



CHADLI BENDJEDID UNIVERSITY
FACULTY OF SCIENCE AND
TECHNOLOGY



THE 1ST INTERNATIONAL CONFERENCE ON AUTONOMOUS SYSTEMS AND THEIR APPLICATIONS

ICASA'22

ABDELMADJID BENMACHICHE

ALI ABDELLATIF BETOUIL

CHAOUKI CHEMAM

KHADIJA RAIS



**CHADLI BENDJEDID
UNIVERSITY**



24 MAY 2022

Conference Contributions

Paper Title	Authors
A new Deep Learning Algorithm for Face Emotional Recognition	D. A. Ouali , A. Achour, and M. Sahraoui
A new tracking and monitoring smart system based on intelligent algorithms for schoolchildren and mental patients	S. A. H. Belkhira , M. Rouissat, M. Belkheir, and M. Bouziani
A Novel Artificial Intelligence-Assisted Meta-heuristic Method for Augmenting the Annual Profit of Wind Farm	P. Bhattacharjee , R. K. Jana, and S. Bhattacharya
A particle swarm optimization approach for a robot in an environment with obstacles	I. Boutabia , A. Benmachiche, and M. Hammadi
A survey on disc and cup segmentation methods of retina fundus images	R. Touahri , N. Azizi, N. Hammami, and F. Benaida
An Approach for transforming Activity Diagram	H. Hachichi , H. Boualam, M. B. Errahmani, and O. Harbouche
Applying the ESQuMo Quality Model to Medical Imaging Embedded Software	Z. Tamrabet , T. Marir, and F. Mokhati
Assessing collaborative learning using collaborative quizzes	N. Gouasmi , L. Mahnane, Y. Lafifi, and A. Lebbad
Automatic real-time road extraction image based on color segmentation polynomial regression model algorithm	M. E. A. Bellebna and A. Boudjella
Bio-inspired optimization algorithms for improving a variable step size Maximum Power Point Tracking method for photovoltaic systems	K. E. Heraguemi and A. Zemmit
Compared methods of object tracking based on deep learning	R. Taglout and B. Saoud
Deep Learning, Big Data and High-Performance Computing to enhance Business and Marketing, while wearying of Fake News: Challenges, Opportunities, and Future Works	Z. Bouzidi , A. Boudries, and M. Amad
Deep learning-based semantic skills matching for resumes selection	M. F. Djouhri and H. Aouarib
Design and FPGA Implementation of Low-complexity Pruned DCT Approximation for Image Compression in Wireless Multimedia Sensors Networks	C. Araar
Extreme Learning Machine for the smart faults regression of turbofan engine	R. Nafissa and W. Rezgui
Extreme Learning Machine vs Back-Propagation for the faults smart classification of Tennessee Eastman Process	W. Rezgui and N. Rezki
Forecasting the New Cases of Tuberculosis in Algeria Using Random Forest as a Machine Learning Model	D. Tliba , N. Frissou, and M. T. Kimour

Paper Title	Authors
Hierarchical routing protocols (Energy-aware) for wireless sensor network: a survey	M. Hammadi and M. Redjimi
Human Facial Expression Recognition by Uniform Local Binary Patterns Approach & Wavelet Transform	N. Zermi , A. Khaldi, and R. Hamami
Hybrid Movie Recommendation System using LDA and PSO approaches	I. Gasmi , F. Anguel, and D. Merzougui
Implementation of the FIR filter under DSP and FPGA (comparative study)	F. Benaida , H. Beleili, N. E. Hammami, and R. Touahri
Integrating Reinforcement Learning with Analytic Hierarchy Process for Autonomous Service Web Composition	S. Zeggada , S. Aouag, and D. Hedjazi
Intelligent Control of Efficient Bridgeless SEPIC PFC converter Powered BLDC Motor	S. E. Halledj and A. Bouafassa
Modeling and Assessment of Energy Harvesting WSNs with Unreliable Sensors for Military Applications	N. Oukas and M. Boulif
Multi-classification of thyroid nodules using Machine learning and deep learning	L. Bellal and K. Khemis
New Inequality Index for Fair Multi-Agent Resource Allocation	A.-A. Betouil and A. Benmachiche
Object Detection and Classification in Videos via Optical Flow and VGGNet	M. Labeni , C. Boufenar, and M. Taffar
Phishing website detection using deep learning and machine learning	R. Zaimi , M. Hafidi, and L. Mahnane
RecTulips : A Python Library for Recommendation System	O. Remadnia , F. Maazouzi, and H. E. Azzag
Robust Cepstral Feature for Automatic Speech Recognition	M. C. A. Korba , H. Doghmane, and F. Maazouzi
Robust tracking system of vehicles on motion using Kalman filter at night mode	S. Bahroun , T. Sari, F. Benabbas, and N. Azizi
Service request monitoring using Petri nets and model at runtime	M. M. Bouhamed , O. Kamel, G. Diaz, and A. Chaoui
Sign Language Recognition	I. Zeroual and H. Azzag
Towards a Conceptual architecture for Enabling Self-Adaptive IoT-Based Systems: A Service-Based Approach	F. Z. Merabet and D. Bemmerzoug
Tumor Detection in Mammography Images using Discrete Wavelet Transform and Evidence Fusion Technique	A. Zitouni , F. Benkouider, F. Chouireb, and M. Reggab

A new Deep Learning Algorithm for Face Emotional Recognition

1st Djamel Amer Ouali

Informatics Department

Mohamed Boudief University of M'sila
M'sila, Algeria

djamelamerouali54@gmail.com

2nd Abdeslem Achour

Informatics Department

Mohamed Boudief University of M'sila
M'sila, Algeria

achabdou45@gmail.com

3rd Mohamed Sahraoui

LIAM laboratory

Mohamed Boudief University of M'sila
M'sila, Algeria

mohamed.sahraoui@univ-msila.dz

Abstract—Automatic recognition of human emotions has received increasing interest from researchers in the field of computer vision, which has led to the proposal of several methods. Many of them relied on handcrafted features and traditional fusion and classification techniques. The use of deep learning techniques to automatically extract powerful features from multimedia information as well as their use for merging and classification are new trends that researchers are currently pursuing. In this work, we define a new accurate facial expression detection algorithm based on a deep learning method, specifically on an intentional convolutional neural network capable of focusing on important parts of the face in an image or video database through the use of more number of needed layers. As a result, our proposed algorithm significantly improves the accuracy rate compared to previously proposed models in several datasets.

Index Terms—Deep Learning, CNN, Computer Vision, Image recognition.

I. INTRODUCTION

According to the remarkable development of smart and powerful devices, several related services are being developed. Almost, these services are designed to alternate human-centric services based on computer vision data extraction. one of the common fields is the automatic verification and recognition of individuals based on their physical appearance, behavioral traits and special traits, through biometric methods such as the face, fingerprint, signature and hand geometry. In particular, facial recognition technology can be used in many ways and in different manners ie. online or offline. One of the latest use of this technology is emotion recognition which can be performed using various features, such as in [1]–[3]. Among these features, facial expressions are the most common due to reason that it contains many features useful for recognizing emotions as it is shown in figure 1.

Emotion recognition is the use of computers to detect human faces and analyze performance characteristics information [4], [5]. The machine realizes the purpose of recognizing and understanding humans from emotional expression.

In order to reduce the false detection, improve the recognition accuracy and offer more robustness, machine learning techniques are introduced in this area and thus for recognizing emotions. Recently, with the use of deep learning in particular Convolutional Neural Networks (CNNs) [6], many features



Fig. 1. (Left to right) the six cardinal emotions (happiness, sadness, anger, fear, disgust, surprise) and neutral.

can be extracted, and learn how to have a decent facial expression recognition system. It should be noted that in the case of facial expressions, a lot of clues come from a few parts of the face, for example, the mouth and eyes, whereas other parts, such as the ears and hair, play a small role in the output of emotional recognition such as happiness, sadness or anger [7].

In this work, we propose a new framework based on deep learning to recognize facial expressions, which takes the above observation into account and uses the attention mechanism to focus on the prominent part of the face. By using an attentional convolutional network and unlike the used CNN algorithms that are performed using at most four layers, our algorithm bases on the use of additional layers capable of achieving a very high-resolution rate. more specifically, this paper makes the following contributions:

- We propose an approach based on the attentional convolutional network, which can focus on the feature-rich parts. However, it outshines the face in modern brilliant works in precision.
- In addition, we train him on different data with an evaluation of the accuracy rate.

In the following sections, we first provide an overview of the related work in Section. Two, the structure of the model and the proposed framework are explained in Section three. We will then present the experimental results, an overview of the databases used in this work, as well as a typical visualization

in Section four. Finally, we conclude the paper in Section Five.

II. RELATED WORKS

Emotion recognition can be divided into two main steps. In the first step, some features are extracted from the image, however in the second step, a classifier is used to detect emotions (e.g. SVM or neural network).

Some common used techniques features extraction in order to recognize facial expressions are directed gradient graph (HOG) [8], [9], local binary patterns (LBP) [10], Gabor waves [11] and Haar features [12].

The authors of [2] have proposed a neural network for FER using two convolutional layers, one pooling layer, and a subnet one.

In [7], a deep learning approach to FER based on the attentional alum comprising four convolutional layers, every two layers followed by a maximal pooling layer and an activation function of the straight linear unit (ReLU). Then they are followed by a drop-down layer and two fully connected layers.

The authors of [13] have worked on CNN deep network where they used 10 tags to rename each image in the data set, used different DCNN cost functions, and achieved sufficient accuracy.

An additional enhancement to CNN (IB-CNN) is proposed in [14] in order to improve the recognition of automatic facial expressions by enhancing neuronal discrimination.

In [15], a realization of conscious identity CNN (IA-CNN) that uses sensitive identity and expression is proposed.

A new framework for deep learning called a spatiotemporal recurrent neural network (STRNN) is proposed in [16] to integrate distinct learning from spatiotemporal information about signal sources in a unified spatiotemporal dependency model.

A new peak experience deep network (PPDN) is proposed in [17] that uses a sample with peak expression (easy sample) to moderate the responses of intermediate traits to a non-peak expression sample (hard sample) of the same type and of the same subject.

All of the above works made significant improvements to the field of recognizing facial emotions, but it seems to be missing a little piece of feature extraction by the fact they focus only on the prominent areas of the face.

III. THE PROPOSED FRAMEWORK

In this section, we propose a comprehensive deep learning framework based on a convolutional neural network for the classification of emotions in visually illustrated facial images as presented in Figure 2.

The neural network is based on adding more layers that facilitate gradient flow in the network (e.g. by adding layer skipping), or better regularizations (e.g. spectral normalization). However, for facial expression recognition, since there are few classes, we show that using a convolutional network with less than 8 layers (trained from scratch) is able to achieve

promising results by multiplying the latest models in several databases.

The feature extraction part consists of four convolutional layers and three fully connected layers, the first convolutional layer contains 64 3x3 filters, the second has 128 5x5 filters, and the third layer has 512 3x3 filters and the last also has 512 3x3 filters. In All convolutional layers, we have stride size 1, batch normalization, dropout, max-pooling, and ReLU as activation functions, the hidden layer in the first FC layer contains 256 neurons, and the second FC layer contains 512 neurons, and the third FC layer contains 128 neurons.

In both fully connected FC layers, as in the convolutional layers, we use batch normalization, dropout and ReLU, and batch size of 64, loss layer after decision layer is used to determine how to penalize deviations between expected and true labels using many types of loss functions such as softmax and SVM and sigmoid cross. This model is trained by optimizing the loss function using the stochastic training approach (more specifically Adam's optimizer), for the loss function in this work, we used the softmax function.

IV. EXPERIMENTAL RESULTS

In this section, we present a detailed empirical analysis of our model in several facial expression recognition databases. We first provide a brief overview of the databases used in this work. Basing on these databases, we then present the performance and compare the obtained results of our model with some promising recent work, basing on the task of categorizing each face based on the emotion shown in the facial expression into one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral).

A. Databases

- The Facial Expression Recognition 2013 (FER2013) database was first introduced in the ICML 2013 workshop on challenges in Representation Learning [18]. The data consists of 48x48 pixel grayscale images of faces with 35887 images. The training set consists of 28709 examples and The public test set used for the leaderboard consists of 3589 examples. The faces have been automatically registered so that the face is more or less centered and occupies about the same amount of space in each image. Four sample images from FER dataset are shown in Figure 3.
- The Extended Cohn-Kanade Dataset (CK+) [19] is a complete dataset for action unit and emotion-specified expression. CK+ dataset contains 593 video sequences from a total of 123 different subjects, ranging from 18 to 50 years of age with a variety of genders and heritage. Each video shows a facial shift from the neutral expression to a targeted peak expression, recorded at 30 frames per second (FPS) with a resolution of either 640x490 or 640x480 pixels. Out of these videos, 327 are labeled with one of seven expression classes: anger, contempt, disgust, fear, happiness, sadness, and surprise. The CK+ database is widely regarded as the most extensively

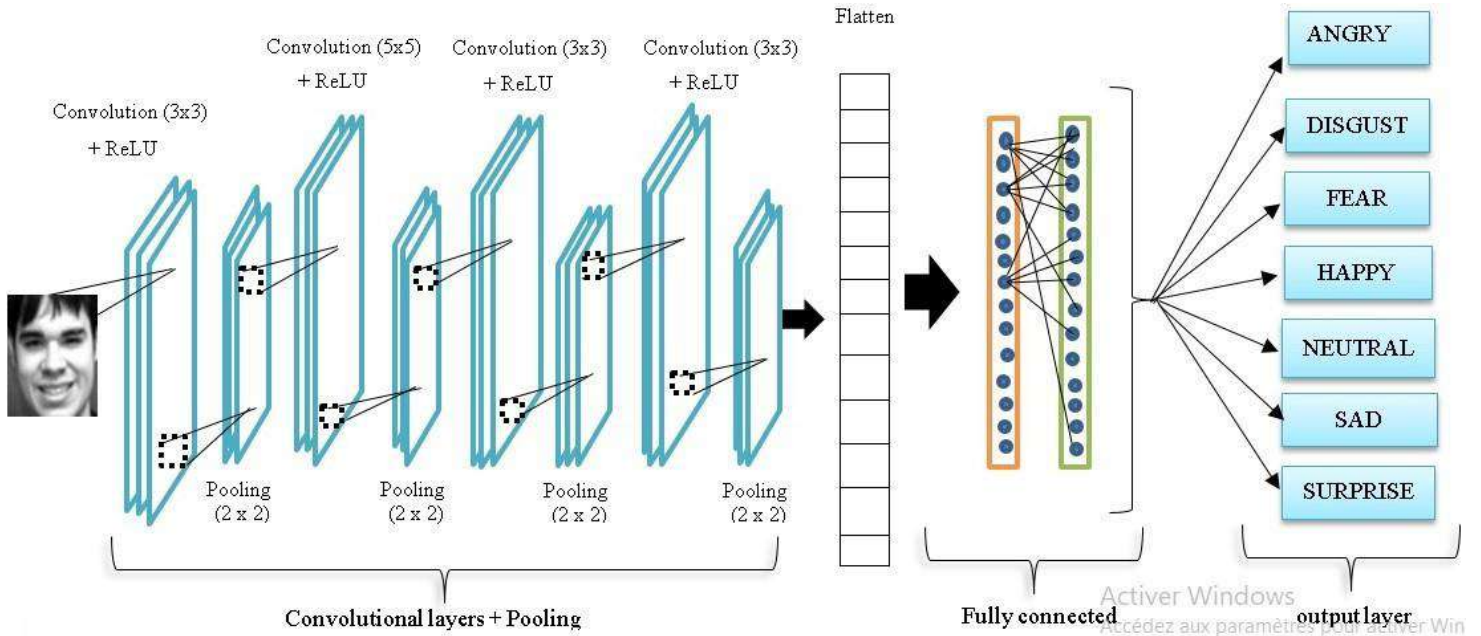


Fig. 2. The proposed Method architecture.



Fig. 3. Four sample images from FER database.

used laboratory-controlled facial expression classification database available, and is used in the majority of facial expression classification methods. Six sample images from this dataset are shown in Figure 4.

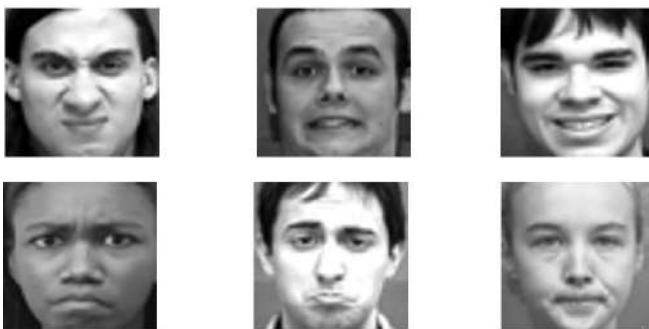


Fig. 4. Six sample images from CK+ database.

- The Japanese Female Facial Expression dataset [20] consists of 213 images of different facial expressions from 10 different Japanese female subjects. Each subject

was asked to do 7 facial expressions (6 basic facial expressions and neutral) and the images were annotated with average semantic ratings on each facial expression by 60 annotators. Sample images from this dataset are shown in Figure 5.



Fig. 5. Six sample images from JAFFE database.

B. Results Comparison and Analysis

For check the Performance of Proposed model on the previous explained datasets, we have trained the different CNN models then reported the accuracy. Therefore, we have trained each model over the three datasets. Furthermore, we have used various hyper-parameters to fine-tune the model and train it on kaggle platform . For FER-2013 dataset the different algorithms have trained for 50 epochs. For JAFFE and CK+ datasets, they have trained for 140 epochs from scratch. For optimisation of our algorithm, we have used Adam optimizer with a learning rate of 0,001 with weight decay. It takes around

4 hours to train our models on FER dataset. For JAFFE and CK+, it takes a few minutes (from 5 to 10 minutes) to train a model since they have much fewer images.

For FER-2013 dataset, we have used around 28,709 images for train the model, 35,888 for validation, and 3,589 for testing. At this end, our algorithm is succeed to achieve the best accuracy through an accuracy rate of around 96.04 %. The comparison of the result of our model with some of previous works on FER 2013 are provided in Table I and Figure6.

TABLE I
OBTAINED ACCURACIES ON FER 2013 DATASET

N°	Algorithm	Accuracy rate
1	Bag of Words [21]	67.4 %
2	VGG+SVM [22]	66.31%
3	GoogleNet [23]	65.2 %
4	Mollahosseini et al [2]	66.4 %
5	Gabor Wavelets [24]	70.02 %
6	Our algorithm	96.04 %

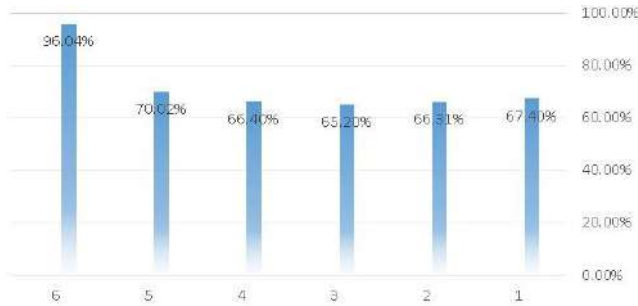


Fig. 6. Obtained Accuracies on FER 2013 dataset.

As shown in the table I, our proposed algorithm realize an improvement of the accuracy rate by an average of 20.02 % compared to the other algorithms. This fact is coming due to the use of multiple enhanced strategies with more number of levels responsible to ensure more accuracy rate.

For JAFFE dataset we have used 120 images for training, 23 images for validation and 70 images for testing. As result, our algorithm is succeed to achieve an accuracy rate of around 96.20The comparison with previous works on JAFFE dataset are shown in Table II as well as in Figure 7.

TABLE II
OBTAINED ACCURACIES ON JAFFE DATASET

N°	Algorithm	Accuracy rate
1	VGG+SVM [22]	89.02%
2	GoogleNet [23]	91.8 %
3	Mollahosseini et al [2]	95.31 %
4	Gabor Wavelets [24]	92.8 %
5	Our algorithm	96.20 %

As shown in the table II, our proposed algorithm realize an improvement of the accuracy rate by an average of 4.3 % compared to the other algorithms.

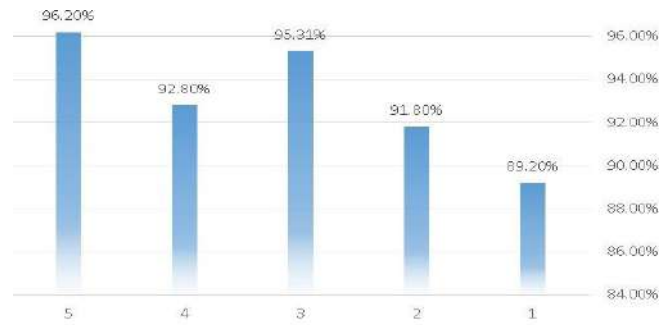


Fig. 7. Obtained Accuracies on JAFFE dataset.

For CK+ dataset we use 70% of images for train the model, 10% for validation, and 20% for testing. We were able to achieve an accuracy rate of around 99.16%. The comparison of our model with previous works on the extended CK dataset are shown in Table I and Figure 8.

TABLE III
OBTAINED ACCURACIES ON CK+ DATASET

N°	Algorithm	Accuracy rate
1	VGG+SVM [22]	95.37%
2	GoogleNet [23]	97.2 %
3	Mollahosseini et al [2]	97.03 %
4	Gabor Wavelets [24]	98.03 %
5	Our algorithm	98.16 %

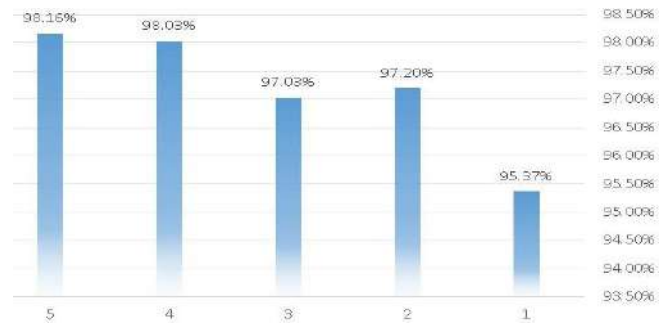


Fig. 8. Obtained Accuracies on CK+ dataset.

As shown in the table III, our proposed algorithm realize an improvement of the accuracy rate by an average of 1.25 % compared to the other algorithms.

V. CONCLUSION

Automatic recognition of human emotions has received increasing interest from computer vision researchers and several solutions have been proposed in this field. Using deep learning techniques, we proposed a new framework for recognizing facial expressions using a convolutional neural network. We conducted an extensive empirical analysis of our proposed algorithm on four well-known databases of facial expressions, which achieved good results and demonstrated good efficiency

and speed in facial expression identification. As a future work, we propose to enhance our solution to be applied in 3D images by integrating more layers with an efficient CNN extension.

REFERENCES

- [1] A. Deepali, A. Colburn, G. Faigin, L. Shapiro, and B. Mones, "Modeling stylized character expressions via deep learning," in *Asian Conference on Computer Vision*, pp. 136-153, 2016.
- [2] M. Ali, D. Chan, and M.H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," in *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on. IEEE*, 2016.
- [3] L. Ping, S. Han, Z. Meng, and Y. Tong, "Facial expression recognition via a boosted deep belief network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1805-1812, 2014.
- [4] K. He, G. Gkioxari, P. Dollr, and R. Girshick, "Mask r-cnn," in *International Conference on Computer Vision*, pp. 2980-2988, IEEE, 2017.
- [5] S. Minaee, A. Abdolrashidiy, and Y. Wang, "An experimental study of deep convolutional features for iris recognition," in *Signal Processing in Medicine and Biology Symposium, IEEE*, 2016.
- [6] M. Polignano, M. Giuseppe de Pinto, P. Lops, and G. Semeraro, "Identification Of Bot Accounts In Twitter Using 2D CNN," in *User-generated Contents Notebook*, University of Bari ALDO MORO, Italy, 2019.
- [7] S. Minaee1, and A. Abdolrashidi, "Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network," *Expedia Group, University of California*, 2019.
- [8] M. Hough, and V.C. Paul, "Method and means for recognizing complex patterns," in *U.S. Patent*, December 1962.
- [9] C. Junkai, Z. Chen, Z.u Chi, and H. Fu, "Facial expression recognition based on facial components detection and hog features," in *International workshops on electrical and computer engineering subfields*, pp. 884-888, 2014.
- [10] S. Caifeng, S. Gong, and P.W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," in *Image and vision Computing*, vol 27.6, pp. 803-816, 2009.
- [11] B., Marian Stewart, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, "Recognizing facial expression: machine learning and application to spontaneous behavior," in *Computer Vision and Pattern Recognition*, vol. 2, pp. 568-573. IEEE, 2005.
- [12] W. Jacob, and C.W. Omlin, "Haar features for faces au recognition," in *7th International Conference on Automatic Face and Gesture Recognition*, pp. 5-12. IEEE, 2006.
- [13] E. Barsoum, C. Zhang, C.C. Ferrer, and Z. Zhang, "Training Deep Networks for Facial Expression Recognition with Crowd-Sourced Label Distribution," in *Proceedings of the 18th ACM International Conference on Multi-modal Interaction*. ACM, 2016.
- [14] H. Shizhong, Z. Meng, A.S. Khan, and Y. Tong, "Incremental boosting convolutional neural network for facial action unit recognition," in *Advances in Neural Information Processing Systems*, pp. 109-117, 2016.
- [15] M. Zibo, P. Liu, J. Cai, S. Han, and Y. Tong, "Identity-aware convolutional neural network for facial expression recognition," in *IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 558-565. IEEE, 2017.
- [16] T. Zhang, W. Zheng, Z. Cui, Y. Zong, and Y. Li, "Spatial-temporal recurrent neural network for emotion recognition," in *IEEE transactions on cybernetics*, vol 99, pp. 1-9, 2018.
- [17] Z. Xiangyun, X. Liang, L. Liu, T. Li, Y. Han, N. Vasconcelos, and S. Yan, "Peak-piloted deep network for facial expression recognition," in *European conference on computer vision*, pp. 425-442. Springer, Cham, 2016.
- [18] P. L. Carrier, A. Courville, I. J. Goodfellow, M. Mirza, and Y. Bengio, "FER-2013 face database," *University of Montreal*, 2013.
- [19] J.R. Lee, L. Wang, and A. Wong, "EmotionNet Nano: An Efficient Deep Convolutional Neural Network Design for Real-Time Facial Expression Recognition," in *Artificial Intelligence*, 3:609673, 2021.
- [20] X. Chang, F. Nie, Z. Ma, and Y. Yang, "Balanced k-Means and Min-Cut Clustering," in *arXiv:1411.6235*, 2014.
- [21] I.R. Tudor, M. Popescu, and C. Grozea, "Local learning to improve bag of visual words model for facial expression recognition," in *Workshop on challenges in representation learning, ICML*, 2013.
- [22] G. Mariana-Iuliana, R.T. Ionescu, and M. Popescu, "Local Learning with Deep and Handcrafted Features for Facial Expression Recognition," in *arXiv preprint arXiv:1804.10892*, 2018.
- [23] G. Panagiotis, I. Perikos, and I. Hatzilygeroudis, "Deep Learning Approaches for Facial Emotion Recognition: A Case Study on FER-2013," in *Advances in Hybridization of Intelligent Methods*. Springer, Cham, pp. 1-16, 2018.
- [24] J. L. Michael, M. Kamachi, and J. Gyoba, "Coding Facial Expressions with Gabor Wavelets," in *arXiv:2009.05938*, 2020.

A new tracking and monitoring smart system based on intelligent algorithms for schoolchildren and mental patients

1st BELKHIRA Sid AHmed Hicham

department of Technology Nour Bachir university center
Elbayadh, Algeria

Belkhira.hichem@gmail.com

2nd ROUISSAT Mehdi

department of Technology Nour Bachir university center
Elbayadh, Algeria

mehdi.m.rouissat@gmail.com

3rd BELKHEIR Mohammed

department of Technology Nour Bachir university center
Elbayadh, Algeria

belkheirm@yahoo.f

4th BOUZIANI Merahi

department of Electronics university of djillali liabes
Sidi Bel Abbe's, Algeria

bouzi.mera@gmail.com

Abstract—In this paper, we present a developed smart system for tracking and monitoring of special category of people, which needs a permanent assistance; schoolchildren, Alzheimer patients and mental deficiency patients. The smart system allows parents to easily monitor and track them, for ensuring their safety; based on original algorithms allowing the learning/definition of safe path/zone. Real-time alerts are sent when the assisted persons get out of the predefined zones/paths. Besides tracking, the system includes others indications, such as: steps counter, speed, body position and falls detection. The proposed system consists of two parts: An electronic embedded system worn by the assisted person and a mobile application installed on the parents Smartphone. Data exchanged between the two parts can be stored in the cloud and managed for further use. The proposed solution has been successfully tested within various environments, and proved its efficiency. The present work is protected by the patent N 10197, delivered by the national Algerian institute of the intellectual property.

Index Terms—Embedded Systems ,tracking GPS, falls detection, Android, remote monitoring, web services.

I. INTRODUCTION

In the last few years, the world witnessed an alarming number of kids losses and kidnapping, which affect this social assisted category. The number of confirmed cases of serious violations against children in has almost tripled from 2010 to 2019 according to the UNICEF report[1]. In Algeria for instance, the police inspection reveals 887 disappearance cases of children in 2017. Also, Between January and August 2015, around 39 girls out of a total of 52 children were kidnapped and raped by their abductors; some cases were followed by murder [2]. On the other hand, the burden of mental disorders like: depression, trouble disorders, dementia... continues to grow with significant impacts on health and society (hundreds millions of patients in the world) [3]. Where, every 3 seconds, a person in the world develops a dementia. Regarding Algeria

for example, nearly 200,000 Alzheimer's cases have been identified nationwide between 2017 and 2018 [4]. The above mentioned statistics encouraged and pushed researchers to find solutions to make the lives of those assisted people easier, like developing smart wearable systems involving recent technologies (microchips, IoT, RFID,...etc.), in order to assist these social categories. The smart wearable systems allow assisted persons to safely circulate and normally achieve their daily tasks, even when they are far away from their parents. Also, via these solutions, parents can be relieved and remotely track their assisted persons having all details about their movements and health states. With the evolution of smart embedded-systems, ICTs (information and communications technology), sensor technologies, and the rapid growth of mobile telecommunications systems (GPS, 5G...etc), the remote monitoring and the help of the assisted persons like elderly, mental deficiency patients and schoolchildren becomes a challenge facing researchers and developers. In the research domain, various solutions have been proposed for monitoring and assisting elderly and schoolchildren. In [5], authors proposed a remote health monitoring solution to efficiently transmit the patient ECG (Electrocardiogram) using a home gateway and a web server for data collect. Collected data, can be used for further management of the patient healthcare. Authors In [6], presented another wearable solution to record the patient breathing rate and the frequency of coughing for COPD (Chronic obstructive pulmonary disease patients). Researchers have proposed other global solutions for schoolchildren transportation. In [7], authors proposed a design of a GPS tracking solution for schoolchildren buses, based on RFID tags to identify schoolchildren where they are in the school bus. Other kids GPS trackers are proposed in the industry like smart watches, Smartphone Apps and so on.

However, these proposed devices are not smart enough, and are dedicated for a common use and they are not fully personalized for special cases. Notifications and alerts sent by the wearable systems in the case for schoolchildren are much different from those sent in the case of elderly or a mental patient device. Moreover, schoolchildren are facing kidnapping and violence acts when they frequent their schools. Alzheimer and mental patients can easily get lost when they are far away from their familiar environment, or when they take means of transport. In addition, data collected by the wearable system should be efficient and reliable, in order to avoid false alerts, and efficiently report the assisted person real state (true falls, smart path learning, path deviation detection). To do so, a smart monitoring and tracking system should implement a smart algorithms for each use case. A recent study results [8], depicts that the implementation of a personalized smart monitoring system with the use of statistical data analysis and data mining systems allow providing efficient healthcare systems, to prevent some catastrophic cases (losses, deaths...etc.), which make the realization of an adequate solution a very big challenge. The rest of the paper is organized as follows: we first describe the modeling and materials used to implement our system. After that, we describe the main functionalities and features, illustrated by implemented algorithms. Finally, we discuss the functionalities and we conclude the paper.

II. METHODS AND MATERIALS

The main objective of the proposed smart system is to efficiently track and monitor the assisted persons. The system consists of two main parts: • A Smart wearable device attached to a belt carried by the assisted person. We opted for a belt to dissimulate the embedded device and avoid disturbing the assisted person. The device contains a set of sensors in order to collect a set parameters for deciding about the assisted person state. The device allows various options, depending on the monitored person, like: falls detection, steps counting, alerts notifications and response to the parent's requests. In addition, it implements original developed algorithms for path tracking. • An Android Application installed in the parent Smartphone, multilingual and easy-to-use, intended to receive alerts from the smart wearable device and send request for requesting the assisted person state.

A. Functional description

The smart aspect of the proposed solution resides in the personal customization and the efficiency of alert notifications, in order to avoid false alerts. This is achieved by implementing algorithms which allow a maximum of precision, when diverting the initial defined path. Fig.1. illustrates the functional scheme of the proposed solution. It also depicts all the features of the system and options for tracking and monitoring. The main supervised parameters are:

- Localization (GPS coordinates)
- Predefined path/zone deviation
- Over speed detection, when taking a transportation mean (vehicle, bus,... etc.) kidnapping case

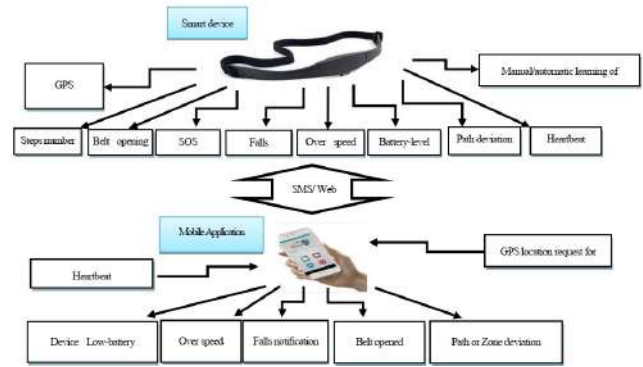


Fig. 1. General description of the smart system

- Falls detection of patient or elderly,
- Body position, (standing, sitting...etc),
- Pedometer, steps counting, to estimate the activity level of the elderly
- Heartbeat measurement
- Belt opening,
- Smart device low-battery level.
- Manual notification in case of danger, (SOS button),
- Remote voice monitoring

The modeling process is very important in each complex project, especially when combining both the hardware and the software in the same solution. With UML (Universal Modeling Language), it's easy to present a visual description of the system and the interactions between all its actors. In addition, UML allows us to efficiently follow the modeling process, starting from the requirement analysis phase to the system final implementation phase [9]. ARGO UML [10] is an open-source software, UML based, we used it to build all system scenarios and description diagrams. In this paper, we simply present the two main UML diagrams to describe the system modeling process for the schoolchildren case. Fig.2.A. illustrates the use case diagram which details all operations made by actors (Parent and assisted person) within the system. After this diagram, we can easily illustrate all system classes' attributes and operations which should be developed, software or hardware, explained by fig.2.B. The mental deficiency and Alzheimer patients cases follow the same modeling process, by including other parameters (Steps number counting, heartbeat measurement...etc.) and falls detection.

B. System Modeling

C. Wearable device Architecture

"Fig. 3" illustrates the block diagram of the smart device worn by the assisted person. It contains the following components: • An Atmega 2560 microcontroller [11], for connecting all sensors and communication modules (GPS/GSM). It represents the system brain. • The SD card is a simple 128 Mo card for path coordinates storage and other useful information for the system. • A NEO-6M GPS module [12], for the acquisition of the current GPS positions of the assisted person.

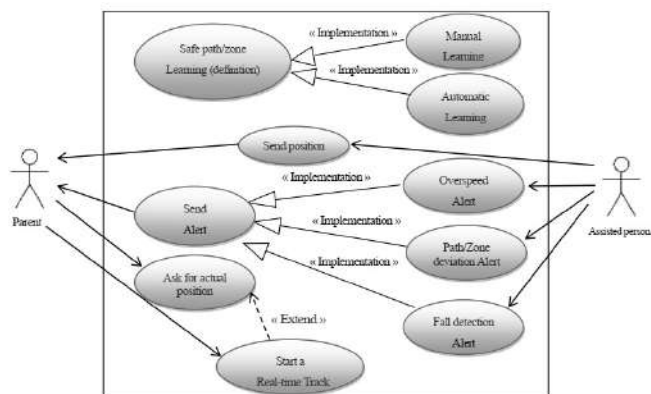


Fig. 2. Schoolchildren UML Use case Diagram

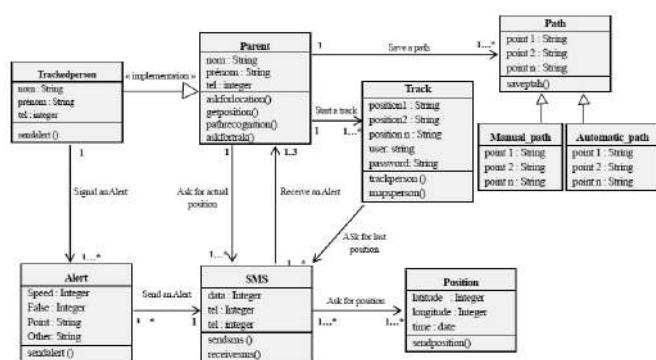


Fig. 3. Schoolchildren UML Class Diagram

This is essential for tracking and detecting safe Path/zone deviation.

- A SIM 800I GSM/GPRS module [13], for sending alerts to the android application and receiving requests from the assisted person parent. It is also used for sending GPS coordinates via the cloud for the real time tracking.
- A 6-Axis Gyro/accelerometer MPU-6050 sensor [14], for falls detection and steps counting.
- A BMP 280 pressure sensor [15], for the pressure measurement. The information gathered from this sensor will be used along with that of the MPU-6050 for efficiently decide about the falls.
- A SOS button, allowing the assisted person to immediately alert the parent when facing a danger.
- A Manual path/zone learning button. This button allows parents to manually register a regular safe path (a set of GPS coordinates). The stored path will be used as reference for determining the path/zone deviation.
- A reset button for resetting the device.
- The device is equipped with a magnetic field sensor for alerting parents when the device is detached from the belt.
- The device is battery-powered. So, it is equipped with a voltage indicator which measures how much energy has left and sends an alert to the parent when the battery-level becomes low.

D. "Smart Tracker" Android Mobile Application

The developed android application is part of the smart system. It is installed in the parent Smartphone. It receives

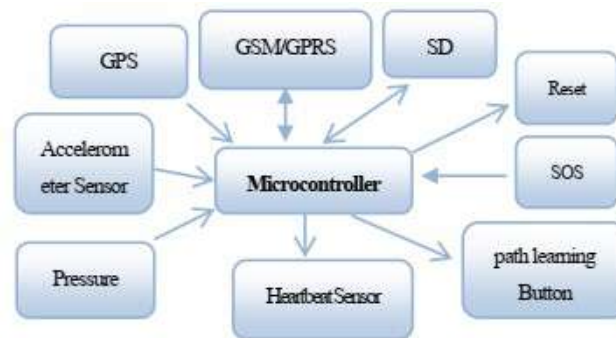


Fig. 4. Bloc Diagram of the Smart device

alerts sent by the smart device and allows the parent to easily monitor and track the assisted person. The application

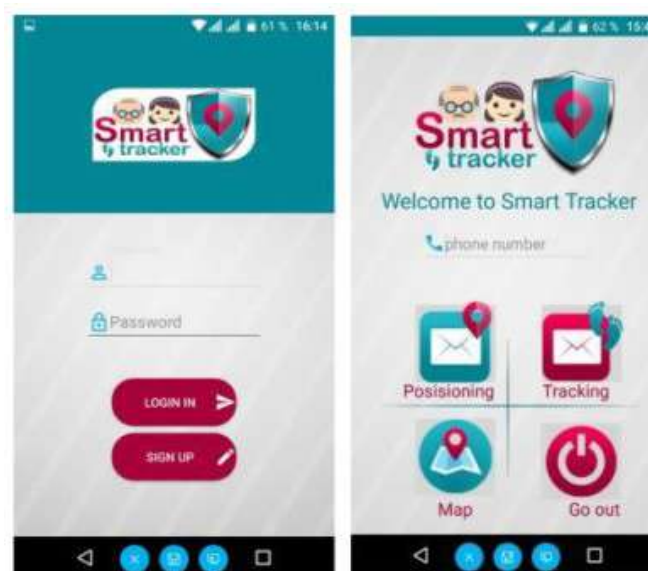


Fig. 5. Smart Tracker interface and options

is named "Smart Tracker" and was designed and developed using Android Studio platform [16]. "Smart tracker" works in a standalone mode and it can connect to the cloud for storing all notifications, alerts and positions received from the smart device. This option seems very important for performing statistics about the assisted person status and activity. This also allows parents or medical staff to follow the health status evolution of the mental deficiency or Alzheimer patients (falls number, activity level, frequented zones or paths...). Fig.4. illustrates the interface and options offered by "Smart tracker". The offered options by the application are:

- Account creation to sign-up for authorized persons.
- Multilingual interface (Arabic, English, French)
- Request for the current position of the assisted person tem Real-time tracking through a map, the map tracking is performed in an offline mode, with no need of internet connection.

- Customized display when receiving SMS notifications and alerts
- Alert triggering and customized ringing for SOS case.
- Acknowledgment of alerts and notifications. This allows the smart device, in case of emergency, to send alerts to another phone number stored in its phonebook when no acknowledgment is received from the parent Smartphone.

In addition, the web server contains the PHP files and MySQL database, developed in order to process data received from the device. Data retrieved are then displayed and managed by the endusers (i.e parents, medical staff...etc.).

III. RESULTS AND DISCUSSIONS

In this section, we explain our proposed algorithms, implemented for the safe path definition and deviation detection, in the case of schoolchildren. A path is defined by a set of GPS coordinates which represents the main points like: cross-roads, turns, pedestrian crossings...etc. Each point defines a predefined virtual zone. The path is then formed by a set of juxtaposed circles (zones). Once the path is defined, the device calculates and stores distances and estimated times between each two successive points. The device is equipped with a 3 selector buttons to select the path learning method:

- Manual path learning: in this method, the parent manually selects the GPS coordinates belonging to the safe path.
- Automatic path learning (based on the distance): the device stores a GPS coordinate each 30m.
- Automatic path learning (based on time): the device stores a GPS coordinate each 30 seconds.

After setting a safe path, the smart device, by using the algorithm illustrated in Fig.5, can easily monitor the schoolchild moves from home until he reaches his school. Two main parameters are taking into account when the schoolchild moves from one point to another which are time and distance. A path can be determined by the formula (1)

$$Path = \bigcup_{i=2}^n P_i \quad (1)$$

Where P_i represents a set of important point defining a path (crossroad, pedestrian...) ig.6. illustrates a practical case to clearly explain the path deviation detection algorithm processing. At first, we define all points defining the path:

- Point 1: Defines the schoolchild home
- Point2: Represents the 1st crossroad
- Point 3: Represents the 2nd crossroad
- Point 4: Represents the 3rd crossroad
- Point 5: Is the schoolchild school

From this, we can easily calculate distances between each two successive points, using haversine formula [17][18]. If we have two successive points P_i (Lat i , Long i), P_{i+1} (Lat $i+1$, Long $i+1$), the distance D_i , is then calculated by the following formula:

$$D_i = 2 * R * a \quad (2)$$

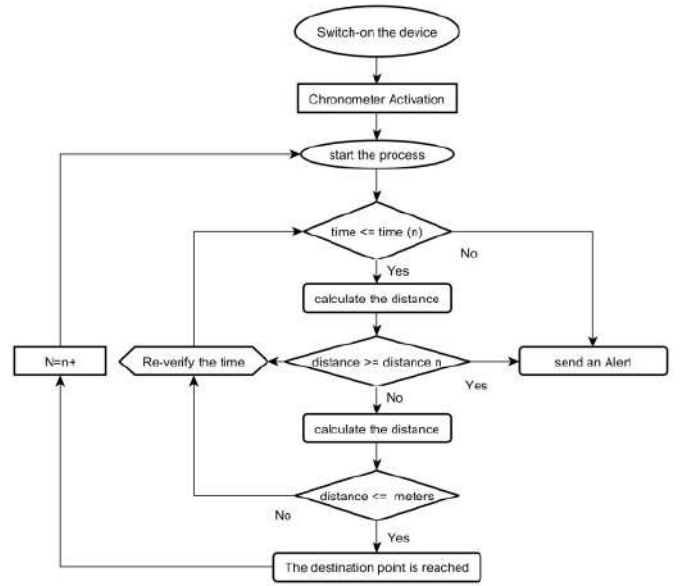


Fig. 6. path deviation detection Algorithm

Where :

$$a = \arcsin \sqrt{\sin^2 b} \quad (3)$$

$$b = \left\{ \frac{Lat_i Lat_{i+1}}{2} + \cos(Lat_i) * \cos(Lat_{i+1}) * \sin^2\left(\frac{Long_i * Long_{i+1}}{2}\right) \right\}$$

- R: is the radius of earth (R= 6371 km)
- D1: Distance between the house and the point2
- D2: Distance between the point 2 and point 3
- D3 :Distance between point 3 and point 4
- D4: Distance between point 4 and the school

Where the distance can be determined by Knowing that the mean pedestrian speed is about 4km/h, we can also calculate the time spent by the schoolchild between each predefined points:

- T1: Time spent to walk D1
- T2: Time spent to walk D2
- T3: Time spent to walk D3
- T4: Time spent to walk D4

When the schoolchild comes out of the house, the smart device starts the calculation of the distance between the device and P2 (point 2). This distance is named Dactual and the elapsed time is Tactual

- if $D_{current} > D_1$ that means the schoolchild has taken the opposite direction or another path then the predefined one. An alert is sent to the parent,
- if $T_{current} > T_1 + 10s$, an alert is sent to the parent,
- if $D_{current} \leq 10m$ we consider that P2 is reached. If P_2 is reached, then the device calculates the distance between itself and P_3 .
- if $D_{actual} > D_2$, this means that the schoolchild has taken

the opposite direction or another path then the predefined one. An alert is sent to the parent,

- if $T_{actual} > T_2 + 10s$, an alert is sent to the parent,
- if $D_{actual} \leq 10m$, we consider that P3 is reached.

This routine is repeated based on the number of points defining the safe path, until the schoolchild reaches the final destination. In addition, the parent can at any time track his child, by requesting the device current location. For this option, the android application communicates primarily with the smart device via the cloud for receiving data about the GPS coordinates of the assisted person in movement. The received data is then converted and displayed on a map which indicates graphically, the real-time movement of the assisted person. As discussed earlier, the alert notifications are fully customized, depending on their potential severity. For each alert, an SMS notification is received, defining the nature of the alert and describing its details.

IV. CONCLUSION

In this paper, we presented the design and the development of a smart solution for assisting the parents of schoolchildren, mental deficiency and Alzheimer. The solution is a wearable device worn by the assisted person and an android application installed in the parent Smartphone.

The solution implements the newer embedded-systems and ICT technologies. It is fully customized for special cases and implements original tested efficient algorithms for path learning and alert notifications. The system has been tested and proved its efficiency.

Data gathered by the system and stored in the cloud, could be used by parents and medical staff monitors for further statistics management and monitoring of the assisted person status and health.

REFERENCES

- [1] United Nations Children's Fund, United Nations, Annual Report on UNCIEF Humanitarian Action, Annual session 2019, E/ICEF/2019/12. April 15th 2019.
- [2] N. Ait Ahcène. Kidnapping des enfants en Algérie, Available from : <https://www.podcastjournal.net/Kidnapping-des-enfants-en-Algerie-un-fleau-inquietant-a23301.html> Published on June 26th 2016. last accessed September 2019.
- [3] World Health Organization, WHO Depression and Other Common Mental disorders, Available from: www.who.int/mental-health/management/depression-prevalence-global-health-estimates/en Published in 2017, last accessed October 2019.
- [4] S. Amalou. Maladie D'Alzheimer, Available from: <http://www.aps.dz/sante-science-technologie/75911-maladie-d-alzheimer-pres-de-200-000-cas-recenses-a-l-echelle-nationale> Published on July 3rd 2018. last accessed September 2019
- [5] K. Guan, M. Shao, S. Wu, A Remote Health Monitoring System for the Elderly Based on Smart Home Gateway, Journal of Healthcare Engineering Volume 2017, Article ID 5843504, <https://doi.org/10.1155/2017/5843504>. October 24th 2017
- [6] T. Elfaramawy, C. L. Fall, M. Morissette, F. Lehouche, B. Gosselin, Wireless respiratory monitoring and coughing detection using a wearable patch sensor network, 2017 15th IEEE International New Circuits and Systems Conference (NEWCAS). IEEE. August 15th 2017
- [7] K.S. Kalid, N. Rosli, The design of a schoolchildren identification and transportation tracking system, 2017 International Conference on Research and Innovation in Information Systems (ICRIIS), 10.1109/ICRIIS.2017.8002454. IEEE. August 10th 2017.
- [8] L. Yu, W.M. Chan, Y. Zhao, K. Tsui, Personalized Health Monitoring System of Elderly Wellness at the Community Level in Hong Kong, IEEE Access, DOI: 10.1109/ACCESS.2018.2848936. July 19th 2018.
- [9] Zheng, Y. Feng, Y. Zhao, A Unified Modeling Language-Based Design and Application for a Library Management Information System, Cybernetics and information technologies – Volume 14, DOI: 10.2478/cait-2014-0050, Sophia 2014.
- [10] M. Vand der Vulp, ARGO UML user manual: A tutorial description, <https://people.cs.pitt.edu/~chang/1635/ArgoUMLman.pdf>, 2009.
- [11] Atmel, 8-bit Atmel Microcontroller with 16/32/64KB In-System Programmable Flash, atmega 640–1280–1281–2560–2561 datasheet.pdf. 2549Q–AVR–02/2014.
- [12] U-blox, NEO–6u–blox GPS modules datasheet, NEO–6 DataSheet (GPS.G6-HW-09005).pdf. April 2019.
- [13] SIMcom, SIM800 Series–Serial Port Application Note V1.01.pdf, 05–08–2014.
- [14] InvenSense, MPU–6000 and MPU–6050 Register Map and Descriptions Revision 4.2 Available from: <https://www.invensense.com/products/motion-tracking/6-axis/mpu-6050/>, 08/19/2013 last accessed December 2019
- [15] Adafruit BMP280 Barometric Pressure +Temperature Sensor Breakout Created by lady ada <https://learn.adafruit.com/adafruit-bmp280-barometric-pressure-plus-temperature-sensor-breakout>, 2021-11-15
- [16] Developer guides: <https://developer.android.com/guide/essential-documentation> last accessed August 2019.
- [17] AGY. Swara, Implementation of haversine formula and best first search method in searching of tsunami evacuation route, International Conference on Environment and Technology. doi : 10.1088/1755–1315/97/1/012004. (IC – Tech) 2017.
- [18] D. Iksari, Widiastuti and R. Andika, "Determine the Shortest Path Problem Using Haversine Algorithm, A Case Study of SMA Zoning in Depok," 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2021, pp. 1-6, doi: 10.1109/HORA52670.2021.9461185.

A Novel Artificial Intelligence-Assisted Meta-heuristic Method for Augmenting the Annual Profit of Wind Farm

Prasun Bhattacharjee

Department of Mechanical Engineering
Jadavpur University
Kolkata, India
prasunbhatta@gmail.com

Rabin K. Jana

Operations & Quantitative Methods
Indian Institute of Management
Raipur, India
rkjana1@gmail.com

Somenath Bhattacharya

Department of Mechanical Engineering
Jadavpur University
Kolkata, India
snb_ju@yahoo.com

Abstract—While climate change is inciting calamitous impacts across the world, wind energy brings forward a germane alternative to well-known hydrocarbon-based fuels for lessening the global greenhouse gas releases. Economic success is a decisive facet of carbon offsetting of electricity generation trades as suggested by the Paris pact of 2015. The current research work insinuates to improve the annual profit of wind farms in the Kayathar area of India. A pioneering modification of the Genetic Algorithm has been proposed. An innovative dynamic scheme for assigning the crossover and mutation possibilities has been put into operation to augment the regarded goal. A reliable air-stream movement configuration has been appointed for calculating the annual profit. The research conclusions corroborate the beneficial competence of the projected state-of-the-art tactic for expanding the economic possibility of wind farms with three arbitrarily selected landscape conditions.

Index Terms—Artificial Intelligence, Genetic Algorithm, Meta-heuristic, Profit Enhancement, Wind Power.

I. INTRODUCTION

Consumption of power is viewed as one of the numerous decisive aspects of the financial development of every contemporary state. While the worldwide electricity generation developed from 8794 Mtoe in 1990 to 14410 Mtoe in 2019 with strengthening of commercial events, the traditional energy generation resources are expending at an exceptional rapidity [1].

The persistent discharge of greenhouse gases because of industrial actions is mounting air temperature and setting off atypical climate change universally [2]. Strikingly, the appliance of renewable power rocketed by 3% in 2020, even though the necessity of hydrocarbon-based fuels weakened for pandemic-linked travel controls [3].

In addition to marginal emission gain, non-conventional power resources like wind power are necessitated to remain justifiable through proposing reasonable generation expenditure [4]. The expense of generating power from the air has contracted radically over the preceding decades transnationally [5].

Researchers are determinedly making a great effort to boost the cost-efficiency of wind farms to help out nations

in achieving their emission curbing aspirations as swiftly as attainable [6].

Artificial Intelligence (AI)-assisted meta-heuristic approaches like Genetic Algorithm (GA) have been employed for a wind farm at the Gökçeada island [7]. GA and Neural Network (NN) have been applied simultaneously in another study for wind power analysis [8]. An NN-empowered technique in the company of Swarm Optimization Algorithm (SOA) and GA has been suggested for the wind power prognostication [9].

SOA and GA have been utilized concurrently for wind power expenditure minimization [10]. Comparative scrutiny of GA and SOA has been presented to curtail wind power expenditures [11].

Even though the most research works were directed toward lessening the wind power expenditure, further attention is to be engaged in the direction of broadening the possibility of financial permanence for wind energy organizations for achieving the 2015 Paris accord assurances made by numerous governments and intercontinental alliances.

This research proposes to accomplish superior annual profit at an Indian wind energy generation location with two arbitrarily chosen territory situations.

On account of the complications of the wind power generation process, outdated optimization tactics are inept to function under such settings.

AI-stimulated meta-heuristic methods have been previously applied in various engineering fields for their adaptability and prompt computation potency [12]–[24].

GA is a distinguished AI-supported meta-heuristic procedure representing the advancement of biotic predilection and obeying the plan of notable discoverer Alan Turing [25]–[27].

GA has been engaged in the present study to maximize the annual profit of a wind farm.

II. PROBLEM STATEMENT

A. Objective Formulation

The power generated by a wind turbine can be estimated as follows.

$$P_T = \frac{1}{2} \rho_{air} A v^3 C_{betz} \cos \beta \quad (1)$$

where P_T signifies the generated wind energy, ρ_{air} stands for the density of the air-stream, A is the swept area, v denotes the wind speed, C_{betz} symbolizes the Betz factor and β shows the angular error of yaw mechanism [11], [28].

The current research is targeted at strengthening the annual profit of a wind farm in the Kayathar area of India. The goal function can be calculated as follows.

$$f_{annual} = [S_P - G_P] \times P_{yearly} \quad (2)$$

where f_{annual} signifies the yearly profit, S_P is the marketing value per unit of wind power, G_P symbolizes the generation cost per unit of wind energy and P_{yearly} is the wind power generated yearly.

The generation expenditure of wind power has been calculated as per the cost function recommended by Wilson *et al.* [29]. In this research work, the wind energy generation capacity of Kayathar, a town in the Indian state of Tamil Nadu, has been deemed.

The wind-flow form of Jafrabad city has been offered in Fig.1.

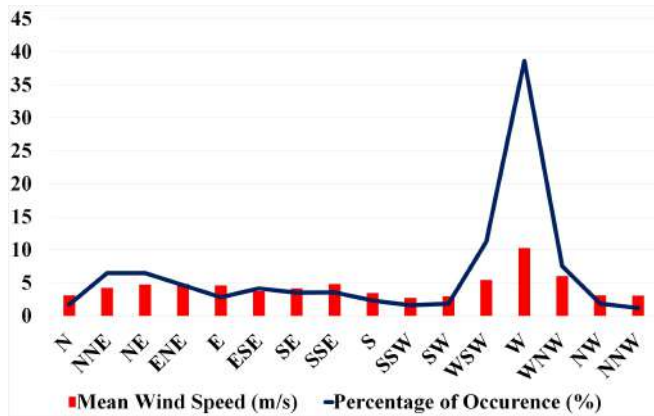


Fig. 1. Air Flow Pattern of Kayathar, Tamil Nadu

B. Terrain Condition

Three randomly selected layout settings have been elected for estimating the annual profit of Wind Power Generation (WPG) System. One of the layouts is with no obstacle and other two layouts have obstacles within their boundaries.

The presence of obstacles has been considered to evaluate its influence on the feasibility of the WPG farm and expand the practicability of the proposed scientific model.

Though the layouts chosen for the current research are square, they can be easily transformed into any rectangular shape consistent with the requirement of the design engineers.

The landscape conditions have been explicitly presented in Figs. 2-4.



Fig. 2. Layout 1 without Obstacle

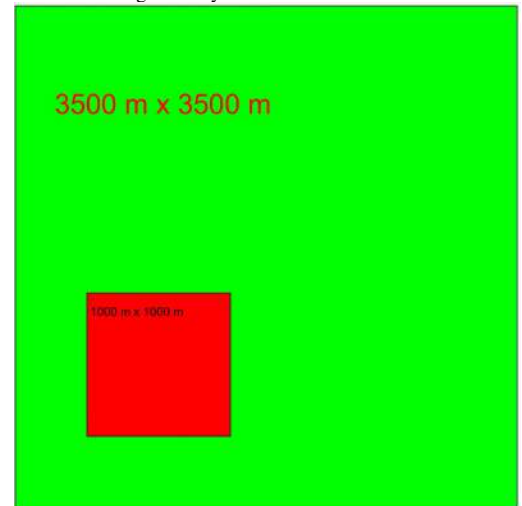


Fig. 3. Layout 2 with One Obstacle of 1000 m x 1000 m

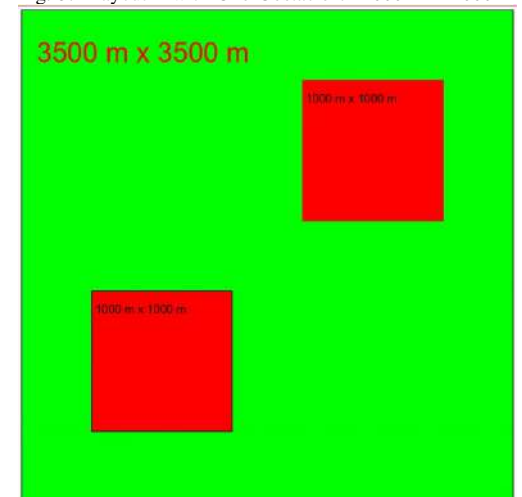


Fig. 4. Layout 2 with Two Obstacles of 1000 m x 1000 m

III. OPTIMIZATION ALGORITHM

GA has been employed in the current research work to assess the optimal annual profit of the WPG farm at Kayathar for three different topography conditions.

The algorithm has been concisely offered as follows [12].

- 1) Establish the basic parameters of GA.
- 2) Prepare the populace randomly.
- 3) Calculate the aptness of all individual entities.
- 4) Accomplish the arithmetic crossover tactic as follows.
 - a) Pick out a fractional value randomly between 0 and 1. If it is less than the factor of the crossover tactic, select it as the parental chromosome.
 - b) Set off the crossover activity.
 - c) Review the appropriateness of the pro-genies.
 - d) If the successor is practicable, combine it into the newest populace.
- 5) Complete the mutation tactic as follows.
 - a) Pick out a fractional value randomly between 0 and 1. If it is less than the factor of the crossover tactic, select it for mutation.
 - b) Set off the mutation activity.
 - c) Evaluate the aptness of the mutated entity.
 - d) If the mutated entity is viable, integrate it into the newest populace.
- 6) Compute the fitness of the fresh entities formed by crossover and mutation processes.
- 7) Elect the finest solution following the choice of the decision-maker.

Accompanied by the well-known method of assigning constant values, this research work has engaged a novel dynamic process for allocating the factors of crossover and mutation.

The dynamic crossover factor has been computed as follows.

$$c_{dynamic} = c_u + \left\{ (c_u - c_l) \left(\frac{R_x}{R_{highest}} \right)^{(1/11)} \right\} \quad (3)$$

where $c_{dynamic}$ denotes the dynamic crossover factor. c_u and c_l signify the limits of the crossover factor. R_x stands for the current recurrence count and $R_{highest}$ is the maximum repetition limit.

The dynamic mutation factor has been calculated as follows.

$$m_{dynamic} = m_u + \left\{ (m_u - m_l) \left(\frac{R_x}{R_{highest}} \right)^{(1/11)} \right\} \quad (4)$$

where $m_{dynamic}$ denotes the growing mutation factor. m_u and m_l stand for the bounds of the mutation factor.

IV. RESULTS AND DISCUSSION

Meta-heuristic approaches have been involved considerably in wind farm layout optimization [18]–[26].

The purpose of the current research is to boost the annual profit of the wind farm at Kayathar with a novel bio-enthused meta-heuristic technique.

The yearly profit of the wind farm at Kayathar has been calculated in US\$. The selling value of wind energy has been considered as US\$ 0.033/kWh [30].

In collaboration with the innovative dynamic method, the standard static assignment tactic for crossover and mutation possibilities has been deemed to calculate comparative competence.

The values of numerous limits allied to the considered optimization process have been presented in Table 1.

TABLE I
VALUES OF SEVERAL FACTORS ASSOCIATED WITH THE PROPOSED META-HEURISTIC PROCESS

Factor	Value
c_u	0.4
c_l	0.3
m_u	0.04
m_l	0.03
Populace Size	20
Highest Generation Count	50
Constant Crossover Factor	0.4
Constant Mutation Factor	0.04
Rated Power of Wind Turbine	1500 kW
Blade Radius	38.5 m
Inter-Turbine Gap	308 m
Minimum Operational Air Flow Speed	12 km/hr
Maximum Operational Air Flow Speed	72 km/hr
Purchasing Price per Turbine	USD 750,000
Expenditure per Sub-Station	USD 8,000,000
Yearly Operational Charge	USD 20,000
Interest	3%
Anticipated Operational Life of Wind Turbine	20 years
Count of Turbines per Sub-Station	30

The optimal locations of Wind Turbines (WTs) for all layouts using both techniques for assigning the factors of crossover and mutation processes of GA have been presented in Figs. 5-10.

The optimization algorithms have been planned to avoid locating any WT within the bounds of the obstacle.

The positions of WTs have been marked with red circular dots.

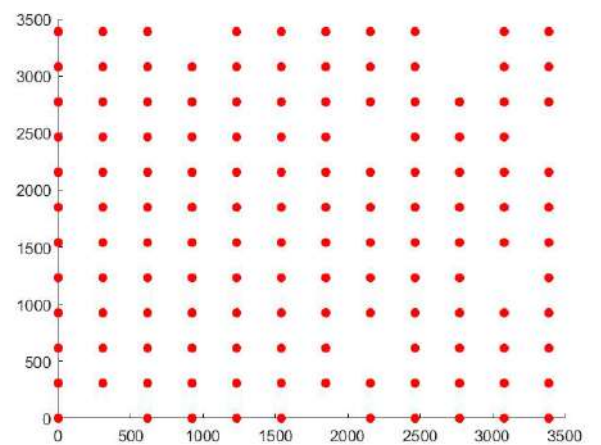


Fig. 5. Optimal Placement of WTs for Layout 1 by the Novel Dynamic Approach for Allocating the Factors of Crossover and Mutation Processes of GA

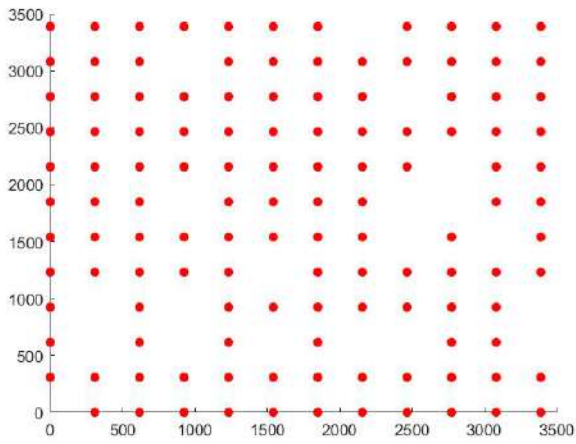


Fig. 6. Optimal Placement of WT's for Layout 1 by the Standard Static Approach for Allocating the Factors of Crossover and Mutation Processes of GA

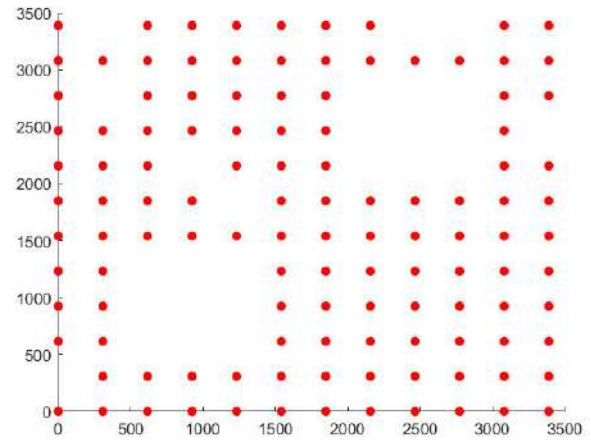


Fig. 9. Optimal Placement of WT's for Layout 3 Engaging the Novel Dynamic Approach for Allocating the Factors of Crossover and Mutation Processes of GA

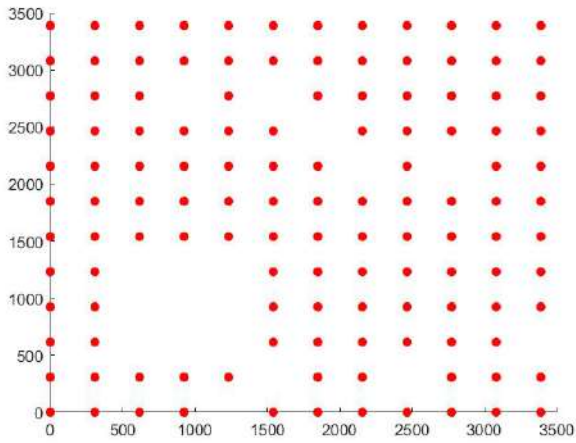


Fig. 7. Optimal Placement of WT's for Layout 2 by the Novel Dynamic Approach for Allocating the Factors of Crossover and Mutation Processes of GA

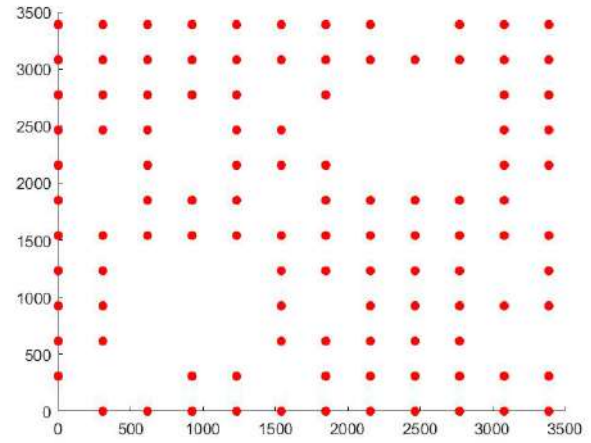


Fig. 10. Optimal Placement of WT's for Layout 3 Engaging the Standard Static Approach for Allocating the Factors of Crossover and Mutation Processes of GA

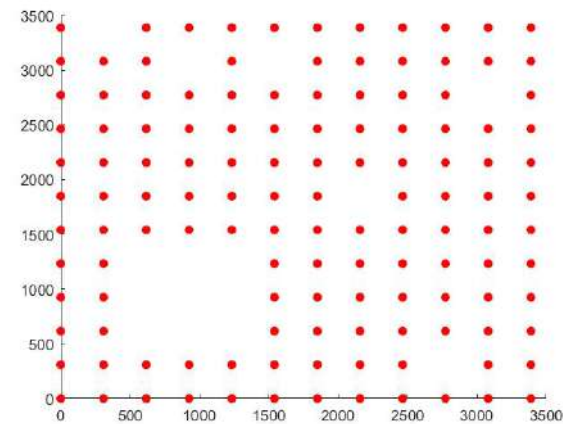


Fig. 8. Optimal Placement of WT's for Layout 2 by the Standard Static Approach for Allocating the Factors of Crossover and Mutation Processes of GA

A relative analysis of the optimal yearly profits achieved by both methods of allocating the factors of crossover and mutation techniques of GA for all three layout designs have been offered in Table 2.

The annual profit for each layout has been computed by Eq. (2) and the generation cost of wind energy has been calculated as per Wilson *et al.* [29].

