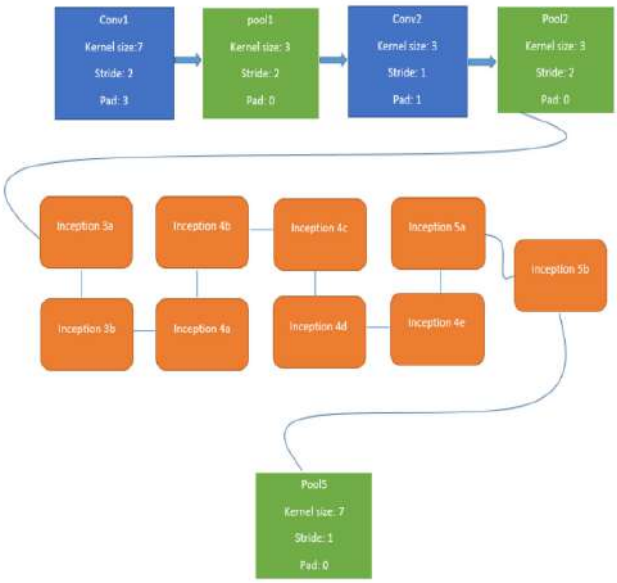


Fig. 4. A simplified architecture of the CNN used in GoogLeNet

GoogLeNet consists of 22 layers, including 9 Inception modules connected to the global average pooling layer. The final output of the network is a softmax layer, which produces the predicted class probabilities. The overall architecture of GoogLeNet is illustrated in Fig. 5.

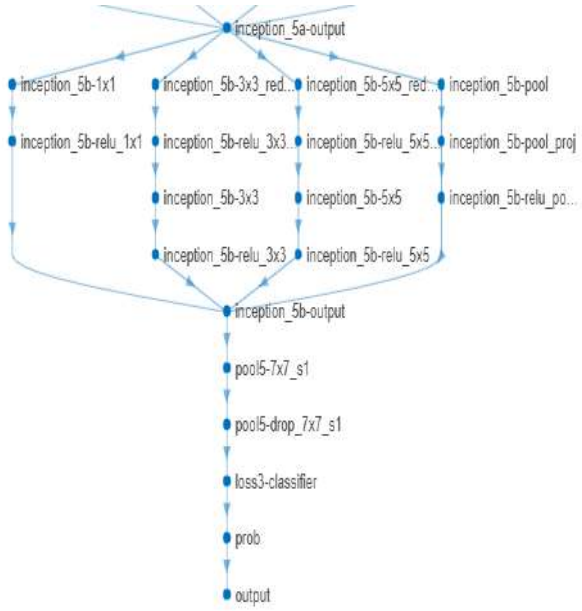


Fig. 5. Last layers of the GoogLeNet before finetuning

Fig. 6 shows the features extracted by GoogLeNet architecture.

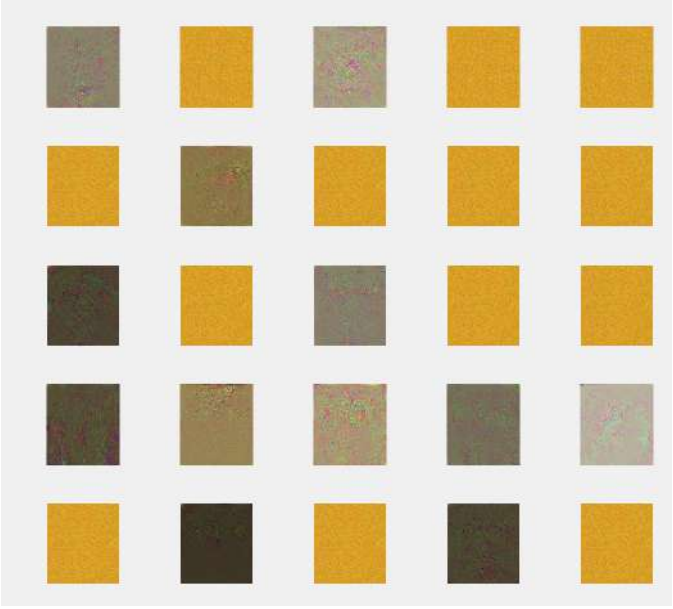


Fig. 6. Features of a Convolutional Neural Network

4) Features selection: To identify the most informative features for classification, we used the F-Score feature selection [14]. The F-score is a statistical method for selecting features that evaluate each feature individually and rank them

accordingly. The higher the F-score value, the more relevant the feature is considered to be [15]. Equation (5) explains the formula of the F-Score.

$$F\text{-}Score(f_i) = \frac{\sum_j (n_j/(c-1))(\mu_j - \mu)^2}{\frac{1}{n-c}\sum_j (n_j - 1)\sigma_j^2} \quad (5)$$

x represents the i feature, x denotes the total number of instances belonging to class j, and n represents the total number of instances across all classes [16]. The variable μ refers to the mean value of the features across all classes, while μ_j indicates the mean value of the features in the j class. The standard deviation of class j is denoted by σ.

5) Classification: We used an SVR for the classification of knee OA. SVR is a type of ML algorithm that is used to predict a continuous output variable. It is a variation of the SVM algorithm, which is commonly used to solve classification problems [17].

SVR's main idea is to find a hyperplane in a high-dimensional space with the greatest possible margin to the data points. The hyperplane is used in SVR to make predictions for new data points. SVR works by reducing the difference between predicted and actual values. A loss function is used to calculate the error, which penalizes the algorithm when it makes large errors. The epsilon-insensitive loss function is the most commonly used loss function in SVR [18].

In this study, we tested other classifiers and compared them to the proposed method. This includes decision Tree (DT), Support Vector Machine (SVM), and KNN.

III. Results and Discussion

The following section presents the key findings from the experiments conducted on high-resolution X-ray images, using various features and ML techniques outlined in the Data section. Specifically, GoogLeNet features were employed, and the experiments were carried out using DT, SVM, KNN, and the SVR model. Fig. 7illustrates the main outcomes. The results indicate that the SVR model utilizing GoogLeNet parameters achieved superior accuracy compared to other classification techniques, including DT, KNN, and SVM.

A confusion matrix is a table that is often used to describe the performance of a classification model. In this case, the rows represent the actual class of the data and the columns represent the predicted class. Fig. 8 shows the confusion matrix. Four different models combining deep learning and machine learning classifiers were used for the classification of knee images: GoogLeNet with SVR, GoogLeNet with KNN, GoogLeNet with DT, and GoogLeNet with SVM. According to the confusion matrix, the GoogLeNet model with the SVR model correctly predicted 44 positive samples (True Positives) and misclassified 5 negative samples as positive (False Positives). On the other hand, it incorrectly predicted 14 positive samples as negative (False Negatives) and correctly predicted 53 negative samples (True Negatives). Overall, this model seems to have performed well with a relatively high true positive rate compared to the false negative rate. GoogLeNet model with KNN correctly predicted 41 positive samples and misclassified 13 negative samples as positive. It also incorrectly predicted 17 positive samples as negative and correctly predicted 45 negative samples. Based on the confusion matrix, this model has a lower true positive rate and a higher false negative rate compared to the previous model.

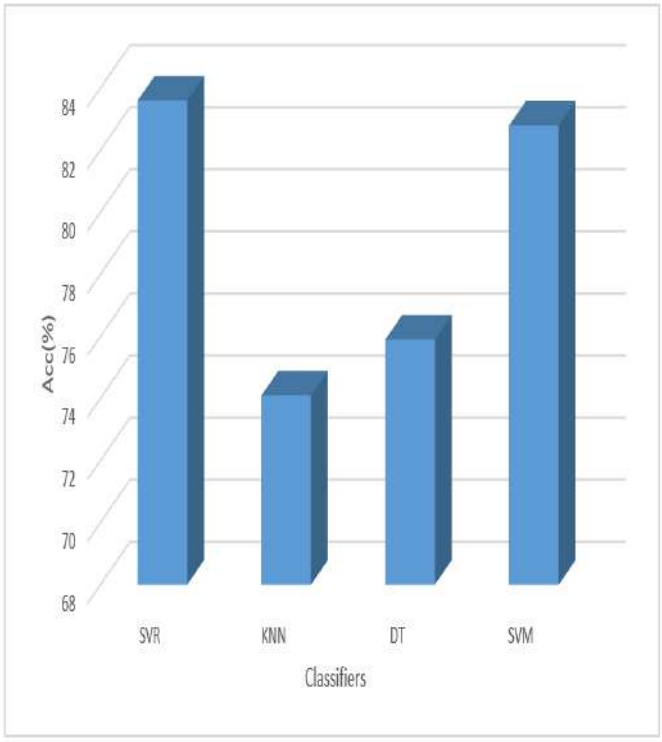


Fig. 7. Comparison of classification performance of the four different classifiers

The GoogLeNet model with DT had similar results to the KNN model. It correctly predicted 42 positive samples and misclassified 12 negative samples as positive. It also incorrectly predicted 16 positive samples as negative and correctly predicted 46 negative samples.

Lastly, the GoogLeNet model with SVM correctly predicted 43 positive samples and misclassified 5 negative samples as positive. It incorrectly predicted 15 positive samples as negative and correctly predicted 53 negative samples. This model appears to have a similar performance to the first model we discussed with a high true positive rate and a relatively low false negative rate.

Overall, it seems that the GoogLeNet model with SVR and the GoogLeNet model with SVM performed the best in terms of correctly predicting positive samples.

The results presented in Table II show the performance of different classification models for knee OA, with and without feature selection using an F-Score. The evaluation metric used is accuracy (Acc), which measures the proportion of correctly classified instances.

The results demonstrate that the feature selection process improves the accuracy of most models. For instance, the

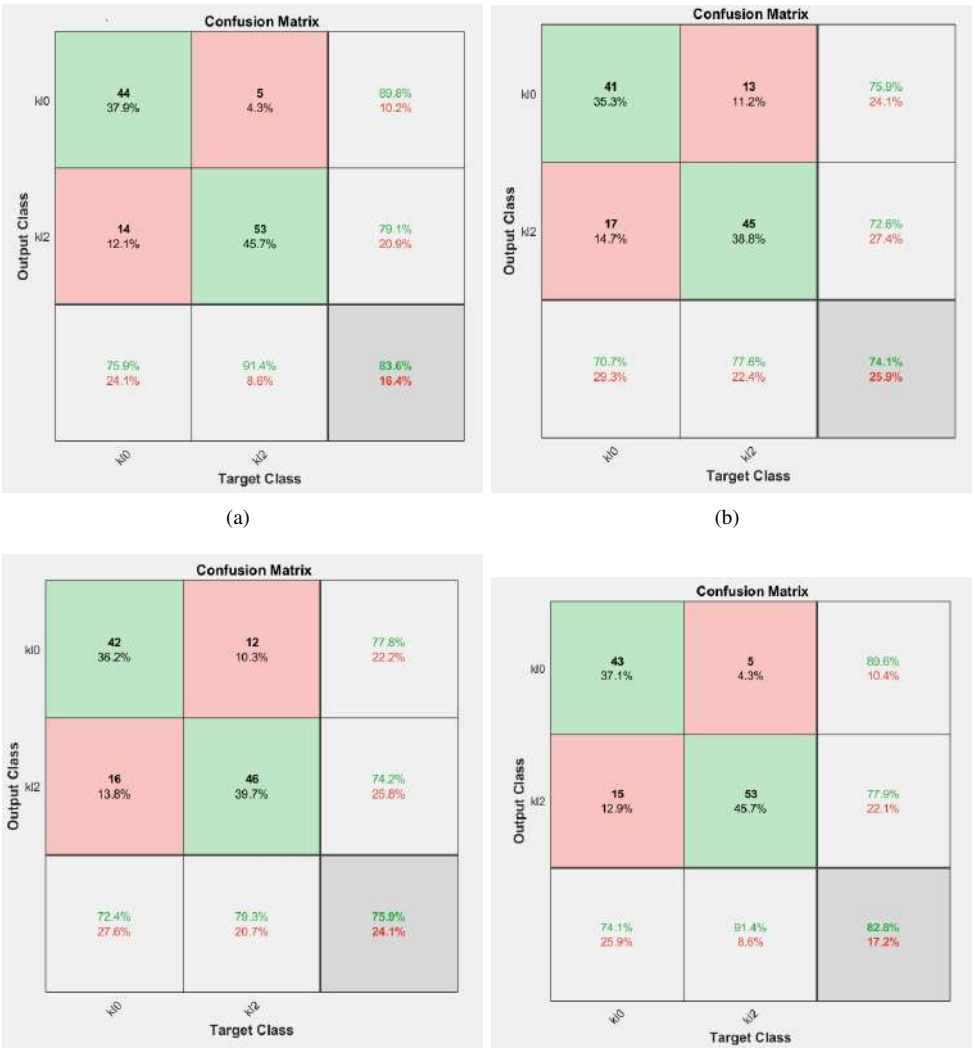


Fig. 8. Confusion matrix for KOA classification: (a) GoogLeNet+SVR, (b) GoogLeNet+KNN,(c) GoogLeNet+DT, and (d) GoogLeNet+SVM

TABLE II
COMPARISON OF CLASSIFICATION RESULTS WITH AND WITHOUT FEATURE SELECTION.

	With features selection	Without selection features
GoogLeNet+SVR	83.6	78.45
GoogLeNet+KNN	74.1	70.5
GoogLeNet+DT	75.9	71.3
GoogLeNet+SVM	82.8	76.65

GoogLeNet+SVR model achieved an accuracy of 83.6% with feature selection, compared to 78.45% without it. Similarly, the GoogLeNet+SVM model achieved an accuracy of 82.8% with feature selection and 76.65% without it.

However, it is worth noting that not all models benefit equally from feature selection. For instance, the GoogLeNet+KNN model only saw a small improvement in accuracy with feature selection (from 70.5% to 74.1%), while the GoogLeNet+DT model saw a more significant improvement (from 71.3% to 75.9%). The results suggest that feature selection using an F-Score can be a useful technique for improving the accuracy of classification models for knee OA. However, the effectiveness of feature selection may depend on the specific model used.

Comparing our results with the most recent methods(Table III), our proposed approach outperforms several other ML-based methods for knee OA detection. For example, the method proposed by Lim et al. achieved an accuracy of 71.97% [6], while our approach achieved an accuracy of 83.6%. Similarly, the method proposed by Jakaite et al [3]. achieved an accuracy of 77.5%. These results demonstrate that our proposed approach is highly effective for detecting knee OA, and represents a significant improvement over the existing methods. In conclusion, the proposed approach combines several innovative techniques and achieves high performance for knee OA diagnosis. The high accuracy of the proposed approach indicates that it can be used as an effective tool

TABLE III
COMPARISON OF CLASSIFICATION RESULTS WITH AND WITHOUT FEATURE SELECTION

Authors	Year	Features extraction method	Classifiers	Acc (%)
Lim et al. [6]	2019	DNN with PCA	/	71.97S
Jakaite et al. [3]	2021	Zernike moment	GMDH	77.5
Cuiva et al. [5]	2022	ResNet-34	/	61
Proposed method	2023	GoogLeNet	SVR	83.6

for the early detection and diagnosis of knee OA, which may contribute to improved patient outcomes and reduced healthcare costs.

IV. Conlusion

The present study proposes a hybrid system that combines deep feature-based methods and Machine Learning techniques for detecting knee osteoarthritis. The proposed approach comprises Gabor filter-based preprocessing, data augmentation, feature extraction utilizing the GoogLeNet model, feature selection via F-Score, and classification using SVR. Empirical findings reveal that the proposed approach surpasses other methods in detecting knee OA. This research presents a promising solution to the challenging task of diagnosing knee OA through the analysis of medical imaging data. Furthermore, it highlights the potential of ML techniques in the field of healthcare to enhance diagnosis and treatment outcomes. Further investigation may explore the feasibility of applying this approach to other stages of OA and imaging modalities, as well as its integration into clinical practice for personalized medicine.

References

[1] A. Cui, H. Li, D.Wang, J. Zhang, Y. Chen, H. Liu, "Global, regional prevalence, incidence, and risk factors of knee osteoarthritis in population-based studies", Clin Med, vol. 29, 100587, 2020.
[2] M. Chalian, F. Roemer, A. Guermazi, "Advances in osteoarthritis imaging", Curr. Opin. Rheumatol., vol. 35, no. 1, pp. 44-54, 2023.
[3] L. Jakaite, V. Schetinin, J. Hladůvka, S. Minaev, A. Ambia, W. Krzanowski, "Deep learning for early detection of pathological changes in X-ray bone microstructures: case of osteoarthritis", Sci. Rep., vol. 11, no. 1, pp. 1-9, 2021.
[4] Y. Wang, X. Wang, T. Gao, L. Du, W. Liu, "An Automatic Knee Osteoarthritis Diagnosis Method Based on Deep Learning: Data from the Osteoarthritis Initiative", J. Healthc Eng., pp. 1-10, 2021.
[5] J.H. Cueva, D. Castillo, H. Espinós-Morató, D. Durán, P. Díaz, V. Lakshminarayanan, "Detection and Classification of Knee Osteoarthritis", Diagnostics (Basel), vol. 12, no. 10, pp.2362, 2022.
[6] J. Lim, J. Kim, S. Cheon, "A Deep Neural Network-Based Method for Early Detection of Osteoarthritis Using Statistical Data", Int. J. Environ. Res. Public Health, vol. 16, no. 7, pp.1281, 2019.
[7] G. Lester, "The osteoarthritis initiative: an NIH public-private partnership", HSSJ, vol. 8, pp. 62-3, 2012.
[8] J. Han, K. Ma, "Rotation-invariant and scale-invariant Gabor features for texture image retrieval", Image Vis. Comput., vol. 25, pp. 1474–1481, 2007.
[9] B.S. Manjunath, W.Y. Ma, "Texture features for browsing and retrieval of image data", IEEE Trans. Pattern Anal. Mach. Intell. vol. 18, no. 8, pp. 837–842, 1996.
[10] H. Hadizadeh, "Multi-resolution local Gabor wavelets binary pat- terns for gray-scale texture description", Pattern Recognit. Lett., vol. 65, pp. 163–169, 2015.
[11] S. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going deeper with convolutions", In Proceedings of the IEEE conference on Computer Vision and pattern recognition 2015, pp. 1-9.
[12] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going Deeper with Convolutions", In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015, pp. 1–9.
[13] C. Tataru, D. Yi, A. Shenoyas, A. Ma, "Deep learning for abnormality detection in chest x-ray images", In IEEE Conference on deep learning, 2017.
[14] B. Manavalan, S. Basith, T.H. Shin, L. Wei, G. Lee, mAHTPred, "A sequence-based meta-predictor for improving the prediction of anti-hypertensive peptides using effective feature representation", Bioinform., vol. 35, no. 16, pp. 2757-2765, 2019.
[15] Y.W. Chen, C.J. Lin, "Combining SVMs with various feature selection strategies", NIPS 2003 feature selection challenge, pp. 1-10, 2003.
[16] L. Wei, S. Luan, L.A.E. Nagai, R. Su, Q. Zou, "Exploring sequence-based features for the improved prediction of DNA N4-methylcytosine sites in multiple species", BioinforM., vol. 35, pp. 1326–1333, 2019.
[17] A.J. Smola, B. Schölkopf, "A tutorial on support vector regression", Stat Comput, vol. 14, pp. 199–222, 2004.
[18] F. Zhang, L.J. O'Donnell, "Support vector regression", Machine learning, pp. 123-140, 2020.

Estimation of Anatomical and Detection System Parameters Effects on MUAP by Kurtosis

Noureddine MESSAOUDI
University of Boumerdes, Faculty of Technology, Department Engineering of Electrical Systems, LIST Laboratory, 35000 Boumerdes, Algeria

Samia Belkacem
University of Boumerdes, Faculty of Technology, Department Engineering of Electrical Systems, LIMOSE Laboratory, Faculty of Sciences 35000 Boumerdes, Algeria

Raïs El'hadi BEKKA
University of Sétif 1, Faculty of Technology, Department of Electronics, LIS Laboratory, 19000 Sétif, Algeria

***Abstract*— Electromyographic (EMG) signal can be used for biomedical applications. It is complicated in interpretation, so it acquires advanced methods for detection, decomposition, processing, and classification. It can be detected on the skin surface above the aimed muscle. The assessment of the effects of anatomical, physiological and detection system parameters on the shape of the detected surface EMG signal is more evident by using the signal generated by a single motor unit (MU). As the motor unit action potential (MUAP) has a non-Gaussian nature, the fourth order moment known as Kurtosis becomes a good indicator of the effects of anatomical and detection system parameters on the shape of the detected signal. The aim of this work was to evaluate the effects of the anatomical and the detection system parameters (the depth of the fibres within the muscle, thicknesses of the fat and skin layers, inter-electrode distance and the radius of the circular electrode) on the Kurtosis of the simulated surface MUAP signal generated in a multilayer cylindrical volume conductor contains the limb muscle. The MUAP is detected on the skin surface with a detection system which is constituted by height (four 1D and four 2D) spatial filters with a grid of nine circular electrodes.**

Key words: Kurtosis, muscle anatomy, non-Gaussian distribution, volume conductor, spatial filter.

I. Introduction

Surface EMG is used extensively to determine the muscle activation patterns of neuromuscular functions such as motor control, posture, and movement [1]. Surface motor unit action potential (MUAP) is the elementary component of surface EMG signal as the motor unit (MU) is the elementary component of the muscle. Modelling of surface MUAP is important for the interpretation of the experimental recordings of the electromyogram (EMG) [2]. Many generation models of the surface EMG signal have been proposed in the past [3], [4]. These models depend on volume conductor description. Analytical [3], [5] and numerical [6], [7] approaches have been proposed for the description of the electrical properties which separate the muscle fibres and the detection system. Cartesian [3], and cylindrical [4], [8], [9] coordinate systems have been adopted for its mathematical description (the Volume conductor). The main steps for modelling the surface MUAP are [3]: a) mathematical description of the volume conductor, b) modelling of the detection system and c) description of the generation, propagation and extinction of the intracellular action potential (IAP). In the 2D spatial filtering techniques approach, the volume conductor and the detection system are described by transfer functions in the two dimensional (2D) spatial frequency domains. According to the number of layers, the cylindrical volume conductor can be consisted of one (muscle only) [10], two (muscle and fat) [4], [8], three (muscle, fat and skin) [8], [9] or multilayer (bone, muscle, fat, skin and air) [4]. Many estimators of the amplitude and the spectral characteristics of the surface MUAP signal are used. The most commonly used estimators of amplitude features are the average rectified value (ARV) and the root mean square (RMS) and the spectral variables commonly used are the mean and the median frequency (MNF and MDF) [11]. Given the non-Gaussian nature of the EMG signal [12], it has been shown that the effects of anatomical [13] and detection system [14] parameters are better evaluated by the Kurtosis.

We analysed, in this article, the effects of the anatomical (the depth of fibres within the muscle, the thicknesses of the isotropic layers) and detection system (inter-electrode distance and the radius of the circular electrode) parameters on the kurtosis of the surface motor unit action potentials (MUAPs) signal generated in a multilayer cylindrical volume conductor contains the limb muscle and detected by height spatial filters. These investigated filters are the longitudinal and transversal single differential (LSD and TSD), the longitudinal and transversal double differential (LDD and TDD), the normal double differential (NDD) the inverse rectangle (IR) and the inverse binomial filter of order two (IB2) [15], [16] and the bi-transversal double differential (BiTDD) [17].

We show that increasing of the depth of the fibres within the muscle decreases the kurtosis of each signal. However, when the thicknesses of the isotropic layers (fat and skin) increase, the kurtosis corresponds to each signal keep almost the same value. The kurtosis of the TDD signal was the least sensitive to the variation of the anatomical parameters. The Kurtosis corresponds to the LSD, LDD, NDD and IB2 signals increased when the inter-electrode distance increase and decreases when the radius of the circular electrode increases.

II. Methods

A. Simulation of the surface MUAP

The surface MUAP was simulated using the analytical model describing the volume conductor as a cylindrical layered medium [4], [8], [9] and which was implemented as

described by Farina et al., [4]. The volume conductor which contains the limb muscle is composed of five layers that are bone, muscle, fat, skin and air as shown in Figure 1.

Parameters of simulation corresponding to the volume conductors are described in the table 1. These parameters are adapted to the studied volume conductor [4], [18], [19]. The simulated signals are obtained by taking into account the generation, propagation, and extinction of the intracellular action potential (IAP), at the neuromuscular junction, along the fibre and at the tendons respectively. These phenomena are described by progressive generation, propagation and extinction of the first derivative of the IAP [3]. The current density source is obtained from the second derivative of the IAP described analytically by Rosenfalck [20]. The source is internal (it is only under the fat and skin layers). It is propagated along the longitudinal direction (along z axis). Surface MUAPs signals were modelled. The conduction velocity was set to $3.8 m/s$. The neuromuscular junction was in the centre of the fibre. The fibre semi-length was $50\,mm$ ($L_1 = L_2 = 50\,mm$) [4].

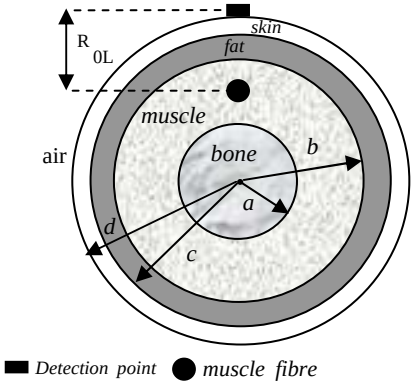


Figure. 1. Non-homogeneous and anisotropic multilayer cylindrical volume conductor describing the limb muscle. The volume conductor is constituted of bone, muscle, fat, skin and air layers. *a*, *b*, *c* and *d* represent the radial distances between bone-muscle, muscle-fat, fat-skin and skin-air interfaces, respectively. The muscle fibres are oriented along the *z* axe (longitudinal direction). The depth of the fibres within the muscle is represented by the distance between the fibre and the detection point [4].

Table I. Selected parameters of the volume conductor model to generate SFAP signal from the limb muscle [4].

Parameter	Description	Value
a	Radius of the bone	20 *mm*
b	Radius of bone + muscle compartment	45 *mm*
c	Radius of bone + muscle + fat	48 *mm*
d	Radius of the volume conductor	50 *mm*
σ_{mz}	Longitudinal conductivity of the muscle	0.5 *S / m*
$\sigma_{m\theta}$	Angular conductivity of the muscle	0.1 *S / m*
σ_b	Conductivity of the bone	0.02 *S / m*
σ_f	Conductivity of the fat	0.05 *S / m*
σ_s	Conductivity of the skin	1 *S / m*
R_{0L}	Distance between fibres and the detection point.	1 *mm*
v	Conduction velocity	3.8 *m / s*
θ_0	Angular distance between the source and the centre of the detection system	-
α	Angle of inclination of the fibre	-

B. The detection system

The detection system is modelled as a two-dimensional spatial filter rotated by an angle with respect to the fibres direction and it is constituted of 1D and 2D spatial filters according to the weights given to the electrodes. The investigated filters are the TSD, TDD, LSD, LDD, NDD, IR, IB2 and the BiTDD.

Figure 2 shows the electrodes arrangement with respect to the muscle fibres orientation and the filter weights for each spatial filter. Electrodes arrangement is adapted to the volume conductor shape [4].

The transverse direction of electrodes arrangement is along the angular direction. The coordinates of the detection point are $(z_0, \theta_0) = (0,0)$ [4], [18]. The steps forward to model the surface MUAP signal (the anatomical and the physiological parameters of the volume conductor such as the depth of the fibres within the muscle, thicknesses of the different layers, conductivities of the anisotropic and the isotropic layers and the detection system parameters such as the inter-electrodes distances, the angle of inclination of the fibres with respect to the centre of the detection system and the shape and dimensions of the detection electrodes) are fully described in the section II.

$$TSD = \begin{vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{vmatrix} \quad TDD = \begin{vmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{vmatrix}$$

$$LSD = \begin{vmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{vmatrix} \quad LDD = \begin{vmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{vmatrix}$$

$$IR = \begin{vmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{vmatrix} \quad BiTDD = \begin{vmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \end{vmatrix}$$

$$NDD = \begin{vmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{vmatrix} \quad IB^2 = \begin{vmatrix} 1 & 2 & 1 \\ 2 & -12 & 2 \\ 1 & 2 & 1 \end{vmatrix}$$

fibre direction ↑

Figure 2. A schematic representation of the spatial filter configurations on a nine-electrode grid arranged with respect to the muscle fibre direction (a). The abbreviations of these spatial filter configurations are: TSD: Transversal Single Differential; TDD: Transversal Double Differential; LSD: Longitudinal Single Differential; LDD: Longitudinal Double Differential; NDD: Normal Double Differential; IR: Inverse Rectangle; IB2: Inverse Binomial of order two ([15], [16]) and BiTDD: Bi-Transversal Double Differential ([5], [17]).

C. What is the Kurtosis?

To estimate the effects of the anatomical parameters on the shape of the surface MUAP signal generated from the limb muscle, we used the Kurtosis as an indicator of evaluation. The normalized fourth moment, known as, kurtosis of the desired signal is defined by [21]:

$$K_4(MUAP) = \frac{E(MUAP^4)}{\left(E[MUAP^2]\right)^2} - 3 \qquad (1)$$

Where, $MUAP$ represent the studied signal.

III. RESULTS

Figure 3 shows the effect of the depth of the fibres within the muscle on the kurtosis of the TSD, TDD, IR and BiTDD signals. When the depth of the fibres within the muscle increased, the kurtosis corresponds to the TSD, TDD and BiTDD signals decreased and the kurtosis of the IR signal increased. This implies that the TSD, TDD and BiTDD signals become broader with increasing of the depth of the fibres within the muscle and the IR signal becomes shorter.

Figure 4 shows the effect of the thickness of the fat layer on the kurtosis of the TSD, TDD, IR and BiTDD signals. When the thickness of the fat layer increased, the kurtosis of the TSD and the BiTDD signals decreased slightly and the kurtosis of the IR signal increased slightly but the kurtosis of the TDD signal remain constant. This result is due to the negligible effect of the thicknesses of the isotropic layers on the shape of the TDD signal.

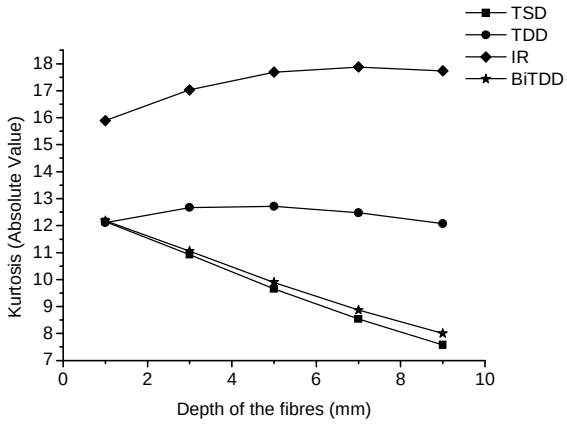


Figure 3. Influence of the depth of fibres within the muscle on the Kurtosis of the signals detected by TSD, TDD, IR and BiTDD filters. These results were simulated by using the parameters shown in the table I.

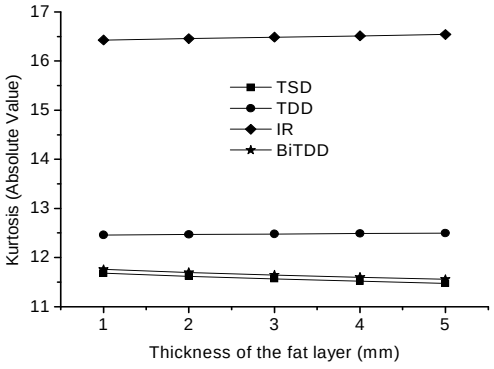


Figure 4. Effect of the thickness of the fat layer on the Kurtosis of the signals detected by TSD, TDD, IR and BiTDD filters. These results are simulated by using the parameters shown in the table I.

Figure 5 shows the effect of the thickness of the skin layer on the kurtosis of the signals detected by the investigated spatial filters and generated in the limb muscle. When the thickness of the skin layer increased, the kurtosis value of the TSD, TDD, IR and BiTDD signals remain almost constant. The only difference resides in the value of the kurtosis corresponds to each signal. This means that the kurtosis of the signals detected by the fourth investigated spatial filters is not sensitive to the variation of the thickness of the skin layer.

It has been indicated that amplitude and spectral variables of surface EMG signals are very sensitive to electrodes orientation [18], [22]. Figure 6 shows that increasing the inter-electrode distance increases the Kurtosis value corresponds to each signal.

The rate of increase of the Kurtosis in relation to the inter-electrode distance is different from an applied detection system to another. It is necessary to mention that when the inter-electrode distance increases, the kurtosis decreases until the inter-electrode distance equal *12 mm*, the kurtosis start with increases.

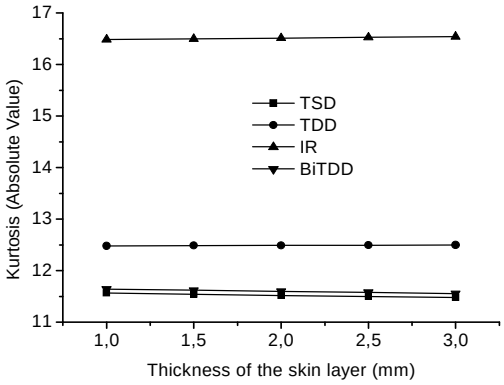


Figure 5. Effect of the thickness of the skin layer on the Kurtosis of the signals detected by TSD, TDD, IR and BiTDD filters. These results are simulated by using the parameters shown in the table I.

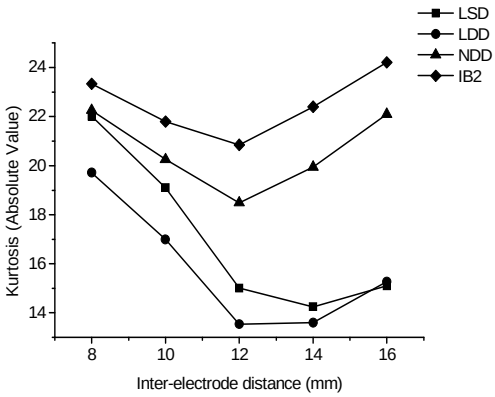


Figure.6. Effect of the inter-electrode distance on the Kurtosis corresponds to the LSD, LDD, NDD and IB2 signals. These results are simulated by using the parameters shown in the table I.

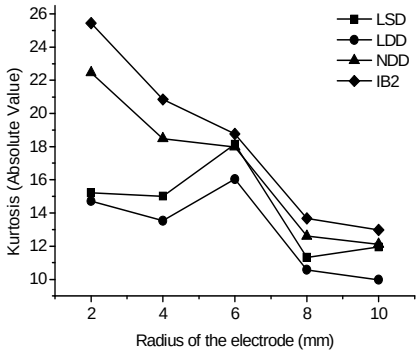


Figure.7. Effect of the radius of the circular electrode on the Kurtosis corresponds to the LSD, LDD, NDD and IB2 signals. These results are simulated by using the parameters shown in the table I.

Fig. 7 shows the effect of the radius of the circular electrode on the kurtosis of the signals detected by the LSD, LDD, NDD and IB2 spatial filters and generated from the limb muscle. When the radius of the circular electrode increased, the kurtosis value decreased. The rate of decrease of the kurtosis is different from one filter to another i.e. this decrease shows the effect of the dimension of the electrode on the selectivity of each spatial filter.