TABLE II
ANALYSIS OF OPTIMAL YEARLY PROFITS FOR THREE CONSIDERED LAYOUT SETTINGS ACHIEVED BY BOTH APPROACHES OF ALLOCATING THE CROSSOVER AND MUTATION FACTORS OF GA

Optimization Tactic	Layout 1	Layout 2	Layout 3
Standard Static	US\$ 26,905	US\$ 26,101	US\$ 24,186
Novel Dynamic	US\$ 27,236	US\$ 26,222	US\$ 24,659

A comparative analysis of the optimal number of WTs achieved by both methods of allocating the factors of crossover and mutation techniques of GA for all three layout designs have been offered in Table 3 .

TABLE III

ANALYSIS OF OPTIMAL NUMBER OF TURBINES FOR THREE CONSIDERED LAYOUT SETTINGS ACHIEVED BY BOTH APPROACHES OF ALLOCATING THE CROSSOVER AND MUTATION FACTORS OF GA

Optimization Process	Layout 1	Layout 2	Layout 3
Standard Static	124	129	109
Novel Dynamic	135	126	118

The research solutions confirm the superiority of the proposed innovative dynamic method of allocating crossover and mutation probabilities over the recognized static approach for both layouts as it realized the greater yearly profit.

As the initial investment of WPG farm is quite high, the proposed optimization technique can help it gain better financial viability.

The improved fiscal stability of the WPG system can support the curtailment of greenhouse gas release for the power generation industries [31].

V. CONCLUSION

Climate change is distressing human societies universally. Soaring sea levels, deviant meteorological settings, and the persistent upsurge in surface air temperature are causing immense loss of exquisite natural resources and human lives everywhere in the world.

International organizations like the United Nation are relentlessly making a great effort in the course of cutting down the greenhouse gas production by the efficient utilization of renewable power generation methods like wind power as envisioned by the Paris accord of 2015 and the COP-26 held in Glasgow.

This study purposes for expanding the yearly profit of the wind farm in the Kayathar region of Tamil Nadu, India with an innovative dynamic approach for allotting the crossover and mutation factors of GA.

The optimization trial conclusions ratify the enhanced suitability of the pioneering dynamic technique over the typical static method for enhancing the wind farm design with the highest yearly profit.

The proposed procedure can assist wind farm businesses to design a realistically practical WPG system with an actual consideration of numerous cost-allied features and flexible air-stream conditions.

The present research work can trigger the inventive possibility for wind farm layout enhancement and permit the de-carbonization of the international electricity generation businesses.

In future studies, different WPG sites and WTs can be considered for the enhanced perception of WPG trades.

ACKNOWLEDGMENT

The first author admits the fiscal support of the TEQIP unit of Jadavpur University for the present research effort.

REFERENCES

- [1] Enerdata, "Global Energy Statistical Yearbook," 2020. [Online]. Available: <https://yearbook.enerdata.net/>. [Accessed 05 September 2020].
- [2] B. Obama, "The irreversible momentum of clean energy," *Science*, vol. 355, no. 6321, pp. 126-129, 2017.
- [3] International Energy Agency, "The impact of the Covid-19 crisis on clean energy progress," 11 June 2020. [Online]. Available: <https://www.iea.org/articles/the-impact-of-the-covid-19-crisis-on-clean-energy-progress>. [Accessed 30 July 2021].
- [4] G. Nicholas, T. Howard, H. Long, J. Wheals and R. Dwyer-Joyce, "Measurement of roller load, load variation, and lubrication in a wind turbine gearbox high speed shaft bearing in the field," *Tribology International*, vol. 148, p. 106322, 2020.
- [5] R. Sitharthan, J. N. Swaminathan and T. Parthasarathy, "Exploration of Wind Energy in India: A Short Review," in 2018 National Power Engineering Conference (NPEC), 2018.
- [6] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "A Relative Analysis of Genetic Algorithm and Binary Particle Swarm Optimization for Finding the Optimal Cost of Wind Power Generation in Tirumala Area of India," *ITM Web of Conferences*, p. 03016, 2021.
- [7] S. Şişbot, Ö. Turgut, M. Tunç and Ü. Çamdalı, "Optimal positioning of wind turbines on Gökçeada using multi-objective genetic algorithm," *Wind Energy*, vol. 13, no. 4, pp. 297-306, 2010.
- [8] S. Saroha and S. K. Aggarwal, "Multi step ahead forecasting of wind power by genetic algorithm based neural networks," in 2014 6th IEEE Power India International Conference (PIICON), 2014.
- [9] D. T. Viet, V. V. Phuong, M. Q. Duong and Q. T. Tran, "Models for Short-Term Wind Power Forecasting Based on Improved Artificial Neural Network Using Particle Swarm Optimization and Genetic Algorithms," *Energies*, vol. 13, no. 11, p. 2873, 2020.
- [10] C. Roy and D. K. Das, "A hybrid genetic algorithm (GA)-particle swarm optimization (PSO) algorithm for demand side management in smart grid considering wind power for cost optimization," *Sādhanā*, vol. 46, no. 2, 2021.
- [11] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "A Comparative Analysis of Genetic Algorithm and Moth Flame Optimization Algorithm for Multi-Criteria Design Optimization of Wind Turbine Generator Bearing," *ADB Journal of Engineering Technology*, vol. 10, 2021.
- [12] R. K. Jana and P. Bhattacharjee, "A multi-objective genetic algorithm for design optimisation of simple and double harmonic motion cams," *International Journal of Design Engineering*, vol. 7, no. 2, pp. 77-91, 2017.
- [13] A. Duggirala, R. K. Jana, R. V. Shesu and P. Bhattacharjee, "Design optimization of deep groove ball bearings using crowding distance particle swarm optimization," *Sādhanā*, vol. 43, no. 1, 2018.
- [14] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "Multi-Objective Design Optimization of Wind Turbine Blade Bearing," *Invertis Journal of Science & Technology*, vol. 14, no. 3, pp. 114-121, 2021.
- [15] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "A Comparative Study of Dynamic Approaches for Allocating Crossover and Mutation Ratios for Genetic Algorithm-based Optimization of Wind Power Generation Cost in Jafrabad Region in India," in *Proceedings of International Conference on "Recent Advancements in Science, Engineering & Technology, and Management"*, Jaipur, India, 2021.
- [16] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "Design Optimization of Cam in Computer-Aided Simulation Applications using Taguchi's Experimentation Method," *International Journal of Electrical and Computer System Design*, 2021.
- [17] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "A Comparative Analysis of Genetic Algorithm and Moth Flame Optimization Algorithm for Multi-Criteria Design Optimization of Wind Turbine Generator Bearing," *ADB Journal of Engineering Technology*, vol. 10, 2021.
- [18] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "A Relative Analysis of Genetic Algorithm and Binary Particle Swarm Optimization for Finding the Optimal Cost of Wind Power Generation in Tirumala Area of India," *ITM Web of Conferences*, p. 03016, 2021.

- [19] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "A Relative Assessment of Genetic Algorithm and Binary Particle Swarm Optimization Algorithm for Maximizing the Annual Profit of an Indian Offshore Wind Farm," in Second International Conference on Applied Engineering and Natural Sciences, Konya, Turkey, 2022.
- [20] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "An Enhanced Genetic Algorithm for Annual Profit Maximization of Wind Farm," *Journal of Information Systems & Operations Management*, vol. 15, no. 2, 2021.
- [21] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "An Enriched Genetic Algorithm for Expanding the Yearly Profit of Wind Farm," in Second International Symposium on Intelligence Design (ISID 2022), Tokyo, Japan, 2022.
- [22] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "An Improved Genetic Algorithm for Yearly Profit Maximization of Wind Power Generation System," in The 31st ACM SIGDA University Demonstration, New York, USA, 2021.
- [23] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "Augmenting the Yearly Profit of Wind Farm," in The 14th Regional Conference on Electrical and Electronics Engineering of Chulalongkorn University, Bangkok, Thailand, 2022.
- [24] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "Maximizing the Yearly Profit of an Indian Nearshore Wind Farm," in First International Conference on Applied Engineering and Natural Sciences, Konya, Turkey, 2021.
- [25] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "Optimizing Offshore Wind Power Generation Cost in India," in Third New England Chapter of AIS (NEAIS) Conference, Boston, Massachusetts, 2021.
- [26] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "Realizing The Optimal Wind Power Generation Cost in Kayathar Region of India," in International Conference on Information, Communication and Multimedia Technology - 2021 (ICICMT - 2021), Madurai, 2021.
- [27] A. Turing, "Computing Machinery and Intelligence (1950)," in *The Essential Turing*, Oxford University Press, 2004.
- [28] Z. Wu and H. Wang, "Research on Active Yaw Mechanism of Small Wind Turbines," *Energy Procedia*, vol. 16, p. 53–57, 2012.
- [29] D. Wilson, S. Rodrigues, C. Segura, I. Loshchilov, F. Hutter, G. L. Buenfil, A. Kheiri, E. Keedwell, M. Ocampo-Pineda, E. Özcan, S. I. V. Peña, B. Goldman, S. B. Rionda, A. Hernández-Aguirre, K. Veeramachaneni and S. Cussat-Blanc, "Evolutionary computation for wind farm layout optimization," *Renewable Energy*, vol. 126, pp. 681–691, 2018.
- [30] U. Bhaskar, "Adani Renewable places lowest bid in SECI's wind auction," *Mint*, 2021.
- [31] P. Bhattacharjee, R. K. Jana and S. Bhattacharya, "Amplifying the Financial Sustainability of a Wind Farm near the Coast of Gujarat with an Augmented Genetic Algorithm," in International Symposium on Information & Communication Technology 2022, Greater Noida, India, 2022.

A particle swarm optimization approach for a robot in an environment with obstacles

Ines Boutabia
Computer Science Department Chadli
Ben Djedid University
El Teref, Algeria
boutabia.ines@gmail.com

Abdelmadjid Benmachiche
Computer Science Department Chadli
Ben Djedid University El Teref, Algeria
benmachiche@hotmail.fr

Maroua Hammadi
Department of computer science,
Embedded systems laboratory Badji
Mokhtar University Annaba, Algeria
marouahammadi@gmail.com

Abstract— The main objective of autonomous mobile robotics is to construct physical systems able to move intentionally and human intervention free in dynamic environments — that is, in real-world conditions that have not been particularly contrived for the robot. The evolution of methods for autonomous robot navigation constitutes one of the significant directions in the current research involving robotics. In this paper, we will be viewing those methods. Our approach today concentrates on the navigation of a robot from point A to point B in an environment with obstacles using particle swarm optimization. The trajectories achievable by the robot follow the pre-established path while bypassing both mobile and fixed obstacles. These obstacles are added to test the robots' ability to avoid, recalculate and alter their position in real-time to better test their efficiency and performance. this study has been implemented and conducted under different conditions.

Keywords— *PSO, Navigation, Autonomous mobile robot, Obstacles avoidance*

I. INTRODUCTION

The main objective of autonomous mobile robotics is to construct physical systems able to move intentionally and human intervention free in dynamic environments — that is, in real-world conditions that have not been particularly contrived for the robot. The evolution of methods for autonomous robot navigation constitutes one of the significant directions in the current researches involving robotics. This trend is inspired by the existing gap between the available technology and the new application demands. On one side, current industrial robots lack flexibility and autonomy: typically, these robots conduct pre-programmed series of processes in highly constrained circumstances, and unable to function in new environments or to face unpredictable situations. On the other, there is a clear emerging market for truly autonomous robots. Potential applications include intelligent service robots for offices, hospitals, and factory floors; maintenance robots operating in hazardous or hardly accessible areas; domestic robots for cleaning or entertainment; semi-autonomous vehicles for help to disabled people; and many more.

II. RELATED WORK

A. Fuzzy Logic Technique for Mobile Robot Navigation

The notion of fuzzy logic was introduced by Zadeh [16], which is broadly applied in multiple engineering applications such as mobile robotics, image processing, etc. This technique has a vital role in the domain of mobile robots. The fuzzy logic approach has been successfully applied by numerous researchers to maintain the position and direction of mobile robots in the environment. Ren et al. [17] have worked on an intelligent fuzzy logic controller to determine the navigation issue of wheeled mobile robots in a strange

and constantly changing environment. Fuzzy logic systems take inspiration from human reasoning, which functions based on perception. In [18], the authors have demonstrated the Gradient method based on optimal Takagi-Sugeno fuzzy controller to adjust the membership function parameters and applied it to mobile robot navigation and obstacle avoidance. Qing-Yong et al. [19] introduced the behavior-based fuzzy architecture for mobile robot navigation in unknown environments. They have created four essential behaviors: goal-seeking behavior, obstacle avoidance behavior, tracking behavior, etc. for mobile robot navigation and experimented with them in different simulation conditions. The eight rule-based fuzzy controllers have been developed by Boubertakh et al. [20] for obstacle avoidance and goal searching conduct of the mobile robot. Muthu et al. [21] have proposed the fuzzy logic controller based on the Atmega microcontroller for the wheeled mobile robot. The controller they suggested trains the mobile robot to navigate in an environment without any human intervention. The controller acquires inputs (obstacle distance) from the collection of sensors to control the right and left motors of the mobile robot.

[31-32] discussed the sensor-based mobile robot navigation in an indoor environment using a fuzzy logic controller. Wu et al. [33] have worked on sensor-based mobile robot navigation in a restricted environment employing fuzzy controllers and genetic algorithms. the fuzzy controller supplies the initial membership function and the genetic algorithm selects the most satisfactory membership value to optimize the fuzzy controller for mobile robot navigation. Obstacle avoidance is significantly necessary for the successful navigation of autonomous mobile robots. Samsudin et al. [34] have united the reinforcement learning method and genetic algorithm to optimize the fuzzy controller for enhancing its execution when the mobile robot proceeds in an unfamiliar environment. for more complicated environments, Beom and Cho [35] demonstrated the Fuzzy reinforcement learning sensor-based mobile robot navigation. Pradhan et al. [36] have operated a fuzzy logic controller with various membership functions for the navigation of a thousand robots in a completely strange environment. The authors have conducted a comparison of the implementation of different membership functions such as triangular, trapezoidal, and Gaussian for mobile robot navigation and concluded that the Gaussian membership function is more efficient for navigation. In [37], the authors have combined the fuzzy genetic algorithm to answer the path planning and control problem of autonomous mobile robot (AMR) using ultrasonic range finder sensor information. Farooq et al. [38] have exemplified a comparative study between the zero-order Takagi-Sugeno and Mamdani-type fuzzy logic models for mobile robot navigation and obstacle avoidance. Inputs (obstacle distance) are sent to both the controllers from the left and right ultrasonic sensors to maintain the left and right speeds of the engines of the mobile robot. While conducting

the comparative study, the authors have discovered that in terms of smoothness Mamdani-type fuzzy model delivers better results. On the other hand, the Takagi-Sugeno fuzzy model takes less memory space in the real-time microcontroller performance.

B. Neural Network Technique for Mobile Robot Navigation

The neural network is one of the significant methods for mobile robot navigation. This neural network technique is inspired by the human brain, which is being applied by multiple researchers in various domains to name a few signal and image processing, pattern recognition, mobile robot path planning, business, etc. Zou et al. [52] have demonstrated the literature survey of neural networks and its applications in mobile robotics. In [53], the authors have integrated the multi-layer feed-forward artificial neural network with the Q-reinforcement learning method to form a robust path-planning algorithm for the mobile robot. Rai & Rai [54] have developed the Arduino Uno microcontroller-based DC motor speed control system with the Multilayer neural network controller and Proportional Integral Derivative (PID) controller. Patino & Carelli [55] used neural network architecture to design the automatic steering controller for a mobile vehicle. Yang & Meng [56] in order to generate a collision-free path in a nonstationary environment, applied the biologically inspired neural network. A biologically motivated neural network-based wall-following mobile robot has been given by Nichols et al. [57]. Online path planning between unknown obstacles in the environment is an intriguing issue in the area of mobile robotics. Motlagh et al. [58] applied neural networks and reinforcement learning to find the target seeking and obstacle avoidance behaviors. Mobile robot navigation using a hybrid neural network has been handled by Gavrilov & Lee [59]. Singh & Parhi [60] were the minds behind a multilayer feed-forward neural network, which basically controls the steering angle of the robot autonomously in both static and dynamic environments. The four-layered neural network has various obstacle distances as inputs, and the outputs are the steering angles. Real-time collision-free path planning becomes even more challenging when the robot is in a dynamic and unstructured environment.

C. Genetic Algorithm for Mobile Robot Navigation

Ghorbani et al. [13] have found a solution to the global path planning issue of a mobile robot in a complicated environment with the help of the genetic algorithm approach. Elshamli et al. [92] have offered a genetic algorithm strategy for solving the path planning problem of a mobile robot in static and dynamic environments. Mohanta et al. [93] have created the Petri-GA method to optimize the navigation path length of numerous mobile robots in a disorganized environment. Kubota et al. [94] have put the fuzzy controller to use in order to guide the mobile robot in a static and dynamic environment, and the established genetic algorithms (GAs) are merged with it, to optimize the navigation course length. Tuncer & Yildirim [95] have suggested a new mutation operator for a genetic algorithm (GA) and involved it in mobile robot navigation in dynamic environments. Furthermore, the authors have tested the developed method in miscellaneous simulation environments and conducted a comparison between it and traditional GA processes and noted that their developed mutation operator-based GA serves better than traditional GA. In [96], the authors have designed a genetic algorithm to pick the best membership parameters from the fuzzy inference system and executed it to handle the steering angle of a mobile robot in a partially unknown environment. Hu et al. [97] have developed a

knowledge-based genetic algorithm for mobile robot navigation between U-shaped obstacles and maze environments.

Liu et al. [98] have presented the optimal path planning technique for a mobile robot using fuzzy logic and genetic algorithm. The fuzzy controllers are applied to modify the moving direction of the mobile robot according to the obstacle distance received from the sensors, and genetic algorithm is used to adjust and tune membership function and rules. Improved genetic algorithm based mobile robot navigation has been proposed by Li et al. [99]. The authors have done many simulation tests in the both static and dynamic environments to show the effectiveness of the proposed algorithm. Qu et al. [100] have developed the improved genetic algorithm instead of a conventional genetic algorithm for global path planning of the multiple mobile robots. The advantages of the improved genetic algorithm are capable of guiding the mobile robots efficiently from the starting node to end node without any collision in the environment. In [101], the authors have implemented Genetic-Fuzzy Controller (GAFLC) to optimize and tune the Gaussian membership function parameters for mobile robot motion control. Castillo et al. [102] have designed Multiple Objective Genetic Algorithm (MOGA) for navigation path optimization of the mobile robot. Arora et al. [103] have presented the single fitness based genetic algorithm for solving the navigation problem in the dynamic environments.

D. Particle Swarm Optimization Algorithm for Mobile Robot Navigation

Particle swarm optimization (PSO) is a population-based stochastic algorithm, which takes motivation from the social conduct of bird flocks. PSO algorithm is used to find an optimal or near-optimal solution to the problem using the fitness function $f(x) = f(x_1, x_2, x_3, \dots, x_i)$ where x_i is a population of the particles. Ahmadzadeh & Ghanavati [1] have demonstrated a navigation method for numerous mobile robots based on the PSO algorithm. In every iteration, these robots proceed according to the global best (g-best) position of a particle. To set an optimal intelligent regulator for a self-reliant wheeled mobile robot, Castillo et al. [2] have developed the hybridization of an Ant Colony Optimization (ACO) algorithm and the Particle Swarm Optimization (PSO) algorithm to optimize the membership function of a fuzzy controller. Zhang et al. [3] have suggested the Multi-Objective Particle Swarm Optimization Algorithm (MOPSO) with the main purpose being to search for a collision-free path in an undetermined dynamic environment. Zhang & Li [4] present an unexplored objective function for mobile robot navigation using PSO. This function operates based on the position of both the obstacles and the target in the environment. Raja & Pugazhenthii [5] successfully put the PSO algorithm in use to optimize the travel time of the mobile robot in dynamic environments. This algorithm -in every iteration- randomly searches for the viable pathway in the environment. Masehian & Sedighzadeh [6] have determined the motion planning issue of the mobile robot by using multi-objective PSO.

Wong et al. [7] developed the optimal fuzzy controller based on the PSO to define the accelerations of the left-wheeled motor and right-wheeled motor of a mobile robot with differential drive. A specialized particle swarm optimization algorithm has been proposed by Li et al. [8] for

global optimum path planning of mobile robots. The authors have performed numerous simulation trials in both simple and complicated environments to display the effectiveness of the presented algorithm. Huang [9] worked on the designing the Parallel Meta heuristic Particle Swarm Optimization (PPSO) algorithm to resolve the global path planning problem of an autonomous mobile robot. The author has executed this PPSO algorithm in real-time using the field-programmable gate array (FPGA) chip. Chung et al. [10] have worked on PSO and fuzzy-based combinatorial algorithms for a mobile robot's intelligent navigation architecture. The PSO algorithm is applied so that the robot escapes from dead-end situations, and the fuzzy algorithm controls the turning angle of a wheeled mobile robot during navigation and obstacle avoidance. Shiltagh & Jalal [11] have researched the application of Modified Particle Swarm Optimization (MPSO) in the domain of mobile robotics to define the shortest possible route from the starting point to the ending in an environment filled with obstacles. The advanced altered PSO improves the intersection rate of the algorithms. Chatterjee & Matsuno [12] have solved the Simultaneous Localization and Mapping (SLAM) issue of mobile robots or vehicles using modified PSO and fuzzy evolutionary algorithms.

Juang & Chang [13] have shown an evolutionary-group-based particle swarm-optimization (EGPSO) for an automatic understanding of fuzzy systems for mobile robot navigation or wall-following control in strange environments. In [14], the writers have altered the robot path planning situation to the minimization problem and prepared a fitness function established on the location of the target and obstacles in the environment. Allawi & Abdalla [15] have submitted the sensor-based PSO-fuzzy type-2 model for the navigation of numerous mobile robots. They have utilized the PSO algorithm to select the optimal input/output membership function parameters and rules for the fuzzy type-2 controller.

III. METHODOLOGIES

A. How it works

PSO is a robust stochastic optimization technique based on the movement and intelligence of swarms. PSO applies the concept of social interaction to problem solving. It uses a number of agents (particles) that constitute a swarm moving around in the search space looking for the best solution. Each particle is treated as a point in a N-dimensional space which adjusts its "flying" according to its own flying experience as well as the flying experience of other particles.

Each particle changes its position according to

- Its Current Velocity
- Its Current Position
- Distance between current position and g_{best} .
- Distance between current position and p_{best} .

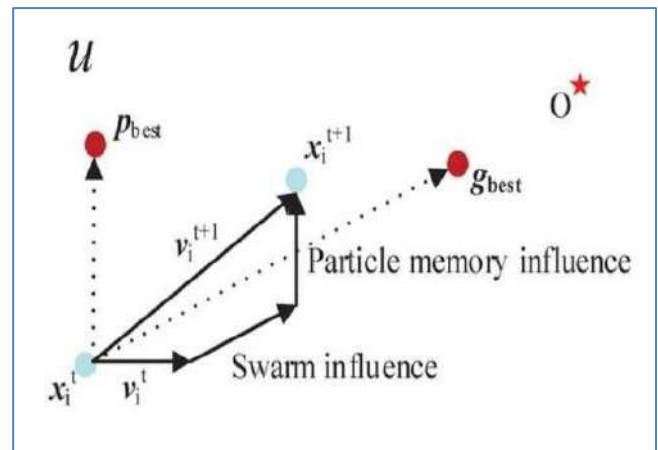


Figure 1: how PSO work

Each Element moves randomly following three simple rules:

- Cohesion: the boids are attracted to the average position of the group;
- Alignment: the boids follow the same path as their neighbours;
- Separation: to avoid collisions, the boids keep a certain distance between them.

B. Swarms and Particles

It became clear during the simplification of the paradigm that the behavior of the population of agents is now more like a swarm than a flock. The expression "swarm" has a foundation in the literature. In particular, the authors use the term in compliance with a piece by Millonas [1], who developed his models for applications in artificial life and enunciated five basic regulations of swarm intelligence. First is the proximity principle: the population should have the ability to conduct simple space and time calculations. Second is the quality principle: the population should be able to respond to quality factors in the environment. The third is the principle of diverse response: the population should not commit its activities along excessively cramped channels. Fourth is the principle of stability: the population should not alter its way of behavior every time the environment changes. Fifth, and lastly, is the principle of adaptability: the population must be able to change behavior mode when it's worth the computational price.

Note that principles four and five are the opposite sides of the same coin. Basic to the paradigm are n-dimensional space calculations taken over a series of time steps. The population is reacting based on the quality factors p_{best} and g_{best} . The allotment of responses between p_{best} and g_{best} guarantees its variousness. The population alters its state (mode of behavior) only when g_{best} changes, thus abiding by the principle of stability. The population is adaptive because it does change according to g_{best} . The term particle was selected as a settlement. arguments can be made that the population members are mass-less and volume-less, and thus could be referred to as "points," it is felt that velocities and accelerations are more suitably applied to particles, even if each is determined to have arbitrarily small mass and volume. Further, Reeves [2] discusses particle systems consisting of clouds of primitive particles as models of diffuse objects such as clouds, fire, and smoke. Therefore the designation authors have decided to define the optimization concept as particle swarm.

IV. DISCUSSION

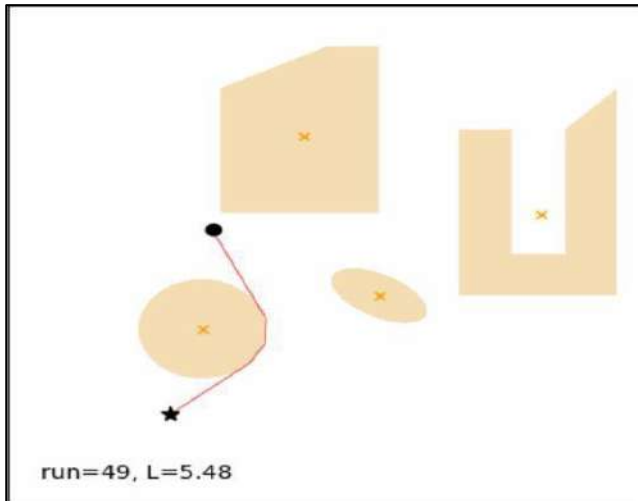


Figure 2: the circle obstacle is moving so the robot recalculates the shortest path to the goal

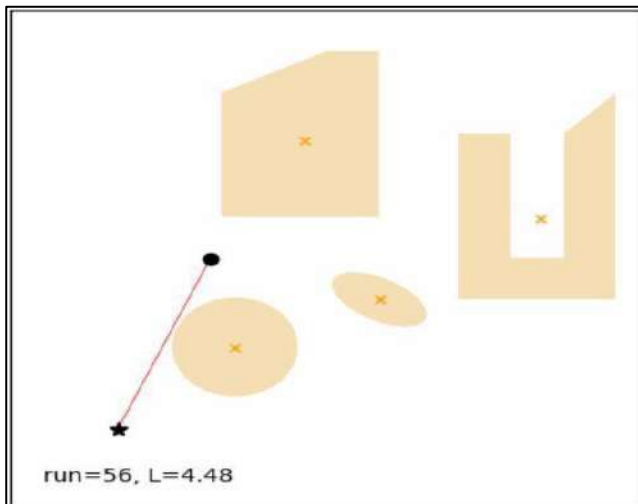


Figure 3: the robot recalculated the shortest path around the moving object

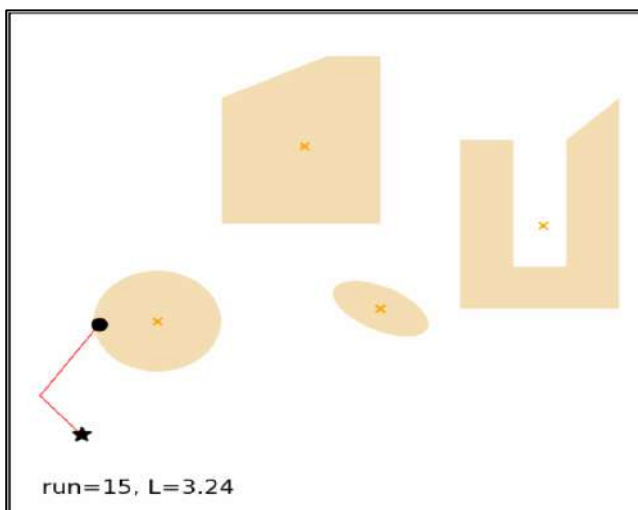


Figure 4: When the moving obstacle contacts the robot, the robot deviates from its current path forcing it to recalculate the shortest path while under the influence

V.

The most exciting part of PSO is there is a stable topology where particles are able to communicate with each other and increase the learning rate to achieve global optimum. The metaheuristic nature of this optimization algorithm gives us lots of opportunities as it optimizes a problem by iteratively trying to improve a candidate solution. Applicability of it will increase more with the ongoing research work. Potential applications include intelligent service robots for offices, hospitals, and factory floors; maintenance robots operating in hazardous or hardly accessible areas; domestic robots for cleaning or entertainment; semi-autonomous vehicles for help to disabled people; and many more.

VI. Bibliography

- [1] M. M Millonas, "Swarms, phase transitions, and collective intelligence," *Artificial Life III*, 1994.
- [2] W. T Reeves, "Particle systems - a technique for modeling a class of fuzzy objects.," *ACM Transactions on Graphics*, vol. 2, no. 2, pp. 91-108, 1983.
- [3] S Ahmadzadeh and M Ghanavati, "Navigation of Mobile Robot Using the PSO Particle Swarm Optimization.," *Journal of Academic and Applied Studies (JAAS)*, vol. 2, no. 1, pp. 32-38, 2012. [Online]. <http://academians.org/>
- [4] O Castillo, R Martinez-Marroquin, P Melin, F Valdez, and J Soria, "Comparative study of bio-inspired algorithms applied to the optimization of type-1 and type-2 fuzzy controllers for an autonomous mobile robot," *ELSEVIER Information sciences*, vol. 192, pp. 19-38, 2012. [Online]. <https://dl.acm.org/doi/10.1016/j.ins.2010.02.022>
- [5] Y Zhang, DW Gong, and JH Zhang, "Robot Path Planning in Uncertain Environment using Multi-Objective Particle Swarm Optimization.," *ELSEVIER Neurocomputing*, vol. 103, pp. 172-185, 2013.
- [6] Q Zhang and S Li, "A Global Path Planning Approach based on Particle Swarm Optimization for a Mobile Robot.," in *International Conference on Robotics, Control & Manufacturing Technology World Scientific and Engineering Academy and Society (WSEAS)*, China, 2007, pp. 263-267. [Online]. <https://dl.acm.org/doi/10.5555/1364367.1364374>
- [7] P Raja and S Pugazhenth, "Path Planning for Mobile Robots in Dynamic Environments using Particle Swarm Optimization.," in *IEEE International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom)*, India, 2009, pp. 401-405. [Online]. <https://ieeexplore.ieee.org/document/5329373>
- [8] E Masehian and D Sedighzadeh, "A Multi-Objective PSO-based Algorithm for Robot Path Planning.," in *IEEE International Conference on Industrial Technology (ICIT)*, Chile, 2010, pp. 465-470. [Online]. <https://ieeexplore.ieee.org/document/5472755>

A SURVEY ON DISC AND CUP SEGMENTATION METHODS OF RETINA FUNDUS IMAGES

Radia Touahri
Computer Science Department
Badji Mokhtar University
Annaba, Algeria
radia_inf@yahoo.fr

Nabiha Azizi
Computer Science Department
Badji Mokhtar University
Annaba, Algeria
azizi@labged.net

Nacer eddine Hammami
Computer Science Department
AL mustaqbal University
Arabie Saoudite,
Nacereddine.hammami@gmail.com

Farid benaïda
Computer Science Department
Badji Mokhtar University
Annaba, Algeria
faridbenaida@yahoo.fr

Abstract— Due to the lifestyle problems and aging, the majority of people suffer from many diseases such as high blood pressure and diabetes; in turn these problems affected many retinal disorders like diabetic retinopathy, hypertensive retinopathy, glaucoma, age-related macular degeneration etc. Most of retinal disorders can lead to vision loss if left untreated, to avoid this problem various researchers are oriented to develop automatic applications to detect these retinal disorders, for the aim of more diagnosis and less time and money. Researchers use the retina images captured by cameras to detect abnormalities.

The interpretation of retinal images combines three essential tasks; classification, detection and segmentation, whatever Deep learning (DL) or machine learning techniques (ML) have achieved high state of art as Computer Aided Diagnosis system (CAD) in both tasks (classification and segmentation) of retinal images to help ophthalmologists to detect and classify diseases.

Hence, this work, presents various traditional methods and the new approaches based on deep learning architectures used to segment retina fundus images, with listing the most datasets used for these tasks.

Keywords: Classification, Segmentation, Convolution neural network, Deep learning, Machine learning

I. INTRODUCTION

Retinal disorders cause severe eye disease that leads to blindness if it is not treated in earlier stages, around the world people suffer from several kinds of eye disease.

Many researches are focused on automated segmentation and classification of fundus images to increase prevention against disease. [3].

Any diagnosis requires regular monitoring that takes more time and needs more money. The diagnosis decision depends on the

domain knowledge of the ophthalmologists, where they may fall into subjective decisions. Therefore automatic diagnosis of retina diseases attracts a lot of attention.

Despite the good performance of traditional methods, Computer Aided Diagnosis system (CAD) Based on deep learning techniques have recently reached a large development in retina images analysis approaches, thanks to its two fundamental advantages, first the automatic generation of features (no need of handcraft features) and second the universality of the CNN, where we can recognize different patterns in several images by the same network [12].

Three retina landmarks used for eye diseases diagnosis: Optic cup and disc segmentation, blood vessels segmentation and fovea and macula localization of.

According to the results obtained in several researches, the optic cup and disc are considered as the most important retina organ that must be segmented from the images to classify it as normal or abnormal diagnosis, because the detection of the area of interest which is the optic disc focuses on the detection of the most liminal area where the recognition of the disease depends directly on the shape of the optic disc and its luminosity.

This paper surveys the state-of-the-art of automatic segmentation of retinal images to assist early diagnosis of the different eye diseases, where we listed the most works of Disc and cup segmentation based on traditional methods and the ones based on deep learning applications. We have also conducted critical evaluation of the existing automatic segmentation methods. Moreover we talk about the most public data sets used in retinal images diagnosis.

II. PUBLIC DATABASES USING FOR RETINAL IMAGES DIAGNOSIS

The availability of publicly accessible image databases opens the door to the development of design strategies,

implementation and testing of algorithms to analyze the morphology of the retina and even to diagnose other diseases. Innovation in new areas of computing, such as deep learning, require a huge amount of data to process. This need has led several research groups to make their own publicly accessible retinal image datasets. This trend results in the generation of databases with a wide variety of pathological signs in high resolution with the segmentation of the retinal land marks (optic disc, blood vessels and macula) accompanied by instructions to facilitate user access. This makes it possible to

compare the performance of the algorithms within the same background image.

In order to encourage people to use public data bases, we have summarized the main features and key information of each retina database in the following table:

Table 1: Summary of publicly available fundus image data bases:

Le nom de la base	Nombre des images	Système de caméra de fond	Domaine d'utilisation
DRIVE (Utrecht Database) – Digital Retinal Images for Vessel Extraction	40	Canon CR5 (Canon Corp., Tokyo, Japan)	blood vessels segmentation
REVIEW – Retinal Vessel Image Set for Estimation of Widths	16	HRIS – Canon 60 UV VDIS – Zeiss fundus camera CLRIS – Zeiss FF 450 (Carl Zeiss AG, Oberkochen, Germany) KPIS – Canon 60 UV (Canon Corp.)	blood vessels segmentation
CHASE_DB1 – Child Heart and Health Study in England	28	Nidek NM-200-D (Nidek Co., Ltd., Gamagori, Japan)	blood vessels segmentation
HRF – High Resolution Fundus Image Database	45	Canon CF-60 Uvi Canon EOS 20D (Canon Corp.)	blood vessels segmentation- Détection de disque optique (OD)
STARE – Structured Analysis of the Retina	81	TopCon TRV-50 (Topcon Corp., Tokyo, Japan)	blood vessels segmentation- Artery/Vein Labeling
VICAVR	58	TopCon camera NW-100 (Topcon Corp.)	Vessel caliber - Artery / Vein Labeling
INSPIRE - AVR	40	Zeiss Fundus Camera (Carl Zeiss AG)	artery / vein labeling OD segmentation
ORIGA(-light) Online Retinal Fundus Image Database for Glaucoma Analysis	650		OD & OC segmentation and Cup-to-Disc Ratio (CDR)
BioImLab	60	TopCon TRC-50 (Topcon Corp.)	Assessment of vessel tortuosity
DRIONS-DB – Digital Retinal Images for Optic Nerve Segmentation Database	110	35 mm Film Color Fundus Camera Digitized using HP-Photo Smart-S20 scanner (HP, Palo Alto, CA, USA)	OD Segmentation
RIM-ONE	475	Nidek AFC-210 (Nidek Co., Ltd.) Canon EOS 5D (Canon Corp.)	OD and OC Segmentation
ImageRet	219	Digital Fundus Camera – Settings Unknown	Diabetic retinopathy (DR) Grading Lesions
MESSIDOR	1200	TopCon TRC NW6 (Topcon Corp.)	Diabetic retinopathy (DR) Macula Edema
E-Ophtha	463	Telemedical Network Various Cameras used	Classification of DR lesions
ARIA – Automated Retinal Image Analysis	450	Zeiss FF450+ Fundus Camera (Carl Zeiss AG)	Segmentation of Optic Nerve Head Limit
Reffuge	1200	Zeiss Visucam 500 Canon CR-2	Classification du glaucome Segmentation du OD et OC Localisation de la fovéa
Drishti-GS	101	----	segmentation OD et OC

III. Related work

Computer aided diagnosis system based deep learning reached high results in classification and segmentation tasks. Optic cup and disc considered as the most important retina

organ that must segmented from the images to classify it as normal or abnormal diagnosis, where doctors have proposed several norms related to OD and OC regions, like: the

position, the center and Ratio (distance between OD and OC) that can facilitate early glaucoma diagnosis[5]

In this section we present an overview of several methods for medical images segmentations, where we have focused on encoder-decoder architectures which have made a large difference in this last years. The most of these works was evaluated by their authors using publicly available datasets. and their results are well detailed in this section, our related works are divided on two important domains :

a) **III.1. Disc and cup segmentation based on traditional methods**

Authors	Approach Used for segmentation		Datasets	Diagnosis characteristic	results
	OD	OC			
Liu et al [12]	Level set approach	Level set approach	Private dataset of 73 images	CDR	Avg. CDR acceptability for VLS-CIT 79%
Zhang et al. [13]	--	Variational level set approach	Private dataset of 71 images	CDR	Accuracy 97.2% Acceptable CDR ratio 0.18
Ayushi et al. [14]	Thresholding	Thresholding	private dataset of 60 images	CDR	Accurac y=90
Noor et al. [15]	Thresholding	Thresholding	DRIVE dataset of 27images	CDR	Accuracy of optic cup 67.2% Accuracy of optic disc 70.9%
Sarkar et al. [16]	Thresholding	Thresholding	RIMONE dataset	CDR	Accuracy 97.58%
Yuji et al. [17]	Active contourmodeling	Active contourmodeling	Local dataset of 50 images	CDR, RDR	Area under curve for cup to disc ratio 0.966
Yin et al. [18]	Active shape modeling	Active shape modeling	ORIGA-light dataset	CDR	Dice metric for optic disc 0.92
Thakur et al. [19]	Clustering	Clustering	Drishiti-GS	CDR	Accuracy of optic cup 91.9 Accuracy of optic disc 94.46
Lim et al. [20]	Hybrid	Hybrid	MESSIDOR, SEEDDB	CDR	Nonoverlap ratio for optic disc 0.08 Nonoverlap ratio for optic cup 0,23

III.2. Disc and cup segmentation based en encoder-decoder architecture

Recently most aided diagnosis studies have turned to deep learning architectures based CNN, to overcome the limited of the earliest sited above, specially the problem of handcraft features extraction [21] and to beneficiate the universality of convolutional neural networks where the same network can generally recognize different patterns in different images and for different objects [26]. Encoder_Decoder approach is novel architecture of deep learning that have achieved the state-of-the-art in several studies by improving the segmentation performance for detecting glaucoma diseases.

Retinal fundus image processing is the way for detecting eye disorders in early stages. optic disc and optic cup, are the most suitable features that can be segmented automatically from the fundus image to analyze the pathological situation of the eye[22]. Where the analysis based on the results of calculation of the most important ocular indicator: Ratio (CDR) to confirm the finally decision diagnosis[23]. many studies have based on U-Net

:

The majority of traditional studies for automated glaucoma diagnosis, were focused on the segmentation of OD and OC criterion , were mostly based on the extraction of hand-crafted features , several approaches were implemented in this filed, such as : Level set approach that proposed by Osher and Sethian in 1988[6], Thresholding based approach [7], Active contour modeling approach [8], Active shape modeling approach proposed by ChrisTaylor and Tim Cootes in 1995[9], Clustering based approach [10]. Component-based approach [11]. some of these works will be summarized in the table below:

Table 2: Disc and cup segmentation approaches based on traditional methods

architecture to improve their systems, UNet was introduced by Ronneberger *et al.* [24]. As more elegant convolutional neural network architecture with U shape , it is specially developed for the segmentation of medical images. After, Many new architectures of glaucoma diagnosis are inspired from UNet : Yu et al, [25] proposed a new version of U-Net architecture ,where U-net undergone several changes for optic disc and cup segmentation thanks to data augmentation. authors used ResNet-34 model as the down-sampling encoder section of the U-Net architecture, but they kept the up-sampling decoder section of the original U-Net structure which formed by a proposed global context attention (GCA) block and squeeze and excitation pyramid pooling (SEPP) block. This architecture trained on RIGA, DRISHTI-GS, RIM-ONE and one local datasets, they achieved significant results. In the same spirit, Sevastopolsky and al, [26] have also investigated from the U-Net architecture for optic disc and cup segmentation , and introduced some modification in the Contracting path that takes the typical architecture of the classification network (VGG-16) with less filters in all convolutional

layers that's gives better results with less training time. moreover Graham et al, [27] proposed M-Net , as a multi-label deep network for joint the segmentation of Optic disc and Optic Cup of fundus images it is an end to end deep-learning system, where authors apply a polar transformation on fundus image .they evaluate this system on ORIGA and SCES datasets to get highest results :areas under curve (AUC) of 0.85 and 0.90 on ORIGA and SCES datasets. Always in the same subject, Al-Bander and al, [28] proposed a new approach for simultaneous segmentation of OD and OC in retina fundus images by combination of fully convolutional network and DenseNet(FC-DenseNet) using large retinal colour fundus dataset (ORIGA, DRIONS-DB, Drishti-GS , ONHSD , and RIM-ONE) to

acheived superiority in their results, The recent related work is the Retinal Fundus Glaucoma Challenge (REFUGE) proposed by Orlando and al,[21] which based on two tasks optic disc/cup segmentation and glaucoma classification using a large data set of 1200 fundus images. 12 teams of work have created a unified framework to compare different models of deep learning architectures.they have obtained good results the table below conducted us deeper into details of these studies;

Table 3: Disc and cup segmentation based en encoder-decoder architecture

Author	year	Approach Used for segmentation	Dataset	Nbre of classes	Results
Yu et al, [25]	2019	Adaptation of the ResNet-34 model into U-Net architecture (encoder – decoder)	DRISHTI-GS, RIM-ONE and RIGA(Magrabia),	2	Dice(Disc)= 96.44 Jacc(Disc)= 93.19 Dice(Cup)= 87.39 Jacc(Cup)= 78.08 } DRISHTI-GS Dataset (%) Dice(Disc)= 94.91 Jacc(Disc)= 90.65 Dice(Cup)= 79.32 Jacc(Cup)= 68.28 } RIM-ONE dataset (%) Dice(Disc)= 97.31 Jacc(Disc)= 94.80 Dice(Cup)= 87.61 Jacc(Cup)= 79.40 } Magrabia dataset (%)
Graham et al, [27]	2018	M-net (encoder – decoder) multi-label deep network	ORIGA and SCES datasets	2	CDR=0.8019 RDAR=0.7981 } ORIGA Dataset CDR =0.8397 RDAR=0.8290 } SCES Dataset
Al-Bander and al, [28]	2018	Fully convolutional DenseNet	ORIGA, DRIONS-DB , Drishti-GS [42], ONHSD [10], and RIM-ONE [43]. Glaucoma	3	<i>optic cup segmentation</i> Dice= 0.8282 Jacc=0.7113 Acc=0.9948 } DRISHTI-GS Dataset <i>optic disc segmentation</i> Dice= 0.949 Jacc=0.9042 Acc=0.9969 } DRISHTI-GS Dataset <i>optic cup segmentation</i> Dice= 0.6903 Jacc=0.5567 } RIM-ONE Dataset <i>optic disc segmentation</i> Dice= 0.9036 Jacc=0.8289 } RIM-ONE Dataset <i>optic disc segmentation</i> Dice= 0.94 Jacc=0.9415 } DRIONS-DB Dataset <i>optic disc segmentati ...</i> Dice= 0.9556 Jacc=0.9155 Acc=0.999 } ONHSD Dataset
Sevastopolsky and al, [26]	2017	Modified U-Net with less filters in all convolutional layers (encoder – decoder)	DRIONS-DB, RIM-ONE v.3 and DRISHTI-GS		<i>optic disc segmentation</i> IOU= 0.89 Dice= 0.94 IOU= 0.89 Dice= 0.95 } DRIONS-DB Dataset <i>optic cup segmentation</i> IOU= 0.75 } RIM-ONE v.3 Dataset } DRISHTI-GS Dataset

					Dice= 0.85 IOU= 0.69 Dice= 0.82
					} RIM-ONE v.3 Dataset
Orlando and al.[21]	2020	Encoder decoder models	REFUGE training set, DRISHTI-GS, HRF, and ORIGA RIM-ONE r3 Kaggle MESSIDOR and IDRiD ORIGA		<i>Results in the paper</i>

V. Discussion

This review outlines both domains traditional approaches and deep-learning methods based optic cup and disc segmentation models from fundus images. Recently, Deep learning techniques has enormous potential for clinical application because it is a robust method that overcomes many of the limits associated with traditional ones especially concerning the manual features extraction . According to our survey, we can say that The most useful deep learning techniques for unsure the optic cup and disc segmentation task are: traditional CNN, and encoder-decoder architecture, where each network architecture has its advantages and disadvantages.

- for the traditional CNN network it needs too much accurately labeled images to offer excellent optic cup and disc segmentation;
- for encoder-decoder networkes that achieved the state-of-the-art to prove the segmentation task thanks to its training simplicity and data efficiency. but, its multi-scale skip connection tends to use unnecessary information that leading to poorer their performance.
- The segmentation task is so difficult comparing by the classification task. It is not easy to compare and identify the best architecture for optic cup and disc segmentation. This challenge leads researchers to use different metrics to evaluate their work using the public annotated datasets.

V. CONCLUSION

This paper has shown different techniques for disc and cup segmentation based on traditional methods and deep learning ones. We have surveyed some literature papers and listed their workings with results.

Disc and cup organs are the most important features that can be segmented from retina funds images because any changement of their shap or liminosity can easily detect any kind of abnormalities in the eye, especially diabetic retinopathy and glaucoma diseases. The survey results confirmed the robustness of encoder-decoder systems based deep learning applications on the automatic segmentation task of retina fondus images.

deep learning approaches needs large dataset to provide their friability, for this most of the experiments cited on this survey are based on publicly available databases.

REFERENCES

- [1]. - S. Anil, E. Issac : "Survey on Various Methods of Detecting Glaucoma ",international Journal of Trend in Scientific Research and Development (IJTSRD),vol.2,pp.2456-2470,2018.
- [2]. - N.S .Chethan Kumar, k.Somashekar:" Image Processing Techniques for Automatic Detection of Glaucoma",vol.VI,pp.1-8,2017.
- [3]- A. Tugashetti, T.C.Manjunath, Pavithra.g:" Design of algorithms for diagnosis of primary glaucoma through estimation of CDR in different types of Fundus Images using IP techniques", International Journal of Innovative Research in Information Security (IJIRIS), Vol. 04, pp.1-8,2017.
- [4] J.Bullock, C. Cuesta-L'azaro, A. Quera-Bofarull," XNet: A convolutional neural network (CNN) implementation for medical X-Ray image segmentation suitable for small datasets", SPIE Medical Imaging, vol. 10953, pp.1-11,2018.
- [5] M. N. Bajwa, M. I. Malik, S. A. Siddiqui, A. Dengel, F. Shafait, W. Neumeier and S.Ahmed:" Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning, BMC Medical Informatics and Decision Making ,pp.1-17, 2019
- [6] Mark Sussman, Peter Smereka, Stanley Osher, A level set approach for computing solutions to incompressible two-phase flow, J. Comput. Phys. 114 (1) (1994) 146–159
- [7] R.C. Gonzalez, Woods R.E. Thresholding, Digital image processing, Pearson Educ. 59 (2002) 595–611
- [8] Michael Kass, Andrew Witkin, Demetri Terzopoulos, Snakes active contour models, Int. J. Comput. Vision 1 (4) (1988) 321–331.
- [9] T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, Active shape models-their training and application, Comput. Vision Image Understanding 61 (1) (1995) 38–59.
- [10] Masulli Francesco, Andrea Schenone, A fuzzy clustering based segmentation system as support to diagnosis in medical imaging, Artif. Intell. Med. 16 (2) (1999) 129–147
- [11] Hanan Samet, Markku Tamminen, Efficient component labeling of images of arbitrary dimension represented by linear bintrees, IEEE Trans. Pattern Anal. Mach. Intell. 10 (4) (1988) 579–586
- [12] J. Liu, D.W.K. Wong, J.H. Lim, X. Jia, F. Yin, H. Li, W. Xiong, T.Y. Wong, Optic cup and disk extraction from retinal fundus images for determination of cup-to-disc ratio, IEEE Proceedings of 3rd Conference on Industrial Electronics and Applications (2008) 1828–1832
- [13] Zhang Zhuo, Jiang Liu, Wing Kee Wong, Ngan Meng Tan, Joo Hwee Lim, Shijian Lu, Huiqi Li, Ziyang Liang, Tien Ying Wong, Neuro-retinal optic cup detection in glaucoma diagnosis, IEEE Proceedings of 2nd International Conference on Biomedical Engineering and Informatics (2009) 1–4
- [14] Ayushi Agarwal, Shradha Gulia, Somal Chaudhary, Malay Kishore Dutta, Carlos M. Travieso, Jesus B. Alonso-Hernandez, A novel approach to detect glaucoma in retinal fundus images using cup-disk and rim-disk

ratio, Proceedings of 4th International Work Conference on Bioinspired Intelligence (2015) 139–144

[15] N.M. Noor, N. Khalid, N.M. Ariff, Optic cup and disc color channel multi-thresholding segmentation, Proceedings of IEEE International Conference on Control System, Computing and Engineering (2013) 530–534

[16] Sarkar Debasree, Das. Soumen, Automated glaucoma detection of medical image using biogeography based optimization, in: Proceedings of Advances in Optical Science and Engineering, Springer Singapore, 2017, pp. 381–388.

[17] Hatanaka Yuji, Yuuki Nagahata, Chisako Muramatsu, Susumu Okumura, Kazunori Ogohara, Akira Sawada, Kyoko Ishida, Tetsuya Yamamoto, Hiroshi Fujita, Improved automated optic cup segmentation based on detection of blood vessel bends in retinal fundus images, Proceedings of 36th Annual International Conference on Engineering in Medicine and Biology Society (2014) 126–129

[18] Yin Fengshou, Jiang Liu, Damon Wing Kee Wong, Ngan Meng Tan, Carol

Cheung, Mani Baskaran, Tin Aung, Tien Yin Wong, Automated segmentation

of optic disc and optic cup in fundus images for glaucoma diagnosis, Proceedings of 25th International Symposium on Computer-based Medical Systems (2012) 1–6.

[19] N. Thakur, M. Juneja, Clustering based approach for segmentation of optic

cup and optic disc for detection of glaucoma, Curr. Med. Imaging Rev. 13 (1)

(2017) 99–105.

[20] Gilbert Lim, Yuan Cheng, Wynne Hsu, Mong Li Lee, Integrated optic disc

and cup segmentation with deep learning, Proceedings of 27th International Conference on Tools with Artificial Intelligence (2015) 162–169.

[21] J. I. Orlando, Huazhu Fu, J. B. Breda, K. van Keer, D. R. Bathula, A. Diaz-Pinto, R. Fang, P-Ann Heng, J. Kim, J. Lee, J. Lee, X. Li, P. Liu, S. Lu, B. Murugesan, V. Naranjo, S.S. R. Phayee, S.M. Shankaranarayana, A. Sikka,

J. Son, A. van den Hengel, S. Wang, J. Wu, Z. Wu, G. Xu, Y. Xu, P. Yin, F. Li, X. Zhang, Y. Xu, H. Bogunovic: “REFUGE Challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs”, Medical Image Analysis, vol.59 .pp. 1-21,2020 .

[22]R.Priyanka,S.J.G.Shoba,A.BTherese, “Segmentation of Optic Disc in Fundus Images using Convolutional Neural Networks for Detection of Glaucoma”, *International Journal of Advanced Engineering Research and Science (IJAERS)*, Vol.4, pp.1-10,2017.

[23] N. Thakur, M. Juneja : “ Survey on segmentation and classification approaches of optic cup and optic disc for diagnosis of glaucoma”, Biomedical Signal Processing and Control ,vol.42, pp.162–189, 2018

24- O. Ronneberger, P. Fischer, T. Brox,” U-Net: Convolutional Networks for Biomedical Image Segmentation ” , vol.9351, Medical Image Computing and Computer-Assisted Intervention – MICCAI, pp.234-241, 2015.

25-S. Yu, D. Xiao, S. Frost, Y. Kanagasingam, “Robust optic disc and cup segmentation with deep learning for glaucoma detection”, Computerized Medical Imaging and Graphics , vol.74 , pp.61–71, 2019.

26- A. Sevastopolsky,” Optic Disc and Cup Segmentation Methods for Glaucoma Detection with Modification of U-Net Convolutional Neural Network”, Pattern Recognition and Image Analysis ,Vol.27,pp.618-624,2017

27-H. Fu, J. Cheng, Y.Xu, D. W. K. Wong, J. Liu, X. Cao,” Joint Optic Disc and Cup Segmentation Based on Multi-label Deep Network and Polar Transformation”, *IEEE Transactions on Medical Imaging* ,Vol. 37 , pp.1-9 , 2018.

28- B. Al-Bander, B. M. Williams , W. Al-Nuaimy , M. A. Al-Taei, H.Pratt , Y. Zheng, “Dense Fully Convolutional Segmentation of the Optic Disc and Cup in Colour Fundus for Glaucoma Diagnosis”, MDPI ,pp.1-16,2018.

29-Z.Chena , D.Lia , H.Shena , H.Moa , Z. Zenga, H.Weia ; : ‘Automated segmentation of fluid regions in optical coherence tomography B-scan images of age-related macular degeneration’ , Optics and Laser Technology,2019.

30- J. Ni, J. Wu , J. Tong , Z.Chen , J. Zhao,’ GC-Net: Global context network for medical image segmentation’, Computer Methods and Programs in Biomedicine,pp.1-10,2019.

31_M.Dash, N.D. Londhe, S. Ghosh, A.Semwal , R. S. Sonawane ,” PsLSNet: Automated psoriasis skin lesion segmentation using modified U-Net-based fully convolutional network”, Biomedical Signal Processing and Control, vol 52,pp. 226–237, 2019.

32 S. Alqazzaz, X. Sun, X. Yang, L. Noke:” Automated brain tumor segmentation on multi-modal MR image using SegNet”, Computational Visual Media, Vol. 5, pp. 209–219, 2019.

33- P. Kumar , P.Nagar, C. Arora , A. Gupta,” U-SEGNET: FULLY CONVOLUTIONAL NEURAL NETWORK BASED AUTOMATED BRAIN TISSUE SEGMENTATION TOOL”, pp.1-5, 2018.

34-M.A. Al-masni, M.A. Al-antari, M.Choi , S.Han , T.Kim,” Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks”, Computer Methods and Programs in Biomedicine , vol.162, pp.221–231, 2018.

35-Z.Zhou, N. Tajbakhsh, M. R. Siddiquee, J. Liang:” UNet++: A Nested U-Net Architecture for Medical Image Segmentation”,Springer Nature Switzerland, pp.1-10 , 2018

36-Zhang, C., Bengio, S., Hardt, M., Recht, B., Vinyals, O., “Understanding deep learning requires rethinking generalization”. arXiv preprint arXiv:1611.03530,2016

[37]N. Tajbakhsh, L.Jeyaseelan, Q. Li, J. Chiang, Z. Wu, X. Ding,“Embracing Imperfect Datasets: A Review of Deep Learning Solutions for Medical Image Segmentation”, arXiv:1908.10454v1 [eess.IV] , 2019

[38] A. M. N. Allam , A. A.Youssif , A. Z. Ghalwash:” Automatic Segmentation of Optic Disc in Eye Fundus Images:”, Electronic Letters on Computer Vision and Image Analysis ,vol.14(1);pp.1-20, 2015.

[39] G.Ramani.L.bramanian : » Macula segmentation and fovea localization employing image processing and heuristic based clustering for automated retinal screening, *Biomedicine, Volume 160*, July 2018, Pages 153-163.

[40] R.Alais, P.Dokl dal, A.Erginay, B.Figliuzzi, E.Decenci re:” Fast macula detection and application to retinal image quality assessment”, *Biomedical Signal Processing and Control* , Vol. 55, pp.101_567, 2020.

[41] F.Girard-C.Kavalec-F.Ch riet: ”Joint segmentation and classification of retinal arteries/veins from fundus images”, *Artificial Intelligence in Medicine*,Vol. 94, pp. 96-109, 2019.

[42] Refuge Data set; <https://ai.baidu.com/broad/introduction>. Last access 26/06/2020

[43] G.P.Gonz lez : “D tection et segmentation des disques optiques » , universit  coruna, 2018

[44] F. Chollet, et al., Keras, 2020, online available (<https://keras.io>)

[45] M. Abadi, et al., TensorFlow: a system for large-scale machine learning, in: 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), 2016, pp. 265–283

An Approach for transforming Activity Diagram

Hiba HACHICHI Halim BOUALEM Mohammed BOUZIANE ERRAHMANI Oussama HARBOUCHE
Djilali Bounaama Khemis Miliana University
Khemis Miliana, Algeria
{h.hachichi, h_boualem, m_bouziane-errahmani, o.harbouche}@univ-dbkm.dz

Abstract— The work presented in this article is a contribution in the field of Model Driven Engineering (MDE). Its main objective is the application of model transformation techniques, and more precisely graphs transformation. Therefore, we propose an approach and a tool for modeling and transforming Activity Diagrams into Timed Automata. The implementation of this approach requires the use of the EMF (Eclipse Modeling Framework), the GMF (Graphical Modeling Framework) and ATL (Atlas Transformation language) as a language of transformation.

Keywords— *Activity Diagrams, Timed Automata, EMF, GMF, ATL, Graph Transformation*

I. INTRODUCTION

The Unified Modeling Language (UML) [14] has become a widely accepted standard in the software development industry. Some diagrams are used to model the structure of a system while others are used to model the behavior of a system. Activity Diagram [12] is an important behavioral diagram in UML diagrams to describe dynamic aspects of the system. This diagram presents a series of actions or flow of control in a system similar to a flowchart. Activity Diagram is also used and extended in other modeling languages such as System Modeling Language (SysML) [15].

However, UML lacks a precise formal semantic, which hinders the formal evaluation and verification of system structure. Timed Automata (TA) [2] is a finite automaton extended with a finite set of real-valued clocks and it is a formal and graphical tool.

In this paper, we propose an approach and a tool for transforming Activity Diagrams into Timed Automata using graph transformation [3], [8]. The aim of this transformation is to bridge the gap between semi-formal models generated by Activity diagrams and the formal model Timed Automata.

Indeed, we propose a meta-model for each model: Activity Diagram meta-model and a Timed Automata meta-model. The Eclipse Modeling Framework (EMF) [6] and the Graphical Modeling Framework (GMF) [7] tools are used to generate automatically a visual Modeling tool to process models in Activity Diagrams and Timed Automata. We also define a graph grammar to translate models presented above.

This paper is organized as follows: section 2 outlines main related work relevant to the paper topic. Section 3 states the basic concepts of Activity Diagrams, Timed Automata and graph transformation. Section 4 describes the proposed approach. In section 5, the proposed approach is

illustrated through an example. The final section concludes the paper and gives some perspectives.

II. RELATED WORK

The most closely topics related to this work are generally dealing with model transformations and especially graph transformations. In the literature, several research works have been conducted about the integration of UML Activity Diagrams and formal methods.

[9] propose a graphical mapping of an Activity Diagram into a colored petri net model. In [1] the author presents a set of rules to translate a UML Activity Diagram into a Petri Nets. This transformation allows to analyze the equivalent Petri Nets and infer properties of the original workflow. [16] transform UML Activity Diagrams to petri nets also they use the model checker of INA tool to verify the properties of the resulted petri nets models. [12] suggest an automatic translation of UML Activity Diagram to Petri nets. Indeed, authors in this work focus on the most frequently used items on Activity Diagram: state action, condition, fork and join. In [11], an approach to transform automatically UML Activity Diagram to Maude specification has been proposed. This approach is based on the combined use of Meta-Modeling and graph grammars and it is carried out using AToM3. [4] propose an approach that automates the mapping between UML Activity Diagrams and pi-calculus using graph transformation and AToM3 tool. [5] propose an approach consists of transforming UML 2.0 Activity Diagrams to the communicating sequential process (CSP) for modeling and verifying software systems. The meta-modeling is realized using the AToMPM tool, while the model transformation and the correctness of its properties are realized using the GROOVE tool.

III. BACKGROUND

In the following we recall some basic notions about Activity Diagram, Timed Automata and graph transformations.

A. Activity Diagram

An Activity Diagram is a behavioral diagram, it is a variant of UML state machines [14]. We use Activity Diagram to illustrate the flow of control in a system, it is represented as a graph composed with nodes connected by control flows and object flows. It represents the operations of a process and their consequences on objects (software or hardware). Modeling can be used to describe the progress of a use case or method.

B. Timed Automata

Timed Automata [2] are finite state machines whose transitions are decorated by clocks constraints. They are widely studied as model in which the control of real time systems is finite. In fact, timed automata are construction of both finite set of locations and finite set of clock variables. Each edge is specified by a label name which contains action (that is going to be executed) and clocks formula. Those are considered as a guard of the edges and set of clocks which are going to be reset. Clock variables, in fact capture the time elapsed since the last clocks rest.

The execution (control) of automaton proceeds along an edge only when the valuation on clocks satisfies the corresponding constraint.

C. Graph Transformation

A transformation between models is the automatic generation of a target model from a source model. This task requires a set of rules that describe how one or more constructs in the source language can be transformed to one or more constructs in the target language.

Graph Grammars [10], [8], [13] are used for model transformation. They are composed of production rules; each one has graphs in their left and right hand sides (LHS and RHS). Rules are compared with an input graph called host graph. If a matching is found between the LHS of a rule and a sub graph in the host graph, then the rule can be applied and the matching sub graph of the host graph is replaced by the RHS of the rule. Furthermore, rules may also have a condition that must be satisfied in order to apply the rule, as well as actions to be performed when the rule is executed. A rewriting system iteratively applies matching rules in the grammar to the host graph, until no more rules are applicable.

IV. THE APPROACH

In order to perform the transformation of Activity Diagram to Timed Automata, we propose two meta-models; the first one associated to the Activity Diagram model and the second one is associated to the Timed Automata model.

We note here that the meta-models are described using UML class diagrams. Then a grammar for the transformation

is proposed using ATL language. Meta-models are implemented in EMF and GMF using java language.

Activity Diagram Meta-Model

This meta-model is a class diagram composed of the following classes and associations (Figure 1):

- **ControlNode** class: it represents an abstract activity node used to coordinate the flows between the nodes of an activity. It is the superclass of the end nodes classes (final node, final flow), initial node, fork node and decision node.
- **Initialnode** class: represents the start of an Activity Diagram. graphically it is represented by a small circle.
- **DecisionNode** class: this class specifies the different possible alternatives; it has one incoming arc and two or more outgoing arcs. An arc can be guarded by conditions.
- **ForkNode** class: this class represents a synchronization node; it has a single incoming arc and several outgoing arcs which must be triggered simultaneously.
- **MergeNode** class: this class gathers several incoming flows into a single outgoing flow.
- **JoinNode** class : it represents a synchronization node which can only be crossed when all the input transitions have been triggered.
- **FinalNode** class: indicates successful termination. It has one or more incoming arcs and no outgoing arcs. It is visually represented by an empty circle containing a small circle.
- **ObjectNode** class: it represents an abstract meta-class for defining the flow of objects in Activity Diagrams.
- **Transition** class: it represents the transitions of Activity Diagrams.

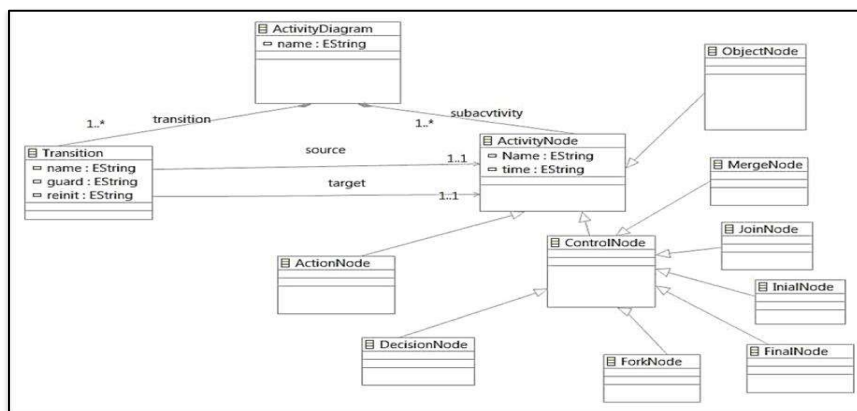


Fig. 1. Activity Diagram Meta-Model

Timed Automata Meta-Model

This meta-model is a class diagram composed of the following classes (Figure 2):

- State class: it represents the states of Timed Automata.
- Transition class: it represents the transitions of Timed Automata.

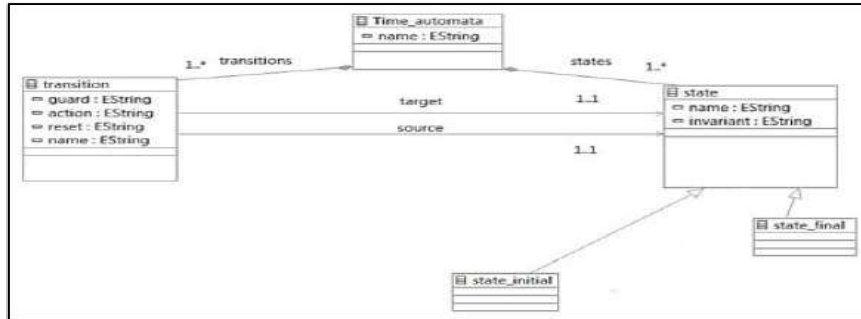


Fig. 2. Timed Automata Meta-Model

Graph Grammar

The proposed graph grammar is composed of the following rules (Figure 3):

- Rule 1: This rule transforms the "ActivityDiagram" class to the "timed_automata" class.

The ActivityDiagram class is composed of two classes: the "Transition" class and the "ActivityNode" class. We transform the "Transition" class of the Activity Diagram to the "Transition" class of the Timed Automata and the "ActivityNode" class of the Activity Diagram to the "State" class of the Timed Automata.

- State_initial class: it represents the initial state of Timed Automata, it inherits the attributes from State class.
- State_final class: it represents the final state of Timed Automata, it inherits the attributes from State class.

- Rule 2: This rule transforms the "InitialNode" class of the Activity Diagram to the "State_Initial" class of the Timed Automata.
- Rules from 3 to 8 transform these classes: "Action", "ObjectNode", "JoinNode", "ForkNode", "DecisionNode" and "MergeNode" of the Activity Diagram to "State class" of the Timed Automata.
- Rule 9: This rule transforms the "FinalNode" class of the Activity Diagram to the "state_Final" class of the Timed Automata.



Fig. 3. Set of proposed rules in our graph grammar

V. EXAMPLE

For monitoring the ambient temperature in a house, we have a program that works as follows:

The process starts by measuring the temperature for (10sec):

- If the temperature is less than or equal to 20° ($\leq 20^\circ$) the house must be heated: the heater operates for 30 minutes, then the temperature is measured again.

- Otherwise, if the temperature is greater than or equal to 20° ($> 20^\circ$): The heating is stopped and the house must be ventilated for (30min), then the temperature is measured again.

Figure 4 presents the Activity Diagram of the proposed example. We have applied our approach on the Activity Diagram and obtained automatically the Timed Automata model (Figure 5).

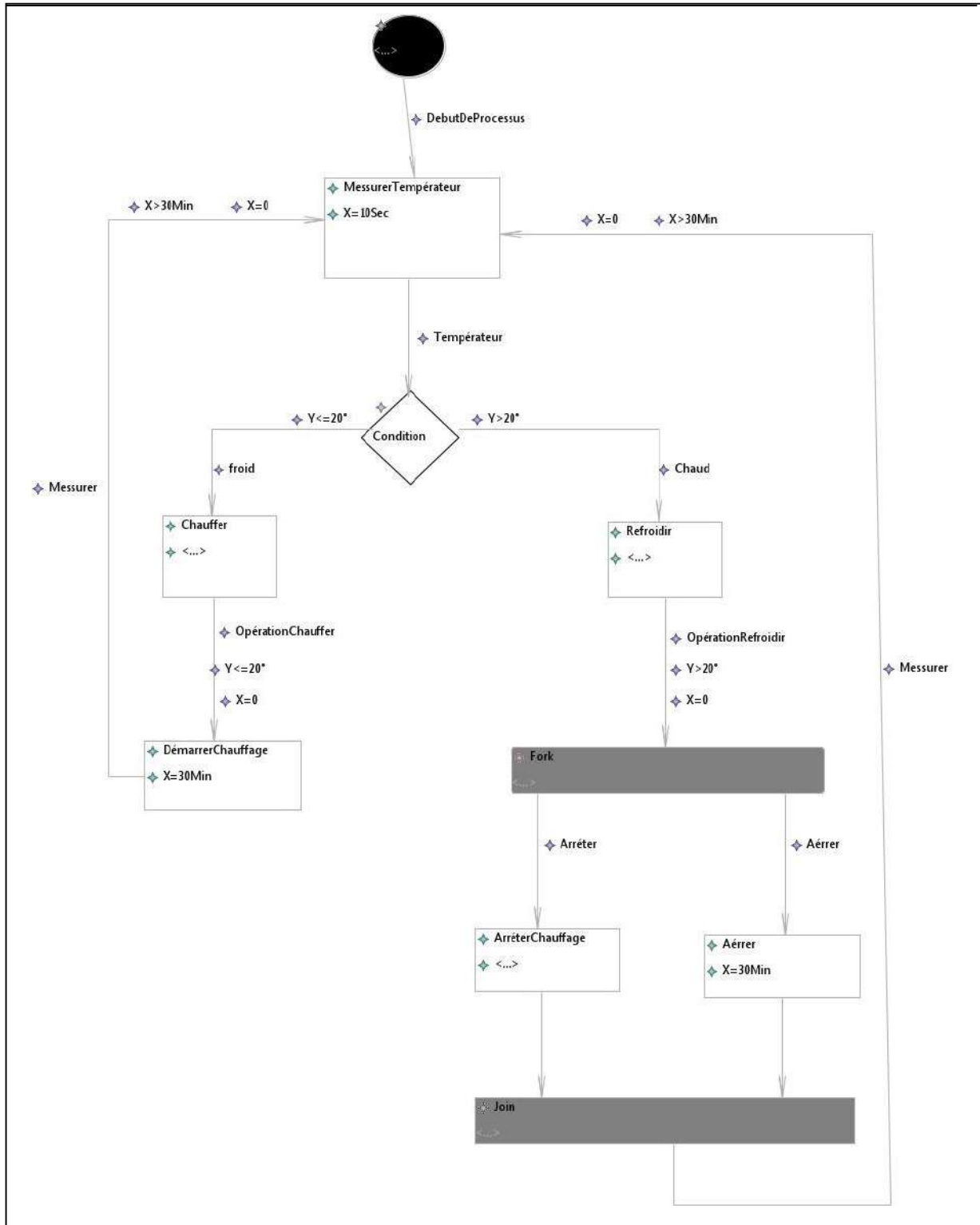


Fig. 4. Activity Diagram model

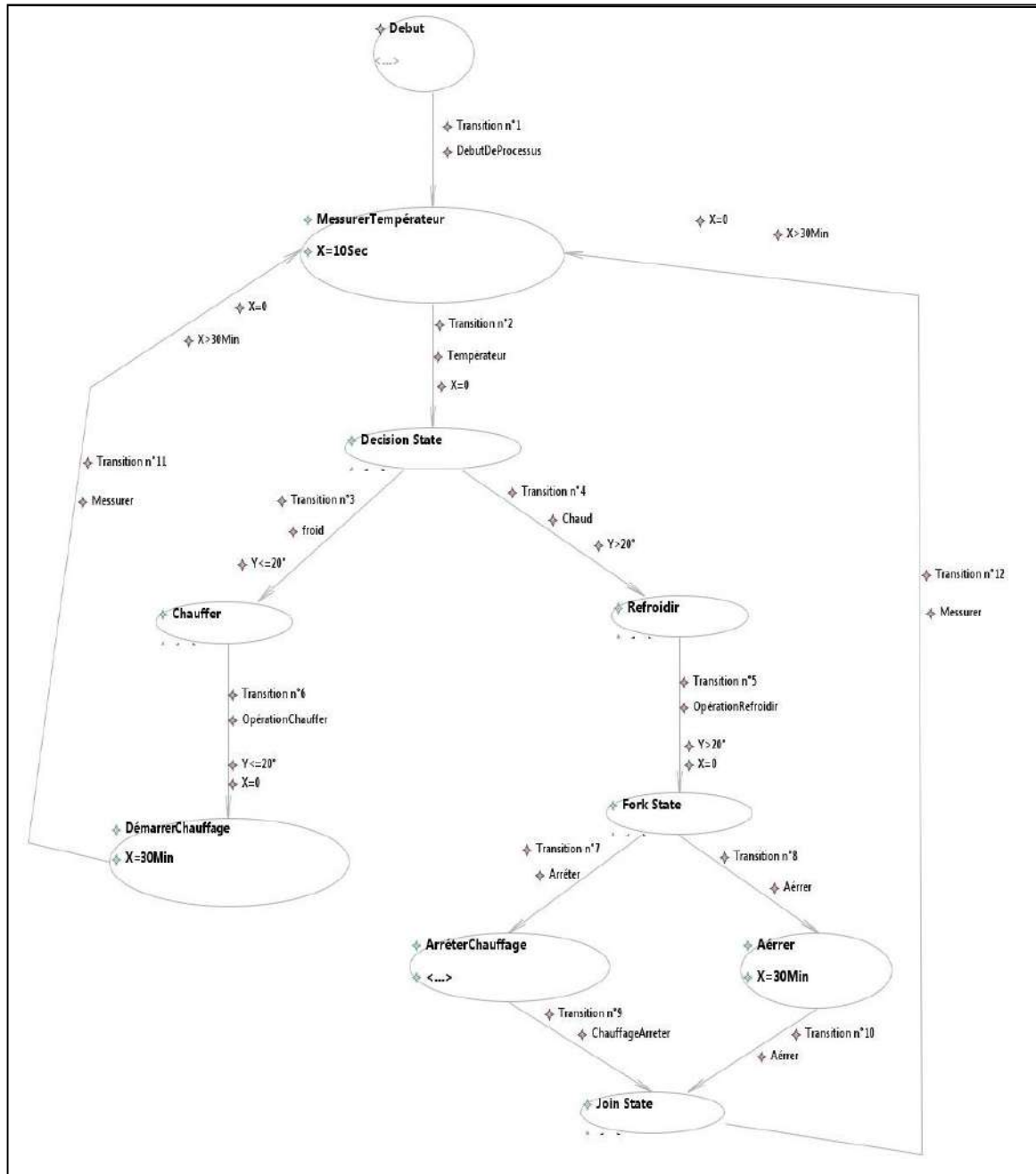


Fig. 5. Timed Automata model

VI. CONCLUSION

In this paper, we have proposed an approach based on combining Meta-modelling and Graph Grammars to automatically generate a visual modelling tool for Activity Diagram and Timed Automata.

The cost of building a visual modelling tool from scratch is prohibitive. Meta-Modelling approach is useful to deal with this problem since it allows the modelling of the formalisms themselves. By means of graph grammars, models manipulations are expressed on a formal basis and in a graphical way.

In this approach, two meta-models are proposed; one for the Activity Diagram model and the second for the Timed Automata model. Based on these meta-models, we have proposed a graph grammar that deals with the transformation process. The EMF (Eclipse Modeling Framework), the GMF

(Graphical Modeling Framework) and ATL (Atlas Transformation language) are used for this purpose.

We have illustrated the proposed approach through an example. In a future work we plan to perform this transformation using graph transformation visual tools like AGG, GROOVE and transform other UML diagrams.

REFERENCES

- [1] B. Agarwal, "Transformation of UML Activity Diagrams into Petri Nets for Verification Purposes", International Journal Of Engineering And Computer Science, ISSN: 2319-7242, Volume 2 Issue 3 March 2013, pp. 798-805, 2013.
- [2] R. Alur, and D. L. Dill, "A theory of timed automata", Theoretical Computer Science (TCS'94), 126, 183–235, 1994.
- [3] L. Baresi, and R. Hekel, "Tutorial Introduction to graph transformation. A software Engineering perspective", Lecture Notes in Computer Science, Vol 3256/2004, Springer Berlin 2004, pp.431-433, 2004.

- [4] A. Belghiat, and A. Chaoui, "A Graph Transformation of Activity Diagrams into Pi-calculus for Verification Purpose", Proceedings of the 3rd Edition of the International Conference on Advanced Aspects of Software Engineering (ICAASE18). Constantine, Algeria, 1,2-December-2018, published at <http://ceur-ws.org>
- [5] R. Elmansouri, S. Meghzili, and A. Chaoui, "A Uml 2.0 Activity Diagrams/Csp Integrated Approach For Modeling And Verification Of Software Systems", Computer Science 22(2) 2021, pp. 209–233, 2021.
- [6] EMF home page, <http://www.eclipse.org/emf/>
- [7] GMF home page, <http://www.eclipse.org/gmf>
- [8] R. Heckel, and G. Taentzer, "Graph Transformation Concepts", In: Graph Transformation for Software Engineers. Springer, Cham, 2020.
- [9] H.T. Jung, and S. H. Joo, "Transformation of an activity model into a colored petri net model", in Trendz in Information Sciences & Computing (TISC), 2010. IEEE, pp. 32–37, 2010.
- [10] G. Karsai, and A. Agrawal, "Graph Transformations in OMG's Model-Driven Architecture", Lecture Notes in Computer Science, vol 3062, Springer Berlin , Heidelberg 2004, pp. 243-259, 2004.
- [11] E. Kerkouche, K. Khalfaoui, and A. Chaoui, "A rewriting logic-based semantics and analysis of UML activity diagrams: a graph transformation approach", Int. J. Computer Aided Engineering and Technology, Vol. 12, No. 2, pp. 237–262, 2020.
- [12] A. V. Markov, A. A. Voevoda, and D. O. Romannikov, "Automatic translation UML activity diagrams to Petri net", 2015 International Siberian Conference on Control and Communications (SIBCON). Electronic ISBN: 978-1-4799-7103-9, 2015.
- [13] S. Meghzili, A. Chaoui, M. Strecker, and E. Kerkouche, "An Approach for the Transformation and Verification of BPMN Models to Colored Petri Nets Models", International Journal of Software Innovation, Vol. 8(1), pp. 17–49, 2020
- [14] O. OMG, "Unified Modeling Language Specification". Object Management Group, 2015
- [15] O. OMG, "Systems modeling language (sysml)", OMG Available Specification, 2007.
- [16] Y. Rahmoune, A. Chaoui, and E. Kerkouche, "A Framework for Modeling and Analysis UML Activity Diagram using Graph Transformation", International Workshop on the Use of Formal Methods in Future Communication Networks (UFMFCN 2015), pp. 612 – 617, 2015.

Applying the ESQuMo Quality Model to Medical Imaging Embedded Software

Zouheyr Tamrabet

Department of Mathematics and Computer Science,
University of Oum El Bouaghi, Algeria
tamrabet.zouheyr@yahoo.fr

Toufik Marir

Department of Mathematics and Computer Science,
University of Oum El Bouaghi, Algeria
marir.toufik@yahoo.fr

Farid Mokhati

Department of Mathematics and Computer Science,
University of Oum El Bouaghi, Algeria
mokhati@yahoo.fr

Abstract— Embedded systems are electronic computer systems dedicated to specific tasks. The increasing complexity of these systems imposes the adoption of sophisticated techniques and tools such as Quality Models, in order to guarantee their quality. However, quality models still suffer from several shortcomings, the most important of which is the inability to pass from the quality definition provided by the quality model to other quality purposes such as quality assessment and quality prediction unless we use another separate model that serves for that. The main contribution of this paper is the proposition of quality measures in the context of Medical Imaging Embedded Software based on ESQuMo, an Embedded Software Quality Model.

Keywords— *Embedded Software, Medical Imaging Systems, ESQuMo, Quality Models, Quality Measures, ISO/IEC 25010, ISO/IEC 25023.*

I. INTRODUCTION

Embedded systems are systems designed to solve specific problems or tasks, but they are not general-purpose computers. Usually, an embedded system is the combination of software and hardware side by side to form an electronic computer system dedicated to a specific task [1]. Recently, these systems are widely used, and this is due to their effectiveness in many fields, such as video games, signal processing, telecommunication, etc. However, these systems are getting more complex than before as they submit to temporal, resources, and environmental constraints [2] [3].

The increasing complexity of these systems needs a sophisticated development process to guarantee their quality, for this reason, quality assurance has taken place. Quality assurance is a systematic process that includes all the techniques and tools that should be followed in order to develop the intended systems. Among these techniques and tools, Quality Models are so far the most promising way in order to improve quality [4].

Although quality models are among the most successful tools of quality assurance, quality models still suffer from several shortcomings especially when it comes to software development. First, the quality models in the context of the software found in the literature do not conform to an explicit meta-model [5]. In addition to their inability to cover both the features of general software and those related to the specific type of software. Moreover, the essential objectives in quality assurance from the definition, assessment to prediction, are usually carried out through separate independent models that there is no means of coordination between them. Therefore, standard quality models are often used that give greater control and cover different quality perspectives all the way through quality assurance stages.

ESQuMo (Embedded Software Quality Model) is a good example of models based on standard quality models [6]. In addition to its definition of what quality is for the embedded software, this model helps to unify and cover the common characteristics of the ordinary software as well as the

relevant characteristics of the embedded software at the same time.

The main contribution of this paper is the application of ESQuMo by the proposition of specific quality measures related to the Medical Imaging Embedded Software. Note that these quality measures are conforming to the international standard ISO/IEC 25023 that works in conjunction with other international standards such as ISO/IEC 25010.

The remainder of this paper is organized as follows: Section 2 introduces the Medical Imaging Embedded Systems as a specific kind of these systems. A brief overview of ESQuMo an Embedded Software Quality Model is presented in section 3. Section 4 presents the set of the quality measures proposed in the context of Medical Imaging Embedded Software, along with their detailed definitions as well as their measurement functions. Conclusion and future works are provided in section 5.

II. MEDICAL IMAGING EMBEDDED SYSTEMS

Nowadays, we do rely more and more on embedded systems compared to general-purpose systems. The reason behind this phenomenon is due to the huge amount of these systems in our daily life on the one hand, and the increased complexity of the present operations on the other hand [3]. Therefore, it is necessary to depend on these systems to gain time and effort while confronting these operations.

Embedded systems can be found in a wide range of fields. For instance, the medical field is littered with many examples such as Defibrillators, Oximeters, Fetal monitors...etc. [7]. These medical devices made the lives of people related to the medical field much easier as well as their profession more efficient. Medical imaging devices are no less important than the aforementioned devices, as medical imaging is so sensitive, that it requires the use of new imaging technologies and relies more on these kinds of systems [3] [7].

Medical imaging devices are considered among the most commonly used embedded systems since they are able to see inside a patient's body without actually having to cut them open. This is extremely important as it allows doctors to identify any abnormalities and it helps to decide the right medical diagnosis as well. However, the real challenge with these types of systems is to ensure that all the provided results are dependable and can be relied on [7]. In addition, the images taken in the medical imaging process must be not only accurate but also in real-time. For example, in ultrasound imaging, any delay in the process of displaying images affects the diagnosis process and may completely conceal the real issue to be imaged in the first place [8].

Furthermore, medical imaging devices are among the embedded devices that necessitate prior knowledge before performing any imaging process. This is because medical imaging devices require certain procedures during imaging that if not followed would completely disrupt the imaging, such as the patient's continuous movement or carrying any type of metal that would perturb the imaging process and at times jeopardize the patient's safety, particularly when using CT or MRI. Therefore, the choice of the right medical imaging modality is related to several points such as the patient's age, the type of injury, the organ or tissue to be imaged, the cost of imaging in addition to the amount of radiation that is sometimes present in the medical imaging process. In the following, we will present the most popular modalities of medical imaging [9].

A. X-ray

An x-ray is an image of a specific area of the body, such as the chest, abdomen, or a specific joint. With an x-ray, highly charged electrons are shouted through a specific area of the body, the image is then formed according to the scatter of these electrons. Thus, high-density objects will appear white because electrons scatter everywhere against them, and low-density objects like air will appear black since there is no scatter at all. As a result, everything in between is going to be some shade of grey. For the radiation of the x-ray, there is some radiation. However, it is very low and no need to worry about it, and in terms of cost, x-rays are among the very cheap tests.

B. CT scan

A CT scan is a little bit the same technology as an x-ray, in which high-energy electrons are shouted through the body, except for taking cross-sectional images in the case of a CT scan. The images are taken from a 360-degree view around the patient and this creates a slice-by-slice, a layer-by-layer map of the patient. This technology allows for visualizing bones, soft tissues, and vessels all at once. CT scan is a very powerful imaging technique. However, it is not without issues. It is a much more radiation-intensive process since it is a series of cross-sectional x-rays. compared to an x-ray, a CT scan is probably three or four times as intense in terms of radiation exposure, and because it is more sophisticated, it costs a little bit more.

C. Ultrasound

The ultrasound is a very useful imaging modality, which contains much interesting physics behind it. The main idea is to take high-frequency sound waves and bounce them all over the tissues, and the way that they echo back to the ultrasound probe is what produces the images used for the density assessment of the tissues. It is a great, easy and effective imaging modality for visualizing body cavities, fluids, tissues, and blood vessels. As sound waves are used, no radiation exposure adds to the usability of an ultrasound. In addition, it is a very cheap test compared to a CT scan or MRI, similar to probably the cost of an x-ray but there is no radiation cost so ultrasound is much more frequently used.

D. MRI

The MRI is also a common imaging modality but it is a little bit different from the other imaging modalities. Instead of using radiation or sound waves, the MRI uses magnetic

fields. Two perpendicular magnetic fields that are acting in opposite directions to each other are used to manipulate hydrogen atoms that are spinning in the body. These atoms release signals that are captured, and then used to determine their location in the body, and display it on a gradient of gray colors indicating the strength of the signal, as well as the different tissues of the body (the brain, the spinal cord, the liver, the pancreas, ... etc.).

Typically, an MRI is used to visualize the tissues of the body in order to detect tissue irregularities or tumors, and because it uses magnets, there is obviously no radiation exposure which is greatly beneficial, especially in the pediatric population. Rather than exposing children to CT scans where they are going to get a significant dose of radiation, the images could be friendly obtained using an MRI. Unfortunately, this magnificent imaging modality comes with a cost, it is definitely the most expensive of the four basic imaging modalities.

In the next section, the ESQuMo quality model is presented, which will be adopted in the rest of our study.

III. ESQuMo: AN EMBEDDED SOFTWARE QUALITY MODEL

Quality models are generally used to define quality for specific products. This definition is of prime importance because it constitutes the basis of quality assessment and prediction [5]. The Embedded Software Quality Model proposed by [6] is one of these definition quality models. The main contribution of ESQuMo is it unifies and covers the common characteristics of the ordinary software through the Functionality characteristic. In addition to the relevant characteristics of the embedded software, which are presented through *Complexity* and *Dependability* as depicted in Fig. 1. Besides, ESQuMo is based on the famous and well-known ISO/IEC 25010 standard quality model. The choice of this latter is justified by its belonging to the SQuaRE series of international standards, which includes other standards that are complementary to each other [10]. This is quite useful, because the actual model not only provides the definition of quality but also offers the possibility to achieve other objectives such as quality assessment and quality prediction through these complementary standards with the maximum coherence possible [11].

The ISO/IEC 25023 is one of these standards [12]. It provides a set of quality measures for the characteristics of system/software products that can be used for specifying requirements, measuring and evaluating the system/software product quality, in conjunction with other standards such as ISO/IEC 25010 [12]. In other words, these quality measures are meant for evaluating system/software product quality in terms of characteristics and sub characteristics defined in ISO/IEC 25010. The special thing about the quality measures in this international standard is they are not intended to be exhaustive and users are very encouraged to refine them if necessary [12].

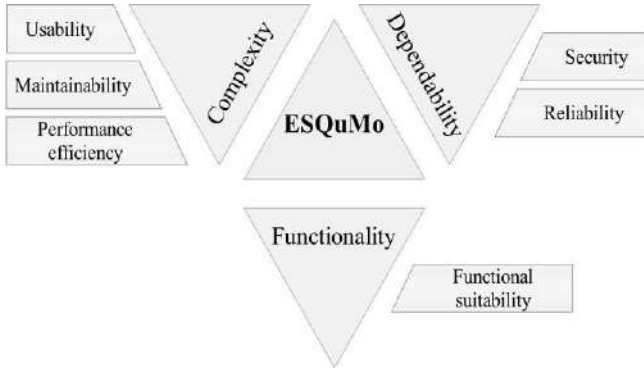


Fig. 1. ESQuMo an Embedded Software Quality Model [6].

Before proceeding to provide quality measurements of the medical imaging embedded software, it is important to note that the SQuaRE provide a general reference model as shown in Fig. 2, to facilitate the navigation through the different complementary standards [11].

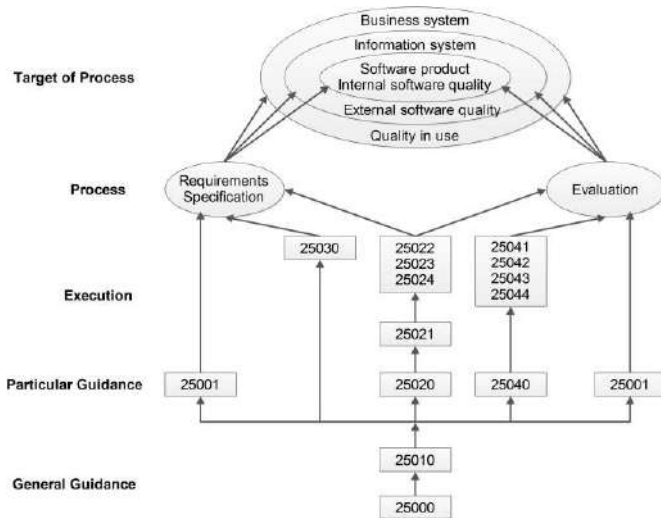


Fig. 2. SQuaRE general reference model [11].

Through the SQuaRE general reference model [11], it seems necessary to review the ISO/IEC 25020 Measurement reference model and guide straightaway after obtaining the definition of quality through the standard quality model ISO/IEC 25010. Note that the ISO/IEC 25020 includes ISO/IEC 25021, ISO/IEC 25022, ISO/IEC 25023, and ISO/IEC 25024 respectively for Quality measure elements, Measurements of internal quality, Measurements of external quality, and Measurements of quality in use. In this paper, we are concerned only with the ISO/IEC 25023 Measurements of external quality, where the software is evaluated and measured according to an external point of view. In the next section, quality sub-characteristics and quality measures related to the medical imaging embedded software are presented along with their detailed definitions, as well as their evaluation methods.

IV. QUALITY MEASUREMENTS OF MEDICAL IMAGING EMBEDDED SOFTWARE

The measuring process is very important as it helps to control and supervise the overall software quality. In this process, before proposing a set of quality measurements for each of the identified characteristics in the Embedded Software Quality Model (ESQuMo), it is important to break

down these characteristics into measurable sub-characteristics. Moreover, as stressed in the SQuaRE main guide [11]; the quality model can be used as a guideline or checklist where the selection of the concerned characteristics is possible. From this point of view, the proposition of quality measurements for all the sub characteristics is not mandatory. In addition, since the Embedded Software Quality Model is based on the ISO/IEC 25010 standard quality model, three distinguished cases are possible for the sub-characteristics:

- Specific sub-characteristics are proposed for medical imaging embedded software, as well as their subsequent measures.
- Some sub-characteristics can be drawn from the ISO/IEC 25010 quality model, extended with new defined measures in order to support the medical imaging embedded software.
- The remaining sub-characteristics of the ISO/IEC 25010 quality model can be overlooked.

A. Specific Sub-characteristics of Medical Imaging Embedded Software

1) Cost of Use

a) *Cost of Use*: “Degree to which a system costs while performing its operations”.

Price has always been an important indicator in any type of product. As we have already mentioned, the modality of medical imaging controls the cost of the imaging process. This sub-characteristic is part of the Usability characteristic, where it can be measured qualitatively as follows:

$$\text{CoU} \in \{\text{Cheap, Economic, Expensive, Extremely Expensive}\} \quad (1)$$

2) Vulnerability of Data

a) *Vulnerability of Data*: “Degree to which a system ensures that data is not influenced by patient factors”.

The medical imaging process may be affected by factors that the patient often unintentionally commits. For example, movement during the imaging process or wearing metallic jewelry may affect the results of imaging and make it vulnerable. This sub-characteristic is part of the Security characteristic, and it is measured as follows:

$$\text{VoD} = \text{the probability of patient factors occurrence.} \quad (2)$$

B. Extended ISO/IEC 25010 Sub-Characteristics

1) Time Behaviour

Medical imaging devices are among the real-time embedded systems. Thus, the time aspect is the most susceptible feature for these kinds of devices. A real-time embedded system is defined as a system in which the time at each produced output is significant. Hence, the lag from input time to output time must be sufficiently small. Therefore, in the context of embedded systems, it is not only the outcome of the output that matters but also the time of its occurrence. From this point of view, the Time Behaviour sub-characteristic is extended with new measures of Timeliness and Responsiveness.

a) *Timeliness*: “is the quality of being done or occurring at a favorable or useful time”.

$$T = 1, \text{ if Response Time} \leq \text{Processing Time} + \text{Latency}; \\ 0, \text{ else} \quad (3)$$

-Latency is the delay incurred in communicating a message.

-Processing time is the amount of time a system takes to process a given request.

-Response time is the total time it takes from when users make a request until they receive a response.

b) *Responsiveness*: “is the ability to detect change of system’s response over time”.

Responsiveness is often expressed as a Standardized Response Mean [13]. In order to calculate the Standardized Response Mean, first, we have to make sure that we already have two groups of response times, and then we apply the next formula, where:

$$SRM = (M1 - M2) / SD \quad (4)$$

-M1 is the mean of the first group.

-M2 is the mean of the second group.

-SD is the standard deviation of the population from which the groups were sampled.

After the SRM index is calculated, we use the Cohen’s benchmarks [13] for better interpretation where:

-Small change: $SRM \leq 0.2$

-Moderate change: $2 < SRM < 0.8$

-Large change: $SRM \geq 0.8$

2) Resource Utilization

Medical imaging devices are among the most energy-consuming embedded systems, and this is mainly due to their use of GPUs as well as CPUs to meet the requirements of the system along with increasing the level of performance. However, this performance comes with a cost. Since the amount of resources used is enormous, it is natural to have a massive energy consumption. This side effect is originally a result of the extensive use of processors in addition to excessive access to the memory. Hence, it is necessary to determine the level of energy consumption of the embedded software. In order to evaluate this latter, the software energy consumption measure is proposed as an extension of the Resource utilization sub-characteristic.

a) *Software Energy Consumption*: “is the amount of energy consumed by the software, given the amount of resources used from processors and memory access”.

$$SEC = EIC \times MAC \quad (5)$$

-Executed Instruction Count (EIC), which corresponds to the number of executed instructions by the processing unit [14].

-Memory Access Count (MAC), which is equal to the number of memory accesses [15].

3) Maturity

The determinism of medical devices is of prime importance because it represents the system’s next behavior. The behavior of a non-deterministic system can never be predicted and is always questionable. From this point of view, we do believe that determinism influences the Maturity of the overall system.

a) *Determinism*: “A deterministic system is a system in which a given initial state or condition will always produce the same results”.

In the following section, we adopt the cyclomatic number as a measure of determinism for medical imaging embedded software:

$$CN = E - N + 2P, \text{ where:} \quad (6)$$

-E = the number of edges of the graph,

-N = the number of nodes of the graph,

-P = the number of connected components.

The cyclomatic number is a famous software measure used to indicate the number of linearly independent paths through a program [16]. This measure is worth at least “1” since there is always at least one path. Mathematically, a cyclomatic number that is too high indicates that the number of decision points of the system is high, which potentially translates to several behaviors of the very same system. Theoretically, a low cyclomatic number indicates that the system is well deterministic.

4) Fault Tolerance

Robustness is considered a special case of Fault tolerance. Since Fault tolerance corresponds to the ability of the system to operate as intended despite the presence of faults, the use of invalid inputs similarly must be tolerated and continue to work normally. The definition of this measure is provided in the following.

a) *Robustness*: “The degree to which a system continues to function in the presence of invalid inputs”. where it can be measured as follows:

$$R = \text{Number of invalid inputs} / \text{Number of stops} \quad (7)$$

V. FORMAT OF THE PROPOSED QUALITY MEASURES

For better conformance during the measuring process, and in order to maintain a unified format with the rest of the quality measurements predefined in the international standard ISO/IEC 25023, we must support the following information [12]:

1) ID: identification code of quality measure; each ID consists of the following three parts:

a) abbreviated alphabetic code representing the quality characteristics as capital X and sub characteristics as one capital followed by lowercase x;

b) serial number of sequential order within quality sub characteristic;

c) G (Generic) or S (Specific) expressing potential categories of quality measure; where, generic measures can be used whenever appropriate and specific measures could be used when relevant in a particular situation;

2) Name: quality measure name;

3) Description: the information provided by the quality measure;

4) Measurement function: mathematical formula showing how the quality measure elements are combined to produce the quality measure.

Table 1 summarizes all of the previously mentioned characteristics, sub-characteristics and measurements in the context of medical imaging embedded software.

VI. CONCLUSIONS AND FUTURE WORKS

Embedded systems are increasingly used in our daily life and this is due to their extreme importance. Thus, these systems should deliver their attended services with quiet quality since their outcome could be critical in some situations, the medical imaging embedded systems as an example. In order to control the quality of such systems, quality models are very promising tools to do that for they provide a clear definition of the software quality of these systems through characteristics and sub characteristics. However, most of the quality models in the literature do not cover the relevant characteristics of the embedded software in addition to the general characteristics of the ordinary software. Moreover, even if it does, there is no means to pass from the quality definition provided by the model to other quality purposes such as evaluation and prediction unless we use another separate model that serves for that. For this reason, ESQuMo is adopted as a quality model that overcomes the previous shortcomings. In this paper, we have proposed a set of quality measures for the Medical Imaging Embedded Software based on ESQuMo (Embedded Software Quality Model). It is important to note that these quality measures are conforming to the international standard ISO/IEC 25023 that works in conjunction with other international standards such as ISO/IEC 25010. In future work, we intend to extend this contribution by proposing a new quality model for the purpose of prediction in order to estimate quality even before the development of the embedded software.

ACKNOWLEDGMENT

This research work was supported by the General Direction of Scientific Research and Technological Development (DGRSDT) of the Algerian Higher Education and Scientific Research Ministry. We would like to thank the DGRSDT for its support in the achievement of this work.

REFERENCES

- [1] Guessi, M., Nakagawa, E. Y., Oquendo, F., & Maldonado, J. C. (2012, June). Architectural description of embedded systems: a systematic review. In *Proceedings of the 3rd international ACM SIGSOFT symposium on Architecting Critical Systems* (pp. 31-40). ACM.
- [2] Garcés, L., Ampatzoglou, A., Avgeriou, P., & Nakagawa, E. Y. (2017). Quality attributes and quality models for ambient assisted living software systems: A systematic mapping. *Information and Software Technology*, 82, 121-138.
- [3] Lee, E. A., & Seshia, S. A. (2016). *Introduction to embedded systems: A cyber-physical systems approach*. MIT Press.
- [4] Galin, D. (2004). *Software quality assurance: from theory to implementation*. Pearson education.
- [5] Wagner, S. (2013). *Software product quality control*. Berlin: Springer.
- [6] Tamrabet, Z., Marir, T., & Mokhati, F. (2022). ESQuMo: An Embedded Software Quality Model. *International Journal of Embedded and Real-Time Communication Systems (IJERTCS)*, 13(1), 1-18.
- [7] Dere, G. (2021). *Biomedical Applications with Using Embedded Systems*. In *Data Acquisition-Recent Advances and Applications in Biomedical Engineering*. IntechOpen.
- [8] Niederhauser, J. J., Jaeger, M., Lemor, R., Weber, P., & Frenz, M. (2005). Combined ultrasound and optoacoustic system for real-time high-contrast vascular imaging in vivo. *IEEE transactions on medical imaging*, 24(4), 436-440.
- [9] Ahmad, H. A., Yu, H. J., & Miller, C. G. (2014). *Medical imaging modalities*. In *Medical imaging in clinical trials* (pp. 3-26). Springer, London.
- [10] International Organization for Standardization. (2011). *ISO/IEC 25010: Systems and software engineering – systems and software quality requirements and evaluation (SQuaRE) – system and software quality models*. Geneva, Switzerland.
- [11] International Organization for Standardization. (2005). *ISO/IEC 25000: Systems and software engineering – systems and software quality requirements and evaluation (SQuaRE) – Guide to SQuaRE*. Geneva, Switzerland.
- [12] International Organization for Standardization. (2016). *ISO/IEC 25023: Systems and software engineering – systems and software quality requirements and evaluation (SQuaRE) – Measurement of system and software product quality*.
- [13] Husted, J. A., Cook, R. J., Farewell, V. T., & Gladman, D. D. (2000). Methods for assessing responsiveness: a critical review and recommendations. *Journal of clinical epidemiology*, 53(5), 459-468.
- [14] Furber, S. B. (2000). *ARM system-on-chip architecture*. pearson Education.
- [15] Chatzigeorgiou, A., & Stephanides, G. (2002). Energy metric for software systems. *Software Quality Journal*, 10(4), 355-371.
- [16] McCabe, T. J. (1976). A complexity measure. *IEEE Transactions on software Engineering*, (4), 308-320.

TABLE I. QUALITY MEASURES OF MEDICAL IMAGING EMBEDDED SOFTWARE.

ID	Characteristic	Sub-Characteristic	Name	Description	Measurement function
UCu-1-S	Usability	Cost of Use	Cost of Use	Degree to which a system costs while performing its operations	Equation (1)
SVu-1-S	Security	Vulnerability of Data	Vulnerability of Data	Degree to which a system ensures that data is not influenced by patient factors	Equation (2)
PTb-5-S	Performance efficiency	Time Behaviour	Timeliness	the quality of being done or occurring at a favorable or useful time	Equation (3)
PTb-6-S	Performance efficiency	Time Behaviour	Responsiveness	the ability to detect change of system's response over time	Equation (4)
PRu-5-S	Performance efficiency	Resource utilization	Software Energy Consumption	the amount of energy consumed by the software, given the amount of resources used from processors and memory access	Equation (5)
RMa-5-S	Reliability	Maturity	Determinism	A deterministic system is a system in which a given initial state or condition will always produce the same results	Equation (6)
RFt-4-S	Reliability	Fault tolerance	Robustness	The degree to which a system continues to function in the presence of invalid inputs	Equation (7)

Assessing collaborative learning using collaborative quizzes

1st Nouredine Gouasmi
Labstic
08 Mai 45 University
Guelma, Algeria

2nd Mahnane Lamia
LRS Laboratory
University of Badji Mokhtar
Annaba, Algeria

3rd Yassine Lafifi
Labstic
08 Mai 45 University
Guelma, Algeria

4th Amira Lebbad
Computer science department
08 Mai 45 University
Guelma, Algeria

gouasmi.nouredine@univ-guelma.dz mahnane_lamia@yahoo.fr lafifi.yacine@univ-guelma.dz lebbad.amira@gmail.com

Abstract—The covid-19 pandemic has demonstrated the importance of e-learning during lock down. But collaboration between learners remains an important element in learning, hence the need of collaborative learning that offers learning activities carried out by a group of learners oriented towards the same goal. But one of the most important questions claimed by collaborative learning is to assess the quality of collaboration for each learner. In this paper, we present a model, derived from software qualimetry, for assessing learner's collaboration during a collaborative activity, called collaborative quiz. The model allows to calculate a *collaboration* factor from four metrics. **Index Terms**—collaborative e-learning, qualimetry, collaboration factor, communication, metrics

I. INTRODUCTION

collaborative learning is an important part in teaching strategies. It is very effective in improving the learning performance of students in e-Learning or in the classroom, and brings many benefits to students. It aims to allow learners to develop social skills, to share and exchange ideas, in addition to the acquisition of useful skills.

In collaborative learning, learners work together in a group to achieve a common goal: *the learning goal*. This learning strategy enables students to acquire new knowledge and skills through social interaction with peers. In addition, obtaining good learning issue is allowed by the presence of well-structured groups and good interactions between peers [1].

The main advantage of collaborative learning is that it encourages students to help each other in solving problems. This grants learners with cooperation and communication skills, and also, the possibility to share resources, experience and ideas [2]. Another benefit of collaborative learning is that it helps students develop the ability to communicate and share ideas clearly [3]. In the other hand, teamwork helps to find creative solutions that reflect the general views of students [4], and stimulate students to have a positive attitude towards their teammates and to establish cohesive group relationships [5].

In collaborative learning the composition of the learning group is the most important factor in the success of the collaboration, in particular the individualities constituting a group. Therefore a good grouping strategy is important because integrating elements into a team implies cooperation to achieve goals [6].

Group formation is a complex process depending on different factors. These factors are conditioned by characteristics such as group members, the context of formation or the technique used to form groups [7]. According to Bekele [8], the process of forming a group lay on three challenges: determine the size of the group based on the learning objectives, choose which learner to assign to which group depending on the nature of the group formation, and which nature adopt between homogenous or heterogeneous grouping.

The process of grouping learners depends on a number of characteristics describing the learner's profile that are categorized into classes, such as communication (e.g. number of posts), activities (e.g. grade on activities), psychology (e.g. sociability), etc. [9].

while these characteristics, mainly those related to collaboration, are used to form groups of learners (for example [10], [11], [12]), they are not used for evaluating collaborative works done by learners. The evaluation of learners depends only on the result of the collaborative activity.

Our objective is to evaluate a collaborative activity done by a group of learners. The activity considered is collaborative quizzes, and the evaluation takes into account the process of collaboration between the learners during the activity alongside with their grades.

The paper is organised in three main parts. The first one introduces the qualimetry domain and the software quality evaluation. The second part presents the proposed qualimetry model for assessing learner's collaborative work within collaborative quizzes. An experimentation is described in the last part, by describing the methodology used in this work, and after, by showing results of data analysis.

II. QUALIMETRY

The term *qualimetry* was originally applied to a scientific discipline studying the methodology and problems of quantitatively assessing the quality of various objects, mainly industrial products [13]. Its main goal is to generalize the quantification of quality to any domain and type of object [14].

A. Software qualimetry

There are many aspects that must be taken into account when creating a software. One of them is its quality. The

purpose of software quality testing is to find out whether the software implementation is as needed [15].

There have been two notable models of software quality: *Boehm model* and *McCall model*. Boehm and McCall both described quality using a decomposition approach.

Boehm's (1978) model is based on a wide range of characteristics and incorporates 19 criteria. These criteria are not independent, they interact with each other [16].

McCall's Software Quality Model (The GE Model, 1977) [17] includes 11 criteria encompassing product operation, product review, and product transition. This model is presented below.

B. McCall's software quality model

The quality of software is represented through a hierarchy of three levels as shown in figure 1.

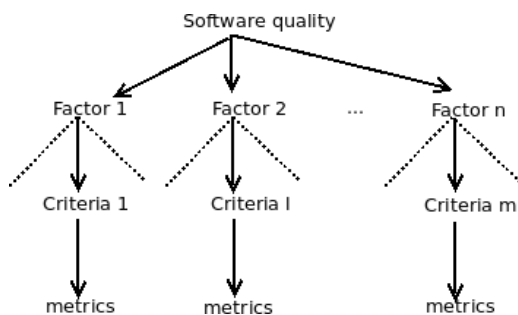


Fig. 1. McCall's model

At the top level are the *quality factors* from a customer or user point of view. These factors are: accuracy, reliability, efficiency, integrity, ease of use, maintainability, testability, flexibility, portability, reusability and interoperability. At the second level are the *quality criteria* which represent the translation of factors into technical concepts. At the third level are the *quality metrics*, which measure the attributes of software products [16].

McCall suggests the following steps for applying the quality pattern [16]:

- 1) Deduce quality factors according to the characteristics of the system.
- 2) Choose the quality factors according to the needs of users.
- 3) Deduce quality criteria and measures.
- 4) Apply metrics at the specification, design, coding and testing stages of the software lifecycle.

The path from the end user to the developer is made up of three levels: the factors represent the needs, each of them is broken down into criteria which are in turn combinations of [18] metrics.

The *factors* and *criteria* are qualitative characteristics of the software. The first is seen from the user's point of view and the second is handled by the developer. The criteria represent constraints unknown to the end user (e.g. software that is

easy to maintain for the user means for the developer modular programs that are easy to test).

The *metrics* are measures that can be obtained from the source code (e.g. the total number of operators contained in a program).

A *normalization function* will allow a quality value to be assigned to a factor from the measured metrics.

III. LEARNER ASSESSMENT USING COLLABORATIVE QUIZZES

The main goal of the developed system is to evaluate learners during collaborative work which consists of answering quizzes.

The system offers to learners two essential features:

- 1) Answer quizzes in a collaborative way.
- 2) Communicate with the elements of its group, during work, through instant messaging.

It also offers to both learners and teachers the possibility to communicate, in a social media, with posts.

The evaluation of learners in a group will be done through the answers to quizzes and their collaboration during work.

A. System design

The system consists of two essential parts (figure 2). The first part is the posts features management system. The second part is for the management of quizzes features.

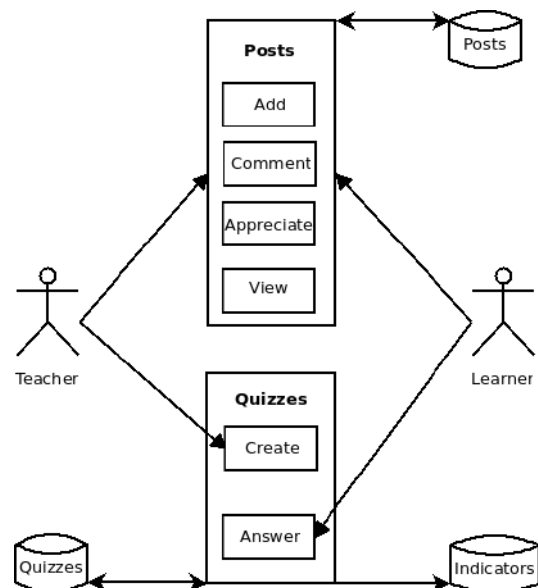


Fig. 2. System functional architecture

To evaluate learners, we use collaborative quizzes. A collaborative quiz consists of a web page composed of two sections:

- 1) A *quiz* section containing the statements of the quiz to which the learner must answer.
- 2) A *collaboration* section consisting of a messaging system through which he can discuss, with the members of his group, while answering the quiz.

A quiz page example is shown in figure 3, and some other pages of the system are presented in figures 4, 5, and ??.

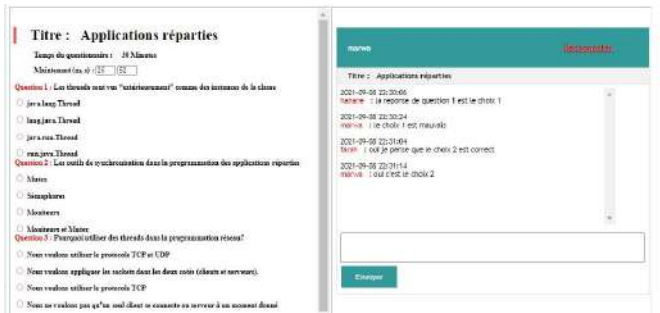


Fig. 3. Quiz page



Fig. 4. Proposed quizzes (Teacher's page)



Fig. 5. Proposed quizzes (Learner's page)

After the response has been validated, the system evaluates the score to be assigned to each member of the group who responded to the quiz. The evaluation is done by applying a qualimetry model.

B. Design of a qualimetry model for the evaluation of answers to collaborative quizzes

Our goal is to assess the level of learner involvement in the collaboration to respond to a quiz in a collaborative way, by observing the interactions of learners during collaborative work.

In a context of classic use of the system, a teacher offers a questionnaire to a group of learners. The latter must answer it at the same time, and can communicate together, through an

instant messaging system integrated into the questionnaire, to choose the correct answer.

The evaluation of the questionnaire focuses on the quality of the collaboration between the members of the group, in particular by answering questions such as: do they collaborate to answer the quiz? do they communicate with each other? etc.

The three levels of the proposed model are as follows:

1) Factor:

In our system, the item to be evaluated is: *Collaboration between learners*.

2) Criteria:

To assess the above factor, we identified four criteria: the *presence of learner* during the test, the *level of communication* of learner with the other members of the group, the *sentiment* of learner for the group, and his *knowledge* evaluated by his score on the quiz.

3) Metrics:

The above criteria are measured as follows:

- The **presence** of the learner is evaluated by comparing the connection times to the quiz. The connection time of the majority of learners of the group is considered the connection time reference. If a learner is logged in at this reference time, then its presence equals 1, otherwise, it equals 0.
- The **communication level** is evaluated from the number of messages exchanged in the group during the test period. The more messages the learner sends, compared to the total number of messages sent, the higher his level of communication. The value of the metric is in the range [0, 1].
- We assessed the learner's **sentiment** for the group using the AFINN [19] method. It is a method of sentiment analysis based on a value (positive or negative) assigned to each word [20]. When a word is present in a message, if its value is positive (respectively, negative), then the sentiment expressed is positive (respectively, negative). This method was chosen for its simplicity of implementation, even if it has a certain number of lacks, in particular not taking into account pairs of words or sentences. For example, between the sentences "it's pleasant" and "it's not pleasant". The word *pleasant* has a positive value (+2), which must be reversed in the second sentence.

The indicator for the **sentiment** criterion is calculated by summing values of the positive and negative words extracted from the messages and compared to the AFINN dictionary [19], which contains 2193 words with, for each word, a sentiment value (positive or negative) enclosed between the interval [-5, +5]. If the sum is greater than 0, then the sentiment is positive (indicator = 1), otherwise negative (indicator = 0). This indicator is considered for


```

1 input mq: messages exchanged by learner l
2 in quiz q
3 output sentimentq: sentiment of learner l
4 in quiz q
5
6 sum = 0
7 foreach message in mq do
8   decompose message into list of words
9   foreach word in the list do
10     search for word in the AFINN dictionary
11     if word belongs to AFINN
12       then add word's sentiment value to sum
13   end
14 end
15 if sum > 0 then sentimentq = 1
16 else sentimentq = 0

```

Listing 1. Sentiment indicator calculus

TABLE I
SIMULATED VALUES

	Learner's Id	Grade (/100)	Number of messages (Total=23)	Present or Not at the same time	Positive sentiment	Negative sentiment
1	H1	69	5	YES	25	-19
2	H2	26	0	NO	0	0
3	H3	74	43	YES	3	-1
4	H4	54	47	YES	22	-12
5	H5	61	23	YES	12	-22

each learner concerned by the quiz. The principle of sentiment indicator calculus is explained by listing 1.

- The learner's **knowledge** is assessed using his score on quiz, which is the number of correct answers out of to the total number of questions (score belonging to the interval [0, 1]).

In the section below, we describe an experimentation of the proposed model.

IV. EXPERIMENTATION

To test the qualimetry model for the evaluation of collaborative quizzes, we simulated data for a group of learners made up of 5 students who answered a questionnaire composed of 100 statements.

We generated randomized data for a group. The values are shown in table I.

The metrics are obtained as follows:

- The **grade** is calculated through the sum of the correct answers out to the total number of questions.
- The **number of messages** (representing the communication in a group) is the number of messages sent in the group, by the learner, compared to the total amount of messages exchanged during the test.
- The **relative communication** represents the weight of messages exchanged in a group relatively to all messages

TABLE II
METRICS VALUES (HIGH COMMUNICATION LEVEL)

id	Grade	Communication	Presence	Sentiment
H1	0,69	0,04	1	1
H2	0,26	0	0	0,5
H3	0,74	0,36	1	1
H4	0,54	0,4	1	1
H5	0,61	0,2	1	0

sent for all groups. Its the total amount of messages for a group divided by the total number of messages for all groups.

- When learner answer the quiz at the same time as his group, the indicator of **presence** has the value 1, otherwise the value is 0.
- The metric for **positive and negative sentiment** will allow us to calculate the *sentiment* criterion with the following formulas:

- sentiment = 1, if (positive + negative) > 0
- sentiment = 0.5, if (positive + negative) = 0
- sentiment = 0, if (positive + negative) < 0

Table II show metrics values obtained from the simulated data.

From the metrics, we calculated the values of the factor *collaboration* for each learner, by applying three different normalization functions.

The proposed functions for a learner *l* answering a quiz with a group *g*:

- Function based on greater importance given to the learner's grade:

$$collaboration_{ig} = co_g * \frac{(2 * gr_i) + me_i + pr_i + se_i}{5}$$

- Function based on greater importance given to communication:

$$collaboration_{ig} = co_g * \frac{gr_i + (2 * me_i) + pr_i + se_i}{5}$$

- Function based on an equality between the criteria:

$$collaboration_{ig} = co_g * \frac{gr_i + me_i + pr_i + se_i}{4}$$

Where co_g is the **relative communication** of the group *g*, gr_i the **grade** of learner *i*, me_i the number of messages, pr_i the **presence**, and se_i the **sentiment**.

A. Results

We will discuss the values of the *collaboration* factor for the simulated values, with respect to each normalization function (see table III).

a) *Grade based function*: Learners H1, H3 and H4 (table III), have high ranks, but we notice that the levels of H1 and H4 are very close, even though the grades are distant. This is due to the inclusion of communication in the assessment of learners. H1 has a better grade but does not communicate, while H4 has a lower grade but communicate more.

TABLE III
FACTOR *Collaboration*

Collaboration factor			
id	function a	function b	function c
H1	0,57	0,46	0,57
H2	0,17	0,13	0,16
H3	0,64	0,58	0,65
H4	0,58	0,56	0,62
H5	0,41	0,34	0,38

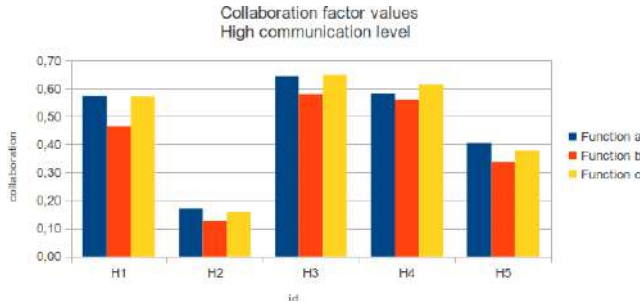
b) *Communication based function*: The collaboration values of this function are distant to the values of function (a), because this function is very sensitive to the communication metric.

As a general rule, the collaboration values for function (a) are always higher than for function (b), the communication occurring more in function (b) and being a relative rather than an absolute metric.

c) *Function based on criteria equivalence*: The main remark concerning the table function (c), is the proximity of the results of this function with those of the function a. Which is consistent with the importance of communication in the *collaboration* factor. In this case, the grade has relative importance, even if it is also present in the evaluation of the overall factor.

B. Discussion

Histograms in figure 6 present values of *collaboration* factor obtained with each specified normalization function.

Fig. 6. Histogram for *collaboration* factor

From the results obtained by simulating a test on a collaborative quiz, taken by a group of 5 learners, we observe that the normalization functions a and c offer very similar *collaboration* values. This is explained by the fact that in the function b the emphasis is on communication, which is a relative value, which therefore reduces the level of collaboration within this function.

In conclusion, we observe that when evaluating collaboration, the *communication* criteria is important. In the other hand, the *sentiment* criteria was not sufficiently exploited, it will be necessary to propose another metric to measure it.

We propose the use of the function (a) when it is preferable to give importance to grade in the assessment of learners,

and the function (b), when communication is important. The function (c) will be abandoned because it provides the same results as the function (a).

V. CONCLUSION

This paper presents a model, based on qualimetry, used to assess learners in a collaborative learning activity. The measured factor is *collaboration*, and it is calculated from criteria, evaluated with metrics, associated in normalization functions. Three function were proposed and tested, the results show that, the *communication* metric and *grade* have more influence in each of normalization functions.

As perspective for this work, we propose to add new metrics, primarily to better measure the learner's sentiment for the group, as well as his presence or absence during collaborative work. We also propose to use the collaboration factor to dynamically group learners.

REFERENCES

- [1] Zheng, Y.; Liu, Y.; Lu, W. & Li, C. A hybrid PSO-GA method for composing heterogeneous groups in collaborative learning 2016 11th International Conference on Computer Science & Education (ICCSE), 2016, 160-164
- [2] Pang, Y.; Xiao, F.; Wang, H. & Xue, X. A clustering-based grouping model for enhancing collaborative learning 2014 13th International Conference on Machine Learning and Applications, 2014, 562-567
- [3] Chen, C.-M. & Kuo, C.-H. An optimized group formation scheme to promote collaborative problem-based learning *Computers & Education*, Elsevier, 2019, 133, 94-115
- [4] Bernacki, J. & Kozierkiewicz-Hetmanska, A. The comparison of creating homogeneous and heterogeneous collaborative learning groups in intelligent tutoring systems *Asian Conference on Intelligent Information and Database Systems*, 2015, 46-55
- [5] Zheng, Y.; Li, C.; Liu, S. & Lu, W. An improved genetic approach for composing optimal collaborative learning groups *Knowledge-Based Systems*, Elsevier, 2018, 139, 214-225
- [6] Odo, C.; Masthoff, J. & Beacham, N. Group formation for collaborative learning *International Conference on Artificial Intelligence in Education*, 2019, 206-212
- [7] Maqtary, N.; Mohsen, A. & Bechkoum, K. Group formation techniques in computer-supported collaborative learning: A systematic literature review *Technology, Knowledge and Learning*, Springer, 2019, 24, 169-190
- [8] Bekele, R. *Computer-Assisted Learner Group Formation Based on Personality Traits* Bekele, R. (2005). Doctoral dissertation, Staats-und Universitätsbibliothek Hamburg Carl von Ossietzky, 2005
- [9] Gouasmi, N., Lamia, M., & Lafifi, Y. (2022). Extracting Learner's Model Variables for Dynamic Grouping System *Networking, Intelligent Systems and Security*, Springer, 2021, 479-493
- [10] balaji2017effective) Balaji, K. & Poorni, G. An Effective Recommendation System for E-Learning Using Fuzzy Tree *African Journal of Basic & Applied Sciences*, 2017, 9, 126-131
- [11] Yamamoto, N. & Uchida, N. Dynamic Group Formation for an Active Learning System Using Smartphone to Improve Learning Motivation *International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, 2018, 183-189
- [12] Garshasbi, S.; Mohammadi, Y.; Graf, S.; Garshasbi, S. & Shen, J. Optimal learning group formation: A multi-objective heuristic search strategy for enhancing inter-group homogeneity and intra-group heterogeneity *Expert Systems with Applications*, Elsevier, 2019, 118, 506-521
- [13] Azgaldov, G.; Kostin, A. & Omiste, A. P. The ABC of Qualimetry. *The Toolkit for Measuring Immeasurable* Ridero, 2015
- [14] Argotti, Y.; Baron, C. & Esteban, P. Quality quantification in systems engineering from the qualimetry eye 2019 *IEEE International Systems Conference (SysCon)*, 2019, 1-8
- [15] Juniawan, F. P.; Sylfania, D. Y.; Sulaiman, R.; Putra, R. R. C.; Pradana, H. A.; Sugihartono, T. et al. E-Voting Software Quality Analysis with McCall's Method 2020 8th International Conference on Cyber and IT Service Management (CITSM), 2020, 1-5

The 1st International Conference on Autonomous Systems and their Applications (ICASA'22)

- [16] Swarup, M. B. & Ramaiah, P. S. An approach to modeling software safety *2008 Ninth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*, 2008, 800-806
- [17] Walters, G. F. & McCall, J. A. Software quality metrics for life-cycle cost-reduction *IEEE Transactions on Reliability*, IEEE, 1979, 28, 212-220
- [18] Lac, C.; Raffy, J.-L. *et al.* A tool for software quality *Proceedings of the Second Symposium on Assessment of Quality Software Development Tools*, 1992, 144-150
- [19] Nielsen, F. A. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs arXiv preprint arXiv:1103.2903, 2011
- [20] Silge, J., & Robinson, D. *Text mining with R: A tidy approach* O'Reilly Media, 2017

Automatic real-time road extraction image based on colour segmentation polynomial regression model algorithm

Amine Bellebna
Ecole National Polytechnique (ENP)
 Alger, 10 Rue des Frères OUDEK, El
 Harrach 16200, Algiers, Algeria
m.a.bellebna@gmail.com

Aissa Boudjella
dept. Electrical Engineering
 Bircham International University
 Miami, USA
boudjellaa@bircham.edu

Abstract—In this investigation we have developed a road extraction computer vision algorithm, a technique for road recognition of autonomous cars. The system is implemented and simulated in python, and its performance is tested on real road images based on a color segmentation polynomial regression linear model, which detects the road from the markings lines using image processing toolbox. The system consists of the following main steps: 1) Image acquisition and preprocessing; 2) Optimal thresholding and marking lines identification; 3) Bird's-Eye View Transforms (BEVT); 4) Segmentation; 5) Inverse Perspective Mapping (IPM) and, 8) Second-degree smoothing regression. The present results show that the computation time t_p decreases and the recognition rate R_{rate} increases as the threshold T_c increases, which gives a better recognition rate of 90% in the threshold interval of [220~245] with an average computation time of approximately 0.010 seconds making this method robust and faster. Based on the magnitude of recognition rate R_{rate} , the Error rate E_{error} , and the average computation time t_p , we can define four (04) dominant regions to evaluate the performance of the system: The recognition rate R_{rate} reaches its minimum value zero ($R_{rate}=0\%$, $t_p=[0.0256-0.029]$, $E_{error}=100\%$) in the threshold interval $T_c=[100-115]$. When T_c increases from 130 to 155, the recognition increases with an average rate equals approximately to 50 % ($R_{rate}=50\%$, $t_p=0.014$, $E_{error}=50\%$). In the interval of $T_c=[170-195]$, R_{rate} increases and the processing time remains constant ($R_{rate}=80\%$, $t_p=0.010$ seconds, $E_{error}=20\%$). Finally, in the interval of $T_c=[200-245]$, the accuracy reaches its highest value 90% and low processing time $t_p=0.010$ seconds. This method shows a significant higher recognition rate with reduced processing time.

Keywords—Autonomous car, Road extraction, Transformation in bird's eye view

I. INTRODUCTION

In recent years, major automobile brands (Volvo, Mercedes, BMW, Audi, and others: even the American company Google) have been looking for solutions to make their vehicles safer in terms of safety, by increasingly assisting the driver using state-of-the-art electronic equipment, radars, and sensors, to perform various tasks better and more efficiently than the driver. Many research works have been published [1,2,3,4,5,6,7,8,9]. These various innovations gradually make automatic cars completely autonomous. Even if cars are not yet developed, they are at

the beginnings of the autonomous. In addition, competition between different car companies is about to increase intelligent car autonomy using artificial intelligence. In fact, the implementation of new tools will later allow the development of intelligent cars. They will take over from the human driver and allow safer roads. With the technology development, the autonomous car is on its way to replacing, the present no autonomous cars.

The world of mobility is undergoing profound change. The deployment of the autonomous vehicle will not only transform our modes of transport, but also will have an impact on the evolution of society in terms of safety, environment, town planning, and on the automotive industry. Even, to produce vehicles in which telecommunications and artificial intelligence will play an important role.

According to WHO statistics, the number of road accident victims around the world is increasing, in most cases due to human mistakes. However, if the vehicle no longer depends on the driver but on a computer, the risk of an accident may decrease. The main reason for building autonomous intelligent vehicles is to improve safety conditions fully or partially automating driving tasks. Among these tasks, road detection continues to play an important role in driver assistance systems (ADAS). These new techniques open up research development perspectives for new generation cars. In this study fixed cameras on-board car for automatic extraction of road in real-time, allowing the vehicle to determine its trajectory, based on color lines segmentation, and experimental test using KITTI-Road dataset [10].

METHODOLOGY

KITTI road database used in this work is a secondary dataset. The dataset contains 384 images divided on three several types: unmarked road, marked road, multiple marked road. We will use the multiple marked road in this paper. The detection of the road requires real-time processing of images supplied by an embedded camera in the car, to extract useful information, namely the markings on the ground, which is considered as a reference point for locating car position on the road lane. Cameras will replace the user's eyes allowing the car to see the world around it. An autonomous car can be equipped with different types of

cameras such as mono-vision cameras, stereo vision cameras, and infrared cameras.

Computer vision is based on the use of camera mounted on top of the windshield to provide an open view of the road. With a camera, you can detect the road by extracting markings from the ground. This processing makes automatic driving possible so that a car follows its lane trajectory on the road without the intervention of the driver. In case the driver is not paying attention, the image captured by a camera can also be used to process the segmentation of road signs to indicate to the driver a notification on the detected plate. Another artificial intelligence treatment with a computer vision is to shape recognition treatment, and identify what is a vehicle types (car, truck, bus, motorcycle ...), or a pedestrian crossing the road. The reaction then depends on the level of danger, either it is in the form of sound signals to alert the driver, or it is the electronic system which takes control of the car.

A lane departure warning requires the execution of three processes as indicated in Fig.1: 1) Acquisition of road image; 2) detection of the road by calculating the lane line curve; 3) Vehicle control based on the previous processing.

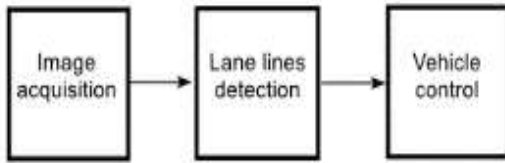


Fig.1. Lane departure warning process.

II. SYSTEM MODEL OVERVIEW

There are several algorithms proposed for road detection. In this investigation, we have developed a road extraction algorithm to develop computer vision, which detects the road from the marking lines. The detection is based on color segmentation which uses image processing toolbox implemented on python consists of the following steps:

The first step of this algorithm is the segmentation of the marking lines in a grayscale image after applying a threshold to extract the white color, which is the color of these lines. The functions used at the level of the proposed algorithm is the OpenCV library is a computer-based library that offers excellent functions in terms of real-time processing and performance. We have divided the program into four (04) functions based on the functions of the OpenCV library. The program is structured as follows:

A. The Segmentation function

The segmentation function will allow segmenting the marking lines, through three (03) very simple steps: a) going through the gray level using the function; b) histogram equalization with the function, and) segmentation of the marking lines using an optimal threshold. All the pixels between the threshold and the maximum value will have the maximum value 1 otherwise 0.

B. The function of the IPM

IPM is a function of the Opencv library. The segmented image to be transformed into a bird's view, and homograph matrix makes the projections from the source zone to the destination zone Fig.2. This matrix is calculated by calling

the of the Opencv library function getPerspective Transform either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as "3.5-inch disk drive".

C. The detection lines function

The detection lines function is the function that detects by sliding windows. It will return an array of coordinates of the right line and an array of coordinates of the left line. Fig.3

D. Second degree smoothing function

At the level of this function, we calculated the parameters of the mathematical model of the lines using the poly fitting function of the NumPy extension, which allows refining using the least-squares method. The function receives parameters (the vector of coordinates of the straight line, and the degree of the polynomial, which is of order two). In return, we receive a vector of coefficients of the second-degree polynomial "a, b, c ". To plot the curve represented by $y_i = ax^2 + bx + c$, we use cv2.polylines () method: (Fig. 6).

III. EXPERIMENTAL RESULTS

A. Thresholding of Making Lines

On the grayscale image as we have seen previously, the marking lines are present in white color; an intuitive method that allows the extraction to be done is to set a threshold that segments these lines.

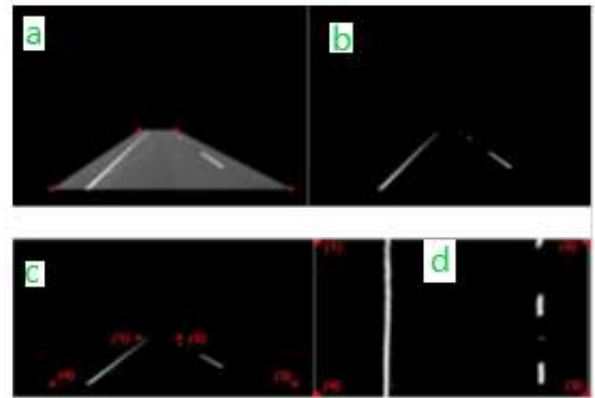


Fig.2. The area of interest (a), Image segmentation (b), transform the camera view image to the BEV bird view (d) from the image (c).

B. Transform in Bird View

After segmentation, the next step transforms the camera view into an aerial view Fig.2(d). It is also called the bird's-eye view (BEV). This technique removes the perspective effect of the image, because in a camera view, the lines seem to converge at a point called the vanishing point, but in reality, the lines are parallel. The goal is to have an image with the two marking segments in parallel to facilitate the line detection processing for the next step. This transform requires a homography matrix [11], which is used to make the correspondences between the source image and the destination image. The bird's-eye view is the perspective projection of the original image is indicated in Fig.4. The relation between the pixel of the destination image as the

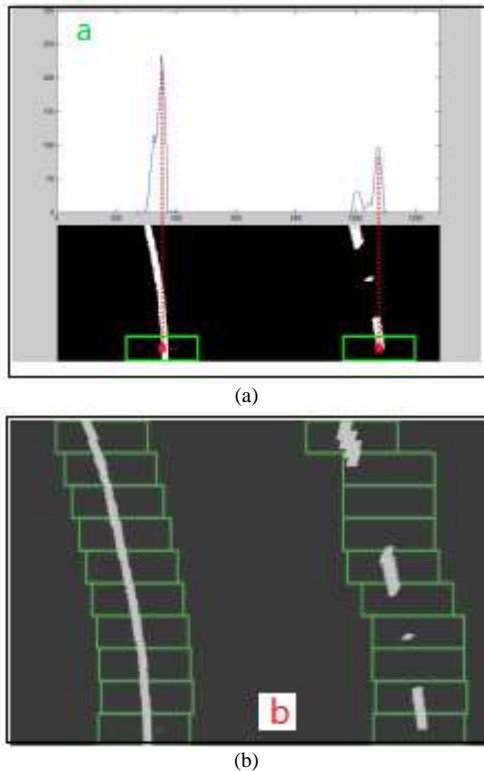


Fig.3 (a) Thresholding of the marking lines. Original image (b) The extraction of the coordinates of points by the sliding windows

crow flies and the pixel of the source image can be formulated by the homography matrix [11]. The homography matrix is calculated so that the area bounded by the four points on the source image (the four points of the area of interest), will be projected a new area on the destination image as indicated in Fig.2.

C. Identification of Marking Lines

In this step, we identify the right and the left lines. The technique used in this case is based on a sliding window, It allows to follow the line and extract the position of each point that is detected inside this window (white dots segmented). Two (02) windows will be placed at the very bottom of the image, which has well-determined dimensions, the first on the right line and the second on the left line, and extract the set of points that are inside of each window (coordinates x , linked to the frame of the image), then they will move above its first position to perform the next extraction. To determine the initial position for each window, it is necessary to calculate the histogram which represents the number of white pixels according to the position on the horizontal axis on the image, in other words, the sum of the pixels on each column of the image. This processing is only done on the lower half of the image but not on the entire image. Figure 3.1 shows an example of setting detection windows. The next position of the window will depend on the position calculated above. It is a displacement by a step equal to the width of the window on the ordinate axis, and the next abscissa will be the average of the abscissa of the points, which are extracted by the window if it has succeeded in making the detection.

However, if no point is inside the window, it will keep its position on the x -axis Fig. 3(b).

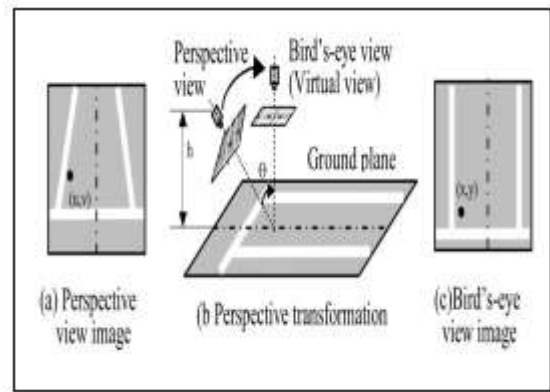


Fig.4 Transformation in bird's eye view.[11]

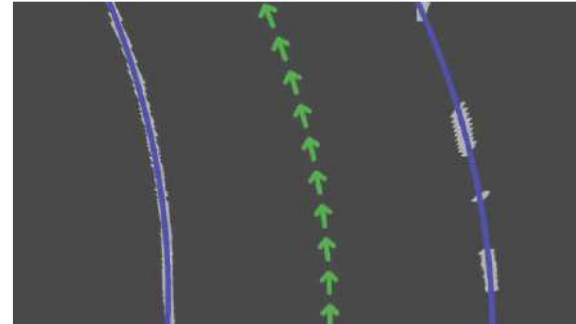


Fig.5 Result after second order polynomial regression

D. Polynomial Regression

The windows of the lines will extract the straight line, store them in a vector by their Cartesian coordinates, and the same thing for the left side, and it remains only to adjust the series of this set of coordinates of points (the pixel coordinate in the image coordinate system). The curves used to model roads are the (2nd) second-degree curves. We will fit this set of points detected on a given image to an arc of a parabola of the equation. Once we have found the coefficients " a , b , c " for each side of the road from the polynomial equation, we can plot the function, which follows the model of the parabolic equation on the image, and we will have the result of the detection. Fig.6 shows the result of the parabolic regression of all the points detected. The coefficients a , b , and c of the parabolic equations are indicated in table 1.