IV. Conclusion

The non-Gaussian nature of the generation and detection phenomenon of the EMG signals us to allow using parameters of random processes for analyze and interpret the effects of the parameters related to the anatomy and the physiology of the muscle, thus the detection system. Among the random parameters, the kurtosis is an important parameter.

We show that the effect of the depth of the fibres within the muscle on the kurtosis is more important than the effects of the thicknesses of the fat and the skin layers. When the thicknesses of the fat and skin layers increase, the kurtosis of the simulated signals remain almost constant, this is not the case when the depth of the fibres increases. The increase of the inter-electrode distance increases the kurtosis corresponds to each signal and the increase of the radius of the circular electrode decrease them.

References

[1] Östlund. N, Yu. J, Roeleveld. K and Karlsson. J. S, "Adaptive spatial filtering of multichannel surface electromyogram signals," Med. Biol. Eng. Comput., vol. 42, pp. 825-831, 2004.

[2] Mesin. L and Farina D, "Simulation of surface EMG signals generated by muscle tissues with inhomogeneity due to fiber pinnation," IEEE Trans. Biomed. Eng., vol. 51, pp. 1521-1529, 2004.

[3] Farina D, Merletti R. A novel approach for precise simulation of the EMG signal detected by surface electrodes. IEEE Trans Biomed Eng 2001; BME-48:637-46.

[4] Farina D, Mesin L, Martina S, Merletti R. Surface EMG generation model with multilayer cylindrical description of the volume conductor. IEEE Trans Biomed Eng 2004; 51:415-26.

[5] Dimitrov. G. V and Dimitrova. N. A, "Precise and fast calculation of the motor unit potentials detected by a point and rectangular plate electrode," Med. Eng. Phys., vol. 20, pp. 374-381, 1998.

[6] Lowery. M. M, Stoykov. N.S, Taflove. A, and Kuiken. T.A, "A multiple-layer finite-element model of the surface EMG signal," IEEE Trans. Biomed. Eng., vol. 49, pp. 446-454, 2002.

[7] Mesin. L, Joubert. M, Hanekom. T, Merletti. R, and Farina. D, "A Finite Element Model for Describing the Effect of Muscle Shortening on Surface EMG," IEEE Trans. Biomed. Eng., vol. 53, pp. 593-719, 2005.

[8] Gootzen THJM. Muscle fibre and motor unit action potentials. A biophysical basis for clinical electromyography. PhD. Dissertation, Univ Nijmegen, Nijmegen, the Netherlands, 1990.

[9] Blok JH, Stegeman DF, Van Oosterom. Three-layer volume conductor model and software package for applications in surface electromyography. Ann Biomed Eng 2002; 30:566-77.

[10] Merletti. R, Lo Conte. L, Avignone. E, and Guglielminotti. P, "Modelling of surface myoelectric signals – part I: model implementation," IEEE Trans. Biomed. Eng., vol. 46, pp. 810-820, 1999.

[11] Farina. D and R. Merletti, "Comparison of algorithms for estimation of EMG variables during voluntary isometric contractions," J. Electromyo. Kinesiol., vol. 10, pp. 337-349, 2000.

[12] Zazula D. Experience with surface EMG decomposition using higher-order cumulants. Signal processing 2001: 19-24.

[13] Messaoudi N, R.E BEKKA. Estimation of the Effects of the Anatomical Parameters on the MUAP Signal by Kurtosis. ICEEA 2nd-3rd November 2010 at Bejaia.

[14] Messaoudi N, R E BEKKA. Use kurtosis for the Assessment of the Effects of the Detection System Parameters on the MUAP Signal. ISPA 6th-8th December 2010 at Biskra.

[15] Reucher H, Silny J, Rau G. Spatial filtering of non-invasive multi-electrode EMG: Part II-Filter performance in theory and modelling. IEEE Trans Biomed Eng 1987; BME- 34:106-13.

[16] Farina D, Shulte E, Merletti R, Rau G, Disselthorst-Klug C. Single motor unit analysis from spatially filtered surface electromyogram signals. Part I: Spatial Selectivity. Med Biol Eng Comput 2003; 41:330-37.

[17] Disselhorst-Klug C, Silny J, Rau G. Improvement of spatial resolution in surface-EMG: A theoretical and experimental comparison of different spatial filters. IEEE Trans Biomed Eng 1997; BME-44: 567-74.

[18] Farina D, Cescon C, Merletti R. Influence of anatomical, physical and detection system parameters on surface EMG. Biol Cybern 2002; 86:445-56.

[19] Merletti R, Bottin A, Cescon C, Farina D, Gazzoni M, Martina S, Mesin L, Pozzo M, Rainoldi A, Enck P. Multichannel surface EMG for the invasive assessment of the anal sphincter muscle. Oasis Progress Report Digestion 2004; 69:112-22.

[20] Rosenfalck P. Intra and extracellular filelds of active nerve and muscle finres. Acta Physiol Scand 1969; 32:1-49.

[21] Zhi-Lin Zhang_, Zhang Yi. Extraction of a source signal whose kurtosis value lies in a specific range. Neurocomputing 2006 ; 69 : 900–904.

[22] Messaoudi N, Bekka RE, Belkacem S. Effects of detection system parameters on cross-correlations between MUAPs generated from parallel and inclined muscle fibres. Polish Journal of Medical Physics and Engineering 2021; 27(1): 87–97.

A Survey on Software Defined 5G Internet of Vehicles networks

Nadjet Azzaoui
Department of Computer Science and Information Technology, Kasdi Merbah University,
Ouargla, Algeria.
azzaoui.nadjet@univ-ouargla.dz

Ahmed Amine Fetni *, Fatima Djafour †
Computer Science Department and Information Technology, Kasdi Merbah University,
Ouargla, Algeria.
*ahmed.amine.fetni@gmail.com
†fatmadjafour20@gmail.com

Laid Kahloul
Computer Science Department LINFI Laboratory
Biskra, Algeria
l.kahloul@univ-biskra.dz

***Abstract*—The rapid increase in the number of vehicles and road congestion led to heterogeneous, dynamic, and large-scale vehicular networks, which are making it hard to meet the strict requirements, such as low latency, high mobility, and huge connectivity of 5G networks. Namely, 5G Vehicular Communication Networks are a potent instrument that makes a variety of vehicle data services and applications possible. With the advent of 5G, operators will be able to take advantage of newly available infrastructure capabilities. Traffic control, accident avoidance, ease of travel, and convenience are supported by an infrastructure that can offer a wide range of capabilities and is supported by Software Defined Networking (SDN) for vehicle networks. Low latency, more bandwidth, and more prolonged error-free connections can be achieved by utilizing 5G-IoV networks built on SDN. SDN keeps a positive novelty in network programmability, allowing for more intelligent network management. In this essay, we give a summary of earlier wireless communication-based works. First, a brief explanation of three key ideas: the IoV, 5G, and SDN. The second section offers a review of different SDN-5G vehicular network applications, with a focus on analyzing and contrasting Software-defined Network (SDN) operations across many parameters. The difficulties and requirements of SDN-based 5G Internet of Vehicle networks will also be covered. This survey serves as a catalyst for improving the emerging robustness, connectivity, and latency of SDN-IoV architectures in the future.**

***Index Terms*—SDN, 5G, IoV, Vanet, SDN-5G vehicular networks.**

I. Introduction

It is challenging to meet the demanding requirements of 5G networks, such as low latency, high mobility, and massive connections, due to the heterogeneous, dynamic, and large-scale Internet of Vehicles that have emerged as a result of the rapid growth in vehicle numbers and traffic congestion ([1]). In contrast to conventional vehicular networks (VANets [2]), the Internet of Vehicles (IoV) or Connected Vehicles (CV) offers drivers services that increase traffic safety and road comfort. It does this by combining Intelligent Transportation Systems (ITS) and the Internet of Things (IoT). This new concept supports data exchange between vehicles and everything (V2X), including five communication types: vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), vehicle-to-roadside unit (V2R), and vehicle-to-sensor (V2S), vehicle-to-personal device (V2P), via various communication technologies (e.g. 3G/4G/Long Term Evolution (LTE)-A/5G)) [3]. A powerful instrument for allowing multiple vehicular data services and applications is the 5G vehicular communication network. With the advent of 5G, operators will be able to benefit from upgraded infrastructure thanks to new features and technologies. Its capacity to provide the infrastructure for a variety of facilities that rely on software-defined networking (SDN) based on vehicle networks for traffic control and accident avoidance is one of its distinguishing features. [4]. The development of software-defined networking has had an impact on how new vehicular network architectures are designed, as well as how V2X services are implemented in the next intelligent transportation systems. Low latency, better bandwidth, and longer error-free connections can be provided via SDN-based 5G-IoV networks, enabling network management with superior intelligence. [5].

In this paper, we provide an overview of previous works based on wireless communication. First, a brief summary of current research and technologies: IoV, 5G, and SDN. Second, some SDN based 5G vehicular network applications are outlined, focusing on examining and comparing the works of software-defined networking on several parameters. In addition, the challenges and requirements of SDN-based 5G IoV networks are discussed. This survey serves as a catalyst for improving latency, and connectivity, and the robustness of routing protocols for future SDN-IoV architectures.

The remainder of the paper is organized as follows: the current research and technologies based on wireless communication are detailed in Section 2. A comparative study of previous works of SDN based are discussed in Section 3. Analysis and discussion about SDN-based 5G-IoV is detailed in Section 4. Finally, we conclude our work in Section 5.

II. Background and overview

In this section, we provide a brief background and overview of IoV (in Section II-A) and SDN (in Section II-B) technologies, 5G (section II-C) and SDN based 5G-IoV (in Section II-D) as follows.

A. *Internet of Vehicles networks*

In order to enhance traffic safety, provide services for drivers, and increase comfort on the road, researchers have proposed to combine ITS and the Internet of Things leading to what we call Internet of Vehicles networks. IoV is a new concept, and it differs from the well-known concept of Vehicular Ad-Hoc Networks. Indeed, IoV offers facilities allowing vehicles to connect to any object, hence resolving problems that occur on the road, such as accidents, crimes, and any health aspects (road traffic deaths, road deaths, and injuries). Thus, IoV aims basically to ensure more safety and so help drivers reach their destination by providing different entertainment services during the journey. On another hand, the Internet of Vehicles is considered a dual platform that combines network and intelligence. In a such platform, vehicles can use several communication technologies (such as 5G, 4G/LTE, Wave, etc.) to ensure connections between them and their environment. Based on that communication, six new communication modes are developed V2V, V2I, V2P, V2S, V2R, and V2C. In fact, the design of IoV encourages the use of cloud computing and fog computing capabilities to enhance different computing tasks such as processing and storage [6]. (See Fig. 2)

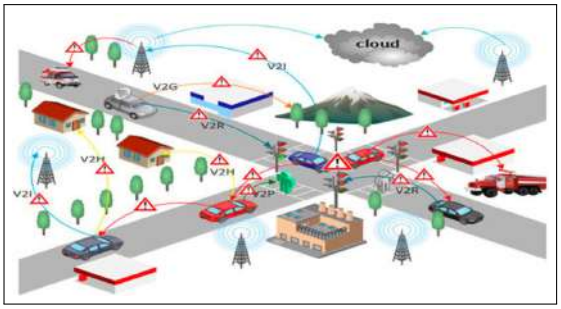


Fig. 1: V2X Communications in IoV.

B. *Software Defined Networks (SDN)*

One of the newest networking approaches, software-defined networks (SDN), is distinguished by the separation of the control plane and data plane. A controller, a logically centralized entity, is employed in the control plane to keep an eye on and control network resources. By maximizing the use of network resources, controllers are intended to enhance overall network performance (i.e., effective communication and flow control). The network infrastructure used for data forwarding, known as the data plane, consists of forwarding hardware and wired or wireless communication channels. Using standardized interfaces like OpenFlow, SDN enables devices from various manufacturers to communicate with one another. This provides programmable and flexible network architecture design and facilitates network monitoring and administration. [7]. (See fig. 2)

SDN makes the implementation of SDN application network services easier by abstracting the data plane from applications and enabling them to perform dynamic requests through logic

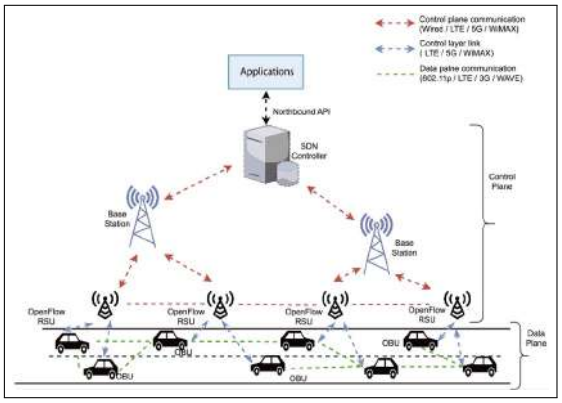


Fig. 2: SDN based 5G-IoV networks.

Data plane entity for the centralized controller. SDN greatly simplifies network management by using global network information to perform efficient resource utilization. Although SDN has many benefits, its fundamental features—such as SDN-based programmable switches, constrained southbound channel bandwidth, and constrained SDN controller resources—also provide new security risks. Following, we provide a brief summary of the most significant benefits and drawbacks of employing SDN technologies. [8]:

- Support heterogeneity and improve resource utilization.
- Routing and Spectrum Management Optimization.
- Improved network security.
- Bad information travels slowly.
- Energy management in 5g in Vehicular networks.

C. *The fifth generation (5G)*

One of the most crucial topics of discussion in the area of cellular mobile communications is the fifth generation (5G) network, a promising candidate technology for the next vehicular networks. With its remarkable connection speeds, 5G communications are modernizing networks and increasing their market penetration. High-speed connectivity, up to 10 Gbps, low latency or even less than 1 ms, and expanded coverage are the main benefits of the 5G network [9].

D. *SDN based 5G-IoV*

The most recent wireless communication system to be created is 5G or fifth-generation wireless technology. It was developed with high data rates in mind (up to 20 Gbps). The technology ensures 1ms latency for real-time applications while bringing new breakthroughs including vehicle-to-vehicle (V2V) connectivity and SDN. 5G will provide improved bandwidth, lower latency, connectivity to larger devices, higher data rates, and dependable Quality of Experience (QoE) delivery thanks to these technological developments. Together with increasing bandwidth and lowering latency, network optimization may also be a crucial part of the 5G infrastructure. Moreover, Vehicular Ad-Hoc Networks' issue with supporting a large number of nodes is resolved by 5G. Furthermore, because of their interdependence, mobile network processes,

and technology cannot be addressed separately. The objective is to concentrate on a connection model that incorporates existing technologies, complements them, and meets a variety of application requirements in a flexible, trustworthy, and heterogeneous way. In this case, 5G is a compelling choice for a variety of circumstances. VANET uses it, so it is not an exception. It is possible to support numerous simultaneous connections in VANET, often known as the "internet of vehicles networks". 5G could offer flexible security advantages from a security standpoint. SDN control and multimodal security based on 5G are two well-known advancements that are essential. Thus, 5G facilitates data protection above the user level and allows customization of security parameters [10].

III. LITERATURE REVIEW

Santosh et al. [11] suggested a software-defined network-based load-balancing strategy for IoT which uses an integrated whale optimization algorithm to reduce latency and IoT tasks through cloud and edge computing devices. To implement IoV, the authors in [12] suggested a 5G-V2X ecosystem. The idea of "Software Defined Networking" serves as the foundation for the proposed ecosystem (SDN). If you think of consumer entertainment venues as being in automobiles, the network infrastructure needs to be quite large to guarantee delivery and quality. To that purpose, they assess vehicle-to-vehicle (V2V) communications in urban and rural settings, as well as Internet video service traffic. The authors of [13] developed a seamless delivery system in citing 13 that uses mobile computing (MEC) and software-defined networking (SDN) technologies to react to dynamic structure change in VANETs. When a changeover occurs, the installation of SDN enables a stable transport layer connection, preventing network changes from affecting the performance of the upper layer. It also gives an overview of the network topology and centralizes control. The data is cached before delivery through the usage of the MEC server, allowing the vehicle to quickly resume the normal connection. In [14], a layered SDN method and a vehicle-based Fog computing architecture were combined to create the infrastructure model known as VISAGE for upcoming 5G VANET systems. Vehicle QoS becomes more effective in terms of computation time and communication latency by using automobiles as fog infrastructure and integrating them with regional SDN controllers. uses the network edge to offload computing services from the cloud, ample vehicle resources, and local network control decisions to explain this. In [15], Congestion Free Path (CFP) and Optimized CFP (OCFP), two heuristic algorithms, were suggested for the SD-VANET architecture. The suggested approach reduces end-to-end latency for vehicles in the VANET while simultaneously resolving the issue of traffic congestion. In [16], by utilizing a software-defined hierarchical network, the authors created a novel IoV-Fog network environment (SDN). Also, they presented a routing protocol based on fog and SDN dubbed Delay-Efficient Routing Method for IoV. (DRIFS). SDN Orchestra uses DRIFS to create powerful tracks. It also considers the bandwidth that is still available, as well as the position and speed of the cars.

In order to lower the task/message failure rate, the authors of [17] offered heuristic algorithms and deployed a fault tolerance system. They suggested a software-defined v-car network with cloud computing that is QoS-aware and fault-tolerant (QAFT-SDVN). Via SDN nodes mounted atop fog nodes, the proposed model receives messages from vehicles. While receiving messages from neighboring SDN entities, the SDN console gives them two different levels of priority. The message type is one, while the date and size method of message priority is another. Messages are divided into secure and unsecured categories by the SDN console before being forwarded to their destination. They check their acknowledgment after the message has been sent to the intended recipient; if the recipient has received it, nothing is done; otherwise, they employ a fault tolerance mechanism. For IoV networks, the authors of [18] presented a hierarchical architecture built on software-defined networks (SDN) and fog computations. IoV vehicles, a cloud computing framework, semi-centralized SDN controllers, and centralized SDN control planes make up this system. Also, they suggested a routing method for the Internet of Things (IoT) dubbed the Delay Active Routing Method based on SDN and fog computing (DRSFI). DRSFI is used by SDN controllers to determine the least delay trajectory while taking the bandwidth constraints, vehicle position, and speed into consideration. To manage the enormous data traffic and prevent transmission storm issues with fast packet delivery, a Software Defined Internet of Vehicles (SD-IoV) system was created in [19]. The suggested broadcast targeting method effectively broadcasts notifications by using selective transmission and the identification of surrounding vehicles, avoiding traffic congestion and shortening travel times. The SD-IoV system's transmission algorithm is launched as soon as the on-board unit (OBU) in the cars recognizes the event. The authors of [20] introduced an innovative fifth-generation SDN (SDN)-based architecture for intelligent transportation systems that will enhance ITS capabilities (5G). The sensor layer, the relay layer, and the core network layer are the three functional levels that make up the proposed system architecture. Continuous access is made possible with flexible and adaptable functionality because of the SDN capabilities. The suggested 5G design also provides high data rates and throughput. In [21], it was recommended to use a unique hierarchical 5G VANET design. This end-to-end architecture's main objective is to integrate 5G connectivity technologies with the centralization and robustness of SDN and C RAN for efficient resource allocation from a wide angle. Also, it has been recommended that in order to avoid often switching between cars and MSWs, a cyber fog structure composed of zones and clusters be used. The load control, throughput, and transmission latency of the controller are analyzed and compared to those of rival architectures.

IV. ANALYSIS AND DISCUSSION

With the development of Intelligent Transportation Systems, vehicular networks have to face rigorous performance

requirements in the future. 5G mobile communications, cloud computing, and SDN technologies provide solutions for future vehicular networks. In this survey, we presented a summary of recent research studies on SDN-IoV based 5G, and suggested architectures, characteristics, and challenges of SDN based 5G-IoV are included in this article. Table I illustrates the summarizing of the reviewed related works.

Although a number of architectural designs have been put out in the literature to enhance communication security and reliability in vehicular networks, a thorough analysis to determine the viability and accuracy of these designs is still lacking. It is important to pay close attention to the new security and privacy vulnerabilities that result from the integration of new technologies with the current IoV networks. For instance, academics should not only analyze how SDN may be used to enhance the design of vehicular networks but also look into and address the new problems that SDN has introduced, such as service latency, mobility, and SDN controller security. Furthermore, it is really critical to consider how dynamic real-time change, scalability, and integration of service context will be crucial in enabling effective deployment and preventing performance visibility gaps in SDVNs. Vehicular network architecture will be continually changing in the near future to meet the quickly expanding needs of its new applications.

V. CONCLUSION

The traditional network model is not appropriate for the vehicular network due to its extremely dynamic nature and the requirement for effective real-time communication. IoV is considered as a crucial approach that can be used to achieve a variety of services, including traffic control, passenger safety, and travel convenience and comfort. As part of the developing 5G mixed with SDN, IoV is now being proposed as the major 5G enabler. This article has reviewed and summarized a number of articles based on the problems they sought to address, the papers purposes, and the outcomes they were able to produce. The use of SDN-based 5G networks provides lower latency, higher frequency bandwidth, and connectivity. Longer, more users should be able to complete more tasks in less time to solve congestion problems, the purpose of this document is to understand the importance of a certain network through 5G programs for traffic maintenance and passenger safety.

REFERENCES

[1] F. Adelantado, M. Ammouriova, E. Herrera, A. A. Juan, S. S. Shinde, and D. Tarchi, "Internet of vehicles and real-time optimization algorithms: Concepts for vehicle networking in smart cities," *Vehicles*, vol. 4, no. 4, pp. 1223–1245, 2022.

[2] E. Benalia, S. Bitam, and A. Mellouk, "Data dissemination for internet of vehicle based on 5g communications: A survey," *Transactions on Emerging Telecommunications Technologies*, vol. 31, no. 5, p. e3881, 2020.

[3] J. Wang, Y. Shao, Y. Ge, and R. Yu, "A survey of vehicle to everything (v2x) testing," *Sensors*, vol. 19, no. 2, p. 334, 2019.

[4] Y. Yang and K. Hua, "Emerging technologies for 5g-enabled vehicular networks," *IEEE Access*, vol. 7, pp. 181 117–181 141, 2019.

[5] W. B. Jaballah, M. Conti, and C. Lal, "Security and design requirements for software-defined vanets," *Computer Networks*, vol. 169, p. 107099, 2020.

[6] T. Adbeb, D. Wu, and M. Ibrar, "Software-defined networking (sdn) based vanet architecture: Mitigation of traffic congestion," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 3, 2020.

[7] W. Jaballah, M. Conti, and C. Lal, "Software-defined vanets: Benefits, challenges, and future directions," *IEEE Commun*, pp. 1–17.

[8] J. Bhatia, Y. Modi, S. Tanwar, and M. Bhavsar, "Software defined vehicular networks: A comprehensive review," *International Journal of Communication Systems*, vol. 32, no. 12, p. e4005, 2019.

[9] N. Azzaoui, A. Korichi, B. Brik, and H. Amirat, "A survey on data dissemination in internet of vehicles networks," *Journal of Location Based Services*, pp. 1–44, 2022.

[10] S. Askar, "Sdn based 5g vanet: a review," *Available at SSRN 3963000*, 2021.

[11] S. A. Darade, M. Akkalakshmi, and D. N. Pagar, "Sdn based load balancing technique in internet of vehicle using integrated whale optimization method," in *AIP Conference Proceedings*, vol. 2469, no. 1. AIP Publishing LLC, 2022, p. 020006.

[12] C. R. Storck and F. Duarte-Figueiredo, "A 5g v2x ecosystem providing internet of vehicles," *Sensors*, vol. 19, no. 3, p. 550, 2019.

[13] R. Duo, C. Wu, T. Yoshinaga, J. Zhang, and Y. Ji, "Sdn-based handover scheme in cellular/ieee 802.11 p hybrid vehicular networks," *Sensors*, vol. 20, no. 4, p. 1082, 2020.

[14] A. Soua and S. Tohme, "Multi-level sdn with vehicles as fog computing infrastructures: A new integrated architecture for 5g-vanets," in *2018 21st Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN)*. IEEE, 2018, pp. 1–8.

[15] T. Adbeb, D. Wu, and M. Ibrar, "Software-defined networking (sdn) based vanet architecture: Mitigation of traffic congestion," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 3, 2020.

[16] A. Kadhim and J. Naser, "Routing protocol for iov-fog network supported by sdn," *Telecommunications and Radio Engineering*, vol. 79, no. 5, 2020.

[17] S. A. Syed, M. Rashid, S. Hussain, F. Azim, H. Zahid, A. Umer, A. Waheed, M. Zareei, and C. Vargas-Rosales, "Qos aware and fault tolerance based software-defined vehicular networks using cloud-fog computing," *Sensors*, vol. 22, no. 1, p. 401, 2022.

[18] A. J. Kadhim, S. A. H. Seno, and R. A. Shihab, "Routing strategy for internet of vehicles based on hierarchical sdn and fog computing," *Journal of University of Babylon for Pure and Applied Sciences*, vol. 26, no. 10, pp. 309–319, 2018.

[19] G. Raja, P. Dhanasekaran, S. Anbalagan, A. Ganapathisubramaniyan, and A. K. Bashir, "Sdn-enabled traffic alert system for iov in smart cities," in *IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2020, pp. 1093–1098.

[20] S. Din, A. Paul, and A. Rehman, "5g-enabled hierarchical architecture for software-defined intelligent transportation system," *Computer Networks*, vol. 150, pp. 81–89, 2019.

[21] A. A. Khan, M. Abolhasan, and W. Ni, "5g next generation vanets using sdn and fog computing framework," in *2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 2018, pp. 1–6.

TABLE I: Summarizing of the Reviewed Studies.

reference	year	proposition	goal	tools	scenario	environment
[11]	2022	software-defined network-based load balancing strategy	load balancing and minimization of latency	MATLAB	highway	IoV
[12]	2018	5G V2X ecosystem to provide IoV	Analyze and compare data throughput, transmission latency, and packet delivery rate (PDR)	NS-3	Urban	IoV
[13]	2020	Transparent deployment plan based on Software Defined Networking (SDN) and Mobile Computing (MEC) technologies	Reduce data loss during transmission and shorten the time it takes to restore normal network connections	Omnet++, INET, SUMO	Highway	Vanet
[14]	2018	A Multi-level SDN architecture for fog enabled vehicular networks, called VISAGE	The improvement in communication and computational capabilities that can be achieved by the VISAGE architecture	/	Highway	Vanet
[15]	2020	Heuristic algorithms called Congestion-Free Path (CFP) and Optimize CFP (OCFP), in SD-VANET architecture	To determine the amount of traffic on each road, the SDN controller in the VANET gathers data on the vehicles on each route	SUMO/NS3 OpenStreetMap	Urban, Highway	Vanet
[16]	2020	Routing protocol named Delay-Efficient Routing approach for IoV depending on Fog and SDN (DRIFS)	Calculate the best path based on the proposed mathematical model.	OMNeT++, SUMO	Highway	IoV
[17]	2022	QoS aware and fault tolerance-based software-defined vehicular networks using Cloud-fog computing (QAFT-SDVN)	decreased by up to 4% the execution times of the safety and non-safety messages	/	Highway	Vanet
[18]	2018	Delay-Efficient Routing strategy based on SDN and Fog computing for IoV	Calculation of the optimal route for the vehicle, driving reactions with the lowest latency and high network bandwidth	OMNeT++, SUMO	Highway	IoV
[19]	2020	Broadcast routing mechanism to selectively hand over and detect vehicles by neighbors to effectively send alerts	Avoiding congestion and reducing travel time, High packet delivery ratio with minimal delay	OMNeT++, SUMO	Urban	IoV
[20]	2018	5G enabled SDN for ITS	Provides the flexibility, improve overall network performance	Simulation using C language and Hadoop ecosystem	Urban	Vanet
[21]	2018	A fog computing framework (comprising of zones and clusters)	Combine 5G communication technologies with the centralization and flexibility of Software Defined Networking (SDN) and Cloud-RAN (CRAN)	/	Highway	Vanet

Fault Diagnosis for PV Systems: From the Classic Analysis Methods to the Artificial Intelligence Ones

Dr. Ismahan MAHDI
LREEI University Mhamed Bougara Boumerdes Algeria
Department of Electrical System Engineering, Faculty of Technology, University Mhamed Bougara
Boumerdes, Algeria
ismahane.mahdi@univ-boumerdes.dz

Dr. Yasmine GUERBAI
Department of Electrical System Engineering, Faculty of Technology, University Mhamed Bougara
Boumerdes, Algeria
y.guerbai@univ-boumerdes.dz

Pr. Bouchra NADJI
LREEI University Mhamed Bougara Boumerdes Algeria
b.nadji@univ-boumerdes.dz

***Abstract*—In order to respond to the ever-growing energy needs, other renewable and sustainable sources, such as solar energy, had to be chosen. Since Algeria has a large solar reservoir, given its geographical location, it must be exploited effectively. In the case of a photovoltaic (PV) system, the presence of a failure may affect the operation of the system, and therefore the effect is either a degradation in the efficiency or the total shutdown of the system. The reliability study of such systems is little or almost absent, due to the lack of software tools and simulators, or the lack of necessary information and data. Over the recent years, many fault diagnosis techniques and methods have been developed, that should be able to detect, identify and locate the position of the fault. For a small-size PV system, it is relatively easy to diagnose the system using manual or semi-automatic methods. However, the localization of faults in large-size PV plants is quite a challenging issue. Our work is divided into two parts: the first part presents the results found by using the Stochastic Petri Nets (SPN); and the second part sheds light on the methods of artificial intelligence (AI) used in the same theme, in order to bridge the boundaries of the classical methods used in the first part. It is therefore a contribution to the safety study of a typical PV system. It is based on the estimation of reliability, availability, and durability over a period of twenty-five years, in three different configurations and taking into account failure modes that have high criticality on the overall operation of the studied system. Through the simulations, we carried out using the SPN, the results obtained showed the impact of the arrangement of the PV modules on the reliability, availability, and durability of the system. The availability and the value of the mean time between failures (MTBF) increase in the case where the PV modules are arranged in series, as well as the output power of the panel whose value increases and will be close to the value of the Maximum Power Point Tracking (MPPT). The application of AI techniques in PV systems was shown the ability of such techniques to solve some problems such as sizing and optimization, output power forecasting, control, and MPPT. Nowadays, AI-based methods for fault detection, identification, and localization in PV systems have attracted many researchers, due to the promising results of AI in this field. The main objective of this second part is to present a short review of the recent application of AI techniques, including machine learning (ML) and recently deep learning (DL) in fault detection, identification, and fault classification of PV systems.**

Keywords— Fault diagnosis, Stochastic Petri Nets, Reliability, MTBF, Artificial intelligence techniques, Neural Networks, Machine learning, Deep learning

I. INTRODUCTION

As mentioned in [1], the introduced PV limit all over the planet toward the end of 2021 was around 800 GW. A similar source [1] demonstrated that all of the PV frameworks introduced all through the world are presently ready to cover around 20% of worldwide power request. PV plants are dependent upon various sorts of failures (in the PV modules, in the wiring, inverter, and security hardware), their yield can drop fundamentally, particularly in desert areas under extreme conditions. Many fault diagnosis techniques and methods have been developed by researcher [2]. They are used to detect, identify and locate the fault which causes the failure. For a PV system used in domestic applications, it is mostly easy to make a diagnostic using simple methods. However, in large-size PV plants is quite challenging issue.

The application of the AI techniques in PV systems was given the ability to solve some problems such as: output power optimization, control, and MPPT [3]. These methods have attracted many researchers, due to the promising results of AI in this field, to detect, identify and locate the fault in PV systems. The aim of this paper is to present a review on the latest application of the AI techniques, in the reliability assessment of PV systems. Our work is organized as follows: section 2 presents a background of simulation tools; In section 3 we propose a PV system for studying and we assess its reliability by using SPN; Application of AI in fault diagnosis is discussed in section 4.

II. BACKGROUND, STATE OF THE ART AND SIMULATION TOOLS

The failures detected in PV systems have proven that the manufacturers don't give much interest to the analysis of reliability and the estimation of the lifetimes of their solar modules. We can even see the appearance of breakdowns during the operation of PV installations, and even at the beginning of their lifespan. The reliability study of these systems has become the key problem in the industrial performance and the financial competitiveness of these installations. Hence, the use of methods and tools for the study of the systems dedicated to renewable energies has become necessary.

A. Works of Fara L. and Craciunescu D.

A contribution dedicated to the analysis of a reliable PV system for specific applications was proposed by the authors in [4]. The reliability study was based on: RAMS model (Reliability, Availability, Maintenance, and Safety) applied to a PV system using the SYNTHESIS simulation platform developed by ReliaSoft; Simulation of the PV system using the TM-21 software developed by Energy Star. The objective of this analysis was to obtain stable and sustainable operation of a PV system with regard to reliability, maintainability, availability and system degradation. In order to increase the reliability of the PV system and its components, it is necessary to develop an elaborate outage management strategy that

involves knowing the initial level of reliability and developing effective solutions to assess probable failures and risks.

B. Works of Huffman D. L. and Antelme F.

In [5, 6, 7], the authors have done a lot of research on the study of the reliability of solar energy systems. They started by choosing to model these systems using reliability block diagrams, the proposed method is easily scalable for additional branches and microinverter units per branch. Using this model, the authors could measure various system performance parameters, such as point availability, average availability, capacity, state probabilities, and losses associated with performance degradation. The authors then analyzed the reliability and security of these systems. Several tools have been used for the analysis of reliability and safety, such as FMECA (Analysis of Failure Modes, their Effects and their Criticality) and FT (Fault Tree). AI methods have thus been initiated by the authors, they are leaning towards automatic learning or machine learning. This technique offers a complementary approach to maintenance planning by analyzing large sets of performance data from each part of the system under study and environmental variables, identifying failure signatures and patterns, and providing a prediction usable from the failure of the elementary parts of the system.

C. Works of Tetsuyuki I., Takumi T. and Kenji O.

In [8], the authors showed the results of an outdoor exposure test in which the performance of fourteen PV modules consisting of five different types manufactured by six different manufacturers were measured since July 2004. The average performance is calculated each year from 2005 to 2008, and performance degradation is assessed quantitatively. The results found showed that the magnitude of performance degradation can be clearly classified by the types of PV modules. The performance difference of monocrystalline silicon PV modules between 2005 and 2008 is 1.9% to 2.8%. Polycrystalline silicon PV modules show a performance degradation of 0.7% to 1.4%. The performance of amorphous silicon / crystalline silicon decreased by 0.7%. Although a pair of a-Si PV modules had already been exposed to the sun for about six months, the pair of modules showed 4.4% performance degradation. More than half of the performance degradation occurred during the initial period from 2005 to 2006. This indicates that it takes about two years for the performance of a-Si-based modules to be stable. The performance is quite stable beyond 2006. Interestingly, the performance of copper indium gallium modules in 2008 is about 0.8% higher than in 2005.

D. Works of Sayed A., El-Shimy M., El-Metwally M. and Elshahed M.

The authors made a reliability, availability and maintainability analysis of a solar grid-connected system. This analysis has been elaborated and presented from the sub-set level to the sub-system level, then the global system [9]. Additionally, an improved reliability block diagram was presented by the authors to estimate the performance of seven PV grid-connected systems. The required input data was obtained from worldwide databases of failure and repair rates, of various subsets including various weather conditions. A new approach was initiated by the authors in order to estimate the best density function for each subset. Monitoring the critical sub-sets of a PV system will increase the possibility not only to improve the availability of the system, but also to optimize maintenance costs. In addition, it will inform the operators of the status of the various sub-systems forming the overall system. The results show that the best density function for some sub-sets, such as PV modules, connectors and charge controller, is the exponential law function. But the best density function for sub-sets, such as DC-AC converter, bypass diode, and switch, is the Weibull's law function. And finally, the best density function for the rest of the sub-sets of the PV system is the log-Normal law function. In the reliability analysis, the estimated lifetime of PV modules without encapsulation defects reaches 43.73 years. While the expected lifetimes of the converter, the inverter and the storage system are respectively 30.77 years, 8.3 years and 10.31 years.

E. Works of Obeidat F. and Shuttleworth R.

The authors in [10], studied the reliability of a PV micro-inverter with a power of 250 W. The reliability of: the grid driver, the power supply, the current and voltage sensors and the circuits of the microprocessor were studied; the failure rate and the MTBF were thus calculated. The sum of the component failure rates is equal to the complete failure rate of the PV microinverter. Taking into account the effects of temperature, the component failure rate was calculated for each inverter operating temperature and multiplied by the percentage of occurrence of that operating temperature to obtain a weighted failure rate. A similar procedure was used to calculate the failure rate of the main circuits of a 4.6 kW and 4.5 kW multi-string inverter. All calculations and methods are based on data from reference MIL-217F N2. The MTBF for the analyzed microinverter is estimated at 123938 hours (14.15 years) for a high-quality system, and it is about 19699 hours (2.25 years) for a low-quality system. Unfortunately, there is no evidence that PV microinverters fail at these rates, as manufacturers don't provide data or information on the failure rate of their microinverters. Failure rates were calculated only for the main feeder circuits of the two multi-string inverters, as details of the auxiliary circuits were not available. The MTBF for the first multi-string inverter was analyzed and estimated at 53594 hours (6.1 years), and it is about 48279 hours (5.5 years) for the second low quality factor multi-string inverter.

F. Works of Laronde R. and Charki A.

The authors have several research works which consist in developing a methodology allowing to estimate the reliability, the availability and the durability of a PV system [11, 12, 13]. To simulate the dysfunctional behavior of a PV system, the capitalization of the knowledge of each component of the system is essential. For this, the authors used methods and data that already exist to assess the reliability of these components. Those found in PV modules require extensive development efforts to assess their reliability and to estimate their lifetime: an accelerated test protocol has been proposed by the authors. They therefore developed a methodology for estimating reliability from accelerated tests: damp heat tests (to reproduce module corrosion) and UV exposure tests (to reproduce discoloration of the encapsulant). The authors also proposed to integrate the effect of degradation on the different components of the system, using several nested SPN (Stochastic Petri Nets) on different levels. This methodology takes into account the fact that an inverter only transforms the DC into AC only from a certain power threshold. The simulation of the degradation of some components makes it possible to estimate the evolution of the power of the PV system as a function of time.

On the basis of the main related research works cited above, we can make an overall synthesis summarizing the different methods, tools, and software used in the study of the reliability of PV systems, illustrated in the table below:

TABLE I. SUMMARIZED RELATED WORKS

Ref	Studied System	Methods / Simulation tools	Objectives Targeted / Achieved
[4]	PV System	- Simulation Platform SYNTHESIS of ReliaSoft - TM-21 Software of ENERGY STAR	To obtain a stable and a sustainable PV system with respect to reliability, maintainability, and system availability
[5, 6, 7]	PV System	- Reliability Block Diagrams - Accelerated Tests	To estimate a system performance parameters, such as point availability, average availability, capacity, status probabilities, and losses associated with performance degradation
[8]	PV System	Outdoor exposure tests of 14 PV modules composed of five different types manufactured by six different manufacturers, during the period from July 2004 to September 2008	The average performance is calculated each year from 2005 to 2008, and the performance deterioration is assessed quantitatively. The results showed that the magnitude of performance degradation can be clearly classified according to the types of PV modules
[9]	PV System	Reliability Block Diagrams	The results show that the best probability density function for PV modules is the exponential one. In the reliability analysis, the expected lifetime is around 30 years
[10]	PV Inverter	Probabilistic calculations based on the basis of MIL_Hbk 217F N2 database	The MTBF is estimated at around 15 years for high systems and 3 years for low systems.
[11, 12, 13]	PV System	- Stochastic Petri Nets - Accelerated tests	The estimation of the reliability, availability and durability of a PV system

III. MODELING AND RELIABILITY ASSESSMENT USING SPN

In our study, we have considered the simplest and most common case of a PV system. It is composed of three basic components connected in series: set of PV panels, AC/DC wires and PV inverters as shown in Fig. 1.

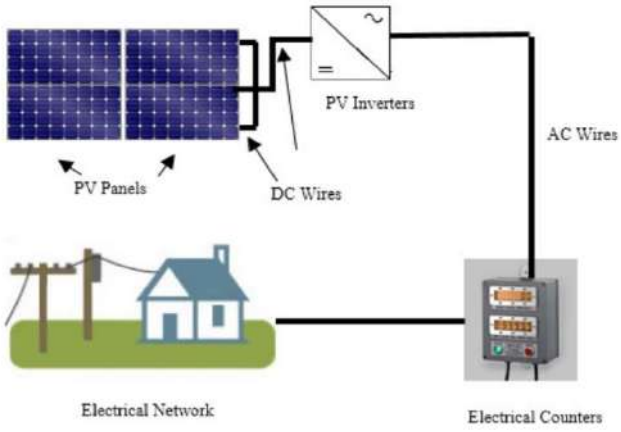


Fig. 1. PV System with its components connected in series

SPN method allows making a quantitative evaluation of the reliability and availability of our PV system in the presence of intermittent failures designed by discrete processes [14]. The availability A(t) and reliability R(t) are given by:

$$A(t) = 1 - \prod_{i=1}^{n}(1 - A_i^{m_i}(t)) \quad (1)$$

$$R(t) = 1 - \prod_{i=1}^{n}(1 - R_i^{m_i}(t)) \quad (2)$$

m_i: set of components connected in series in each subsystem i

n: subsystems connected in parallel to form the global system (parallel-series)

$\lambda_i(t)$: fault rate variations

$\mu_i(t)$: repair rate variations

$A_i(t)$: availability for each component i given by:

$$A_i(t) = \frac{\mu_i}{\mu_i + \lambda_i} + \frac{\lambda_i}{\mu_i + \lambda_i} . e^{-(\mu_i + \lambda_i)t} \quad (3)$$

The SPN of the PV system using Petri net module of Grif-TOTAL (http://grif-workshop.com/) is:

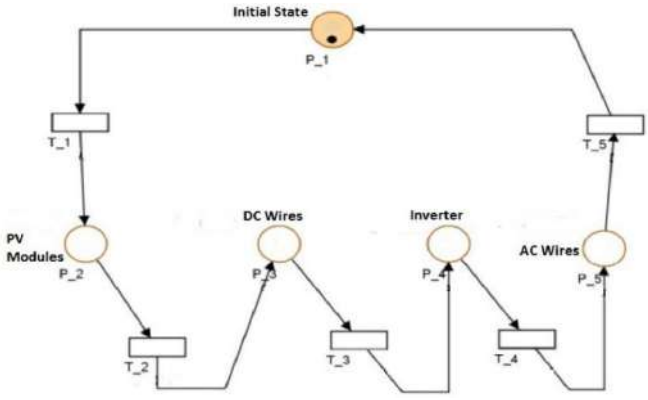


Fig. 2. SPN of working PV System

The PV system is composed of a set of PV modules, the inverter and AC/DC wires, which are connected on series. If one of these components fails, the system will don't work or give a degraded power. The SPN of working PV system is shown in Fig. 2. The delay associated with transition T1 is estimated to one hour to separate the simulation in discrete time steps [14].

The non-working PV system is illustrated in the figure below, by using the same software which was used to simulate the case of a working system (http://grif-workshop.com/). Each component can pass from the initial state (safe case) to: failing mode (T1, T4), degraded – failing mode (T2, T3, T4) and degraded – repaired mode (T2, T5) (Fig. 3) [14].

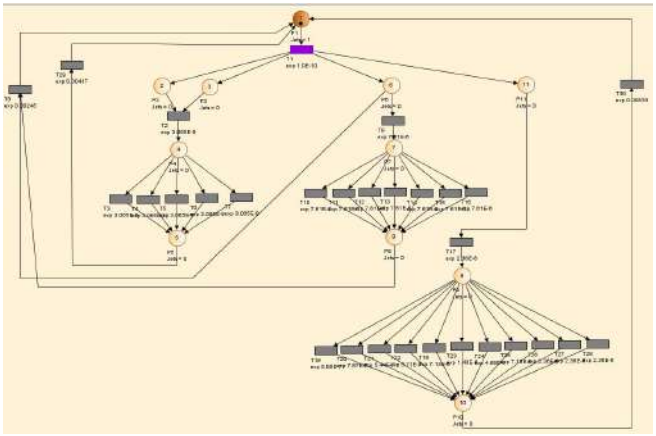


Fig. 3. SPN of non-working PV System

We have estimated the availability and the lifetime of our PV system using SPN method and be considering the time distribution to failure and to repair. We have used some necessary data for the availability assessment shown in tables below [15]:

TABLE II. PV COMPONENTS FAILURE DATA

Component	*Failure mode*	*Distribution law*	*Failure rate*
PV module	Hot spots	Exponential	$\lambda = 7.13 \cdot 10^{-7}\ h^{-1}$
	Bypass diode	Exponential	$\lambda = 5.85 \cdot 10^{-7}\ h^{-1}$
	Junction box failures	Exponential	$\lambda = 7.87 \cdot 10^{-7}\ h^{-1}$
	Encapsulant delamination	Exponential	$\lambda = 5.44 \cdot 10^{-7}\ h^{-1}$
	Broken glass	Exponential	$\lambda = 5.71 \cdot 10^{-7}\ h^{-1}$
	Broken cells	Exponential	$\lambda = 7.13 \cdot 10^{-6}\ h^{-1}$
	Solder bond failures	Exponential	$\lambda = 4.84 \cdot 10^{-6}\ h^{-1}$
	Broken interconnects	Exponential	$\lambda = 4.68 \cdot 10^{-6}\ h^{-1}$
	Corrosion	Exponential	$\lambda = 1.43 \cdot 10^{-6}\ h^{-1}$
Inverter		Exponential	$\lambda = 7.61 \cdot 10^{-6}\ h^{-1}$
AC wire		Exponential	$\lambda = 1.30 \cdot 10^{-8}\ h^{-1}$
DC wire		Exponential	$\lambda = 4.83 \cdot 10^{-8}\ h^{-1}$

TABLE III. PV COMPONENTS REPAIR DATA

Component	*Distribution law*	*Repair rate*
PV module	Exponential	$\mu = 8.33 \cdot 10^{-3}\ h^{-1}$
Inverter	Exponential	$\mu = 2.45 \cdot 10^{-3}\ h^{-1}$
AC wire	Exponential	$\mu = 4.17 \cdot 10^{-3}\ h^{-1}$
DC wire	Exponential	$\mu = 4.17 \cdot 10^{-3}\ h^{-1}$

The study of the reliability of the PV system is estimated for a period of 25 years (i.e. 219,000 hours). Its performance, during this period, shows the effect of the degradation and failure modes on it. The time to repair is generally estimated after the first failure. We have supposed that our system is composed of 24 PV modules in three different configurations:

- **Configuration 1:** three strings of eight modules (m = 8, n = 3)
- **Configuration 2:** two strings of 12 modules (m = 12, n = 2)
- **Configuration 3:** one string of 24 modules (m = 24, n = 1)

The reliability of PV system decreases by up to **50%** after **10 - 13 years** of operation regardless of PV system configuration and of PV module failures. It is highly recommended to take the necessary repair and maintenance measures beyond 15 years of operation, which are as follows:

- A good selection of the equipment from the start while respecting the specifications established for the installation of the system
- The installation must be carried out by a team composed of technicians specialized in the field
- Scheduling a regular maintenance calendar for the installation

The availability of these three cases is shown in Fig. 4, 5, and 6:

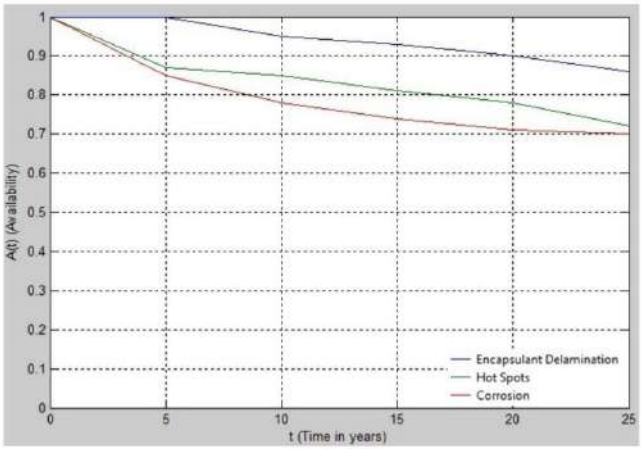


Fig. 4. PV system availability (Config-1 m = 8, n = 3)

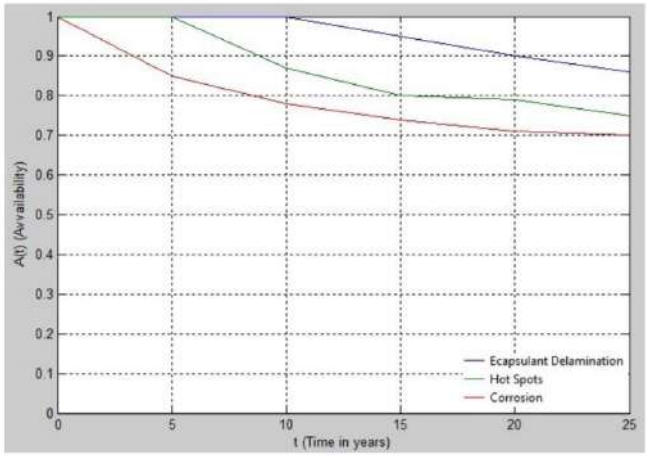


Fig. 5. PV system availability (Config-2 m = 12, n = 2)

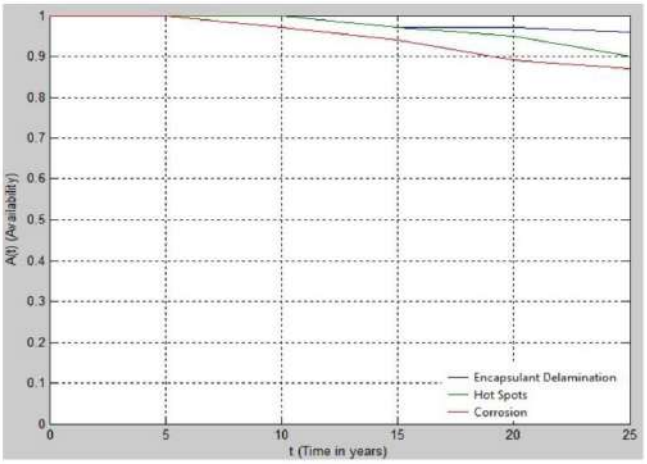


Fig. 6. PV system availability (Config-3 m = 24, n = 1)

The PV system in the third configuration (1 string of 24 modules) has a better availability: between **88 %** to **95 %** when reaching 25 years. The availability in the first and second configurations (3 strings of 8 modules and 2 strings of 12 modules) passes from **100%** to **76% - 77%** after 25 years of service in the case of the encapsulant delamination and in the presence of hot spots on the PV modules. For the corrosion failure, the PV system availability decreases by up to **75%** after 15 years of operation, and **70%** after 25 years of service for the first and second configurations. The storage of the PV modules has an impact on the availability of the system: the configuration with 1 string of 24 modules is the most optimal to gain power in the vicinity of the MPPT.

The MTBF difference, as shown in table 4, between the first configuration and the third one is between **05 years** (43,800 hours) and **07 years and 06 months** (65,700 hours). We can conclude that the string configuration in the third case is more reliable than the multi-string one. As shown previously, the availability in configuration 3 is more important during the first fifteen years of use. These results show that the MTBF can be increased by adjusting the configuration of the PV system, and to make a decision which one seems the most optimal for PV plants and to help the economic analysis of such systems.

TABLE IV. MTBF VALUE IN DIFFERENT CONFIGURATIONS

Configuration	$A(t)_{t\to+\infty}$	*MTBF*
Config-1 m = 8, n = 3	≈ 76 %	Encapsulant delamination: ≈ 43,800 h
		Hot spots: ≈ 21,900 h
		Corrosion: ≈ 17,520 h
Config-2 m = 12, n = 2	≈ 77 %	Encapsulant delamination: ≈ 87,600 h
		Hot spots: ≈ 43,800 h
		Corrosion: ≈ 17,520 h
Config-3 m = 24, n = 1	≈ 91 %	Encapsulant delamination: ≈ 109,500 h
		Hot spots: ≈ 122,640 h
		Corrosion: ≈ 61,320 h

IV. Artificial Intelligence Techniques in Fault Diagnosis

The most used definition of the AI (Artificial Intelligence) is '*imitating intelligent human behaviour*' [16]. We can consider that intelligent computing technologies are the approaches to conventional techniques of integrated systems. There are different types of AI techniques and the main one used in PV systems are: Evolutionary Algorithms EA (with meta-heuristic methods: Genetic Algorithm GA, Particle Swarm Optimization PSO and Ant Colony Optimization ACO), Machine Learning ML (including Support Vector Machine SVM, k-Nearest Neighbour k-NN and Decision Trees DT), Neural Networks NNs (MLP, RNN and RBFN), Fuzzy Logic (FL), and Deep learning (DL).

For some applications of AI techniques in fault diagnosis, detection and localization in PV systems, it can be clearly observed that researchers are more and more attracted by ML and recently DL. It has been shown that supervised learning is the most widely used algorithm. The most advanced methods-based ML require a dataset of electrical parameters (currents, voltages…) and meteorological data (solar irradiance, air temperature, …). The most investigated techniques in this field are NNs and FL, mainly used for fault classification in order to distinguish between faults that have the same signatures and classify type of faults. The most available fault diagnosis techniques have been evaluated offline and only for small-size PV plants. In addition, the most investigated faults are the ones appeared in a PV array (e.g., short / open-circuit, corrosion and hot spots). Table 5 lists some selected applications of AI techniques in fault diagnosis, detection and localization in PV systems:

TABLE V. SUMMARIZED RELATED WORKS USING AI IN FAULT DIAGNOSIS

Ref	Year	Works
[17]	2003	**Expert System-primarily based on Learning Method** is advanced so that we can diagnosis a PV system. The designed approach is used to diagnose the shading impact. The technique became simulated and established experimentally. In factor of view complexity, we will say that the technique is easy and powerful however most effective withinside the case of shedding impact
[18]	2009	**Fault diagnosis technique-based on NNs** is advanced, for a few everlasting faults. The technique is carried out for a small-scale grid-linked PV system, and it may be carried out without difficulty with a medium cost
[19]	2010	**An Artificial NNs** is used to construct a smart fault diagnosis in a small-scale PV system. The proposed technique can find and localize a temporally fault as it should be with much less time
[20]	2011	A type of **NNs called Bayesian** is introduced to locate temporally permanent failures in a grid-linked PV system. The technique is simulated for offline application. It is centered especially on data-collection. At this stage, the technique isn't capable to diagnose the fault. It required an expensive device
[21]	2011	**A FL method** used to detect faults in PV modules. The technique is simple to be implemented, however it charges because the system needs a lot of sensors to localize the faults
[22]	2014	**SVM and k-NN** techniques used to detect and classify short circuit failure in a PV panel. The technique was examined and simulated using MATLAB/Simulink. Only short circuit is investigated; However, the technique needs some knowledge in ML
[23]	2015	**FL** used for comparing the diagnostic rules. The designed approach may be used for detecting partial shading, elevated series-resistance losses and potential-

		prompted degradation in a PV system. The fundamental characteristic of the techniques is rapid and low computation. It may be used for each type of faults: permanent and temporal
[24]	2016	**NNs** model designed to localize fault in a PV system. The technique became examined for a temporal fault (shading impact). It is so simple to implement and it was tested with MATLAB/Simulink for a PV panel
[25]	2017	**Technique based on I-V curves evaluation and FL** for fault detection in DC facet of a 1.1 kWp grid-linked PV system was done. The research fault is partial shading impact on PV modules. The technique is easy to be carried out and it was verified experimentally
[26]	2018	**FL and radial function networks** used in comparison for detecting fault in PV systems. The techniques had been used to localize and classify permanent and temporally faults
[27]	2019	A fault diagnosis approach combines a **Long Short-Time Memory Network (LSTMN)** and **Softmax Regression Classifier (SRC)** is designed. The LSTMN used to extract features, and the remaining one is used as enter to a regression classifier. The model is used to check hot spots and line-to-line faults with proper accuracy
[28]	2019	**ML and thermography** used to categorize hot spot fault in PV modules. The technique outperforms the type strategies and presents satisfying results. The technique requires an expensive equipment

V. Conclusion

In this paper, the literature on reliability assessment and fault diagnosis of PV systems using classical analysis methods and AI techniques is briefly reviewed. Through the simulations using SPN, the results obtained for the various configurations studied showed the impact of the PV modules storage on the reliability, availability and durability of the system. The availability and the MTBF increase in the case where the PV modules are ranged in series, thus the output power of the panel increases and will be close to the value of the MPPT. In addition, the failures that occur in PV modules have a significant impact on system reliability. The proposed methodology can be used for optimizing the performance of the installation of a PV system. The influence of each failure mode can be analyzed in order to improve each component of the PV system. The AI techniques have recently attracted many researchers, and many efforts have been deployed in order to design and implement new advanced methods for fault diagnosis of PV systems. It should be noted that fault diagnosis accuracy using ML algorithms is relatively not scalable, particularly with a large-size database. However, DL algorithms are the most appropriate in this case, due to their capability to manage a huge number of data. The application of AI techniques in fault diagnosis of PV systems, particularly DL, will continue to progress in the near future.

References

[1] https://www.pv-magazine.com/2022/03/15/humans-have-installed-1-terawatt-of-solar-capacity/

[2] Mellit, A., Tina, G.M., Kalogirou, S.A. "Fault detection and diagnosis methods for photovoltaic systems: a review". Sustain. Energy Rev. 91, 1–17 (2018)

[3] Mellit, A., Kalogirou, S.A. "Artificial intelligence techniques for photovoltaic applications: a review". Prog. Energy Combustion Sci. 34(5), 574–632 (2008)

[4] Fara, L., Craciunescu D. "Reliability analysis of photovoltaic system for specific applications". Book Chapter, DOI: http://dx.doi.org/10.5772/intechopen.85688. IntechOpen (2019)

[5] Huffman D. L., Antelme F. "Availability analysis of solar power system with graceful degradation". Annual Reliability and Maintainability Symposium, DOI: 10.1109/RAMS.2009.4914701 (2009)

[6] Huffman D. L., Antelme F. "What we can learn about reliability and safety analyses from different industries". Proceedings Annual Reliability and Maintainability Symposium (RAMS), DOI: 10.1109/RAMS.2013.6517694 (2012)

[7] Huffman D. L., Antelme F. "Emerging trends in risk assessment and evaluation". Proceedings Annual Reliability and Maintainability Symposium (RAMS), DOI: 10.1109/RAMS.2015.6517691 (2014)

[8] Tetsuyuki I., Takumi T., Kenji O. " Long - term performance degradation of various kinds of photovoltaic modules under moderate climatic conditions " . Wiley Edition, DOI : https://doi.org/10.1002/pip.1005 (2011)

[9] Sayed A., El-Shimy M., El-Metwally M., Elshahed M. "Reliability, Availability and Maintainability Analysis for Grid-Connected Solar Photovoltaic Systems". Energies Journal, DOI : https://doi.org/10.3390/en12071213 (2019)

[10] Obeidat F., Shuttleworth R. "Photovoltaic Inverters Reliability Prediction". World Applied Sciences Journal 35 (2): 275-287, DOI : 10.5829/idosi.wasj.2017.275.287 (2017)

[11] Laronde R., Charki A. "Reliability of photovoltaic modules based on climatic measurement data". Int. J. Metrol. Qual. Eng., DOI : https://doi.org/10.1051/ijmqe/2010012 (2010)

[12] Laronde R., Charki A. "Photovoltaic system lifetime prediction using Petri networks method". Proceedings Volume 7773, Reliability of Photovoltaic Cells, Modules, Components, and Systems III - 777306, DOI : https://doi.org/10.1117/12.856110 (2011)

[13] Laronde R., Charki A. "Lifetime Estimation of a Photovoltaic Module Subjected to Corrosion Due to Damp Heat Testing". Journal of Solar Energy Engineering, DOI : https://doi.org/10.1115/1.4023101 (2013)

[14] Mahdi I., Nadji B., Simeu-Abazi Z. "Contribution to The Reliability Study of Photovoltaic Systems Using Static and Dynamic Analysis Methods". International Journal of Renewable Energy Technology (IJRET), Volume 11 N° 01. DOI: 10.1504/IJRET.2020.106515 (2020)

[15] "Military handbook : reliability prediction of electronic equipment". MIL-HDBK-217F, December 1991

[16] Russell, S.J., Norvig, P. "Artificial intelligence: a modern approach". 3rd ed. Prentice-Hall Inc, USA (2009)

[17] Yagi, Y., Kishi, H., Hagihara, R., Tanaka, T., Kozuma, S., Ishida, T., et al. Diagnostic technology and an expert system for photovoltaic systems using the learning method". Solar Energy Mater Solar Cells 75, 655–663 (2003)

[18] Wu, Y., Lan, Q., Sun, Y. "Application of BP neural network fault diagnosis in solar photovoltaic system". In: IEEE international conference on mechatronics and automation, pp. 2581–2585 (2009)

[19] Chao, K.-H., Chen, C.-T., Wang, M.-H., Wu, C.-F. "A novel fault diagnosis method based on modified neural networks for photovoltaic systems". In Advances in swarm intelligence. Springer, Berlin, pp. 531–539 (2010)

[20] Coleman, A., Zalewski, J. "Intelligent fault detection and diagnostics in solar plants". In: The 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), pp. 948–953 (2011)

[21] Cheng, Z., Zhong, D., Li, B., Liu, Y. "Research on fault detection of PV array based on data fusion and fuzzy mathematics". In: IEEE Asia-Pacific power and energy engineering conference, pp. 1–4 (2011)

[22] Rezgui, W., Mouss, L.H., Mouss, N.K., Mouss, M.D., Benbouzid, M. "A smart algorithm for the diagnosis of short-circuit faults in a photovoltaic generator". IEEE First Int Conf Green Energy ICGE 2014, 139–143 (2014)

[23] Spataru, S., Sera, D., Kerekes, T., Teodorescu, R. "Diagnostic method for photovoltaic systems based on light I-V measurements". Sol. Energy 119, 29–44 (2015)

[24] Mekki, H., Mellit, A., Salhi, H. "Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules". Simul Modelling Pract Theor 67, 1–13 (2016)

[25] Dhimish, M., Holmes, V., Mehrdadi, B., Dales, M., Mather, P. "Photovoltaic fault detection algorithm based on theoretical curves modelling and fuzzy classification system". Energy 140, 276–290 (2017)

[26] Dhimish, M., Holmes, V., Mehrdadi, B., Dales, M. "Comparing Mamdani Sugeno fuzzy logic and RBF ANN network for PV fault detection". Renew Energy 117, 257–274 (2017)

[27] Appiah, A.Y., Zhang, X., Ayawli, B.B.K., Kyeremeh, F. "Long short-term memory networks based automatic feature extraction for photovoltaic array fault diagnosis". IEEE Access 7, 30089–30101 (2019)

[28] Kurukuru, V.B., Haque, A., Khan, M.A. "Fault classification for photovoltaic modules using thermography and image processing". In: IEEE industry applications society annual meeting, pp. 1–6 (2019)

Speech Analysis-Synthesis using Deep Neural Networks: A Review

Youcef Tabet
LIST Laboratory
Institute of Electrical and Electronic Engineering
University of M'Hamed Bougara
Boumerdes, Algeria
y.tabet@univ-boumerdes.dz

Amel Boustil
LIMOSE Laboratory
University of M'Hamed Bougara
Boumerdes, Algeria
a.boustil@univ-boumerdes.dz

Karim Baiche
LIST Laboratory
University of M'Hamed Bougara
Boumerdes, Algeria
kbaiche@univ-boumerdes.dz

***Abstract*—The goal of this paper is to provide a short but a comprehensive overview of speech synthesis methods. The most important speech synthesis methods will be described starting from the earlier ones up until the last ones, in order to highlight major improvements over these methods. First two rule-based synthesis techniques (formant synthesis and articulatory synthesis) are explained then the concatenative synthesis is explored. Unit selection synthesis, which is a kind of concatenative synthesis, is also described in this paper. Next Hidden Markov model (HMM) synthesis is introduced. Finally three artificial neural network speech synthesis techniques are described. Emphasis will be given to the artificial neural network speech synthesis techniques, since they seem to be more promising and robust speech synthesis techniques. Speech Analysis-Synthesis, HMM Synthesis, Deep Neural Network Synthesis, End to End Synthesis**

I. Introduction

Speech synthesis is the process of producing artificial speech by computers with the aim to obtain a synthetic speech understandable and indistinguishable from normal human speech [1]. Hence, during the last decades, synthetic speech has been developed steadily in order to improve the intelligibility and naturalness of the speech synthesis system output.

In current speech synthesis systems, several techniques have been developed with the aim of generating natural sounding speech. Hence, the primary goal of this paper is to give an overview of the most popular and important approaches used in speech synthesis by highlighting major improvements over these techniques.

There are two main approaches to speech synthesis: Conventional and artificial neural networks speech synthesis methods. Conventional speech synthesis methods include on one hand the following: formant synthesis[4,5], articulatory synthesis[9], and concatenative synthesis[9]. Formant synthesis models the frequencies of speech signal. Formants are the resonance frequencies of the vocal tract. The speech is synthesized using these estimated frequencies. Articulatory synthesis generates speech by direct modeling of human articulator behavior. On the other hand, concatenative speech synthesis produces speech by concatenating small, prerecorded units of speech, such as phonemes, diphones and triphones to construct the utterance. In case that not just one, but hundreds of realizations of each phonetic speech unit are present in an inventory, a unit selection process [9] must take place in order to create the final synthetic unit sequence. Such speech synthesis method is also called corpus based speech synthesis. The HMM (Hidden Markov model) synthesis system [11, 12] is introduced in order to obtain less memory to store the parameters of the models and more variations are allowable.

Deep neural networks speech synthesis methods include on the other hand the following: deep neural net work [14], WaveNet [15], Tacotron [16] and deep voice [17] synthesis methods. The use of the Deep Neural Network [14] can address some limitations of the conventional approaches. WaveNet [15], is a deep neural network for generating raw audio waveforms. The model is fully probabilistic and autoregressive, with the predictive distribution for each audio sample conditioned on all previous ones; nonetheless we show that it can be efficiently trained on data with tens of thousands of samples per second of audio. When applied to text-to-speech, it yields state-of-the-art performance. Tacotron[16], is an end-to-end generative text-to-speech model that synthesizes speech directly from characters. Given ¡text, audio¿ pairs, Tacotron can be trained completely from scratch with random initialization. Deep Voice, is a production-quality text-to-speech system constructed entirely from deep neural networks. Deep Voice [17] lays the groundwork for truly end-to-end neural speech synthesis.

The remainder of the paper is organized as follows. First we give an overview of the Formant synthesis. After that the Articulatory synthesis is depicted. Next, the

Concatenative synthesis is described. This is followed by a survey of the Unit Selection and the HMM synthesis. Finally, the artificial neural network synthesis is shown. The conclusion is given at the end of the paper.

II. Conventional speech synthesis methods

A. Formant synthesis

This method is based on the assumption that the transfer function of the vocal tract can be modeled satisfactorily by simulating the formant amplitudes and frequencies. Synthesis therefore consists of the artificial reconstruction of the characteristics of the formants to be produced. The technique consists of exciting a set of resonators with a voice source or noise generator to obtain the desired voice spectrum, and controlling the excitation source to simulate either voice or the absence of voice. The addition of a set of anti-resonators furthermore allows the simulation of nasal tract effects, fricatives and plosives. For a satisfactory restitution of the speech signal, a specification of about 20 or more of such parameters is sufficient. The advantage of this technique is that its parameters are highly correlated with the production and propagation of sound in the oral tract. The main currentdrawback of this approach lies in the widespread dissatisfaction with automatic techniques for specifying the formant parameters, and that therefore the majority of the parameters still have to be optimized manually [2]. Formant synthesis relies on rules written by linguists to generate the parameters that will allow speech synthesis and to manage the transition from one phoneme to another, i.e., coarticulation and it does not use any human speech samples. To write the rules, linguists have studied spectrograms and derived the rules of evolution of formants. However, we do not yet know the optimal rule to do this [3]. Moreover, the speech waveform is naturally produced in such a complex process that, currently, rules can only model the features of the speech waveform. Therefore, the synthesized speech has an artificial, robotic sound, and the goal of naturalness is not reached. However, the rule-based synthesized speech is very intelligible, even at high speeds, which is quite useful for visually impaired for quickly navigating computers using a screen reader. Moreover, when memory and processing costs are limited, such as in embedded systems, these synthesizers are more interesting because they don't have a database of speech samples. The formant synthesis approach has been implemented in MITalk [4, 5], in KlatTalk [6], and in DECTalk [7].

B. Articulatory synthesis

Articulatory synthesis generates speech by direct modeling of the human articulator behavior, so in principle it is the most satisfying method to produce high-quality speech. In practice, it is one of the most difficult methods to implement. The articulatory control parameters include lip aperture, lip protrusion, tongue tip position, tongue tip height, tongue position and tongue height [8]. There are two difficulties in articulatory synthesis. The first difficulty is acquiring data for articulatory model. This data is usually derived from Xray photography. X-ray data do not characterize the masses or degrees of freedom of the articulators [3]. The second difficulty is to find a balance between a highly accurate model and a model that is easy to design and control. In general, the results of articulatory synthesis are not as good as the results of formant synthesis or the results of concatenative synthesis.

C. Concatenative synthesis

The main limitation of formant synthesis and articulatory synthesis is not so much in generating speech from parametric representation, but the difficulty is in finding these parameters from the input specification that was created by the text analysis process. To overcome this limitation, concatenative synthesis follows a data driven approach. Concatenative synthesis generates speech by connecting natural, prerecorded speech units. These units can be words, syllables, half-syllables, phonemes, diphones or triphones. The unit length affects the quality of the synthesized speech. With longer units, the naturalness increases, less concatenation points are needed, but more memory is needed and the number of units stored in the database becomes very numerous. With shorter units, less memory is needed, but the sample collecting and labeling techniques become more complex [9]. The most widely used units in concatenative synthesis are diphones. A diphone is a unit that starts at the middle of one phone and extends to the middle of the following one. Diphones have the advantage of modeling coarticulation by including the transition to the next phone inside the diphone itself. The full list of diphones is called diphone inventory, and once determined, they need to be found in real speech. To build the diphone inventory, natural speech must be recorded such that all phonemes within all possible contexts (allophones) are included, then diphones must be labeled and segmented. Once the diphone inventory is built, the pitch and duration of each diphone need to be modified to match the prosodic part of the specification.