TABLE I. COMPARISON OF PARABOLIC REGRESSION CURVES OF ROAD LINES (FIGURE 6)

Right line	$a = -3.2e-4$	$b = 0.58$	$c = 868.9$
Left line	$a = -4.8e-4$	$b = 0.47$	$c = 117.8$

The middle of the lane of the road is shown in Fig.6 above in green arrows. It is calculated by the average of the model coefficient of the right and the left lines, and then the decision-making is based on this middle line. If it is located in the middle of the image so the car is in the middle of the road, otherwise an attempt to leave the lane can be detected by the distance between the calculated middle line and the middle of the road image. It is assumed that the middle of the image is a reference if the camera is properly calibrated



Fig.6 (a) the source image (b) detection.



Fig.7 (a) Result after returning to the camera view. (b) Superposition of the result on the original image algorithm/

E. Results

After the detection of the marking lines, we do an inverse transform from, the aerial view to the initial view of the camera. As we are working on a segmented image i.e. black & white colors, the aerial view will be superimposed on the original image in order to visualize the result of the detection. The following aerial view shows the result of the road detection (Fig.6 and Fig.7).

F. Test Images Detection on the Road

In order to compare the results of the proposed algorithm to the same source image, we calculate the corresponding execution time. The test is performed on images taken on the national road (Algeria). The detection result is shown in Fig.7.

Fig.8 shows the result of the detection of the road with the proposed technique, which uses the extraction of lines by sliding windows using ten (10) images. This method performs better for nine images, but only the first image gave a bad result. Therefore, the success rate is 9/10. In addition to the detection of images the processing time of each image indicated in Table.II.

This method based on the thresholding of the white color is very sensitive to the change of lighting. According to the results of the tests on the images taken from the KITTI database, the proposed algorithm performs better in terms of success rate and processing time. It has a fast execution; and can estimate the model of the channel by a polynomial curve. Therefore, it offers a possibility of detecting turns and estimating the speed limit according to the radius of curvature, which opens up another subject of research. The only downside is the rate of illumination, which affects the threshold range. Twenty (22) of images were used for testing to assess preliminary classifier performance. First, the image features are extracted and then the transformation coefficient of the wavelet function is obtained by wavelet decomposition. These functions are tested using the proposed algorithm. Fig.11 shows the results of the detection of the road with several threshold values on two different road images (110, 145, 185).

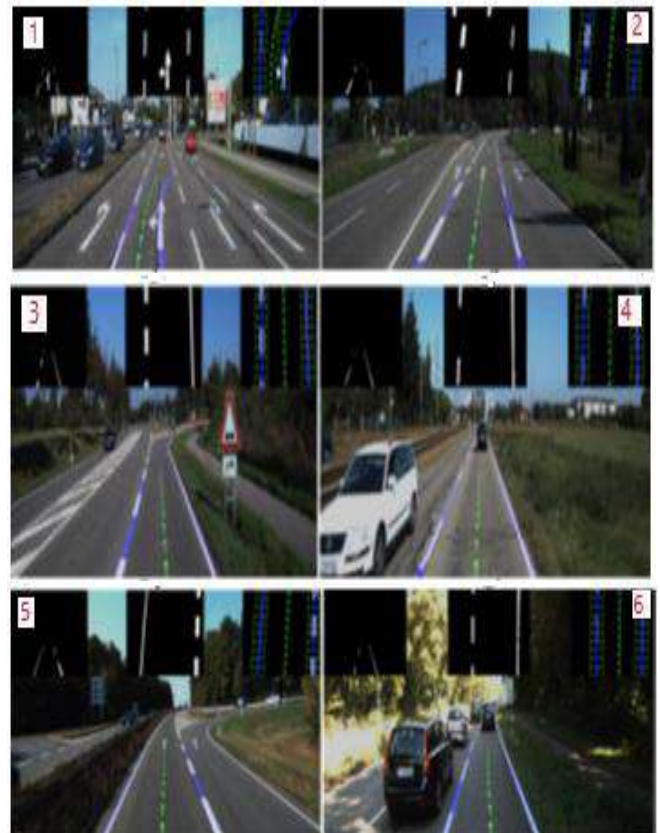


Fig. 8 Test result with the first algorithm on 10 images of KITTI

TABLE II. PROCESING TIME CALCULATED USING PYTHON ON SEVERAL IMAGES IF KITTI DATASET TC=[200~245]

Image	Processing time of each function during the test (seconds)				
	Segmentation	BEV	Sliding window	Polyfitting	Total time
1	0.00098	0.0029	0.0119	0.0059	0.021
2	0.00097	0.0029	0.0129	0.0059	0.022
3	0.00099	0.0029	0.0119	0.0059	0.021
4	0.00096	0.0029	0.0119	0.0059	0.021
5	0	0.0020	0.0129	0.0059	0.020
6	0.00097	0.0032	0.0130	0.0059	0.023
7	0	0.0019	0.0119	0.0059	0.020
8	0.00098	0.0038	0.0109	0.0059	0.021
9	0.00096	0.0029	0.0129	0.0059	0.022
10	0	0.0024	0.0119	0.0059	0.020

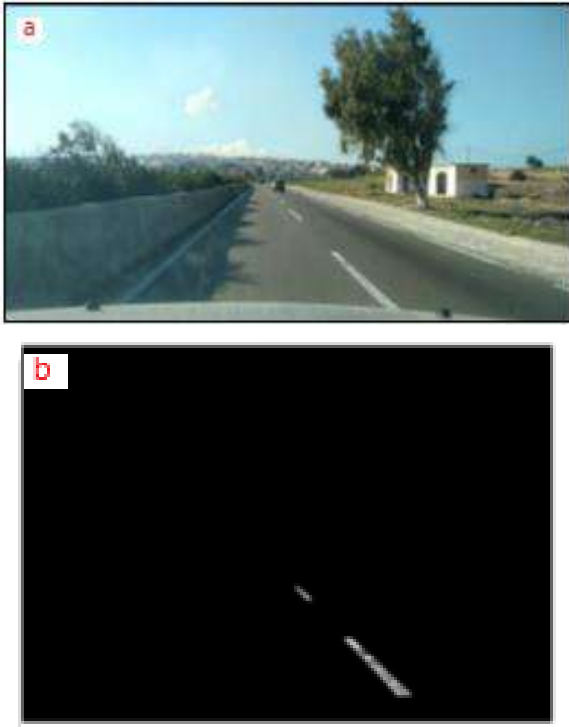


Fig. 9. (a)The test image, (b)Segmentation by the proposed algorithm

The road rate recognition or accuracy is characterized by the recognition rate or accuracy represented by

$$R_{rate} = V/W \quad (1)$$

Where V: The total number of recognized images, and W: The Total number of image to be read.

The error rate is given by

$$E_{error} = K/W \quad (2)$$

Where K :The total number of non-recognized images. Fig.12. Shows the recognition rate R_{rate} , the error rate E_{error} , and the processing time t_p versus the threshold T_c . The processing time decreases as TC increases. The highest value (0.029) is observed when TC =100 and the lowest value (0.010) at TC =255. In the interval of TC =[100~115], the accuracy reaches its minimum value zero. When TC goes from 130 to 155, the recognition rate remains approximately constant with an accuracy of Rrate=50%. As TC becomes larger than 155, in the interval of [170~195], the recognition rate increases with the average value equal to 80%.Increasing

TC from 200 to 245, the threshold does not affect significantly the recognition rate, Rrate=90%.Based on the magnitude of recognition rate, we can define four (04) dominant regions: In the threshold interval of TC=[100~115], the recognition rate reaches its minimum value zero (Rrate=0%, tp=[0.0256~0.029], Error=100%). In the interval of TC = [130~155] the recognition increases with an average rate equals to 50 % (Rrate=50%, tp=0.014, Error =50%). In the interval of TC = [170~195], the accuracy increases (Rrate=80%), and the processing time remains approximately constant tp=0.010 seconds with the error rate Error =20%). Finally, in the interval of TC = [200~245], the accuracy reaches its highest value 90% and low processing time tp=0.010seconds.

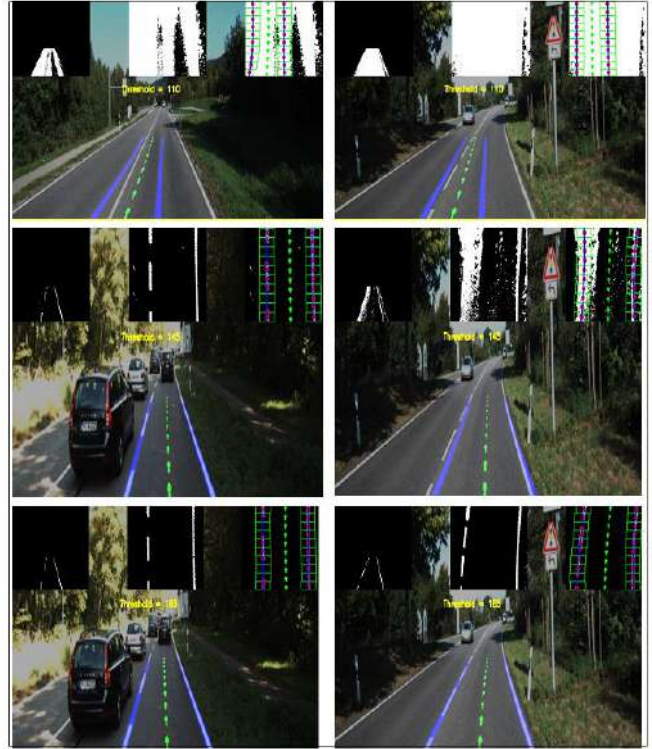
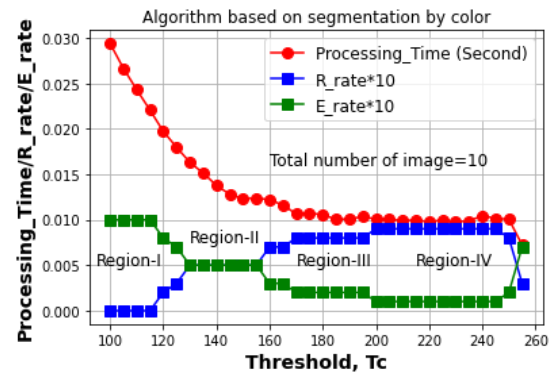


Fig.10 Polynomial regression-graphs (blue color) in the intervals [100~255]

Fig.11 Recognition rate, error rate, and processing time versus threshold. Abbreviation X₁=segmentation processing time versus the threshold X₂=BEV, X₃=detection line, X₄=Polyfit, and total time.

IV. CONCLUSION

Road recognition for automatic real-time road extraction based on color segmentation algorithm is proposed to improve the safety and the efficiency of automatic driving-

assistance system to reduce the workload of security personnel. When applying the road markings detection color segmentation algorithm, a car can follow its lane on the road without the intervention of the driver. Before implementing the system, previously study is necessary. A comparison of the performance characteristics under fixed total number of images 10 is carried out to examine the recognition rate accuracy when the image threshold T_C changes from 0 to 250. Based on the magnitude of the recognition rate R_{rate} , the Error rate E_{error} , and the average computation time t_p , we can define four (04) dominant regions to evaluate the performance of the system:

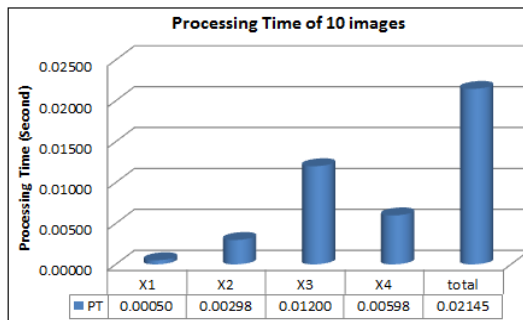


Fig.12. Illustrates the Recognition rate R_{rate} , the error rate E_{error} , and the processing time versus threshold. Abbreviation X_1 =segmentation processing time, X_2 =BEV, X_3 =detection line, X_4 =Polyfit, and total time. PT: Processing Time.

- 1) Region I:** In the threshold interval of the threshold $T_C = [100 \sim 115]$, the system works badly with a higher error 100% and lanes with the highest processing time $t_p = [0.0256 \sim 0.029]$ could not recognize road lanes.
- 2) Region II:** When T_C increases from 130 to 155, the recognition rate increases with an average error equals to 50% and the processing time $t_p = 0.014$. Under these conditions, the implemented system need the intervention of the drive.
- 3) Region III:** In the interval of $T_C = [170 \sim 195]$, the accuracy of recognizing the road lanes increases with lower error of 20% and constant time processing %, $t_p = 0.010$ seconds. Under this situation the system with lower risk is not autonomous. The driver should pay intention to avoid an accident.
- 4) Region IV:** Finally, in the interval of $T_C = [200 \sim 245]$, the system shows the best scenario. The accuracy reaches its highest value 90% and a low processing time $t_p = 0.010$ seconds. The system is not perfect and is not totally autonomous, but it helps the driver to minimize the risk

To conclude, the performance of the system can be improved by adjusting T_C in the range of $[200 \sim 245]$. The recognition rate can be optimized in the region IV with a lower error rate 10%. In this region V, the working system was tested successfully, which detects and recognizes the road on real images. The experimental results of the proposed method shows the effectiveness of identifying the road on real images with high accuracy 90%.

The image threshold combined with color segmentation algorithm can be used to control the recognition rate and the computation time to propose an optimal solution. This technique is faster with reduced processing time and maybe effective helping the driver to pay intention under certain

condition to take decision as accident prevention. This method serves as a performance metrics guideline that tells us how much better the proposed model is detecting the road lines. It also helps programmers to select the optimal model that generates the best performance scenario.

The purpose of our future work is to introduce machine learning computer-aid technology methods to perform a road recognition to improve the accuracy larger than 95%. The system will be implemented and simulated in Anaconda, and its performance will be tested on a real dataset that contains 6 features and three (03) classes defined by the characteristics of road [10]. The features of the numerical dataset such mean, variance, standard deviation, entropy, skewness, kurtosis etc. will obtained from the image proprieties.

REFERENCES

- [1] Manuel Martinez, Alina Roitberg, Daniel Koester, Rainer Stiefelhofen, Boris Schauerte. Using Technology Developed for Autonomous Cars to Help Navigate Blind People. Proceedings of the IEEE Int. Conf. on Computer Vision (ICCV), 2017, pp. 1424-1432J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [2] Alexander Hars. .Autonomous cars: The next revolution looms. Autonomous cars: The next revolution looms. Inventio Innovation Briefs 2010-01, Nuremberg, 2010, www.inventio.com/innovationbriefs/2010-01.
- [3] Ingrid Pettersson, I.C. MariAnne Karlssonl Setting the stage for autonomous cars: a pilot study of future autonomous driving experiences. The Institution of Engineering and Technology 2015.IET Intell. Transp. Syst., Vol. 9, Iss. 7, 2015, pp. 694–701..ISSN 1751-956X.doi: 10.1049/iet-its.2014.0168. www.ietdl.org.
- [4] Rasheed Hussain; JooYoung Lee; SheraliZeadally. Autonomous Cars: Social and Economic Implications.IEEE Nov.- Dec. 1 -2018 . pp. 70 –77. DOI: 10.1109/MITP.2018.2876922
- [5] Daniel Hein et al, What Drives the Adoption of Autonomous Cars?Thirty ninth International Conference on Information Systems, San Francisco 2018, pp. 1-17.6.
- [6] Zanchin, Betina Carol, Rodrigo Adamshuk, Max Mauro Santos, and Kathya Silvia Collazos. "On the instrumentation and classification of autonomous cars." In 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), , 2017 , pp. 2631-2636. IEEE.
- [7] Tamás Tettamanti1*, István Varga1, Zsolt Szalay1. Impacts of Autonomous Cars from a Traffic Engineering Perspective.Periodica Polytechnica Transportation Engineering. 44(4), 2016 pp. 244-250. DOI: 10.3311/PPtr.9464
- [8] Juan Contreras-Castillo, SheraliZeadally, Juan Guerrero-Ibáñez.Autonomous Cars: Challenges and Opportunities. IT Professional, Volume: 21, Issue: 6, Nov.-Dec. 1 2019, pp. 6 - 13. DOI: 10.1109/MITP.2018.2876930M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
- [9] J. A. Guerrero-Ibanez, S. Zeadally and J. Contreras-Castillo,"Integration challenges of intelligent transportation systems with connected vehicle cloud computing and Internet of things technologies", IEEE Wireless Commun., vol. 22, no. 6, pp. 122-128, Dec. 2015.Bb
- [10] Bihao Wang, Vincent Fremont, Sergio Alberto Rodriguez Florez. Color-Based Road Detection and Its Evaluation on the KITTI Road Benchmark. IEEE Intelligent Vehicles Symposium (IV 2014), Jun, Dearborn, United States. 2014, pp.31-36. fhal-01010998f
- [11] Luo LB, Koh IS, Min KY, Wang J, Chong JW. Low-cost implementation of bird's-eye view system for camera-on-vehicle. In2010 Digest of Technical Papers International Conference on Consumer Electronics (ICCE) Jan 9, 2010,pp. 311-312. IEEE.

Bio-inspired optimization algorithms for improving a variable step size Maximum Power Point Tracking method for photovoltaic systems

KamelEddine Heraguemi
Informatics departement
University of M'Sila
M'Sila, Algeria
Kameleddine.heraguemi@univ-msila.Dz

Abderrahim Zemmit
Electrical Engenering Departement
University of M'Sila
M'Sila, Algeria
Abderrahim.zemmit@univ-msila.dz

Abstract—In this paper, two novel intelligent maximum power points tracking of Photovoltaic systems (MPPT) developed and investigated. The proposed methods take advantage of Whale optimization algorithm (WOA) and grey wolf optimization (GWO) to tune the PID controller gains, which optimizes the PWM duty cycle driving the DC/DC boost converter. The biggest drawback of old MPPT algorithms is the oscillation around the MPP and fail accuracy under fast variable isolation. Therefore, to defy the previous MPPT limitations; an objective fitness function including simultaneously: accuracy, ripple, and overshoot is proposed and optimized by two different recently developed methods, whale optimization algorithm (WOA) and grey wolf optimization (GWO), with the aim to provide the most relevant variable step size. In order to validate our proposal; several testes are carried on a photovoltaic system based on Solar-ex MSX-60 panel and DC/DC boost converter. The performance of the methods is compared to classical fixed step size Perturbation and Observation (P&O) MPPT algorithm. The results show a clear superiority of our proposals against the classical one in terms of power overshoot and ripple. Moreover, the analysis shows huge benefits in response time in case of atmospheric condition rapid variation.

Keywords— *fixed and variable step size algorithms, Perturbation and observation (P&O) MPPT algorithm, Photovoltaic Cell modeling, optimization methods, grey wolf optimizer (GWO), whale optimization algorithm with (WOA) tracking accuracy, overshoot, Ripple.*

I. INTRODUCTION

Recently, the energy consumption has reached unexpected high levels, in which it surpassed all expectations, besides, continuous increasing demand and the high cost of conventional energies have forced many countries and institutions to find and develop new energy source able to replace the fossil energy gradually with better performance especially in term of abundance, pollution, price, efficiency, etc. Hence, most researchers have opted to investigate renewable energy as an alternative solution, which can be the most promising option while it is renewable and naturally replenished. Renewable energy is the energy produced from natural resources like geothermal, wind turbines, tides, rain, and sun. In this context, many studies have demonstrated that photovoltaic (PV) energy can be considered as one the most important renewable energy sources while it exhibits better performances, which is clean, free and abundant in the most of the world as well as simplicity in design, low maintenance, and low cost. However Photovoltaic still suffers some limitations, particularly it's low conversion efficiency which is only in the range of 9 - 17 % and the nonlinear characteristics. Hence, several studies and researches on the subject of the

development and improvement of PV systems have been updated continuously in terms of efficiency, cell materials, DC/DC converters, MPPT methods, etc.

To increase the efficiency of the PV system, many studies have been focused on maximum power point tracking MPPT algorithms [5 – 8]. These methods can be classified into two categories: conventional and intelligent methods. The conventional include perturb and observe (P & O) [5], Incremental Conductance (IC) [6], Hill Climbing (HC) [7], fractional open-circuit voltage [8], fractional short-circuit current [9], the intelligent methods uses neural network [20], fuzzy logic, Grey wolf optimization (GWO) [21], and genetic algorithms [22], Particle Swarm Optimization PSO[22]. the P–V curve of the module displays various local MPPs (LMPP's). In general bypass diodes, are used to decrease the hotspots on PV modules. Placement of these diodes in return generates various LMPP's, and at the same time, the controller has to track the global MPP (GMPP) despite various LMPP. Dhimish, et al. [30] provides the novel hot-spot mitigation technique using MOSFET's connected panel, and forward-looking infrared in thermal camera is used to observe these hot-spots. The power loss due to the hot-spot effect by these diodes has a prominent effect to consider. Because of different LMPP's, it is tough for the traditional MPPT's to distinguish the GMPP on the P–V curve. Thus, to accomplish this most extreme productivity from PV frameworks, some best MPPT algorithms are utilized [31]. To mitigate these limitations, especially in PSC's, analysts have proposed few strategies and techniques in the literature. In [32], Ghasemi et al., introduced a two-step methodology based on the GMPP tracking algorithm in which the proposed algorithm tracks effectively than particle swarm optimization (PSO) algorithm in PSC's. Huang et al. [33] introduced a methodology which predicts the MPP at a very faster rate based on a natural cubic-spline-based prediction model and this proposed algorithm is incorporated into the iterative search process. Coelho et al. [34] gives the new temperature based MPPT sensor, which is sophisticated in design aspects. This method utilizes the fact that the module voltage directly depends on the PV panel surface temperature. In [35], Karami et al., detailed at least 40 methods, includes advanced classical methods like three-point weight comparison method, parasitic comparison, β method, intelligent, and optimized techniques. However, this paper only confined to comparing the five parameters, mainly on algorithm tracking. Podder et al. [36] made a comparison depending upon the tracking nature. Nearly 50 techniques are reviewed, stating their merits and demerits. Comparison of different mathematical calculation/meta heuristic-based MPPT is made by [37]. Choosing an MPPT for a particular application is important during the design

process. References [38, 39] distinguished some of the MPP parameters in the literature. However, the review papers mentioned above lack the real-time implementation procedures of these MPPT techniques. Therefore, this paper furnishes the hardware information of the particular technique by different authors done in various platforms with their tracking speeds and efficiencies.

Although the great advance in MPPT algorithms, several limitations are inescapable such as the oscillation around the MPP and poor accuracy under fast variable atmospheric conditions. In this defy, two intelligent methods based on whale optimization algorithm (WOA) and Grey wolf optimization (GWO) have been proposed and implemented to guarantee better performances, in which both methods tune the variable step size providing the adaptive duty cycle of the DC/DC boost converter. Analysis and comparative study between the optimization algorithm (WOA) and grey wolf optimization (GWO) have been presented employing a boost converter connected to a Solar-ex MSX-60 model. Many scenarios and schemes for temperature and irradiation have been considered to validate the efficiency of the proposed methods.

II. MODELING OF PHOTOVOLTAIC CELL

Photovoltaic is the direct conversion of light into electricity. It uses materials which absorb photons of lights and release electrons charges. It can be used for making electric generators. The equivalent model of a PV cell is shown in Fig. 1. It consists of a light generated current source, a single diode, a series resistance R_s and a shunt resistance R_{sh} [3,10-11].

The solar cell terminal current can be expressed as a function of photo-generated current, diode current and shunt current.

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

Where:

I_{ph} is the current generated by the incident light (it is directly proportional to the Sun irradiation),

I_d is the current through the diode;

I_{sh} is the current through the parallel resistor R_{sh} .

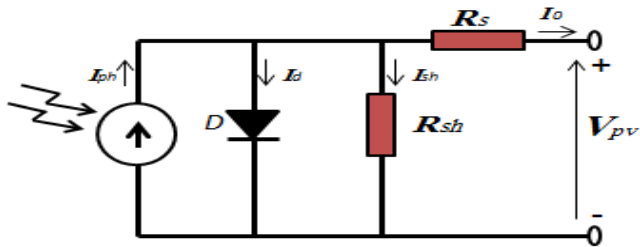


FIGURE 1: SIMPLIFIED EQUIVALENT CIRCUIT OF A PHOTOVOLTAIC CELL

The output current of a PV array is given by following equation:

$$I_o = N_p I_{ph} - N_p I_{rs} \left[e^{\left(\frac{q(v + R_s I_o)}{A k T N_s} \right)} - 1 \right] - N_p \left(\frac{q(v + R_s I_o)}{N_s R_{sh}} \right) \quad (2)$$

where:

I_{rs} : cell reverse saturation current;

q is the electron charge ($1.60217646 \times 10^{-19}$ C),

k is the Boltzmann constant ($1.3806503 \times 10^{-23}$ J/K),

n is the diode ideality constant.

R_s : series resistance of cell (0.001Ω).

T : reference cell operating temperature (20°C).

V : cell output voltage(V).

R_s and R_p are the series and shunt resistors of the cell, respectively.

The generated photo-current I_{ph} is related to the solar irradiation by the following equation:

$$I_{ph} = [I_{rs} + k_i(T - T_r)] \frac{S}{1000} \quad (1)$$

where

I_{sc} : cell short circuit current at reference temperature and irradiation

k_i : Short-circuit current temperature coefficient

T_r : Cell reference temperature

G : Solar irradiation in W/m²

III. PERTURB & OBSERVE (P&O) METHOD

At present, photovoltaic (PV) has received much interest as a secondary energy source. Due to nonlinear characteristics and low efficiency of photovoltaic arrays, tracking the maximum power point (MPP) of a photovoltaic array is an essential part of a PV system. Many MPPT techniques have been proposed and implemented. These methods include such as perturb and observe [1,2], incremental conductance [3], parasitic capacitance [4], constant voltage [5].

The mechanism of the classical P&O method is based on the perturbation of the module operating point and observation of the change in module output power. The polarity of the module output power defines the direction of the coming perturbation. For positive polarity, the next voltage perturbation can increase or decrease in the voltage, the same as that for the previous perturbation. For negative polarity, the On the Improvements of Perturb-and-Observe-Based MPPT.

The Perturbation and observation is one of the most commonly used MPPT methods for its simplicity and ease of implementation [2-4]. In this method, the array voltage is slightly disturbed (increase or decrease) then the actual value of the power $P(k)$ is compared to the previous obtained value $P(k-1)$. If the power panel is increased due to the disturbance, the following disturbance will be made in

the same direction. While if the power decreases, the new perturbation is made in the opposite direction. The flowchart of the Perturbation and observation method is illustrated in Fig.2.

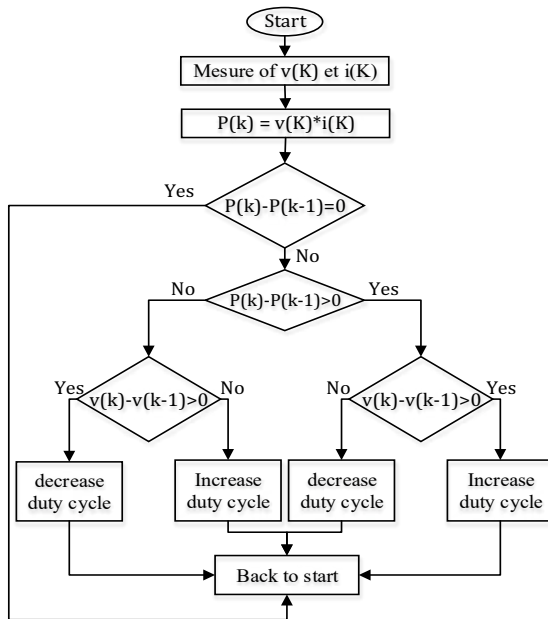


Figure 1: Flowchart of the conventional P&O algorithm

IV. BACKGROUND ON BIO-INSPIRED ALGORITHMS

In the last decade, the progress of optimization approaches encouraged through nature or interaction of animals to find food sources has increased [29], which results in hundreds of algorithms that proved their efficiency in various applications in many domains. Grey Wolf Optimizer [30] (GWO) and Whale Optimization Algorithm [31] (WOA) are two recently developed algorithms inspired by grey wolves and whale behaviors, respectively.

Based on the literature, GWO and WHO are simple and robust population-based algorithms. They are advantageous over other meta-heuristics because of the reduced number of random parameters and user-selected parameters. These advantages cause GWO and WOA to be appropriate algorithms for solving different constrained or unconstrained optimization problems for practical applications without a structural reformation in the algorithm. Due to their huge benefits, we investigate them in the MPPT method for PV Systems. This section will summarize the main inspiration and the mathematical model of GWO and WOA.

A. Grey wolf optimizer

GWO is population based stochastic algorithm that mimics the natural leadership hierarchy and hunting strategy of grey wolves. It was developed in 2014 by Mirjalili et al. This technique is essentially inspired by searching an hunting mechanism of Grey wolves groups. In the nature, wolves prefer to be in pack with a steady hierarchy. Each pack is divided into four categories relays on type and responsibilities of each individual, which helps in the process of hunting. The four categories are called:

alpha, beta, delta and omega. The leader of a group is called alpha (α), which is responsible for guiding and directing

the whole group during hunting mission, feeding, and migration. Beta (β) wolf stands on the second level of social pyramid. This category can substitute α when they are killed or cannot lead the group anymore. The next class are delta wolves (δ) and the rest of population is called omega (ω).

With a view to mathematically modeling the social hierarchy of wolves when designing the GWO, the first three that have the fittest solutions in the population are regarded as: α , β and δ . The rest of search agents are declared as ω . During the optimization process ω group is guided and directed by α , β and δ toward the promising search space in hope to find the best optimal solution.

Basically, the mathematical model of GWO is described in three main phases which are: encircling, hunting and attacking the prey and they are detailed as follows:

4.1.1. Encircling

When a prey is found, the iteration begins ($t = 1$). Therefore, α , β and δ wolves leads the rest of search agents to pursue and eventually encircle the prey, this behavior of grey wolves is expressed as:

$$\vec{X}(t+1) = \vec{X}_p(t) + \vec{A} \cdot \vec{D} \quad (2)$$

Where \vec{X} is search agents (wolves) positions, \vec{X}_p referenced the prey position. I is the iteration number. And \vec{A} is a vector of coefficient at the $(t+1)^{th}$ iteration, Whereas, \vec{D} is another coefficient that can be described as following:

$$\vec{D} = |\vec{C}, \vec{X}_p(t) - \vec{X}(t)| \quad (3)$$

The parameters \vec{A} and \vec{C} are combinations of controlling parameters a and random numbers \vec{r}_1 and \vec{r}_2 which can be mathematically expressed as follows:

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \quad (4)$$

$$\vec{C} = 2 \cdot \vec{r}_2 \quad (5)$$

where components of \vec{a} are linearly decreased from 2 to 0 over the course of iterations and \vec{r}_1, \vec{r}_2 are random vectors in $[0,1]$.

4.1.2. Hunting the prey

Grey wolves hunting behavior change position of each pack in the group by approaching to the prey, this behavior is mathematically described as follows: α is considered as the leader and the finest solution, β and δ are expected to know more information about prey's possible positions. Therefore, ω pack will pursue them and obliged to change locations in the light of α , β and δ in the next iterations. Position updating or hunting behavior is described by the following equations:

$$\begin{aligned}\vec{D}_\alpha &= \left| \vec{C}_1^t \cdot \vec{X}_\alpha^t - X^t \right| \\ \vec{D}_\beta &= \left| \vec{C}_1^t \cdot \vec{X}_\beta^t - X^t \right| \\ \vec{D}_\delta &= \left| \vec{C}_1^t \cdot \vec{X}_\delta^t - X^t \right|\end{aligned}\quad (6)$$

$$\begin{aligned}\vec{X}_1^t &= \vec{X}_\alpha^t - A_1^t \cdot \vec{D}_\alpha^t \\ \vec{X}_2^t &= \vec{X}_\beta^t - A_2^t \cdot \vec{D}_\beta^t \\ \vec{X}_3^t &= \vec{X}_\delta^t - A_3^t \cdot \vec{D}_\delta^t,\end{aligned}\quad (7)$$

$$X^{t+1} = \frac{X_1^t + X_2^t + X_3^t}{3} \quad (8)$$

4.1.3. Attacking the prey

The attacking process is controlled by the parameter \vec{a} changes the vector \vec{A} and conduct the omega wolves to approach or run a way from the prey (solution), theoretically, if $|\vec{A}| > 1$ wolves run a way to explore more search space. Else, they approach to dominants which means that omega wolves will follow the dominants which exploit the small search space. $\vec{D} = |\vec{C}, \vec{X}_p(t) - \vec{X}(t)|$ are linearly decreased from 2 to 0 over the course of iterations are carried on:

$$\vec{a} = 2(1 - t/N) \quad (9)$$

Where N is the total number of iterations and t is the actual iteration number. The basic steps of the Grey wolf optimization can be shown Fig. 3

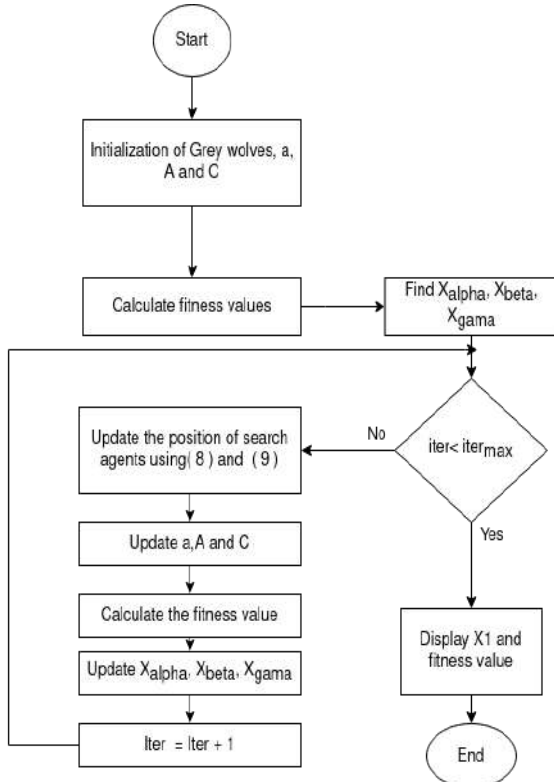


FIGURE 2: FLOW CHART OF GWO ALGORITHM

Whale Optimization Algorithm

Whale Optimization Algorithm (WOA)[31] belongs to the family of stochastic population-based algorithms, It is recently developed by Mirjalili et al. In 2016. This method is mainly inspired from the social behavior of humpback whales, considered as the biggest mammals in the world. More precisely, WOA is mimics the bubbles-net feeding in the foraging behavior of the humpback whales. The bubbles-net created when the whale swim in a 6-shipped path. **Fig.4** presents the Bubble-net feeding behavior of humpback whales.

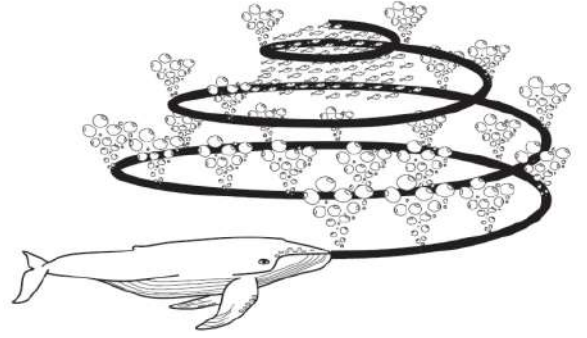


FIGURE 3: BUBBLE-NET FEEDING BEHAVIOR OF HUMPBAC WALÉS

Generally, the algorithm is composed in three main phases; Encircling prey, Bubble-net attacking method and Search for prey. The first two phases are described as exploitation phase, whereas the last is the exploration phase. The mathematical model of WOA is described in the following sections.

4.2.1. Exploitation phase (Encircling prey, Bubble-net attacking)

To hunt a prey, humpback whales first encircle it. Eqs. (12) and (13) can be used to mathematically model this behavior [31] .

$$\vec{D} = |\vec{C} \cdot \vec{X}'(t) - \vec{X}(t)| \quad (10)$$

$$\vec{X}(t+1) = \vec{X}'(t) - \vec{A} \cdot \vec{D} \quad (11)$$

where t indicates the current iteration, X' represents the best solution obtained so far, X is the position vector, In addition, A and C are coefficient vectors that are calculated as in Eqs. (14) and (15)

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \quad (12)$$

$$\vec{C} = 2 \cdot \vec{r} \quad (13)$$

where a decreases linearly from 2 to 0 over the course of iterations (in both exploration and exploitation phases) and it is a random vector generated with uniform distribution in the interval of [0,1]. Search agents update their positions based on the best-known solution. The solution location is controlled by the adjustments of r and values.

The hum-pack hunting method is based on shrinking encircling mechanism and a spiral-shaped path toward the prey. The shrinking behavior is formulated as shown in Eq (16).

$$a = 2 - t \frac{2}{\text{MaxIter}} \quad (14)$$

where t is the iteration number and $MaxIter$ is the maximum number of allowed iterations. The spiral-shaped path is calculated by the distance between the actual solution and the best position by Eq (17),

$$\vec{X}(t+1)=D'e^{bl}.\cos(2\pi l)+\vec{X}'(t) \quad (15)$$

Where $D' = |\vec{X'}(t) - \vec{X}(t)|$ describe the distance of i^{th} whale from the prey (The best solution obtained so far). A random coefficient p between 0 and 1 is used to choose between the two mechanisms (shrinking encircling mechanism and the spiral-shaped path) with probability of 50% during the optimization process. So that if $p < 0,5$ the shrinking encircling is used to update the position, else the spiral-shaped path will be used.

4.2.2. Exploration phase (Search for prey)

Whales also have a certain probability of searching for prey when they are constructing bubble-network. Mathematically, searching a prey enhance WOA exploration, this phase is based on the change of A coefficient. If A exceeds the range of $[-1, 1]$, the distance data \vec{D} is updated randomly. At this time, whales will deviate from the original optimal fitness, so that the algorithm has a certain global search-ability, which is formulated as follow:

$$\vec{D} = |\vec{C} \cdot \overrightarrow{X_{\text{rand}}} - \vec{X}| \quad (16)$$

$$\vec{X}(t+1) = \overrightarrow{X_{\text{rand}}} - \vec{A} \cdot \vec{D} \quad (17)$$

$I_{\alpha} = I_{pk} - I_d - I_{sk}$ Where, $\overrightarrow{X_{rand}}$ is random location information of a whale selected from this iteration. The flowchart of WOA technique is depicted in Fig. 5.

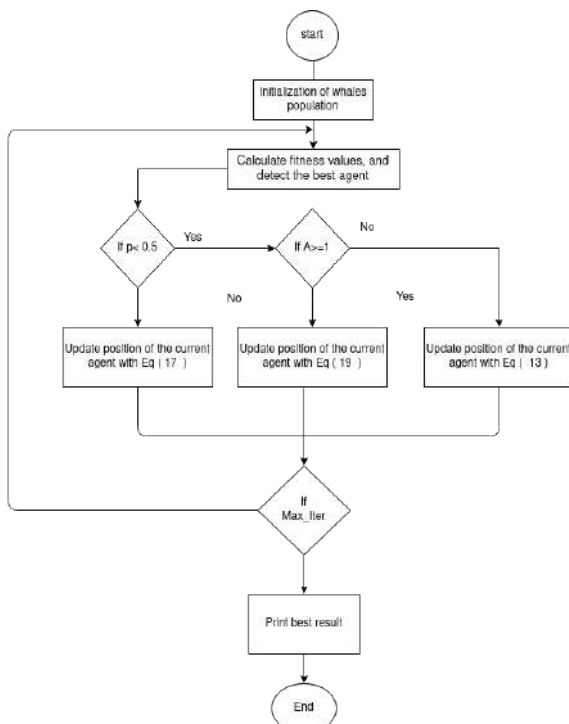


FIGURE 4: WHALE OPTIMIZATION ALGORITHM FLOWCHART

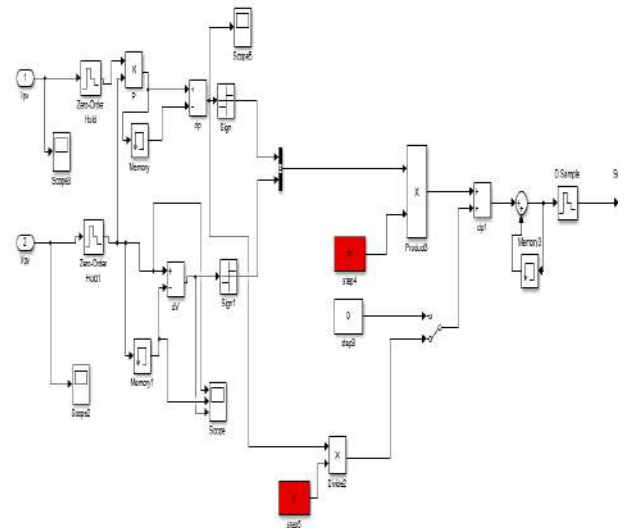


FIGURE 5: VARIABLE STEP SIZE MPPT METHODS SIMULINK MODEL

GWO/WOA are applied in the tuning of the step size m , y to ensure optimal control MPPT performance at nominal condition for the PV. The block diagram for the entire system is given in Fig. 7.

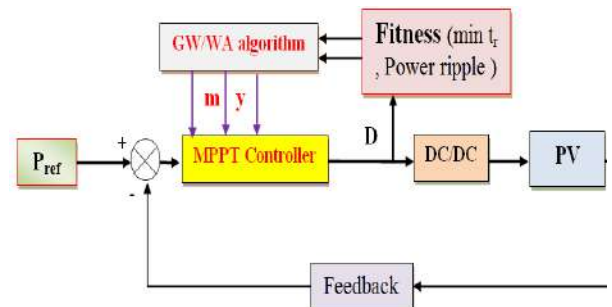


FIGURE 6: OPTIMIZATION OF STEP SIZE USING GWO AND WOA

V. VARIABLE STEP SIZE MPPT ALGORITHM USING GWO AND WOA

B. System implementation

As mentioned previously, the conventional MPPT methods based on fixed step size have a good performance. However, they are characterized by major drawbacks like slow convergence, oscillations around the MPP point and failing to track the MPP point with fast atmospheric conditions changes. Speedy tracking can be achieved with larger step size but excessive steady state oscillations is unavoidable. While smaller step size can reduce the oscillations with slower dynamics.

In order to solve these dilemmas, many contributions have been introduced using variable step size and significant progress has been made, where the algorithm changes the step size automatically according to the PV array

characteristics [1, 2, 9, 17]. Depending on each operational condition, step size should make a satisfactory trade-off between the dynamics and oscillations. This study proposes a new variable step size MPPT algorithm characterized by more simplicity, faster response time, and fewer oscillations. Fig. 6 shows the variable step MPPT developed using Simulink. The variable step-size method proposed is given as follows:

$$D(k) = D(k-1) \pm N^{\pm} dP \quad (18)$$

where

$D(k)$: is the duty cycle and coefficient;

N : is the scaling factor adjusted at the sampling period to regulate the step size (adjusted manually);

dP : is the PV array output power derivation.

C. Objective function

In each optimization issue, the evaluation of the results can be performed using the objective function. In which, the evaluation measures are included. The successful choice of function means more accurate and better results. In this work, a combination of various criteria is introduced in the evaluation function, in order to be optimized and get the best gains in our main goals such as Ripple, Overshoot, and system response time.

Firstly, integral square error (ISE)[29] criterion is given as a measure for ripple, in which the difference between the theoretical and produced power is calculated. ISE is given by the equation:

$$ISE = \int_0^{\tau} (P_{ref} - P_{out})^2 dt \quad (19)$$

also, the overshoot criteria employed as second part of the fitness function, which is defined as follows:

$$Overshoot = \max(P_{out}) - Pref \quad (20)$$

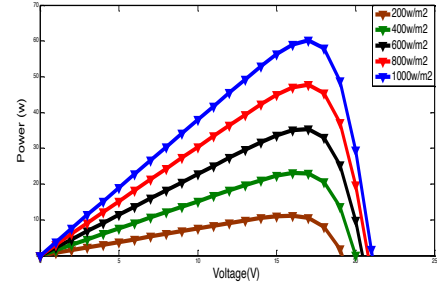
Base on these two measurements; we define our fitness function as follows:

$$F = \alpha \cdot ISE + \beta \cdot Overshoot \quad (21)$$

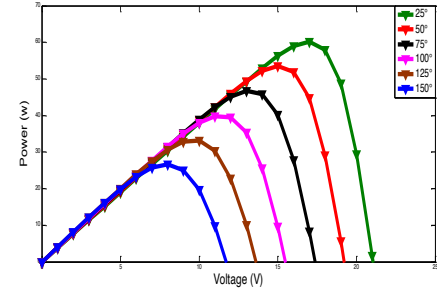
Where α and β are two empirical parameters used for balancing between the measurements. In our case, we have no preference between the two objectives, we choose ($\alpha=\beta=0.5$).

VI. SIMULATION RESULTS AND DISCUSSION

For simulations and analysis of proposed method, the Solarex MSX-60 photovoltaic module with 36 solar cells was chosen [1]. the electrical properties of the Solarex MSX-60 module for $S=1000W/m^2$ and $T=25^\circ C$ are illustrated in table. 1. Moreover, Figure.7 shows the P-V characteristics of the PV cell from a typical Solarex MSX-60 PV panel. For more details; Figure. 7.A show the P-V voltage characteristics of the PV cell for different irradiation levels (0.2, 0.4, 0.6, 0.8, 1 Kw/m^2) at fixed temperature $T=25^\circ C$, respectively. Whereas, the figure. 7,B reports the impact of the temperature variation ($T=25, 50, 75, 100, 125$ and $150^\circ C$) with constant irradiation ($S=1 Kw/m^2$) on the P-V voltage characteristics of the PV cell.



(A)



(B)

FIGURE 7: P-V CHARACTERISTICS UNDER VARIOUS INSOLATION LEVELS A) P-V CURVES FOR VARIOUS IRRADIATION LEVELS (SOLAREXMSX-60, $S=0.2, 0.4, 0.6, 0.8, 1 Kw/m^2$, $T=25^\circ C$). B) P-V CURVES FOR VARIOUS TEMPERATURES. (SOLAREXMSX-60, $S=1 Kw/m^2$, $T=25, 50, 75, 100, 125$ AND $150^\circ C$).

Description	MSX-60
Maximum Power (Pm)	60W
Voltage Pmax (Vm)	17.1 V
Current at Pmax (Im)	3.5 A
Short Circuit current (Isc)	3.8 A
Open Circuit voltage (Voc)	21.1

TABLE 1. ELECTRICAL CHARACTERISTICS OF SOLAREX MSX -60 ($1kw/m^2$, $25^\circ C$)

In order to prove our proposals efficiency, with the quick variation of irradiation due to weather changing or cloud passing, three irradiation steps signals were simulated using MATLAB/Simulink software as described by As stated in the background on bio-inspired algorithms section, the main advantages of GWO and WOA is the smaller number of parameters entered by the user. Mainly, three parameters are used in both algorithms, which are number of iterations, number of tables.2. the obtained results are compared to P&O classical algorithm and even compared between them.

Irradiation (W/m ²)	Time (s)
1000	0 - 0,5
600	0,5 - 1
800	1 - 1,5
600	1,5 - 2
1000	2 - 2,5

TABLE 2. TEST PATTERN SIGNALS.

As stated in the background on bio-inspired algorithms section, the main advantages of GWO and WOA is the smaller number of parameters entered by the user. Mainly, three parameters are used in both algorithms, which are number of iterations, number of agents and optimized variables number, such as, the last one representing the step size (m and y) in our model. Table.3 reports the utilized values for each parameter in the algorithms.

Description	Parameters	
	WAO	GWO
Number of search agents	10	10
Maximum number of iterations	20	20
Number of variables	2	2

TABLE 3. GWO AND WAO SETUP PARAMETERS.

As the first comparison, with the aim to forecast the best method that gives more minimization of the fitness function between GWO and WOA. Figure.8 reports the results of fitness function in terms of number of iterations. It's observed that whale optimization algorithm performs the grey wolf's optimization algorithm in terms of minimization, which means more efficiency in terms of overshoot and response time. This result is due to the good trad-off between exploitation and exploration in WOA. However, Gray wolf's optimization algorithm starts with less fitness function in the first few iterations, which effect directly on the ripples in our model, due to the hierarchical search method that makes GWO algorithm powerful in terms of exploration. Table. 4 give the results of 4 executions of the proposed methods with the same global objective with different performances (ripple, overshoot and response time). The best results, which will be used in the rest of our study, are highlighted in bold.

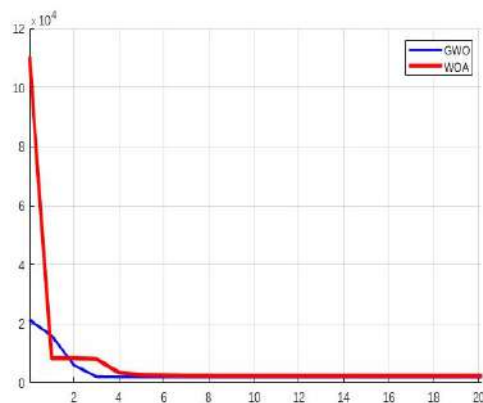


FIGURE 8: EVOLUTION CURVES OF THE FITNESS VALUE

Algorithm		M	y
GWO	Test 1	-0,0150	0,0010
	Test 2	- 0.0016	0,0000
	Test 3	-0,0190	0,0042
	Test 4	-0,0180	0,0052
WOA	Test 1	-0,0160	0,0049
	Test 2	-0,0200	0,0012
	Test 3	-0.005	0.0016
	Test 4	-0,0180	0,0013

TABLE 4: OPTIMUM SET OF CONTROLLERS GAINS

To illustrate the efficiency of the proposed GWO MPPT and WOA MPPT methods, a comparative study between fixed and variable step size MPPT have been illustrated. Three improvements have been demonstrated: a) tracking accuracy, b) ripple. c) overshoot. Figure 10 shows the GWO and WOA performance in training offline step. While figure 12 shows the obtained results using the trained optimal of fixed and variable step size MPPT.

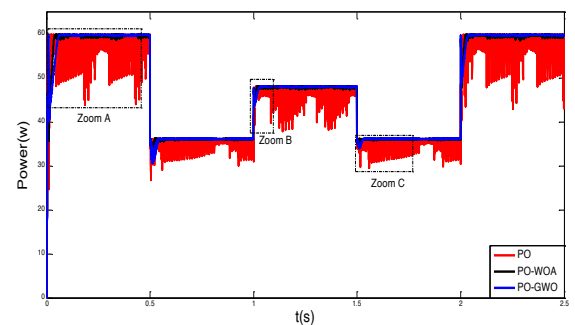
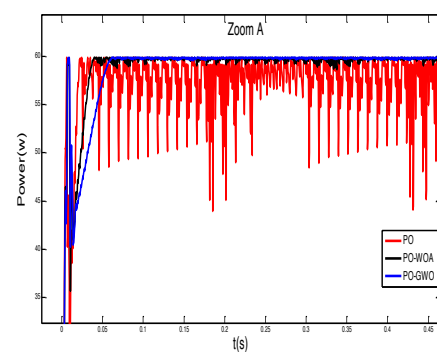
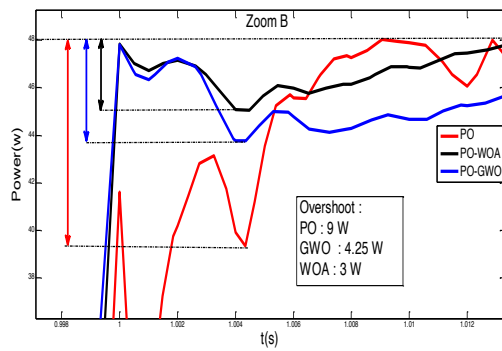


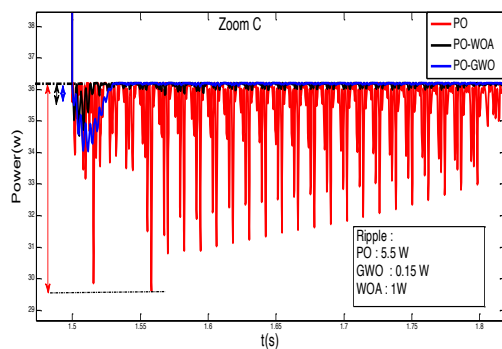
FIGURE 9: OUTPUT POWER OF MPPT USING GWO AND WOA



(a)



(b)



(c)

From simulation results we can see that:

- **MPPT tracking:** both fixed and variable step size MPPT algorithms have acceptable accuracy. The power values in both cases are very close to the theoretical value corresponding to irradiation levels (Fig. 12).
- **Overshoot:** the power peak value of the overshoot in case of suddenly changing atmospheric conditions is more important with the fixed value GWO/WOA MPPT compared to overshoot using the proposed variable step size GWO/WOA MPPT controller (Fig. 11.b);
- **Ripple:** The improvement of variable step MPPT method regarding ripple is undeniably clear. It can be observed that the quality of the output power P_{PV} (Ripple) with variable step GWO/WOA MPPT algorithm are obviously better than this with fixed step size MPPT algorithm (Fig. 11.c). variable step.

VII. CONCLUSION

In this paper a bio-inspired Optimization Algorithms (WOA and GWO) for MPPT controllers have been proposed, where the MPPT controllers are designed in two steps: The offline step used for testing and optimization fixed step size version which shows the same drawbacks of conventional MPPT methods characterized by slow convergence; oscillations in the PV power around the MPPT. To do simulation, we have used the Simulink

environment and we have detailed different aspects of the model design and parameters. The simulation results were divided into transient, steady state, and dynamic response. The reduction of the ripple, the overshoot and the response time as well as the good ability of the proposed PO-GWO/WOA algorithms to follows the MPPT point especially in fast changing environment conditions, which resulting in an overall reduction of lost energy, are the main contributions in this work. In addition, we plan to extend the present work

studying the biggest disadvantages of both proposed algorithms GWO / WAO when they used to optimize their parameters online and change the PID gains according the selected decisive criteria.

VIII. REFERENCES

- [1] D. Bonkougou, Z. Koalaga, and D. Njomo, "Modelling and Simulation of photovoltaic module considering single-diode equivalent circuit model in MATLAB," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 3, no. 3, pp. 493–502, 2013, doi: 10.3390/electronics2030192.
- [2] A. Safari, S. Mekhilef, "Simulation and hardware implementation of incremental conductance MPPT with direct control method using cuk converter," *IEEE Trans. Ind. Electron.*, vol. 58, no. 4, pp. 1154–1161, Apr. 2011.
- [3] V. Salas, E. Olías, A. Barrado, A. Lázaro, "Review of the maximum power point tracking algorithms for stand-alone photovoltaic systems". *Solar Energy Materials and Solar Cells*, 90(11), 1555-1578 (2006) .
- [4] T. Markvart, 'Solar Electricity', John Wiley & Sons, Chichester, 1994.
- [5] S. Leva, D. Zaninelli, Technical and Financial Analysis for Hybrid Photovoltaic Power Generation Systems, *WSEAS Transactions on Power Systems*, vol.5, no.1, May 2006, pp.831-838.
- [6] D. Rekioua , E. Matagne, Optimization of Photovoltaic Power Systems: Modelization, simulation and control , Springer-Verlag London 2012.
- [7] L. Bangyin, D. Shanxu, and C. Tao, "Photovoltaic DC-building-module-based BIPV system-concept and design considerations," *IEEE Trans.*
- [8] *Power Electron.*, vol. 26, no. 5, pp. 1418–1429, May 2011.
- [9] O. Wasynczuk, "Dynamic behavior of a class of photovoltaic power systems," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-102, no. 9, 1983, pp. 3031-3037.
- [10] D. Sera, R. Teodorescu, J. Hantshel, and M. Knoll, "Optimized maximum power point tracker for fast-changing environmental conditions," *IEEE Trans. Ind. Electron.*, vol. 55, no. 7, pp. 2629–2637, Jul. 2008.
- [11] J.-H. Park, J.-Y. Ahn, B.-H. Cho, and G.-J. Yu, "Dual-module-based maximum power point tracking control of photovoltaic systems," *IEEE Trans.Ind. Electron.*, vol. 53, no. 4, pp. 1036–1047, Aug. 2006.
- [12] Mei Qiang, Shan Mingwei, Liu Liying, Guerrero Josep M. A novel improved variable step-size incremental-resistance MPPT method for PV systems. *IEEE Transactions on Industrial Electronics* 2011;58(6):2424–2434
- [13] W. Xiao, W. G. Dunford, P. R. Palmer, and A. Capel, "Application of centered differentiation and steepest

- descent to maximum power point tracking," IEEE Trans. Ind. Electron., vol. 54, no. 5, pp. 2539–2549, Oct 2007.
- [14] N. Femia, D. Granozio, G. Petrone, and M. Vitelli, "Predictive & adaptive MPPT perturb and observe method," IEEE Trans. Aerosp. Electron. Syst., vol. 43, no. 3, pp. 934–950, Jul. 2007.
- [15] T. Noguchi, S. Togashi, and R. Nakamoto, "Short-current pulse-based maximum-power-point tracking method for multiple photovoltaic-and converter module system," IEEE Trans. Ind. Electron., vol. 49, no. 1, pp. 217–223, Feb. 2002.
- [16] A. Reza Reisi, M. Hassan Moradi, S. Jamasb, "Classification and comparison of maximum power point tracking techniques for photovoltaic system: a review. Renew Sustain Energy Rev 2013;19:433-443.
- [17] A. Pandey, N. Dasgupta, and A. K. Mukerjee, "High-performance algorithms for drift avoidance and fast tracking in solar MPPT system," IEEE Trans. Energy Convers., vol. 23, no. 2, pp. 681–689, Jun. 2008.
- [18] F. Liu, S. Duan, F. Liu, B. Liu, and Y. Kang, A Variable Step Size INC MPPT Method for PV Systems, IEEE Trans. Ind. Electron., vol. 55, no. 7, 2008, pp. 2622–2628.
- [19] N. Femia, G. Petrone, G. Spagnuolo, and M. Vitelli, "A technique for improving P&O MPPT performances of double-stage grid-connected photovoltaic systems," IEEE Trans. Ind. Electron., vol. 56, no. 11, pp. 4473–4482, Nov. 2009.
- [20] T. Esmar and P. L. Chapman, "Comparison of photovoltaic array maximum power point tracking techniques," IEEE Trans. Energy Convers., vol. 22, no. 2, pp. 439–449, Jun. 2007.
- [21] G. N. Femia, G. Petrone, G. Spagnuolo, M. Vitelli, "Optimization of perturb and observe maximum power point tracking method," IEEE Trans. Power Electron., vol. 20, no. 4, pp. 963–973, Jul. 2005.
- [22] Li G, Wang HA. Novel stand-alone PV generation system based on variable step size INC MPPT and SVPWM control. IEEE 6th International Power Electronics and Motion Control Conference, IEEE-IPEMC'09; 2009, p.2155e60.
- [23] F. Liu, Y. Kang, Y. Zhang, S. Duan, "Comparison of P&O and hill climbing MPPT methods for grid-connected PV converter," 3rd IEEE Conference on Industrial Electronics and Applications, 2008 (ICIEA 2008), pp.804-807, 3-5 June 2008.
- [24] N. A. Kamarzaman, C. W. Tan, "A comprehensive review of maximum power point tracking algorithms for photovoltaic systems", Renewable and Sustainable Energy Reviews, Vol(37), 2014, Pages 585-598
- [25] W. Xiao and W. G. Dunford, "A modified adaptive hill climbing MPPT method for photovoltaic power systems," in Proc. 35th Annu. IEEE Power Electron. Spec. Conf., 2004, pp. 1957–1963.
- [26] Ishaque, K. ; Salam, Z. ; Amjad, M. ; Mekhilef, S. An Improved Particle Swarm Optimization (PSO)-Based MPPT for PV With Reduced Steady-State Oscillation, IEEE Trans on Power Electron, Vol. 27, no. 8, pp. 3627 - 3638
- [27] Al-Diab A. and C. Sourkounis. 2010. Variable Step Size P&O MPPT Algorithm for PV System. 2010. Presented at the OPTIM 12th International Conference on Optimization of Electrical and Electronic Equipment. 1097–1102.
- [28] Kundur P., Morison G.K., Wang L « Techniques For On-Line Transient Stability Assessment And Control», Trans.s on Power Systems, 2000, pp. 46-51.
- [29] Chokri Ben Salah, Mohamed Ouali, "Comparison of fuzzy logic and neural network in maximum power point tracker for PV systems", Electric Power Systems Research, Vol. 81, 2011, pp. 43-50.
- [30] Chakraborty, K., Mukherjee, R. R., & Mukherjee, S. (2013). Tuning of PID controller of inverted pendulum using genetic algorithm. International Journal of Soft Computing and Engineering (IJSCE), 3(1), 21-24.
- [31] Dhimish, M., Holmes, V., Mather, P., et al.: 'Novel hot spot mitigation technique to enhance photovoltaic solar panels output power performance', Sol. Energy Mater. Sol. Cells, 2018, 179, pp. 72–79
- [32] Kamarzaman, N.A., Tan, C.W.: 'A comprehensive review of maximum power point tracking algorithms for photovoltaic systems', Renew. Sustain. Energy Rev., 2014, 37, pp. 585–598
- [33] Ghasemi, M.A., Foroushani, H.M., Parniani, M.: 'Partial shading detection and smooth maximum power point tracking of PV arrays under PSC's', IEEE Trans. Power Electron., 2016, 31, (9), pp. 6281–6292
- [34] Huang, C., Wang, L., Yeung, R.S.C., et al.: 'A prediction model-guided Jaya algorithm for the PV system maximum power point tracking', IEEE Trans. Sustain. Energy, 2018, 9, (1), pp. 45–55
- [35] Coelho, R.F., Concer, F.M., Martins, D.C.: 'A MPPT approach based on temperature measurements applied in PV systems'. 2010 IEEE Int. Conf. on Sustainable Energy Technologies (ICSET), Kandy, 2010, pp. 1–6
- [36] Karami, N., Moubayed, N., Outbib, R.: 'General review and classification of different MPPT techniques', Renew. Sustain. Energy Rev., 2017, 68, (1), pp. 1–18
- [37] Podder, A.K., Roy, N.K., Pota, H.R.: 'MPPT methods for solar PV systems: a critical review based on tracking nature', IET Renew. Power Gener., 2019, 13,(10), pp. 1615–1632
- [38] Rezk, H., Fathy, A., Abdelaziz, A.Y.: 'A comparison of different global MPPT techniques based on meta-heuristic algorithms for photovoltaic system subjected to partial shading conditions', Renew. Sustain. Energy Rev., 2017, 74, (6), pp. 377–386
- [39] Bhatnagar, P., Nema, R.K.: 'Maximum power point tracking control techniques: state-of-the-art in photovoltaic applications', Renew. Sustain. Energy Rev., 2013, 23, pp. 224–241

Compared methods of object tracking based on deep learning

Ramdane Taglout

Department of Computer Sciences, LIMPAF, University of
Bouira University of Bouira Bouira 10000, Algeria
r.taglout@univ-bouira.dz Bilal

Bilal Saoud

Electrical Engineering
Departement Bouira University University of Bouira Bouira ,
Algeria
bilal340@gmail.com

Abstract—Object trackers can find the relevant objects in moving images like videos. Object localization can be very useful in many fields. Many trackers have been proposed over years, some of them are based on correlation and deep neural networks. Object tracking is still a challenging problem due to many factors such as illumination, pose, scale, deformation, motion blur, noise, and occlusion. In this article, we study some methods based on correlation and deep learning. We have compared the following trackers ECO, SaimRPN, ATOM, DiMP, TRASFUST and TREG. These trackers have been developed based on deep neural networks and are very recent. Performances of trackers have been evaluated on OTB-100, UAV123, VOT 2019, GOT-10k and LaSOT dataset. Results prove the effectiveness of deep neural networks to cope up with object tracking in videos.

Keywords—Deep neural network, object tracking, benchmarks, trackers, simulation

I. INTRODUCTION

Today digital images are very used and motion analysis in videos has proven to be an important tool for many applications like video surveillance, video compression, medical imaging, robotics[1], human-computer interaction[2], analysis of sports footage...etc. Localization of an object in video is object tracking operation or process. It is a field of research, which is first works date back to the end of the 1980s and whose great progress has been made in recent years. Nowadays, the tracking of objects in a video sequence is ranked among the most active research topics. The central purpose of tracking is to estimate over time, the location of the target object in each frame of a video sequence.

Object tracking is a challenging problem and a complex task due to many factors related to the limitations of vision sensors (low frame rate, low resolution, low dynamic range per pixel, color distortions, noise, etc.), objects (non-rigid objects, number of objects varying over time, occlusions between objects, small object sizes, etc.), the application scenarios' requirements (real-time operation, high system reliability, etc.) and the environment (lighting variation, occlusion which is caused by the environment, etc.). In addition, trackers try to give a precise location of a target. In order to reach this goal, many object tracking methods have been proposed to cope with these challenges and to ensure good tracking quality. Furthermore, we can find some trackers use generator models [3] while others use discriminative models [4]. Finally, there is not a tracker which can be successfully applied to all scenarios.

Numerous research have been offered to perform computer vision using the developments in deep learning mechanisms. This has driven the implementation of deep learning algorithms for the tracking of a single object. Danelljan et al. [5], took a first step in reducing the feature space by considering the linear combination of raw deep features; nevertheless, the approach still cannot operate in real-time and the redundancy of deep features was not totally suppressed. SiamRPN[6] incorporates RPN into the SiamFC[7] in order to improve target bounding box accuracy. Then, ATOM[8] and DiMP[9] are currently between top trackers since they use the most advanced IoUNet for accurate object localization.

On the other hand, machine learning approaches such as transfer learning[10] and domain adaptation are commonly utilized to overcome these challenges. Adapting knowledge from one domain to another is the goal of this approach, which includes an additional offline learning step that utilizes a few instances from the target domain. In addition, Cui et al. [11] suggested Target Transformed Regression (TREG) to move a regression component from detection to tracking. They establish a pair between the target and the search area. Enhancing regression requires the resulting target. In this paper we have studied these trackers in order to locate objects in videos. A comparison between these trackers has been made according the results of trackers on some very used dataset. We have chosen to compare them on OTB-100[12], UAV123[13], VOT 2019[14], GOT-10k[15] and LaSOT[16] dataset, which are very challenging for trackers. The results that there is not a robust tracker for all scenarios and prove the need for new trackers. Our paper will be organized as follows; some concepts about object tracking will be presented in the second section. Then, we will present some trackers based on deep learning. The evaluation of these trackers will be illustrated in the fourth section with some well-known datasets. Our paper will be concluded by a conclusion and some perspectives.

II. BACKGROUND

The goal of visual object tracking is to locate a specific target in all frames of a video. This is accomplished via the use of tracking methods, which extract certain features from a template of target appearance and a search frame. After that, repeatedly match these features in order to identify the object. For the purpose of keeping the effective target templates, the appearance of the object in the first frame is considered as the initialization and is continually updated during the tracking procedure. A manual design and refinement approach is used throughout the overall tracking process, in contrast, to create the matching framework. The present trackers can typically be divided into two categories: generative techniques and discriminative approaches [17,18]. The search for areas that are the most similar to the tracked item is the focus of generative approaches, which include template-based, subspace-based, and sparse representation, to mention a few examples. Discriminative trackers, on the other hand, see tracking as a classification issue that separates the targeted objects from their immediate surroundings.

In order to tackle the tracking problem, several standard machine learning approaches have been tried, including boosting[19], support vector machine[20], naive Bayes[21], random forest[22], and so on.

Among the most efficient trackers we find trackers based on deep learning. These trackers used deep learning because it has proven that deep learning gives great results in image processing such as which image classification, object detection, semantic segmentation, image captioning, pose estimation, saliency detection, edge detection, and others. These techniques can be exploited in object tracking. Deep neural networks (DNNs) firstly proved their superior

learning skills on image classification, which was one of the primarily computer vision tasks on that they were tested. According to [23] deeper networks with more advanced designs [24] have been developed, resulting in improved classification accuracy. Ioffe et al. [25] present a Batch Normalization approach to reduce the problem of exploding/vanishing gradients. This method is capable of speeding up network training while also improving final performance by minimizing covariate shift. An additional research field in which DNN-based approaches have attained state-of-the-art performance has been the detection of objects. DNN-based detectors are often implemented in two steps, with the first step producing a number of candidate areas and the second step using DNNs to categorize them into the background or object categories. Fast R-CNN [26] creates the ROI pooling layer, that extracts features using shared convolutional feature maps, which enhances performance.

Object detection and visual tracking are fundamentally different in that object detection seeks to discriminate between objects belonging to distinct categories, whereas visual tracking is meant to find objects of interest in a manner that is agnostic to their classification. Yet, they are also very closely connected to one another. For example, several current visual tracking approaches [5] pre-train networks using object detection data sets, which is a common trend. Others [26] make use of item detection findings or region proposals to make online tracking more accurate.