D. Unit selection synthesis

In concatenative synthesis, diphones must be modified by signal processing methods to produce the desired prosody. This modification results in artifacts in the speech that can make the speech sound unnatural. Unit selection synthesis (also, called corpus-based concatenative synthesis) solves this problem by storing in the unit inventory multiple instances of each unit with varying prosodies. The unit that matches closest to the target prosody is selected and concatenated so that prosodic modifications needed on the selected unit is either minimized or not necessary at all. Since multiple instances of each unit are stored in the unit inventory, a unit selection algorithm is needed to choose the units that best match the target specification. This selection is based on minimizing two types of cost

functions, which are target cost and join cost. In the case of automatic unit selection, the coarticulatory influence isn't limited to the last phonemeThe database is much larger (1-10 hours) and comprises several occurrences of each acoustic unit, captured under various contexts (like its neighboring phonemes of course, but also its pitch, its duration, its position in the syllable, etc.). As a result, the sequence of phonemes to synthesize leads to a lattice of acoustic units, in which the best corresponds to the expected contexts (prosody, phonetics, etc) but also minimizes the spectral and prosodic discontinuities. Consequently, automatic unit selection requires much less modification of the speech units, which leads to an overall quality of the synthesized speech much more natural than with diphones based synthesis. Apart from this naturalness, unit selection techniques have several disadvantages. They rely on a very large database, which implies, on the one hand, considerable development time and cost to collect and label the data, and on the other hand, large memory resource requirements to store the data. The second drawback is incorrect labeling and occurrence of unseen target contexts lead to fragments of synthesized speech of extremely poor quality. This phenomenon of unseen contexts may well never be fully overcome with concatenative synthesis as [10] suggest that rare events will always occur in language.

E. Hidden Markov model (HMM) synthesis

In unit selection synthesis, multiple instances of each phone in different contexts are stored in the database. To build such a database is a time consuming task and the database size increases in an enormous way. Another limitation of the concatenative approach is that it limits us to recreate what we have recorded. An alternative is to use statistical parametric synthesis techniques to infer specification to parametric mapping from data. These techniques have two advantages: firstly, less memory is needed to store the parameters of the models than to store the data itself. Secondly, more variations are allowable for example; the original voice can be converted into another voice. One of the most usable statistical parametric synthesis techniques is the hidden Markov model (HMM) synthesis. It consists of two main phases, the training phase and the synthesis phase. At the training phase, it should be decided which features the models should be trained for. Mel frequency cepstral coefficients (MFCC) and their first and second derivatives are the most common types of features used. The feature are extracted per frame and put in a feature vector. The Baum-Welch algorithm is used with the feature vectors to produce models for each phone. A model usually consists of three states that represent the beginning, the middle and the end of the phone. The synthesis phase consists of two steps: firstly, the feature vectors for a given phone sequence have to be estimated. Secondly, a filter is implemented to transform those feature vectors into audio signals. The quality of the HMM generated speech is not as good as the quality of the speech generated from unit selection synthesis. The modeling accuracy can be improved by using hidden semi-Markov models (HSMMs) [11], trajectory HMMs [12], and stochastic Markov graphs [13].

III. Deep neural network speech synthesis methods

A. Speech synthesis using deep neural network model

Conventional approaches to statistical parametric speech synthesis typically use decision tree-clustered context-dependent Hidden Markov models (HMMs) to represent probability densities of speech parameters given texts. Speech parameters are generated from the probability densities to maximize their output probabilities, and then a speech waveform is reconstructed from the generated parameters. This approach is reasonably effective but has a couple of limitations, e.g. decision trees are inefficient to model complex context dependencies. The authors in [14] examines an alternative scheme that is based on a deep neural network (DNN). The relationship between input texts and their acoustic realizations is modeled by a DNN. The use of the DNN can address some limitations of the conventional approach. Experimental results showed that the DNN-based systems outperformed the HMM-based systems with similar numbers of parameters.

Hence, The authors in [14] examined the use of the DNNs to perform speech synthesis. The DNN-based approach has a potential to address the limitations in the conventional decision tree-clustered context-dependent HMM-based approach, such as inefficiency in expressing complex context dependencies, fragmenting the training data, and completely ignoring linguistic input features which did not appear in the decision trees. The objective evaluation showed that the use of a deep architecture improved the performance of the neural network-based system for predicting spectral and excitation parameters. Furthermore, the DNN-based systems achieved better preference over the HMM-based systems with a similar numbers of parameters in the subjective listening test. These experimental results showed the potential of the DNN-based approach for statistical parametric speech synthesis.

One of the advantages of the HMM-based system over the DNN-based one is the reduced computational cost. At synthesis time, the HMM-based systems traverse decision trees to find statistics at each state. On the other hand, the DNN-based system in [14] performs mapping from inputs to outputs which includes a number of arithmetic operations at each frame.

B. Speech synthesis using WaveNet model

The authors in [15] introduceD WaveNet, a deep neural network for generating raw audio waveforms. The model is fully probabilistic and autoregressive, with the predictive distribution for each audio sample conditioned on all previous ones; nonetheless we show that it can be efficiently trained on data with tens of thousands of samples per second of audio. When applied to text-to-speech, it yields state-of-the-art performance, with human listeners rating it as significantly more natural sounding than the best parametric and concatenative systems for both English and Mandarin. A single WaveNet can capture the characteristics of many different speakers with equal fidelity, and can switch between them by conditioning on the speaker identity. When trained to model music, the authors found that it generates novel and often highly realistic musical fragments. they also showed that it can be employed as a discriminative model, returning promising results for phoneme recognition.

C. Speech synthesis using Tacotron model

A text-to-speech synthesis system typically consists of multiple stages, such as a text analysis frontend, an acoustic model and an audio synthesis module. Building these components often requires extensive domain expertise and may contain brittle design choices. In [16], the authors presented Tacotron, an end-to-end generative text-to-speech model that synthesizes speech directly from characters. Given (text, audio) pairs, the model can be trained completely from scratch with random initialization. the authors presented several key techniques to make the sequence-to-sequence framework perform well for this challenging task. Tacotron achieves a 3.82 subjective 5-scale mean opinion score on US English, outperforming a production parametric system in terms of naturalness. In addition, since Tacotron generates speech at the frame level, it's substantially faster than sample-level autoregressive methods.

D. Speech synthesis using deep voice model

In [17] Deep Voice is presented. The model is a production-quality text-to-speech system constructed entirely from deep neural networks. Deep Voice lays the groundwork for truly end-to-end neural speech synthesis. The system comprises five major building blocks: a segmentation model for locating phoneme boundaries, a grapheme-to-phoneme conversion model, a phoneme duration prediction model, a fundamental frequency prediction model, and an audio synthesis model. For the segmentation model, the authors in [17] proposed a novel way of performing phoneme boundary detection with deep neural networks using connectionist temporal classification (CTC) loss. For the audio synthesis model, we implement a variant of WaveNet that requires fewer parameters and trains faster than the original. By using a neural network for each component, the deep voice system is simpler and more flexible than traditional text-to-speech systems, where each component requires laborious feature engineering and extensive domain expertise. Finally, the authors in [17] showed that inference with their system can be performed faster than real time and describe optimized WaveNet inference kernels on both CPU and GPU that achieve up to 400x speedups over existing implementations.

IV. Conclusions and future work

In This paper, we have presented a survey of several speech synthesis techniques. Formant synthesis and articulatory synthesis are less used today but these techniques can be suitable for applications that require less memory and low processing cost. The unit selection and HMM speech synthesis methods allow for more natural-sounding modifications of the signal. The parametric representation of speech using HMM model provides a straightforward way of smoothing discontinuities of acoustic units around concatenation points. The main limitation of the unit selection synthesis is high processing cost and fewer variations are allowable on the recorded data. Hidden Markov Model Synthesis is statistical methods that allow more variations on the recorded data but the focus nowadays is on the deep neural network speech synthesis methods due to their ability to provide more natural and intelligible synthetic speech compared to the conventional speech synthesis techniques' outputs.

References

[1] P. Taylor, "Text-to-Speech Synthesis". Cambridge Univesity Press. 2009

[2] T .Styger, and E. Keller. "Formant synthesis," In E. Keller (ed.), Fundamentals of Speech Synthesis and Speech Recognition: Basic Concepts, State of the Art, and Future Challenges (pp. 109-128). Chichester: John Wiley.1994

[3] D.H. Klatt "Review of text-to-speech conversion for English," Journal of the Acoustical Society of America, vol. 82(3), 1987.

[4] J. Allen, S. Hunnicutt, R. Carlson and B. Granstrom, "MITalk-79: The 1979 MIT text-to-speech system," Speech communications papers presented at the 97th meeting of the acoustical society of america, Cambridge, USA, pp. 507- 507,1979.

[5] J. Allen, S. Hunnicutt and D.H. Klatt, "From Text-tospeech: The MITalk System," Cambridge University Press, Cambridge, 1987

[6] D.H. Klatt, "The klattalk text-to-speech conversion system," Proceeding on the international conference on acoustic, speech and signal processing, Paris, pp. 1589-1592, 1982.

[7] D.H. Klatt, "DecTalk user's manual," Digital Equipment Corporation Report, 1990.

[8] B. Kroger , "Minimal Rules for Articulatory Speech Synthesis," Proceedings of EUSIPCO92, pp.331-334,1992.
[9] T. Dutoit, "High-Quality Text-to-Speech Synthesis: an Overview," Journal of Electrical & Electronics Engineering,Australia: Special Issue on Speech Recognition and Synthesis, vol. 17, pp. 25-37, 1999.
[10] B. Mobius, "Rare events and closed domains: Two delicate concepts in speech synthesis," Proceedings of the 4th ESCA Workshop on Speech Synthesis, Perthshire, Scotland, 2001.
[11] K. Tokuda et al., "Hidden semi-Markov model based speech synthesis," in Inter speech, pp. 1185–1180, 2004.
[12] K. Tokuda et al., "An introduction of trajectory model into HMM-based speech synthesis," in ISCA SSW5, 2004.
[13] M. Eichner et al., "Speech synthesis using stochastic Markov graphs," in ICASSP, pp. 829–832, 2001.
[14] Ze, Heiga, A. Senior, and M. Schuster. "Statistical parametric speech synthesis using deep neural networks." 2013 ieee international conference on acoustics, speech and signal processing. IEEE, 2013.
[15] Oord, Aaron van den, et al. "Wavenet: A generative model for raw audio." arXiv preprint arXiv:1609.03499. 2016.
[16] Wang, Yuxuan, et al. "Tacotron: Towards end-to-end speech synthesis." arXiv preprint arXiv:1703.10135 2017.
[17] Arık, Sercan Ö., et al. "Deep voice: Real-time neural text-to-speech." International conference on machine learning. PMLR, 2017.

A hybrid shuffled frog leaping algorithm-back propagation algorithm for adaptive channel equalization

Saidani Samir[(1)], Ghadjati Mohamed[(2)] and Moussaoui Abdelkrim [(2)]

[(1)]Advanced Control Laboratory (LABCAV), Université 8 Mai 1945 Guelma, Algeria.
[(2)] Laboratory of Electrical Engineering of Guelma (LGEG), Université 8 Mai 1945 Guelma, Algeria.

e-mail:saidani.tlc@gmail.com

***Abstract*— In this paper, we have proposed a hybrid training algorithm for the problem of channel equalization, also referred to as SFLA–BP algorithm, which introduces BP to SFLA to update the weights of artificial neural network (ANN) in order to accelerate the standard SFLA. The proposed training algorithm combines the characteristic of gradient search algorithm because of its strong local-search ability and the characteristic of stochastic search algorithms due to its global-search ability. Its performance is investigated with nonlinear channel equalization, and the results of experiments show improved performance obtained by the proposed hybrid SFLA-BP algorithm in convergence speed, robustness and stability, as compared to the modified SFLA (the MSFLA) and standard BP algorithm.**

***Keywords*—Artificial Neural Network (ANN), Shuffled Frog Leaping Algorithm (SFLA), Back-propagation (BP), Adaptive Channel Equalization, Digital Communication.**

I. INTRODUCTION

Adaptive channel equalizers play an important role in digital communication systems; its main function is returning the transmitted signal after being distorted by the channel effect and noise. Different ways have been suggested to decrease the effects of channel nonlinearity and additive gaussian noise on signal transmission [1]. Artificial neural networks (ANN) have been used in the field of channel equalization to combat nonlinear distortion in digital communication systems. Neural networks-based equalizers provide an important performance improvement in a different of communication channels compared to conventional equalization [2]. The success of ANNs for a particular problem depends on the adequacy of the training algorithm regarding the necessities of the problem [3], [4]. The existing gradient-based techniques, in particular the Back-propagation algorithm which is one of the most popular training algorithms designed to minimize the minimum squared error (MSE) between the output of ANN and its desired output. However, gradient search algorithm perform local searches, they are susceptible to the local minimum problem [3], [4].

The use of stochastic algorithms is an important alternative for ANN training, because they are characterized by performing global searches [4]. Among them, we highlight the GAs, Differential Evolution (DE), particle swarm optimization (PSO) and Shuffled Frog Leaping Algorithm (SFLA), which have been also used in training ANNs.

Moreover, in spite of the fact that they are efficient to run global searches, they are insufficient to follow more local searches [4]. In this light, we implemented a new hybrid algorithm, called SFLA-BP, which combines the advantages of both gradient search algorithm and stochastic search algorithms. The Shuffled Frog Leaping Algorithm is a memetic metaheuristic that is designed to seek a global optimal solution by performing an informed heuristic search using a heuristic function [5]. It is based on evolution of memes carried by interactive individuals and a global exchange of information among the population [5].

In SFLA-BP algorithm, SFLA ensures that the search converges faster because of its global-search ability, while BP makes the search passes over the local optima to obtain the global optimum because of its strong local-search ability. Compared with SFLA and BP, the proposed algorithm is shown to be more superior in performance in channel equalization because it combines them both.

This paper is organized as follows. Section 2 briefly introduces the BP based neural network. The shuffled frog

leaping algorithm is considered in section 3. The proposed new hybrid SFLA–BP algorithm is described in Section 4. Subsequently, Section 5 provides the simulation results. Finally, our conclusion is presented in Section 6.

II. BP NEURAL NETWORK

Back-propagation (BP) [6] algorithm is the best known and one of the most common learning algorithms used in neural networks. It is a generalized LMS algorithm that minimizes the mean-squared error between the actual and desired outputs of the network [7].

The architecture of a BP neural network as shown in Fig.1, consists of an input layer, one or more hidden layers and an output layer. All input nodes are connected to all hidden nodes, and all hidden nodes are connected to output node. From Fig. 1, it can be seen that the outputs of network is expressed in the following way:

$$y_j = f\left(\sum_{i=1}^{I} w_{ji}^{(1)} x_i + b_j\right) \quad (1)$$

$$o = h\left(\sum_{j=1}^{J} w_j^{(2)} y_j + b\right) \quad (2)$$

Where y_j is the output of the jth hidden node, I is the number of the input node, $w_{ji}^{(1)}$ is the connection weight from the ith node of input layer to the jth node of hidden layer, x_i is the ith input node, $w_j^{(2)}$ is the connection weight from the jth node of hidden layer to the output node of output layer and b is the threshold level.

In the above expressions, h represents the output activation function (linear in this case), and f is non linear activation function of the hidden layer.

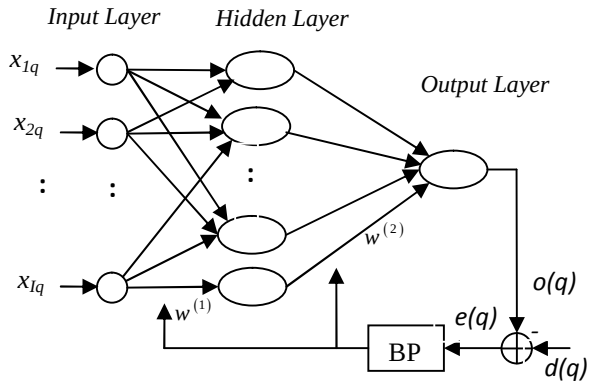


Fig.1 Neural network trained with BP algorithm.

For a standard BP algorithm, the output of the node in the output layer is compared with the desired response, resulting an error signal. This error signal is propagated layer by layer from the output layer to the input layer to adaptively adjust all weights in neural networks (hence the name back propagation algorithm for this training process). This training procedure continues until the weights of the network are updated sufficiently.

The BP algorithm adjusts the network weights so as to minimize a cost function, e.g. the squared error function given by:

$$E = \frac{1}{2}\sum_{q=1}^{Q} (e^{(q)})^2 \quad (3)$$

$$e^{(q)} = \left(d^{(q)} - o^{(q)}\right) \quad (4)$$

Where Q is the number of patterns, d is the desired output, and o is the network output.

The basic BP algorithm calculates the gradient of E and updates the weights by moving them along the gradient-descendent direction. This can be summarized with the expression [8]:

$$w_{ji}^{(1)} = w_{ji}^{(1)} - ¦ \frac{\partial E}{\partial w_{ji}^{(1)}} \quad (5)$$

$$w_j^{(2)} = w_j^{(2)} - ¦ \frac{\partial E}{\partial w_j^{(2)}} \quad (6)$$

Where the parameter ¦ is the learning rate that controls the learning speed.

III. SHUFFLED FROG LEAPING ALGORITHM

The shuffled frog-leaping algorithm, developed by Muzaffar Eusuff and Kevin Lansey in 2003 [5], is a memetic meta-heuristic that is described to search a global optimal solution of the problem. It combines the advantages of the genetic-based memetic algorithm (MA), and the social behaviour-based particle swarm optimization (PSO) [9]. SFLA consists of a population of frogs partitioned into different subgroups called memeplexes, in which individual frogs represent a set of possible solution. In each memeplex a local independent search (uses the search strategy of PSO) is performed by the set of frogs [10]. Each frog can effect on the idea of the other frogs through a process of memetic evolution [5]. Selecting frogs using a triangular probability distribution provides a competitive advantage to better ideas [11]. Moreover the ideas can be exchanged among memplexes via a shuffling process. Then a local search and shuffling processes will be continued until the termination criteria met.

The flowchart of the SFLA is illustrated in Figure 2. At first, the initial individuals of P frogs are produced randomly. A fitness function is defined to evaluate the frog's position. Then the frogs are sorted in a descending order according to their fitness, and the frog with the global best fitness is identified as X_g.

Then, the entire population is divided into m memeplexes, *n* frogs in each memeplex ($P = m \times n$). At this stage, the first frog goes to the first memeplex and the second frog goes to the second memeplex, *m*th frog goes to the *m*th memeplex, and *(m+1)*th frog back to the first memeplex, etc.

Within each memeplex there is a sub-memeplex including *q* frogs generated by a triangle probability distribution expressed in Equation (7). The sub-memeplex selection strategy is to give higher weights to frogs that have higher performance values and less weight to those with lower performance values [11].

$$P_j = \frac{2(n+1-j)}{n(n+1)}, \quad j = 1,.....,n \tag{7}$$

Where *j* is the fitness sorting number of the current frog in the population according to the fitness value in decreasing order. More explanation about this process reviewed in [11].

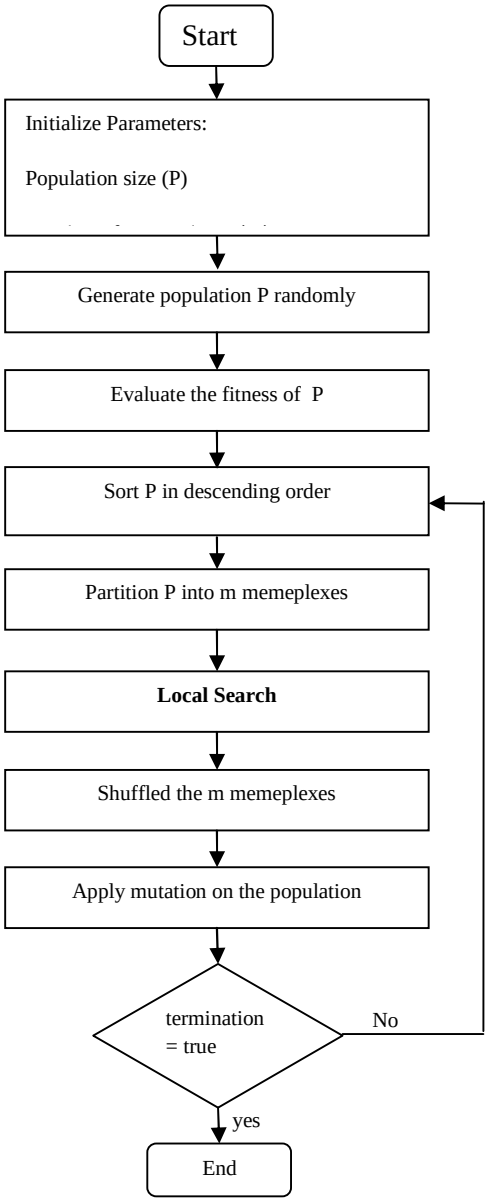


Fig.2. Flowchart of the algorithm SFLA

The fundamental function of the algorithm is to update the position of the worst frog through iterative operation in each sub-memeplex. The new position of the worst frog is updated as follows:

$$D_i = rand(\ ).(X_b - X_w) \tag{8}$$

$$X_w^{(new)} = X_w + D_i, \ (-D_{\max} \le D_i \le D_{\max}) \tag{9}$$

Where X_w and X_b are the worst frog position and the best frog position respectively in the sub- memeplex. *rand( )* is a random number in the range of [0,1]. D_{max} is the maximum allowed change of frog's position in each leaping. If the new position of the worst frog is not better than before, the calculations in Equations (8) and (9) are repeated with replacement of X_b by X_g. If this process still can't obtain the better performance, the position is randomly generated for the worst frog by the following formula:

$$X_w^{(new)} = X_{\min} + rand(\ ).(X_{\max} - X_{\min}) \tag{10}$$

Where X_{max} and X_{min} represent the maximum and minimum search range.

After finishing the local search processes of *m* memeplexes, the memeplexes are shuffled to enhance the exchange of global information.

IV. HYBRID SFLA-BP AMGORITHM

In the local search of the standard SFLA, the possible new position of the worst frog is limited to the range between the current position and the best position (X_b or X_g) [12]. At each iteration, the worst frog attempts to change its position in order to become closer to the best frog [12],[13].

When the difference in position between the worst frog X_w and the best frogs (X_b or X_g) becomes small, the change in the position of X_w frog will be very small, and thus it might stagnate at a local optimum, and decreases the convergence velocity [12], [13].

To expand the possible search range, a modified form of the SFLA has been introduced in [13]. The authors add an acceleration factor C into the formulation of the original algorithm, which is named MSFLA in this study. It is noted that C cannot be too small, or the success rate for finding the global optimum decreases [13], and it cannot be too large, or the frogs make large leaps in the search space without finding a promising location [13]. The author proposed that C is set to 1.3-2.1.

Based on the above reasons, this paper presents a new method to accelerate the convergence and improve the stability and global search ability by combining shuffled frog leaping algorithm with BP algorithm which allows to explore the ability of SFLA and to exploit ability of BP. The flowchart of the proposed hybrid SFLA-BP in local search is illustrated in Figure 3. To introduce BP into SFLA, BP operation process can be added into every sub-memeplex each iterative operation in SFLA.

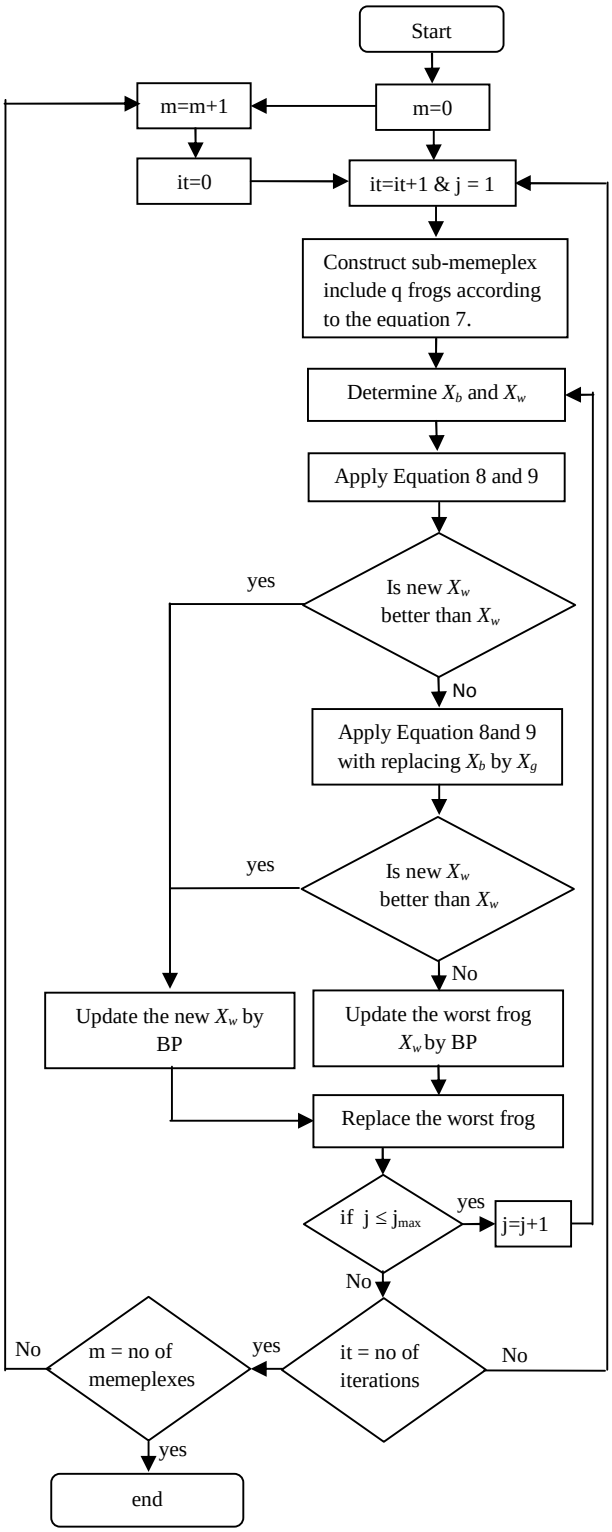


Fig. 3 Flowchart of Local Search for SFLA-BP

In every sub-memeplex, X_w is first updated with X_b. If this process produces a better solution, X_w is replaced by new_X_w. This latter is updated again by BP algorithm using Equations 5 and 6. Otherwise, X_b is replaced by X_g, in this case, if the fitness is improved, then X_w is replaced by new_X_w and updated new_X_w with BP algorithm. On the other hand, if X_w cannot be improved by X_b or X_g, in this case, the worst frog X_w is updated with BP algorithm using Equations 5 and 6.

For SFLA each frog represents all weights of an ANN structure. For example, for ANN with the structure of 4-2-1, the corresponding encoding style for each frog can be represented as:

$$Frog(i)=[w_{51}\ w_{52}\ w_{53}\ w_{54}\ w_{61}\ w_{62}\ w_{63}\ w_{64}\ w_{75}\ w_{76}] \quad (11)$$

However, when calculating the output of the ANN, we need to decode each frog into weights matrix. The encoding strategy can be written as:

$$w^{(1)}=\begin{bmatrix} w_{51} & w_{52} & w_{53} & w_{54} \\ w_{61} & w_{62} & w_{63} & w_{64} \end{bmatrix} \quad (12)$$

$$w^{(2)}=\begin{bmatrix} w_{75} & w_{76} \end{bmatrix} \quad (13)$$

The simulation results will demonstrate that the SFLA-BP algorithm can improve the performance of the standard SFLA algorithm.

V. EXPERIMENTS AND RESULTS

To examine the convergence and effectiveness of the proposed hybrid, SFLA-BP is compared with classical BP and MSFLA presented in the literature [13].The ANN structure containing a single hidden layer of three nodes. Supposed that the input layer has five samples and output layer has 1 node. The hidden transfer function is tangent sigmoid function expressed in Equation (14), and the output transfer function is a linear activation function. Since all algorithms above are used in the same structure.

$$f(x)=\text{tansig}(x)=\frac{2}{(1+\exp(-2.n))-1} \quad (14)$$

The SFLA parameters were defined as: population size P = 25, number of memeplexes m = 5, number of frogs per memeplex n=5. Number of iterations within each memeplex it=5. Number of iterations within each sub-memeplex it=3. For the BP, The learning parameters ¦ = 0.07. The parameters setting of SFLA-BP are equal to SFLA and BP algorithm.

Supposed that every initial frog was a set of weights generated in the range of [-0.95, 0.95], and all thresholds in the network were 0 s.

The digital message applied to the channel is made of uniformly distributed bipolar random numbers {-1, 1}. The channel noise is taken to be additive white Gaussian noise with a signal to noise ratio (SNR) of 20 dB.

Simulations were conducted for the channel with transfer function:

$$H(z)=0.3482+0.8704z^{-1}+0.3482z^{-2} \quad (15)$$

To examine the effect of nonlinearity on the equalizer performance, we consider a nonlinear channel models composed of a linear channel followed by a nonlinearity defined by:

$$z(n) = y(n) + 0.1y^2(n) + 0.05y^3(n) + v(n) \quad (16)$$

Where *y(n)* is output of the linear part of the channel, *z(n)* is the nonlinear channel output and *v(n)* is the additive Gaussian noise.

During this part of the simulations, the performance measure is obtained through the use of signal constellations, eye diagrams, learning curves and BER curves. For MSFLA, we have chosen C = 1.7 because it shows the best results.

Signal constellations and eye diagrams are plotted in figures 4, 5, 6, 7, 8 and 9. These figures are determined by taking 10 iterations and with a sequence which has a length of 500 symbols.

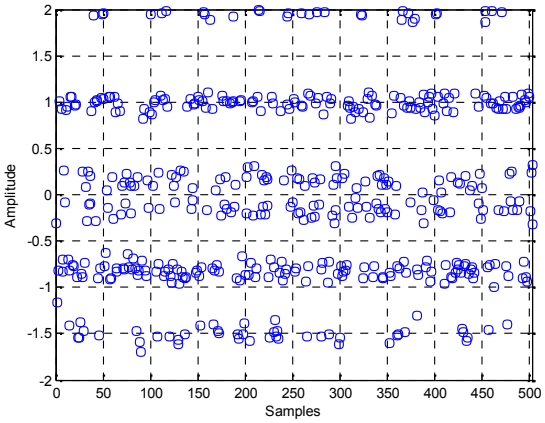


Fig.4 Signal constellation at the input of the equalizer.

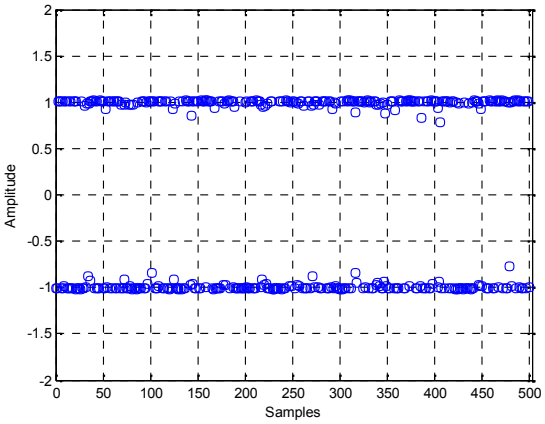


Fig.5 Signal constellation after equalization for SFLA-BP algorithm (Number of iterations = 10).

Fig. 4 shows the unequalized data, while Fig. 5 shows the equalized version obtained through the use of the proposed algorithm. So, the equalizer performance is much improved and the symbols are seen to converge closer to their original positions.

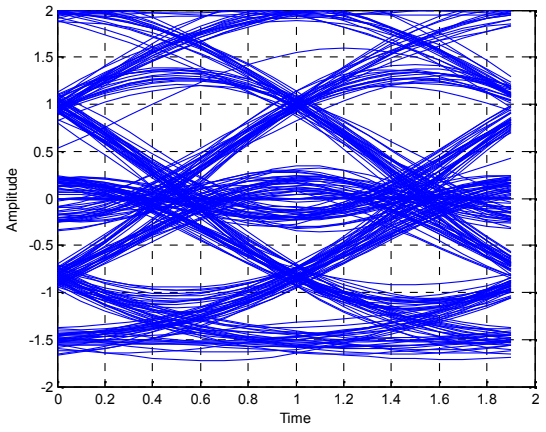


Fig.6 Eye diagram at the input of the equalizer.

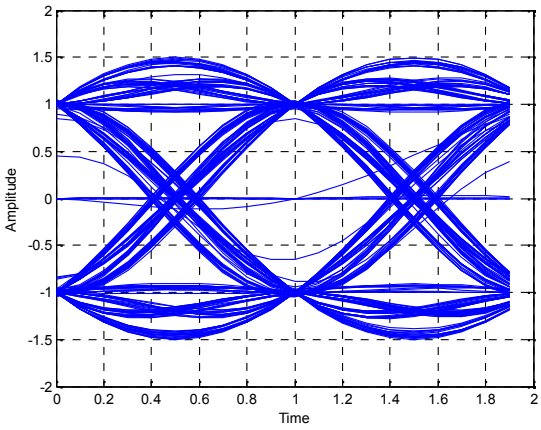


Fig.7 Eye diagram after equalization for SFLA-BP algorithm (Number of iterations = 10).

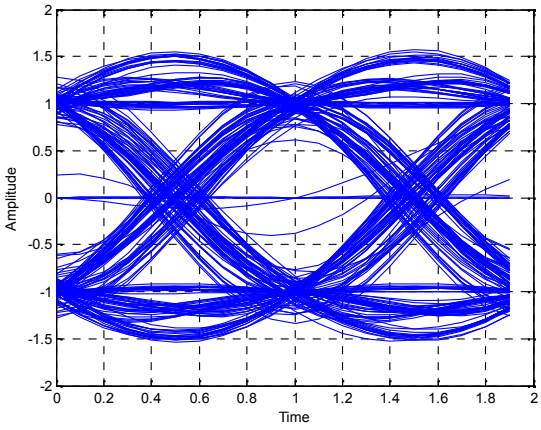


Fig.8 Eye diagram after equalization for MSFLA (Number of iterations = 10 and C = 1.7).

Figures 6, 7, 8 and 9 represent the eye-diagram of the equalizer input and output signals. Fig.6 shows the input signals of the equalizers distorted by the effect of the channel and noise. Whereas, the other figures show respectively the outputs of the equalizers using SFLA-BP, MSFLA and BP algorithms. It is clearly seen from these figures that the output signals of the proposed equalizer are the best one.

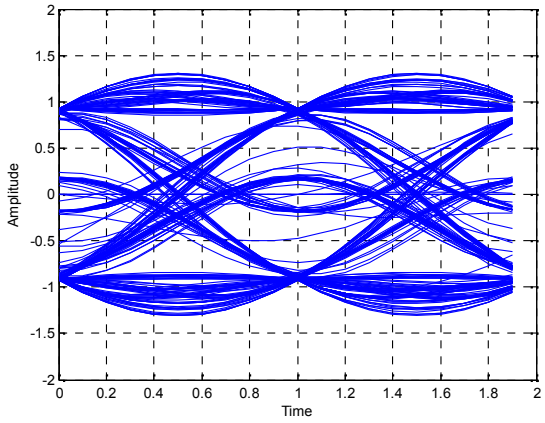


Fig.9 Eye diagram after equalization for BP algorithm (Number of iterations = 10).

The mean squared error (MSE) is one of the most useful measures for the performance of an equalizer. Here, the MSE performance is determined by taking an average of 10 individual runs, each of which involves a different random sequence and has a length of 500 symbols. All the symbols are used for training in the evaluation of the MSE. Fig.10 shows the convergence curves of the three algorithms, these curves shows a clear improvement in both the convergence time and the steady state MSE when using the SFLA-BP algorithm. It is also clearly shown that SFLA-BP algorithm performs better than MSFLA and BP algorithm.

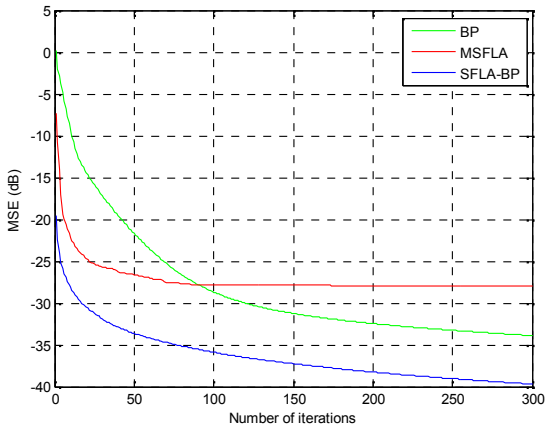


Fig.10 Learning curves for the different algorithms.

Bit error rate (BER) is another performance measure popularly used in channel equalization and is defined as:

$$\text{BER}=\frac{\text{Number of error bit at the equalizer output}}{\text{Total number of bits sent}} \quad (17)$$

Fig.11 shows the BER performance of the equalizer with different algorithms. The BER performance is determined by taking an average of 10 individual runs, each of which involves a different random sequence and has a length of 500 symbols. The number of iteration used for each individual runs is 500 iterations. The first 200 iterations are used for training and the rest which involves a different random sequence for each iteration are used for testing. The results indicate that the SFLA-BP algorithm yields a much lower BER than the MSFLA and BP algorithm

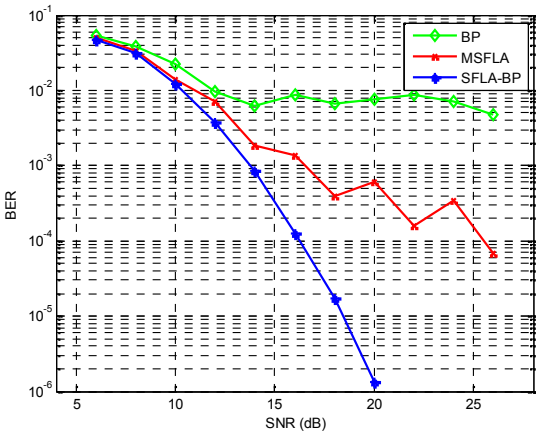


Fig.11 BER curves of different algorithms.

VI. CONCLUSION

In this paper, we have proposed a hybrid training algorithm for the problem of channel equalization, which introduces BP to SFLA to train the weights of artificial neural network (ANN). The BP algorithm is used to update the position of the worst frog in order to accelerate the standard SFL algorithm. This strategy enables to speed up convergence by adding a search learning coefficient that pulls the worst frog faster toward the best solution. The experimental results have shown that SFLA-BP has superior performance in comparison with MSFLA and standard BP algorithm.

References

[1] S. Qureshi. Adaptive Equalization, *Proceeding of IEEE*, Vol.73, No.9, 1985, pp. 1349-1387.

[2] S. K. Padhy, S. P. Panigrahi, P. K. Patra. A Novel Channel Equalizer Using Large Margin Algebraic Perceptron Network, *Journal of Communications*, Vol.5, No.8, 2010, pp. 637-645.

[3] M. R. Peyghami, R. Khanduzi. Predictability and forecasting automotive price based on a hybrid train algorithm of MLP neural network, *Neural Computing and Applications*, Vol.21,No.1, 2012, pp. 125-132.

[4] R. B. C. Prudencio, T. B. Ludermir. *Between Data Science and Applied Data Analysis*, Springer Berlin Heidelberg, 2003.

[5] M. M. Eusuff, K. E. Lansey. Optimization of Water Distribution Network Design Using the Shuffled Frog Leaping Algorithm, *Journal of Water Resources Planning and Management*, Vol.129, No.3, 2003, pp. 210-225.

[6] D. Rumelhart, G. Hinton, R. Williams, Learning representations by back-propagation errors, *Nature*, 323, 1986, pp. 533–536.

[7] R. S. Scalero, N. Tepedelenlioglu, A fast new algorithm for training feedforward neural networks, *IEEE Transactions on Signal Processing*, Vol.40, No.1, 1992, pp. 202-210.

[8] E. Alba, J. F Chicano. *Genetic and Evolutionary Computation*, Springer Berlin Heidelberg, 2004.

[9] D. Tang, Y. Cai, J. Zhao. An Improved Shuffled Frog Leaping Algorithm with Single Step Search Strategy and Interactive Learning Rule for Continuous Optimization, *Journal of Computers*, Vol.9, No.6, 2014, pp. 1300-1308.

[10] A. Tripathi, T. K. Sharma, V. Singh. Bespoke Shuffled Frog Leaping Algorithm and its Engineering Applications, *I.J. Intelligent Systems and Applications*, Vol.7,No.4, 2015, pp. 41-46.

[11] M. Eusuff. K. Lansey. F. Pasha. Shuffled frog-leaping algorithm: a memetic meta-heuristic for discrete optimization, *Engineering Optimization*, Vol.38,No.2, 2006, pp. 129-154.

[12] J. Zhao, L. Lv. Two-Phases Learning Shuffled Frog Leaping Algorithm, *International Journal of Hybrid Information Technology*, Vol.8,No.5, 2015, pp. 195-206.

[13] E. Elbeltagi, T. Hegazy, D. Grierson. A modified shuffled frog-leaping optimization algorithm: applications to project management, *Structure and Infrastructure Engineering*, Vol.3,No.1, 2007, pp. 53-60.

Electromagnetic Coupling of Band Stop Filter and CSRRs for Biosensors Applications in Medical

1st Mr Abdelkader Serhane
phD,Networks and telecommunications
University of Mascara
Laboratory E.P.O, University of Sidi-Bel-Abbes, Algeria
abdelkader.serhane@univ-mascara.dz

2nd Dr. Mohammed Berka
Professor, Department of electrical engineering, University of Mascara, Laboratory E.P.O, University of Sidi-Bel-Abbes, Algeria
Email:mohammed_76_berka@yahoo.fr

3rd Pr. Zoubir Mahdjoub
Professor, Department of Electronics, University of Sidi-Bel-Abbes
Laboratory E.P.O, University of Sidi-Bel-Abbes, Algeria
Email:mahdjoubz@yahoo.com

Abstract—**in this paper, we design and simulate a low-cost Biosensor in the microwave ranges capable of detecting solids and liquids of unknown physical properties.**
The proposed biosensor is composed of a split ring resonator (SRR) of modified polygonal shape with two rings in a Symmetrical fashion of an electromagnetic behavior of the Band-stop-filter and the Complementary split ring resonator (CSRR) Circular cells etched in the ground plane.
This work is a contribution that can subsequently design a new Class of sensors which are miniaturized microwave biosensors.

The simulation was performed using Ansoft HFSS (High Frequency Structure Simulator)

Keywords— ***Band-Stop-Filte, Metamaterial, Biosensor, Complementary split ring resonator (CSRR), HFSS.***

I. INTRODUCTION

Metamaterial-based biosensing technologies have attracted considerable attention from microwave to optical frequency due to their cost-effective, label-free biomolecule detection and cost-effectiveness. [1].

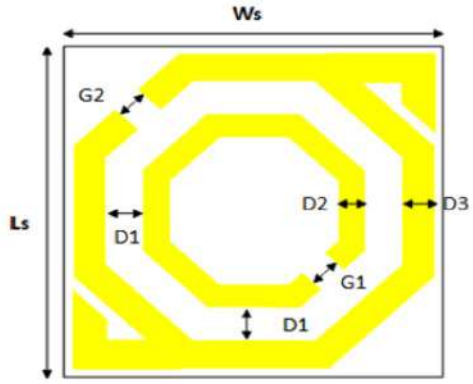
In recent years, various biosensors have been designed using structures based on metamaterials and with technical definitions of detection (Extraction of effective parameters of metamaterials-Change in dielectric constant -change in resonance frequency-modulation of coupling.. ...),

In this work, we will propose a new class of sensors which are the miniaturized microwave biosensors capable of detecting solids and liquids of unknown physical properties. This genre of biosensors is based on the effect of electromagnetic coupling between microwave filters band stops and metamaterial resonators CSRR .
Desired results can subsequently offer factors of qualities and also high sensitivities for our biosensors,

II. THE PROPOSED BAND-STOP FILTERE

The Design and dimensions of the modified polygon are shown in Figure 1.
The prospective composition of our proposed model is based on two polygonal shaped rings of 8 segments spaced by D1. The outer ring is D3 wide, with a G2 wide opening in the eighth segment. The inner ring is D2 wide, with a G1 wide opening in the fourth segment.

(a)



(b)

parameters	Value mm
W_S	5.8
L_S	5.8
G_1	0.5
G_2	0.6
D1	0.4
D2	0.5
D3	0.6

Figure 1: (a) Design of BSF (b) geometric parameters

The structure is simulated for the frequency band [1-10] GHz; the substrate used for the simulation is Rogers RO3003 which has a relative permittivity of 3, tangential losses of around 0.001 and a thickness of 1.52 mm ,and this substrate use for all simulations.

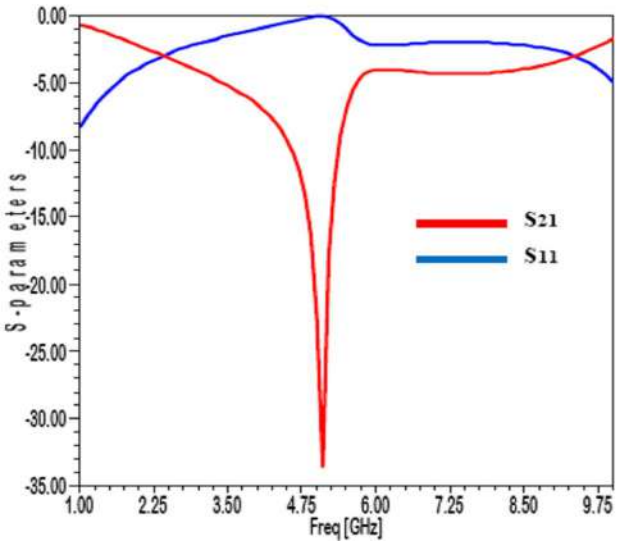


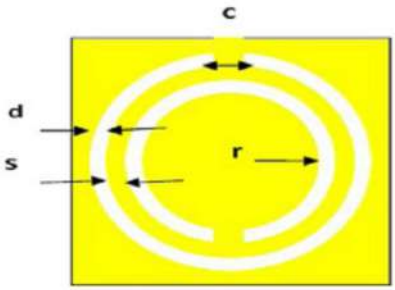
Figure 2: S-parameters results of band stop filter

Figure 2 shows that the polygonal cell and an electromagnetic behavior of the band stop filter in the band [4.4 GHz – 5 GHz] with very low efficiency and insertion losses.
and must be very narrow, to see its displacement in dependence with the materials or the liquid under test. The band-stop filters thus constituting the proposed biosensors make it possible to eliminate the undesirable harmonics.

III. BIOSENSOR DESIGN CONCEPT

The proposed biosensor contains band-stop filter on the upper side and (CSRR) Circular cells etched in the ground plane.

(a)



(b)

parameters	Value mm
d	0.15
c	0.15
r	0.85
s	0.2

Figure 3: (a) (CSRR) Circular cells
(b) Geometric Parameters

The lated
for tl CSRR
and 4CSRRs .

The characteristics of CSRRs have been studied by several searchers [2].The orthogonal electric field excites the CSRR that is comparable to an electric dipole which is placed along the axis of the ring.
The coupling between CSRR and band stop filter capacitance coupling [3], through the ring slot and the split of the outer ring produces magnetic coupling. The retrieved results from S-parameters shown in Figure 4 .

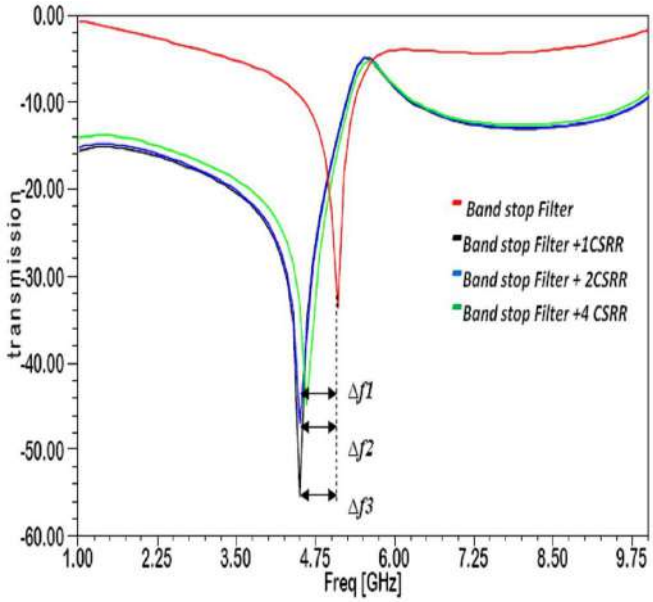


Figure 4: Numbers CSRRs are etched on the ground plane

The results of the simulation in Figure 4 show resonance frequency shifts because of the Presence CSRR in structure Every time we add CSRR the resonance frequency is increasing The quality factor for general resonators, Q can be written [4]:

$$Q = \Delta f_r / f_r$$

The resonance frequency of the proposed biosensors can be shifted to lower frequencies due to interaction with the material under test (MUT) which is the basic principle of microwave biosensors. At the resonance frequency of the biosensors, the stored electric field (E0) and magnetic field (H0) are equal to each other.

When a MUT interacts with the biosensors, it disrupts the Equilibrium of the stored electromagnetic fields and generates the new electric field (E_1) and magnetic field (H_1) which causes it to change the resonance frequency.

This change in resonance frequency (Δfr) depends on the change in permittivity (Δε), permeability (Δμ) and volume (v) of the MUT, which can be expressed mathematically as given in reference [5]:

$$\frac{\Delta f_r}{f_r} = \frac{\int_v (\Delta\varepsilon E_1.E_0 + \Delta\mu H_1.H_0)dv}{\int_v (\Delta\varepsilon_0 |E_0|^2 + \Delta\mu_0 |H_0|^2)dv}$$

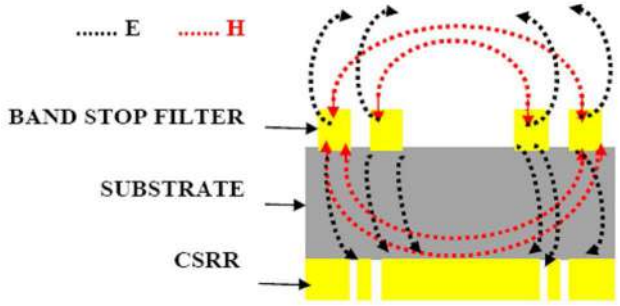


Figure 5: Cross section of band stop filter with a single CSRR and a schematic electromagnetic field distribution.

In order to study the capacity of our biosensor for the detection of biological or chemical species, we deposited a thin layer of polyamide of variable thickness, having an electrical permittivity of 3.5, on the surface band stop filter.

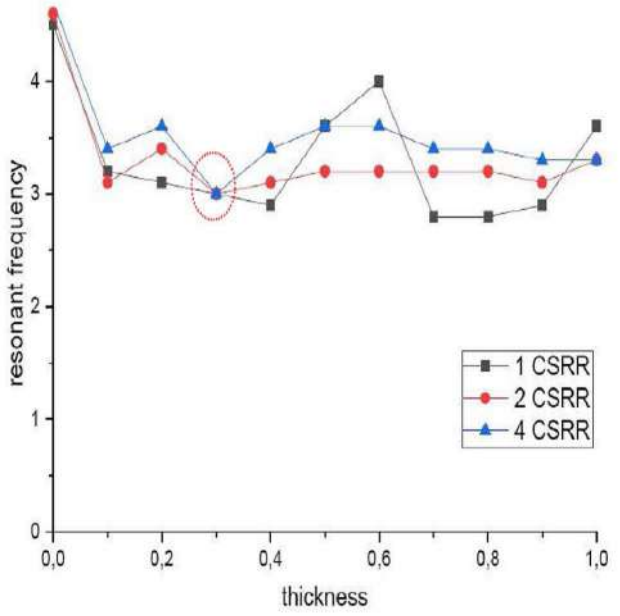


Figure 6: the resonance frequency as a function of the thickness

Have resonance frequency stability for d=0.3 the miniaturization of electronic devices is a major challenge in the context of microwave telecommunications.

In this proposed structures, we are interested in using a single CSRR .

The next step in the design of the biosensor The sensitivity,

In order to calculate the sensitivity, we varied the relative permittivity of the polyimide layer from εr = 1 to 3.5 for the thickness: d = 0, 3

TABLE I. EFFECT OF RELATIVE PERMITTIVITY OF MUT ON RESONANCE FREQUENCY

Material Under Test (MUT)				Resonance Frequency	
εr	μr	$tan\delta e$	$tan\delta m$	*GHz*	*dB*
1	1	0.008	0	4.62	-50.72
1.5	1	0.008	0	4.25	-42.79
2	1	0.008	0	3.75	-48.81
2.5	1	0.008	0	3.65	-36.20
3.5	1	0.008	0	3.50	-42.12

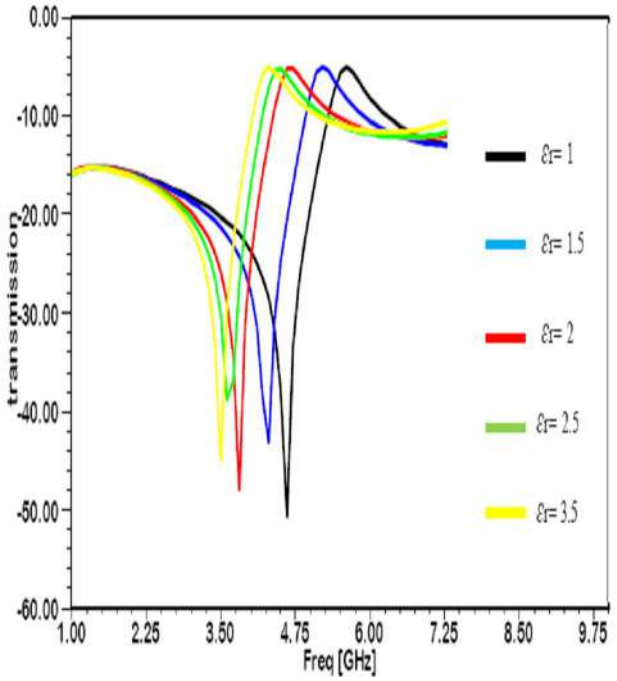


Figure 7: Simulated transmission coefficient (S21) of the biosensor due to interaction with the MUT with different values of permittivity (εr).

From the figure 7, it can be seen that the relative permittivity is inversely proportional to the resonance frequency when the permittivity of MUT increases, the resonance frequency decreases due to the increase in the total capacity of the biosensor

where fr , MUT is the resonance frequency of the sensor due to interaction with the MUT with relative permittivity εr. The relative sensitivity of various state of the art sensors [66–70] is tabulated in Table 2.

TABLE II. TABLE COMPARISON OF FREQUENCY SHIFT WITH VARIOUS STATE OF ART SENSORS.

Ref	Resonating Structure	Resonance Frequency (GHz)	Permittivity Range Studied	Shift in Resonance Frequency
[6]	SIR	1.91	10–80	38%
[7]	CSRR	2.4	10–80	36%
[8]	SRR	1.72	10–80	34%
[9]	Open-Loop Resonators	2.6	10–140	36%
[10]	SRR	1.8	2.42–22.52	38%
[11]	CSRR	2.5	2.93–3.64	50%
[12]	CSSSR	15.12	2.1–3	50%
This Work	BSF+Single CSRR	3.5–4.62	1–3.5	60%

The analysis of biosensor is done in microwave region.Small change in dielectric constant produces shift in resonant peak of CSRR this biosensor can be used for the detection of different types of cancer cells like Hela, PC12, MDA-MB-231, MCF-7, Jurkat. The advantages of this biosensor are low cost manufacturing and level free sensing with appropriate sensitivity [13].

TABLE III. DIALECTRIC CONSTANT OF NORMAL CELL AND DIFFERENT CANCER CELL

Cancer cell	Dielectric constant
Normal cell	1.822500
Hela	1.937660
PC12	1.946025
MDA-MB-231	1.957201
MCF-7	1.962801
Jurkat	1.932100

CONCLUSION AND PERSPECTIVES

This paper presents an extremely sensitive microwave biosensor that is based on a band stop filter and Complementary split ring resonator (CSRR),

The proposed bisensor is sensitive for the parameters of MUT (εr) which is demonstrated by electromagnetic simulation.

It is concluded that the fringe electromagnetic fields of biosensor interact with the MUT and real relative permittivity cause a decrease in the resonance frequency

Sensitivity for different cancer cells are achieved with only which make the propose biosensor cost effective so it may be used for different cancer cell detection in biosensing application .

REFERENCES

[1] Tao Chen, Suyan Li ; Hui Sun. "Metamaterials Application in Sensing", Sensors Journal , 2012.

[2] KUMAR, D. Shashi; SUGANTHI, S. "Miniaturization of Microstrip Antenna with Enhanced Gain Using Defected Ground Structures. " International Conference on Data Science and Communication (IconDSC). IEEE, p. 1-5 , 2019.

[3] RAMBE, Ali Hanafiah, et al. " Miniaturization of rectangular patch microstrip antenna by using complementary split ring resonator. In: IOP Conference Series: Materials Science and Engineering. IOP Publishing,. p. 012132, 2018.

[4] Salim, A; Lim, S. "Review of Recent Metamaterial Microfluidic Sensors". Sensors, 2018.

[5] Pozar, D.M. " Microwave Engineering; John Wiley & Sons, Inc.: Hoboken, NJ, USA; ISBN 97898965408213, 2012.

[6] Ebrahimi, A.; Scott, J.; Ghorbani, K. "Ultrahigh-Sensitivity Microwave Sensor for Microfluidic Complex Permittivity Measurement" . IEEE Trans. Microw. Theory Tech., 67, 4269–4277 , 2019. [CrossRef]

[7] Ebrahimi, A.;Withayachumnankul,W.; Al-Sarawi, S.; Aboott, D. "High-Sensitivity Metamaterial-Inspired Sensor". IEEE Sens. J. 2014, 14, 1345–1351 , 2014 .[CrossRef]

[8] Rowe, D.J.; Abduljabar, A.A.; Porch, A.; Barrow, D.A.; Allender, C.J. "Improved Split-Ring Resonator for Micro Fluidic Sensing". IEEE Trans. Microw. Theory Tech., 62, 689–700 , 2014 .[CrossRef]

[9] Abdolrazzaghi, M.; Daneshmand, M.; Iyer, A.K. Strongly Enhanced Sensitivity in Planar Microwave Sensors Based on Metamaterial Coupling". IEEE Trans. Microw. Theory Tech., 66, 1843–1855 , 2018. [CrossRef]

[10] Galindo-Romera, G.; Javier Herraiz-Martínez, F.; Gil, M.; Martínez-Martínez, J.J.; Segovia-Vargas, D. "Submersible Printed Split-Ring Resonator-Based Sensor for Thin-Film Detection and Permittivity Characterization". IEEE Sens. J, 16, 3587–3596, 2016 [CrossRef]

[11] Su, L.; Mata-Contreras, J.; Vélez, P.; Fernández-Prieto, A.; Martín, F. "Analytical method to estimate the complex permittivity of oil samples". Sensors, 18, 984. 2018 [CrossRef]

[12] Tanveerul Haq; Cunjun Ruan; Xingyun Zhang; Shahid Ullah. "Extremely Sensitive Microwave Sensor for Evaluation of Dielectric Characteristics of Low-Permittivity Materials". Journal Sensors 20, 20, 1916; doi:10.3390/s20071916.2020

[13] Ankit Singh,"Split ring resonator biosensor-an innovative design and analysis" , IEEE 8th International Conference on Photonics (ICP) , 2020.

Blind and robust image watermarking algorithm Based on DWT-DCT and Edge Insertion

Razika Souadek, Nassir Guellil
Department of Electronic,
LIS laboratory,
University of Setif 19000,
Republic Algeria.
razikaese@gmail.com

***Abstract*— Blind and robust watermarking image algorithm is implanted in this paper, we assemble two transforms discrete wavelet transform (DWT) and discrete cosines transform (DCT) accompanied with Edge insertion to obtain an efficiency algorithm. We embedded the binary watermark image in the coefficients of DCT after applying the first level of DWT. The component used to insert the watermark is HH. The advantage of hybrid technique gives more robustness against various attacks, and the watermarked image achieves a good quality in term invisibility of the watermark. The proposed algorithm detailed step by step all instructions are used and the experimental results determine the efficiency of our algorithm. The evaluation metrics used in this algorithm are the parameters PSNR (Peak Signal to Noise Ratio) and NC (Normalized Correlation).**

***Keywords*— watermarking image, Discrete Wavelet Transform, and Discrete cosine transform.**

I. INTRODUCTION

The world becomes a small town with the evolution of communication technology as service web and multimedia applications. However, it gives rise to problem of security and authentication of digital multimedia contents. Therefore, it is important to employ algorithms that provide enough securities for multimedia contents while providing protection against piracy and duplicate of the original work. Nowadays, the most researchers are interests to domain of watermarking image where here motive are the authentication [1].

The watermarking image is the operation to embed information called the watermark image into original image such that this watermark can be extracted to later [2, 3]. The level of the watermarking scheme decomposed into two domains, the first the spatial domain and the second the frequency domain. In the spatial domain the watermark is embedded in the cover image directly in the LSB bit (Last Significant Bit) [4, 5], it is simple and rapid. However, the major disadvantage of this process is the fragility against geometrical and image processing attacks [6]. In contrast, the frequency domain is embedded the watermark image in the frequency coefficients of original image [7]. The known frequency transforms are: Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT) and Discrete Fourier Transform (DFT). Consequently, the uses of these transforms are given more robustness against various attacks, invisibility of watermark and augment the capacity of insertion. The design of an effective watermarking scheme determined by response to the three exigencies among following: invisibility, robustness and capacity of insertion. For this reason the modelling mathematical in the watermarking image algorithm has come to give more of robustness like as the technique of the singular value decomposition (SVD) [8]; the researchers have combined the SVD with the others transforms well-known to create hybrid techniques. The same algorithms have given perfect invisibility and higher robustness as noted in [9], where Rajesh Mehta and al. have proposed the algorithm of watermarking image based on DWT-SVD; in this scheme the authors embed the watermark image in the component LL of DWT transform done the SVD technique. Musrrat Ali and al. in [10] have proposed the watermarking scheme in DWT-SVD domain using the differential evolution technique (DE). The step of insertion determine in all sub-band of DWT transform of one level done in the singular value decomposition with a secret key calculated by the DE, it is different for each sub-band. A. Manjunath and al. [11] presented the method of watermarking image based on the stationary wavelet transform (SWT), this system performed the high frequency component are omitted, and the low frequency are kept with few alternations in the stage of watermark insertion. O. jane and al. in [12] explained in their paper a robust watermarking method based on the DWT-SVD, the LU decomposition are used to obtain a perfect watermarking algorithm ensure the robustness. In addition, Hung-Hsu Tsai and al. in [13] proposed the blind watermarking image algorithm based on hybrid technique DWT-SVD done by SVR and PSO technique to embed the watermark image. The proposed algorithm in [14] based on DWT of two levels in order to compact a higher energy in component LL1, and Contrast Sensitivity Function (CSF) to improve the invisibility and robustness, the Function of Pixel Movement (PMF) is applied to increase the security properties. The experimental results give a good robustness against various attacks.

The proposed algorithm is based on the hybrid technique DWT-DCT and Edge insertion; we embed the binary watermark in the sub band HH of the DWT done by the DCT

transform and the key k. where the pixel bit of watermark image equal to one we insert this bit multiplying to k, else we subtract the bit one multiplying to k. The Robustness of the watermarking algorithm is demonstrated by the Normalized Correlation (NC) between watermark image and extracted watermark and the Peak Signal to Noise Ratio (PSNR) measured between the cover image and the watermarked image.

The rest of this paper is organized as follows: description of transform domain presented in section two. The proposed algorithm is detailed in section three. The experimental results and their discussion are considered in section four. Finally, we conclude by a conclusion.

II. TRANSFORMS DOMAIN

A. Discrete Wavelet Transform

The DWT is compacted to a high energy and used in various domains well known. This theory based on separate the frequency into four sub-bands, a lower resolution approximation component (LL) and three other corresponding to horizontal (HL), vertical (LH) and diagonal (HH) detail component. The LL sub-band is the result of low-pass filtering and contains major information of the cover image. The HH sub-band component is high pass filtering and contains the high frequency along the diagonal. The HL and LH of the image are the result of low pass filtering and high pass filtering in one direction, vertical or horizontal. The most information of cover image concentrates in sub-band LL and the details of the image demonstrated in the component LH and HL [15]. Lastly, the LL sub-band can be decomposed to obtain another level as it shown in figure 1.

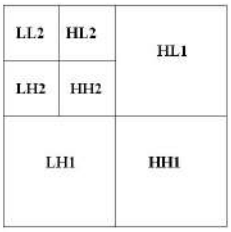


Fig. 1 Two levels of discrete wavelet transform.

B. Discrete cosine transform

The DCT transform is a form of signal decomposition that converts images from spatial domain to frequency domain [16]. The equation of the discrete cosines transform of two dimensions is presented as following:

The 2-D DCT for a block of 8* 8 pixels is given by (1) [16, 17].

$$F(u,v) = \frac{k(u)k(v)}{4}\sum\sum x(i,j)\cos\left(\frac{(2i+1)u\pi}{16}\right)\cos\left(\frac{(2j+1)v\pi}{16}\right)$$

$$k(u), k(v) = \begin{cases} \frac{1}{\sqrt{2}}, & if\ u,v = 0 \\ 1, & otherwise \end{cases}$$

Where x(i,j) is the intensity value of pixel in location of (i,j). Therefore, we have [17, 18].

$$F(0,0) = \frac{1}{8}\sum\sum x(i,j) \quad (2)$$

In the second formula shown, the coefficient $F(0,0)$ of the DCT transform is measured by the average value of the 8×8 pixels of one block of the image and is used for the extraction of colour data [19]. The rest pixels are show the frequency feature of the image where are called AC [20].

III. PROPOSED METHOD

In this section, the proposed watermarking algorithm will present. The watermark image embedded into the sub-band HH of DWT transform done by DCT transform and key k in the original image. The watermarking algorithm evaluated by using the following metrics PSNR (Eq. 3) and NC (Eq. 4). These parameters help us to the evaluation process in term robustness against the attacks and invisibility of watermark.

$$PSNR = 10\ log_{10}\left(\frac{(I_{max})^2}{1/(m\times n)\sum_i\sum_j (I(i,j) - \mathrm{I_w}(i,j))^2}\right) \quad (3)$$

$$NC(w,\mathrm{w}') = \frac{\sum_i\sum_j w(i,j)\ \mathrm{w}'(i,j)}{\sqrt{\sum_i\sum_j w(i,j)^2}\sqrt{\sum_i\sum_j \mathrm{w}'(i,j)^2}} \quad (4)$$

Where I, I_w present original and watermarked images respectively.

Where w, w' present original and extracted watermark respectively.

B. Embedded process

To embed the watermark image w of size 32*32 in the cover image I of size 512*512 pixels, we used the embedding algorithm summarized in the following steps:

- Divide the host image to the following sub bands done by the first level of DWT: {LL, HL, LH, HH}
- Choose the sub band HH and divided into blocks B_i, each block of size 8*8; then apply DCT transform to each block to obtain a new block called (B_i^*).

$$B_i^* = T * B_i * T' \quad (5)$$

- Embed the watermark w into the blocks B_i^*

$$\begin{cases} if\ w(i,j) = 1\ then \\ B_w = B_i^* + k \times\ ones(8) \\ else \\ B_w = B_i^* -\ k \times ones(8) \\ end \end{cases} \quad (6)$$

Where, B_w represents the blocks of component watermarked HH_w, and k represents the key.

- Apply the IDCT transform to B_w as following:

$$B_w^* = T' * B_w * T \quad (7)$$

- Apply the first level of IDWT transform to HH_w to obtain the watermarked image.

C. *Extracted process*

We use the watermarked image for extracted the watermark, the extraction process explained as following:

- Divide the watermarked image into sub-bands by done the first level of DWT
- Choose the component HH_w.
- Divide the component into block B_w^*, each block of size 8*8; then apply the DCT transform to each block to obtain B_w.

$$B_w = T * B_w^* * T' \quad (8)$$

- Extract the watermark image w by using the equation (9, 10)

$$S = \sum_i \sum_j B_w \quad (9)$$

$$w(i,j) = \begin{cases} if\ S > 0 \\ w(i,j) = 1 \\ else \\ if\ S < 0 \\ w(i,j) = 0 \end{cases} \quad (10)$$

IV. Result and Discussion

In this section, we use the images Lena and Airplane with size 512*512 pixels to test our algorithm and we use a watermark image of size 32*32 pixels. Figure two represent the original image and watermarked image (WI) Lena with PSNR equal to 43.939 dB and the normalized correlation value NC=1.

(a) (b) (c) (d)

Fig. 2: (a) original image Lena, (b) watermark image, (c) watermarked image (WI) Lena with PSNR 38.9153 dB, (d) extracted watermark (EW) with NC =1.

We validate our proposed method of watermarking image by several experiments to ensure the robustness. For this reason we use the various attacks following in Figure 3 (a) Salt and pepper (SP) with 1% variance, (b) Gaussian noise (GN) with zero mean and 1% variance, (c) Median filtering (MF) using 3×3 pixel's neighborhood, (d) Average filtering (AF) using 3×3 pixel's neighborhood, (e) sharpening filter (SH), (f) Gaussian low-pass filtering (GF), (j) Gamma correction of 0.6 (GC) with 0.6, (h) JPEG compression(JPEG) with quality factor 60.

The proposed method examined against different attacks to provide her resistance against them (Figure 3 (a) to (h)). Our experiments indicate that the proposed watermarking scheme exhibits high imperceptibility and good robustness. The parameter NC is maintained the good level of similarity between watermark before and after the attack.