III. TRACKERS DEEP VISUAL TRACKING

In this section, some well-known trackers will be presented. We have based on trackers developed based on convolutional neural network (CNN), Recurrent Neural Network (RNN), and other networks

A. CNN-based trackers

Because of its remarkable abilities in feature extraction and image classification, the CNN model [23] is an excellent choice for constructing robust appearance models in the tracking task. Besides traditional trackers, CNN-based tracking algorithms can be classified as either discriminative or generative, depending on how they are implemented. Using the CNN model, the discriminative technique attempts to conduct a binary classification of the tracked object in order to successfully identify it from its adjoining background.

The generative technique focuses on constructing a robust similarity metric that is used to effectively identify the object template inside a given search zone.

B. RNN-based trackers

When it comes to sequence modeling, the recurrent neural network (RNN) model is ideal since the neuron's output may be applied directly to itself in the next time. Numerous efforts have been made in visual tracking to exploit semantic information across spatial configurations and temporal connection among frames, which has been driven by research studies on speech recognition [27] or handwriting recognition [28]. For the production of confidence maps, Cui et al. [29] suggest an RNN-based

technique that they apply to a correlation filter. The RNN model is trained through four separate directions, which makes the appearance model more resistant to partial occlusion and robust.

C. Others networks trackers

Additionally, several researchers have sought to construct strong tracking algorithms utilizing other deep networks. According to Wang et al. [30], they have developed the first deep-learning-based tracker, which trains a stacked denoising autoencoder offline in order to learn generic image features. These characteristics have been shown to be robust with appearance changes. As a result of [30] work, Zhuang et al. [31] utilize an autoencoder network to generate deep feature representations and the shallow subspace model to overcome partial occlusions or other unexpected noises. In comparison to the classic DLT algorithm, this technique successfully mixes both shallow and deep models and provides better results.

IV. SIMULATION AND RESULTS

In this section we will present our evaluation of ECO[5], SaimRPN[6], ATOM[8], DiMP[9], TRASFUST[32] and TREG[11] trackers. First of all, benchmarks will be illustrated. Evaluation metrics will be shown for each benchmark. Then, results of trackers on these benchmarks will be presented.

A. Evaluation benchmark

There are many benchmarks to evaluate trackers. Among the most useful benchmark we find OTB-100, UAV123, VOT 2019, GOT-10k and LaSOT benchmark. In the rest of this section we are going to give a description of each benchmark.

1) OTB-100

The OTB-100 dataset provided by Wu et al. [12] has been frequently utilized in the assessment of online visual trackers for many years. The dataset contains 100 video clips annotated with several features, including Illumination Variation, Scale Variation, Occlusion, Deformation, Motion Blur, Fast Motion, In-Plane Rotation, Out-of-Plane Rotation, Out of View, Background Clutters, and Low Resolution (LR). For example, we may use the 11 criteria to determine how well trackers function.

Quantitatively, the 23 trackers' performance is measured using two metrics: distance precision (DP) (percent) and overlap success (OS) (percent) at a threshold of 20 pixels. Center location error (i.e., Euclidean distance between tracked target's center and ground truth's center) is a percentage that falls below a specified threshold in a given sequence. The OS value is determined by the percentage of monitored frames that were successful. As long as the tracking bounding box RT and the ground truth (RG) overlap by more than a pre-defined criterion (such as 0.5), the target is considered to have been tracked.

2) UAV123

Unmanned aerial vehicle tracking dataset UAV123 is a well-known aerial video-based tracking dataset that contains 123 video sequences, 115 of which were produced by a UAV platform and 8 of which were generated via the use of UAV simulation software.

UAV123 includes a variety of situations, including fields, streets, cities, suburbs, seas, and other locations. The dataset contains a total of 12 attributes: Aspect Ratio Change, Background Clutter, Camera Motion, Fast Motion, Full Occlusion, Illumination Variation, Low Resolution, Out-of-View, Partial Occlusion, Similar Object, Scale Variation, and Viewpoint Change. The attributes are as follows: Aspect Ratio Change, Background Clutter, Camera Motion, Fast Motion, Full Occlusion, Illumination Variation, Low Resolution, Out-of-View, Partial Occlusion.

1) VOT 2019

The vot 2019 dataset comprises 60 sequences, with each target identified by a bounding box. Across all videos, six attributes are annotated: illumination changes, motion changes, size changes, occlusion, camera motion, and an unassigned property. The VOT2019 benchmark is evaluated based on three criteria: robustness, accuracy, and expected average overlap. The accuracy metric estimates the average overlap that occurs during a successful tracking session.

The robustness, also known as the failure rate, is a measure of how often the tracker fails to locate the target (the case in which the overlap rate is zero). While the expected average overlap (EAO) criteria examine the overall performance of a tracker, which averages the IoU without requiring a reset operation, the expected average overlap (EAO) criterion analyzes the overall performance of a tracker. All of these are beneficial in providing insight about the behavior of a tracker.

1) GOT-10k

GOT-10k is a massive, high-diversity, one-shot tracking database with an unprecedentedly extensive coverage of real-world moving objects. It is the largest tracking database ever created. GOT-10k captures over 10,000 videos of 563 object classes and manually annotates 1.5 million tight bounding boxes. GOT-10k has collected over 10,000 videos of 563 object classes. It is also the first tracking dataset to be created using the one-shot approach, which is intended to encourage generality in tracker design.

1) LASOT

In the field of Large-scale Single Object Tracking, LaSOT is a high-quality benchmark. More than 3.5 million frames are included inside LaSOT's 1,400 sequences. Each frame in these sequences has been meticulously and personally annotated with a bounding box, resulting in LaSOT being the most densely annotated tracking benchmark available to us, to the best of our understanding. Every sequence in LaSOT is more than 2,500 frames long on average, and each sequence has a variety of challenges derived from the natural world, in which target items may disappear and re-appear in the view repeatedly.

B. Simulation and results

In this experiment, we gather 6 deep visual trackers, whose source code or benchmark results are already publicly available. These approaches have produced the best results on the selected benchmarks described above.

Fig. 1. Comparison over GOT-10K benchmark sequences

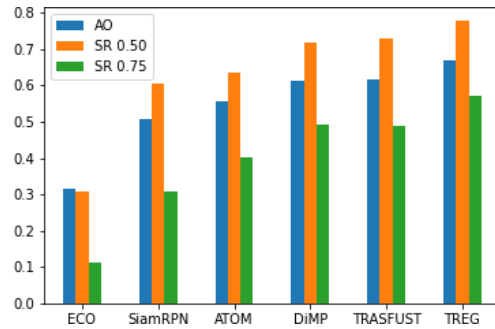
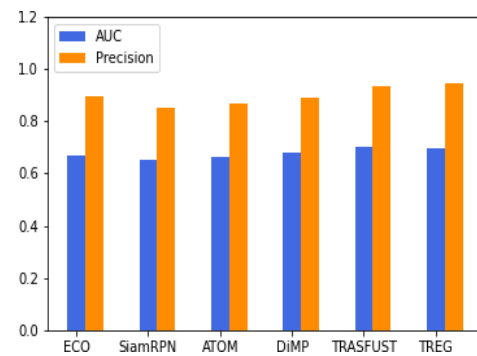


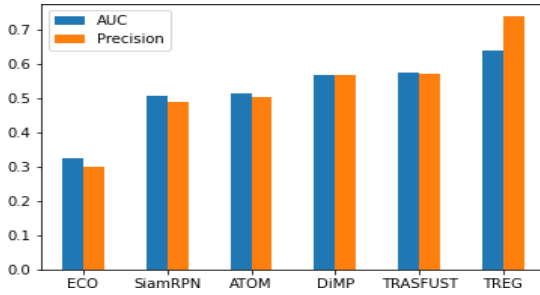
Figure 1 shows the obtained results for each tracker on GOT-10K benchmark based on OA, SR 0.5 and SR 0.75. From figure 1 we find that TREG has the best scores which is 0.668 in term AO metric and 0.778 , 0.572 respectively in term SR 0.50 and SR0.75. Furthermore, DiMP and TRASFUST have approximately the same results with some superiority of TRASFUST. However, ECO has the worst results on this benchmark.

Fig. 2. Comparison over OTB100 benchmark sequences



Trackers results on OTB100 are shown in figure 2. The evaluation metrics for this benchmark are AUC and Precision. All trackers have well performed in this benchmark, where they have scored more then 60% for AUC and more then 80% for Precision. It can be seen as a balanced performance across OTB100. We mentioned that TREG scores the highest value in term of Precision, which is which 0.945, and TRASFUST performs better by 0.701 in AUC.

Fig. 3. Comparison over LaSOT benchmark sequences



Evaluation of trackers on LaSOT are presented in figure 3. Results are different for each tracker, where ECO has the worst results in term of AUC and Precision metric. However, TREG has located objects with a great Precision and scored a good value of AUC. In addition, the rest of trackers, which are SaimRPN, ATOM, DiMP and TRASFUST, have approximately similar results in term of AUC and Precision.

Fig. 4. Comparison over UAV123 benchmark sequences

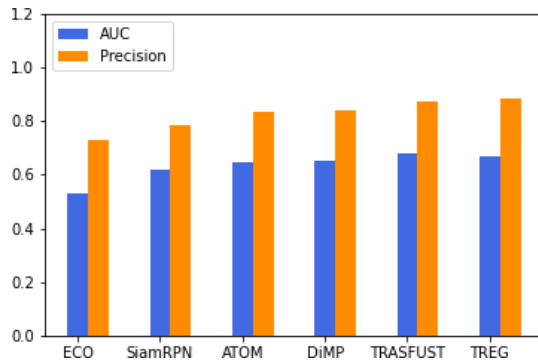
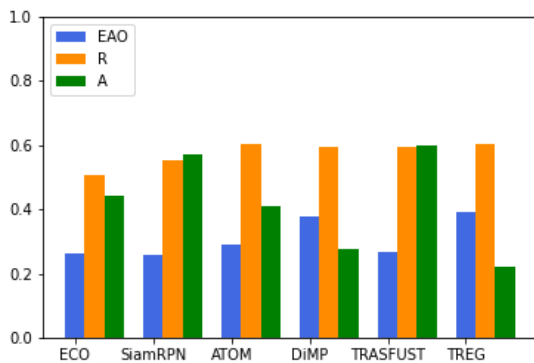


Figure 4 gives Success (AUC) and Precision score for tracker on UAV123 benchmark. TREG tracker obtains the highest results of AUC and Precision which 0.669 and 0.884 respectively. The rest of trackers have also scored acceptable results in term of AUC and Precision.

Fig. 5. Comparison over VOT2019 benchmark sequences



benchmark VOT2019. As shown from this figure, the robustness of TRASFUST tracker is the worst one.

However, TREG outperforms others in term of accuracy, robustness and EAO which are 0.221, 0.603, and 0.391 respectively.

V. CONCLUSION

In this paper, we have analyzed some recent visual object trackers based on deep learning. We have compared some trackers based on CNN, RNN and others networks. In order to conduct our comparisons, we have illustrated the most useful datasets (benchmarks). Where OTB-100, UAV123, LaSOT, GOT-10K and VOT2019 benchmarks have been presented. Then, we have used performances results provided by each method. From the results we can see that some trackers perform well in some benchmarks but not all the benchmarks, which prove the need to develop new tracker in order to improve the accuracy of localization in different scenarios.

Deep learning prove its performance and superiority in many domains and it can still be used in object tracking to improve performances of trackers by proposing new architecture or exploiting some existing architectures like VGG16, VGG19.

REFERENCES

- [1] Junliang Xing, Haizhou Ai, and Shihong Lao. Multiple human tracking based on multi-view upper-body detection and discriminative learning. In ICPR, 2010.
- [2] Liwei Liu, Junliang Xing, Haizhou Ai, and Xiang Ruan. Hand posture recognition using finger geometric feature. In ICPR, 2012.
- [3] Fernando, Tharindu, Simon Denman, Sridha Sridharan, and Clinton Fookes. "Tracking by prediction: A deep generative model for multi-person localisation and tracking." In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1122-1132. IEEE, 2018.
- [4] Bolme, D.S.; Beveridge, J.R.; Draper, B.A.; Lui, Y.M. Visual Object Tracking Using Adaptive Correlation Filters. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 2544–2550.
- [5] Danelljan, Martin, Goutam Bhat, Fahad Shahbaz Khan, and Michael Felsberg. "Eco: Efficient convolution operators for tracking." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 6638-6646. 2017.
- [6] Li, Bo, Junjie Yan, Wei Wu, Zheng Zhu, and Xiaolin Hu. "High performance visual tracking with siamese region proposal network." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 8971-8980. 2018.
- [7] Bertinetto, Luca, Jack Valmadre, Joao F. Henriques, Andrea Vedaldi, and Philip HS Torr. "Fully-convolutional siamese networks for object tracking." In *European conference on computer vision*, pp. 850-865. Springer, Cham, 2016.
- [8] Danelljan, Martin, Goutam Bhat, Fahad Shahbaz Khan, and Michael Felsberg. "Atom: Accurate tracking by overlap maximization." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4660-4669. 2019.
- [9] Bhat, Goutam, Martin Danelljan, Luc Van Gool, and Radu Timofte. "Learning discriminative model prediction for tracking." In Proceedings of the IEEE/CVF international conference on computer vision, pp. 6182-6191. 2019.
- [10] Dunnhofer, Matteo, Niki Martinel, and Christian Micheloni. "Tracking-by-trackers with a distilled and reinforced model." In Proceedings of the Asian Conference on Computer Vision. 2020.
- [11] Cui, Yutao, Cheng Jiang, Limin Wang, and Gangshan Wu. "Target transformed regression for accurate tracking." arXiv preprint arXiv:2104.00403 (2021).

- [12] Y. Wu, J. Lim, and M. H. Yang. Object tracking benchmark. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9):1834–1848, Sept 2015.
- [13] M. Mueller, N. Smith, and B. Ghanem. A Benchmark and Simulator for UAV Tracking, pages 445–461. Springer International Publishing, Cham, 2016.
- [14] Kristan, M., Matas, J., Leonardis, A., Felsberg, M., Pflugfelder, R., Kamarainen, J.K., ˇCehovin Zajc, L., Drbohlav, O., Lukezic, A., Berg, A. and Eldesokey, A., 2019. The seventh visual object tracking vot2019 challenge results. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops* (pp. 0-0).
- [15] Lianghua Huang, Xin Zhao, and Kaiqi Huang. Got-10k: A large high-diversity benchmark for generic object tracking in the wild. *arXiv preprint arXiv:1810.11981*, 2018.
- [16] Heng Fan, Liting Lin, Fan Yang, Peng Chu, Ge Deng, Sijia Yu, Hexin Bai, Yong Xu, Chunyuan Liao, and Haibin Ling. Lasot: A high-quality benchmark for large-scale single object tracking. In *CVPR*, 2019.
- [17] D.A. Ross, J. Lim, R. Lin, M. Yang, Incremental learning for robust visual tracking, *Int. J. Comput. Vis.* 77 (1–3) (2008) 125–141.
- [18] Z. Kalal, K. Mikolajczyk, J. Matas, Tracking-learning-detection, *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (7) (2012) 1409–1422..
- [19] S. Avidan, Ensemble tracking, *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (2) (2007) 261–271.
- [20] S. Avidan, Support vector tracking, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (8) (2004) 1064–1072.
- [21] A. Saffari, C. Leistner, J. Santner, M. Godec, H. Bischof, On-line random forests, in: *Proceedings of the IEEE International Conference on Computer Vision*, 2009, pp. 1393–1400.
- [22] K. Zhang, L. Zhang, M. Yang, Fast compressive tracking, *IEEE Trans. Pattern Anal. Mach. Intell.* 36 (10) (2014a) 2002–2015.
- [23] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: *Proceedings of the Advances in Neural Information Processing Systems*, 2012, pp. 1106–1114.
- [24] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S.E. 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.
- [25] S. Ioffe, C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift, in: *Proceedings of the International Conference on Machine Learning*, 2015, pp. 448–456.
- [26] R.B. Girshick, Fast R-CNN, in: *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1440–1448.
- [27] S. Kim, T. Hori, S. Watanabe, Joint ctc-attention based end-to-end speech recognition using multi-task learning, in: *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, 2017, pp. 4835–4839.
- [28] D.C. Ciresan, U. Meier, L.M. Gambardella, J. Schmidhuber, Convolutional neural network committees for handwritten character classification, in: *Proceedings of the International Conference on Document Analysis and Recognition*, 2011, pp. 1135–1139.
- [29] Z. Cui, S. Xiao, J. Feng, S. Yan, Recurrently target-attending tracking, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1449–1458.
- [30] N. Wang, D. Yeung, Learning a deep compact image representation for visual tracking, in: *Proceedings of the Advances in Neural Information Processing Systems*, 2013, pp. 809–817.
- [31] B. Zhuang, L. Wang, H. Lu, Visual tracking via shallow and deep collaborative model, *Neurocomputing* 218 (2016) 61–71.
- [32] Dunnhofer, Matteo, Niki Martinel, and Christian Micheloni. "Tracking-by-trackers with a distilled and reinforced model." In *Proceedings of the Asian Conference on Computer Vision*. 2020.

Deep Learning, Big Data and High Performance Computing to enhance Business and Marketing, while warying of Fake News: Challenges, Opportunities and Future Works

Zair Bouzidi
Laboratory LMA, Commercial
Sciences Dept
Faculty of Economy,
Bejaia University
Béjaia, 06 000, Algeria
zair.bouzidi@univ-bejaia.dz

Abdelmalek Boudries
Laboratory LMA, Commercial
Sciences Dept
Faculty of Economy,
Bejaia University
Béjaia, 06 000, Algeria
abdelmalek.boudries@univ-bejaia.dz

Mourad Amad
LIMPAF Laboratory
Computer Science Dept
Faculty of Science and Applied Science
Bouira University
Bouira 10 000, Algeria
mourad.amad@univ-bouira.dz

Abstract— The Gathered or deducted knowledge managed to be applied to multiple domains including notably social networks, forecasting, marketing, business, detecting fraud and financial time series prediction, using deep learning paradigms, constitute our major challenge. Many contents during events convey timely actionable and precious information. First responders, political and economic decision makers, and even the public can use certain information, as awareness, to gain insight into the situation as it unfolds. However, processing information from big data (including smart data and social media) to obtain such information, involves solving multiple challenges, notably parsing informal contents, handling information overload and prioritizing types of relevant information, while being wary of fake news.

Keywords—Big data, Business, Deep learning, Fake news, Marketing, Supercomputing

I. INTRODUCTION

Since recent years, social networks have become more and more ubiquitous in our daily lives. They have a potential resource to communicate, detect, track and extract information to improve the management of our daily life [1, 2], with the ability to share a message with a potentially large audience. However, the amount of information that is now released, especially during critical times (crisis situations), can make it difficult to locate precious, useful and usable information, while being wary of fake news [3, 4]. During the various crisis events in recent years, social media has enabled the population to quickly release a considerable amount of information about current events to the public: this will also serve to help managers make good and quick decisions [2]. That became standard practice for communicating information in a timely manner [1].

The link between social media and notoriety is the subject of increasing [6] attention from researchers. Many studies have been carried out in recent years [6] for the automatic identification of different types of content. With the proliferation of social media, an ongoing event is being discussed across all of these channels. To get a complete view of the event, even to avoid certain fake news, it is important to retrieve information from multiple sources that usually contain qualitative differences, and possibly even detect fake news. However, when it comes to retrieving information posted on social networks, the task is daunting.

The Web has several sources of information on which an ongoing event is discussed. Several automated systems have been designed to help managers identify and filter useful information from multiple sources [7–9]. Additionally, Twitter is usually the main source of information for these systems.

Data is often available from different sources, namely databases, log files, online web applications, and social media networks. Similarly, data, in the area of Bank and tax department, are generated from numerous sources, including laboratory experiments, accounting datasets, bank records, insurance and crisis data which are accessible online.

The rest of the paper is organized as follows. Chapter 2 contains background with the latest surveys and related works. In chapter 3, we begin this survey with raw data and their storage, big data and smart data. Chapter 4 illustrates an overview on supercomputing. A complete study on the different models of neural learning constitutes chapter 5. Chapter 6 relates all that is improvement of the economic world with all the advantages and very few disadvantages of big data, supercomputing and in particular of deep learning. We end with a conclusion and many future works.

II. BACKGROUND AND RELATEDWORKS

The table 1 shows many surveys for Big Data [10–15], and in particular in the field Smart Data and Social Media [10, 16–20]. In Deep Learning, we have [16,21–26]. Even the Parallel Computing Environment has been covered [13, 20, 27–31]. Some studies have also touched on the Evolution of FPGA [30], Human Computing and Machine Learning to Make Sense of Big Data [32], Decision Making and emerging Big Data [33], Quality of Social Media Data in Big Data [34] and Big Data privacy in public Social Media [35].

There are also studies that addressed data acquisition for Recent Neural Learning Approaches of Marketing, Business, Finance and Fraud detection, from Twitter only [36–38], Twitter & Facebook [7, 8], all the Web [7, 8] and Connected Objects (Smart Data) and All the Web [9, 19, 39–41].

Other studies have focused on Economic Tasks such as Detect novel frauds [42], Trading performance [43], Exchange rate prediction [44], Stock prediction [45], Trade on the stock market [46], Company stock prices [47], Forecasting of financial time series [48] and Fraud detection and Financial Time Series Prediction [41, 49, 50].

For fake news studies, we have review of Detecting and Fighting Fake News [3]. All of these studies only looked at work done only on Big Data, Deep Learning, and Supercomputing separately.

Only our investigation will cover, at the same time, the various models of Deep Learning, including the Capsule Network model and even the updated model of Spiking Neural Network with all the Big Data, Smart Data up to the models of connected objects to arrive at Supercomputing with all the models of Marketing, Corporate Marketing, Business strategies, Fraud detection and Forecasting of financial time series.

A. Recent Surveys

Table 1. Classification of Recent Surveys.

Tasks	Surveys
Deep Learning	
Deep Learning	[16,21-26]
Fake Account Detection in Social Media Using	[96]
Big Data Analytics	
Big Data	
Big Data	[10-15]
Smart Data & Social Media	[10,16-20]
Supercomputing	
Parallel Computing Environment	[13,20,27-31]
Evolution of FPGA	[30]
Human Computing and Machine Learning	[32]
Decision-Making and emerging Big Data	[33]
Quality of Social Media Data in Big Data	[34]
Big Data privacy in public Social Media	[35]
Economic Tasks	
Detect novel frauds	[42]
Trading performance	[43]
Exchange rate prediction	[44]
Stock prediction	[45]
Trade on the stock market	[46]
Company stock prices	[47]
Forecasting of financial time series	[48]
Fraud detection& Financial Time Series Prediction	[41,49,50]
Deep Time-Series Clustering	[51]
Fake news	
Detecting and Fighting Fake News	[3, 52-57]
Detecting Fake News Spreaders	[58]
Study of Detecting and Fighting Fake News with Big Data, Deep Learning and Supercomputing	Our Survey

B. Related Works

The next subsections provide instructions on how to insert figures, tables, and equations in your document. The table 1 shows many surveys for Big Data [13-18], and in particular in the field Smart Data and Social Media [13,19-

23]. In Deep Learning, we have [19,24-29]. Even the Parallel Computing Environment has been covered [16,23,30-33]. Some studies have also touched on the Evolution of FPGA [33], Human Computing and Machine Learning to Make Sense of Big Data [34], Decision-Making and emerging Big Data, Quality of Social Media Data in Big Data and Big Data privacy in public Social Media [35].

There are also studies that addressed data acquisition for Recent Neural Learning Approaches of Marketing, Business, Finance and Fraud detection, from Twitter only [36-38], Twitter & Facebook [4], all the Web [5] and Connected Objects (Smart Data) and All the Web [22,39-41].

Other studies have focused on Economic Tasks such as Detect novel frauds, Trading performance, Exchange rate prediction, Stock prediction, Trade on the stock market, Company stock prices, Forecasting of financial time series and Fraud detection and Financial Time Series Prediction [41].

All of these studies only looked at work done only on Big Data, Deep Learning, and Supercomputing separately. Only our investigation will cover, at the same time, the various models of Deep Learning, including the Capsule Network model and even the updated model of Spiking Neural Network with all the Big Data, Smart Data up to the models of connected objects to arrive at Supercomputing with all the models of Marketing, Corporate Marketing, Business strategies, Fraud detection and Forecasting of financial time series.

III. FROM RAW DATA TO SMART DATA VIA BIG DATA

Data ingestion is the process of collecting raw data from various silo databases or files and integrating it into a data lake on the data processing platform, e.g., Hadoop data lake. Today, several data ingestion tools are available for ingesting a variety of data onto Hadoop.

Data processing platforms are Data Lake and Data Warehouse where the first one is a storage repository that holds a huge amount of raw data in its native format. It has the schema-on-read characteristic. It is thus more agile as data can be easily configured and reconfigured following different models during the analysis, as Hadoop but the data structure and requirements are not defined until the data is used and typically stores data using a flat architecture. While the second one is a storage repository that holds a huge amount of raw data in its native format.

A. Big Data

Big Data refer to all the digital data produced by the use of new technologies for personal or professional purposes, including [12, 32, 35] corporate data (emails, documents, databases, business processor history, etc.) as well as data from sensors, content published on the web (images, videos, sounds, texts), e-commerce transactions, exchanges on social networks, data transmitted by connected objects (electronic labels, smart meters, smartphones, etc.), geolocated data, etc.

They include information published on social networks (age and contact details, reviews, comments, photos, videos), Web browsing (through the famous cookies) and online purchases. Massive data duplication is one of the keystones of the big data architecture. They provide very interesting clues about consumer behavior and market trends, making it possible to know your profile, but also your overall behavior (frequency of use of social networks and of your online purchases (history of transactions, expenses made), channels used, hours of connection, etc.). Cloud computing, hybrid supercomputers (high performance computing or HPC) and distributed file systems (DFS) are among the main storage models currently available. Big data is an essential tool for BtoB and BtoC companies. The data collected helps them to design personalized marketing campaigns adapted to the needs, preferences and behavior of consumers. This information helps improve the customer experience, attract prospects and retain existing customers. Improved targeting makes marketing campaigns more effective and reaches the desired segment, the one that is most likely to be interested in the company's products and / or services. Big data is also a competitive advantage for professionals who hold a multitude of data, because they can anticipate changes in behavior and better understand why consumers have turned to a particular service provider. Big data is a valuable tool in vast private and public fields ranging from online sales to scientific research, including culture, politics (electoral campaigns), transport, insurance, industry, the banking sector and energy.

Table 2. Comparative table of all techniques and methods used in Models.

Ref	Identification Methods	Used Methods
[7,8]	Based on Artificial Neural Network	Social Media (Twitter & Facebook)
[9]	Based on FeedForward Neural Net	Social Media (All the Web)
[41]	Based on Hybrid of Deep CNN-LSTM (DeepCNN-LSTM)	Social Media (All the Web)
[39]	Based on Recurrent Neural Net	Smart Data & Social Media (Entire Web)
[40,59]	Based on LSTM	Smart Data & Social Media (Entire Web)

Online Listening Tools tracking automatically Channels to collect Contents are collected from all online channels tracked automatically by the Online Listening Tool, namely Radian6 or any its competitor, such as Awario, Brand24.com, Brandwatch, Mention, Keyhole, Socialert.net, SocialPilot.co, Simplify360, etc. from websites to social media, such as Twitter, Facebook, LinkedIn, Instagram, Google+, Youtube and so on [17, 60]. Actually, many networking platforms allow access to their data via Application Programming Interface (API) [16].

B. Smart Data

Knowledge is thus generated from the data, based on technologies making complex processing and generating knowledge, a source of added value to become the foundation of the new data economy [13, 29]: that's the reason for which big data must be transformed into smart data. It is the information produced every day, in all situations of life (at workplace, doctor, behind the wheel of car or baker, via smartphone, smartwatch or computer). It can be textual data, images or sensors. If this data can be analyzed, processed and used, we can really speak of it as raw material. These masses of data, big data, contain immense potential. Streaming is a way of broadcasting and reading streaming content (widely used on the Internet). It involves being connected to internet server, opposed to the file download requiring to recover all the data of a file.

Table 3. Smart Data application areas.

No	Smart Data Application Areas	Smart Data Application
1	Financial services	Fraud detection and prevention
2	Retail	Allowing brands to: <ul style="list-style-type: none"> - Analyzing the sentiment of their customers, and - Offering personalized and contextual promotions
3	Telecommunications industry	Possibility of <ul style="list-style-type: none"> - Better allocating bandwidth based on real-time needs, and - diagnosing the condition of antennas
4	Manufacturing	Preventive maintenance
5	Healthcare	- Monitoring patient vital signs and - Reducing readmission rates
6	Oil industry	- Proactive repairing infrastructures, and - Balancing the power delivered according to consumption
7	Public sector	- Detecting and preventing intrusion attempts on the network, and - Predicting the risk of epidemics
8	Transport sector	Possibility of detecting risky conduits

Smart data is based primarily on real-time data analysis. It refers to an approach to data analysis analyzing directly data at the source, as a different concept of Big Data and without the need to transmit it to a centralized system. Big and smart data are transversal technologies: will impose themselves in all areas of life by, changing them for life. Data are collected, converted, placed in databases and processed in waves, but usually out of date once analyzed. They refer to the analysis of data directly at the source to enable immediate decisions, without sending them to a centralized system. Processing time is reduced to just seconds. While Smart referring to the intelligence, thanks to

processing less data (statistical models are responsible for determining the relevant variables, even most correlated). It enables to control and qualify data: used for establishing a coherent marketing strategy. Generating leads to increase sales in building customer loyalty by providing optimal experience and satisfaction, unlike Big Data.

The possibilities offered by Smart Data are therefore numerous. Table 3 shows some among Smart Data application areas. Virtually all areas of society and the economy can benefit from the intelligent use of data. Big data technologies are already widely used in finance and marketing. Data scientists, marketers, and manufacturers are experimenting with ways to use edge computing and smart data devices to generate more revenue, improve decision-making, and spot equipment issues before they go. Smart Data can also be useful in many industries.

IV. SUPERCOMPUTING

Supercomputing is now as a continuum of computing. There is definitely a big need for high capacity systems to enable applications that need precision and capacity. For these high-capacity models to work it needs composable systems that can permit to compose a solution by automating workflows, using all models of supercomputing. That means supercomputing is becoming even more useful to research, among others. The term supercomputing refers to the processing of massively complex or data-laden problems using the concentrated compute resources of multiple computer systems working in parallel (i.e. a supercomputer). Supercomputing involves a system working at the maximum potential performance of any computer. Sample use cases include weather, energy, life sciences and manufacturing, health care, social networks, mining, recommendation systems, image processing, pattern recognition and all the predictions.

An information technology forecasts that supercomputing will have a significant impact on the data processing operations of corporations and be a competitive necessity actually [31]. In current supercomputers, storage is typically provided by parallel distributed file systems for hot data and tape archives for cold data.

These file systems are often compatible with local file systems due to their use of the POSIX interface and semantics, easing development and debugging because of easily running both on workstations and supercomputers. There is a wide variety of file systems, each tuned for different use cases and implementing different optimizations. However, the overall application performance is often held back by I/O bottlenecks due to insufficient performance of file systems or I/O libraries for highly parallel workloads. Performance problems are dealt with using novel storage hardware technologies as well as alternative I/O semantics and interfaces. These approaches have to be integrated into the storage stack seamlessly to make them convenient to use.

Table 4. Comparative table of all models using Supercomputing.

Models	Identification Usage of Supercomputing
[20, 13, 27–31]	Parallel Computing Environment
[30]	Evolution of FPGA : FPGAs inherent reconfigurable nature & low power consumption
[32]	Human Computing and Machine Learning to Make Sense of Big Data
[33]	Decision-Making and emerging Big Data
[34]	Quality of Social Media Data in Big Data
[35]	Big Data privacy in public Social Media

Upcoming storage systems abandon the traditional POSIX interface and semantics in favor of alternative concepts such as object and key-value storage; moreover, they rely on new technologies such as Non-Volatile Memory (NVM) emulation and burst buffers to enhance performance. Additional tiers of storage will increase the importance of hierarchical management. These changes will be disruptive and require developers to modify their approaches to data management and I/O. A thorough understanding of today's storage infrastructures, is crucially important for designing and implementing scalable storage systems suitable for needs of exascale computing [28]. This work integrates the core ideas of deep learning and its applications in application domains of Marketing corporate, Business Strategies, Fraud detection and Financial time series prediction, to be accessible to all scholars and researchers [28]. Field Programmable Gate Arrays (FPGAs) inherent reconfigurable nature and their low power consumption have become so complementary to microprocessors: they are included in few mainstream computer systems for accelerating application specific performance [27, 30].

Deep Learning and Parallel Computing Environment for Bioengineering Systems delivers a significant forum for the technical advancement of deep learning in parallel computing environment across bio-engineering diversified domains and its applications.

V. DEEP LEARNING

In recent years, Deep learning methodologies have achieved impressive results in computer vision [61], speech recognition, image processing, and handwritten recognition of characters. They are beginning to interest the economic world with some very important applications whether in marketing, commerce, fraud and especially finance. While they are currently in their infancy in fault diagnosis [26].

It can also overcome limitations of shallow networks that prevent efficient training and abstractions of hierarchical representations of multi-dimensional training data [26] in many instances. Deep learning offers a set of units such as [62] convolution unit, [63] recurrent unit, and long-short term memory unit for feature extraction on samples with distinct features, delving into the math behind training algorithms used in recent deep networks, explaining existing shortcomings, [26] improvements and implementations.

5.1 Discussions about Deep Learning Models Neural

learning is carried out by Feedforward or Feedback neural network.

Table 5. Classification of Social Media-based Recent Neural Learning Approches of Marketing, Business, Finance and Fraud detection.

Social Media-based Approaches	References
Twitter	[36–38]
Twitter & Facebook	[7, 8]
All the Web	[9, 19]
Connected Objects (Smart Data) and All the Web	[9, 19, 39–41]

In Feedforward, we have supervised learning such as Feedforward neural network itself for classification [9], convolutional neural network [36, 37] for image recognition/ classification or Residual neural network (ResNets) [64] for image recognition, and unsupervised learning such as Autoencoder [63] for Dimensionality reduction and encoding, Generative Adversarial Network [38] network for generating realistic fake data, reconstruction of 3D models or image improvement and with supervised or unsupervised learning such as Restricted Boltzmann machine [63] for dimensionality reduction, feature learning, topic modeling, classification, collaborative filtering or many body quantum mechanics. In Feedback, we have only supervised learning such as Recurrent neural network [39, 40, 59, 63] for sequences recognition as precise timing, Bidirectional Recurrent Neural Network [38] for natural language processing (NLP), Long Short-Term Memory [63] for temporal data such as stock market values over a period of time and video frames, Fully Connected-LSTM [65] for learning non-linear and complex processes in hydrological or meteorological modeling and Bi-Directional-LSTM [37] through time-natural language processing and language translation. Neural learning can be trained in either supervised/unsupervised ways by Radial Basic Function Network [66] for M-means clustering, Least square function, function approximation and time series prediction or unsupervised ways by Kohonen Self Organizing Network [66] for dimensionality reduction, optimization problems or clustering analysis.

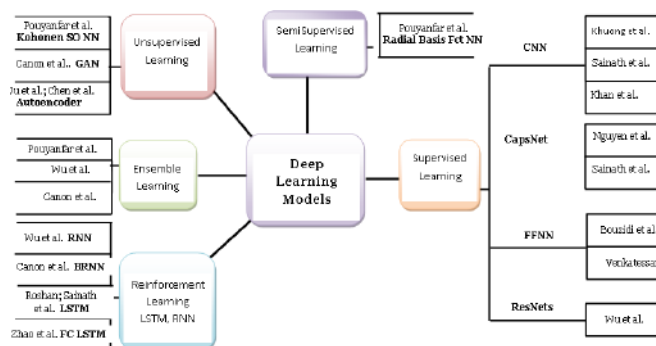


Fig. 2: Classification of Deep learning Models.

While feedforward neural networks are not capable of extrapolation [67], ResNets have increased architectural complexity as a drawback. They are highly dependent on implementing batch normalization layers, as well as adding jump level connections where dimensionality is essential between the different layers. As a CNN variant designed for computer vision image classification tasks, ResNets is proposed to overcome the problems of VGG-style CNNs: stacking convolutional layers (CNNs) to make the model deeper does not guarantee increased the precision of the validation. The ResNet architecture was introduced to ignore connections: by increasing the number of convolutional layers, the generalizability and accuracy of validation decreases. CNN does not encode object position and orientation and lacks ability to be spatially invariant to input data: capsule neural networks (CapsNet) are introduced to solve this CNN problem with the dynamic routing algorithm [68]. Autoencoders are based on deep learning regression for the discovery of slow modes in dynamical systems. However, a rigorous analysis of nonlinear autoencoders is still lacking: they cannot properly discover slow [69] modes. Spiked neural networks are computationally more powerful than other neural network models, in terms of the number of neurons needed [70]. Abiodun et al. (2018) [21] recommend that future research can focus on combining various deep learning models in a single network-scale application, depending on our needs and the characteristics of the different models of deep learning.

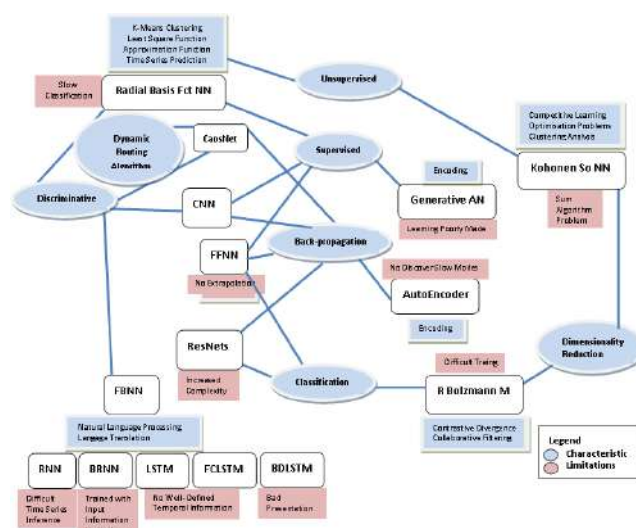


Fig. 1: Classification of Deep learning with features and limitations.

VI. FAKE NEWS

The concept of fake news is defined as misleading content including conspiracy theories, rumors, clickbaits, fabricated news, and satire [57] (See Figure 3). The spread

of fake news has become a global issue that needs to be attended immediately [71].

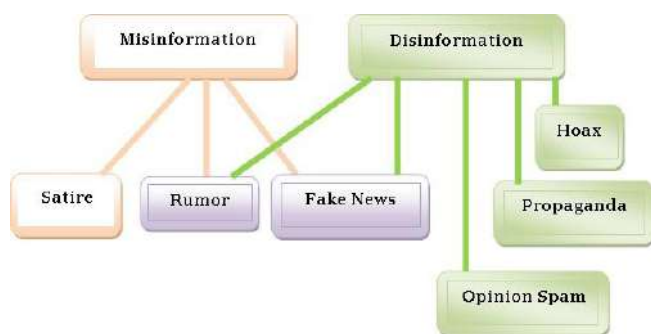


Fig. 3: Components of Online Fake News.

This concept drew major attention after the US elections of 2016 [72]. Fake news is defined by as misleading content including conspiracy theories, rumors, clickbaits, fabricated news, and satire [57]. It is defined as misinformation and disinformation both, including false and forged information, that is spread on purpose to mislead people or to fulfill a propaganda. It is considered as a vehicle of purposely targeted fabricated news spread to affect the cognitive activities of a user through user-content interaction by indirectly affecting his unconscious behavior. This unconscious behavior, can further strengthen confirmation bias among users and aid in further spread of fake news, notably humans have always been attracted to sensationalism and controversies [73]. It is the case of the recent spread of false information about COVID 19 vaccines (and dangerous scientific treatment methods) [74], political smear campaigns during elections. In order to clearly understand the spread of fake news, it is important examining components that can be divided into four main categories namely creator/spreader, target victims, content and social context [75] (See Figure 4).

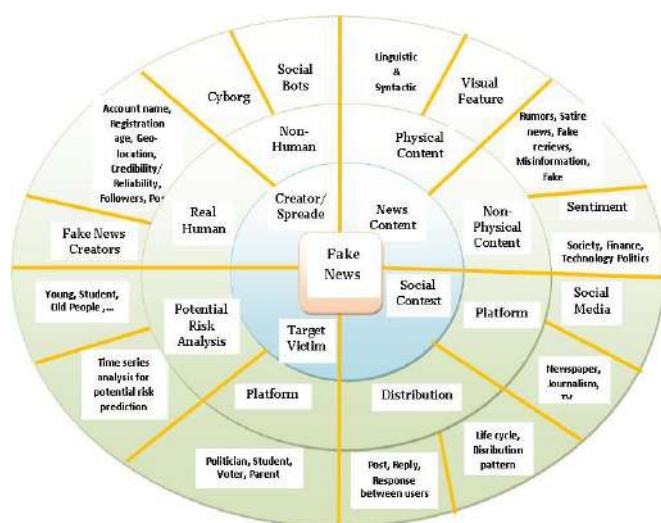


Fig. 4: Characteristics of Online Fake News.

Figure 2 shows an overview of online fake news, beginning by creator/spreader, News Content, Social Context and Target Victims. Creator/Spreader may be a human (as Design Author or Publisher) or Non-human (Bot or Cyborg). New Content may be Physical (headline, body text or image/video) or Non-physical (sentiment, rumor or new topic). Social context may be platform (social media or main streaming) or distribution (Community of users or Broadcast pattern). Target Victims may be online or main streaming users, young, students, temporal-based analysis.

1.1 Types of Fake News

Fake news may be stance, satire, multi-modal, deep fake, and disinformation. Stance can be classified into four types, (agree, disagree, discuss and unrelated) [77]: stance (headline of fake news), whereas a disagreeing stance (contradictory information), Satire (humor and mockery) [78] and Multi-modal (spread of fake news) [79] using multiple means (videos, images, audio, text etc.). Deep fake [80] is spread through manipulated video clips, images and recordings (generated by using deep learning techniques). Disinformation is misleading and false information to deceive people (political agendas or creating social havoc).

1.2 Features Used in Detecting Fake News

Uncovering and recognizing fake news and developing automatic fake news detection systems are quite a challenging task as newer methods are usually came up as new solutions [52] to mislead people. The key features in detecting fake news are user-based features, temporal analysis, sentiment features, linguistic analysis, social context-based analysis, and network features.

- User-based features include unique characteristics of users' profiles that can be analyzed to find out if the person is a fake news spreader. They are divided into profile analysis (based on username, age, profile picture, Geo-location and account verification status [81]) and credibility analysis (number of friends and followers): bot accounts have more followers but follow less of them [53].
- Temporal analysis (timing and frequency of posts, user engagement) helps to identify bot accounts [54]: Bot accounts have a more active engagement.
- Sentiment features (triggering emotional response): Bot accounts use out-of context misleading facts (highly polarized and exaggerated) to provoke emotions [3, 4].
- Linguistic analysis involves determining the writing patterns and formats [82]: the excessive use of bold letters in headings and paragraphs, presence of suspicious tokens such as URLs, tags, and excessive uppercase words.

- Knowledge-based analysis involves verifying suspicious information from credible resources such as using verified websites (manually or by using AI algorithms).

- Social context-based analysis includes user network analysis (engagement patterns between online accounts) and distribution network analyses (distribution of information) [55].

- Network features have two types of networks analyzed: Homogeneous networks and heterogeneous networks. Homogeneous networks include stance networks (users' stance on a specific idea or news, with classifying as agreement or disagreement between the main headline and body of the news) and propagation networks [56] (relationship between posts and re-posts): fake news gets re-posted excessively. Heterogeneous networks involves analyzing relationships between multiple nodes (articles, publishers, users and posts) [57].

1.3 Fake News Dissemination Studies Based on Spreaders Accounts Features

Fake News Spreads are categorized based on three categories, namely source, propagation and target based features, as in Figure 5.

Source-Based Account Detection : A source of fake news can be a human, bot or cyborg. Features of their source are classified into three main categories, namely: feature of personality, historical and of credibility.

Propagation-Based Accounts Detection A propagator, disseminating fake news widely to increase its reach to maximum victims [60], has features classified into three main categories, namely user engagement, time dynamics and platform-based features.

Target-Based Accounts Detection The target features identify end users that are affected by the fake news. A target can be a human, bot or cyborg depending on the nature and domain of fake news. Although fake news can reach almost all the users through social media, an easy target will be those people that are more vulnerable and prone to get influenced by the fake news [76]. Victim dynamics mean thoroughly understanding the details of the end user. The details can include age, gender, education history, account creation history, network of followers, location etc. Generally new users with limited exposure to social media are targets of fake news spreaders, as they tend to believe anything presented to them due to lack of exposure [83], as well as Teenagers and aged people with limited knowledge of possibilities of fake news on social media are an easy target. Similarly, people with low qualifications and coming from rural areas are more prone to be the victims of fake news [84].

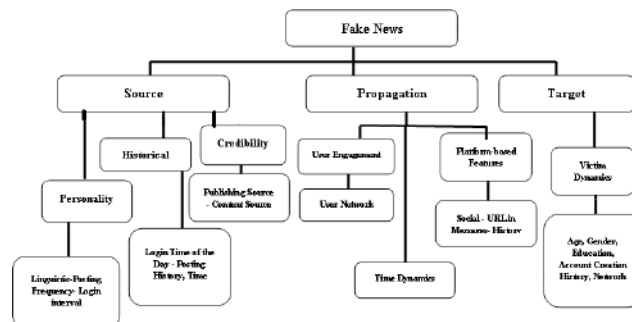


Fig. 5: Spreaders Accounts Features-based Dissemination of Fake News

VII. ECONOMIC TASKS: MARKETING, BUSINESS STRATEGY, FRAUD DETECTION AND FINANCIAL TIME SERIES PREDICTION

With a wide variety of products and buying behaviors, the shelves, on which products are presented, are the most important resources in the retail environment. Retailers can not only increase profits, but also reduce costs by properly managing shelf allocation and product display [85]. Using learning models in the organization of shelves in supermarkets by grouping products that are usually bought together; we can extract the following relation: customers who buy the product X at the end of the week, during the summer, generally also buy the product Y.

There is also the credit institution that permits to decide whether or not to grant credit based on the applicant profile, his / her request, and past loan experiences. The overbooking (optimization of the number of seats in planes, hotels, etc.), the targeting of offers (organization of advertising campaigns, promotions) and the analysis of business practices, strategies and their impact on sales. This knowledge, unknown at first, may be correlations, patterns, or general trends in that data. Table 1 shows a comparison of works with various economic tasks. Experimental data are necessary to verify the correction of the system or the estimation of some difficult parameters to mathematical modeling. Our purpose is to discover, in large amounts of data, valuable information that can help understand the data or predict the behavior of future data. We can use several tools for statistics and artificial intelligence to achieve its objectives: an essential component of Big Data technologies, large data analysis techniques and recently data smart streaming.

Table 6. Comparative table of all economic tasks used..

Economic Tasks	References
Detect novel frauds [42]	[42]
Trading performance	[43]
Exchange rate prediction	[44]
Stock prediction	[45]
Trade on the stock market	[46]
Company stock prices	[47]
Forecasting of financial time series	[48]
Fraud detection and Financial Time Series Prediction	[41,49,50]
Deep Time-Series Clustering	[51]

It is often defined as the process of discovering new knowledge. By examining large amounts of data (stored in warehouses or in streaming) using pattern recognition

technologies as well as statistical and mathematical techniques used for the study, we describe the method used for the data collections.

There are also other specific tools, such as the model where some sequence exploration problems lend themselves to the discovery of frequent item sets and their order, traditionally, used in marketing applications to detect patterns among competing elements in large transactions, for example, by analyzing customer shopping basket transactions in a supermarket, according to Han et al. [86].

The models can have many real applications because the data is encoded as sequences in many areas such as bio-informatics such as genomics and Proteomics [87]. We also see the development of market basket analysis which consists of studying sales (Sales receipt analysis) [23].

Thanks to the huge amount of information available online, the World Wide Web is a fertile ground for research. Its research is at the crossroads of research conducted by several research communities, such as databases, information retrieval and within Artificial Intelligence (AI), particularly the sub-domains of learning and natural language processing [22]. We can use the processing of text corpora to analyze the content and extract knowledge. The main tasks to be accomplished are the recognition of the information presented in the document and its interpretation.

VIII. CONCLUSION

This paper exemplifies the comprehensive study of data, of all data: starting from raw to Smart data to Big Data, while emphasizing all economic tasks. This study is supplemented by a overview of supercomputing with a close look at Deep Learning, its evolution and much more about the future with Deep Learning.

In this review we have not forgotten the social networks with their advantage with its popularization in all the tasks of the daily life of all the populations, but above all the weight of the fake news that they convey: where the truth will be more more difficult to decipher, recognize, read, and so on, that will be our research perspectives and future works.

ACKNOWLEDGMENT

The preferred spelling of the word —acknowledgmentl in America is without an —el after the —gl. Avoid the stilted expression —one of us (R. B. G.) thanks ...l. Instead, try —R. B. G. thanks...l. Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

- [1] Olteanu, A., Vieweg, S., & Castillo, C. (2015), What to Expect When the Unex pected Happens, Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15. doi:10.1145/2675133.2675242
- [2] Bouzidi Zair (2022), Social networks for Disaster management, Classical Doctorate Thesis, University of Bouira, Algeria, February 06th, 2022
- [3] Bouzidi Z., Boudries A. & Amad M., (2022b), Enhancing Detecting and Fighting Fake News with Big Data and Deep Learning: Challenges, Opportunities and Fu ture Works, 6th International Conference on Information System and Data Mining, ICISDM 2022, Silicon Valley, CA, United States, May 27-29, 2022, www.icisdsm.org, Accepted
- [4] Bouzidi, Z., Boudries, A. and Amad, M., (2022c), Big Data, Deep Learning and Supercomputing to Improve Financial Time Series Forecasting, Fraud Detection, Business and Marketing: Survey, First edition of the International Colloquium on Marketing Management COLI'MM22, Algiers, Algeria, Accepted
- [5] Imran, M., Castillo, C., Lucas, J., Meier, P., & Rogstadius, J. (2014b) Coordinating human and machine intelligence to classify microblog communications in crises. Pro ceedings of (ISCRAM) The 11th International Conference on Information Systems for Crisis Response and Management, (ISCRAM 2014), pp. 712-721, The Pennsylv ania State University, Pennsylvania, USA
- [6] Vulic, I. and Moens, M.-F., (2015), —Monolingual and Cross-Lingual Information Retrieval Models Based on (Bilingual) Word Embeddings—, SIGIR'15 Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, Pages 363-372, Santiago, Chile — August 09 - 13, 2015
- [7] Bouzidi, Z., Boudries A. & Amad M. : A New Efficient Alert Model for Disaster !! Management, in Proceedings of Conference AIAP'2018 : Artificial Intelligence and Its Applications, El-Oued, Algeria
- [8] Bouzidi, Z., Amad, M. & Boudries, A. : Intelligent and Real-time Alert Model for Disaster Management based on Information retrieval from Multiple Sources—, International Journal of Advanced Media and Communication. 7(4), 309-330, doi:10.1145/253260.253325 (2019)
- [9] Bouzidi, Z., Boudries, A. & Amad, M.: Towards a Smart Interface-based Auto mated Learning Environment Through Social Media for Disaster Management and Smart Disaster Education. Advances in Intelligent Systems and Computing. SAI 2020. 1228. Springer, Cham, 443-468. https://doi.org/10.1007/978-3-030-52249-0_31 (2020)
- [10] Saroj, Anita & Pal, Sukomal, (2020), Use of social media in crisis man agement : A survey, International Journal of Disaster Risk Reduction, doi:10.1016/j.ijdr.2020.101584
- [11] Emmert-Streib, F., Yang, Z., Feng, H., Tripathi, S. and Matthias, D., (2020), An Introductory Review of Deep Learning for Prediction Models With Big Data, doi:10.3389/frai.2020.00004
- [12] Saranya, M. and Prema, A., (2017), Survey On Big Data Analytcis using Hadoop ETL, International Journal of Computer Trends and Technology (IJCTT), vol. 48, No. 1, doi:10.14445/22312803/IJCTT-V48P105
- [13] Alexandrov, A., Bergmann, R., Ewen, S., Freytag, J.-C., Hueske, F., Heise, A., Kao, O., Leich, M., Leser, U., Markl, V., Naumann, F., Peters, M., Rheinl'ander, A., Sax, M. J., Schelter, S., Hoger, M., Tzoumas, K. and Warneke, D., (2014), The stratosphere platform for big data analytics, The VLDB Journal, vol. 23, No. 6, pp. 939-964, Springer Berlin Heidelberg
- [14] Sanjeet Singh, Jagmeet Bawa, Gagan Deep Sharma, (2018), The Goal of Business A Review Paper, Global Journal of Management and Business Research, Volume XV, Issue XI, Version I, 2015
- [15] Mininath R. Bendre & Vijaya R. Thool, (2016), Analytics, challenges and appli cations in big data environment: a survey, Pages 206-239, Journal of Management Analytics, Volume 3, 2016 - Issue 3, <https://doi.org/10.1080/23270012.2016.1186578>
- [16] Imran, M., Ofli, F., Caragea, D. & Torralba, A.: Using AI and Social Media Multimodal Content for Disaster Response and Management: Opportunities, Chal lenges, and Future Directions. Information Processing & Management, 57(5), 1-9. <http://sci-hub.tw/10.1016/j.ipm.2020.102261> (2020)
- [17] Ruggiero, A. & Vos, M., (2014), Social Media Monitoring for Crisis Communica tion: Process, Methods and Trends in the Scientific Literature, In Online Journal of Communication and Media Technologies, Vol 4, No. 1
- [18] Poddar, S., Mondal, M. & Ghosh, S., (2020), A Survey on Disaster: Understand ing the After-effects of Super-cyclone Amphan and Helping Hand of Social Media, Computer Science, Computers and Society
- [19] Bouzidi, Zair and Boudries, Abdelmalek and Amad, Mourad, (2021), Deep Learn ing and Social Media for Managing Disaster : Survey, Intelligent Systems Conference 2-3 September, 2021, (IntelliSys 2021), Amsterdam, The Netherlands, Volume 294, Intelligent Systems and Applications, IntelliSys 2021, Vol 1, LNNS 294, Chapter 2, DOI : 10.1007/978 - 3 - 030 - 82193 - 72, Springer Nature

- [20] Bouzidi, Z., Boudries, A. & Amad, M., (2021c), Enhancing Crisis Management because of Deep Learning, Big Data and Parallel Computing Environment : Survey, ID:443, Proceedings of the 3rd International Conference on Electrical, Communication and Computer Engineering (ICECCE), 12-13 June 2021, Kuala Lumpur, Malaysia, doi:10.1109/ICECCE52056.2021.9514189
- [21] Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A. & Arshad, H. (2018). State-of-the-art in artificial neural network applications : A survey. In Heliyon, 4(11), DOI:10.1016/j.heliyon.2018.e00938
- [22] Kosala, R. and Blockeel, H., (2000), Web Mining Research : A Survey, ACM SIGKDD Explorations Newsletter, vol. 2, No. 1, ACM
- [23] Masegla, F., Teisseire M. & Poncelet, P. (2005). Sequential pattern mining: A survey on issues and approaches. Encyclopedia of Data Warehousing and Mining
- [24] LeCun, Y., Bengio Y. and Hinton G., (2015), Deep Learning: Review, Nature, Vol 521, May 28st
- [25] Nash, W., Drummond, T. and Birbilis, N., (2018), A review of deep learning in the study of materials degradation, Materials Degradation, doi:10.1038/s41529-018-0058-x
- [26] Shrestha, A. and Mahmood, A., (2019), Review of Deep Learning Algorithms and Architectures, IEEE Access, vol. 7, pp. 53040-53065, doi:10.1109/ACCESS.2019.2912200
- [27] Sangaiah, A. K., (2019), Deep Learning and Parallel Computing Environment for Bioengineering Systems, ISBN 9780128167182, publisher Academic Press
- [28] Luttgau, J., Kuhn, M., Duwe, K., Alforov, Y., Betke, E., Kunkel, J. and Ludwig, T., (2018), SURVEY OF STORAGE SYSTEMS FOR HIGH-PERFORMANCE COMPUTING, SUPERCOMPUTING FRONTIERS AND INNOVATIONS, vol. 5, No. 1, doi:10.14529/jfsf180103
- [29] Stillger, M., Lohman, G. M., Markl, V. and Kandil, M., (2001), LEO-DB2's learning optimizer, Very Large DataBases (VLDB), vol. 1, pp. 19-28
- [30] Awad, M., (2009), FPGA supercomputing platforms: A survey, FPL 09: 19th International Conference on Field Programmable Logic and Applications, pp. 564- 568, doi:10.1109/FPL.2009.5272406
- [31] Amini, M. M. and Schooley, R., (1993), Supercomputing in corporate America: A sample survey, Information & Management, vol. 24, No. 6, pp. 291-303, doi:10.1016/0378-7206(93)90027
- [32] Ofli, F. and Meier, P. and Imran, M. and Castillo, C. and Tuia, D. and Rey, N. and Briant, J. and Millet, P. and Reinhard, F. and Parkan, M. and Joost, S., (2016), Combining Human Computing and Machine Learning to Make Sense of Big (Aerial) Data for Disaster Response, Big Data, vol. 4, No. 1
- [33] Horita, Flavio E.A., Albuquerque, Joao Porto (de), Marchezini, Victor and Mendiondo, Eduardo M., (2017), Bridging the gap between decision-making and emerging big data sources: An application of a model-based framework to disaster management in Brazil, Decision Support Systems, vol. 97, pp. 2-22, doi:10.1016/j.dss.2017.03.001
- [34] Immonen, A. and Paakkonen, P. and Ovaska, E., (2015), Evaluating the Quality of Social Media Data in Big Data Architecture, IEEE Access, vol. 3, pp. 2028- 2043, doi:10.1109/ACCESS.2015.2490723
- [35] Smith, M., Henne, B., Szongott, C. and Voigt, G. von, (2012), Big data privacy issues in public social media, 6th IEEE International Conference on Digital Ecosystems Technologies, pp. 1-6, DOI:10.1109/DEST.2012.6227909
- [36] Nguyen, D. T., Al-Mannai, K., Joty, S. R., Sajjad, H., Imran, M. and Mitra, P., (2017b), Robust classification of crisis-related data on social networks using convolutional neural networks, ICWSM, pp. 632-635
- [37] Kabir, Md Yasin & Madria, Sanjay Kumar, (2019), A Deep Learning Approach for Tweet Classification and Rescue Scheduling for Effective Disaster Management, Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, (SIGSPATIAL'19), pp. 269-278, doi:10.1145/3347146.3359097
- [38] Canon, M. J., Satuito, A. & Sy, C., (2018), Determining Disaster Risk Management Priorities through a Neural Network-Based Text Classifier, 2018 International Symposium on Computer, Consumer and Control (IS3C), Taichung, Taiwan, 2018, pp. 237-241, doi:10.1109/IS3C.2018.00067
- [39] Bouzidi, Z., Amad, M. & Boudries, A. : Deep Learning-based Automated Learning Environment using Smart Data to improve Corporate Marketing, Business Strategies, Fraud Detection in Financial Services and Financial Time Series Forecasting. International Conference on Managing Business through Web Analytics - (ICMBWA2020), Khemis Miliana University, Algeria
- [40] Bouzidi, Z., Boudries, A. and Amad, M., (2021b), LSTM-based automated learning with smart data to improve marketing fraud detection and financial forecasting, Proceedings of 5th International on Economics & Management (Eman 2021) March 18, 2021, Series 49
- [41] Bouzidi, Z., Boudries, A. & Amad, Mourad, (2021e), A Hybrid of Deep CNN LSTM Model to improve Warning, Situational Awareness and Disaster Education in Managing Emergency: Case study of Covid-19, International Journal of Business Information Systems Strategies, Inderscience Publishers, Submitted
- [42] Fiore U., et al. —Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. Information Sciences 479 (2019): 448- 455.
- [43] Sermpinis G., et al. —Neural networks in financial trading. Annals of Operations Research (2021): 293-308.
- [44] Calvez A (le) and Cliff D. —Deep learning can replicate adaptive traders in a limit-order-book financial market. 2018 IEEE Symposium Series on Computational Intelligence, (SSCI 2018): 1876-1883.
- [45] Kodogiannis V and Lolis A. —Forecasting Financial Time Series using Neural Network and Fuzzy System-based Techniques. Neural Computing and Applications 11 (2002): 90-102.
- [46] Fischer J., Jiang W. and Moran S., (2012), 'AtomicOrchid: A mixed reality game to investigate coordination in disaster response', In Proceedings of the 11th International Conference on Entertainment Computing (ICEC 2012), p. 572-577
- [47] Pinheiro, et al. —Stock market prediction with deep learning: A character-based neural language model for event based trading. Proceedings of the Australasian Language Technology Association Workshop (2017): 6-15.
- [48] Mohammad., et al. —Hybrid deep learning model for stock price prediction. IEEE Symposium Series on Computational Intelligence SSCI. IEEE (2018): 1837-1844.
- [49] Ryman-Tubb, N. F., Krause, P. & Garn, W. (2018). How artificial intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark. Engineering Applications of Artificial Intelligence, 76, 130-157. <http://www.sciencedirect.com/science/article/pii/S0952197618301520>
- [50] Cavalcante, R. C., Brasileiro, R. C., Souza, V. L., Nobrega, J. P. & Oliveira, A. L. (2016). Computational intelligence and financial markets: A survey and future directions. Expert Systems with Applications, vol. 55, 194-211
- [51] Ali Alqahtani, Mohammed Ali, Xianghua Xie and Mark W. Jones, Deep Time-Series Clustering: A Review, Electronics 2021, 10(23), 3001; <https://doi.org/10.3390/electronics10233001>, <https://www.mdpi.com/journal/electronics>
- [52] A. Figueira and L. Oliveira, The current state of fake news: Challenges and opportunities, Proc. Comput. Sci., vol. 121, pp. 817-825, Jan. 2017
- [53] K. Warwick, (2003), Cyborg morals cyborg values cyborg ethics, Ethics Inf. Tech nol., vol. 5, no. 3, pp. 131-137
- [54] S. Castillo, H. Allende-Cid, W. Palma, R. Alfaro, H. S. Ramos, C. Gonzalez, et al., (2019), Detection of bots and cyborgs in Twitter: A study on the Chilean presidential election in 2017, Proc. Int. Conf. Hum.-Comput. Interact., pp. 31
- [55] A. Balestrucci and R. D. Nicola, (2020), Credulous users and fake news: A real case study on the propagation in Twitter, Proc. IEEE Conf. Evolving Adapt. Intell. Syst. (EAIS), pp. 1-8, May 2020
- [56] M. Orabi, D. Mouheb, Z. Al Aghbari and I. Kamel, (2020), Detection of bots in social media: A systematic review, Inf. Process. Manage., vol. 57, no. 4
- [57] C.-C. Wang, (2020), Fake news and related concepts: Definitions and recent research development, Contemp. Manage. Res., vol. 16, no. 3, pp. 145-174, Sep. 2020
- [58] W. Shahid, Y. Li, D. Staples, G. Amin, S. Hakak and A. Ghorbani, Are You a Cyborg, Bot or Human?—A Survey on Detecting Fake News Spreaders, in IEEE Access, vol. 10, pp. 27069-27083, 2022, doi: 10.1109/ACCESS.2022.3157724.

- [59] Zair Bouzidi, Abdelmalek Boudries and Mourad Amad, (2021), Deep LSTM-based Model using Smart Data to Improve Awareness and Education in Marketing, Business Strategy and Financial Forecasting, MC Medical Sciences (MCMS) journal, Volume 1, Issue 5, December 2021, <https://themedicon.com/medicalsciences/volume-1-issue-5>, doi:10.55162/MCMS.2021.01.034
- [60] Young, S. D., Rivers, C. and Lewis, B., (2014), Methods of using real-time social media technologies for detection and remote monitoring of HIV outcomes, Preventive Medicine, vol. 63, pp. 112-115
- [61] Krizhevsky, A. and Sutskever, I. and Hinton, G., (2012), ImageNet classification with deep convolutional neural networks, Proceedings of Advances in Neural Information Processing Systems, vol. 25, pp. 1090-1098
- [62] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., (1998), Gradient-based learning applied to document recognition, Proceedings of the IEEE, vol. 86, No. 11, pp. 2278- 2324, doi:10.1109/5.726791
- [63] Wu, Q., Ding, K. & Huang, B., (2018), Approach for fault prognosis using recurrent neural network. Journal of Intelligent Manufacturing. 1-13. Doi:10.1007/s10845-018- 1428-5
- [64] He, K., Zhang, X., Ren, S. & Sun, J., (2016b), Deep Residual Learning for Image Recognition, Proceedings of Computer Vision and Pattern Recognition (CVPR), IEEE, Las Vegas, NV, pp. 770-778, doi:10.1109/CVPR.2016.90
- [65] Zhao, J., Deng, F., Cai, Y. & Chen, J., (2018), Long short-term memory - Fully connected (LSTM-FC) neural network for PM2.5 concentration prediction, Chemosphere, vol. 220, doi:10.1016/j.chemosphere.2018.12.128
- [66] Pouyanfar, S., Tao, Y. and Tian, H., Chen, S.-C. & Shyu, M.-L., (2018), Multi modal deep learning based on multiple correspondence analysis for disaster management, World Wide Web, vol. 22, pp. 1893-1911, doi:10.1007/s11280-018-0636-4
- [67] Haley, P. J. & Soloway, D., (1992), Extrapolation limitations of multi layer feedforward neural networks, Proceedings 1992 IJCNN International Joint Conference on Neural Networks, vol. 4, Baltimore, MD, USA, pp. 25-30, doi:10.1109/IJCNN.1992.227294
- [68] Mazzia, V., Salvetti, F. & Chiaberge, M. Efficient-CapsNet: capsule network with self-attention routing. Sci Rep 11, 14634 (2021). <https://doi.org/10.1038/s41598-021-93977-0>
- [69] Chen, W., Sidky, H. and Ferguson, A., (2019), Capabilities and Limitations of Time-lagged Autoencoders for Slow Mode Discovery in Dynamical Systems, The Journal of Chemical Physics, vol. 151, No. 6, doi:10.1063/1.5112048
- [70] Wolfgang Maass, (1997), Networks of spiking neurons: The third generation of neural network models, Neural Networks, Volume 10, Issue 9, Pages 1659-1671, ISSN 0893-6080, [https://doi.org/10.1016/S0893-6080\(97\)00011-7](https://doi.org/10.1016/S0893-6080(97)00011-7)
- [71] D. M. J. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, et al., (2018), The science of fake news, Science, vol. 359, no. 6380, pp. 1094-1096
- [72] Creech, B., (2020), Fake news and the discursive construction of technology companies' social power. Media, Culture & Society, 42(6), 952-968. <https://doi.org/10.1177/0163443719899801>
- [73] J. M. Burkhardt, (2017), History of fake news, Library Technol. Rep., vol. 53, no. 8, pp. 5-9, 2017.
- [74] S. B. Naeem, R. Bhatti and A. Khan, (2021), An exploration of how fake news is taking over social media and putting public health at risk, Health Inf. Libraries J., vol. 38, no. 2, pp. 143-149, Jun. 2021.
- [75] X. Zhang and A. A. Ghorbani, (2020), An overview of online fake news: Characterization detection and discussion, Inf. Process. Manage., vol. 57, no. 2
- [76] E. Ferrara, O. Varol, C. Davis, F. Menczer and A. Flammini, The rise of social bots, Commun. ACM, vol. 59, no. 7, pp. 96-104, Jul. 2016.
- [77] M. Umer, Z. Imtiaz, S. Ullah, A. Mehmood, G. S. Choi and B.-W. On, (2020), Fake news stance detection using deep learning architecture (CNN-LSTM), IEEE Access, vol. 8, pp. 156695-156706.
- [78] G. Baym and J. P. Jones, News Parody Political Satire Across Globe, Evanston, IL, USA:Routledge, 2013.
- [79] S. Singhal, R. R. Shah, T. Chakraborty, P. Kumaraguru and S. Satoh, ISpotFake: A multi-modal framework for fake news detection, Proc. IEEE 5th Int. Conf. Mul. multimedia Big Data (BigMM), pp. 39-47, Sep. 2019.
- [80] D. Guera and E. J. Delp, IDepfake video detection using recurrent neural network, Proc. 15th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS), pp. 1-6, Nov. 2018
- [81] S. Cresci, M. Petrocchi, A. Spognardi and S. Tognazzi, IBetter safe than sorry: An adversarial approach to improve social bot detection, Proc. 10th ACM Conf. Web Sci., pp. 47-56, Jun. 2019
- [82] R. D. Nicola, M. Petrocchi and M. Pratelli, (2021), On the efficacy of old features for the detection of new bots, Inf. Process. Manage., vol. 58, no. 6, Nov. 2021
- [83] Berglund, M., Raiko, T., Honkala, M., Karkkainen, L., Vetek, A. & Karhunen, J., (2015), Bidirectional Recurrent Neural Networks as Generative Models, MIT Press, Cambridge, MA, USA
- [84] Mujeeb S. & Gupta S. (2022) Fake Account Detection in Social Media Using Big Data Analytics. In: Reddy A.B., Kiranmayee B., Mukkamala R.R., Srujan Raju K. (eds) Proceedings of Second International Conference on Advances in Computer Engineering and Communication Systems. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-16-7389-4_57
- [85] Aloysius G and Binu D., (2013), —An approach to products placement in super markets using Prefix Span algorithm, Journal of King Saud University-Computer and Information Sciences 25.1: 77-87
- [86] Han J.-H., et al., (2007), —Neuronal Competition and Selection During Memory Formation, Science 316.5823 : 457-460
- [87] Fournier-Viger P. et al., (2017), —A Survey of Sequential Pattern Mining, Data Science and Pattern Recognition, Ubiquitous International 1.1: 54-78

deep learning based semantic skills matching for resumes selection

DJOUHRI Mohamed Fouzi 1st

Laboratoire de l'Intelligence Artificielle et des
Technologies de l'Information Faculté des nouvelles
technologies de l'information et de la communication,
Université Kasdi Merbah Ouargla
Ouargla 30 000, Algeria
mohamedfouzidjoughri@gmail.com

AOUARIB Houssam Eddine 2nd

Laboratoire de l'Intelligence Artificielle et des Technologies
de l'Information Faculté des nouvelles technologies de
l'information et de la communication, Université Kasdi
Merbah Ouargla
Ouargla 30 000, Algeria
housameddine.aouarib@gmail.com

Abstract—E-recruitment systems has become an emerging resource for talent acquisition in recent years. Choosing the relevant candidates for a given job offer can be sometimes difficult for the recruiters, since they always seek to select the best candidates, in a short period of time. While candidates exploit technical language on their profiles documents, recruiters try to find the relevant candidates for jobs using their own terminology which can make the selection very difficult. To deal with these problems, we propose in this paper, a candidate recommendation system, which proposes the suitable candidate curriculum vitae for the giving job offer. Our system is based on semantic parsing of candidates profiles, to extract meaning from text to enable natural language understanding, and documents semantic classification, using deep learning, word embedding and machine learning clustering. Firstly, the word embedding representation is used to analyse the textual content of the candidate's profile, Then we performed a clustering process, to regroup similar candidate's profile on the basis of their shared semantic skills terms Thanks to the machine learning. Finally, a list of top ranked curriculum vitae are recommended for job opportunities. Our experiments are tested on a candidates resumes dataset, coming from a Kaggle platform.