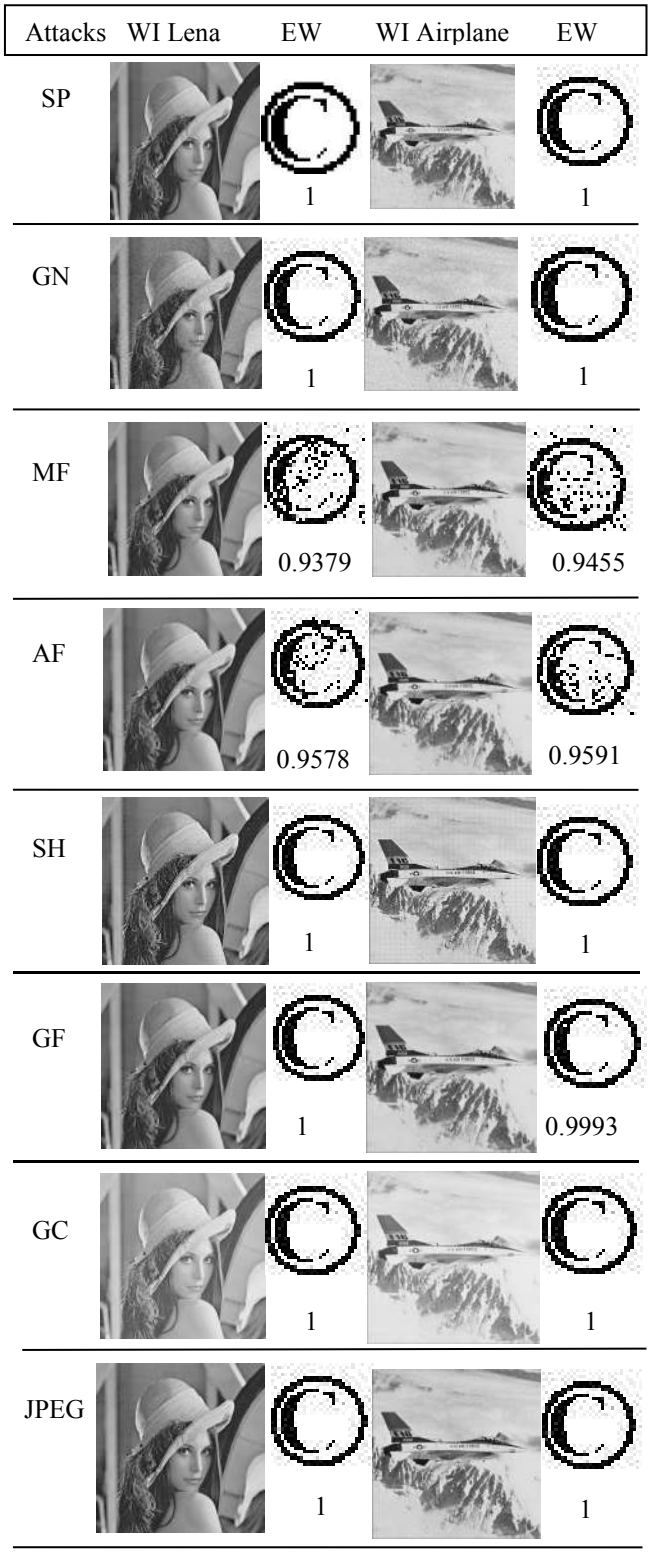


Fig. 3 : Distorted watermarked image (WI) Lena, Airplane and their corresponding extracted watermarks (EW) after attacks indicating NC value, (a) Salt and pepper, (b) Gaussian noise, (c) Median filtering, (d) Average filtering, (e) sharpening filter, (f) Gaussian low-pass filtering, (j) Gamma correction, (h) JPEG compression.

The table one discusses the comparison of our algorithm against anther published work. The obtained results proved the robustness against various attacks with a good NC values compared with the results in [21].

attacks	Proposed Algorithm in [21]		Our algorithm	
	Lena	Airplane	Lena	Airplane
SP 5%	1	1	0.9974	0.9940
BG 5%	0.9917	0.9867	0.9921	0.9888
MF 3*3	0.9967	0.9980	0.9379	0.9455
AF 3*3	1	1	0.9578	0.9591
SH 0.8	1	1	1	1
GF 3*3	0.9275	0.9190	1	1
GC	1	1	1	1
HE	1	1	1	1
C-JPEG 50%	1	1	1	1

Table 1: The NC values of comparison.

V. CONCLUSION:

The blind watermarking algorithm realized in this paper based on hybrid technique DWT-DCT give a high robustness. The three exigencies of watermarking image are realized as following the robustness of system, invisibility of watermark and capacity of insertion. The value of PSNR given by 38.9153 dB for Lena and Airplane images demonstrate the good quality of the watermarked image compared by the original image. The values of NC for the two watermark images given by 1 indicate the similarity between watermark image and extracted watermark. The mathematics technique of Singular Value Decomposition as perspective for evaluates the robustness of the algorithm.

Reference:

[1] V. Tomar, A.Kumar, A. Choudhary. Conception & Implementation of a novel digital image watermarking algorithm using cascading of DCT and LWT. 2014 International conference on reliability, Optimization and information Technology. ICROIT 2014 Feb 6-8 (2014) 501 – 505.

[2] J.L. Liu, D.C. Lou, M.C. Chang, H.K. Tso. A robust watermarking scheme using self reference image. Computer Standards & Interfaces 28 (2006) 356 – 367.

[3] D.C. Lou, H.K. Tso, J.L. Liu. A copyright protection scheme for digital images using visual cryptography technique. Computer Standards & Interfaces 29 (2007) 125 – 131.

[4] N. Nikolaidis, I. Pitas. Robust image watermarking in the spatial domain. Signal Process, 66 (1998) 385 – 403.

[5] J.-C. Lui, S.-Y.Chen. Fast two-layer image watermarking without referring to the original image and watermark. Image Vis. Comput. 19 (2001) 1083 – 1097.

[6] M. Ali, C.W. Ahn, M. Pant. A robust image watermarking technique using SVD and differential evolution in DCT domain. Optik 125 (2014) 428 – 434.

[7] A. Umaamaheshvari, K. Thanushkodi. Robust Image Watermarking Based on Block Based Error Correction Code. International conference on current trends in engineering and technology, ICCTET'13, (2013) 34 – 40.

[8] R. Liu, T. Tan, An SVD-based watermarking scheme for protecting rightful ownership, IEEE Transactions on multimedia 4 (2002) 121 – 128.

[9] R. Mehta, N. Rajpal. A Hybrid Semi-Blind gray scale image watermarking algorithm based on DWT-SVD using human visual system model. IEEE (2013) 163 – 168. doi:978-1-4799-0192-0

[10] M. Ali, C.W. Ahn. An optimized watermarking technique based on self-adaptative DE in DWT-SVD transform domain. Signal processing 94 (2014) 545 – 556.

[11] Nagarjuna P.V, K. Ranjeet, Robust blind digital image watermarking scheme based on stationary wavelet transform. 2013 IEEE.

[12] O. jane, H.G. Ilk, E. Elbasi. A Secure and robust watermarking algorithm based on combination of DWT, SVD and LU decomposition with Arnold's Cat Map approach. Electrical and Electronics Engineering (ELECO), 2013 8th International Conference on, 306-310.

[13] H.-H. Tsai, Y.-J. jhuang, Y.- Shoulai. An SVD-based image watermarking in wavelet domain using SVR and PSO. Applied soft computing 12 (2012) 2442 – 2453.

[14] R. Souadek, N.E. Boukezzoula, New Image Watermarking Algorithm Based on DWT and Pixel Movement Function PMF. The International Arab Journal of Information Technology, Vol. 17, No. 1 (January) 2020 pp. 1-7.

[15] M. Ali, C. W. Ahn, M. Pant. An Optimized watermarking Technique Based on DE in DWT-SVD domain. IEEE (2013) 99 – 104. doi:978-1-4673-5873-6

[16] Shinfeng D. Lin , Shih-Chieh Shie , J.Y. Guo : "Improving the robustness of DCT-based image watermarking against JPEG compression" Computer Standards & Interfaces 32 (2010) 54–60.

[17] Z. Lu, S. Li, H. Burkhardt, " A content-based image retrieval scheme in JPEG compressed domain," Innovative Computing, Information and Control 2,pp.831–839, 2006

[18] J. Jiang, A. Armstrong, G.C. Feng, "Direct content access and extraction from JPEG compressed images," Pattern Recognition 35, pp. 2511–2519, 2002.

[19] M. Bajaj, J.A. Lay, "Image indexing and retrieval in compressed domain using color clusters," in: IEEE Symposium of Computational Intelligence, Image Signal Processing, pp. 271–274, 2007.

[20] P. Poursistania, H. Nezamabadi-poura, R. Askari Moghadamb, M. Saeeda, "Image indexing and retrieval in JPEG compressed domain based on vector quantization," Elsevier Ltd, pp. 1005-1017, 2011.

[21] Henri Bruno Razafindradina and Attoumani Mohamed Karim. Blind and Robust Images Watermarking Based on Wavelet and Edge Insertion. International Journal on Cryptography and Informaion Security (IJCIS); 23-30, 2013.

On The Comparative Study Between LEACH and VSG-LEACH Routing Protocols

1st TERMECHE Hayet
LIST laboratory
University of Boumerdes
Boumerdes,Algeria
h.termeche@univ-boumerdes.dz

2nd LECHANI Taous
LREEI laboratory
University of Boumerdes
Boumerdes,Algeria
t.lechani@univ-boumerdes.dz

3rd RAHMOUN Faycal
LIST laboratory
University of Boumerdes
Boumerdes,Algeria
faycal.rahmoune@univ-boumerdes.dz

***Abstract*—The aim of this work is to compare the performance of LEACH (Low Energy Adaptive Clustering Hierarchy) routing protocol with a modified version called VSG_LEACH (Virtual Square Grid Low Energy Adaptive Clustering Hierarchy) protocol in order to create a generalized view about these protocols and thus to enhance the performance of power consumption of sensors used in WSNs (Wireless Sensor Networks).**

Keywords—WSNs, LEACH, network lifetime, residual energy

I. INTRODUCTION

Wireless Sensor Network (WSN) has received a lot of attention during the last decades because they cover a wide range of applications such as military, environment, medicine and industry. WSNs are usually used to control unattended zones where the batteries of sensors are unchangeable and not replaceable [1]. WSNs nodes are deployed randomly in specific area to collect information on the environment like temperature, pression, pollution and send it to the base station (BS) [2].

In recent years, researchers work mostly on routing protocols to increase the lifetime of networks. The purpose of routing protocols is the optimal routing of data from source to destination taking into account material constraints and energy consumption [3].

Hierarchical routing protocol is a very important topic that divides sensors into different layers in order to reduce energy consumption. The typical hierarchical routing technique is clustering, where the network is divided into clusters [4]. LEACH grouped nodes into clusters and each cluster selects one node to become a cluster head (CH) and the other nodes become a cluster member [2]. In each cluster, a leader CH ensures the aggregation of data from ordinary nodes (ONs) and transmission of data to the base station (BS) [5].

Among the drawbacks of LEACH is the random generation of the cluster heads and that LEACH deviates from the optimal value of CHs. Various protocols have been developed to elevate this drawback. In this paper we will study one of these protocols called VSG_LEACH and we will compare its performance to the LEACH protocol.

II. PRESENTATION OF LEACH ALGORITHM

LEACH is a protocol developed by Heinzelman et al [6]. It works in rounds; each round is divided into two phases: set-up phase and steady-state phase [7].

A. Set-up phase

To select the CH, each node generates a random number between 0 and 1 if it is different than CH. If this number is less than the threshold value, the node is elected as CH [8]. The threshold value *T(n)* is calculated as follows:

$$T(n) = \begin{cases} \dfrac{P}{1 - P * \left(r \bmod \frac{1}{P}\right)} & : \text{if } n \in G \\ 0: \text{ otherwise} \end{cases} \quad (1)$$

Where :

- P: the percentage of nodes to become CHs.
- r: is the number of rounds that have elapsed.
- G: is the set of nodes that have not elected as CH in the previous 1/P rounds.

Each CH broadcast a message to all nodes, each node joins the CH with the high signal strength of the message. The CH creates a TDMA (Time Division Multiple Access) schedule to indicate to each node exactly which slot use to transmit its data to prevent collisions and increases the network lifetime.

If the CH loses its energy, all the nodes belong to the CH cannot communicate with the rest of the network. LEACH therefore integrates a random rotation of the cluster head position for all cluster members [6].

B. Steady-state phase

In this phase all nodes send data to their CHs according to the TDMA planification to conserve the overall energy consumption of network. The CHs aggregate and forward a data to the base station (BS) in a single hope transmission. If a failure occurs all the data captured by the nodes are lost [9].

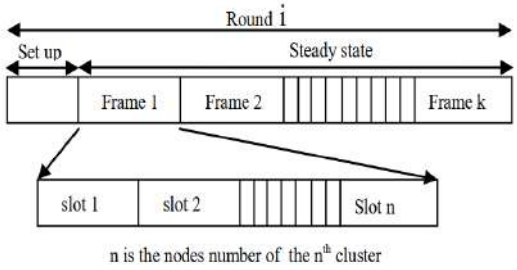


Fig. 1. Round in LEACH

III. Presentation of VSG-LEACH

VSG-LEACH (Virtual Square Grid LEACH) is a protocol developed by Lechani et al [10] in order to eliminate the disadvantages associated with the random formation of CH. VSG-LEACH protocol partitions the area to be monitored into several squares that have the same area. Nodes in the same square form a cluster. The optimal number of CHs is set to 5% of the total number of nodes. Fixing the clusters allows the distribution of CH over the entire monitored area and the energy consumed during cluster formation is conserved.

In VSG-LEACH each round is divided into two phases: set-up phase and steady-state phase.

A. *Set-up phase*

In order to select the CHs a weight W(i) assigned to each node. In each cluster the node having the highest weight is selected to be the CH. The weight value W(i) is calculated as follows:

$$W(i) = \alpha . Energy_{res}(i) + \beta / Dist(i) \quad (2)$$

Where

- $Energy_{res}(i)$: is the residual energy of the node(i).
- Dist(i) : is its distance to the centre of the square.
- α, β **:** are weights assigned to residual Energy and Distance to square centre according to its importance.

B. *Steady-state phase*

LEACH and VSG-LEACH have the same steady state phase. All nodes send data to their CH according to the TDMA schedule. The CHs aggregate and forward a data to the BS.

IV. Results and Discussions

In this section we will perform a series of simulations in order to compare the two protocols LEACH and VSG_LEACH. For this reason, the algorithms have been tested in two different scenarios. In scenario 1, the nodes are deployed randomly in a 100×100 area and BS is placed at the position (50, 175). In scenario 2, the nodes are deployed randomly in 200 × 200 areas and BS is placed at the position (100,275).

Initially all nodes have the same amounts of energy and each node sends one packet per time unit to the CH. The simulation is implemented using MATLAB R2020a, all the parameter used and their values are shown in the TABLE1. The same parameters are used for LEACH and VSG_LEACH algorithms.

TABLE I. The Parameters Used

Parameter	Scenario 1	Scenario 2
Network size	100x100 m	200x200 m
Base station location	(50,175)	(100,275)
Number of nodes	100	100
Data packet size	4000bit	4000bit
Aggregation energy	5nJ	5nJ
Initialy energy of nodes	2J	2J
Percentage of CHs	5%	5%
Maximum number of rounds	10000	10000

Fig. 2, Fig.3 and Fig.4 show the clustering in LEACH and VSG-LEACH, nodes of the same cluster are designed by one color. In LEACH the number of clusters is randomly generated and deviates strongly from the optimal value five (5 clusters in round 1, 7 clusters in round 2) in the same scenario. This can explains the load imbalance in LEACH protocol.

In order to fix the number of clusters during all the lifetime of the network, the area in VSG-LEACH is devided into equal size squares (4 square in scenario 1) as shown in Fig.4. The nodes of each square form a cluster.

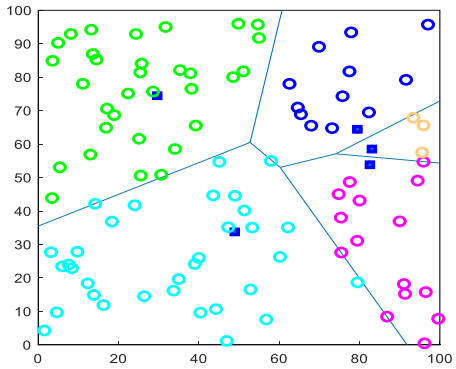


Fig. 2. Clustering in LEACH in 100x100 area in round 1

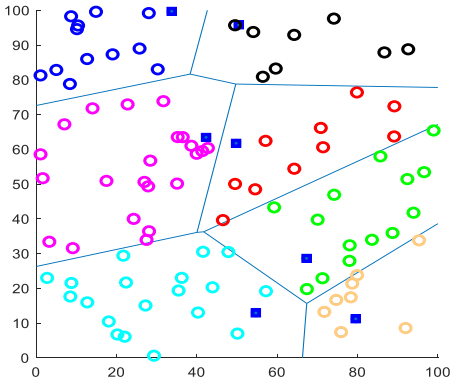


Fig. 3. Clustering in LEACH in 100x100 area in round 2

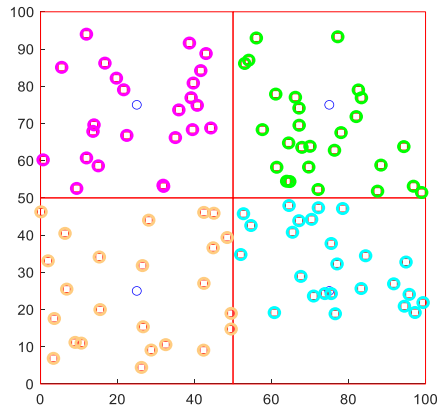


Fig. 4. Clustering in VSG-LEACH in 100x100 area

In this study, we will take the most important factor network lifetime and residual energy to compare the performance of LEACH and VSG-LEACH.

Fig. 5 and Fig. 6 show the number of alive nodes per round for LEACH and VSG_LEACH in scenario 1 and scenario 2 successively. In scenario 1 the first dead node in LEACH was in round 476 and the last dead node was in round 3286. While in VSG_LEACH the first dead node was in round 795 and the last dead node was in round 4386. In scenario 2 the first dead node in LEACH was in round 8 and the last dead node was in round 2380. While in VSG_LEACH the first dead node was in round 8 and the last dead node was in round 3652.

Therefore, VSG-LEACH has better network lifetime than LEACH in the two scenarios.

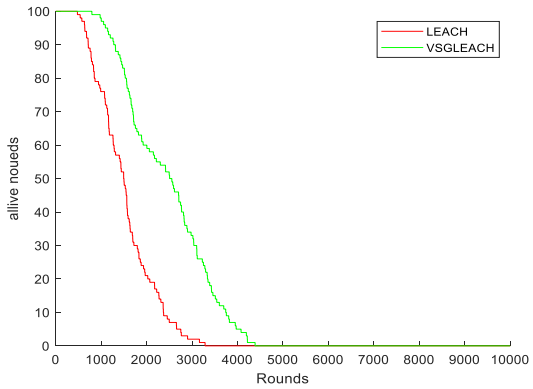


Fig. 5. Alive nodes per round in LEACH and VSG-LEACH in (100x100) area

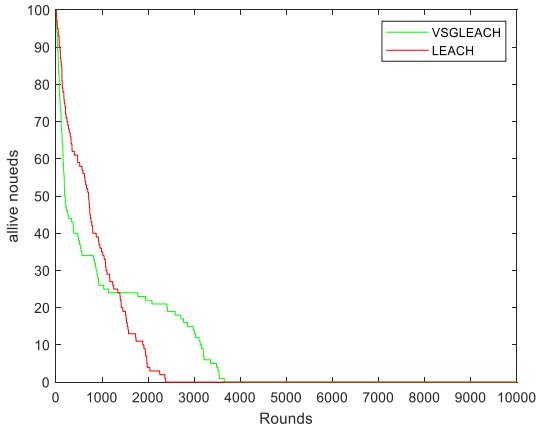


Fig. 6. Alive nodes per round in LEACH and VSG-LEACH in (200x200) area

Fig. 7 and Fig. 8 show the residual energy per round for LEACH and VSG_LEACH In scenario 1 and scenario 2 successively.

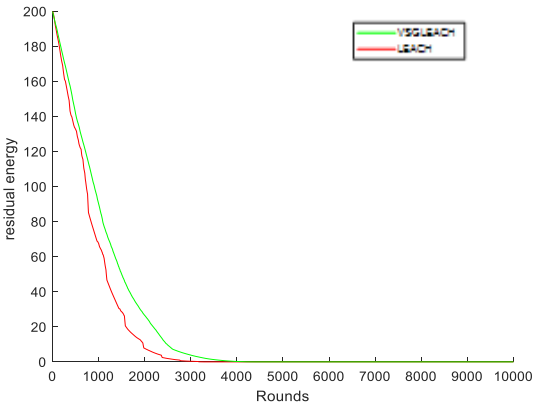


Fig. 7. Residual energy per round in LEACH and VSG-LEACH in (100x100) area

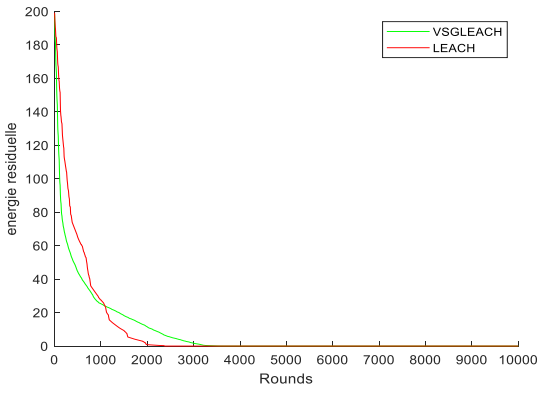


Fig. 8. Residual energy per round in LEACH and VSG-LEACH in (200x200) area

TABLE II. COMPARISON OF NETWORK LIFETIME OF PROTOCOLS

Scenario	% Dead node	LEACH	VSG-LEACH	Rate VSG-LEACH
Scenario 1	1(FND)	476	795	67,01%
	100(LND)	3286	4386	33,47%
Scenario 2	1(FND)	8	8	0%
	100(LND)	2380	3652	53,44%

TABLE III. COMPARISON OF RESIDUAL ENERGY OF PROTOCOLS

Scenario	% Energy consumption	LEACH	VSG-LEACH	Rate VSG-LEACH
Scenario 1	50	733	908	23,87%
	100	3286	4386	33,47%
Scenario 2	50	270	111	-58,88%
	100	2380	3652	53,44%

V. CONCLUSION

In this paper, we have study two protocols LEACH and VSG_LEACH focusing on network lifetime in order to compare the performance of each protocol. In the first scenario VSG_LEACH increase the network lifetime by 67.01% and residual energy with 33,47 compared with the basic LEACH protocol. In the second scenario VSG_LEACH increase the network lifetime by 0% and residual energy with 53,44 compared with the basic LEACH protocol. The results show that the network with VSG_LEACH survives longer than with LEACH which mean that the VSG_LEACH protocol is more efficient than the LEACH protocol.

REFERENCES

[1] Sun, Haibin, et al. "LPLL-LEACH: A study of a low-power and low-delay routing protocol based on LEACH." *Ad Hoc Networks* 140 (2023): 103064.

[2] Daanoune, Ikram, Baghdad Abdennaceur, and Abdelhakim Ballouk. "A comprehensive survey on LEACH-based clustering routing protocols in Wireless Sensor Networks." *Ad Hoc Networks* 114 (2021): 102409.

[3] Hossan, Arif, Sharmin Akter, and Pallab K. Choudhury. "Distance and energy aware extended LEACH using secondary cluster head for wireless sensor networks." *Telematics and Informatics Reports* 8 (2022): 100029.

[4] El Khediri, Salim. "Wireless sensor networks: A survey, categorization, main issues, and future orientations for clustering protocols." *Computing* 104.8 (2022): 1775-1837.

[5] Liu, Xuxun. "Atypical hierarchical routing protocols for wireless sensor networks: A review." *IEEE Sensors Journal* 15.10 (2015): 5372-5383.

[6] Heinzelman, Wendi Rabiner, Anantha Chandrakasan, and Hari Balakrishnan. "Energy-efficient communication protocol for wireless microsensor networks." *Proceedings of the 33rd annual* Hawaii international conference on system sciences. IEEE, 2000.

[7] Kaur, Amandeep, and Rajneesh Kumar. "A Comparative Analysis of Improvements in Leach Protocol: A Survey." Mobile Radio Communications and 5G Networks: Proceedings of MRCN *2020*. Springer Singapore, 2021.

[8] Yadav, Rajesh Kumar, and Monica Gupta. "Analysis of modified LEACH protocol for enhancing network lifetime in WSN." 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA). IEEE, 2020.5

[9] Saxena, Madhvi, et al. "Comparison of different multi-hop algorithms to improve the efficiency of LEACH protocol." *Wireless Personal Communications* 118 (2021): 2505-2518.

[10] Lechani, Taous, Victor Tourtchine, and Saïd Amari. "A new modification of LEACH for efficient energy in WSN." Indonesian Journal of Electrical Engineering and Computer Science 20.3 (2020): 1495-1506.

Bearing Fault Diagnosis Based on Multi-Scale CNN and Hilbert transform

Toumi Yassine
Laboratory of Advanced Electronic Systems, University Yahia Farès of Medea, Algeria
toumiyacine12@gmail.com

Bengherbia Billel
Laboratory of Advanced Electronic Systems, University Yahia Farès of Medea, Algeria
bil.elec@gmail.com

Meghraoui djamila
Laboratory of Speech Communication and Signal Processing, University USTHB, Algeria
djamila.meghraoui@gmail.com

Toumi Hayet
Laboratory of Mechanics, Physics and Mathematical, University Yahia Farès of Medea, Algeria
Allae_atchi@yahoo.fr

Bouchareb Nariman
Laboratory of Advanced Electronic Systems, University Yahia Farès of Medea, Algeria
narimanbouchareb1998@gmail.com

Rebaia Mohamed
Laboratory of Advanced Electronic Systems, University Yahia Farès of Medea, Algeria
rebaiamd@gmail.com

Abstract— **One of the most crucial components of a rotating machine is bearings. Mechanical failure, financial loss, and even personal injury are all possible consequences of bearing failure. For these reasons, it became necessary to find a reliable approach to identify and classify bearing faults. An intelligent fault classification model that combines Hilbert Transform (HT) and one-Dimensional Convolutional Neural Network (1D-CNN) is proposed to classify bearing fault. In the beginning, the original vibration signal is segmented into multiple samples, and we apply the HT on each sample. Finally, the resultant of the HT is used to design and train two parallel convolutional neural networks with the same filter length but different filter sizes. The Paderborn University bearing dataset is used to prove the performance of the proposed method. The results show the high accuracy of the proposed method with an accuracy rate of 99.47%.**

Keywords—rolling bearing, one-dimensional convolutional neural network, Hilbert transform.

I. INTRODUCTION

Recently, rotating machines are become widely used in different industrial applications, which leads to the emergence of deteriorating operating conditions. Almost, 44% of the rotating machines faults result due to the malfunctioning of the bearings [1]. Consequently, bearing faults detection and diagnosis in rotating machinery becomes important. Bearings of rotating machines usually generate vibration signals. The occurrence of faults in these bearings will cause a change in the normal bearing vibration signal, as a result, vibration analysis considers the most widely used technique in preventive maintenance [2].

Bearing faults identification includes two steps, feature extraction and faults classification. In the case of feature extraction, several signal processing techniques are used to extract features, including statistical analysis, Fast Fourier Transform (FFT), envelope analysis, Hilbert Transform (HT), wavelet transformation (WT) and empirical mode decomposition (EMD). In the case of fault classification, many machine learning techniques have been used. In recent years, deep learning (DL) technologies have been widely used in the field of fault diagnosis of bearing because of their ability to overcome the shortcomings inherent in traditional machine learning methods [3].

With the motivation to improve the capability of bearing fault detection and classification, a signal analysis method based on Hilbert transform and one-Dimensional Convolutional Neural Network is proposed in this study. The remainder of this document is structured as follows: The proposed method, including the HT and the 1D-CNN are presented in the second section. The proposed method was validated using the Paderborn university bearing dataset, which is described in part three. The proposed architecture of the CNN is shown in section four. The experimental results are presented in section five. In the final section, some conclusions are made.

II. PROPOSED METHOD

The flowchart in figure 1 illustrates the suggested method. The first step is to segment the raw vibration signal into segments. The HT of each segment is calculated as a second

step. Finally, the bearing defects are classified using a 1D-CNN classifier.

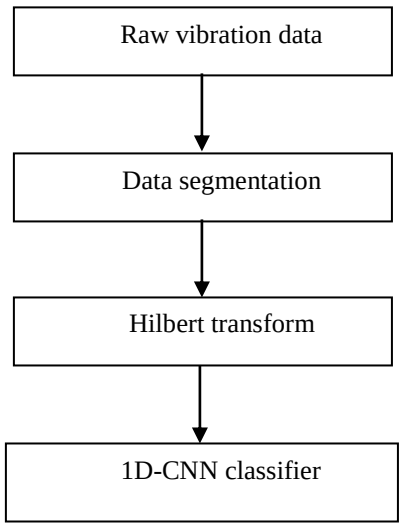


Fig1. Flowchart of the proposed method

A. *Hilbert Transform*

The Hilbert transform is used to facilitate the formation of the analytic signal $x_a(t)$, as follows:

$$x_a(t) = x(t) + j\hat{x}(t) \quad (1)$$

$\hat{x}(t)$ is the Hilbert Transform of $x(t)$, which is given by the convolving the signal $x(t)$ with $\frac{1}{\pi t}$, as giving in equation 2.

$$H\{x(t)\} = x(t) * \frac{1}{\pi t} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau \quad (2)$$

where t and τ are time and translation parameters respectively.

There is no equivalent between the result of the HT and the original signal. In other words, the phase spectrum of the signal is changed and shifted by ±90°. However, the magnitude spectrum does not change. Also, the HT of a signal does not change the domain [4]. Moreover, HT is an adaptive method used for both non-linear and non-stationary signals. As a result, it is one of the preferred advanced signal analysis techniques for fault diagnosis [5].

B. *Convolutional Neural Networks*

The Convolutional Neural Network CNN is one of the representative deep learning algorithms, which is frequently utilised in disciplines like defect diagnostics, computer vision, and natural language processing. A CNN architecture typically comprises four different kinds of layers: convolution, pooling, activation, and fully connected.

The convolution layer is the fundamental way that the filter is applied to an input. Firstly, the size of the filter window is defined, this filter window moves progressively from left to right and from top to bottom by a certain number of steps defined beforehand (stride) until it reaches the end of the input. At each portion of the input encountered, a convolution calculation is carried out making it possible to obtain a feature map as output which indicates where the features are located in the input. Once the feature map is computed, each value is passed to an activation function. Several nonlinear activation functions exist, but the most common function used presently is the rectified linear unit (ReLU). The ReLU function replaces all negative values received as inputs with zeros, which allows the model to perform better, learn from the training data faster, and overcomes the vanishing gradient problem [6].

Convolution layers are often followed by another block of pooling layers. This layer is a sample-based discretization process. Its purpose is to down sample an input representation by reducing its dimension. Moreover, its interest is that it reduces the computational cost by reducing the number of parameters to be learned, which helps control the overfitting of the network. After the convolutional part of a CNN, comes the classification part, which corresponds to the fully connected layer. These layers are placed at the end of the CNN architecture and are fully connected to all output neurons. After receiving an input vector, the fully Connected layer successively applies a linear combination and then an activation function with the final aim of classifying the input. Finally, it returns as output a vector corresponding to the number of classes in which each component represents the probability for the input to belong to a class.

III. DATASET DESCRIPTION

Paderborn bearing dataset [7] was provided from the mechanical engineering research centre of Kat-Data Centre. The test rig of this dataset is displayed in Figure 2.

The dataset has been generated using different types of bearings, including healthy bearings, artificially damaged bearings and real damaged bearings. Drilling, electric discharge machining (EDM), and manual electric-sculpting methods are techniques used to create artificial damage on the bearing. For the real damaged bearing, two different techniques, named fatigue and plastic deformation, were applied. For each bearing, various kinds of data, including vibration, current, speed, and temperature, were recorded. In our work, we used vibration data. The dataset is generated with multiple combinations of rpm, torque, and load as shown in table 1. For the purpose of this study, we used the first combination of the signals illustrated in table 2. Furthermore, each bearing is used 20 times to generate 20 signals with one fixed combination. The signal produced has a sample frequency of 64 kHz and lasts for 4 seconds. This indicates that there are 256,000 data points in a signal. We divide each type of signal into 125 segments and each segment contains 2048 points, which means a total of 22500 segments are used. 70% of these segments were used for the training and the remaining signals were for the test.

Table 2: working conditions

NO	Relational speed (rpm)	Load torque (Nm)	Radial force (N)	Signal name
1	1500	0.7	1000	N15_M07_F10
2	900	0.7	1000	N09_M07_F10
3	1500	0.1	1000	N15_M01_F10
4	1500	0.7	400	N15_M07_F04

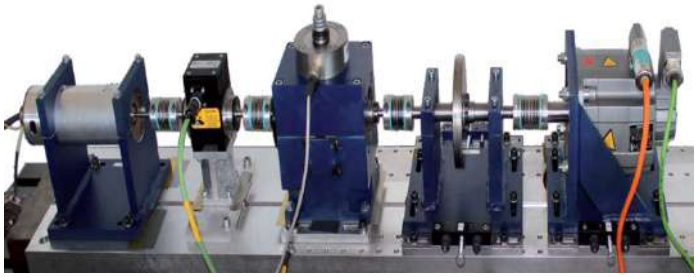


Fig.2. The experimental test rig.

Table 2: List of data sets used

Bearings faults	Healthy	Inner	Outer
Signals code	K001, K002, K003	KA01, KA04, KA07	KI01, KI04, KI07

IV. PROPOSED ARCHITECTURE

Figure 3 represents our 1D-CNN architecture comprising two parallel 1-D CNNs followed by the concatenating layer. Indeed, convolutional streams in the two 1D_CNN models produce feature maps then their fusion provides the input of the fully connected layers. In each convolutional layer for both 1D-CNN models, the number of filters is fixed but their sizes are different, filters of size 3 with a stride of 2 with padding were used for the first 1D-CNN model, and filters of size 5 with a stride of 2 with padding were used for the second model. For all the max pooling layers, a stride of 2 was used. The proposed model uses different hyperparameters, as an example, for the optimization algorithm we use Adam Optimizer, and sparse categorical cross-entropy is used for the loss function.

V. EXPERIMENTAL RESULTS

Table 3: The Accuracy, Recall and F1-score rates of each faults type

Classes	0	1	2
Precision	0.9947	0.9907	0.9938
Recall	0.9929	0.9969	0.9938
F1-score	0.9998	0.9942	0.9965
Accuracy	99.47%		

The loss and accuracy curve for the proposed model are provided in figure 4 and figure 5 respectively. The accuracy curves indicate that the model achieves an accuracy of 100% for the training data and 99,47% for the testing data. In order to further check the performance of the proposed model, the precision, recall and F1 score rates are calculated and presented in table 3. From table 3, we can see that all the rates are close to 1, which confirms the ability of the proposed model to classify bearing faults.

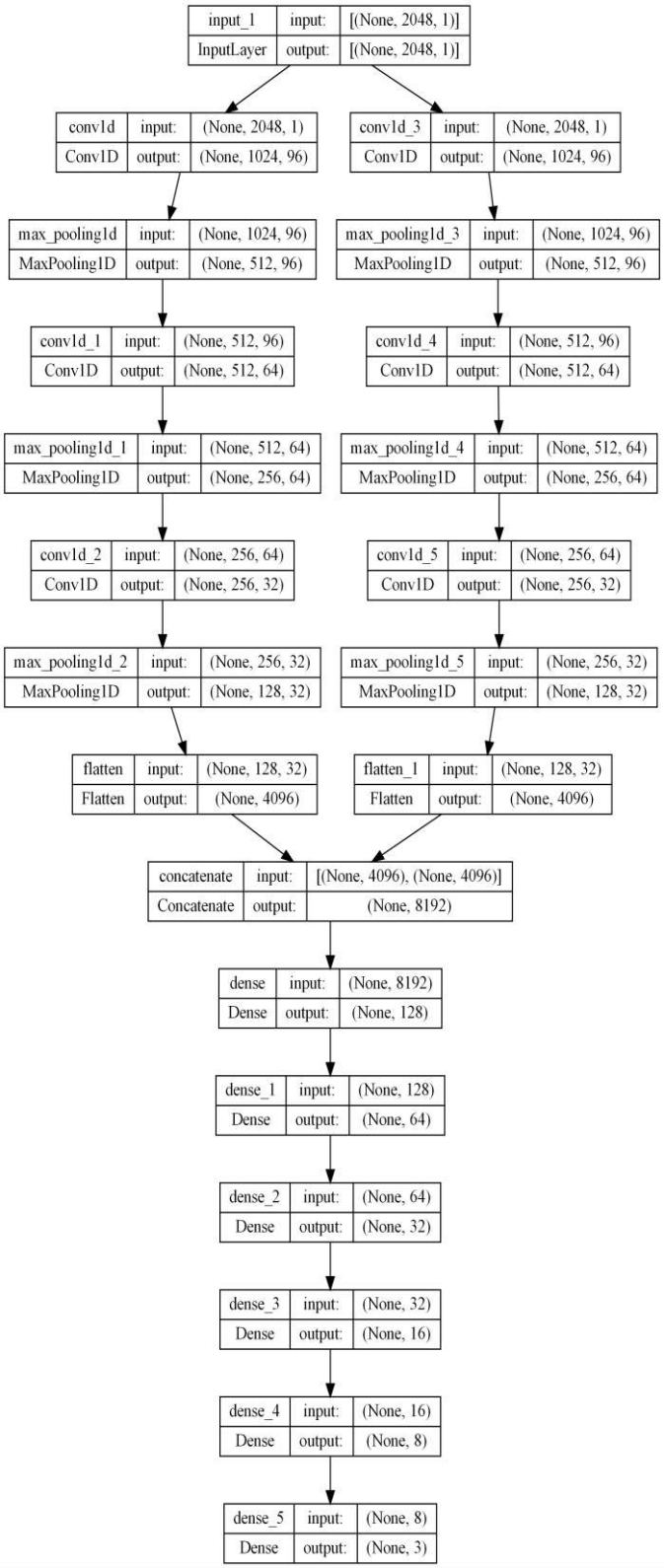


Fig.3. Proposed method architecture.

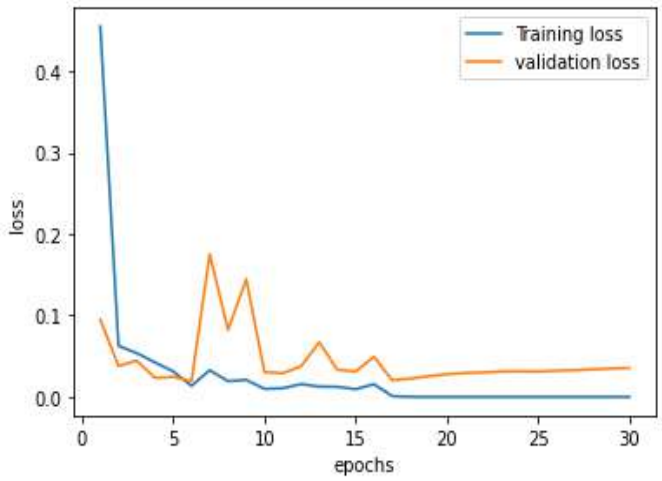


Fig.4. Loss curve of the 1D-CNN

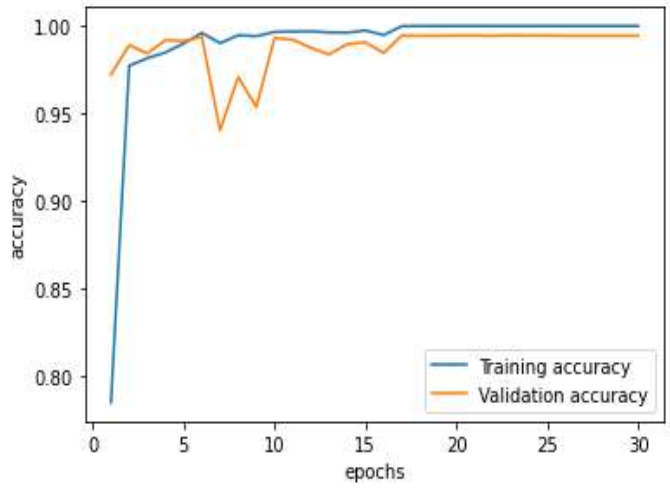


Fig.5. Accuracy curve of the 1D-CNN

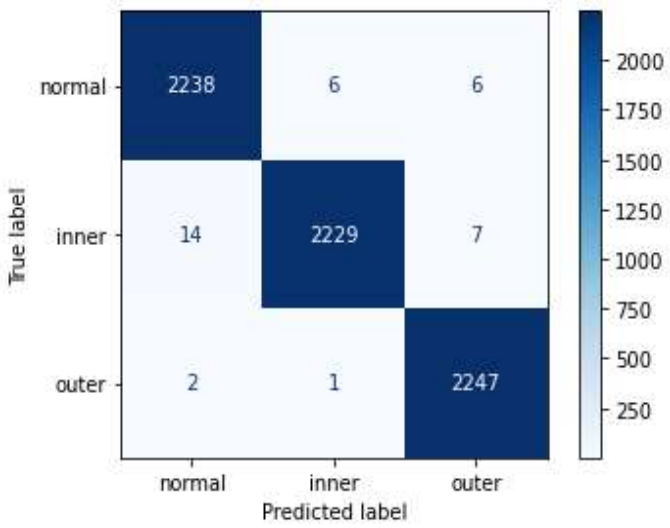


Fig.6. Confusion matrix of the 1D-CNN.

VI. Conclusion

This work has proposed an approach for bearing fault identification and classification. In this context, the Hilbert transform was used to extract features from the bearing signals vibration. For classification purposes, two parallel one-dimensional convolution neural network was applied. The findings demonstrate the efficacy of the suggested method for bearing faults classification with an accuracy of 99.47%.

References

[1] I. Attoui, B. Oudjani, N. Boutasseta, N. Fergani, M. Bouakkaz, and A. Bouraiou, "Novel predictive features using a wrapper model for rolling bearing fault diagnosis based on vibration signal analysis," *Int. J. Adv. Manuf. Technol.*, vol. 106, pp. 3409–3435, 2020, doi: 10.1007/s00170-019-04729-4.

[2] G. Tingarikar and A. Choudhury, "Vibration Analysis - Based Fault Diagnosis of a Dynamically Loaded Bearing with Distributed Defect," *Arab. J. Sci. Eng.*, 2021, doi: 10.1007/s13369-021-05862-7.

[3] S. Han, S. Oh, and J. Jeong, "Bearing Fault Diagnosis Based on Multiscale Convolutional Neural Network Using Data Augmentation," *J. Sensors*, vol. 2021, 2021, doi: 10.1155/2021/6699637.

[4] P. Gangsar and R. Tiwari, "Signal based condition monitoring techniques for fault detection and diagnosis of induction motors: A state-of-the-art review," *Mech. Syst. Signal Process.*, vol. 144, p. 106908, 2020, doi: 10.1016/j.ymssp.2020.106908.

[5] S. K. Ramu, G. C. Raj Irudayaraj, S. Subramani, and U. Subramaniam, "Broken rotor bar fault detection using Hilbert transform and neural networks applied to direct torque control of induction motor drive," *IET Power Electron.*, vol. 13, no. 15, pp. 3328–3338, 2020, doi: 10.1049/iet-pel.2019.1543.

[6] E. U. H. Qazi, A. Almorjan, and T. Zia, "A One-Dimensional Convolutional Neural Network (1D-CNN) Based Deep Learning System for Network Intrusion Detection," *Appl. Sci.*, vol. 12, no. 16, pp. 4–17, 2022, doi: 10.3390/app12167986.

[7] "Konstruktions- und Antriebstechnik (KAt) - Data Sets and Download (Universität Paderborn)." https://mb.uni-paderborn.de/en/kat/main-research/datacenter/bearing-datacenter/data-sets-and-download (accessed Feb. 21, 2023).

Curvelets Transform versus Histogram of Gradient for Handwriting Gender Recognition System

1st Yasmine Guerbai
Communicating and Intelligent System Engineering Laboratory Faculty of Electrical Engineering, USTHB
Electrical Systems Engineering department, Faculty of technology,
University M`hammed bougara, Boumerdes, .guerbai@univ-boumerdes.dz

2nd Ramzi Halimouche
Faculty of Electrical Engineering, USTHB

3rd ismahan Mahdi
LREEI , University M`hammed bougara, Boumerdes, UMBB

4 th Youcef chibani
Communicating and Intelligent System Engineering Laboratory Faculty of Electrical Engineering, USTHB
ychibani@usthb.dz

***Abstract*—** In this paper we propose the design of an intelligent system for the prediction of soft-biometric characteristics through offline handwritten texts. The goal is to present a study that allows the gender determination of individuals from their handwriting. The idea is based on the extraction of a set of characteristics from samples of man and woman writers, so that it can distinguish between the two categories. Writing attributes such as orientation, curvature, texture, and readability are estimated by calculating characteristics. Classification is done using One Class Support Vector Machine (OC-SVM). Both proposed methods were evaluated using IAM datasets, an interesting results were recorded.

***Keywords*—** One-Class support Vector machines, gender recognition, Curvelets transform, Histogram of Oriented Gradients.

I. Introduction

Handwritings can be classified into many categories including gender, age, handedness, and nationality. This type of classification has several applications. For example, in the forensic domain, handwriting classification can help the investigators to focus on a certain category of suspects. Additionally, processing each category separately leads to improved results in writer identification and verification applications. [1], gender based on the face mode in order to use for publicity card, which can serve to detect whether the product is designed for Woman or Man [2], [3] and gender based on handwriting mode [4] for Parkinson disease [5]. In this context, a research group on computer vision and artificial intelligence at Bern University developed the IAM handwriting dataset, which is devoted to writer identification like gender and handedness prediction [4]. Using this dataset, Liwicki et al [6], used the gender and handedness prediction. Specifically, Bi-Class Support Vector Machine (BC-SVM) and Generalized Markov Model (GMM) classifiers are used in order to achieve the classification task. For the off-line characterization, they applied a sliding window in the writing direction in order to calculate features. On-line and off-line features yield accuracy equal to 64.25% and 55.39%, respectively. For more improvement of the obtained results, a combination of both modes of features yields an accuracy of 67.57%. In related work, Siddiqi et al [7], proposed an alternative system which is based on extracting a set of specific features such as slant, curvature, texture and legibility, which are estimated from writing samples of Woman and Man writers. The classification step is carried out using artificial neural networks and support vector machine to discriminate between the Man and Women writing style. More recently, Bouadjenek el al [8] provided local descriptors in order to improve off-line handwriting gender, a set of different features is used, involving Grid features, Local Bianary Pattern (LBP) and Histogram of Oriented Gradients (HOG), in order to identify the handwritten style is whether for Man or Woman a BC-SVM is used, From the obtained results, HOG descriptor shows its superiority to achieve 74% while the density and LBP descriptors yield 72% and 70%, respectively. Guerbai et al [9] used combining different classifiers through the minimum rule of decision of OC-SVM classifier, MRRs are increased whatever the selected threshold. Indeed, when selecting the hard threshold, the MRR are 72.00%. Whilst, for soft threshold, the MRRs are 77.33%. This proves the need to adjust the threshold and the usefulness to combine classifiers in order to achieve better performances. Furthermore, we can notice that the threshold is very sensitive and it should be adjusted carefully according to the requirements of the HGRS.

The main contribution of this paper is to investigate the effective use of OC-SVM classifier for HGRS by using curvelets transform and Histogram of Gradient (HOG). The remainder of the paper is organized as follows. Section 2 describes a review a features generation and OC-SVM classifier. Section 3 presents the design of individual and combined HGRS system. The experimental and statistical results are summarized in Section 4. Finally, the conclusion is provided in Section 5.

II. Review of features generation and OC-SVM classifier

A. *Features generation*

1) Histogram of Oriented Gradients

The Histogram of Oriented Gradients (HOG) feature was introduced by Dalal and Triggs for human detection [10]. HOG is designed to characterize the shape of objects based on the distribution of local intensity gradients or edge directions. This descriptor has shown high performance in various applications [11]

2) Curveltet transform

The curvelet transform is a new multi-scale transform developed by Candés and Donoho [12], [13]. The curvelet transform has been developed specifically for the representation and other singularities along curves, which makes it much more efficiently than traditional transforms, i.e. using many fewer coefficients for a given accuracy of reconstruction. The curvelet transform has been widely used in several applications such as pattern recognition [14] and image compression [15]. More recently, the curvelet transform has been successfully used for offline handwritten signature retrieval, where the feature vector is generated using the energy and the standard deviation of the curvelet coefficient [16]. In our system, we only use the energy of the curvelet coefficient computed from the entire of the handwritten document image. The feature vector is composed of 44 components.

B. Novelty detection

The OC-SVM (also known as single-class classification or novelty detection) introduced by Schölkopf et al [17] algorithm relies on mapping the training samples belonging to S into a high dimensional feature space leading to find the maximal margin hyperplane, which best separates the training samples from the origin. For a test pattern x, a decision function namely $f(x)$ is performed for evaluating in which side of the hyperplane it falls. Formally, the decision function takes the following form:

$$f(x) = \{\sum_{i=1}^{n} \alpha_i K(x, x_i) - \rho\} \quad (1)$$

$K(.,.)$ defines the OC-SVM kernel that allows projecting data from the original space to the feature space. n is the number of training samples and α_i are the Lagrange multipliers computed by optimizing the following equations:

$$\min_{\alpha} \left\{\frac{1}{2} \alpha_i \alpha_j K(x_i, x_j)\right\} \quad (2)$$

[13] proposed an alternative way using a logarithmic transformation of the decision function calculated as follows:

$$f_j(x) = -log\left[\sum_{i=1}^{n_j} \alpha_j k_j(x, x_{ij})\right] + log(\rho_j) \quad (3)$$

where n_j and ρ_j are the number of training samples and the distance from the origin associated to each OC-SVM classifier, respectively. Hence, a pattern is assigned to the predefined class according to the decision rule:

$$x \in \theta_i \text{ if } f_i(x) = min\{f_j(x), 1 \le j \le M\} \quad (4)$$

The problem of maximizing decision functions using simple decision rule becomes a minimization problem when using the logarithm transformation. The logarithm function has the ability to attenuate the high values and increase small values for calibrating the outputs of the OC-SVMs.

III. Design of the System

The Handwritten Gender Recognition System (HGRS) is designed as any pattern recognition system for which two modules are required for developing a robust system: feature generation and classification. However, for HGRS, a specific set of features and classification method should be selected for fitting the application requirements. This design can be conducted into two steps by considering the individual system for classifying the Man and Woman gender. The second step is performed by combining multiple individual HGRS by means of the combination rule in order to improve the classification robustness.

The HGRS is designed as any pattern recognition system for which two modules are required for developing a robust system: feature generation and classification. However, for HGRS, a specific set of features and classification method should be selected for fitting the application requirements. This design can be conducted into two steps by considering the individual system for classifying the Man and Woman gender. The second step is performed, by combining multiple individual HGRS by means of the combination rule—in order to improve the classification robustness.

$$\text{Subject to } 0 \le \alpha_i \le \frac{1}{\nu n}$$
$$\sum_i^n \alpha_i = 1 \quad (5)$$

ν is the percentage of samples considered as outliers taking values in the range [0, 1]. A pattern x is then accepted when $f(x) > 0$. Otherwise, it is rejected. Various kernel functions can be used as polynomial or Radial Basis Function (RBF) or multilayer perceptron [18]. Usually, the RBF kernel is the most used, which is defined as:

$$K(x, x_i) = exp\left(-\gamma\, d(x, x_i)\right) \quad (6)$$

Such that $d(x, x_i)$ is the distance between the pattern x and the training sample x_i, whereas γ is the kernel parameter.

C). Extension for Multi-class Classification

The OC-SVM was originally designed for one-class classification [12]. Its extension to multi-class implementation is currently an active research [10]. The usual approach consists to represent each sample set of an object or class by its corresponding decision function of the OC-SVM. This approach builds M models (classifiers) of OC-SVM (OC-SVM for each class). The i^{th} of OC-SVM is learned with all the training examples of θ_i class that are indexed with positive labels (+1). This classifier is a maximal distance from the origin that separates it from $M-1$ other classes.

Let $f_i(x)\ \{i = 1, \dots, M\}$ be un decision function associated to each OC-SVM classifier, a straightforward assignation of a pattern x to a predefined class is formulated as follows:

$$x \in \theta_i \text{ if } f_i(x) = max\{f_j(x), 1 \le j \le M\} \quad (7)$$

When using this equation, a major problem arises related to the combination of several classifiers designed separately (each classifier learned forms a class by ignoring those of other classes in the training phase). For this, Rabaoui et al

A. Design of the Individual System

1) Architecture of the System: The design of the HGRS can be performed into three steps: selecting a set of writers, deriving the HGRS classification model and finding the optimal decision threshold. More precisely, the set of samples is randomly selected from a dataset each one having N samples. To get a robust HGRS classification model, the set of dataset is divided into two subsets namely Subset 1 and Subset 2 as depicted in Figure 1. The first subset (Subset 1) containing G_p samples is used for finding the parameters of the OC-SVM during training step. While the second subset (Subset 2) containing G_t samples is used for finding the optimal decision threshold. The proposed individual HGRS is designed to automatically classify writers into two classes that are "Man" or "Woman" ("Male or Female"). The proposed HGRS as shown in Figure 2 can be designed through three main steps: feature generation using the curvelet transform, the classification using the OC-SVM classifier and the decision rule. The parameters of the HGRS can be computed according to the following steps: selecting a set of writers, feature generation using the curvelet transform, building the handwritten gender recognition models using the OC-SVM classifiers and finding the optimal decision threshold. Firstly, the handwriting document is transformed to features using the Curvelet transform, which are submitted to the OC-SVM classifier in order to decide whether the gender of the person. Let f_M and f_W are the decision functions associated to the respective Man and Woman OC-SVM namely OC-SVM$_M$ and OC-SVM$_W$, respectively. Then the decision rule consists to select the best response provided by both OC-SVM classifiers in order to assign a document to the corresponding class.

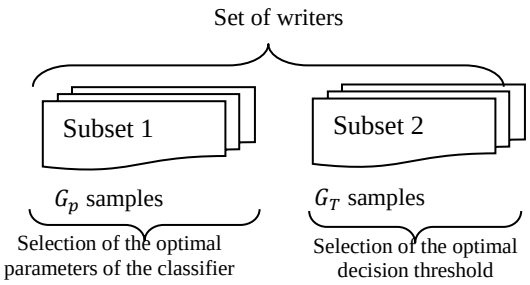


Fig. 1. Design of the HGRS

3) Threshold Adjustment: The threshold adjustment of the OC-SVM is similar as described in the previous paper [9]. For HGRS like in handwritten signature verification, a pattern x is theoretically correctly classified and thus accepted when the decision function f_M or f_W is positive. Hence, the threshold is implicitly fixed to zero. This approach can be considered as a hard thresholding. Indeed, some handwritten samples near to the hyper-plane into the feature space are not accepted and thus rejected. Thus, a threshold adjustment method is proposed to get better flexibility of the decision function of OC-SVM for performing more accurate classification.

The optimal decision threshold (t_{opt}) is deduced from RRM and RRW curves as depicted in Figure 3 using the Half Total Recognition Rate (HTRR) defined as:

$$HTRR = \frac{\text{RRM+RRW}}{2} \tag{8}$$

such as RRM and RRW are the Recognition Rates associated to the Man and Women, respectively. Figure 3 shows the principle for selecting the optimal decision threshold from RRW and RRM curves.

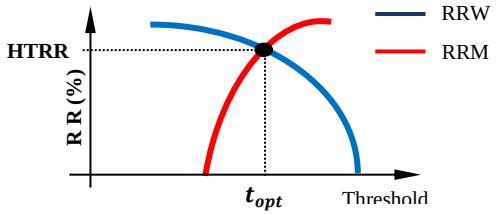


Fig. 3. Selection of the optimal threshold from RRM and RRM curves versus threshold.

4) Distance metrics: The training of the OC-SVM requires defining the appropriate distance used into the RBF kernel defined as follows:.

$$K(x, x_i) = \exp\left(-\gamma d(x, x_i)\right) \tag{9}$$

Such that $d(x, x_i)$ is the distance into the kernel and γ is a user-defined parameter that controls the RBF radius.

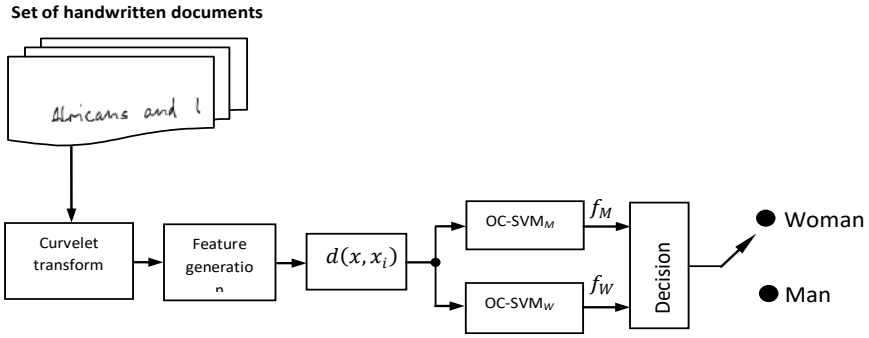


Fig. 2. The overall architecture of the proposed individual system for gender recognition.

Usually, the Euclidean distance is the most used metric. However, this metric does not provide satisfactory results in all cases. Hence, various distances can be selected, which are defined in the literature [19]. In our case, we use five distance metrics which are Euclidean, Cityblock, Chebychev, Correlation and Spearman [20].

B. Design of the Combined Systems

For the design of the combined systems, two types of combination are possible. The first one (System 1) as depicted in Figure 4 consists to combine the Man and Woman classifiers for each distance using a decision rule. Then, all decision functions are combined through a combination rule. The second system (System 2) as depicted in Figure 5, which is the standard system of combination, consists to combine the OC-SVM distances for gender (Man or Woman) and combine them with a combination rule. After, a decision rule is used in order to select the select response [21,22,23].

The training of the OC-SVM requires defining the appropriate distance used into the RBF kernel. Usually, the Euclidean distance is the most used metric. However, this metric does not provide satisfactory results in all cases. Hence, we propose to combine several OC-SVM each one has its own distance. Several combination rules are possible to achieve a robust system like majority vote, static weighting, dynamic weighting, product, average, maximum and minimum rule [21]. For both proposed combined HGRS, we select the minimum rule as a combination rule for its simple implementation.

IV. EXPERIMENTAL RESULTS

A. Dataset Description

The IAM dataset (Institut für Informatikund Angewandte Mathematik) was developed by a group of researchers in computer vision and artificial intelligence from the University of Bern, in particular for the recognition of gender and manuality. Using this dataset, several researchers [29] and [30] have accompanied it with different classifiers and descriptors in order to accomplish the task of online and offline classification of handwriting.

The classification task aims to identify the writer gender (Woman or Man) for each handwritten text. For a fairly comparison obtained from the state-of-art [6], we use 75 samples for each class. For training the OC-SVM classifiers, we randomly selected 50 samples for each class and the remaining 25 samples for testing.

Figure 4 shows some samples from the IAM datasets.

Fig. 4. Some samples from the IAM dataset.

B. Experimental protocol

In order to evaluate the performance of OC-SVM descriptors for the proposed systems, we consider a set of samples for the design of HGRS using a single class (Male or Female) with the aim of finding the parameters and the optimal threshold. . In this way, each class has its own OC-SVM mode l[24]. Samples not enrolled in the system are used to assess the robustness of the proposed system .

We evaluate our descriptors individually and then using combinations.

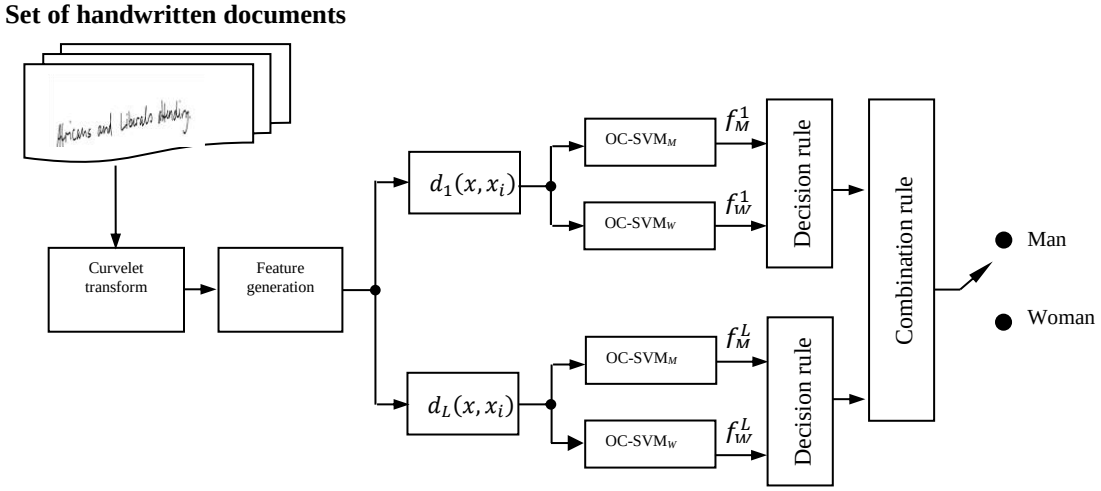


Fig. 5. The overall architecture of the proposed combined System 1 for gender recognition

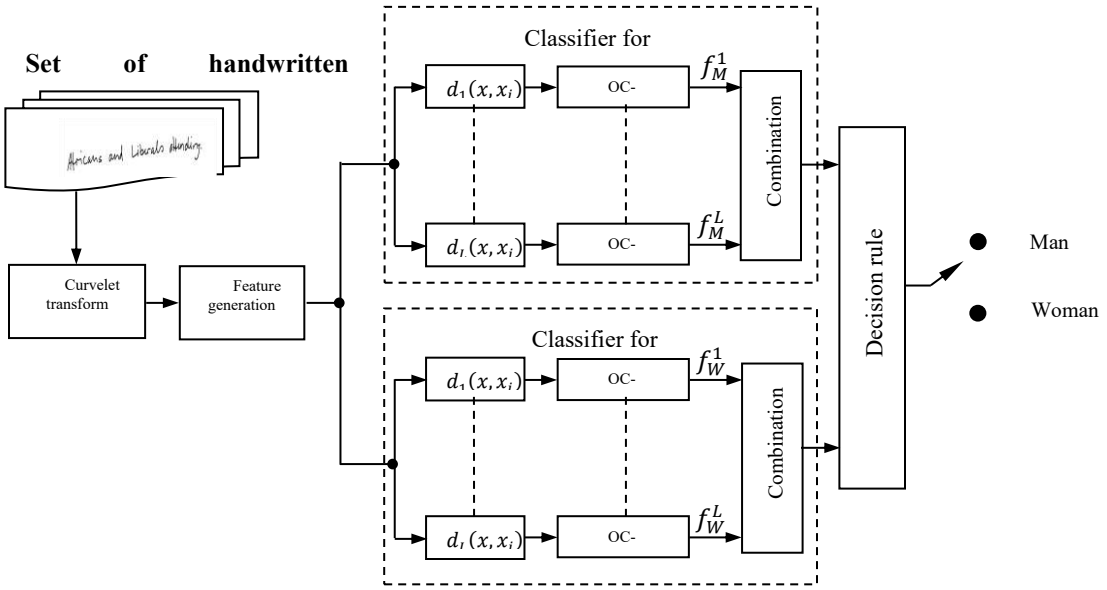


Fig.6. The overall architecture of the proposed combined System 2 for gender recognition

1) Evaluation of the individual system

a) Decision threshold selection

The choice of the optimal decision threshold is a crucial step for the design of the robust HGRS. In our case, it is selected from the RRW and RRM curves using male and female handwriting samples. In this section, the OC-SVM classifiers are trained separately. To find the optimal threshold, we use all the values of the decision functions provided by different OC-SVMs to calculate the mean and the standard deviation, respectively.

Tables 1 and 2 report the recognition rates by class (Male and Female) and the average recognition rate provided by the different distances used in the OC-SVM core for the hard ($t = 0$) and soft ($t = t\ opt$). The results obtained show different precision values provided by each classifier for all Male and Female data depending on the chosen threshold.

TABLE1. THE AVERAGE OF THE RECOGNITION RATE (%) FROM THE DIFFERENT DISTANCES USED IN THE OC-SVM CORE FOR HARD AND SOFT THRESHOLDS USING THE INDIVIDUAL SYSTEM FOR THE CURVELET TRANSFORM

Distances	Hard threshold			Soft threshold		
	Man	Woman	Mean	Man	Woman	Mean
Euclidienne	50,66	53,66	52,16	52	57,66	54,33
Cityblock	50	52,16	51,08	51,66	55,33	53,5
Chebychev	62	63,66	61,83	63,41	64	63,7
Correlation	53,33	57,33	55,33	58,76	61,33	60
Spearman	56,93	58,66	57,8	58,33	59,76	59,05

As shown in Table 1 when the HGRS is trained with male samples, the RRM varies between 50 and 62% for the five distances. On the other hand, when the threshold value is soft, the RRM varies between 51.66 and 63.41% for the five distances. Therefore, adjusting the threshold improves the RRM regardless of the distance used.

- For Histogram of Oriented Gradients

TABLE 2. THE AVERAGE RECOGNITION RATE (%) FROM THE DIFFERENT DISTANCES USED IN THE OC-SVM KERNEL FOR HARD AND SOFT THRESHOLDS USING THE INDIVIDUAL SYSTEM FOR ORIENTED GRADIENT HISTOGRAM

Distances	Hard threshold			Soft threshold	
	Man	Woman	Men	Man	Woman
Euclidienne	50,66	59,33	51	51,33	52,33
Cityblock	50,33	58	54,16	55,33	56,66
Chebychev	55,33	56,66	56	55,33	58,33
Correlation	49,69	49,66	49,68	52,33	53,66
Spearman	52,66	63,69	53,18	54,33	54,66

As shown in Table 2 when the HGRS is trained with male samples, the RRM varies between 50.66 and 55.33% for the five distances. On the other hand, when the threshold value is soft, the RRM varies between 51.33 and 55.33% for the five distances. Therefore, adjusting the threshold improves the RRW regardless of the distance used.

2) Evaluation of the combined system

a) Selection of the decision threshold

In order to find the optimal threshold, we use all the values of the decision functions provided by the different OC-SVM classifiers to calculate the mean and the standard deviation,

value of K and the optimal threshold (t opt) are chosen from RRW and RRM.

C. Classification results

- Using the curvelet transform

TABLE 3. THE AVERAGE OF THE RECOGNITION RATE (%) FROM THE DIFFERENT DISTANCES USED IN THE OC-SVM KERNEL FOR HARD AND SOFT THRESHOLDS USING THE COMBINED SYSTEM FOR THE CURVELET TRANSFORM.

			Hard threshold		
Man	Woman	Mean	Man	Woman	Mean
88,45	90	89,25	88,89	90	89,44

- Using the Histogram of Oriented Gradients

TABLE 4. THE AVERAGE OF THE RECOGNITION RATE (%) FROM THE DIFFERENT DISTANCES USED IN THE OC-SVM KERNEL FOR HARD AND SOFT THRESHOLDS USING THE COMBINED SYSTEM FOR ORIENTED GRADIENT HISTOGRAM.

Soft thr	eshold		Hard threshold		
Man	Woman	Mean	Man	Woman	Mean
88,45	90	89,25	88,89	90	89,44

When combining different classifiers using the maximum rule, the MRRs increase regardless of the selected threshold. As shown in Tables 3 and 4 respectively, when the threshold is hard, the MRR is 83.1% for the HOG and 88.45% for the Curvelet transform while for the soft threshold, the MRR is 83.54% for the HOG and 88.89% for the Curvelet transform. This proves the necessity of adjusting the threshold and the usefulness of combining classifiers to obtain better performance. Also, we can notice that the threshold is very sensitive and needs to be adjusted carefully according to the HGRS requirements.

3. Comparative study between the Curvelet transform and the Histogram of Oriented Gradients

After carrying out evaluations on the two methods adopted, we note that for the individual systems as for the combined systems, the Curvelet transform provided an accuracy of 89.44% for the combined system the HOG provided an accuracy of 84.09%. These results are very surprising especially for the Curvelet transform which, as we can see, is more efficient since it allows the system to be more robust.

4.6. Comparative analysis

Table 5 reports a comparative analysis in terms of MRR with the state-of-art. For individual system, we can note that the best system is provided by Bouadjenek et al [8]. When combining individual systems, our proposed system is better comparatively to other systems.

TABLE 5. MEAN RECOGNITION RATE (%) PROVIDED BY THE PROPOSED SYSTEM COMPARATIVELY TO OTHER SYSTEMS USING INDIVIDUAL AND COMBINED SYSTEMS

Type of the system	References	MRR (%)
Individual	Liwicki et al [4]	55,39
	Bouadjnek et al [8]	74
	Guerbai et al [9]	62,49
	Our system	**63,7**
Combined	Liwicki et al[4]	67,57
	Guerbai [9]	77,33
	Our system	**89,44**

V. CONCLUSION

In this paper, we presented a system that classifies the gender of writers from their handwritings using the Curvelet transform and the Histogram of Oriented Gradients (HOG) with the OC-SVM classifier. The main advantage of the proposed gender recognition system is that the OC-SVM classifier allows designing the HGRS using a single target class. However, we can note that the threshold must be carefully adjusted in order to obtain better results. To further improve the HGRS, a combined system based on several distances used in the OC-SVM kernel is proposed and gives better results compared to the state of the art HGRS using the curvelet transform.

REFERENCES

[1] D Voyer, D.:Sex differences in dichotic listening, Brain and Cognition Journal, Vol.76 (2011) 245–255

[2] Abel. S., deBruin. M., and Nowak. A.: Women adverdting and representation, beyond familiar paradigms cressikll. Nj. Hampton (2010)

[3] Lee, J, D., Lin, C, Y., and Huang, C, H.: Novel features selection for gender classification, IEEE International Conference on in Mechatronics and Automation (ICMA) (2013) 785–790

[4] Liwicki, M., Schlapbach, A., Loretan, P., and Bunke, H.: Automatic detection of gender and handedness from on-line handwriting, in Conference of the International Graphonomics Society, (2007) 179–183

[5] Shulman, L, M.: gender differences in parkinson's disease, Gender Medicine, Vol. 4, No. 1 (2007)8-18

[6] Liwicki, M., Schlapbach, A., and Bunke, H.: Automatic gender detectionusing on-line and off-line information, Pattern Analysis Aplication, Vol. 14 (2011) 87–92

[7] Siddiqi, I., Djedd; C., Raza, A.: Automatic analysis of handwriting for gender classification, Pattern Anal Applic (2014)

[8] Bouadjenek, N., Nemmour, H., and Chibani, Y.: Local Descriptors to Improve Off-line Handwriting-based Gender Prediction, International Conference of Soft Computing and Pattern Recognition (2014)

[9] Guerbai, Y., Chibani, Y., and Hadjadji, B., Handwriting Gender Recognition System Based on the One-Class Support Vector Machines, International Conference on Image Processing Theory, Tools and Applications, IPTA, Canda (2017)

[10] Cha, S., and Srihari, S, N.: On measuring the distance between histograms, Pattern Recognition, Vol. 35, No. 6, (2002) 1355-1370

[11] Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, Vol. 1,(2005). 886–893.