Keywords— Recommender system, Information retrieval, Semantic Matching, Selection process, Word Embedding

I. INTRODUCTION

Since the last two decades, the use of Internet for recruitment purposes has grown considerably. This recruitment process, also known as “e-recruitment” is based on the use of information technology and communications. In this context, the expansion of the Web has led to an increase in the number of web platforms for the diffusion of job offers (also called job board) and consequently the number of candidates that can be contacted through these intermediary tools. [1]. Currently, the Internet has been proved to be an essential medium for recruitment, CareerBuilder as an example operates one of the largest job boards in the world. It has millions of job Vacancies, more

than 60 million actively-searchable resumes, over one billion searchable documents, and receives several million searches per hour [2]. According to the LinkedIn report, only on LinkedIn, there are 706 million Users and 20 million Job postings, and a total of about 36 thousand skills listed. [3]. However, despite the widespread use of existing platforms of e-recruitment, the main problem faced by recruiters is to find the best profiles (i.e. the most talented candidates) for a given job offer. In a strategic and economic context where cost control is fundamental, the selection of the most suitable candidate curriculum vitae (CV) regarding an offer is very important for the optimization of this fully digital recruitment process. Concentrating on matching job offers with CVs, the job offers and the CVs contain technical aspects of the language, hence in many cases, there is a deep gap between the job offer and the language which is used by candidates in their CVs in the same business category. As an example, consider the “Mobile Application Developer” offer. As candidates of the Mobile Application Developer community do not use the term itself directly in their CVs and instead they use some terms like “Android” and “iOS” which have a direct relation with “Mobile Application Developer”, we can say that a deep gap exists between the “Mobile Application Developer” as main skills required and the language used by candidates. In a HR staff, each job offer has one or more keywords which indicate the required skills to having right candidates. These keywords can basically be considered as skill areas which recruiters are interested in. For example, consider the following job description on offer, “ expertise in Java and Object-Oriented Design and development principles (e.g. XML, JSON) ”. This description can be tagged by “xml”, “json”, “java-ee” which are important skill areas in Java programming language. In this work, our main goal is to provide a tool which can help recruiters to (i) to detect and rank candidates who are skillful in a given skill area (tag) with respect to the qualifications and duties of the job offer which are given, (ii) select more effectively talented profiles, that is to say in the right place at the right time, (iii) provide tools to connect candidate's CVs and job offers automatically.

To meet the above objectives, In this paper, we propose here a candidate CVs recommendation system based on natural language processing and machine learning techniques for automatically clustering CVs into profiles and identify the relevant candidates CVs for a given new job offer. The semantic matching of CV and job offer requires a textual analysis of the content of both the job offer and CV using models to interpreted a given skill area (e.g. “Android”, “JavaScript” and etc.) to a set of relevant words. These interpretations can be useful to improve the matching between recruiters queries and the technical textual evidence (i.e. CVs) associated with each candidate. They can also be used independently by recruiters to detect important aspects of each skill area in business categories. Our model is based on a word embedding method utilizing the specific structure to interpret a skill area to its relevant words After finding the appropriate interpretations, we have used scoring approaches to find the final ranking of candidates CV for a given job offer.

The remainder of this paper is organized as follows. Background concepts and related work are introduced in Section 2. The proposed model will be presented in section 3 The experiments and the evaluation result in sections 4 and 5, respectively. Finally, we provide conclusions and future work as well as the limitations of the proposed approach are discussed in Section 5.

II. RELATED WORKS

A. Recommender Systems

The development of the Internet and web technologies has expanded the range of items available in various areas such as entertainment, e-commerce, and education. However, it is becoming difficult to find suitable items that match the preferences of users. For this reason, recommender systems are expected to guide users to find items or product that might be of interest to them. The main goal of recommender system is to assist people in decision-making situations under contexts of “information overload”. The recommender systems have been extremely used in different areas, such as recommending movies [10], books [11], user-generated videos [12], products [13], documents [14], and information on the Web [15]. Various approaches have been proposed for generating recommendations, including: content-based filtering, collaborative filtering and hybrid recommendation systems [16,17,18].

The Content-based filtering (CBF) approach analyzes users’ preferences according to the attributes of an item, in order to build up a personal feature profile and then predict which items the user will like [19,20]. items . On the other hand, the collaborative filtering (CF) may be the most commonly used approaches focus on identifying the relationship between items or users .Generally, The existing CF approaches can be divided into two types: user-based CF and item-based CF [16]. The user-based CF clusters users

into different groups or finds the neighborhood of target users first, and then observes the behavior of neighbors to make recommendations [20,10]. The item-based CF approach analyzes similarities between items on the basis of a user’s ratings among items. These item similarities are then used to generate recommendations for a user by finding items that are similar to those the user has previously liked [21]. Hybrid approaches have been proposed to overcome drawbacks of the CF and CBF methods [22,18]. They combine content-based filtering and collaborative filtering to improve recommendation the prediction performance . Furthermore, matrix factorization-based methods have been proposed to discover latent factors of items and users for enabling effective recommendation [23,24]. Other recent work have addressed the diversity of recommendations [25,26].

B. E-recruitment recommender systems

Recommender systems in automatic recruitment platforms allow HR agents to search and identify the best candidates in a small temporal period . To this aim, there are many e-recruitment systems consider a important part of the recruitment process, concentrating on matching job offers with CVs. In this context, several recommendation systems have been presented in the literature In particular, Kessler et al. [27] presented the project e-gen to assist a recruitment consultant. The aim, perform analysis and categorization of job offers and the candidates' responses. E- gen approach based on vectorial and probabilistic models to solve the issue of profiling applications according to a specific job offer. The system transformed the document into a vector with weights characterizing the frequency of terms Tf and Tf-idf. Then compares candidatures and a particular job offer using measures of similarity such as cosine similarity and Minkowski distance, to rank all candidatures in relation to a job offer as well, they make use of a Relevance Feedback method that consists of enriching the job offer dictionary by concatenating already analyzed relevant CVs from the same job offer. In the same spirit, authors in [28] proposed extensions to improve the system's performance by adding a statistical automatic summarization system, namely Cortex, as a powerful filter to eliminate irrelevant information in candidate resumes and job offers . In [29] the authors introduced "prospect" system as decision support tool to shortlist candidates for job openings . Prospect uses a resume miner to extract meaningful information from resumes and job offers using a conditional random field model to segment and label the CV. Then, to rank the CV based on the job offer, the system compares the information from both documents using three established retrieval models: Okapi BM25 , Kullback–Leibler divergence algorithm or Lucene scoring. We note in the literature another system to match job offers and candidates CVs . In [30] the authors proposed system namely Expert system based on ontology to represent individual CVs and job offers. the system ranks the CVs based on proximity to the

job offer ontology with respect to each resume ontology. The CVs with ontologies most similar to those of the job offer are ranked near the top. In addition, a framework for job matchmaking is proposed by [31] where the system strategy is based on ontologies used to enhance information from resumes and job offers. The framework is a web-based system developed to match professional competencies from CVs and job offers. The ranking of a resume with respect to a job is determined through semantic similarity. Finally, another approach which is essentially based on ontology was proposed in [32]. This work considered the problem information extraction process from CVs. Thanks to ontologies, the system processes a CVs by first converting the CV documents into plain text, sentence segmentation, applying an ontology knowledge base to find concepts in each sentence, term normalization, and finally classifying each sentence.

III. THE PROPOSED RECOMMENDATION APPROACH

In this section, we introduce the architecture for the proposed semantic recommendation system of the relevant candidates resumes for a giving job offers. The architecture of the proposed framework is illustrated in Fig. 1. The framework can be divided into two modules namely: learning the predictive model step (preparation, in the top frame of Fig. 1), and the recommendation step (execution, in the down frame of Fig. 1).

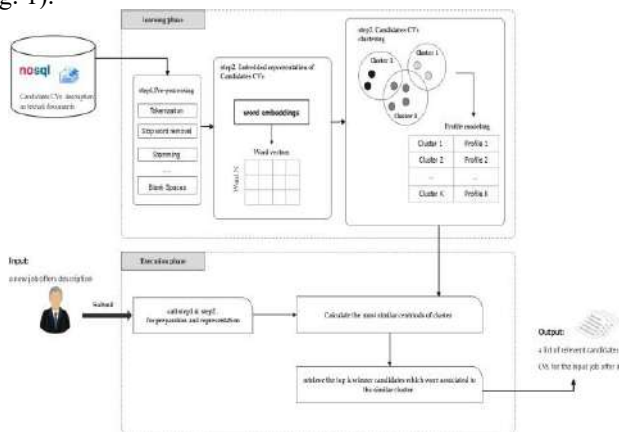


Fig. 1. The architecture of the curriculum vitae recommender system.

In the learning stage, there are three main steps. Firstly, the system concerns the preparation and the formatting of these heterogeneous data. As a second step, we will use embedded layers of deep neural networks to represent the textual resumes documents in a sub dimensional word embedding space. Then we will apply a clustering procedure to discover similar classes among this representation of the resumes. The idea is to create a topology of resumes classes, that belong to similar resumes categories. In the second phase, the recommendation comes in two stages. Firstly, the system concerns the embedded representation of the job offer textual documents and projected in the learned embedded space model to extracting latent topics and identifying the nearest pre-generated cluster of resumes, compared to this new offer in

term of its similarity, with the centroid of the cluster. Since we already have an association for each cluster with its corresponding winner resumes, thus we recommend to the input job offer. In the following sub-sections, we describe in more detail the different steps of the proposed method.

A. Document Preprocessing

In this work, we need to perform some preprocessing steps to prepare the candidates CVs contents. These preparations represent natural language processing techniques such as tokenization, stop word removal and stemming. Tokenization is the process of mapping a stream of text content into words, phrases, symbols or other significant elements, called tokens. These tokens, when filtered, can be used to build the language dictionary. Most of the words in the documents occur very often, but they are basically meaningless words as they are used to connect the words well organized to form a sentence they are commonly words with little useful information like “and”, “or”, “but”, “will” and others. Generally, it is assumed that stop words, which does not denote content or context of text documents. Stemming is the process of combining the different types of a word into a typical illustration, the stem. The main goal of stemming is to reduce the derive of words frequency. Other preprocessing used are converting all text to lowercase aimed to enable the model to recognize better different cases and to avoid confusion during processing. Removing White Spaces, special characters, punctuation from the text like “.” and “!” and also the symbols like “@#” and overlaying useless terms for text features such as Numbers and links.

B. Deep learning-based word embedding model

Once all NLP pre-processing terminated we used the corpus of candidates CVs to represent each document in the vector space model. The development of deep learning drives continuous innovation of text representation. To that aim, we have used a deep learning to produce the embedding representation of CVs documents. Natural language modelling technique like Word Embedding is used to map words or phrases from a vocabulary to a corresponding vector of real numbers. As well as being amenable to processing by Machine Learning (ML) algorithms, this vector representation has two important and advantageous properties: (i) Dimensionality Reduction – it is a more efficient representation, and (ii) Contextual Similarity – it is a more expressive representation. In preceding works on Bag of Words (BoW) approach have shown that it often produces huge, very sparse one-hot encoded vectors, where the dimensionality of the vectors representing each document is equal to the size of the supported vocabulary [4,5]. Word Embedding aims to create a vector representation with a much lower dimensional space. These are called Word Vectors. For our purposes, Word Embedding is used for semantic parsing of candidates CVs, to extract meaning from text to enable natural language understanding, and documents semantic

classification. For a language model to be able to predict the meaning of text, it needs to be aware of the contextual similarity of words. The vectors created by Word Embedding preserve the similarities, thus words that regularly occur nearby in text will also be in close proximity in vector space. Fig. 8 gives an example of an intuitive representation of some candidates profiles in a word embedding space. Theoretically, candidates that belong to the same fields are supposed to be close to each other in the produced space model. Later, candidates represented with these profiles are to be categorized in the same clusters (e.g. data scientist and deep learning expert are in the same cluster of computer scientist profiles). This is the main purpose of word embedding implementation in the first stage of our learning algorithm.

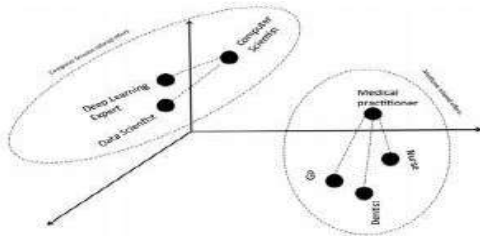


Fig. 2. Example representation of resume using Word2vec.

1. Word2Vec (CBOW and skip-gram model)

One of the best known algorithms for producing word embedding models is Word2vec. This framework initially developed by Mikolov et al. [6,7–8] for representing words as vectors and implemented by Google, which released a version of the code for academic purposes. The Word2Vec model is a shallow neural network-based model which is used for word embedding. This model, which was trained with a large corpus of text, creates a unique vector for each word in a high dimensional space. That is, word2vec can be used to find the semantic relationships between words. Word2Vec uses two main model architectures to create word vector representation: Continuous Bag-of-Words (CBOW) and Skip-gram. Figures 2 and 3 show the difference between the two models. The CBOW architecture model predicts a word based on the surrounding context words while the Skip-gram architecture model uses the current word to predict the surrounding words in a fixed size window ;

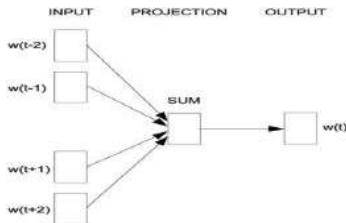


Fig. 3. Continuous bag-of-words (Mikolov et al., 2013)

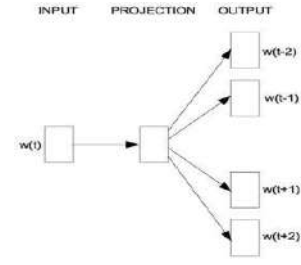


Fig. 4. The Skip-gram model architecture.

In CBOW, the model uses encoders neural networks to generate embeddings of a target word from an input context. While a language model depends on the preceding words in the corpus for its predictions. Mikolov et al. thus use both the n words preceding and following the target word w_t to predict it. The basic idea of CBOW is to maximize the following average log-likelihood :

$$J_{\theta} = \frac{1}{T} \sum_{t=1}^T \log p(w_t | w_{t-n}, \dots, w_{t-1}, w_{t+1}, \dots, w_{t+n})$$

instead of feeding n previous words into the model, the model receives a window of n words around the target word w_t at each time step t . In the deep neural network, this probability is computed through the softmax layer(exp):

$$p(w_t | w_{t-n}, \dots, w_{t-1}, w_{t+1}, \dots, w_{t+n}) = \frac{\exp(h^T v_{w_t})}{\sum_{w_i \in V} \exp(h^T v_{w_i})}$$

where, h is the intermediate state vector which is the word embedding v_{w_t} of the input word w_t of a vocabulary V

The skip-gram model allows to do the inverse of CBOW. The goal of skip-gram is to predict the context words by computing the the log probabilities of the surrounding n words to the left and to the right of the target word w_t to produce the following average log-likelihood:

$$J_{\theta} = \frac{1}{T} \sum_{t=1}^T \sum_{-n \leq j \leq n, j \neq 0} \log p(w_{t+j} | w_t)$$

Similarly the skip-gram model computes $p(w_{t+j} | w_t)$, with the softmax layer as:

$$p(w_{t+j} | w_t) = \frac{\exp(v_{w_t}^T v_{w_{t+j}})}{\sum_{w_i \in V} \exp(v_{w_t}^T v_{w_i})}$$

2. Doc2Vec

In 2014, Doc2vec that is an extension of Word2Vec, was presented by Mikolov [6,7–8] as a set of approaches to represent documents a fixed-sized continuous vector that are document embeddings. It is an unsupervised learning algorithm that represents an input large document in the paragraph vector with PV-DM (Paragraph Vector with

Distributed Memory) and PV DBOW (Paragraph Vector with Distributed Bag-of-Words) forms . The difference between these two methods lies in the prediction task employed for learning the document vectors . The PV-DM method is similar to the continuous bag-of words (CBOW) for the Word2Vec model [33,34] which attempts to predicts the target word using given words sampled from a paragraph and the identifier of the paragraph , whereas PV-DBOW predicts the words contained in each paragraph using only the identifier of the paragraph. Le and Mikolov (2014) showed that PV-DM performed better in embedding words and documents , but this method is slower to train as it has more training sets than PV-DBOW for a given corpus. Thus, as it focuses more on the quality of document embeddings rather than training time .Although Le and Mikolov (2014) recommended using the concatenation of PV-DM and PV-DBOW in a difficulties situation, The PV-DM performed well when used by itself in a less difficult case.

C. Expertise domain extraction

Once resumes vectors are embedded into a semantic space, we followed the primary step of our learning algorithm, by classifying the documents vectors in numerous clusters, in-order to extract the emerging topics from the datasets, during which the CV vectors in neighboring semantic space are clustered into a typical cluster. Since each document is represented through a numerical embedding vector, we've tested two of the foremost widely known clustering algorithms to discover correlated documents in the database which are hierarchical and K means clustering algorithms. As deep learning based word embeddings models maximizes the vector product between the embedding vectors and also the context vectors, cosine distance has been chosen as an appropriate metric for measuring distances between word vectors in semantic space and clustering nearby word vectors into a standard concept.

D. Candidates CVs Recommendation

Once the steps of the learning phase terminated, then the model can be used in the recommendation phase. Our recommendation approach recommends CVs documents that are similar to those offers preferred by the recruiters based on the semantic similarity between documents. The first step consists of finding the cluster of candidate's profiles to which job offer belonging to, compared to this offer in terms of its similarity, with the centroid of the cluster (since a centroid is a candidate's profile). For the semantic cluster previously calculated, the recommender system produces candidate CVs , scores and ranks the CVs for the input job offer. In the actual calculations, we used the cosine similarity measure [35] to select one corresponding CVs for giving job offer . The cosine similarity (cos) is given by the equation bellow :

$$\cos(\vec{P}, \vec{J}) = \frac{\vec{P} \cdot \vec{J}}{|\vec{P}| |\vec{J}|}$$

where P represents the candidate CV and J the job offer . The two documents are converted into vectors by some vector representation (e.g., Doc2vec). The cosine similarity value is a number in the interval [0, 1], where the value 0 indicates that the vectors are completely different and the value 1 indicates that. Finally, the recommendation list of CVs documents based on the semantic recommendation is derived. In this list, the recommended CVs documents are ranked in descending order of the recommendation score. The most important candidates are those that are at the top of the ranked list, the users of the tool can then select the candidates that are having an increasing score

IV. EXPERIMENTAL RESULTS

A. Description of the dataset

To evaluate the proposed approach, we need an unstructured database of CV documents. The literature on publicly available datasets is fairly rare in this context. In fact, these datasets may contain personal data (addresses, phone numbers, etc.), and regulations such as the GDPR make it more difficult to distribute datasets for research purposes as well. In addition, collecting and labeling datasets with such information is costly in terms of both the task itself and the knowledge require. Experimental data are collected from the online data science platform Kaggle [39], Kaggle is a platform that hosts data science competitions for business problems, recruitment, and academic research purposes The members of a Kaggle community can collect the useful benchmark dataset to better evaluate their approaches. In this work, 25 business classes are represented in the dataset which consists of different domains: computer, medicine and finance and so on. The data set contains 1219 CV written in English. We analyzed a total of 1202 resumes because some resumes appear to be corrupted or unavailable. Each instance is represented by the plain text of the CV and a category indicating the work sector pertaining to the CV.

B. Evaluation Mertic

The system performance is evaluated in terms of precision, recall, and F-measure [9,36,37]. These metrics have been commonly used in information retrieval. Precision is determined by the fraction of the recommended profiles documents that are actually found to be interesting for a giving job offers. On the other hand, Recall involves the fraction of the interesting profile documents that are correctly identified by the recommendation process. The equations of precision and recall are measured in the following formulas:

$$\text{Precision} = \frac{\text{number of correctly recommended Profile documents}}{\text{number of recommended Profile documents}}$$

$$\text{Recall} = \frac{\text{number of correctly recommended Profile documents}}{\text{number of collected Profile documents}}$$

F-measure use both precision and recall to compute the score for balancing the trade-off between precision and recall. More specifically, The F-measure (or F1 score) is the harmonic mean of precision P and recall R as shown in the following formula:

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

I. RESULTS AND DISCUSSION

The result is under development

II. CONCLUSION

In this paper, we have presented candidate recommendation system based on semantic knowledge using deep learning, word embedding and clustering algorithms. The objective is to select the best candidates for a job offer. The system is divided into two main stage, the learning stage and the recommendation stage. Concerning the learning stage, we used some preprocessing steps to prepare the candidates documents contents the we used word embedding representation of candidates documents. Finally the we applied an unsupervised machine learning algorithm in order to regroup similar documents on the basis of their textual content. The proposed system suggests the recommendation of candidates in the top ranked for the disseminated job offer in which may be at best match to the job. As future work, we will use author semantic representation technique such as fastText and Glove.

REFERENCES

- [1] J. Sgula, Fouille de donnees textuelles et systmes de recommandation appliqués aux offres d'emploi diffusés sur le web, Ph.D. thesis, CEDRIC Laboratory, Paris, France, 2012.
- [2] K. AlJadda, M. Korayem, T. Grainger, C. Russell, Crowd sourced query augmentation through semantic discovery of domain-specific jargon, in: 2014 IEEE International Conference on Big Data, IEEE, 2014, pp. 808–815.
- [3] LinkedIn Pressroom Statistics <https://news.linkedin.com/about-us#1>.
- [4] F. Chollet, Deep Learning with Python, Manning Publications Company, 2017.
- [5] Y. Bengio, R. Ducharme, P. Vincent, C. Janvin, A neural probabilistic language model, *J. Mach. Learn. Res.* 3 (2003) 1137–1155.
- [6] T. Mikolov, W.-T. Yih, G. Zweig, Linguistic regularities in continuous space word representations., in: Proceedings of the HLT-NAACL, 2013, pp. 746–751.
- [7] T. Mikolov, I. Sutskever, K. Chen, G.S. Corrado, J. Dean, Distributed representations of words and phrases and their compositionality, in: Proceedings of the Advances in Neural Information Processing Systems, 2013, pp. 3111–3119.
- [8] R. Pascanu, T. Mikolov, Y. Bengio, On the difficulty of training recurrent neural networks, in: Proceedings of the International Conference on Machine Learning, 2013, pp. 1310–1318.
- [9] D. Billsus, M.J. Pazzani, Learning collaborative information filters, in: Proceedings of the Fifteenth International Conference on Machine Learning, 1998, pp. 46–54.
- [10] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, J. Riedl, GroupLens: an open architecture for collaborative filtering of netnews, in: Proceedings of ACM Conference on Computer Supported Cooperative Work, Chapel Hill, North Carolina, United States, 1994, pp. 175–186.
- [11] G. Linden, B. Smith, J. York, Amazon.com recommendations: item-to-item
- [12] Z. Wang, L. Sun, W. Zhu, S. Yang, H. Li, D. Wu, Joint social and content recommendation for user-generated videos in online social network, *IEEE Trans. Multimedia* 15 (2013) 698–709
- [13] D.-R. Liu, Y.-Y. Shih, Integrating AHP and data mining for product recommendation based on customer lifetime value, *Inform. Manage.* 42 (2005) 387–400.
- [14] D.-R. Liu, C.-H. Lai, Y.-T. Chen, Document recommendations based on knowledge flows: a hybrid of personalized and group-based approaches, *J. Am. Soc. Inform. Sci. Technol.* 63 (2012) 2100–2117.
- [15] H. Ma, I. King, M.R.-T. Lyu, Mining web graphs for recommendations, *IEEE Trans. Knowl. Data Eng.* 24 (2012) 1051–1064.
- [16] G. Adomavicius, A. Tuzhilin, Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions, *IEEE Trans. Knowl. Data Eng.* 17 (2005) 734–749.
- [17] M. Balabanovi, Y. Shoham, Fab: content-based, collaborative recommendation, *Commun. ACM* 40 (1997) 66–72.
- [18] R. Burke, Hybrid recommender systems: survey and experiments, *User Model. User-Adap. Inter.* 12 (2002) 331–370.
- [19] M. Pazzani, D. Billsus, Content-based recommendation systems, in: The Adaptive Web, LNCS, vol. 4321, Springer, Berlin/Heidelberg, 2007, pp. 325–341.
- [20] J.A. Konstan, B.N. Miller, D. Maltz, J.L. Herlocker, L.R. Gordon, J. Riedl, GroupLens: applying collaborative filtering to Usenet news, *Commun. ACM* 40 (1997) 77–87.
- [21] B. Sarwar, G. Karypis, J. Konstan, J. Riedl, Analysis of recommendation algorithms for e-commerce, in: Proceedings of the 2nd ACM Conference on Electronic Commerce, Minneapolis, Minnesota, United States, 2000, pp. 158–167
- [22] A.B. Barragáns-Martínez, E. Costa-Montenegro, J.C. Burguillo, M. Rey-López, F.A. Mikic-Fonte, A. Peleteiro, A hybrid content-based and item-based collaborative filtering approach to recommend TV programs enhanced with singular value decomposition, *Inf. Sci.* 180 (2010) 4290–4311.
- [23] Y. Koren, R. Bell, C. Volinsky, Matrix factorization techniques for recommender systems, *IEEE Comput.* 42 (2009) 30–37.
- [24] J. Liu, C. Wu, W. Liu, Bayesian probabilistic matrix factorization with social relations and item contents for recommendation, *Decis. Support Syst.* 55 (2013) 838–850.
- [25] G. Adomavicius, Y. Kwon, Improving aggregate recommendation diversity using ranking-based techniques, *IEEE Trans. Knowl. Data Eng.* 24 (2012) 896–911.

- [26] M. Gan, R. Jiang, Improving accuracy and diversity of personalized recommendation through power law adjustments of user similarities, *Decis. Support Syst.* 55 (2013) 811–821.
- [27] Kessler, R., Béchet, N., Torres-Moreno, J.-M., Roche, M., & El-Bèze, M. (2009). Job offer management: How improve the ranking of candidates. In *Foundations of intelligent systems: Proceedings of 18th international symposium on methodologies for intelligent systems (ISMIS 2009)*. In *Lecture Notes in Computer Science*: 5722
- [28] Kessler, R., Béchet, N., Roche, M., Torres-Moreno, J.-M., & El-Bèze, M. (2012). A hybrid approach to managing job offers and candidates. *Information Processing & Management*, 48(6), 1124–1135. doi:10.1016/j.ipm.2012.03.002.
- [29] Singh, A., Rose, C., Visweswariah, K., Chenthamarakshan, V., & Kambhatla, N. (2010). PROSPECT: A system for screening candidates for recruitment. In *Proceedings of the 19th ACM international conference on information and knowledge management (CIKM 2010)* (pp. 659–668). Toronto, Canada: ACM. doi:10.1145/1871437.1871523 .
- [30] Senthil Kumaran, V., & Sankar, A. (2013). Towards an automated system for intelligent screening of candidates for recruitment using ontology mapping (EXPERT). *International Journal of Metadata, Semantics and Ontologies*, 8(1), 56–64. doi:10.1504/IJMSO.2013.054184
- [31] Montuschi, P., Gatteschi, V., Lamberti, F., Sanna, A., & Demartini, C. (2014). Job recruitment and job seeking processes: How technology can help. *IT Professional*, 16(5), 41–49. doi:10.1109/MITP.2013.62
- [32] Kmail, A. B., Maree, M., & Belkhatir, M. (2015). MatchingSem: Online recruitment system based on multiple semantic resources. In *12th international conference on fuzzy systems and knowledge discovery (FSKD 2015)* (pp. 2654–2659). doi:10.1109/FSKD.2015.7382376.
- [33] Seungwan Seo, Deokseong Seo, Myeongjun Jang, Jaeyun Jeong, Pilsung Kang, Unusual customer response identification and visualization based on text mining and anomaly detection, *Expert Systems with Applications*, Volume 144, 2020, 113111, ISSN 0957-4174.
- [34] Daeseong Jeon, Joon Mo Ahn, Joram Kim, Changyong Lee, A doc2vec and local outlier factor approach to measuring the novelty of patents, *Technological Forecasting and Social Change*, Volume 174, 2022, 121294, ISSN 0040-1625.
- [35] Salton, G., & McGill, M. J. (1986). *Introduction to Modern Information Retrieval*. New York, NY, USA: McGraw-Hill, Inc.
- [37] J. Cho, K. Kwon, Y. Park, Collaborative filtering using dual information sources, *IEEE Intell. Syst.* 22 (2007) 30–38.
- [38] C.J.V. Rijsbergen, *Information retrieval*, second ed., Butterworth-Heinemann, London, 1979.
- [39] <https://www.kaggle.com/maitrip/resumes>.

Design and FPGA Implementation of Low-complexity Pruned DCT Approximation for Image Compression in Wireless Multimedia Sensors Networks

Chaouki Araar
Electronics Departement
University Frères Mentouri Constantine
1325 Ain El Bey Way, Constantine,
25017, Algeria
chaouki_wil23@live.fr

Abstract—A pruning-based discrete cosine transform (DCT) approximation is proposed. It is based on pruning the parametric eight-point transform. The new proposed transform is investigated in terms of processing time, energy consumption and image quality for image compression in the context of wireless multimedia sensor networks (WMSNs). Our experimental results showed that the proposed algorithm provided a significant savings in processing time and energy with competitive performance in image quality. Digital hardware architecture is proposed and implemented on 40-nm Xilinx Virtex-6 XC6VSX475T-2ff1156 FPGA device. The designed architecture presents high efficiency in terms of clock speed, hardware resources utilization and power consumption. Therefore, a large set of WMSN applications requiring low bitrates can benefit from the proposed scheme when embedded into multimedia sensor nodes.

Keywords— WMSNs, image compression, approximate DCT, pruning approach, energy conservation, VLSI, FPGA

I. INTRODUCTION

Advances in low-cost CMOS camera have fostered their integration into sensor node in wireless sensor networks (WSNs), which then led to the emergence of a new class of WSNs, referred to in the literature as wireless multimedia sensor networks (WMSNs) [1]. WMSNs consist of a network of tiny, battery-powered camera nodes with a limited embedded processing, storage and communication capabilities. Each camera node has the ability to capture large volumes of images from the monitoring area, process them and transmit wirelessly processed data over the network to the sink. Unfortunately, the huge amount of visual flows manipulated by the resources-constrained camera node requires higher power consumption and higher bandwidth usage. In other words, the sensor node will spend more energy which causes the battery power to drop faster and hence, affect the whole network lifetime. One of the most effective solutions to reduce energy dissipation and to use the bandwidth in more efficiency way is to compress images at source. Habitually, image compression is an ideal choice to overcome the bandwidth and energy dissipation bottlenecks.

Therefore, the well-known JPEG standard compression algorithm can be used in this context.

Over the past years, JPEG compression method has gained widespread acceptance; and it is most often used in image compression. Nevertheless, this method is very time and energy consuming in WMSNs context [2]. This is due to the encoder complexity; and mainly to the DCT transform step which consumes more than 60% of the computation requirements of the JPEG algorithm [3]. Thus, adapting the JPEG encoder to the resource-constrained WMSNs by reducing the DCT complexity seems to be the most perfect solution to gain in time and energy. In this way, we proposed a low complexity DCT transform which combines an approximate DCT with a pruning approach. The aim of the former is to reduce the draconian computational demands of the DCT in terms of number of arithmetic operations, whereas the latter aims to compute only coefficients that are likely to be significant. Such a DCT design, leads to a drastic reduction in the computational cost of the compression process and thereby the energy dissipated by the camera sensor node.

In this article, our goals are: (1) to focus on the adaptation of the standard JPEG compression algorithm to the requirement of WMSNs, by proposing a considerable reduction of the computational cost of the DCT transform; and (2) to design an efficient and low-complexity 2-D pruned DCT digital hardware architecture aiming to relieve the main processor of the camera node from the computational complexity of the DCT.

II. BACKGROUND

A. Discrete Cosine Transform

The forward one-dimensional N-point DCT [4] of N input samples $x(0), \dots, x(N-1)$ is defined as seen in (1):

$$X(k) = \sqrt{\frac{2}{N}} \alpha(k) \sum_{n=0}^{N-1} x(n) \cdot \cos \frac{(2n+1)k\pi}{2N} \quad (1)$$

with $k = 0, 1, \dots, N-1$ and $\alpha(k)$ is defined as follows:

$$\alpha(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } k = 0 \\ 1 & \text{otherwise} \end{cases}$$

The above-mentioned equation can be easily formulated as matrix multiplication according to:

$$X = C \cdot x \quad (2)$$

where x is an $N \times 1$ input vector and C is an $N \times N$ matrix called the kernel matrix of the transformation whose basis vectors are sampled cosines, given by:

$$C(k, n) = \sqrt{\frac{2}{N}} \cdot \alpha(k) \cdot \cos\left[\frac{(2n+1)k\pi}{2N}\right], \text{ for } k, n = 0, 1, \dots, N-1$$

It can be easily verified that the kernel matrix C is orthogonal, i.e., it fulfill the condition $C^{-1} = C^T$. So, the inverse 1-D DCT transformation which is defined as a mapping from the frequency-domain to the spatial domain can be also written as matrix multiplication:

$$x = C^T \cdot X \quad (3)$$

Since the DCT is separable transform, the 2-D DCT transformation can be obtained by means of two separate 1-D DCTs. First, 1-D DCT transform is carried out over every row to get the 1-D DCT coefficients. Then, another 1-D DCT is run again over each column which yields to the 2-D DCT coefficients, see Fig. 1. This process is called row-column decomposition and it can be easily formulated in matrix notation as follows:

$$F = C \cdot X \cdot C^T \quad (4)$$

where X is an $N \times N$ input matrix, C is the $N \times N$ kernel matrix, F is the corresponding transform-domain $N \times N$ output matrix and superscript “ T ” denotes the matrix transpose operation.

By virtue of orthogonality, the reverse 2-D DCT transformation can be expressed as presented in the following:

$$X = C^T \cdot F \cdot C \quad (5)$$

B. DCT Approximation

A DCT approximation is a transformation which provides a close estimation of the exact DCT computation at low-complexity requirements. It has the following mathematical structure:

$$C = D \cdot T \quad (6)$$

where D is a scaling diagonal matrix typically constituted of irrational numbers in the form of $1/\sqrt{n}$ (n is a small positive integer value), and T is a low-complexity matrix which contains elements only in the set $\{0, \pm 1/2, \pm 1, \pm 2\}$ [5,6].

The irrational numbers in the diagonal matrix requires high computational complexity which makes the DCT approximation computationally expensive. Fortunately, the diagonal matrix D does not introduce any extra computational complexity, since it can be absorbed into the quantization step [6-9]. Therefore, the low-complexity matrix T is the only source of the computational complexity in the approximation.

According to (6), the forward 2-D DCT approximation can be written as below:

$$F = C \cdot X \cdot C^T = D \cdot (T \cdot X \cdot T^T) \cdot D \quad (7)$$

The same idea can be used for the inverse 2-D DCT approximation, thus, in accordance with (3), it can be expressed as:

$$X = C^T \cdot F \cdot C = T^T \cdot (D \cdot F \cdot D) \cdot T \quad (8)$$

III. PROPOSED PRUNED DCT APPROXIMATION

A. Pruned Bouguezal-Ahmad-Swamy approximate DCT

In [8], a one-parameter 8-point approximate transform is proposed and it is referred to as BAS-2011. Its transform matrix is obtained by introducing an arbitrary parameter in the kernel matrix of the transformation reported in [7], and by carrying out some row permutations to improve the transform energy compaction capability. In the case when the arbitrary parameter a is taken equal to $1/2$, the resulting transform is given by:

$$C^{(a)} = D^{(a)} \cdot T^{(a)} = D^{(0.5)} \cdot T^{(0.5)} \quad (9)$$

where D is a diagonal matrix given by:

$$D = \text{diag}\left(\frac{1}{\sqrt{8}}, \frac{1}{2}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{8}}, \frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{5}}\right) \quad (10)$$

and

$$T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & -1 & -1 \\ 1 & 1/2 & -1/2 & -1 & -1 & -1/2 & 1/2 & 1 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 1 & -1 \\ 1/2 & -1 & 1 & -1/2 & -1/2 & 1 & -1 & 1/2 \end{bmatrix} \quad (11)$$

The associated fast algorithm of the transform matrix T requires only 18 additions and 2 bit-shifts operations, see Fig. 2(a).

The proposed k -pruned version of the transform presented in (9) is expressed as follows:

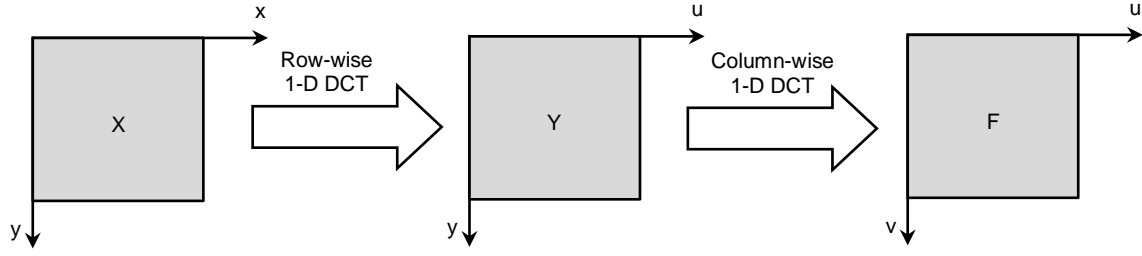


Fig. 1. Block diagram of the 2-D DCT computation using row-column decomposition.

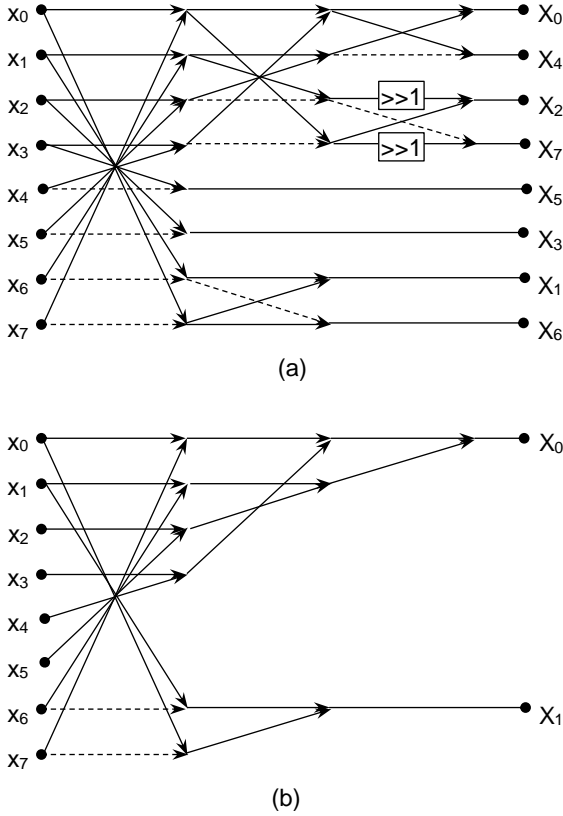


Fig. 2. Signal flow graph for the BAS-2011 [8] and proposed pruned BAS-2011 ($k = 2$) transform matrices. (a) Original (18 adds. and 2 bit-shifts). (b) Proposed (10 adds.). Dashed arrows mean multiplications by -1.

$$C_k = D_k \cdot T_k \quad (12)$$

where D_k is a $k \times k$ scaling diagonal matrix, T_k is a semi-orthogonal $k \times 8$ pruned transform matrix, given respectively by:

$$D_k = \text{diag}(D_{0,0}, D_{1,1}, \dots, D_{k-1,k-1}) \quad (13)$$

and

$$T_k = \begin{bmatrix} t_{0,0} & t_{0,1} & \dots & t_{0,7} \\ t_{1,0} & t_{1,1} & \dots & t_{1,7} \\ \vdots & \vdots & \ddots & \vdots \\ t_{k,0} & t_{k,1} & \dots & t_{k,7} \end{bmatrix} \quad (14)$$

with, $t_{i,j}, \{i, j = 0, 1, \dots, 7\}$ are the entries of T_k and $0 < k \leq 7$

B. Choice of the Pruning Value k

The problem to be addressed in this section is: what value of k has one to choose? In other words, how far it is possible to decrease the pruning value k to meet the WMSNs constraints without severely degrading the transform performance, i.e., the visual image quality?

With an 8×8 image block, after applying the 2-D DCT transform, the important low frequency components gathers towards the top-left corner while the less important high frequency components are at the bottom-right corner. That's means that most of important AC coefficients are in the 2×2 square-shape around the DC coefficient. Thusly, by means of analyzing the grey-scale set of 512×512 standard test images available from [10], we noticed that the 2-D BAS-2011 DCT transform [8] can concentrate 96.05 % of the average total image energy in the 2×2 low frequency transform-domain coefficients. These four significant low-frequency coefficients contain an important fraction of the total image energy and they are almost sufficient to represent the entire image information. Moreover, in image compression applications, after quantization process, long sequences of zeros are produced, especially at the higher-frequencies [11]. Further, since many of WMSN applications do not strongly require high quality imagery and due to the limited transmission energy and bandwidth, images are compressed at high compression ratio, i.e., at low bitrates (≤ 0.5 bpp) [12, 13]. The more an image is compressed at low bitrates, the more DCT coefficients are considered less important (close or equal to zero). So, by judiciously throw away the unimportant information (i.e., high frequency components) and restrict the computation to the 2×2 low-frequency coefficients, computational efforts may be saved which also results in low power consumption with less bandwidth usage and thus, extending the network lifetime. On the basis of the above facts, the pruning value k equal to 2 is an appropriate choice, which is also suggested in other works [14, 15]. Consequently, the proposed pruned BAS2011 in (12) will be:

$$C_2 = D_2 \cdot T_2 \quad (15)$$

where D_2 is a 2×2 scaling diagonal matrix, T_2 is a semi-orthogonal 2×8 pruned transform matrix, given respectively by:

$$D_2 = \text{diag}\left(\frac{1}{\sqrt{8}}, \frac{1}{2}\right) \quad (16)$$

and

$$T_2 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & -1 & -1 \end{bmatrix} \quad (17)$$

Fast algorithm for the proposed T_2 transform matrix is illustrated graphically as shown in Fig. 2 (b).

Based on the row-column decomposition method presented in Fig. 1, the 2-D $k \times k$ pruned DCT approximation transform can be carried out with two stages of 1-D k -pruned approximate DCT transforms, by performing eight row-wise calls of 1-D k -pruned DCT to the 8×8 input block followed by k column-wise calls of 1-D k -pruned DCT to the intermediate result, see Fig. 3. Therefore, the proposed forward and inverse 2-D $k \times k$ pruned DCT approximations ($k = 2$) are given respectively by:

$$F_2 = C_2 \cdot X_8 \cdot C_2^T = D \cdot (T_2 \cdot X_8 \cdot T_2^T) \cdot D \quad (18)$$

$$X_2 = C_8^T \cdot F_2 \cdot C_8 = T_8^T \cdot (D \cdot F_2 \cdot D) \cdot T_8 \quad (19)$$

C. Complexity Assessment

Let $AC_{1-D}(T_k)$ be the arithmetic complexity of T_k . Therefore, the arithmetic complexity of the 2-D $k \times k$ pruned DCT can be given by:

$$AC_{2-D}(T_k) = 8 \cdot AC_{1-D}(T_k) + k \cdot AC_{1-D}(T_k) \quad (20)$$

For the particular case when k equal to 2, we can derive the following expression:

$$AC_{2-D}(T_2) = 10 \cdot AC_{1-D}(T_2) \quad (21)$$

Comparison in terms of arithmetic complexity of the proposed algorithm with pruned and non-pruned DCT transforms is given in Table 1. Selected methods include: (1) the BinDCT-C8 [16]; (2) the Bouguezel–Ahmad–Swamy (BAS2011) [8]; and (3) the zonal BinDCT-C8 [15]. We also assessed the complexity of the exact DCT according to Arai-Agui-Nakajima (AAN) fast algorithm [17].

It is obvious from Table 1 that the proposed pruned 2-D BAS2011 DCT exhibit lower additive and bit-shift complexities than the original 2-D BAS-2011 DCT [8]. The proposal requires 68.757% less operations overall, while it maintains good energy compaction (96.05% of the average total image energy). Moreover, the proposed method outperforms the recently proposed zonal BinDCT-C8 presented in [15], requiring 100% less bit-shift operations and 16.67% less arithmetic additions for both 1-D and 2-D versions.

TABLE I. COMPUTATIONAL COMPLEXITY ASSESSMENT OF DIFFERENT 1-D AND 2-D DCT METHODS.

Method	1-D			2-D		
	Mult.	Add.	Shift	Mult.	Add.	Shift
AAN [17]	5	29	0	80	646	0
BinDCT-C8 [16]	0	24	5	0	384	80
BAS2011($\alpha = 0.5$) [8]	0	18	2	0	288	32
Zonal BinDCT-C8 [15]	0	12	1	0	120	10
Proposed	0	10	0	0	100	0

IV. PERFORMANCE EVALUATION

The evaluation was carried out by considering the microcontroller AMTEL ATmega128 of Mica2 and MicaZ motes, operating at 8 MHz with an active power consumption of 22mW [18]. Furthermore, we consider the quantization table and the Huffman encoder given in the Annexes of the JPEG standard [19].

A. Image Quality

Our quantitative evaluation is performed using the peak signal-to-noise ratio (PSNR) as figures of merit. To this end, a set of well-known gray-scale 512×512 test images namely, Lena, Baboon, Peppers, Barbara and Living room taken from [10] are used.

PSNR is given in decibels (dB) and is defined as:

$$PSNR = 10 \times \log \left(\frac{d^2}{MSE} \right) \quad (22)$$

where d is the maximum possible pixel value of the uncompressed image (255 when gray-scale image is considered). MSE is computed using (23).

$$MSE = \frac{1}{N \times M} \sum_{i=1}^M \sum_{j=1}^N (x_{ij} - y_{ij})^2 \quad (23)$$

Table 2 reports the PSNR measure for the selected images encoded at very low bitrate (0.26 bpp) according to the proposed pruned BAS-2011 and zonal BinDCT-C8 [15].

Fig. 4 shows a visual image quality evaluation of the proposed algorithm. PSNR measure is also included for comparison. Reconstructed Lena, Barbara, Peppers, Baboon and Living room test images encoded at 0.26 bpp are depicted in the figure, from which we can clearly see that the proposed method has a competitive visual quality performance as

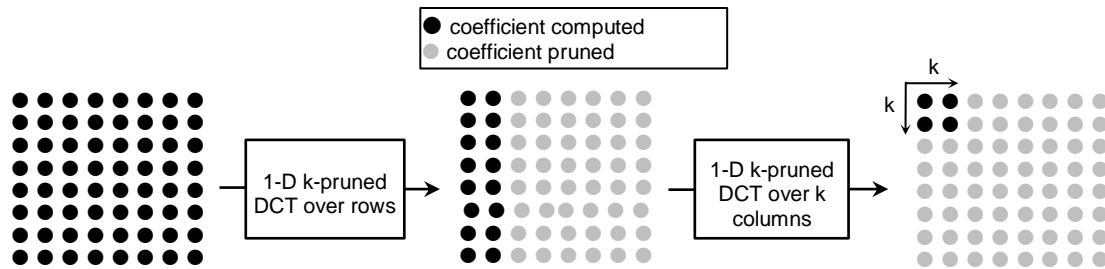


Fig. 3. The 2-D $k \times k$ pruned DCT with row-column decomposition (pruning value $k = 2$).



Fig. 4. Reconstructed Lena, Barbara, Peppers, Baboon and Living room test images given at bitrate of 0.26 bpp. (a) Exact DCT [17]. (b) Zonal BinDCT-C8 [15]. (c) Proposed method.

TABLE II. PSNR OF GREY-SCALE 512×512 TEST IMAGES GIVEN AT 0.26 BPP FOR EXACT DCT [17], ZONAL BinDCT-C8 [15] AND PROPOSED SCHEME.

Image	Lena	Peppers	Living room	Baboon	Barbara
Exact DCT [17]	30.98	30.48	27.20	24.12	24.65
Zonal BinDCT-C8 [15]	27.59	26.03	24.14	22.78	22.90
Proposed	28.00	26.16	24.27	20.70	22.90

compared with the exact DCT [17] and zonal BinDCT-C8 [15].

B. Processing Time and Energy Consumption

The processing time and energy consumption can be computed according to (24) and (25):

$$T = C/F \quad (24)$$

$$E = T \times P \quad (25)$$

where C is the computation cycles given by simulation, F and P expresses respectively, the frequency and the active power consumption of the microcontroller.

Computation cycles C is obtained using WinAVR20100110; an open source software development tools for the Atmel AVR series of RISC microprocessors. Compilation is performed with « -O3 » optimization setting.

Table 3 gives respectively, computation cycles, execution time and energy consumption related to 8×8 block of pixels according to the proposed method and other DCT transforms.

As expected, the software-based simulation results proved clearly the efficiency of our scheme not only on arithmetic costs, but in both execution time and energy consumption. Performance evaluation showed up to 50%, 60.87% and 91.18% execution time and energy savings when compared to other methods, namely zonal BinDCT-C8 [15], zonal BinDCT-C7 [14] and JPEG (IJG fast) [18].

V. PROPOSED 2-D PRUNED DCT DIGITAL HARDWARE ARCHITECTURE

A. System Hardware Architecture

In this section, efficient and low-complexity 2-D pruned DCT digital hardware architecture is designed. The proposed

TABLE III. DCT CYCLE COUNTS, EXECUTION TIME AND ENERGY CONSUMPTION RELATED TO BLOCK ON THE ATMEGA128 MICROCONTROLLER.

Method	DCT (cycles)	Execution Time (ms)	Energy (μ J)
JPEG (IJG fast) [18]	8131	1.02	22.44
Zonal BinDCT-C7 [14]	1840	0.23	5.06
Zonal BinDCT-C8 [15]	1437	0.18	3.96
Proposed	713	0.09	1.98

architecture employs two realizations of 1-D pruned DCT approximation module and one transpose buffer block in-between, all directed by an FSM-based (finite state machine) control and command unit. A basic sketch of the proposed system is depicted in Fig. 5.

The two 1-D pruned DCT modules can be either the proposed pruned BAS2011 approximate DCT described earlier in the paper or the competitor zonal BinDCT-C8 [15]. Fig. 6 depicts the designed digital hardware architecture for the proposed pruned BAS2011 DCT approximation and zonal BinDCT-C8 [15], which accepts both 8 parallel data inputs and produces 2 parallel data outputs. The fine-grain pipeline is adopted in the designed architectures in order to increase the throughput rate. Three 1-stage pipeline registers are inserted at strategic places within each fast algorithm for purpose to minimizing the critical path delay.

B. FPGA Implementation

The 2-D pruned DCT proposed hardware architecture was described in VHDL, and the Xilinx ISE tool 14.7 version is used to synthesize and implement the design on a 40-nm Xilinx Virtex-6 XC6VSX475T-2ff1156 FPGA device.

The FPGA implementation was evaluated for hardware complexity, power consumption and clock speed performance considering the following metrics: the number of occupied configurable logic blocks (CLB), look-up table (LUT) and Flip-Flop (FF) counts, the minimum clock period (T_{min}), in ns. The maximum clock frequency, which was determined by the inverse of the clock period as $F_{max} = (T_{min})^{-1}$, in MHz. Reported values were obtained after the place-and-route (PAR) process from Xilinx device utilization summary and post-PAR static timing report. Using the Xilinx XPower Analyzer, we estimated the frequency normalized dynamic power (D_p) consumption, in mW/MHz, with 100 Mhz clock frequency. We also reported area-time (AT) complexity and area-time-square (AT^2) complexity for comparison, where, area is measured as the number of consumed CLBs and time as the minimum clock period. The AT-complexity and AT^2 -complexity are considered to be important measures with which VLSI circuits are assessed. The values for each design are listed in Table 4.

By analyzing the obtained results, it can be seen that the design associated to the proposed pruned BAS2011 DCT approximation presents the best performances in terms of hardware resources utilization, clock speed and power consumption. These results are pinpointed by all the considered metrics quoted in Table 4.

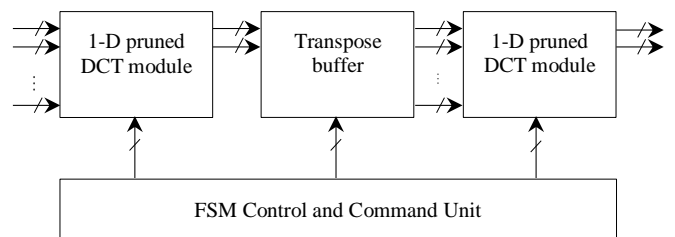
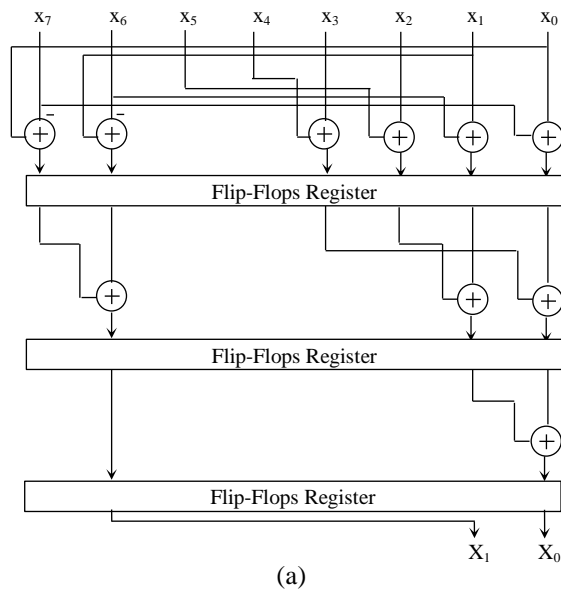
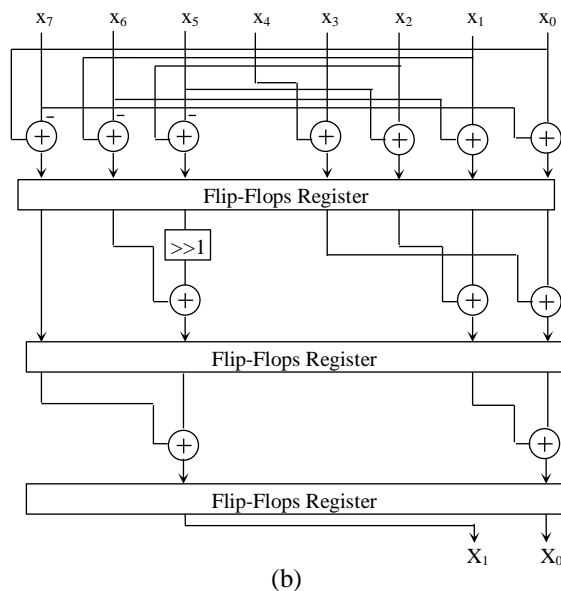


Fig. 5. Digital architecture of 1-D pruned DCT approximation. (a) Proposed pruned BAS2011. (b) zonal BinDCT-C8 [15].



(a)



(b)

Fig. 6. Digital architecture of 1-D pruned DCT approximation. (a) Proposed pruned BAS2011. (b) zonal BinDCT-C8 [15].

It appears clear from the table that the proposal is the fastest hardware architecture operating at clock frequency of 438.982MHz and consumes 20.6% less CLBs, 16.1% less LUTs and 11.1% less FFs as compared to the architecture linked to the zonal BinDCT-C8 [15]. We noticed also an 8% savings in power consumption metric which is considered as a main major parameter to consider in the design of WMSNs. In regard to AT and AT^2 -complexity, the design based on the proposed pruned BAS2011 approximate DCT exhibits the best results.

VI. CONCLUSION

In the present article, a low complexity DCT transform which is suitable for energy conservation in WMSNs is introduced. The proposed transform which was based on the combination of a DCT approximation with a pruning technique, achieved computational savings by computing only the most low frequency coefficients. Quantitative and qualitative image quality assessments were performed, showing the competitiveness of the proposed method. The experimental works conducted on the Atmel ATmega128 microcontroller revealed that our algorithm is very promising and guarantees significant savings in processing time and energy consumption. We designed dedicated digital hardware architecture for the proposed pruned DCT approximation and provided its implementation on 40-nm Xilinx Virtex-6 FPGA technology. An examination of the VLSI implementation of the proposed method in terms of hardware resources utilization, clock speed and power consumption provided the best performance results.

Application Specific Integrated Circuit (ASIC) implementation in CMOS 45 nm technology is planned in the future in order to demonstrate the practicability of the proposed transform and the designed architecture in digital VLSI realization.

REFERENCES

- [1] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A survey on wireless multimedia sensor networks," *Comput. Netw.*, vol. 51, no. 4, pp. 921–960, 2007., doi: 10.1109/MWC.2007.4407225.
- [2] Ferrigno, L. Ferrigno, S. Marano, V. Paciello, and A. Pietrosanto, "Balancing computational and transmission power consumption in wireless image sensor networks," *Proc. IEEE Int. Conf. VECIMS 2005*, Giardini Naxos, Italy, July 2005, pp. 6, doi:10.1109/VECIMS.2005.1567564.
- [3] C. N. Taylor, D. Panigrahi, and S. Dey, "Design of an adaptive architecture for energy efficient wireless image communication," *Lecture notes in computer science*, pp. 260–273, 2002, https://doi.org/10.1007/3-540-45874-3_15.
- [4] K. R. Rao and P. Yip, "Discrete cosine transform: algorithms, advantages, applications," San Diego, CA, USA: Academic, 1990.
- [5] T. I. Haweel, "A new square wave transform based on the DCT," *Signal Processing*, vol. 82, pp. 2309–2319, 2001, [https://doi.org/10.1016/S0165-1684\(01\)00106-2](https://doi.org/10.1016/S0165-1684(01)00106-2).
- [6] K. Lengwehasatit and A. Ortega, "Scalable variable complexity approximate forward DCT," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 11, pp. 1236–1248, Nov. 2004, doi: 10.1109/TCSVT.2004.835151.
- [7] S. Bouguezel, M. O. Ahmad and M. N. S. Swamy, "Low-complexity 8×8 transform for image compression," *Electronics Letters*, vol. 44, pp. 1249–1250, Sept. 2008, doi: 10.1049/el:20082239.
- [8] S. Bouguezel, M. O. Ahmad and M. N. S. Swamy, "A low-complexity parametric transform for image compression," *2011 IEEE International Symposium of Circuits and Systems (ISCAS)*, 2011, pp. 2145–2148, doi: 10.1109/ISCAS.2011.5938023.
- [9] R. J. Cintra and F. M. Bayer, "A DCT approximation for image compression," *IEEE Signal Processing Lett.*, vol. 18, no. 10, pp. 579–582, Oct. 2011, doi: 10.1109/LSP.2011.2163394.

TABLE IV. HARDWARE RESOURCES UTILIZATION, CLOCK SPEED AND POWER CONSUMPTION ON 40-NM XILINX VIRTEX-6 XC6VSX475T-2FF1156 DEVICE

method	CLB	LUT	FF	T_{min} (ns)	F_{max} (Mhz)	D_p (mW/Mhz)	AT	AT^2
Zonal BinDCT-C8 [15]	136	349	442	2.446	408.831	0.25	332.66	813.68
Proposed	108	293	393	2.278	438.982	0.23	246.00	560.44

- [10] J. Kominek, Waterloo BragZone, University of Waterloo. Available at <http://links.uwaterloo.ca/Repository.html>
- [11] K. R. Rao, P.C. Yip, "The Transform and Data Compression Handbook," CRC Press LLC, Boca Raton, 2001.
- [12] M. Tausif, N. R. Kidwai, E. Khan and M. Reisslein, "FrWF-Based LMBTC: Memory-Efficient Image Coding for Visual Sensors, " in *IEEE Sensors Journal*, vol. 15, no. 11, pp. 6218-6228, Nov. 2015, doi: 10.1109/JSEN.2015.2456332.
- [13] N. R. Kidwai, E. Khan and M. Reisslein, "ZM-SPECK: A Fast and Memoryless Image Coder for Multimedia Sensor Networks, " in *IEEE Sensors Journal*, vol. 16, no. 8, pp. 2575-2587, April15, 2016, doi: 10.1109/JSEN.2016.2519600.
- [14] A. V. Phamila, R. Amutha, "Low complexity energy efficient very low bit-rate image compression scheme for wireless sensor network, " *Inf. Process. Lett.*, vol. 113, no. 18, pp. 672-676, 2013, doi:10.1016/j.ipl.2013.06.008.
- [15] G. Suseela, A. V. Phamila, "Low-complexity low-memory energy-efficient image coding for wireless image sensor networks, " *The Imaging Science Journal*, vol. 66, no. 2, pp. 125-132, 2018, doi: 10.1080/13682199.2017.1385175.
- [16] J. Liang and T. D. Tran, "Fast multiplierless approximations of the DCT with the lifting scheme," in *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 3032-3044, Dec. 2001, doi: 10.1109/78.969511.
- [17] Y. Arai, T. Agui, and M. Nakajima, "A fast DCT-SQ scheme for images," *Trans. IEICE*, vol. E-71, no. 11, pp. 1095-1097, Nov. 1988.
- [18] D. Lee, H. Kim, M. Rahimi, D. Estrin and J. D. Villasenor, "Energy-Efficient Image Compression for Resource-Constrained Platforms," in *IEEE Transactions on Image Processing*, vol. 18, no. 9, pp. 2100-2113, Sept. 2009, doi: 10.1109/TIP.2009.2022438.
- [19] ISO/IEC 10918-1/ITU-T Recommendation T.81, "Digital compression and coding of continuous-tone still images," available at <http://www.jpeg.org/jpeg/>.

Extreme Learning Machine for the smart faults regression of turbofan engine

N. REZKI, W. REZGUI

wail.rezgui@enst.dz, nafissa.rezki@enst.dz,

* LTI Laboratory, ENST, Algiers, Algeria

Abstract— In recent decades, Huang et al. proposed a new learning algorithm for single-layer neuronal networks (SLFN) called Extreme Learning Machine (ELM). Compared to classic learning algorithms, ELM offers an extremely fast learning speed, better generalization performance and minimal human intervention. The ELM has been successfully applied either for classification and even regression problems.

In this work, we implemented the Single Layer Feed-Forward (SLFF) neuron networks for the regression and prediction of defects affecting Turbofan Engine. For the learning of this network, we used the ELM algorithm. For the results comparison, we implemented in parallel the same network type but with another learning algorithm which is back-propagation. As evaluation criteria, we have chosen the accuracy and time required for learning. Finally, the results show that the ELM learning of a SLFF is more perform than the BP learning of a SLFF.

Index Terms— Feed-forward neural network, regression, Extreme learning machine, Back-propagation, Turbofan Engine, Accuracy, Time learning and testing.

I. INTRODUCTION

Neuronal network is the combination of a set of formal neurons, presented in a graph. This network is distinguished either by architecture, or by the level of complexity such as the number of neurons and also the presence or absence of feedback loops, or also by the type of neurons i.e. the activation functions, or by the target objective: supervised learning or not, optimization, dynamic systems.

Feed-forward neuronal is a very well-known type of neuronal networks in recent decades, have been widely used in many areas through their benefits, such as complex non-linear mappings approximation capacity directly from learning data, and also ability to offer models for many natural and artificial phenomena, which are difficult to deal with conventional parametric techniques. However, this type of network requires a setting of all its parameters following the existing dependence between its different layers, which make the learning process of this slower network.

Single-Layer-Feed-Forward, the purpose of this work, is one of the most basic and popular feed-forward neuronal networks through their learning and fault tolerance capabilities. Despite the advantages of this network, but the problem still lies in its learning phase even if we use the most popular learning algorithms such as Forward Propagation, Computational Graphs, Backward Propagation with Gradient Descent [9], Quasi-Newton Method [10], Conjugate Gradient [11], and Levenberg-Marquardt Algorithm [12-13]. So the learning phase of this network remains relatively slow in the best situations, since in some cases it can also easily remain blocked in a local minimum.

Extreme Learning Machine (ELM) is a new fast learning algorithm developed recently to improve the effectiveness of SLFNs. With this new ELM algorithm, the learning settings (weight and bias) hidden nodes can be determined in a random and independent manner. The output weights of the network can be determined analytically by a simple general reverse operation. The ELM has been successfully applied for classification and *regression* problems with many applications in the real world [14-17].

Vs. The classical learning algorithms, the major difficulty lies in the manual setting of the control parameters as the learning rate, the learning periods, etc. and also local minima. On the other hand, the ELM put an automatic adjustment of these parameters without any users' intervention. This feature makes ELM faster compared to other traditional methods. However, the classification of the ELM cannot be optimal because the learning parameters of the hidden nodes are assigned randomly in the beginning, then they remain unchanged during the learning phrase.

The main objective of this work is to do a comparative study to choose the best learning algorithm of a Single Layer Feed-Forward neuron network between ELM and Back-propagation. The criteria on which we will choose the best algorithm are learning time and accuracy. Finally, for the validation and test of the chosen algorithms, we will use the data from Turbofan Engine.

II. SINGLE LAYER FEED FORWARD NEURAL NETWORK FOR REGRESSION

A. SLFF-ELM for regression

For a regression problem, if we have a given set of learning samples $S = \{(x_i, y_i) \text{ with } i = 1, \dots, N\} \subset \mathbb{R}^d \times \mathbb{R}$, then, the output of a SLFN trained by the ELM can be represented as follows [1-4]:

$$f_L(x) = \sum_{i=1}^L h_i(x)\beta_i = h(x)\beta, \quad \text{where } x \in \mathbb{R}^d \quad (1)$$

With:

- L Number of hidden nodes.
- $h_i(x)$ with $i = 1$ to L : the function of the hidden node.
- β_i with $i = 1$ to L : is the output weight between the hidden node and the output node.

Remarks:

- Hidden calculation nodes can be sigmoid, additive and radial-based functions (RBF)...
- To minimize the error of the least squares, we can get the objective loss function:

$$\min_{\beta} \|H\beta - Y\|^2 \quad (2)$$

Where:

$$H = [h(x_1)^T, h(x_2)^T, \dots, h(x_N)^T]^T \quad (3)$$

$$Y = [y_1, y_2, \dots, y_N]^T \quad (4)$$

Two solutions of the objective loss function are provided by Huang et al., Where they are equivalent on the basis of the Woodbury identity:

$$\beta = \begin{cases} (H^T H)^{-1} H^T Y, & \text{for } N \geq L \\ H^T (H H^T)^{-1} Y, & \text{for } N < L \end{cases} \quad (5)$$

The ELM is easily adjusted on the basis of the principle of minimizing the empirical risk (ERM) in the objective loss function. The complexity of the ELM is controlled by limiting the output weights β to small values. Thus, the objective function is:

$$\min_{\beta} : \frac{1}{2} \|\beta\|_2^2 + C \frac{1}{2} \sum_{i=1}^N \|\xi_i\|_2^2 \quad (6)$$

$$\text{s.t.} : h(x_i)\beta = y_i - \xi_i, \quad \forall i$$

Where c is the regularization parameter that compensates for the standard of output weights and less square learning errors.