[12] Donoho, D, L., and Duncan, M, R.: Digital curvelets transform: strategy, implementation and experiments, Stanford University (1999)

[13] Candès, E, J.: Ridgelets: Theory and Applications, Ph.D. thesis, Department of Statistics, Stanford University (1998)

[14] Feng, K., Jiang, Z., He, W., and MA, B.: A recognition and novelty detection approach based on Curvelet transform, nonlinear PCA and SVM with application to indicator diagram diagnosis, Expert Systems with Applications, Vol. 38, No. 10, (2001) 12721-12729

[15] Li, Y., Yang, Q., and Jiao, R..: Image compression scheme based on curvelet transform and support vector machine, Expert Systems with Applications, Vol. 37, No. 4, (2009) 3063-3069

[16] Schölkopf, B., Platt, J., Shawe-Taylor, J., Smola, A., andWilliamson, R.: Estimating the support of a high dimensional distribution, Neural Computation, Vol. 13, No. 7, (2001). 1443-1472

[17] Shirdhonkar, M, S., and Kokare, M.: Off-line handwritten signature retrieval usingcurvelet transforms, International Journal of Computer and Engineering, Vol. 3 , No. 4, (2001) 1658-1665

[18] Vapnik, V.: The nature of statistical learning theory. Springer Verlag, New York, USA, (1995),

[19] Bergamini, C., Oliveira, L, S., A. L. Koreich, A, L., and Sabourin, R.: Combining different biometric traits with one-class classification, Signal Processing, Vol. 89, (2009). 2117- 2127

[20] Rabaoui, A., Davy, M., Rossignol, S., Ellouze, N.: Using one-class svms and wavelets for audio surveillance, IEEE Transaction on Information Forensic and Security, Vol. 3 No 4, (2008), 763-775

[21] Cha, S., and Srihari, S, N.: On measuring the distance between histograms, Pattern Recognition, Vol. 35, No. 6, (2002) 1355-1370

[22] Guerbai, Y., Chibani, Y., and Hadjadji, B.: The effective use of the one-class SVM classifier for handwritten signature verification based on writer-independent parameters, Pattern Recognition , Vol. 48, (2015) 103–113

[23] Kuncheva, L, L., 2004.: combining pattern classifiers :methods and algorithms , Wiley-Interscience publication, New Jersey, USA (2004)

[24] Guerbai, Y., Chibani, Y. Meraihi, Y., 2022, Techniques for Selecting the Optimal Parameters of One-Class Support Vector Machine Classifier for Reduced Samples, Vol 13 N°1

Bearing fault classification based on Fast Fourie Transform and 1D-CNN

Toumi Yassine
Laboratory of Advanced Electronic Systems, University Yahia Farès of Medea, Algeria
toumiyacine12@gmail.com

Meghraoui djamila
Laboratory of Speech Communication and Signal Processing, University USTHB, Algeria
djamila.meghraoui@gmail.com

Bengherbia Billel
Laboratory of Advanced Electronic Systems, University Yahia Farès of Medea, Algeria
bil.elec@gmail.com

Bouchareb Nariman
Laboratory of Advanced Electronic Systems, University Yahia Farès of Medea, Algeria
narimanbouchareb1998@gmail.com

Toumi Hayet
Laboratory of Mechanics, Physics and Mathematical, University Yahia Farès of Medea, Algeria
allae_atchi@yahoo.fr

Hamza Benyezza
Laboratory of Advanced Electronic Systems, University Yahia Farès of Medea, Algeria
hamzabenyezza@yahoo.fr

Abstract— **Bearings are one of the most widely used components of rotary machines. Identifying defective bearings in a timely manner may avoid unexpected faults that may cause damage to humans and materials. In recent years, Convolutional Neural Network (CNN) has proven to be a useful technique to identify and classify rolling bearing defects. In this work, we propose an automatic classification of bearing faults using one-Dimensional CNN (1D-CNN) algorithm based on the Fast Fourier transform (FFT) to extract features from the vibration signals. The proposed method is tested on the Paderborn University bearing dataset with 99.98% accuracy. In addition, 5-fold cross validation is used to validate the performance of the proposed model, which showed an average accuracy of 99.89%.**

Keywords—rolling bearing, one-dimensional convolutional neural network, Fast Fourier transform, k-fold cross validation.

I. INTRODUCTION

Rotating machinery is widely used in different industrial fields, including electromechanical drive systems and mechanical power transmission systems, which leads to the deterioration and failure of these machines for several reasons, such as loads. Rolling element bearings are essential components in rotating machinery, and their deterioration can account for up to 44% of defects in rotating machines [1]. To explore the bearing status, various signals are collected and used, among these signals, the vibration signals are the most used [2].

Bearing faults identification includes two steps, feature extraction and faults classification. Many signal processing techniques are used to feature extraction including, Time-domain statistical analysis, Fast Fourier Transform (FFT), envelope analysis, wavelet transformation (WT), and empirical mode decomposition (EMD). Common Machines learning (ML) includes K-Nearest Neighbors (KNN), Random Forest (RF), Support Vector Machine (SVM) and Neural Network approaches used for the bearing fault classification [3]. Recently, Deep Learning (DL) methods have been used in the case of bearing faults classification because of their ability to extract information of interest and to ensure robust decision making [4].

In this work, FFT is used as a technique for the feature extraction of vibration data and the 1D-CNN as a classifier to identify bearing faults. The remainder of this document is structured as follows: The proposed method, including the FFT and the 1D-CNN, are presented in the second section. The proposed method was validated using the Paderborn university bearing dataset, which is described in part three. The proposed architecture of the CNN is shown in section four. The experimental results are presented in section five. In the final section, some conclusions are made.

II. PROPOSED METHOD

A. Fast Fourier Tronsform

Fast Fourier Transform is a crucial measurement technique in various domains. It provides frequency information about the signal by converting a signal into individual spectral components. in reality, the FFT is an optimized algorithm for the implementation of the "Discrete Fourier Transformation" (DFT). A signal is sampled over a period of time and divided into its frequency components. These components are single

sinusoidal oscillations at distinct frequencies each with its own amplitude and phase [5]. the DFT transformation is calculated by the following equation:

$$X(K) = \sum_{n=0}^{N-1} x(n)\, e^{-i2\pi n \frac{K}{N}} \quad (1)$$

B. One Dimensional Convolutional Neural Networks

Convolutional Neural Networks (figure1) are a class of feedforward neural networks consisting of two main parts, the convolution part and the fully-connected part. The convolution part comprises a convolutional layer, activation layer, pooling layer and batch normalization if required. The convolution layer is composed of many convolution kernels (filters), these kernels convolve with the input data to generate feature maps and mapped with the activation function to the new feature information. In addition, the Rectified Linear Unit (ReLU) is often used to perform the activation process, in the purpose to prevent gradient explosion or gradient dispersion [5]. The second major step of the convolution part is the polling layer, which is used to reduce and refine the output feature information via the computational statistics method. The fully-connected part is the classifier, its goal is to map the feature information learned by the network to the label space of the samples.

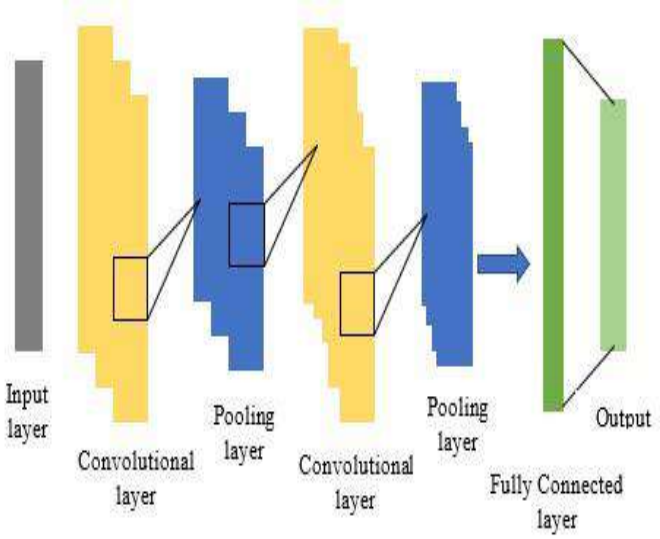


Fig.1. Design of the One-Dimensional Convolutional Neural Network

III. DATASET DESCRIPTION

The Paderborn bearing dataset [6] was provided by the Paderborn University Bearing Data Center. The test rig of this dataset is displayed in Figure 2. This dataset has been generated using 32 bearings, including six healthy bearings, 12 bearings with artificial damages and 14 bearings with real damages generated from accelerated lifetimes tests. The artificial defects have been produced on both, the inner and outer races by using different techniques including drilling, EDM, and electric engraving machine. The vibration data were collected under various working conditions as shown in table 1 with a sampling frequency of 64 kHz and a sampling time of 4s, which means there are 256,000 data points. Furthermore, each bearing is used 20 times to generate 20 signals with one fixed combination. For the purpose of this study, we use the first combination of the signals shown in table 2, each signal has been divided into 125 sub-signal of 2048 points. In total 22500 signals have been used for classification, 70% of these signals used for the training and the remaining signals for the test.

Table 1: working conditions

NO	Relational speed (rpm)	Load torque (Nm)	Radial force (N)	Signal name
1	1500	0.7	1000	N15_M07_F10
2	900	0.7	1000	N09_M07_F10
3	1500	0.1	1000	N15_M01_F10
4	1500	0.7	400	N15_M07_F04

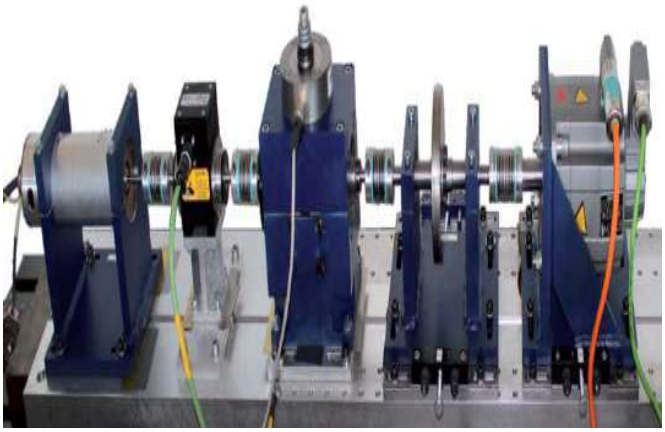


Fig.2. The experimental test rig.

Table 2: List of data sets used

Bearings faults	Healthy	Inner	Outer
Signals code	K001, K002, K003	KA01, KA04, KA07	KI01, KI04, KI07

IV. PROPOSED ARCHITECTURE

The proposed architecture consists of 2 convolution layers, 2 pooling layers, and 5 dense layers as presented in table 1. The first CNN layer is a convolutional layer consisting of 32 filters with a kernel size of 3 along with two strides and padding with ReLU as an activation function. The second convolutional layer is composed of 16 filters with the same kernel size and two strides with padding and ReLU as an activation function. The pooling layers are max pooling layers having a pool size of 2 with two strides, and each max pooling layer has been added after a convolution layer. The results are then flattened using a flattening layer. Afterwards, five dense layers with sizes of 64, 32, 16, and 8 respectively were employed using the ReLU activation function. In the end, a dense layer with the size of 3 and SoftMax activation function was used, in order to classify the three types of bearing faults.

Table 2: 1D-CNN model details.

Layer type	Output shape	Parameters number
Conv1D	(1024,32)	128
Max pooling	(512,32)	0
Conv1D	(512,16)	1552
Max pooling	(256,16)	0
Flatten	4096	0
Dense	64	262208
Dense	32	2080
Dense	16	528
Dense	8	136
Dense	3	27

V. EXPERIMENTAL RESULTS

Table 3: The Accuracy, Recall and F1-score rates of each faults type

Classes	0	1	2
Precision	0.9996	1	1
Recall	1	1	0.9996
F1-score	0.9998	1	0.9998
Accuracy	99.98%		

Table 3 and figure 3 mention the Accuracy, Recall and F1-score rates of each fault type, where all the classes have an accuracy rate of 99.98%, and F1 score close to 1, which shows a perfect classification result. The loss and accuracy curve for the 1D-CNN are provided in figure 4 and figure 5 respectively. The accuracy curve indicates that the 1D-CNN model achieves an accuracy of 100% for the training and 99,98% for the testing models.

In order to further check the performance of the proposed model for the bearing faults classification, a 5-fold cross validation procedure is used. the testing accuracy for the 5 folds is presented in figure 6. The average accuracy is 99.89%, demonstrating another successful classification.

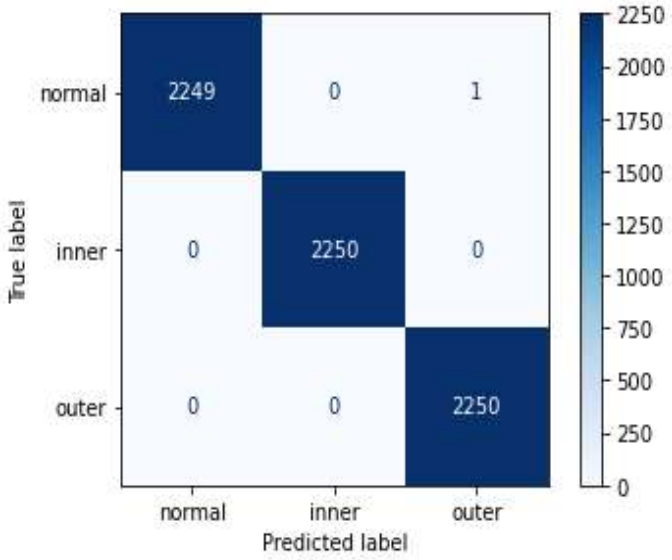


Fig.3. Confusion matrix of the 1D-CNN.

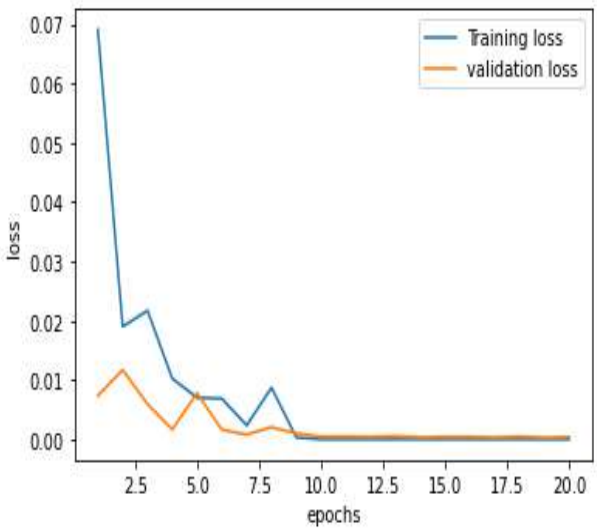


Fig.4. Loss curve of the 1D-CNN

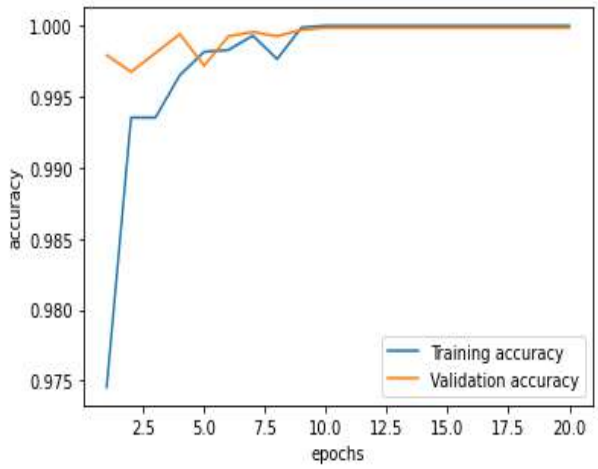


Fig.5. Accuracy curve of the 1D-CNN

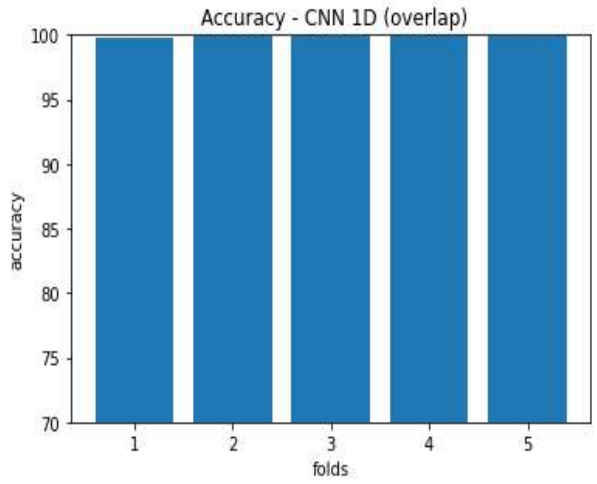


Fig.6. Testing accuracy for the 5 folds.

VI. Conclusion

This work has proposed an approach for bearing fault identification and classification. In this context, the Fast Fourier Transform was used to extract features from the bearing signals vibration. For classification purposes, a one-dimensional convolution neural network was applied. The findings demonstrate the efficacy of the suggested method for bearing faults classification with an accuracy of 99.98%. Furthermore, the cross validation under different 5 folds proves the effectiveness of the proposed model for the bearing faults classification.

References

[1] I. Attoui, B. Oudjani, N. Boutasseta, N. Fergani, M. Bouakkaz, and A. Bouraiou, "Novel predictive features using a wrapper model for rolling bearing fault diagnosis based on vibration signal analysis," *Int. J. Adv. Manuf. Technol.*, vol. 106, pp. 3409–3435, 2020.

[2] G. Tingarikar and A. Choudhury, "Vibration Analysis - Based Fault Diagnosis of a Dynamically Loaded Bearing with Distributed Defect," *Arab. J. Sci. Eng.*, 2021.

[3] C. Liu, A. Cichon, G. Królczyk, and Z. Li, "Technology development and commercial applications of industrial fault diagnosis system: a review," *Int. J. Adv. Manuf. Technol.*, 2021.

[4] A. Shenfield and M. Howarth, "A novel deep learning model for the detection and identification of rolling element-bearing faults," *Sensors*, vol. 20, no. 18, pp. 1–24, 2020.

[5] Y. Qin and X. Shi, "Fault Diagnosis Method for Rolling Bearings Based on Two-Channel CNN under Unbalanced Datasets," *Appl. Sci.*, vol. 12, no. 17, 2022.

[6] "Konstruktions- und Antriebstechnik (KAt) - Data Sets and Download (Universität Paderborn)." [Online]. Available: https://mb.uni-paderborn.de/en/kat/main-research/datacenter/bearing-datacenter/data-sets-and-download. [Accessed: 21-Feb-2023].

Solar tracking system based on fuzzy logic control

Fadhila LACHEKHAB
Automatic applied and signal processing laboratory
M'Hamed Bougara University
Boumerdes, Algeria
f.lachekhab@univ-boumerdes.dz

Sid Ahmed TADJER
Electrification of industrial entreprises laboratory
M'Hamed Bougara Boumerdes
Boumerdes, Algeria
s.tadjer@univ-boumerdes.dz

Dalila ACHELI
Automatic applied and signal processing laboratory
M'Hamed Bougara University
Boumerdes, Algeria
d.acheli@univ-boumerdes.dz

Abstract— **To increase the power efficiency of a photovoltaic module or an array of photovoltaic modules, an electronic controller is integrated between the PV generator and the load. Its main role is to maximize the power delivered by the photovoltaic panels by continuously tracking the position of the sun (solar tracker) and the maximum power point. This technique is commonly called MPPT (Maximum Power Point Tracking).**

In this work, we focus on the control of the solar tracker using a classical controller (P, PI, PID) on one hand, and a fuzzy controller (FIS) on the other hand, and we subsequently compare the performance of the system in each case. To determine the parameters of the classical controller, we used the Evans' root locus method. This method consists of studying the position of the characteristic polynomial roots as a function of a loop parameter, namely the gain k. Specifically, when the characteristic polynomial of the system has a variable parameter, the plot of the root locus as a function of this parameter allows us to know the evolution of the roots of the characteristic equation in the complex plane. Since the stability region is the unit circle centered at the origin, we can easily deduce from the root locus plot the parameters ensuring stability.

For the fuzzy controller, we used a base of 24 fuzzy rules. The results obtained using Simulink in Matlab showed that the fuzzy controller provides better performance than those obtained using the classical controller in terms of maximum overshoot, rise time and steady-state error.

Keywords— Renewable energies, photovoltaic solar energy, solar tracker, Classical controller, fuzzy controller, system performance.

I. INTRODUCTION

Currently, a large part of global energy production is derived from fossil fuels. However, the consumption of these sources results in greenhouse gas emissions and contributes to increasing pollution [1]. Furthermore, excessive consumption of natural resources endangers future generations, particularly in the face of multiple economic and oil crises. As a result, science has turned its attention to renewable resources.

Renewable energies are those that renew themselves quickly enough to be considered practically inexhaustible. The sun delivers the equivalent of 105 billion tons of oil per day, making solar energy an attractive option as it is renewable, non-polluting, and free. This energy is converted from sunlight into electricity through the use of semiconductor materials such as silicon.

However, the incident photovoltaic energy on cells is not optimal depending on the time of day and season. The efficiency of a photovoltaic module is relatively low (around 15%) [5], so it is essential to maximize the power generated by reducing the energy losses from the sun. Studies have shown that the optimal efficiency is obtained when the cells are perpendicular to the incident sunlight. To maintain this perpendicularity for as long as possible, we have opted for a sun tracking system [2].

The main objective of this work is to study and describe a sun tracking system to maximize the amount of absorbed light and, consequently, increase the efficiency of photovoltaic panels by following the movement of the sun throughout the day [7].

II. MODELING of photovoltaic conversion system

The production of electricity by a photovoltaic system depends on the solar radiation received by the panels that make up the system. This has prompted experts in this field to conduct several studies and research to improve the conversion efficiency of the panels, particularly by allowing them to receive more radiation by constantly tracking the movement of the sun in the sky. For this purpose, they have invented a system that enables them to track the position of the sun: the solar tracker.

The photovoltaic conversion is carried out according to the chain described in figure 1 :

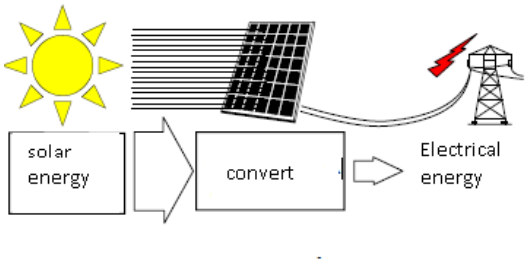


Fig. 1. The principle of photovoltaic conversion [2]

A solar tracker is a device that allows a solar power installation to follow the sun according to the principle of the heliostat, which enables solar panels to be oriented throughout the day. Indeed, many studies focus on sun tracking to maximize sunlight by keeping the active surface of the solar module perpendicular to the solar radiation. [24]

There are two main types of solar trackers: passive and active, which include single-axis and dual-axis trackers.

In this work, we are interested in the type 1 axis tracking system, which is an electromechanical device with one degree of freedom, allowing it to follow the path of the sun during the day.

It consists of two mechanisms: a light detection circuit that determines the position of the sun and an electric motor that rotates the solar panel [6].

A DC motor is represented by the following diagrams:[3]

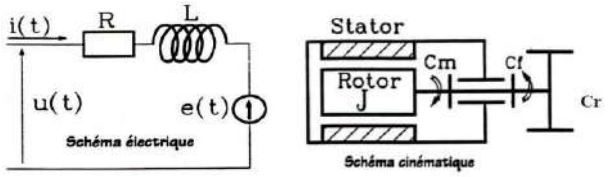


Fig. 2. The electrical and kinematic diagram of a DC electric motor.

$$\begin{cases} u(t)=e(t)+Ri(t)+L\dfrac{di}{dt} \\ j\dfrac{dw}{dt}=C_m(t)-C_r(t) \\ C_m(t)=K_c i(t) \\ C_r(t)=aw(t) \end{cases} \quad (1)$$

As we are interested in the position, we will add a differential equation to the equations presented earlier. Therefore, the model of a solar tracker is given by the following system of equations:

$$\begin{cases} u(t)=K_e w(t)+Ri(t)+L\dfrac{di}{dt} \\ j\dfrac{dw}{dt}=K_c i(t)-aw(t) \\ w(t)=\dfrac{d\theta}{dt} \end{cases} \quad (2)$$

The model of a photovoltaic cell is given by:

$$I_{cell}=I_{ph}-I_s(\exp(V_D/nV_0)-1) \quad (3)$$

According to the electrical diagram:

$$V_D=V_{cell}+R_s I_{cell} \quad (4)$$

Hence, the current of a photovoltaic cell is given by:

$$I_{cell}=I_{ph}-I_s(\exp((V_{cell}+R_s I_{cell})/nV_0)-1) \quad (5)$$

The voltage of the cell is given by:

$$V_{cell}=V_D-R_s I_{cell}$$

Where

$$V_D = nV_0 \ln\left(\frac{I_{ph}-I_{cell}+I_s}{I_s}\right)$$

$$V_{cell} = -R_s I_{cell} + nV_0 \ln\left(\frac{I_{ph}-I_{cell}+I_s}{I_s}\right) \quad (6)$$

III. CONTROL OF PHOTOVOLTAIC SOLAR PANELS

After developing the solar tracker model, we proceeded to simulate it using Matlab's Simulink. The figure below represents the developed diagram :

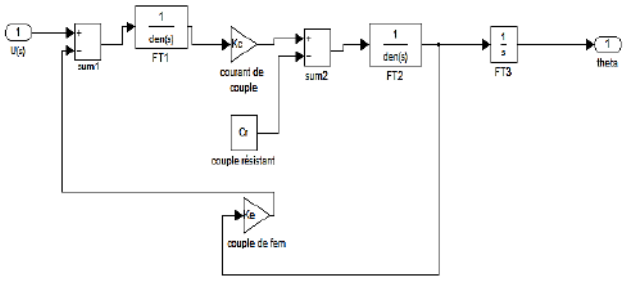


Fig. 3. Simulation diagram of a solar tracker using Matlab

The control of the solar tracker is based on regulating the position of the solar panel according to the position of the sun. We used a DC motor for this regulation. For the initial simulations, we used a classic proportional (P) controller, then a proportional-derivative (PD) controller, and finally a proportional-integral-derivative (PID) controller. The system responses obtained are presented in the figures below:

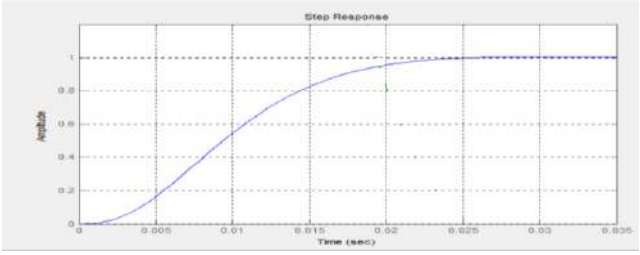


Fig. 4. La réponse du suiveur solaire avec régulateur P

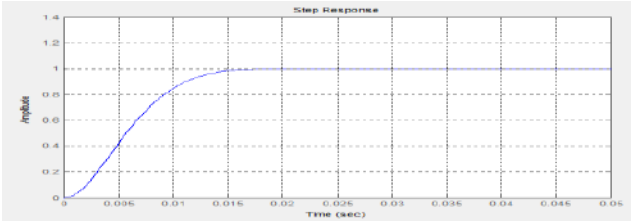


Fig. 5. La réponse du suiveur solaire avec régulateur PD

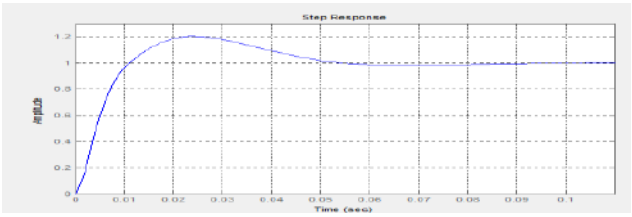


Fig. 6. La réponse du suiveur solaire avec régulateur PID

The following table summarizes the system performance values (response time tr, steady-state error ess, maximum overshoot) for each type of action of the classical controller used:

TABLE I. THE VALUES OF THE PERFORMANCE

	Action		
	P	***PD***	***PID***
tr	0.0256*s*	0.0187*s*	0.0111*s*
ess	0	0	0
D max	0	0	20%

For the control synthesis, we used the Evans root locus method. This method consists of studying the position of the characteristic polynomial roots as a function of a loop parameter, in this case, the gain k. This technique gave rise to the idea of pole placement principles, which allow for the adjustment of the overall dynamics. [3]

The plot of a root locus is based on two main conditions:

- Angle condition
- Magnitude condition

The following figures show the root locus plots for each type of controller:

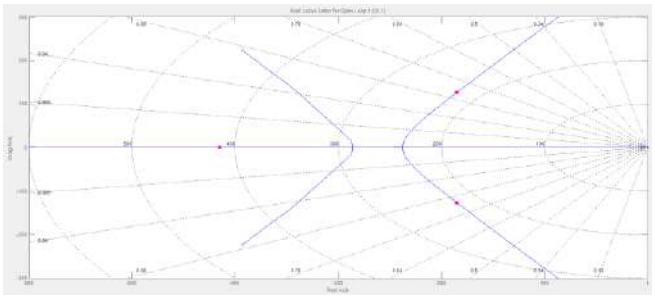


Fig. 7. Root locus plot of the solar tracker with a P controller

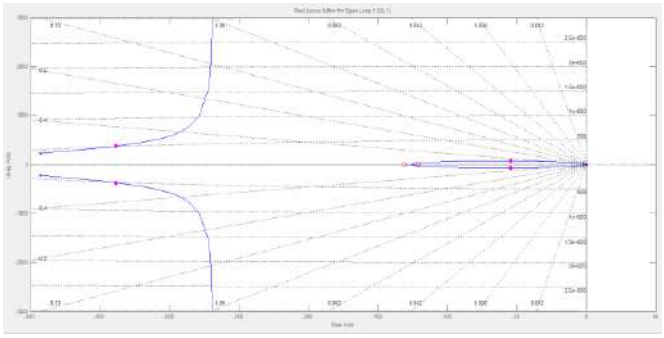


Fig. 8. Root locus plot of the solar tracker with a PD controller.

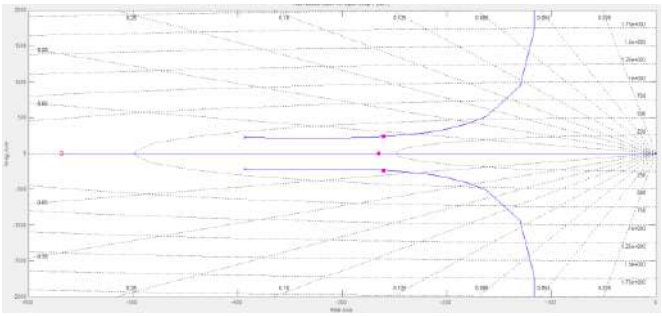


Fig. 9. Root locus plot of the solar tracker with a PID controller.

IV. FUZZY CONTROL

The regulation of the solar tracker using fuzzy logic is based on fuzzy [4]. We programmed this controller in MATLAB by introducing the membership functions of the inputs and outputs, as well as the inference rules.

The solar tracker regulation loop based on fuzzy logic in MATLAB is given by:

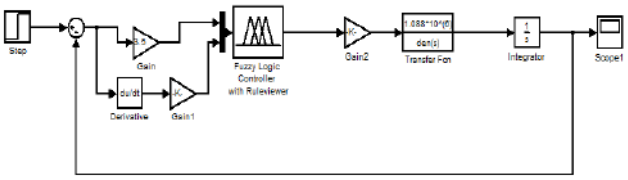


Fig. 10. The solar tracker regulation loop based on fuzzy logic in MATLAB.

V. THE INFERENCE RULES BASE

The fuzzy controller has two inputs, e and de respectively, and one output, u. The inference rules base is given as follows:

1. If (e is BN) and (de is BN) then (u is BP)
2. If (e is BN) and (de is SN) then (u is BP)
3. If (e is BN) and (de is Z) then (u is MP)
4. If (e is BN) and (de is SP) then (u is SP)
5. If (e is BN) and (de is BP) then (u is Z)
6. If (e is SN) and (de is BN) then (u is BP)
7. If (e is SN) and (de is SN) then (u is MP)
8. If (e is SN) and (de is Z) then (u is SP)
9. If (e is SN) and (de is SP) then (u is Z)
10. If (e is SN) and (de is BP) then (u is SN)
11. If (e is Z) and (de is BN) then (u is MP)
12. If (e is Z) and (de is SN) then (u is SP)
13. If (e is Z) and (de is Z) then (u is Z)
14. If (e is Z) and (de is SP) then (u is SN)
15. If (e is Z) and (de is BP) then (u is MN)
16. If (e is SP) and (de is BN) then (u is SP)
17. If (e is SP) and (de is Z) then (u is SN)
18. If (e is SP) and (de is SP) then (u is MN)
19. If (e is SP) and (de is BP) then (u is BN)
20. If (e is BP) and (de is BN) then (u is Z)
21. If (e is BP) and (de is SN) then (u is SN)
22. If (e is BP) and (de is Z) then (u is MN)
23. If (e is BP) and (de is SP) then (u is BN)
24. If (e is BP) and (de is BP) then (u is BN)

The simulation of this regulation loop allows us to obtain the following response:

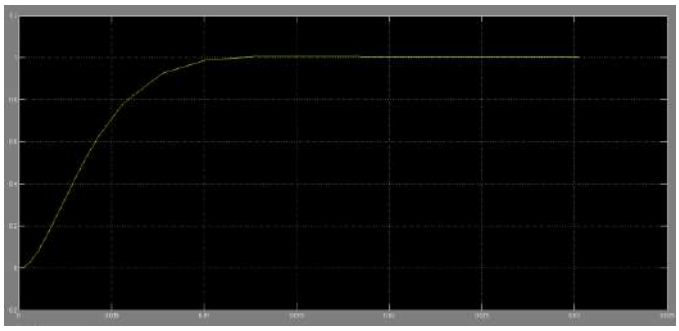


Fig. 11. The response of the solar tracker system

The system performance is : tr=0.012s ess=0 D=0%

After studying the "solar tracker" system, we found that the fuzzy controller provides better system performance compared to the performance given by different classical controllers.

CONCLUSION

In this work, after reviewing the fundamental concepts of solar tracking technology, we noticed through the simulation results that the use of a mobile system could significantly increase the amount of PV energy produced when solar tracking is performed along the East-West axis.

Based on the results obtained, we found that the fuzzy controller gave better performance to the system than the classical controller:

- Zero overshoot
- Small response time
- Zero steady-state error

REFERENCES

[1] S.Kumar. Design, "development and performance test of an automatic two-Axis solar tracker system". 2011 Annual IEEE India Conference, Hyderabad, India, , pp. 1-6, 2011. DOI 10.1109/INDCON.2011.6139586.

[2] M. Rebhi, M. Sellam, A. Belghachi and B. Kadri, "Conception and Realization of Sun Tracking System in the South-West of Algeria", Applied Physics Research, vol. 2, no. 1, 2010.

[3] M. Hossein, K. Alireza, J. Arzhang, A. Karen and S. Ahmad, "A Review of principle and sun-tracking method for maximizing solar system output", Renewable and Sustainable Energy Review, vol. 13, no. 8, pp. 1800-1818, 2009.

[4] Hung-Ching Lu and Te-Lung Shih, "Fuzzy system control design with application to solar panel active dual-axis sun tracking system", IEEE international conference on systems man and cybernetics, pp. 1878-1883, 2010.

[5] Huang, Cong-Hui, Heng-Yau Pan, and Kuan-Chen Lin, "Development of intelligent fuzzy controller for a two-axis solar tracking system." Applied Sciences, 2016.

[6] Abadi, Imam, Adi Soeprijanto, and Ali Musyafa. "Design of single axis solar tracking system at photovoltaic panel using fuzzy logic controller." 5th Brunei International Conference on Engineering and Technology (BICET), 2014

[7] Zhai, Pei, and Eric D. Williams. "Analyzing consumer acceptance of photovoltaics (PV) using fuzzy logic model." Renewable Energy, pp. 350-357. 2012.