Finally, the solution of the objective function is:

$$\beta = \begin{cases} (H^T H + \frac{1}{c})^{-1} H^T Y, & \text{for } N \geq L \\ H^T (H H^T + \frac{1}{c})^{-1} Y, & \text{for } N < L \end{cases} \quad (7)$$

B. SLFF-BP for regression

The single layer is the most basic and successful neuronal network architecture with feed-forward network topologies in industrial applications, while the most used supervised learning technique for training this type of networks is the back-propagation algorithm. The back-propagation term means the process by which the network error derivatives are returned to the network and used to adjust the weights, and decrease the error by each iteration. In this way, back-propagation offers a method of minimizing errors between the calculated outputs and the desired outputs [5-6].

There are generally four steps to developing a BP neural network: *Step 1* for pre-processing where input data and targets are defined. *Step 2* for building the network where the dataset has been divided into training, validation and test subsets. *Step 3* for the network learning phase. *Step 4* for the test phase where the trained network was used to simulate the test data.

The generalized neuronal regression network is a neuronal network architecture that can solve any function approximation problem with sufficient data. The architecture of this network consists of four layers: an input layer, a hidden layer, a summation layer and an output layer.

The expected value of the output y considering the input vector x is given by:

$$E \left[\frac{y}{x} \right] = \frac{\int_{-\infty}^{+\infty} y f(x, y) dy}{\int_{-\infty}^{+\infty} y f(x, y) dy} \quad (8)$$

If the density $f(x, y)$ is unknown, it must be estimated from a sample of x and y observations, with n is the number of observations of the sample, p is the vector dimension of the variable x and σ the smoothing factor.

$$\hat{f}(x, y) = \frac{1}{n(2\pi)^{p+1/2} \sigma^{p+1}} \sum_{i=1}^n \left[\exp \left[-\frac{(x-x_i)^T (x-x_i)}{2\sigma^2} \right] \exp \left[-\frac{(y-y_i)^2}{2\sigma^2} \right] \right] \quad (9)$$

The squared distance between the input vector x and the x_i learning vector is defined as:

$$D_i^2 = (x - x_i)^T (x - x_i) \quad (10)$$

Finally, the output is determined by:

$$\hat{f} = \frac{\sum_{i=1}^n y_i \exp \left(-\frac{D_i^2}{2\sigma^2} \right)}{\sum_{i=1}^n \exp \left(-\frac{D_i^2}{2\sigma^2} \right)} \quad (11)$$

The coefficient of determination R^2 between the modeled output and measures of the training and testing data set with observed value T_i , predicted value A_i and average value T is:

$$R^2 = \frac{\sum_{i=1}^n (A_i - \bar{T})^2}{\sum_{i=1}^n (T_i - \bar{T})^2} \quad (12)$$

III. SIMULATION RESULTS

A. Case study

For the implementation of the selected algorithms, we have chosen as a case study the aircraft engine turbo machinery. A benchmark of this system, called C-MAPSS dataset, is available on the official site of NASA (National Aeronautics and Space Administration). This C-MAPSS database is divided into four subsets from FD001 to FD004 of several multivariate time series, and each subset is divided into two parts, one for learning and the other for testing phases. Each subset has 26 columns: motor number, time step, three operational sensor settings and 21 sensor measurements. Each subassembly is derived from the motors operating under normal starting conditions. But, they are different in terms of manufacture and also in degrees of initial wear, which is unknown to the data analyzer. Over the course of the time series and at some point, all engines begin to degrade. The main objective is to predict the correct *RUL* value for each engine in the test sets. Figure 1 shows the aircraft engine turbo machinery with its main elements [7-8].

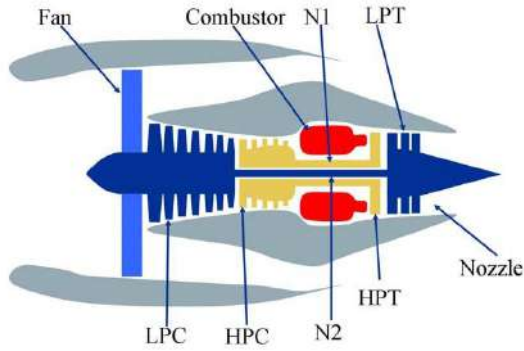


Figure 1. Simplified diagram of engine simulated in C-MAPSS [7].

B. For SLFF-ELM-LR

The implementation of the SLFF algorithm trained by ELM for linear regression results the following figures:

Figure 2 shows the learning phase of the chosen algorithm. The objective of this step is to propose a function by the chosen algorithm, where the learning base is defined as an input of this function, and the Turbofan engine *RUL* base as an output. Linear regression based on the SLFF-ELM algorithm results this function:

$$\text{Output} = 0.3 * \text{Target} + 33$$

Where:

- ✓ Output: is the predefined *RUL* of Turbofan engine.

- ✓ Target: is the training data from sensors installed on Turbofan engine
- ✓ Regression Slope is 0.3
- ✓ Regression Offset is 33
- ✓ Determination coefficient (which has to measure the prediction quality of a linear regression) is 0.6. Therefore, we have a good determination coefficient since it is closer to one compared to zero.

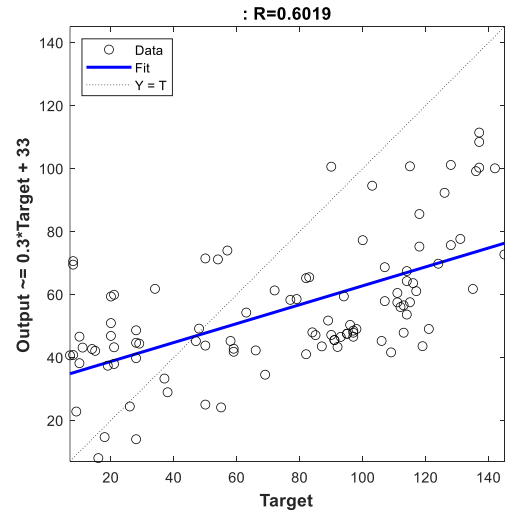


Figure 2. Learning phase of SLFF-ELM-LR

Figure 3 aims to graphically show the test phase of the chosen algorithm that is trained in the previous step. This new step aims to propose a function by the chosen algorithm, where the test base is defined as an input of this function. The displayed output presents the calculation made by the algorithm chosen to determine the *RUL* of the Turbofan engine. The linear regression based on the SLFF-ELM algorithm in the test phase results this function:

$$\text{Output} = 0.43 * \text{Target} + 22$$

With:

- ✓ Output is the calculated *RUL* of the Turbofan engine
- ✓ Target: test data from sensors installed on Turbofan engine.
- ✓ Regression slope is 0.43
- ✓ Regression Offset is 22
- ✓ Determination coefficient is 0.7, therefore, we conclude that we have a good determination coefficient since it is closer to one than to zero.

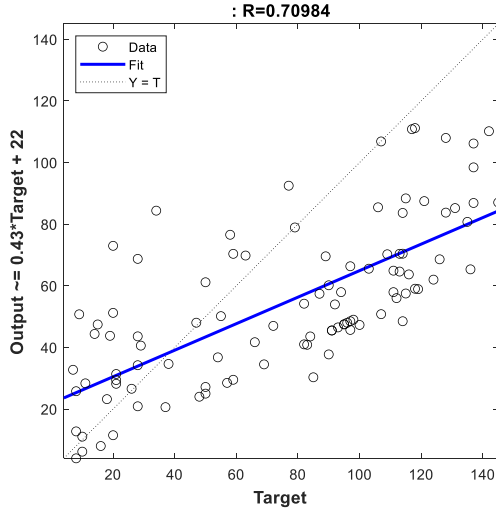


Figure 3. Test phase of SLFF-ELM-LR

C. For SLFF-BP-LR

Figure 4 shows the learning phase of the selected algorithm *SLFF-BP-LR*, which has the objective of determining the mathematical function that defines the relationship between the cloud of the learning base as an input, and the base of the RUL of the Turbofan engine as output. The linear regression based on the SLFF-BP algorithm results this function:

$$Output = 0.0063 * Target + 76$$

With:

- ✓ Output is RUL of the Turbofan engine
- ✓ Target: data from sensors installed on Turbofan engine
- ✓ The slope of the regression is 0.0063
- ✓ The regression lag is 76
- ✓ The coefficient of determination is 0.1. We therefore have a poor coefficient of determination since it is closer to zero than to one.

The objective of this figure 5 is to test the linear regression model based on *SLFF-BP* trained in figure 4. Then, the resulting function, which is well presented in the figure 5, has as an input the test base of the Turbofan engine, and then it provides an output that presents the calculated *RUL* of the Turbofan engine. The function has the following formula:

$$Output = 0.006 * Target + 76$$

With :

- ✓ Output is the calculated *RUL* of Turbofan engine.

- ✓ Target: test data from the sensors installed on Turbofan engine
- ✓ The regression slope is 0.006
- ✓ The regression Offset is 76
- ✓ The determination coefficient is 0.12. We therefore have a poor coefficient of determination because it is closer to zero than to one.

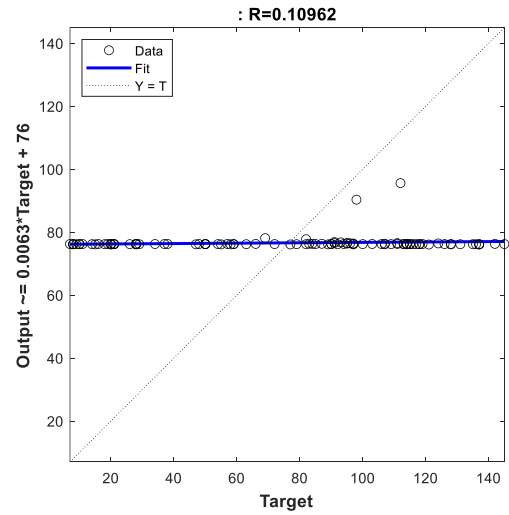


Figure 4. Learning phase of SLFF-BP-LR

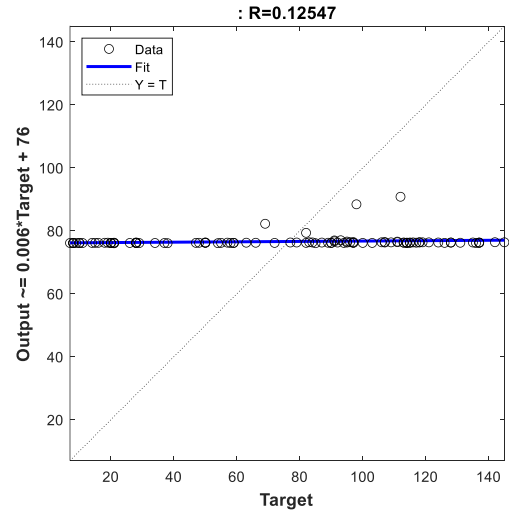


Figure 5. Testing phase of SLFF-BP-LR

D. SLFF-ELM-LR vs SLFF-BP-LR

This section aims to make a comparative study between the implementation of the two algorithms *SLFF-ELM-LR* and *SLFF-BP-LR*. We used as evaluation criteria the time needed for learning and

testing, the accuracy, and the coefficient of determination. We can summarize the results obtained in the following table:

	SLFF-ELM-LR		SLFF-BP-LR	
	Training	Testing	Training	Testing
Time	0.0207	0.0047	2.0968	0.1101
Accuracy	39.0478	36.3620	41.3763	41.3671
r	0.6	0.7	0.1	0.12
m	0.3	0.43	0.0063	0.006
b	33	22	76	76

With a simple *comparison of the results*, we notice that:

1) For the time criterion

For our study case, the simulation of the two algorithms bring out a very important difference in terms of learning and testing time between the two algorithms, such that: for the SLFF-ELM-LR algorithm, we found that it takes 0.0207 s for learning and 0.0047 s for testing. In the other side, for SLFF-BP-LR algorithm, we found that it takes for learning 2.0968 s and for testing 0.1101 s. Therefore, we noticed that for both learning and testing phases, the SLFF-ELM-LR algorithm is the fastest compared to the SLFF-BP-LR algorithm.

2) For the accuracy criterion

The implementation of both algorithms showed us that the accuracy results for both algorithms are really very close. So, we scored for the SLFF-ELM-LR algorithm, a learning accuracy reaching 39.04 % and a testing accuracy reaching 36.36 %. In the other side, the accuracy of the SLFF-BP-LR algorithm, for learning reach 41.37% and for testing reach 41.36%. Thus, for this criterion, we notice that for our case of study, the SLFF-BP-LR algorithm keeps the best results in terms of accuracy.

3) For the determination coefficient

- For *SLFF-ELM-LR*, the new function developed in the test phase is with an increase in the slope value, a decrease in the offset value, on the other hand the most important thing is that the new function developed in the test phase is with a determination coefficient closer to one, this means that the chosen algorithm is efficient thanks to good learning achieved.
- For *SLFF-BP-LR*, the new function developed in the test phase kept almost the values of its parameters determined in the learning phase such as the regression slope, the offset regression, and even the coefficient of determination. And the

problem lies in this last parameter, since its value is close to zero than one, therefore, it means that the chosen algorithm is not performing well due to a bad learning performed.

IV. CONCLUSION

In this paper, we have performed a comparative study between two learning algorithms: Extreme Learning Machine and Back-propagation of the same type of neural networks, which is a Single Layer Feed-Forward for regression. For this comparative study, we have chosen Turbofan Engine as a case study, and as evaluation criteria: time required for learning and testing phases, accuracy, and coefficient of determination.

The obtained results show that the SLFF trained by BP achieves the best accuracy rate compared to the SLFF trained by ELM. On the other hand, the SLFF trained by ELM, either for learning or testing, firstly achieves the best coefficient of determination, and secondly, is faster than the SLFF trained by BP.

Finally, we can conclude that regression by SLFF-ELM is more efficient in terms of time and quality of prediction (strong). On the other hand, we found that SLFF-BP is more efficient in terms of precision.

V. REFERENCES

- [1]. G-B. Huang, H. Zhou, X. Ding and R. Zhang, Extreme Learning Machine for Regression and Multiclass Classification, IEEE Transactions on systems, man, and cybernetics—part b: cybernetics, pp. 513-529, Vol. 42, April 2012.
- [2]. S. Ding, X. Xinzhen, R. Nie, Extreme learning machine and its applications, Neural Comput & Applic, pp. 549–556, Vol. 25, 2014.
- [3]. K. Chena, Q. Lva, Y. Lub, Y. Doua, Robust Regularized Extreme Learning Machine for Regression Using Iteratively Reweighted Least Squares, Neurocomputing, pp. 345-358, Vol. 230, March 2017.
- [4]. Q. Yu, Y. Miche, E. Eirola, M-v. Heeswijk, E. Severin, A. Lendasse, Regularized extreme learning machine for regression with missing data, Neurocomputing pp. 45–51, Vol.102, 2013.
- [5]. A. Konaté, H. Pan, N. Khan & J. H. Yang, Generalized regression and feed-forward back propagation neural networks in modelling porosity from geophysical well logs, Journal of Petroleum Exploration and Production Technology, pp. 157–166, Vol. 5, 2015.
- [6]. T. D. Parsons, A. A. Rizzo, and J. G. Buckwalter, Backpropagation and Regression: Comparative Utility for Neuropsychologists, Journal of Clinical and Experimental Neuropsychology, pp. 95–104, Vol. 26, 2004.
- [7]. A. Saxena, K. Goebel, D. Simon, N. Eklund, Damage Propagation Modeling for Aircraft Engine Run-to-Failure

- Simulation, in the proceeding of International Conference on Prognostics and Health Management, 2008.
- [8]. A. L. Ellefsen, E. Bjørlykhaug, V. Æsøy, S. Ushakov, H. Zhang, Remaining useful life predictions for turbofan engine degradation using semi-supervised deep architecture, *Reliability Engineering & System Safety*, Vol. 183, pp. 240-251, March 2019.
 - [9]. M. Andrychowicz, M. Denil, S. G. Colmenarejo, M. W. Hoffman, D. Pfau, T. Schaul, B. Shillingford, N. de Freitas, Learning to learn by gradient descent by gradient descent, in the proceeding of 30th Conference on Neural Information Processing Systems, Barcelona, Spain, 2016.
 - [10]. A. Likas, A. Stafylopatis, Training the random neural network using quasi-Newton methods, *European Journal of Operational Research*, pp. 331–339, Vol. 126, 2000.
 - [11]. A. Wanto, M. Zarlis, Sawaluddin, D. Hartama, J. T. Hardinata, H. F Silaban, Analysis of Artificial Neural Network Backpropagation Using Conjugate Gradient Fletcher Reeves In The Predicting Process, in the proceeding of International Conference on Information and Communication Technology, 2017.
 - [12]. G. Lera and M. Pinzolas, Neighborhood Based Levenberg–Marquardt Algorithm for Neural Network Training, *IEEE Transactions on neural networks*, Vol. 13, No. 5, September 2002.
 - [13]. A. Reynaldi, S. Lukas, H. Margaretha, Backpropagation and Levenberg-Marquardt Algorithm for Training Finite Element Neural Network, in the procedening of UKSim-AMSS 6th European Modelling Symposium, 2012.
 - [14]. X. Ding ,J. Liu, F. Yang , J. Cao, Random compact Gaussian kernel: Application to ELM classification and regression, *Knowledge-Based Systems*, Vol. 217, 6 April 2021.
 - [15]. Y.Miche, A. Sorjamaa, P. Bas, O.Simula; C.Jutten, A. Lendasse, OP-ELM: Optimally Pruned Extreme Learning Machine, vol. 21, No. 1, January 2010.
 - [16]. H. Rong, G. Huang, N. Sundararajan; P. Saratchandran, Online Sequential Fuzzy Extreme Learning Machine for Function Approximation and Classification Problems, Vol. 39, No. 4, August . 2009.
 - [17]. O. Altay, M. Ulas, K. Alyamac , DCS-ELM: a novel method for extreme learning machine for regression problems and a new approach for the SFRSCC, *PeerJ Computer Science*, 12 March , 2021

Extreme Learning Machine vs Back-Propagation for the faults smart classification of Tennessee Eastman Process

W. REZGUI, N. REZKI

wail.rezgui@enst.dz, nafissa.rezki@enst.dz,

LTI Laboratory, ENST, Algiers, Algeria

Abstract— Currently, Neural Network is a very important tool for the data classification, regression and clustering. In the literature, there are several types of this network. The two most known are recurrent and feed-forward. The recurrent is especially for the regression, and feed-forward is for the classification. The learning procedure of the neural network can be subdivided into two broad categories: supervised (adv. Back-propagation, Extreme learning machine...) and unsupervised (adv. Propagation) learning.

In this paper, we have presented a comparative study between two learning algorithms: the Extreme Learning Machine (ELM) and Back-Propagation (BP) of the Single Layer Feed-Forward (SLFF) type of neural network. For the implementation of this study, we have chosen as a case study the Tennessee Eastman Process (TEP) and as evaluation criteria: the time required and the accuracy of both the learning and testing phases.

Index Terms— Feed-forward neural network, Extreme learning machine, Back-propagation, Tennessee Eastman Process, Accuracy, Time learning and testing.

I. INTRODUCTION

An artificial neural network (ANN) model has a structure similar to the human brain that computes parallel information processing [11]. A neural network system consists of interconnected neurons. The network architecture is a set of input nodes, computing units and output nodes. The input nodes are the only nodes that input information into the network, but they do not execute any computational paradigm. In this network presented in figure 1, the input nodes are called input layer, the set of output nodes is called output layer, while the remaining layers without direct connections from or to outside are called hidden layers. In a layered network, all nodes in one layer are connected to all other nodes in the next layer. The ANN is a powerful tool for solving the non-linearity imposed by the various constraints and complexity of industrial systems. ANN based classifiers can incorporate both structural and statistical information and achieve better performance than minimum distance classifiers. In the literature, several types of neural networks exist, among them

we cite: Stochastic Neural Network [12], Self-Organizing Map [13], Recurrent Neural Networks [14] and feed-forward Neural Networks [15], the object of our work.

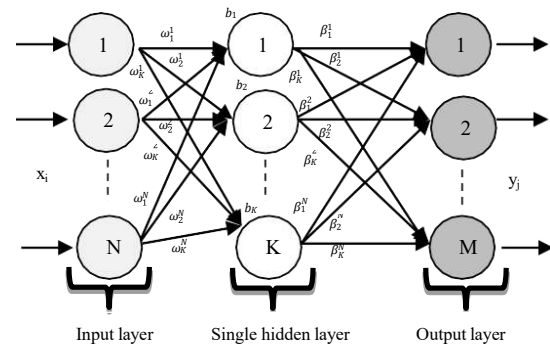


Figure. 1. Single hidden layer feed forward neural network

A feed-forward neural network is a biologically inspired classification algorithm. It consists of a number of simple neuron-like processing units organized in layers. Each unit in a layer is connected to all units in the previous layer with weights. The weights on these connections encode the knowledge of a network. Data enters at the inputs and travels through the network, layer by layer, until it reaches the outputs. In normal operation, i.e. when acting as a classifier, there is no feedback between the layers. Two known architectures in this type of neural networks: Single layer feed-forward [16] and Multi-layer feed-forward [17].

Generally, for Single Layer Feed-Forward neural network (SLFF) learning, we mainly use the back-propagation (BP) algorithm. It is used for supervised learning in neural networks. It is designed to minimize the mean square error (MSE) between the actual output of a SLFF and the desired output. Using this type of learning algorithm can help industrial systems to learn efficiently on a small learning set, since BP is gradient descent based learning, it leads to slow convergence during learning for a large scale learning set.

Also for SLFF learning, a new method called extreme learning machine (ELM) is proposed in [1,4]. This method is a new supervised learning algorithm spatially for single hidden layer feed-forward where the weights of connections between input variables and neurons in the hidden layer and the bias of neurons in the hidden layer are randomly assigned. But, on the other hand, the weights of the connections between the hidden layer neurons and the output neurons are obtained using the generalized Moore-Penrose (MP) inverse, considering that the activation function of the output neuron is linear.

The main objective of this work is the implementation of the SLFF algorithm trained by ELM on a benchmark of the TEP system. For the comparison of the results, another parallel implementation of the SLFF algorithm trained by Back-propagation is made. To make a good comparative study, we will develop for each algorithm: the learning phase, the test phase, confusion matrix and error histogram.

II. EXTREME LEARNING MACHINE

A. ELM definition

ELM (Extreme Learning Machines) is invented in 2006 by G. Huang. It's a learning algorithm for a feed-forward neural network type, which can tends to provide good generalization performance at extremely fast learning speed, because it uses Moore-Penrose generalized inverse to set its weights [1-6].

B. ELM algorithm

The ELM algorithm consists of these steps:

- Step 1. Randomly assign weights ω and bias b .
- Step 2. Calculate hidden layer output Matrix H

$$H = \begin{pmatrix} g(\omega_1^1 x_1 + b_1) & \cdots & g(\omega_K^1 x_1 + b_K) \\ \vdots & & \vdots \\ g(\omega_1^N x_N + b_1) & \cdots & g(\omega_K^N x_N + b_K) \end{pmatrix} \quad (1)$$
- Step 3. Calculate output weight matrix $\beta = H^\dagger Y$, where:
 - Y : Training data target matrix.
 - H^\dagger : Moore-Penrose generalized inverse of a matrix H .
- Step 4. Calculate the new target matrix Y :
$$Y = H\beta \quad (2)$$
- Step 5. Calculate the optimization objective function.

$$\|H\beta - Y\| = \min_{\beta} \|H\beta - Y\| \quad (3)$$

III. BACK-PROPAGATION NEURAL NETWORK

A. Back-propagation definition

Back-propagation is a type of artificial neural network learning algorithm. It is based on the fine-tuning of the neural network weights according to the error rate obtained in the previous iteration, whose the objective is to reduce this error rates and make the model reliable by increasing its generalization [9-10].

B. Back-propagation algorithm

The Back-propagation algorithm consists of these steps:

- Step 1. Randomly assign weights ω , β and bias b .
- Step 2. Calculate the output for every neuron from the input layer, to the hidden layer, to the output layer.

$$S_j = \sum_{i=1}^Z g(p_i x_i + b_i) \quad (4)$$

Where:

- $Z = K$ or M
- g activation function
- $p = \omega$ or β
- Step 3. Calculate the error ε between calculated outputs Y_c and desired outputs Y_d .
$$\varepsilon = Y_c - Y_d$$
- Step 4. Travel back from the output layer to the hidden layer to adjust the weights such that the error is decreased.
- Step 5. Keep repeating the process until the desired output is achieved

IV. CASE STUDY

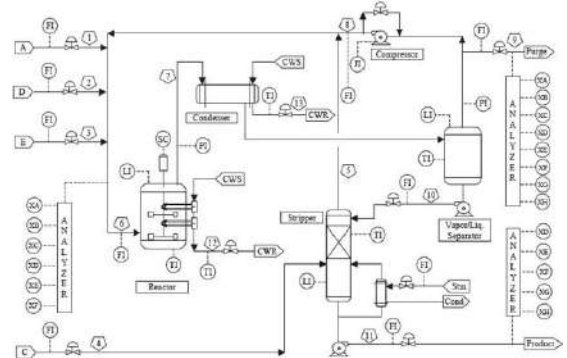
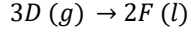
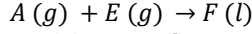
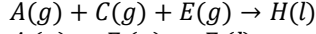
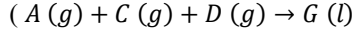


Figure 2. Tennessee Eastman control problem [18].

The authors of [18] offer the Tennessee Eastman process as a simulated model for evaluating industrial complex process monitoring technologies. A

condenser, a separator, a reactor, a compressor, and a stripper are the five main components (Figure 2).

The reactor receives four gaseous reactants (A, C, D, and E) as well as inert B. It generates two components (G and H) as well as the undesired by-product F. The following are the reaction equations:



The reactions are all irreversible, exothermic, and roughly first-order in terms of reactant concentrations. The reaction rates are expressed as a temperature-dependent Arrhenius function. The reaction that produces G has higher activation energy than the reaction that produces H, resulting in greater temperature sensitivity.

The TEP presented in [18] is open loop unstable and should be used in a closed loop configuration. The procedure was proposed by the authors of [19] as a plant-wide control mechanism. This control structure is used in this article to evaluate the performance of Extreme Learning Machine and Back-Propagation on fault diagnosis.

The condenser cools the reactor product stream before it is supplied to a vapor-liquid separator. Through a compressor, the vapor exits the separator and is recycled to the reactor feed. To prevent inert and by-product accumulation, a portion of the recycle stream is purged. The separator's condensed component is transferred to a stripper, which removes the residual reactants. G and H are transferred to a downstream process which is not depicted in the diagram after exiting the stripper's base.

Finally, the inert and by-products are vaporized and purged from the vapor-liquid separator. XMEAS (1) to XMEAS (41) and XMV (1) to XMV (12) are the 41 measured and 12 modified variables, respectively, in the process.

Twelve pre-programmed faults (IDV (1) to IDV (12)) that have been pre-programmed are used to depict various process operation situations.

V. SIMULATION RESULTS

This section aims to present the results obtained after the implementation of two algorithms SLFF trained by ELM and SLFF trained by BP. The comparative study between the two chosen algorithms is based on the development of four stages: 1) learning phase, 2) test phase, 3) confusion matrix, and 4) error histogram. The realization of these four steps helped us to determine two evaluation criteria: 1) the time required for the learning and testing phases, and 2) the accuracy obtained also for the learning and testing phases.

A. For Learning phase

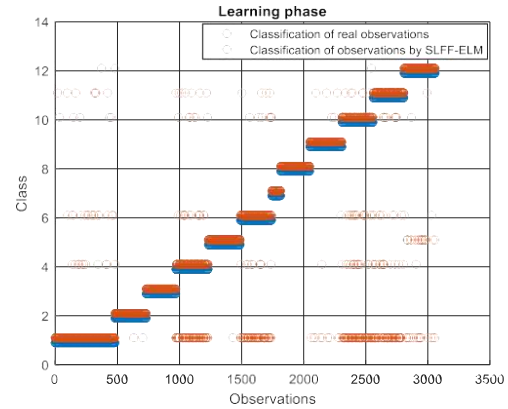


Figure 3. Learning phase of SLFF-ELM

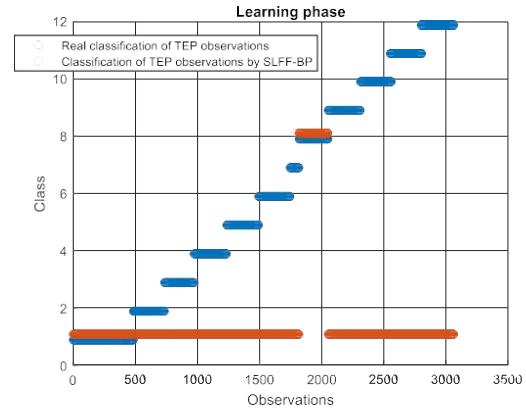


Figure 4. Learning phase of SLFF-BP

Figures 3-4 present successively the learning phase of the SLFF algorithm trained by ELM and SLFF trained by BP. The x-axes show the TEP observations, and the y-axes show the 12 TEP classes (pre-programmed faults). The cloud of blue points shows the actual classification of observations distributed over 12 classes. The cloud of red dots shows the classification proposed by the chosen algorithms of the observations distributed over 12 classes. For more details:

- In figure 3, we see that the first 500 observations (the x-axis) are really assigned to the first class, on the other hand, we found that the SLFF-ELM algorithm proposed another classification where most of the first 500 observations are classified to the first class with a few in classes 4, 6, 10, 11, 12 (see the first column in Figure 3).
- In figure 4, we see that the first 500 observations which are really assigned to the first class, the SLFF-BP algorithm has also assigned these first 500 observations to the first predicted class (good classification), but the problem lies in the other classes where he classified most of the

observations of the other classes into his first and eighth predicted classes (wrong classification).

B. For testing phase

After the learning phase, the objective now is to test the two chosen algorithms. Figures 5 and 6 present the results of the classification of the TEP test base. The x-axes show the TEP test observations, and the y-axes show the actual classes for the blue dot cloud and the predicted classes for the red dot cloud.

So the difference between the two figures is very clear. Therefore, this result is expected, since this phase is based on the learning phase. So the bad learning of SLFF-BP algorithm gives bad classification of test base. For the SLFF-ELM algorithm, it makes a good classification for classes 1, 2, 3, 7, 8 and 9.

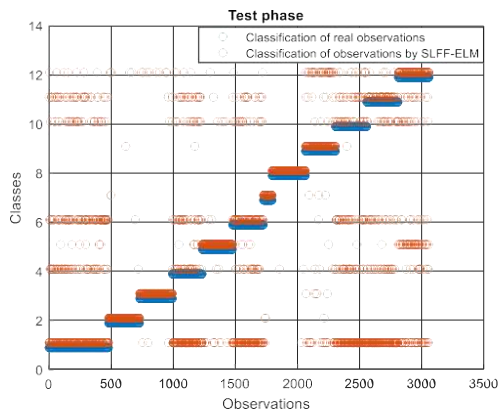


Figure 5. Test phase of SLFF-ELM

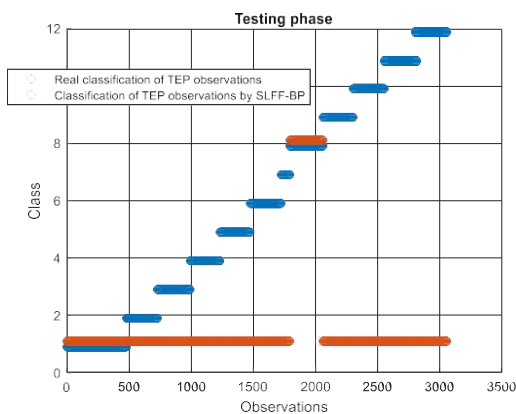


Figure 6. Testing phase of SLFF-BP

C. For Confusion Matrix

Figures 7-8 present the confusion matrices of the classification proposed successively by the SLFF-ELM and SLFF-BP algorithms. These matrices are numerical presentations of the graphs of the test phases. Each column of these matrices presents the classification by one of the two algorithms of the

observations of the real class, which has the same index as this column, on the different existing classes. Each line represents the set of observations (of the different real classes) assigned by one of the two algorithms to this new class which has the same index as this line. Also, the last column and the last line are reserved for the calculation of the percentages of well and badly classified observations.

Confusion Matrix													
Output Class	1	2	3	4	5	6	7	8	9	10	11	12	
1	256	0	3	129	10	114	1	20	127	130	42	80.7%	8.4%
2	0	241	0	0	0	0	2	0	1	0	0	98.8%	0.0%
3	0	0	249	0	0	0	0	0	5	0	0	98.0%	0.0%
4	57	1	0	31	2	22	0	2	4	27	25	13	16.8%
5	7	0	0	6	223	5	0	0	1	1	1	48	76.4%
6	81	1	0	31	2	72	0	0	1	36	21	11	28.1%
7	0	1	0	0	0	0	67	0	3	0	0	0	94.4%
8	0	0	0	0	0	0	0	264	0	0	0	0	100%
9	0	1	0	1	0	0	0	0	163	1	0	0	98.2%
10	32	0	1	16	4	18	0	0	5	26	21	15	8.8%
11	39	0	2	22	1	16	0	0	6	22	50	15	28.9%
12	7	5	8	4	1	2	0	0	33	11	5	105	8.0%
	53.4%	36.4%	4.7%	12.9%	1.8%	28.9%	5.7%	38.9%	37.4%	40.4%	19.8%	42.2%	67.2%
	46.6%	3.6%	5.3%	8.7%	8.2%	71.1%	4.3%	1.1%	32.6%	39.6%	20.2%	57.8%	42.8%
Target Class	1	2	3	4	5	6	7	8	9	10	11	12	

Figure 7. Confusion Matrix by SLFF-ELM testing

Confusion Matrix													
Output Class	1	2	3	4	5	6	7	8	9	10	11	12	
1	479	11	263	240	243	249	51	1	238	251	253	249	18.9%
2	0	0	0	0	0	0	0	0	0	0	0	0	NaN%
3	0	0	0	0	0	0	0	0	0	0	0	0	NaN%
4	0	0	0	0	0	0	0	0	0	0	0	0	NaN%
5	0	0	0	0	0	0	0	0	0	0	0	0	NaN%
6	0	0	0	0	0	0	0	0	0	0	0	0	NaN%
7	0	0	0	0	0	0	0	0	0	0	0	0	NaN%
8	0	0	0	0	0	0	0	0	0	0	0	0	NaN%
9	0	0	0	0	0	0	0	0	0	0	0	0	NaN%
10	0	0	0	0	0	0	0	0	0	0	0	0	NaN%
11	0	0	0	0	0	0	0	0	0	0	0	0	NaN%
12	0	0	0	0	0	0	0	0	0	0	0	0	NaN%
	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	24.4%
	0%	100%	100%	100%	100%	100%	100%	0%	100%	100%	100%	100%	75.6%
Target Class	1	2	3	4	5	6	7	8	9	10	11	12	

Figure 8. Confusion Matrix by SLFF-BP testing

For more details, for Figure 7, we note that:

- For the first column, we have 479 observations actually assigned to the first class (the x-axis). These particular observations are classified by the SLFF-ELM algorithm as follows: (256→C1), (57→C4), (7→C5), (81→C6), (32→C10), (39→C11), (7→C12). According to the last box of the first column, the percentage of correctly

classified observations expects 53.4% (256/479), and also the percentage of misclassified observations expects 46.6% (223/479).

- For the first row, the SLFF-ELM algorithm assigned to its first predicted class 833 observations where it took (256 from C1), (3 from C3), (129 from C4), (10 from C5), (114 from C6), (1 from C7), (1 from C8), (20 from C9), (127 from C10), (130 from C11), (42 from C12). According to the last box of the first line, the percentage of correctly classified observations expects 30.7% (256/833), and also the percentage of misclassified observations expects 69.3% (577/833).

Therefore, we see that:

- The SLFF-ELM algorithm gives a good classification reaching 57.2%, and a bad classification reaching 42.8%.
- The SLFF-BP algorithm gives a good classification reaching 24.4%, and a bad classification reaching 75.6%

D. For Error Histogram

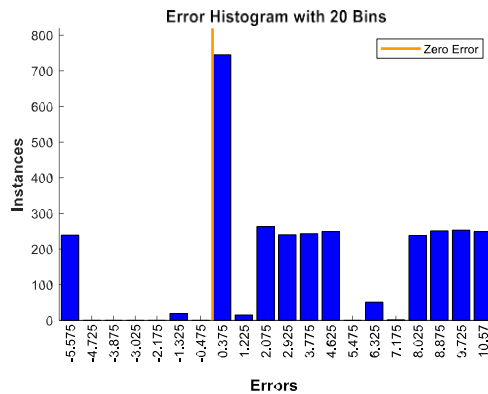


Figure 9. Error Histogram of SLFF-ELM testing

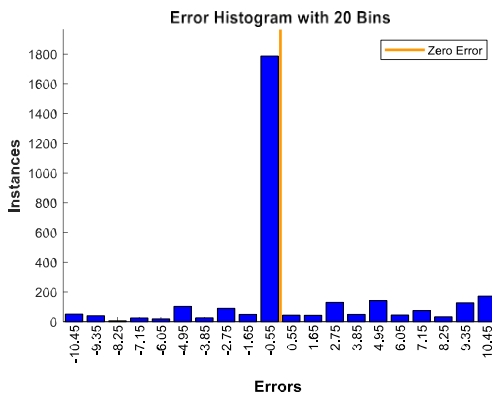


Figure 10. Error Histogram of SLFF-BP testing

Figures 9-10 successively present the error histograms of the test phases of two algorithms

SLFF-ELM and SLFF-BP. These are histograms of the errors between the actual classes and the predicted classes after the SLFF learning phases either by ELM or by BP. The error value indicates how much the predicted classes differ from the actual classes, so it can be negative. The Y-axis shows the number of observations from the TEP test base, grouped by error values. Zero error line corresponding to the zero error value on the error X-axis. In our case, the zero error point falls below the box with the center 0.55 for SLFF-ELM and 0.275 for SLFF-BP. Therefore, we see:

- For the classification by SLFF-ELM, the error ranges from -10.45 to 10.45, but the dominant error value is -0.55 which is for most misclassified observations.
- For the classification by SLFF-BP, the error varies from 0.275 to 10.73, but there are several dominant values of the error among them 10.37, 10.18, 9.075... which are really very important for most of the misclassified observations.

E. SLFF-ELM vs SLFF-BP

This point aims to summarize the work done previously in the following table. It clearly presents the two criteria chosen: the time and the precision of the learning and testing phases, to carry out the comparative study and choose the best classification algorithm between SLFF-ELM and SLFF-BP. As a result:

- For the classification by SLFF-ELM, we have for the learning time 0.7354s, and for the test time 0.282s, with an accuracy reaching 84.17% for the learning and 57.17% for the test.
- For the classification by SLFF-BP, we have for the learning time 0.3527s, and for the test time 0.0518s, with an accuracy reaching 23.42% for the learning and 24.38% for the test.

Learning Algorithm	Learning time	Testing time	Learning Accuracy	Testing Accuracy
SLFF-ELM	0.7354 s	0.2820 s	84.17%	57.17 %
SLFF-BP	0.3527 s	0.0518 s	23.42 %	24.38 %

The results presented in this table show that the learning of SLFF by BP is faster than its learning by ELM. On the other hand, ELM achieves the best accuracy rate, either for learning or testing.

VI. CONCLUSION

In this modest work, we carried out a smart classification of a very complicated TEP benchmark. This smart classification is made by one type of neural network which is Single Layer feed-forward trained by two algorithms: Back-propagation and Extreme Learning Machine. The objective to be achieved is to select the best algorithm according to

the two evaluation criteria: the time and the precision of the learning phase and also of the test phase. In order to simplify the determination of these two criteria, we have developed for each algorithm: the learning phase, the test phase, confusion matrix and error histogram. Finally, we can conclude that for the classification of TEP observations, the results show that, SLFF-BP provides a faster classification than SLFF-ELM for both phases (learning and testing). On the other hand, SLFF-ELM achieves better accuracy rate than SLFF-BP.

VII. REFERENCES

- [1]. Q-Y. Zhu, A.K. Qin, P.N. Suganthan, G-B. Huang, Evolutionary extreme learning machine, *Pattern Recognition*, pp. 1759–1763, Vol. 38, 2005.
- [2]. G-B. Huang, Q-Y. Zhu, C-K. Siew, Extreme learning machine: Theory and applications, *Neurocomputing*, pp. 489–501, Vol. 70, 2006.
- [3]. G-B. Huang, D. H. Wang, Y. Lan, Extreme learning machines: a survey, *International Journal of Machine Learning and Cybernetics*, pp. 107–122, Vol. 2, 2011.
- [4]. G-B. Huang, H. Zhou, X. Ding and R. Zhang, Extreme Learning Machine for Regression and Multiclass Classification, *IEEE Transactions on systems, man, and cybernetics—part b: cybernetics*, pp. 513–529, Vol. 42, April 2012.
- [5]. S. Ding, X. Xinzhen, R. Nie, Extreme learning machine and its applications, *Neural Comput & Applic*, pp. 549–556, Vol. 25, 2014.
- [6]. Y. Wang, Z. Li, L. Feng, C. Zheng, and W. Zhang, Automatic Detection of Epilepsy and Seizure Using Multiclass Sparse Extreme Learning Machine Classification, *Journal of Computational and Mathematical Methods in Medicine*, Vol. 2017, Jun 2017.
- [7]. M. Andrychowicz, M. Denil, S. G. Colmenarejo, M. W. Hoffman, D. Pfau, T. Schaul, B. Shillingford, N. de Freitas, Learning to learn by gradient descent by gradient descent, in the proceeding of 30th Conference on Neural Information Processing Systems, Barcelona, Spain, 2016.
- [8]. A. Likas, A. Stafylopatis, Training the random neural network using quasi-Newton methods, *European Journal of Operational Research*, pp. 331–339, Vol. 126, 2000.
- [9]. G. Lera and M. Pinzolas, Neighborhood Based Levenberg–Marquardt Algorithm for Neural Network Training, *IEEE Transactions on neural networks*, Vol. 13, No. 5, September 2002.
- [10]. A. Reynaldi, S. Lukas, H. Margaretha, Backpropagation and Levenberg-Marquardt Algorithm for Training Finite Element Neural Network, in the procedening of UKSim-AMSS 6th European Modelling Symposium, 2012.
- [11]. Oludare Isaac Abiodun, Aman Jantan, Abiodun Esther Omolara, Kemi Victoria Dada, Nachaat AbdElatif Mohamed, Humaira Arshad, State-of-the-art in artificial neural network applications: A survey, *Heliyon*, Vol. 4, Issue 11, November 2018.
- [12]. P. C. Bressloff, Stochastic neural field theory and the system-size expansion, *Society for Industrial and Applied Mathematics*, pp. 1488–1521, Vol. 70, No. 5, 2009.
- [13]. T. Kohonen, Essentials of the self-organizing map, *Neural Networks*, pp. 52–65, Vol. 37, 2013.
- [14]. A. Malhi, R. Yan, and R. X. Gao, Prognosis of Defect Propagation Based on Recurrent Neural Networks, *IEEE Transactions on instrumentation and measurement*, VOL. 60, NO. 3, March 2011.
- [15]. F. Haddadi, S. khanchi, M. Shetabi, V. Derhami, Intrusion Detection and Attack Classification Using Feed-Forward Neural Network, in the proceeding of the Second International Conference on Computer and Network Technology, Bangkok, Thailand, 23–25 April 2010.
- [16]. S. M. B. Baruah, D. Nandi, P. Gogoi & S. Roy, Primate vision: a single layer perception, *Neural Computing and Applications*, pp. 11765–11775, Vol. 33, 2021.
- [17]. M. Oczak and All, Classification of aggressive behaviour in pigs by activity index and multilayer feed forward neural network, *Biosystems Engineering*, pp. 89–97, Vol. 119, March 2014.
- [18]. V. Sylvain, T. Teodor, K. Abdessamad, Réseaux bayésiens pour l'identification de variables hors-contrôle. In the proceeding of the 5ème Conférence Internationale Francophone d'Automatique (CIFA'08), Bucarest, Roumanie, 2008.
- [19]. M. Douglas, Hawkins. Regression adjustment for variables in multivariate quality control, *Journal of Quality Technology*, pp. 170–182, Vol. 25, 1993.

Forecasting the New Cases of Tuberculosis in Algeria Using Random Forest as a Machine Learning Model

Djalila Tliba
Research Center on Environment
Annaba, Algeria
t-djali23@hotmail.com

Nadjet Frissou
Research Center on Environment
Annaba, Algeria
n.frissou@yahoo.fr

Mohamed Tahar Kimour
Research Center on Environment
Annaba, Algeria
rahatkimm@gmail.com

Abstract: Tuberculosis is a serious public health problem that remains one of the world's deadliest infectious diseases. It is difficult to control this disease that requires special strategies for the execution of immediate response and intervention actions. In this paper, we present a forecasting approach using random forest time series forecasting aiming to provide public health managers in Algeria a tool to rely on accurate predictions that can detect the evolving incidence phenomena of Tuberculosis. Recently, Random Forest (TB) Time series forecasting using machine learning algorithms has gained popularity. According to the results obtained, our proposed tool presented a high performance in relation to the error metrics, consisting of the best model compared to the most other forecasting techniques.

Keywords—*Forecasting, Tuberculosis, Time Series, Random Forest.*

I. INTRODUCTION

Tuberculosis (TB) is an airborne infectious disease caused by a bacteria named “Mycobacterium Tuberculosis” that usually affects the lungs, although it can affect almost any part of body. The Mycobacterium Tuberculosis is an acid-alcohol resistant bacillus (BAAR) that can be acquired by breathing contaminated air droplets coughed or sneezed by a person nearby who has active Tuberculosis [3] [13].

An estimated 9.9 million people have developed tuberculosis worldwide with a total of 1.5 million people died of TB in 2020 (214,000 according to WHO [16].

Tuberculosis remains a health issue in Algeria [4,12,17]. 20649 cases of tuberculosis were declared in 2019, a predominance of extra-pulmonary tuberculosis is notable with an incidence of 32.18 by 100.000 persons and 6392 cases of pulmonary tuberculosis with an incidence of 11.82 by 100.000 persons [17]. Thus, an accurate forecast of infectious tuberculosis based on predictive models is crucial for public health authorities to clearly understand its epidemic characteristics, track updates in advance, and choose the main response

actions related to the surveillance of the TB disease and the deployment of emergency support.

Recently, mathematical models have been used to project the future infectious diseases, and its impact on the future development of infectious diseases related morbidity and mortality [2]. In healthcare systems, decision maker needs efficient models and tools to facilitate the prevention and to predict the health resources and capacities for its handling. To this end, some notable works have been proposed in the literature, such as the grey system [5] the Holt-Winter methods [6-7], and the Box-Jenkins approach [9-12].

In this paper, we propose a new approach using random forest technique as an efficient tool to forecast TB in Algeria, while selecting Medea province as a pilot study area. Having proven to be the most efficient algorithm for prediction and classification, random forest algorithm [1] relies on multiple decision trees and accepts the results of the predictions from each tree. Based on the majority votes of predictions, it determines the final result.

The remainder of the paper is organized as follows: section 2 describes the material and methods, through the detailing the data and the process to handle it, and the development of the methods and tools we used to build the forecasting model. Section 3 presents the application of the proposed forecasting model on the collected data and highlights the findings. Finally, conclusion and future works are provided at section 4.

II. MATERIALS AND METHODS

A. Study area and Data collection

Medea is a province, in northern Algeria, localized at Lat.: 36°.28'N Lon.: 2°.73'E, Height: 1030 m. For Medea province, data for the monthly incidence of TB were obtained from January 2008 to December 2020, from various national data sources and research works such as those from [17] [4].

B. Statistical method

We developed a TB forecasting model using Random Forest regression that allowed over-dispersion of data. Mean monthly predicted cases were estimated through regression on multiple independent variables

that include monthly retrospective TB cases, mean temperature and precipitation. The forecasting model was developed using three processes: 1) data preprocessing and its descriptive analysis, 2) model construction and training using data from 2008–2018; and 3) model validation by forecasting cases in 2018–2020. Our statistical analysis was conducted with our implemented tool using Python 3.9.7.

C. Random Forest Method

Random forest is a supervised machine learning algorithm that constructs several decision trees. The final decision is made based on the majority of decision tree. Decision tree suffer from low bias and high variance. Random forest converts high variance to low variance. A random forest eradicates the limitations of a decision tree algorithm. It alleviates the overfitting of datasets while improves precision.

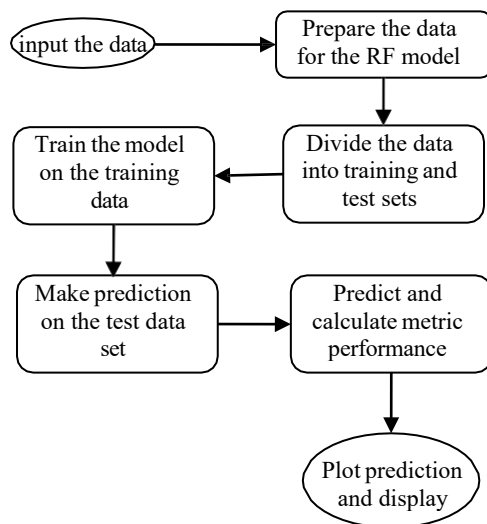


Fig.1. Machine learning process flow.

Random Forest (RF) constitutes a valuable tool for time series forecasting. It requires that the time series dataset be transformed into a supervised learning form first. The walk-forward validation is a specialized technique for evaluating the RF model.

RF model is an ensemble machine learning method proposed by Breiman [1]. It creates multiple CARTs, wherein each tree is trained on a bootstrap sample of the original training data using a randomly selected subset of input variables, and taking the average outputs of the CARTs as the final prediction. One of its most important features is the calculation of the variable importance, which measures the association between a given variable and the accuracy of the

prediction, based on the percentage of increase in the mean square-error.

A random forest algorithm is based on many decision trees. It generates such trees that constitute a ‘forest’. The algorithm is trained by means of bagging or bootstrap aggregating. Bagging has as function to improve the accuracy of machine learning algorithms through the increase of the number of trees to increase the precision of the outcome. Based on the predictions of the decision trees, the random forest algorithm establishes the outcome.

As a machine learning technique, Random Forest completes the task of learning from data with specific inputs to the machine. The random forest process starts with inputting training data into the selected algorithm. Training data being known or unknown data to develop the final random forest algorithm (Fig.1). The type of training data input does impact the algorithm, and that concept will be covered further momentarily.

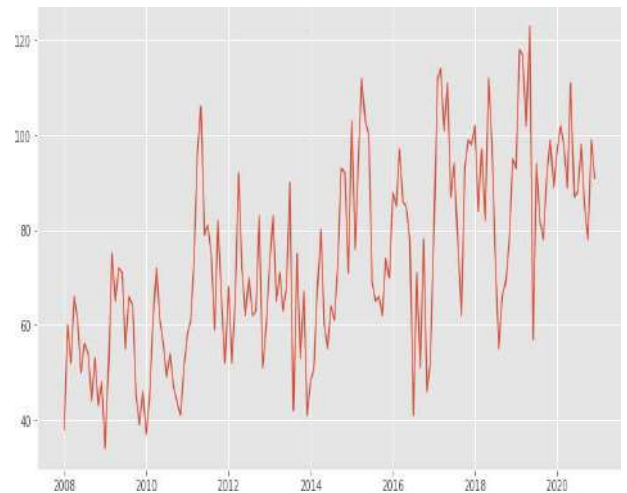


Fig.2. observed TB incidences.

III. RESULTS

From January 2008 to December 2020 there were 11475 tuberculosis notified cases in Algeria with a monthly average 74 cases. Fig.2 depicts the diagram of the tuberculosis incidence from 2008 to 2020. It shows in increase of such incidence from January 2014 and the peak is reached at May 2019 while the minimum number of cases is declared at January 2009.

We apply our model on the collected data as follows:

- 1- Data are pre-processed to deal with the missing values and the outliers.
- 2- Data are splitted into a training data (80%) and test data (20%).
- 3- The RF model is built using the training data.

- 4- The RF model is evaluated using the test data.
- 5- also called the Root Mean Square Deviation RMSE measures the average magnitude of the errors. It is concerned with the deviations from the actual value. When RMSE value is near the zero this indicates that the model has a perfect fit. When the RMSE is low, this implies that the model is better and its predictions are with good precision. When the RMSE is high, this indicates the presence of a large deviation from the residual to the ground truth. RMSE helps in figuring out if the feature is improving the model's prediction or not, thus, it can be used with different features as it.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

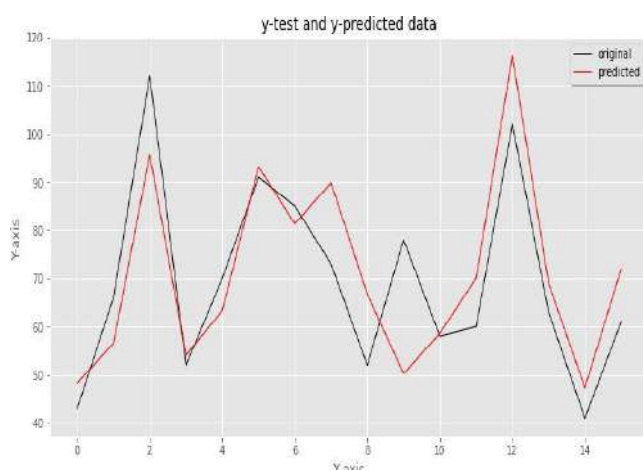


Fig.3. observed vs. predicted TB incidences.

We constructed a Random Forest model, which delivered a good accuracy in predicting the incidence of infectious TB with a RMSE of approximately 15%. Fig.3 depicts the observed vs. predicted incidences and shows that the two measures are very closed. The model may be used as an important tool by public health authorities.

IV. CONCLUSION AND FUTURE WORK

Tuberculosis remains a health public issue in Algeria. We have built a forecasting tool of tuberculosis cases by means of random forest as a most notable machine learning method in order to provide useful information with public healthcare managers to help them for formulating appropriate prevention and preparedness measures. Timely and accurate tuberculosis infectious disease forecasts could help public health responses for informing key preparation and mitigation efforts. Built on the collected TB

dataset, the random forest model demonstrated a high forecasting accuracy. It can be used to forecast the epidemic level, highlighting its potential and practical applicability. As future works, we plan to conduct exploring a random forest model for prediction TB incidences while taking into account the environmental and other variables, and for modeling other major infectious diseases.

REFERENCES

- [1] Breiman, L., 2001, Random forests. *Mach. Learn.*,45, 5–32
- [2] Juang W-C, Huang S-J, Huang F-D, et al. Application of time series analysis in modelling and forecasting emergency department visits in a medical centre in southern Taiwan. *BMJ Open* 2017;7: e018628
- [3] Gursimrat K Sandhu, Tuberculosis: Current Situation, Challenges and Overview of its Control Programs in India, *J Glob Infect Dis.* 2011 Apr-Jun; 3(2): 143–150.
- [4] L'Hadj, M., Selmane, S., "Time Series Analysis of Tuberculosis in Medea Province in Algeria". *Journal of Engineering Technology and Applied Sciences* 4 (2) 2019 : 85-94.
- [5] Zhao Y-F, Shou M-H, Wang Z-X. Prediction of the number of patients infected with COVID-19 based on rolling grey Verhulst models. *Int J Environ Res Public Health* 2020; 17:4582.
- [6] Guan P, Wu W, Huang D. Trends of reported human brucellosis cases in mainland China from 2007 to 2017: an exponential smoothing time series analysis. *Environ Health Prev Med* 2018; 23:23.
- [7] Zhang Y-Q, Li X-X, Li W-B, et al. Analysis and predication of tuberculosis registration rates in Henan Province, China: an exponential smoothing model study. *Infect Dis Poverty* 2020; 9:123.
- [8] Martínez-Bello DA, López-Quílez A, Torres-Prieto A. Bayesian dynamic modeling of time series of dengue disease case counts. *PLoS Negl Trop Dis* 2017;11: e0005696.
- [9] Wang Y, Xu C, Zhang S, et al. Temporal trends analysis of human brucellosis incidence in mainland China from 2004 to 2018. *Sci Rep* 2018; 8:15901.
- [10] Wang Y, Xu C, Li Y, et al. An advanced data-driven hybrid model of SARIMA-NNAR for tuberculosis incidence time series forecasting in Qinghai Province, China. *Infect Drug Resist* 2020; 13:867–80.
- [12] Scherhazad Selmane, Forecasting the Incidence of Tuberculosis in Algeria, *International Journal of Mycobacteriology*, 10(5), 2021.

- [13] Madhukar Pai et al., Tuberculosis, nature reviews disease primers, 27 October 2016
- [14] Permatasari Siltinga, Beti E. Dewib Alhadi Bustamam, Herleya Shaori Al-Ash, Evaluation of Dengue Model Performances Developed Using Artificial Neural Network and Random Forest Classifiers, Procedia Computer Science Volume 179, 2021, Pages 135-143
- [15] Anggraeni W, Nurmasari R, Riksakomara E, Samopa F, Wibowo RP, Condro L, et al. Modified Regression Approach for Predicting Number of Dengue Fever Incidents in Malang Indonesia, Procedia Computer Science (2017), p. 124
- [16] <https://www.afro.who.int>
- [17] <https://www.insp.dz>

Hierarchical routing protocols (Energy-aware) for wireless sensor network: a survey

Maroua Hammadi

department of computer science, Embedded systems laboratory
Badji Mokhtar University
Annaba, Algeria
marouahammadi@gmail.com

Mohammed Redjimi

department of computer science, Université 20 Août 1955
LICUS Laboratory
Skikda, Algeria
m.redjimi@Univ-skikda.dz

Abstract— Wireless sensor networks use hundreds or even thousands of micro sensors that capture information in their neighborhood and send it to the base stations for processing. These sensors sometimes operate in hostile environments out of human reach and they use batteries for functioning. Energy consumption and management in these sensors is a crucial challenge. Thus, several protocols propose algorithms management to save energy in these sensors. LEACH protocol (Low Energy Adaptive Clustering Hierarchy), the first hierarchical protocol, uses the notion of node's clustering of the network. In this paper we present a survey of the state-of-the-art of routing techniques in WSNs, the most recent research works, the advantages and disadvantages of each study

Keywords—*wireless sensor networks, routing protocols, hierarchical routing protocols, energy efficiency.*

I. INTRODUCTION

The number and scope of WSN applications has been expended thanks to their recent rapid development [1]. Smart grids and renewable energy systems are two popular examples of systems that require vast amounts of data to be collected. WSNs have a variety of uses, including assisting with peak load management and optimizing power generation resources. With the expanding size of smart grids, maintaining network performance, dependability, stability, and economy of scale have become increasingly difficult [1, 2]. Comparing to traditional communications in electrical power networks it is fair to conclude that WSNs have numerous advantages over them. In order to improve different aspects of electric power systems with applications in the three phases of electricity generation, distribution and consumption, the technology of WSNs is becoming more widely [3]. That makes them an important component of all complex electrical power systems. Sensor node devices collect data and transfer it to the sink, from where the user can subsequently access over the internet, as depicted in Figure 1. The challenge is to develop secure, energy efficient and adaptable WSN systems. Wireless sensors can be installed in client appliances and electric power meters to create a network capable of supplying real-time data on electricity consumption to customers. The challenge is to handle power use in the most optimized and efficient way [4]. A well placement of WSN nodes is impractical and comes with a high cost. That is why, once deployed, the wireless sensor nodes must be able to self-organize and integrate into an efficient wireless communication network [5]. Many IoT application turn to use WSNs to collect data thanks to their low cost and deployment flexibility [1, 6]. The properties and limits of WSNs must

be properly understood before they can be used successfully. The use of energy remains the most challenging constraint [7]. There is other WSN metrics such as memory storage efficiency, processing power, and data throughput. But since energy consumption and the lifetime of WSN are the most important restrictions, other metrics are used to evaluate the worth of WSN network protocols and algorithms. Numerous studies have been carried out to solve the WSN energy limits [8,9].

We compared numerous hierarchical routing techniques. Due to the importance of routing in WSNs, we have attempted to give a detailed review of various routing algorithms and protocols, as well as their effects on WSN performance. These protocols are classified into chains, trees, grids, and area-based networks. We examined and discussed numerous hierarchical routing protocols for WSNs and their impact on the performance of the WSN networks. LEACH is a hierarchical-based wireless sensor network where the cluster head is picked randomly from the entire area and will deliver data to the base station immediately via single hop transmission. It is worth mentioning that LEACH does not fall into the chain, tree, grid, or area-based network categories even though it is classified as hierarchical-based wireless sensor network. Therefore, this survey compares and contrasts various parameters applied in cluster creation for several upgraded LEACH improvements in order to determine the aspects to consider while forming a cluster.

Various hierarchical routing algorithms for WSNs are examined, and their performance is discussed for further investigation. Finally, we'll wrap up this essay by highlighting the most pressing difficulties and suggestions for further studies in WSN.

The rest of this article is structured as follow: Section 2 highlights the most relevant routing protocol surveys. Section 3 focuses on providing background information on WSNs, while Section 4 presents a classification of the WSN routing protocols. Chain-based, tree-based, grid-based, and area-based hierarchical routing algorithms are examined and discussed in Sections 5, 6, 7, and 8, respectively. Finally, Sect. 9 brings this study to a close with a conclusion.

II. ROUTING PROTOCOLS

Routing protocols are categorized into data centric, hierarchical, location based and multipath based routing protocol [14, 18].

When creating any WSN protocol, numerous aspects are taken in consideration such as fault tolerance, energy efficiency, scalability, latency, power consumption, and network topology. Hierarchical routing algorithms TEEN and APTEEN's energy preservation technique is to select

an appropriate cluster head and manage the frequency of data transmission [18]. To save energy, the location-based protocol LAP maintains a table of neighbor nodes to choose the best transmission channel for forwarding data packets. The real time protocol RPAR satisfies a packet transfer deadline while maintaining the intended transmission rate.

In this survey, each algorithm is presented with its advantages and disadvantages to finish with the limitations of WSN. In the heterogeneous clustering algorithm m-BEENISH nodes have five distinct energy levels, and nodes with higher energy levels have a better probability of becoming cluster heads [12]. LEACH is a hierarchical routing algorithm that helps nodes save energy while communicating data. The latest LEACH versions and improvements are divided into two primary groups: (1) single hop transmissions and (2) multi-hop transmissions. Knowing the geolocation of sensor nodes is expensive since data extraction and communication use a lot of energy, when it comes to cluster head selection, energy consumption is also a crucial consideration [12]. There are some security concerns on LEACH because several sensor nodes are linked together to transmit sensitive and private data. Many ways are suggested for securing LEACH, including hop-by-hop, end-to-end data aggregation, and other security procedures [13]. By clustering protocols, we mean dividing a large network into groups in order to transmit data in a reasonable amount of time. Many applications may benefit from knowing where nodes are located such as environmental monitoring and mobile applications. To identify their location, a GPS feature can be added to sensor nodes. Combining tree-based network with duty cycles may reduce the delay time.

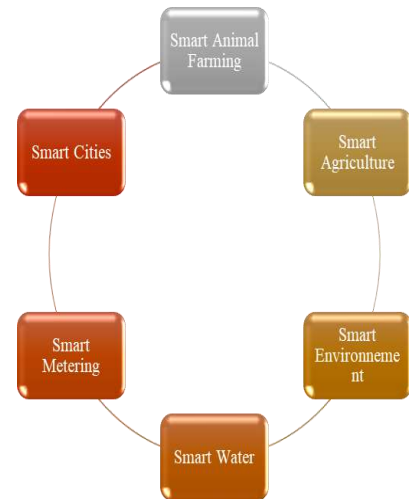
The k-neighborhood algorithm aims to decrease the number of neighborhood links. The neighborhood nodes organize their sleep schedules and routing paths among themselves to save energy. But collisions may happen when numerous nodes send data all at the same time and every node may have its own sleep schedule. That makes the simple broadcast method is not a practical solution. Furthermore, in high-density networks, redundant data will be delivered [15, 16].

In order to decide the multicast strategy, some analysts advocated taking into account the chance that a node rebroadcasts a packet during the current active period, as well as the possibility that the node does not go to the sleep mode and remains on after the active time. On the other hand, a flag may be added to the packet in order to store the quality of links to all of its neighbors and the status of broadcast packets to avoid redundant transmissions. Maintaining the Integrity of the Specifications

III. BACKGROUND OF WIRELESS SENSOR NETWORKS:

WSNs have smart sensors built in, allowing it to provide smart services and applications. They are frequently used because of their sensing, wireless communications, and processing aptitudes. WSN has smart sensors built in,

allowing it to provide smart services and applications. Figure 02 depicts the most significant examples.



. Figure 01. smart WSNs applications

WSNs are made up of a large number of low-cost wireless sensor nodes that operate on low power in areas where traditional networks are unable to compete. The difference between WSNs and other traditional data communication networks is that sensors are widely distributed, and nodes are frequently broken due to extreme weather conditions. The topology of some deployments may shift over time, necessitating the reconfiguration of links between nodes, which may cause some instability and use more energy. That is why WSNs may be unreliable in the field. As a result, maintaining stable WSNs is a tricky task that necessitates sophisticated monitoring and control mechanisms tailored to specific deployment.

Since sensor nodes are deployed in hostile environments, it is impossible to intervene with physical maintenance to recharge or replace sensor batteries. That is why, energy use is the most difficult constraint in WSNs and energy-efficient routing algorithms must be applied. The number of dead nodes can be an indicator of the WSN management. Building an efficient routing protocol that manages well the remaining energy of sensor nodes is a big challenge [19–20]. Furthermore, when batteries are infrequently installed in distant settings, it is often difficult to repair or recharge them. Every node requires electricity to operate its sensor transducer, communicate with other sensor nodes, and perform CPU calculations. Energy is required to sense and process data but also to provide data communication between nodes. And that consumes the majority of the energy. To construct a reliable data transmission network between a Utility Company and its consumers, for example, a significant number of sensor nodes is required. And to maintain the WSN reliability, the most economical use of battery power is required.

A. Robust wireless sensor networks criteria

There are various parameters to consider when determining a reliable WSN, and the importance of each parameter varies depending on the application [21].

Efficient power usage: Sensors get their energy primarily from batteries, which we believe is difficult to charge or refresh due to the large number of sensor nodes in somewhat hostile and hazardous environments. Also, since the sensor nodes are deployed randomly, it is difficult to locate them. Designing energy-efficient algorithms reduces the energy-efficient routing algorithms [19–20].

Scalability: the routing algorithm should be scalable for various network sizes when it is designed [22, 23].

Reliability: Because a dead sensor node cannot communicate any data, reliability is also associated to routing and energy usage. Specially if the dead node is a cluster head. The whole cluster will be less performant. Congestion has an impact on reliability, hence congestion control measures in the routing algorithm are practically always required, depending on the application.

Self-organization: if a sensor node collapses, or the network topology changes, sensors should be able to reorganize themselves. In such dynamic scenarios, routing protocols which are adaptative and able of following real-time topology changes must be used [24, 25].

Adaptability: frequent changes in cluster members may occur, such as sensor nodes joining or exiting a cluster. That is why it is important to design a flexible routing algorithm for sensor networks to accommodate these changes [26].

Security: Personal and private data can be delivered via a sensor network. As a result, for data transfer, a secure data communication network is required to prevent data from being transferred, deleted, or altered along the way. Routing protocols should not operate without considering security. [27, 28].

B. Wireless sensor networks constraints

It's always difficult to find the right balance between ideal criteria for a resilient WSN, especially when individual WSN applications impose extra constraints. [29]. The most important constraints are shown in Fig. 3 and explained further down:

- Limited and unstable energy supply. : Sensor nodes are supplied with batteries as a source of power. That makes the number of dead nodes in a network an indicator of the WSN's energy management. Networks can be employed in situations when changing node batteries is challenging. That is why designing an energy efficient routing algorithm for WSNs that takes in consideration the amount of remaining energy in the entire network is a challenging issue [19–20].
- Massive, random, and varying node deployment : sensor node deployment might be static or random. Supposing that the sensor node distribution changes over time, adaptable routing algorithms must be up to date in order to manage the network in an energy-efficient manner [30].
- Unreliable network environment sensor nodes may be susceptible to collapse due to physical damage or lack of energy. This has an impact on the WSN's performance. Reliable routing algorithms should

give the possibility to sensor nodes to reorganize themselves around a dead or unreliable node [31].

- Scalability : the optimal routing algorithm should adapt to every network scale and changings in physical communication channels between sensor nodes.

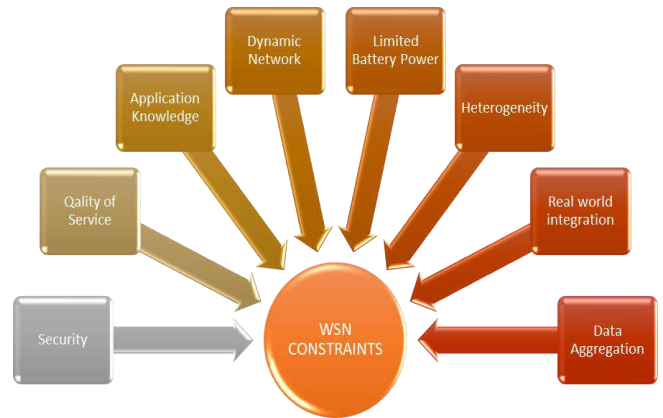


Figure 02. WSNs constraints

IV. CLASSIFICATION OF WIRELESS SENSOR NETWORKS

The routing protocols of the WSN can be classified into three basic structures [14, 17]:

- Flat routing algorithms : Data collecting, functions, transmission, and power consumption are all same among sensor nodes.
- Hierarchical routing algorithms : There are numerous clusters of sensor nodes. Using various well-known criteria to identify the node with the highest energy level in each cluster, it is usually chosen as the cluster head.
- Location-based routing algorithms : sensor nodes must be able to self-localize or calculate their location to be able to send data toward specific regions.

IoT sensor networks most likely choose hierarchical routing protocols. That is why in this survey multiple hierarchical routing protocols are analyzed in details (see Fig. 4).

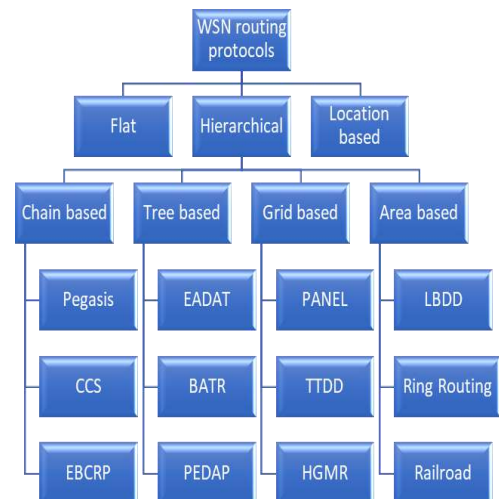


Figure 03. WSNs routing protocols classification

The following main categories can be used to categorize hierarchical routing: (1) chain-based, (2) tree-based, (3) grid-based, and (4) area-based routing.

Table 2 compares hierarchical routing algorithms based on energy consumption. The performance of hierarchical routing algorithms in terms of energy consumption is compared in Table 3

V. HIERARCHICAL CHAIN-BASED ROUTING ALGORITHMS

In chain-based hierarchical routing, the entire WSN is divided into many chains, with a leader node chosen for each chain. Until the packets reach the leader node, each sensor node will transfer them to the next closest node. Data aggregation is done through transmission. Then the leader node sends the aggregated data to the base station. Nodes always send data to the closest node and the chain's formation does not change frequently which makes the chain-based routing simple to run. So, the formation of a chain does not consume a lot of energy. However, in the case where there might be several nodes in a chain, and the source is far from the leader node, the distance that the data takes to reach the leader node will be so long, and that causes time delays. Time-sensitive applications will not be efficient in these cases. Furthermore, data transmissions will always involve nodes that are quite near to the leader node. This puts a lot of strain on these nodes, and their energy consumption rises as a result. The following are the most important chain-based routing algorithms

- Power-efficient gathering in sensor information systems (PEGASIS) [29] every node in the network is aware of the location of other nodes location. The farthest node from the sink will send data to the closest node until it reaches the leader. Leaders then will send the aggregated data to the sink and rotate. Since the leaders can be switched and any node in the network can participate in data transmission, the burden is shared and the load is balanced to some extent. However, data is passed through many intermediate nodes which takes long latency delays. That makes it not suitable for time-critical applications or for large-scale networks.
- Concentric clustering scheme (CCS) [30] In contrary to PEGASIS, which only contains one chain, CCS is constructed with numerous chains. CCS is separated into layers due to its many chains. In each chain, one cluster head will be chosen. Until data reaches the cluster head, all nodes in each chain will send data to the nearest node. The data will then be sent to another cluster head on the next higher layer until it reaches the base station from the cluster head on the farthest chain. Because data is sent among cluster heads up to the base station, the distance traveled is substantially reduced when compared to PEGASIS. As a result, the time delay is reduced, making it more suited to large-scale networks. on the other hand, the cluster head on the chain closest to the base station will have a higher traffic load and will likely be consumed sooner than others. It's worth mentioning that residual energy isn't taken into account when choosing the cluster head. As a result, a low

residual energy node may be chosen, but it will be quickly drained.

- Energy-balanced chain-cluster routing protocol (EBCRP) [31] The entire network is divided into several rectangular portions, and a ladder algorithm is used to create one chain in each rectangle area. Cluster heads are chosen based on residual energy and rotated to distribute the load. Data is collected by the cluster head of each chain from all nodes in its chain and sends it immediately to the base station. The ladder algorithm used by EBCRP is more energy efficient than the greedy algorithm since the total transmission distance may be shorter. The weight and burden of traffic is shared. This prevents any single cluster head node from becoming exhausted. The entire network, however, is separated into rectangles. Within a rectangle, neighboring nodes may not be on the shortest distance path and hence waste more energy. Furthermore, EBCRP is not ideal for large-scale networks due to single-hop transmission.

TABLE I. COMPARISON OF HIERARCHICAL CHAIN-BASED ROUTING PROTOCOLS

Algorithm	Structure	Energy consumption	Taking into account remaining energy for electing a leader	Location information of elected leader
PEGASIS [29]	chain based	High	No	No
CCS [30]	chain based	High	No	Yes
EBCRP [31]	chain based	High	Yes	No

VI. HIERARCHICAL TREE-BASED ROUTING ALGORITHMS:

For tree-based routing, the nodes are split into several branches, leaf nodes, and parent nodes. The data is sent from the leaf node to its parent node, then to the next parent node, and so on, until it reaches the base station. Because the data has been aggregated, some data replications can be eliminated. The tree topology is simple to create; each node simply needs to relay data to the next higher level node that is closer to the base station, with no need for cluster formation. This may lower the amount of energy required for tree routing. The disadvantage of tree formations is that if a tree's parent node fails, all data transmission beneath its branch will be lost. Furthermore, because of the increased data flow, the parent nodes that are extremely close to the base station and have numerous branches attached will consume a lot of energy. Furthermore, if the branch contains a large number of nodes, data transmission may be delayed, and energy consumption may increase. The following are the most important tree-based routing algorithms:

A. Energy-aware data aggregation tree (EADAT) [32]

The sink is where the tree begins to grow. To begin transmission, each node will set a timer for itself, and the time spent waiting is proportional to its remaining energy. The shorter the waiting period is, the bigger the remaining energy is. The node will then choose the node with the most remaining energy and the closest distance as its

parent. When a parent node's remaining energy falls below a certain level, it broadcasts a message to its child. The child can then choose a different parent to transmit to. In EADAT, residual energy is taken into account while choosing a connection node and path. As a result, the likelihood of selecting an exhausted node is lowered, and failure transmission is avoided. Furthermore, by prioritizing the node with the most energy, some load balancing can be achieved, extending the network's life. However, while residual energy is one of the parameters used to choose the way, the final path may not be the shortest, consuming more energy overall and involving many more nodes in data transmissions. As a result, total energy consumption may rise, and time delays may lengthen

B. Balanced aggregation tree routing (BATR) [33]

The base station gathers all node's global position data and generates the routing paths. BATR will build a minimal spanning tree based on energy consumption and determine the number of child nodes beneath the tree to balance the load. Because BATR considers energy usage when constructing the routing path, it can extend the network's lifetime. Furthermore, by evenly dispersing the child nodes among several trees, the loading can be balanced. When making a tree, however, the remaining energy of each node is ignored. As a result, some nodes with low residual energy will be consumed sooner, resulting in transmission failure

C. Power-efficient data gathering and aggregation protocol (PEDAP) [34]

The least spanning tree is also used to calculate energy usage in this procedure. The tree is constructed using data volume and transmission distance. In addition, the remaining energy of nodes will be taken into account during data transmissions. PEDAP can balance the energy consumptions and reduce the delay time after evaluating the distance for transmissions. However, the tree's development could be complicated, and the energy required to calculate the path could be significant. As a result, the setup energy, particularly for large-scale networks, may be rather substantial.

TABLE II. COMPARISON OF HIERARCHICAL TREE BASED ROUTING PROTOCOLS

ABBREVIATIONS:

LIEL= Location information of elected leader

TIAREEL= TAKING INTO ACCOUNT REMAINING ENERGY FOR ELECTING A LEADER

EC = ENERGY CONSUMPTION

Algorithm	Structure	EC	TIAREEL	Organisation	LIEL
EADAT [32]	tree based	Low	Yes	Centralised	No
BATR [33]	tree based	Medium	No	Centralised	No
PEDAP [34]	tree based	Low	Yes	Centralised	Yes

VII. HIERARCHICAL GRID-BASED ROUTING ALGORITHMS

The entire network region is separated into several grids, each with its own leader. All nodes in a grid will send data to their leader, who will then forward it to the

leader of the next grid until it gets the base station. Grid-based routing is easy to set up because it is based on the geographical position of the nodes. Because the grid size is predetermined and only the location information of the grid leader is necessary, it is claimed that the data can be delivered more efficiently. However, if a grid has a large number of nodes, there may be a lot of traffic and the leader node's energy will be consumed quickly. The following are the most important grid-based routing algorithms:

A. Position-based aggregator node election protocol (PANEL) [35]

The aggregator is chosen based on the location information. The entire network is segmented geographically, with the aggregator being the node closest to the reference point. The aggregator will collect data from its cluster members before sending it to the base station. Because every node has the potential to become an aggregator, load balancing can be achieved by sharing the responsibility of communicating with the base station. It's worth mentioning that gathering the location of sensor nodes will consume some energy, which will raise the setup cost.

B. Two-tier data dissemination (TTDD) [36]

The grid is divided into cells, with certain nodes relaying data from the mobile sinks to the source. The mobile sink will send a flood of data requests to the intermediate nodes. Until the data reaches the network's boundary, the source uses a greedy algorithm to select the next relay nodes. The mobile sinks will travel across the grids, extracting data from the source's closest node. For on-demand applications, TTDD is ideal. However, because flooding consumes a considerable amount of power, it is not ideal for large-scale or high-traffic networks. Furthermore, mobile sinks may not move at the same rate as the route generated. There may be a time delay, and retransmission may be necessary, consuming additional energy.

C. Hierarchical geographic multicast routing (HGMR) [37]

The entire network is separated into various cells based on their geographic location in this case. There is an access point for managing location information within each cell. The network is organized in a hierarchical manner. The data will be delivered from the highest level access points to the lowest level access points via the source. Because the accessing points for HGMR can be rotated, different nodes will play corresponding roles in data transmissions. This helps to keep energy consumption in check. The network is appropriate for big scale networks since it is separated into several cells and layers. The transmission from higher level access points to lower level access points, on the other hand, ignores the issue of location. As a result, the path may not be the shortest, resulting in a time delay and the consumption of additional energy.

TABLE III. COMPARISON OF ENERGY CONSUMPTION OF HIERARCHICAL GRID BASED ROUTING PROTOCOLS

Algorithm	Structure	Energy consumption	Taking into account remaining energy for electing a leader
PANEL [35]	grid based	Low	No
TTDD [36]	grid based	High	No
HGMR [37]	grid based	Medium	No

VIII. HIERARCHICAL AREA-BASED ROUTING ALGORITHMS

The entire sensor network is divided into numerous areas, each of which can be customized in size. To gather data, the base station or sink will issue a data request to the closest nodes in the area. The data request will be flooded until the source of the data is identified. The data will subsequently be sent from the source node to the sink. This is appropriate for mobile applications where the mobile sink is constantly moving within a defined area. The following are the most important area-based routing algorithms:

A. Line-based data dissemination (LBDD) [38]

This is an example of a standard area-based routing protocol. A vertical line of nodes divides the entire network into two equal areas. The data will be stored in the nodes on the vertical line in order to serve requests from the sinks. All nodes will be aware of each other's location. The data will be sent from the source node to the line's closest node. A sink will send a perpendicular data request to the line. The receiving node will process the request and proceed to relay it to other nodes on the lines in both directions. Finally, the data-storage node gets the request and transmits the data to the sink directly. The structure's setup and communication are straightforward. If the number of nodes on the connection is minimal, however, they may face a significant strain. These nodes will quickly be depleted. Furthermore, because a flooding technique is applied for data requests, if more nodes are assigned to the line, the energy consumption of all nodes involved on the line would increase. As a result, this protocol isn't appropriate for large-scale networks.

B. Ring routing [39]

A ring topology is proposed. The procedure is similar to that of LBDD. Instead of a line, a ring is formed. The ring's relay nodes can be swapped out for conventional nodes. The ring structure is easy to build. Because the relay nodes on the ring are rotated to protect any one node from overload, it enhances load balancing among nodes. The source node may locate the closest relay node in a shorter distance using the ring topology, which reduces data transmission time and energy consumption. However, if the network is extensive, the ring structure setup costs could be significant, as data requests are sent to all of the ring's associated nodes, causing an overhead.

C. Railroad [40]

Railroad, a data dissemination architecture for large-scale WSNs, was described. One rail divides the network and organizes the data requests. The train is in the heart of the network, allowing for convenient

access to all nodes. The data request will be forwarded to the rail until it reaches the source node, which will then deliver data to the sink immediately. Unicasts, rather than flooding, are used to send data requests from the sink. A rail structure is more flexible than a line or ring construction, and the rail's relay nodes may be easily reached by normal nodes, reducing the distance and time it takes for the source node to deliver data to the railway's relay nodes. However, because the rail is often slightly longer, data requests transmitted along it may take longer and cause delays. In large-scale networks, these delays become more pronounced.

TABLE IV. COMPARISON OF ENERGY CONSUMPTION OF HIERARCHICAL AREA BASED ROUTING PROTOCOLS

Algorithm	Structure	Energy consumption	Taking into account remaining energy for electing a leader
LBDD [38]	area based	Low	No
Ring routing [39]	area based	Low	No
Railroad [40]	area based	Low	No

IX. CONCLUSION

Above, we divided hierarchical routing techniques into four categories: (1) chain-based, (2) tree-based, (3) grid-based, and (4) area-based routing. We went over the main features and characteristics of each type of hierarchical routing algorithms in great depth.

We've come to the end of our survey. Routing strategies in WSNs is a well-studied topic with a plethora of study findings. We offered a detailed review of hierarchical routing approaches in WSNs that have been discussed in the literature in this survey. The goal of all of these strategies is to increase the lifespan of the WSN without sacrificing data transmission performance. Finally, depending on the routing approaches, we categorised hierarchical routing techniques.

REFERENCES

- [1] Kumar, A., Ovsthus, K., & Kristensen, L. (2014). An industrial perspective on wireless sensor networks: A survey of requirements, protocols, and challenges. *IEEE Communications Surveys Tutorials*, 16(3), 1391–1412.
- [2] Aldeer, M. M. N. (2013). A summary survey on recent applications of wireless sensor networks. In *IEEE Student Conference on Research and Development* (pp. 485–490).
- [3] Gungor, V. C., Lu, B., & Hancke, G. P. (2010). Opportunities and challenges of wireless sensor networks in smart grid. *IEEE Transactions on Industrial Electronics*, 57(10), 3557–3564.
- [4] Erol-Kantarci, M., & Mouftah, H. T. (2011). Wireless sensor networks for cost-efficient residential energy management in the smart grid. *IEEE Transactions on Smart Grid*, 2(2), 314–325.
- [5] Prathap, U., Shenoy, P., Venugopal, K., & Patnaik, L. (2012). Wireless sensor networks applications and routing protocols: Survey and research challenges. In *IEEE Symposium on Cloud and Services Computing* (pp. 49–56).

- [6] Gkikopouli, A., Nikolakopoulos, G., & Manesis, S. (2012). A survey on underwater wireless sensor networks and applications. In IEEE Conference on Control Automation (MED) (pp. 1147–1154).
- [7] Blanckenstein, J., Klaue, J., & Karl, H. (2015). A survey of lowpower transceivers and their applications. IEEE Circuits and Systems Magazine, 15(3), 6–17. thirdquarter.
- [8] Miorandi, D., Lowe, D., & Gomez, K. M. (2010). Activationinhibition- based data highways for wireless sensor networks. In E. Altman, I. Carrera, R. El-Azouzi, E. Hart, & Y. Hayel (Eds.), Bioinspired models of network. Information, and computing systems (pp. 95–102). Berlin, Heidelberg: Springer.
- [9] Chang, J.-H., & Tassioulas, L. (2004). Maximum lifetime routing in wireless sensor networks. IEEE/ACM Transactions on Networking, 12(4), 609–619.
- [10] Daanoune, I., Abdennaceur, B., & Ballouk, A. (2021). A comprehensive survey on LEACH-based clustering routing protocols in Wireless Sensor Networks. Ad Hoc Networks, 114, 102409 (540–545).
- [11] Chatap, A., & Sirsakar, S. (2017). Review on various routing protocols for heterogeneous wireless sensor network. In Conference on IoT in Social, Mobile, Analytics and Cloud (pp. 440–444).
- [12] Kumari, J., & Prachi. (2015). A comprehensive survey of routing protocols in wireless sensor networks. In IEEE Conference on Computing for Sustainable Global Development (pp. 325–330).
- [13] Singh, S., Kumar, P., & Singh, J. (2017). A survey on successors of LEACH protocol. IEEE Access, 5, 4298–4328.
- [14] Hao, J., Zhang, B., & Mouftah, H. T. (2012). Routing protocols for duty cycled wireless sensor networks: A survey. IEEE Communications Magazine, 50(12), 116–123.
- [15] Yamunadevi, S. P., Vairam, T., Kalaiarasan, C., & Vidya, G. (2012). Efficient comparison of multipath routing protocols in WSN. In IEEE
- [16] Conference on Computing, Electronics and Electrical Technologies (pp. 807–811).
- [17] Goyal, D., & Tripathy, M. (2012). Routing protocols in wireless sensor networks: A survey. In IEEE Conference on Advanced Computing Communication Technologies (pp. 474–480).
- [18] Baghyalakshmi, D., Ebenezer, J., & Satyamurthy, S. (2010). Low latency and energy efficient routing protocols for wireless sensor networks. In IEEE Conference on Wireless Communication and Sensor Computing (pp. 1–6).
- [19] Sinha, A., & Chandrakasan, A. (2001). Dynamic power management in wireless sensor networks. IEEE Design Test of Computers, 18(2), 62–74.
- [20] Saghar, K., Henderson, W., Kendall, D., & Bouridane, A. (2010). Formal modelling of a robust wireless sensor network routing protocol. In IEEE Conference on Adaptive Hardware and Systems (pp. 281–288).
- [21] Alazzawi, L. K., Elkateeb, A. M., Ramesh, A., & Aljuhar, W. (2008). Scalability analysis for wireless sensor networks routing protocols. In IEEE Advanced Information Networking and Applications Workshops (pp. 139–144).
- [22] Ye, W., Heidemann, J., & Estrin, D. (2004). Medium access control with coordinated adaptive sleeping for wireless sensor networks. IEEE/ACM Transactions on Networking, 12(3), 493–506.
- [23] Horjaturapittaporn, T., & Suntiamorntut, W. (2011). Scalable routing protocol in wireless sensor networks. In IEEE Conference on Communication Software and Networks (pp. 623–627).
- [24] Sohrabi, K., Gao, J., Ailawadhi, V., & Pottie, G. (2000). Protocols for self-organization of a wireless sensor network. IEEE Personal Communications, 7(5), 16–27.
- [25] Sirsakar, S., Chunawale, A., & Chandak, M. (2014). Self-organization architecture and model for wireless sensor networks. In IEEE International Conference on Electronic Systems, Signal Processing and Computing Technologies (pp. 204–208).
- [26] Spadoni, I. M. B., Araujo, R. B., & Marconde, C. (2009). Improving QoS in wireless sensor networks through adaptable mobile agents. In IEEE INFOCOM Workshops 2009 (pp. 1–2).
- [27] Patel, M., & Aggarwal, A. (2013) Security attacks in wireless sensor networks: A survey. In IEEE Conference on Intelligent Systems and Signal Processing (pp. 329–333).
- [28] Gaware, A., & Dhonde, S. (2016). A survey on security attacks in wireless sensor networks. In IEEE Conference on Computing for Sustainable Global Development (pp. 536–539).
- [29] Angurala, M., & Bharti. (2016). A comparative study between leach and pegasis—A review. In IEEE Conference on Computing for Sustainable Global Development (pp. 3271–3274).
- [30] Jung, S. M., Han, Y. J., & Chung, T. M. (2007). The concentric clustering scheme for efficient energy consumption in the PEGASIS. IEEE Conference on Advanced Communication Technology, 1, 260–265.
- [31] Xi-rong, B., Shi, Z., Ding-yu, X., & Zhi-tao, Q. (2010). An energy-balanced chain-cluster routing protocol for wireless sensor networks. IEEE Conference on Networks, Security Wireless Communications and Trusted Computing, 2, 79–84.
- [32] Ding, M., Cheng, X., & Xue, G. (2003). Aggregation tree construction in sensor networks. IEEE Vehicular Technology Conference, 4, 2168–2172. Vol.4.
- [33] Kim, H. S., & Han, K. J. (2005). A power efficient routing protocol based on balanced tree in wireless sensor networks. In IEEE Conference on Distributed Frameworks for Multimedia Applications (pp. 138–143).
- [34] Tan, H. O., & Ko'pceog'lu, I. (2003). Power efficient data gathering and aggregation in wireless sensor networks. ACM SIGMOD Record, 32(4), 66–71.
- [35] Buttyan, L., & Schaffer, P. (2007). PANEL: Position-based aggregator node election in wireless sensor networks. In IEEE Conference on Mobile Adhoc and Sensor Systems (pp. 1–9).
- [36] Luo, H., Ye, F., Cheng, J., Lu, S., & Zhang, L. (2005). TTDD: Two-tier data dissemination in large-scale wireless sensor networks. Wireless Networks, 11(1), 161–175.
- [37] Koutsonikolas, D., Das, S., Hu, Y. C., Stojmenovic, I. (2007). Hierarchical geographic multicast routing for wireless sensor networks. In Conference on Sensor Technologies and Applications (pp. 347–354).
- [38] Hamida, E., & Chelius, G. (2008) A line-based data dissemination protocol for wireless sensor networks with mobile sink. In IEEE Conference on Communications (pp. 2201–2205).
- [39] Bhatti, R., Kaur, G. (2017). Virtual grid based energy efficient mobile sink routing algorithm for WSN. In IEEE Conference on Intelligent Systems and Control (ISCO) (pp. 30–33).
- [40] Shin, J.-H., Kim, J., Park, K., & Park, D. (2005). Railroad: Virtual infrastructure for data dissemination in wireless sensor networks. In ACM Workshop on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks, New York, NY (pp. 168–174).

Human Facial Expression Recognition by Uniform Local Binary Patterns Approach & Wavelet Transform

Zermi.Narima⁽¹⁾, Khaldi Amine⁽²⁾, Hamami Rania⁽¹⁾

⁽¹⁾Laboratory of Automatic and Signals Annaba (LASA) , Department of electronics, Faculty of Engineering, Badji-Mokhtar University, P.O.Box 12, Annaba-23000, Algeria.

⁽²⁾ Computer Science Department, Faculty of Sciences and Technology, Artificial Intelligence and Information Technology Laboratory (LINATI) University of Kasdi Merbah, 30000, Ouargla, Algeria

E-Mail: naili.narima@gmail.com, Khaldi.Amine@univ-ouargla.dz

Abstract— facial expression analysis is rapidly becoming an area of intense interest in computer science and human computer interaction design communities. Psychological studies have suggested that facial motion is fundamental to the recognition of facial expression.

Expression is the most important mode of non-verbal communication between people. Recently, the facial expression recognition technology attracts more and more attention with people's growing interest in expression information. In this paper, we propose LBP histograms based automatic facial expression recognition system to recognize the human facial expression like happy, fear, sad, angry, disgust and surprise. Initially facial image is segmented into three region from which the

uniform local binary patterns (LBP) texture features distributions are extracted and represented as a histogram description. A Support Vector Machine is used to classify different kinds of facial expressions. We have carried our experiments upon Yale face database and JAFFE face database. The Yale Face Database contains 165 grayscale images in GIF format of 15 individuals. JAFFE Database, available at <http://www.kasrl.org/jaffe.html> consisting 213 images posed by 10 Japanese female models. The proposed model reports 95.77% of classification accuracy.

Index Terms- Japanese Female Facial Expression (JAFFE) database; facial, expression; recognition; Matlab; Local Binary Pattern, feature extraction, Yale face

I-INTRODUCTION:

Facial expression analysis plays a significant role for human computer interaction. Automatic analysis of human facial expression is still a challenging problem with many applications. Face identification and recognition has lead to the development of different algorithms for various applications such as automated access control, surveillance, image retrieval...etc [5]. Facial expression analysis has wide range of application in areas such as Human Computer Interaction (HCI), psychological area, image understanding, face animation etc. Humans interact with each other both verbally and non-verbally [6].

Most approaches to automatic facial expression analysis attempt to recognize a principal set of prototypic emotional facial expression .i-e- fear, sadness, disgust, anger, surprise and happiness, from the survey, it was revealed that most of the facial expression systems were based on the Facial Action Coding System (FACS), [3], [10].

It is a system designed for human observers to describe changes in the facial expression in term of visually observable activations of facial muscles. Facial Expression Recognition should not be confused with human emotion recognition as is often done in computer vision. Facial expression recognition deals with classification of facial motion and facial feature deformation in the abstract classes that are purely based on visual information. But emotions are result of many different factors such as emotional voice, pose, gestures, facial expressions etc.

Nowadays, Facial Expression Recognition is getting more and more attention from the research community. Facial

expression plays an important role in smooth communication among individuals. The extraction and recognition of facial expression has been the topic of various researches subject to enable smooth interaction between computer and their users. In this way, computers in the future will be able to offer advice in response to the mood of the users [13]. Recently, a number of new technologies for Facial Expression Recognition have been developed. Our goal is to create a system, which can be trained from training images for the facial expressions and used to recognize the facial expression in the test images.

The remainder of the paper is organized as follows: next section describes wavelet transformation and the filters. Section III describes Methodology. Faces classifier using SVM is given in section IV. Results is describes in section V.

II- WAVELET TRANSFORM

The wavelet transform is similar to the Fourier transform (or much more to the windowed Fourier transform) with a completely different basis functions. The main difference is that Fourier transform decomposes the signal into sinus and cosines, i.e. the functions localized in Fourier space; while the wavelet transform uses functions that are localized in both the real and Fourier space. Generally, the wavelet transform can be expressed by the following equation:

$$F(a, b) = \int_{-\infty}^{\infty} f(x) \psi_{(a,b)}^*(x) dx \quad (1)$$

where the $*$ is the complex conjugate symbol and function ψ is some function. This function can be chosen arbitrarily provided that obeys certain rules.

As it is seen, the Wavelet transform is in fact an infinite set of various transforms, depending on the basis functions used for its computation. This is the main reason, why the term “wavelet transform” is used in very different situations and applications. There are also many ways how to sort the types of the wavelet transforms. Here we show only the division based on the wavelet orthogonality. We can use *orthogonal wavelets* for discrete wavelet transform development and *non-orthogonal wavelets* for continuous wavelet transform development. These two transforms have the following properties [1]:

1. The discrete wavelet transform returns a data vector of the same length as the input. Usually, even in this vector many data are almost zero. This corresponds to the fact that it was decomposed on a set of wavelets that are orthogonal to its translations and scaling. Therefore, we decompose such a signal to a same or lower number of wavelet coefficients spectrum. Such wavelet spectrum is suited for signal processing and compression because it eliminates information redundancy.
2. The continuous wavelet transform, in contrary, returns one dimensional array larger than the input data. For a 1D data, we obtain an image of the time-frequency plane. We can easily see the signal frequencies evolution during the duration of the signal and compare the spectrum with other signals spectra. As it uses non-orthogonal set of wavelets, data are highly correlated and redundant.

In mathematics, a wavelet series is a representation of a square-integrable (real-or complex-valued) function by an orthonormal series generated by a wavelet. Nowadays, wavelet transformation is one of the most popular and commonly used for time-frequency transformations.

Basic idea: The fundamental idea of wavelet transforms is that the transformation should allow only changes in time extension, but not shape. Based on the uncertainty principle of signal processing,

$$\Delta t \Delta \omega \geq \frac{1}{2} \quad (2)$$

Wavelet transform is a popular tool in image processing and computer vision, because of its ability to capture localized time-frequency information of image extraction. The decomposition of the data into different frequency ranges allows us to isolate the frequency components introduced by intrinsic deformations due to expression or extrinsic factors (like illumination) into certain subbands. Wavelet-based methods prune away these

variable subbands, and focus on the subbands that contain the most relevant information to better represent the data.

1-D Continuous Wavelet transform (CWT) can be defined as follows:

$$WT(s, \tau) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} f(t) \psi_{s,t}(t) \quad (3)$$

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}} \left(\frac{t-\tau}{s} \right) dt \quad (4)$$

is basis function

$\psi(t)$ is called mother wavelet. The parameters S and τ , are the scaling and shift parameter, respectively. Discrete Wavelet transform (DWT) is a sampled version CWT. The 2D-DWT, for $m \times n$ images, is defined as follows:

$$DWT(j, k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} f(x) \psi \left(\frac{x}{2} - k \right) dx \quad (5)$$

Where j is the power of binary scaling and k is a constant of the filter. The 2D-DWT is computed by successive low pass-band and high pass-band filtering of the image. Applying 2D-DWT on an image will decompose it in four sub-bands LL, LH, HL and HH corresponding to the approximation of horizontal, vertical and diagonal features respectively. The sub-band denoted by LL is approximately at half the original image. While the sub-bands HL and LH contain the changes of images or edges along vertical and horizontal directions, respectively. The sub-band HH contains the details in the high frequency of the image.

Features based on Gabor filters have been used in image processing due to their powerful properties. Gabor Kernels are characterized as localized, orientation selective, and frequency selective. A family of Gabor kernel is the product of a Gaussian envelope and a plane wave. We have considered the face images with certain degree of orientation and large variations in the facial expressions. We used ‘K’ images with ‘N’ expressions for each face so that ‘KxN’ faces images are used. In the training stage, each N expression from K images is grouped together to form different expression groups which is used as the database.

Each image is convolved with a Gabor kernel of variance and bandwidth as π and a window size of 9×9 . In order to reduce the computational complexity we have used a reduced form of 60 Gabor filter bank. Since expression features are highly described by high frequency components, three frequencies, $v=(0,1,2,3)$ and four orientations $w=(5,6,7,8,9)$ are used. A phase-sensitive similarity function is applied to match the Gabor features of a facial pixels of test face and the images in the database.

III- METHODOLOGY:

The proposed methodology is based on the facial texture information of permanent facial feature components such as eye, eyebrow, nose and mouth to recognize the facial expression.

Initially frontal view image is segmented into three regions, upper region E (eye and eyebrow), middle region N

(nose) and lower region M (Mouth). The uniform rotation invariant LBP texture feature distributions are extracted from these three regions and represented as a weighted 7 dimensional histogram descriptor $h=[h_1, h_2, h_3, \dots, h_7]$.

A- Local Binary Patterns (LBP):

The basic local binary pattern operator, introduced by Ojala et al [13], was based on the assumption that texture has locally two complementary aspects, a pattern and its strength [6]. In that work, the LBP was proposed as a two-level version of the texture unit to describe the local textural patterns. The original version of the local binary pattern operator works in a 3×3 pixel block of an image. The pixels in this block are threshold by its center pixel value, multiplied by powers of two and then summed to obtain a label for the center pixel. As the neighborhood consists of 8 pixels, a total of $2^8 = 256$ different labels can be obtained depending on the relative gray values of the center and the pixels in the neighborhood. See Fig. 1 for an illustration of the basic LBP operator.

At a given pixel position (x_c, y_c) , LBP is defined as an ordered set of binary comparisons of pixel intensities between the centre pixel and its eight surrounding pixels. The decimal form of the resulting 8-bit word (LBP code) can be expressed as follows:

$$LBP(x_c, y_c) = \sum_{n=0}^7 s(t_n - t_0) 2^n \quad (6)$$

Where t_0 corresponds to the grey value of the centre pixel (x_c, y_c) , in to the grey values of the 8 surrounding pixels, and function $s(x)$ is defined as:

$$s(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x \geq 0 \end{cases} \quad (7)$$

To extract the feature of the facial expression, the images are divided into local regions R0, R1, Rm-1 and texture descriptors are extracted from each region independently. For every region a histogram with all possible labels is constructed. This shows that every bin in a histogram represents a pattern and contains the number of its appearance in the region.



(a) Original Image (b) LBP calculation result

FIG 1: Feature extraction using LBP (Local Binary Pattern) Histogram

combines geometric feature and texture feature. To determine the matching degree of textures used for LBP histogram. Feature of LBP algorithm is robust to lighting changes. It divides into $m \times n$ image of a small block for local feature

extraction and calculates LBP histogram in small block. All LBP histogram connections are used as feature descriptors for the entire image [Fig.2] is the result of applying LBP feature descriptor in the JAFFE database.

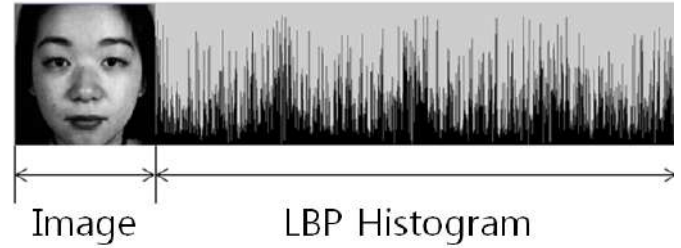


Fig2: SVM input values

We are used 213 images from the database; input matrix for classifier is of size 256×256 . This methodology divided into following steps:

- Extract average face from the original images of database;
- Divide average image in the three blocks such as upper (eyes), middle (nose) and lower (mouth).
- Compute the LBP code for each pixel in a block.
- Compute the histogram over the block.
- Normalized the histogram.
- Concatenate normalized histograms off all blocks.
- Use SVM to classifier these features into different classes such as anger, disgust, fear, happy neutral, sad and surprise.

IV. Faces classifier using SVM approach

Several classifiers have been proposed and used in facial expression recognition systems to handle the issues of non-linearity, dimensionality and generalization, and three main classification can be distinguished, namely, the HMM (hidden Markov Models), the SVM (Support Vector Machines) and the feedforward neural networks (FFNN).

Support Vector Machines (SVM) is a two-class classification method that finds the optimal decision hyper-plane based on the concept of structural risk minimization.

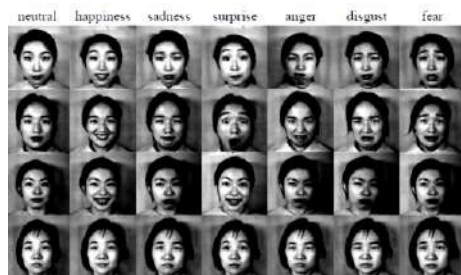
Support vector machines (SVM) are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. The basic SVM takes a set of input data and predicts, for each given input, which of two possible classes forms the output, making it a non-probabilistic binary linear classifier. As the dimensions of face image data is very large, although the satisfactory recognition rate can be obtained, but the processing time need so long as the large dimensions.

The expression classification is done by a SVM classifier. SVM are a set of related supervised learning methods used for classification and regression. It shows better classification accuracy than Neural Networks if the data set is small. SVM rearranges the data points according to a mathematical function or Kernels and transforms it into a feature space which allows

the classification. SVMs exhibit good classification accuracy even when only a modest amount of training data is available.

Based on the analysis above, we proposed a integrated method which is combined with 2D DWT, LBP and SVM approach for facial expression recognition. At first stage, the original images are decomposed into low frequency images by applying wavelets transform, the high-frequency components were ignored, while the low-frequency components which contains the primary information can be obtained. And then LBP histogram is used to deal with feature extraction. After generating feature vector, distance classifier and SVM are used for classification stage. We used “one-against-all” SVM multi-classification for recognizing face, and n SVM classifiers should be trained.

We have carried our experiments upon Yale face database and JAFFE face database. The Yale Face Database contains 165 grayscale images in GIF format of 15 individuals. For each of the subject, there are 11 images, one per different facial expression or configuration: center-light, w/glasses, happy, left-light, w/no glasses, normal, right-light, sad, sleepy, surprised, and wink. The JAFFE face database we use in our experiments contains 213 images of female facial expressions. The head is almost in frontal pose. Original images have been rescaled and cropped such that the eyes are roughly at the same position with a distance of 80 pixels in the final images (resolution: 256pixels*256pixels) the number of images corresponding to each of the 7 categories of expression (Neutral, Happiness, Sadness, Surprise, Anger, disgust and Fear) is roughly the same. A few of them are shown in fig3 [3], [4]



(a)



(b)

Fig.3: Examples of Facial Expression a) JAFFE database b) Yale face database

In this paper, a new facial expression recognition technique based on LBP analysis (LBP), on wavelet subband is proposed. We extract image features of facial images from various wavelet transforms (Haar, Daubechies, Coifet, Symlet, Biothogonal and reverse biothogonal) by decomposing face image in subband 1 to 8. These features are analysed by LBP and SVM.

Facial expression analysis requires multiclass SVM for classification. The one against-all approach constructs M binary SVM classifiers, each of which separates one class from the others.

Our program has been tested using JAFFE face & Yale face Database, consisting of 213 images posed by 10 Japanese female models. We select 130 samples for training; the remaining 100 samples are used as the test set. In our facial expression recognition experiments, we extract image feature of facial images from 49 wavelet transforms by decomposing face image in LL sub-bands at 8 levels (1 to 8). The wavelets used are Haar 1, DB 1 to DB 10, coif 1 to coif 5, sym 1 to sym 10; bio 1.1, 1.3, 1.5, 2.2, and RBIOL 1.5, 1.4, 2.2, 2.4, 2.6 and 2.8.

Table 1 and Fig.4 depict some sample results for db 5 sym10 and bio2.2 wavelets.

	DB5		SYM10		BIOR2.2	
Level	Rec.acc	time	Rec.acc	time	Rec.acc	time
1	95.00	7.00	95.50	7.99	95.00	6.85
2	95.50	5.50	96.00	6.80	95.50	5.30
3	95.50	5.30	96.00	6.50	96.50	5.20
4	91.00	5.90	96.50	6.90	92.50	5.40
5	90.50	6.25	89.00	7.20	87.00	6.00
6	86.00	6.50	81.00	7.65	82.00	6.02
7	83.50	7.41	79.00	8.35	78.00	6.40
8	75.00	7.23	76.50	8.70	74.50	6.80

TABLE 1: RECOGNITION ACCURACY IN PERCENTAGE AND TIME REQUIRED FOR THREE SAMPLE WAVELETS DB5, SYM 10, BIOR 2.2.

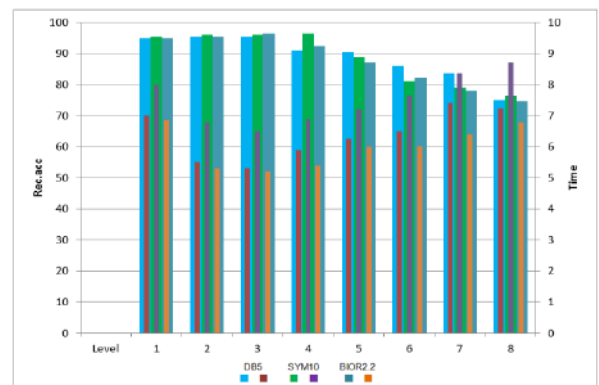


Fig 4: Graph Recognition accuracy in percentage and time required for three sample wavelets DB5, SYM 10, BIOR 2.2.

database	expressions	Average Recognition Rate(%)
Yale Face DataBase	Happy	80.70%
	normal	89.55%
	surprise	92.03%
	sad	88%
JAFFE face DataBase	Happy	85.70%
	normal	93.55%
	surprise	91.03%
	sad	89.55%

Table 2 Performance ratio for two database

VI-CONCLUSION

The facial expression recognition system has been used in communication to make the answer machine more interactive with people. The answer machine has become more intelligent by analysing the client's voice and dealing with the responses according to their emotions.

Local Binary Pattern (LBP) [1] is one of the most popular and efficient appearance-based feature descriptors [3, 4]. It has proven to be highly discriminative and its invariance to monotonic gray-level discrepancies makes it a robust feature descriptor in two-dimensional space. In general, an LBP operator measures each pixel of a given image by thresholding its neighborhood with the value of the center pixel and forms the results into a binary pattern. Then, the occurrence histogram based on the resulted binary patterns can be computed over an image or a region of the image, which is proven to be a powerful feature descriptor. Now, due to its discriminative power and computational simplicity, LBP is broadly utilized in image pattern recognition. Moreover, it receives tremendous success in facial expression recognition as it is insensitive to illumination variations and well describes subtle appearance details of the local features on human faces [5].

As LBP is proven to be of high performance as well as low computational cost, it is frequently applied in pattern recognition on static images. Following the steps of LBP, Zhao *et al.* [6] started to explore appearance-based feature descriptors on dynamic or temporal information, which combines appearance and motion.

The proposed method of facial recognition is innovative in this area with combination of LBP + 2D wavelet transform and SVM for identification of seven different facial expressions sad, disgust, anger, happy, surprise, fear and neutral with reduced features for storage of JAFFE & YALE face database work for both person dependant and independent mode.

We recorded and compared the recognition performance at various levels of the wavelet basis functions. Obtained accuracy with only LBP was about 89% whereas the recorded average

accuracy using LBP with wavelet in sub-bands 2 and 3 was up to 95.77%.

REFERENCE

- [1] Amberg B., Knothe R., Vetter T., "Expression Invariant 3D Face Recognition with a Morphable Model", IEEE International Conference on Automatic Face & Gesture Recognition, pp 1-6, 2008.
- [2] Bouzned kh, N.Zermi, "Use of NLPCA for Sensors Fault Detection and Localization Applied at WTP", Journal of E -Technology Volume 4 Number 3 August 2013.
- [3] Gundimada S., Asari V., "Facial Recognition Using Multisensor Images Based on Localized Kernel Eigen Spaces" IEEE Trans on Image Processing, pp 1314-1325, Vol. 18, No. 6, June 2009.
- [4] L. Chen, L. Zhang, Y. Hu, M. Li, H. Zhang, "Head pose estimation using fisher manifold learning, in:" Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures, in Conjunction with ICCV2003, 2003.
- [5] Pantic, M. & Rothkrantz, J. M. (2000) Automatic analysis of facial expressions: The state of the art. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12), 1424-1445.
- [6] Ma L., "Facial Expression Recognition Using 2-D DCT of Binarized edge Images and Constructive Feedforward Neural Networks", IEEE International Joint Conference on Neural Networks IJCNN (IEEE World Congress on Computational Intelligence), pp4083-4088, 2008 IEEE.
- [7] Moody, J., and Darken, J. (1989). Fast learning in networks of locally tuned processing units. Neural Computation, 1, 281-294.
- [8] Seyed Mehdi Lajevardi, "Facial Expression Recognition in Perceptual Color Space", Digital Object Identifier 10.1109/TSMCB.2012.2191773 IEEE 2012.
- [9] Yue ZENG, Dazheng FENG, Li XIONG "An Algorithm of Face Recognition Based on the Variation of 2DPCA" Journal of Computational Information Systems 7:1 (2011)303-310, 2011 Binary Information Press January 2011.
- [10] M.S.Aswathy R."A literature review on facial expression recognition technique" IOSR, Journal of computer Engineering volume 11, Issue 1 (May-June 2013), pp 61-64.
- [11] Ajit P.sowi, S.R.Khut "facial expression recognition using PCA", International of soft computing and engineering (IJSCE) volume 3, Issue 4, September 2013.
- [12] N.Zermi, ramdani.m " Facial expression recognition using Principal component analysis and artificial neural network" world symposium on computer networks and information security 2014, International Conference on digital Image Processing (ICDIP'2014) 13-15 juin 2014 hammamet Tunisia p51.
- [13] N.Zermi,ramdani.m " Human facial expression recognition based on two dimensional Principal component analysis 2DPCA" world symposium on computer networks and information security 2014 International conference on soft computing, artificial Intelligence and Applications (ICSCAIA'2014), 13-15 juin 2014 hammamet Tunisia p44.

Hybrid Movie Recommendation System using LDA and PSO approaches

Ibtissem Gasmı
Department of Computer Science
Chadli Bendjedid University
El Tarf, Algeria
gasmibtissem@gmail.com

Fouzia Anguel
LISCO Laboratory
Badji Mokhtar University
Annaba, Algeria
fanguel@yahoo.fr

Dalel Merzougui
Chadli Bendjedid University
El Tarf, Algeria
dalel85@yahoo.fr

Abstract— Collaborative Filtering technique is commonly used in the context of recommender systems; however, obtaining better prediction accuracy and overcoming the main limitations of the standard CF recommendation algorithms, such as sparsity. This paper presents a movie context-aware model to alleviate the sparsity problem. The proposed algorithm learns users' latent interests using latent Dirichlet allocation (LDA). Then, by incorporating a weighting mechanism, it combines the explicit contextual parameters and their degree of importance in the prediction process. To learn the weights of these features, the PSO method is used. The experiments showed that the proposed algorithm achieved better performance compared with baseline on MovieLens datasets.

Keywords— Collaborative Filtering, Context, Topic Modeling, PSO, LDA, Sparsity problem.

I. INTRODUCTION

Context-aware recommender systems (CARS) are a promising way of generating more relevant services to users. They produce personalized recommendations in accordance with users' current context [1]. Indeed, in many domains such as electronic commerce, movies, news or restaurant recommenders, considerable contextual information are available along with the rating data. Depending on the application domain, certain contextual information can be very useful for providing better recommendations [2]. Last few years, the importance of contextual information has motivated many researchers to incorporate them in their systems [3], [4]. However, the notion of context is complex and not easily integrated into recommendation process [5].

In the other hand, a good recommender system must learn and integrate the impact of each contextual information to make appropriate suggestions for various situations. Clearly, some contextual features can be more relevant in a given application than others. For example, location is of more importance in restaurant recommendation, whereas gender is the most influential parameters for shoes suggestion.

This paper focuses on item based collaborative filtering algorithm to take advantage of contextual data. The most important step in this technique is to measure the similarity between users or items [6]. However, when items have few co-ratings, similarity based approach suffers from the data sparsity problem, especially for large datasets consisting of a huge number of users and objects. In the other hand, integrating contextual information represents a real challenge in recommender systems. Specifically, data sparsity becomes more severe when user preferences are filtered with contextual factors [7]. In fact, using many contextual parameters increases the data sparsity and few context

factors in recommendations fail to bring the contextual impact [7]. Thus, this study address these two issues by considering the latent properties of items to calculate the similarity scores and by directly incorporating context into the prediction process instead removing items that are irrelevant in a given context.

The main contributions of this paper are summarized as follows: a context-aware collaborative filtering model for movie recommendation is proposed. It combines latent interest and explicit contextual features using Latent Dirichlet Allocation (LDA) and Particle Swarm Optimization (PSO). Firstly, the dimensionality space is reduced by extracting the latent properties from the unstructured text describing items (plot summaries, titles, and genres). The latent contexts are modeled as numeric vectors which are automatically learned through the application of LDA. Moreover, a weighting function is proposed to integrate explicit contextual factors (user's age, user's gender, user's occupation, movie's runtime and evaluation time) with their degree of importance in prediction score calculation. Specially, the time evaluation is introduced into the proposed weighting function to increase the importance of the most recently accessed items because the user's interests change over the time. In addition, the proposed system uses a particle swarm optimization algorithm to assign suitable weights to different contextual features. The efficiency of our model compared to other collaborative filtering methods is demonstrated by checking the implementation on the Movielens 100k data set.

The remainder of this paper is organized as follows: Section 2 reviews related works in context-aware recommender system and the use of LDA and PSO in these systems. In section 3, the proposed model is exposed and a detailed description of how the system employs the LDA and PSO is also provided. Section 4 introduces the dataset, the evaluation settings and metrics. Section 5 discusses the results of the experiments. Finally, section 6 provides conclusion and future works.

II. RELATED WORKS

With the advent of the World Wide Web and big data, recommender systems are becoming more and more popular [8]. The purpose of these systems is to suggest different services to different users. Recently, the research attentions have been raised by context-awareness in the recommender systems. Consequently, the rating function is modeled in a multi-dimensional space as $R: Users \times Items \times Contexts \rightarrow Ratings$ [9]. Particularly, many attempts have been made to exploit the time context for users modeling and recommendation process [10], [11]. Some approaches model the interest drift of users over time as the transition of latent

states within a hidden Markov model (HMM) [4], [12], [13]. Braunhofer and Ricci [14] presented a model which predicts the contextual factors influencing customers when evaluating an item in order to use only relevant contextual information in the recommendation process. In their research work, Zhou et al. [15] selected the optimal features based on the similarity between a feature and a set of other ones to remove the redundancy. They presented a content-context interaction graph model that completely fuses and captures the interactions between content and contexts. Tensor factorization (TF) can also be used to model the time context [16], [17]. It extends the traditional two-dimensional Matrix Factorization problem into a n-dimensional version by integrating contextual information [18]. Ren et al. [19] made a contextual recommendation using location, queries, and web content.

In the other hand, latent context can be very useful in increasing the efficiency of recommender systems. Hence, unsupervised tools such as Latent Dirichlet Allocation (LDA) have been used to learn the latent characteristics [20]. LDA is a machine learning process that associates a text which is seen as an unordered word list with a topic vector [21]. It models each document as random mixtures of latent topics and defines each topic by a word distribution. LDA's popularity is due to its simplicity and modularity [20]. Lin et al. [22] used user reviews from the Airbnb platform for deriving the features of the products and creating customer preferences. The authors implemented the LDA technique to infer both the features and preferences.

A number of researches on recommender systems have exploited Swarm Intelligent technique such as Particle Swarm Optimization (PSO) [23], [24]. PSO is a population based technique firstly developed in 1995 by Kennedy and Eberhart [25]. It is used by many researchers to select feature, or assign weights to different factors in the recommender system [23], [26]. The Work of Ujjin and Bentley [27] applied the PSO to learn users' personal preferences and provide the recommendations based on the adjusted Euclidean distance. Katarya and Verma [28] combined k-means clustering algorithm with Fuzzy c-means and PSO to classify the types of movies according to users. They employed K-Means to provide initial parameters to the PSO algorithm then they use PSO to optimize Fuzzy c-means clustering.

III. SYSTEM OVERVIEW

The present paper aims to address the sparsity problem and improve the prediction in collaborative filtering systems. Consequently, the proposed method exploits the hidden properties of items to compute the similarity matrix using LDA. Then, it directly integrates context in the prediction process by introducing a weighting function which combines contextual information with their degree of importance. The overall framework of the training stage of TMPPO-CF model is presented in Figure 1. In summary, the following steps are performed :

A. Building item neighborhood using LDA

This paper uses the LDA to identify latent properties of items from their textual descriptions (plot summaries, titles and genres). The item similarity scores are then computed in the latent topic space to estimate the movies neighborhood.

The movie's similarity scores are computed through the following steps :

A.1. Preprocessing unit

In the first phase, the textual descriptions of movies are preprocessed by the python NLTK toolkit to filter out noisy data and non-informative words that do not add any distinctive character in the movies. In particular, each sentence is tokenized into a list of words. The stopwords, punctuations, words with low or high inter-document frequency are removed. The upper-case is conversed to lower-case. Finally, the terms stemming from the same lemma having differences due to inflectional morphologies are unified (Lemmatization). Then, all preprocessed words are extracted to build the vocabulary. Thus, each movie is considered as a bag of words (BoW). It is represented as a vector in this term space, with value in each cell indicating the number of occurrences of the corresponding word.

A.2. Preprocessing unit

The preprocessing step is followed by the training of a topic model through LDA to learn the hidden properties in terms of the probability distributions of topics in each movie, and the probability distributions of words over each topic. Given N movies with a textual description of the movie i ($i \in \{1, \dots, N\}$) having V_i words, the generative process is shown in the LDA Algorithm 1. LDA generates the probability distributions of topics $p(T|I)$ for each movie I , where the topics $T = (T_1, T_2, \dots, T_{n_topics})$ are composed of the word probabilities $p(W_j|T_i)$ for words W_j , $j=1, \dots, n_words$ where n_words is the number of words in the vocabulary. Note that n_topics is the pre-defined number of topics. The topics are considered as the features of movies, and the topic density in the model is controled by Dirichlet hyperparameters α and β which play an important role in learning accurate topic models. The pseudo code of LDA algorithm used by the proposed model is shown in the LDA Algorithm 2.

LDA Algorithm 1

Step 1: Choose a multinomial distribution ϕ_t for topic t ($t \in \{1, \dots, T\}$) from a Dirichlet distribution with parameter β .

Step 2: Choose a multinomial distribution θ_i for movie i ($i \in \{1, \dots, N\}$) from a Dirichlet distribution with parameter α .

Step 3: For a word w_j ($j \in \{1, \dots, V_i\}$) in the movie i :

Step 3.1: Select a topic z_j from θ_i .

Step 3.2: Select a word w_j from ϕ_{z_j} .

LDA Algorithm 2

Input: Movielens 1M dataset, CMU Movie Summary Corpus, initial values of α , β and T .

Output: Movie-topic probability distribution, Topic-word probability distribution.

Step 1: Preprocessing textual descriptions of movies.

Step 2: Building the vocabulary.

Step 3: Converting the movies textual descriptions into simple vectors representation (BoW).

Step 4: Training the LDA model using the Gibbs sampling method.

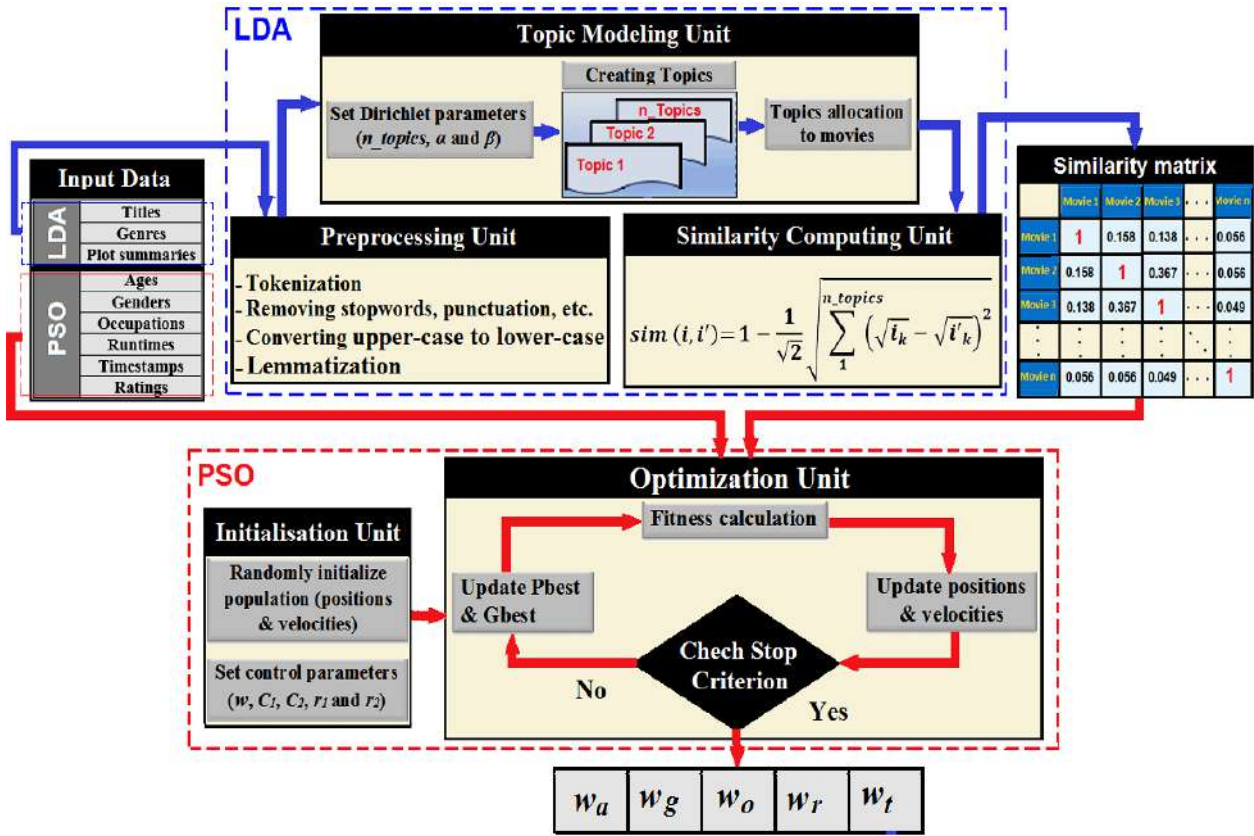


Fig. 1. Training stage of the proposed model

A.3. Similarity computing unit

Each movie is represented in this latent topic space using the movie-topic distribution as its feature vector. The similarity between items i and i' in the latent topic space is from Hellinger distance (HD) between their latent topic distributions. It is expressed by Eq. (1):

$$sim(i, i') = 1 - HD(i, i') = 1 - \frac{1}{\sqrt{2}} \sqrt{\sum_{k=1}^T (\sqrt{t_k} - \sqrt{t'_k})^2} \quad (1)$$

B. Particle Swarm Optimization

This study assumes that ratings, which are most similar in context, are more relevant in making predictions. Thus, the PSO is used to find the most appropriate feature's weights in order to control the contribution of the contextual factors in the recommendation process. In the PSO algorithm, a swarm is a group of particles which represent candidate solutions. For a d-dimensional space, the particles correspond to d-dimensional vectors. The position and velocity of the i th particle are represented by $X_i = (x_{i1}, x_{i2}, \dots, x_{id})$ and $V_i = (v_{i1}, v_{i2}, \dots, v_{id})$, respectively.

B.1. Representation scheme and initial population

In this paper, the positions of particles are viewed as weight vectors of real values in the range of [0, 1]. In addition, the sum of values of each position should be equal to 1. The dimension of the positions is equal to the number of the used features, i.e. 5. A position of one particle can be represented by the vector $(w_a, w_g, w_o, w_r, w_t)$. Where w_a is the weight of the user's age, w_g denotes the weight associated to the user's gender, w_o represents the weight of user's occupation. w_r is the weight corresponding to the movie's

runtime and w_t designates the weight of the evaluation time. Velocity and position of the particles are initialized in a uniform random manner.

B.2. Fitness function

The objective function forms the heart of the PSO algorithm. It is used to quantify the optimality of the weight vector. The contextual feature values used in this study are multiplied by the weight vector to compute the weighting function which is used in the prediction calculation. Thus, the RMSE measure is treated as the fitness of the PSO algorithm used by the proposed model. Consequently, to achieve the lowest fitness, it is necessary to search the optimal position. The fitness function is given by Eq. (2):

$$fitness = \sqrt{\frac{\sum_{(u,i) \in Q} (P_{ui} - P'_{ui})^2}{Q}} \quad (2)$$

$|Q|$ is the cardinality of the training set. P'_{ui} and P_{ui} are real and predicted ratings, respectively. The prediction on an item i for a user u is computed by the sum of the ratings given by the user u on the items similar to i . Each ratings is weighted by the corresponding similarity $sim(i, j)$ between items i and j and by a weighting value $f(u, j)$ which is used to model the context of the user u for the item j . Formally, the predicted rating pu, i is given by Eq. (3):

$$P_{ui} = \bar{R}_u + \frac{\sum_{j=1}^k (R_{uj} \times sim(i, j) \times f(u, i))}{\sum_{j=1}^k (sim(i, j) \times f(u, i))} \quad (3)$$

$Sim(i, j)$ is the degree of similarity between items i and j . Ru, j is the rating of user u for item j . K is the neighborhood size of item i . The proposed weighting function $f(u, j)$ is formulated via a simple linear combination of values and weights to represent the context of ratings. Its values are in the range of 0–1. It is defined according to Eq. (4):

$$f(u, j) = \sum_{l=a,o,r} w_l [h_l(x_l)] + w_g [\beta] + w_i [e^{-t \times w_i}] \quad (4)$$

Where $(w_a, w_g, w_o, w_r, w_i)$ represents the weight vector. x_g and x_o reveal the coded variables that correspond to a user's gender and user's occupation, respectively. x_a and x_r are the values of the user's age and movie's runtime, respectively. The variables $x_l, l=a, g, o, r$ are normalized through the eq. (5):

$$h_l(x) = \frac{x - \min(l)}{\Delta l} \quad (5)$$

Δl represents the interval of variation of the contextual information l , where $l \in [\text{age } (a), \text{gender } (g), \text{occupation } (o), \text{runtime } (r)]$.

$e^{-t \times w_i}$ is defined in the range of [0, 1]. It is a monotonic decreasing coefficient with an exponential form in order to express the changes of user's interests over time by making users' most recent evaluations more significant than the older ones. The parameter t can be expressed by Eq. (6):

$$t = \frac{t_B - t_A}{\Delta t} \quad (6)$$

t_0 represents the time of the last evaluation given by user u . t_d is the time of the last evaluation given by the user u to an item having the same genre as item i . Δt corresponds to the average duration between each two successive rating of user u .

B.3. Particle dynamics and termination condition

At each time step, each particle i keeps the information about its personal best position (pbest $P_i(t)$) and the swarm best position (gbest $g_i(t)$). The particles move throughout the search space by updating their velocities and their positions according to equations (7) and (8), respectively:

$$v_j(t+1) = wv_j(t) + c_1 r_1 (P_j(t) - x_j(t)) + c_2 r_2 (g_j(t) - x_j(t)) \quad (7)$$

$$x_j(t+1) = x_j(t) + v_j(t+1) \quad (8)$$

N is the number of particles in the swarm. The inertia weight w is an important parameter to strike a better balance between global and local exploitation. c_1 and c_2 are positive acceleration con The inertia weight w is an important parameter to strike a better balance between global and local exploitation. The linear decreasing inertia weight has been widely used to update the value of w according to Eq. (9):

$$w(l) = w_{start} - l \times (w_{start} - w_{end}) / Nb_{max} \quad (9)$$

Where Nb_{max} is the maximum number of iterations, l is the iteration index, w_{start} and w_{end} are the maximum and minimum values of the inertia weight, respectively.

c_1 expresses the confidence of a particle in itself, while c_2 expresses the confidence of a particle in its neighbors. r_1 and r_2 are random numbers within the range of [0, 1]. The stop criterion of the PSO algorithm is to reach a maximum number of iterations. At the end of the algorithm, we get the optimal weight vector which will be used to calculate predictions.

IV. EXPERIMENTAL EVALUATION

To demonstrate the effectiveness of the proposed model, several experiments have been conducted on the most popular Movielens¹ dataset and CMU Movie Summary Corpus². All experiments have been performed by randomly splitting Movielens dataset into two subsets: training set (80%) and testing set (20%). Two types of experiments have been performed: The first one concerns the training of the LDA and PSO algorithms. It consists in carrying out numerous tests to select the best configuration of their parameters. The second category concerns the evaluation of the overall quality of the proposed model after learning it.

A. Datasets

Movielens is widely used to evaluate recommender systems. This dataset is provided by the University of American Minnesota GroupLens project group. The ML100K dataset consists of 943 users, 1682 movies and 100,000 ratings. The information of users includes: UserID, Gender, Age and Occupation and those of the items are MovieID, Title and Genres. All ratings are contained in the file ratings.dat. Each line of this file represents one rating of one movie by one user, and has the following format: UserID::MovieID::Rating::Timestamp. The lines within this file are ordered first by UserID, then, within the user, by MovieID. Ratings are made on a 5 star scale. Timestamps represents seconds since midnight Coordinated Universal Time (UTC) of January 1, 1970. The sparsity of the ML100K dataset is 93.7%.

The CMU Movie Summary Corpus is also used in this paper. It contains plot summaries for 42,306 movies and metadata for 81,741 movies such as genre, release date, runtime, languages, etc. Each movie is indexed by a Wikipedia Movie ID.

B. Evaluation Metrics

The proposed model is evaluated from two different perspectives: predictive accuracy and classification accuracy. Predictive accuracy measures the accuracy of a prediction value on a target item. The Root Mean Squared Error (RMSE) is The frequently used metrics in predictive accuracy [29]. The formula and RMSE is defined in equations (10).

$$RMSE = \sqrt{\frac{\sum_{(u,i) \in Q} (P_{ui} - P'_{ui})^2}{Q}} \quad (10)$$

¹ <http://www.grouplens.org/>

² <http://www.cs.cmu.edu/~ark/personas/>

$|Q|$ is the cardinality of a training set. $P'_{u,i}$ and $P_{u,i}$ are real and predicted ratings, respectively.

Classification accuracy evaluates the rate of correct predictions. Precision and Recall are the two popular metrics for classification accuracy. Precision is a percentage of recommended items that are relevant while recall is a percentage of the relevant item that is recommended [29]. Precision and Recall are defined in equations (11) and (12), respectively.

$$\text{Precision} = \frac{\text{Number of correct recommendations}}{\text{Total number of recommendations}} \quad (11)$$

$$\text{Recall} = \frac{\text{Number of correct recommendations}}{\text{Total number of relevant recommendations}} \quad (12)$$

V. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, each user is considered to be an active user and the item ratings for each of them are predicted. The recommendation algorithms are evaluated by comparing the predicted ratings with the actual ratings. The proposed model is compared with some baseline models. Item-CF denotes the item based collaborative filtering algorithm. Context-CF represents the item based collaborative filtering algorithm using the proposed weighting function defined in Eq. (4) with $w_l = 1$ for each $l = a, g, o, r, t$. LDA-CF denotes the item based collaborative filtering using LDA algorithm. PSO-CF is the item based collaborative filtering using the PSO algorithm to estimate the weight vector of the proposed weighting function. Since a different neighborhood size K leads to different recommendation results, the impact of K is also considered in this experiment. Recommendation results are calculated and compared under the condition of different K . The training modules of LDA and PSO algorithms are learned only once, and ratings in the test set have never been used in the training process. The Table 1 summarizes the best settings obtained after the adjusting the parameters of the LDA and PSO algorithms used in the TMP-PSO-CF model. Notice that the parameter ω is linearly decreased from 0.9 to 0.4 during the iterations.

TABLE I. CONFIGURATION PARAMETERS

Algorithm	Parameter	Value
PSO	Number of iterations	80
	Number of particles	40
	Initial inertia weight (w_{start})	0.9
	Final inertia weight (w_{end})	0.4
	c_1	1.5
	c_2	1.5
LDA	Number of topics	50
	α	0.5
	β	0.02

Figure 2 illustrates the evolution of the RMSE through the neighborhood size. It can be seen that the error decreased with the increase of neighbour count for Item-CF approach. However, for the other methods, the errors start with a relatively high value and then they reach the lowest points at $k = 20$. It is also observed that the values of RMSE almost increase with the increase of the neighborhood size in the

range of 20–100. Furthermore, all values of TMP-PSO-CF, PSO-CF, Context-CF and LDA-CF are lower than traditional item-CF. Consequently, models that consider latent interests and/or explicit context perform better than the item based collaborative filtering algorithm.

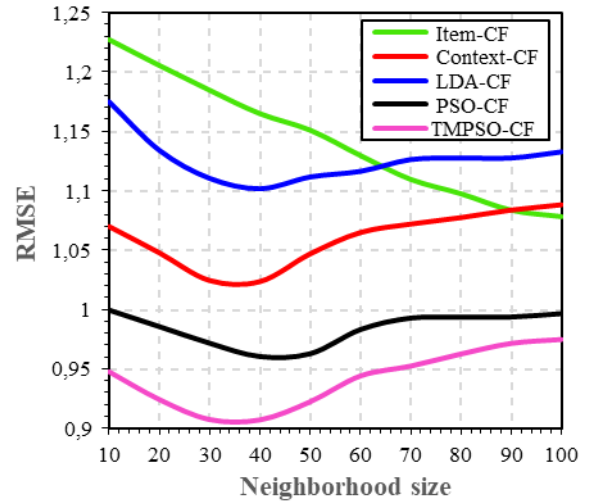


Fig. 2. RMSE with different neighborhood size

Figure 3 outlines the evolution of Precision through the neighborhood size. The results indicate that the proposed method performs best. Moreover, under different neighborhood size, our model mimics PSO-CF, and both perform significantly better compared with the remaining methods. For Item-CF, the Precision follows an increasing trend in the beginning, and it remains stable when the size of the neighborhood is greater than 40. This means that the number of the relevant items increases when more similar items are considered.

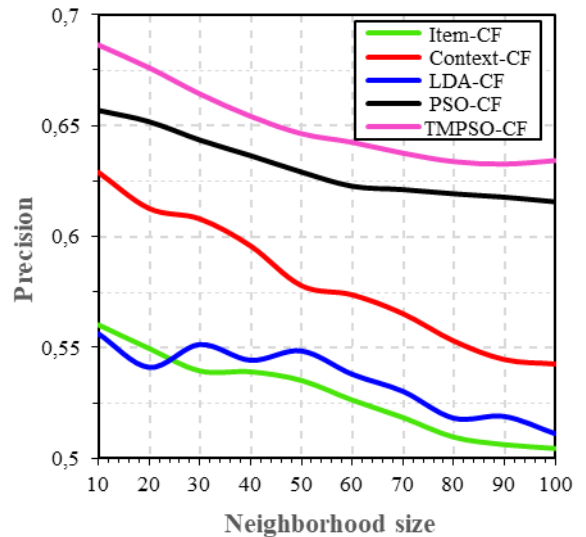


Fig. 3. Precision with different neighborhood size

Figure 4 illustrates the evolution of Recall through the neighborhood size. It is clear that the increase of the neighborhood size decreases the Recall of our model, PSO-CF and Context-CF when $K \geq 20$; and decreases the Recall of LDA-CF when $K \geq 40$. The probable reason is the importance of considering the most similar items for these models. To sum up, the experiments also verify that the

proposed method enhance prediction performance. This is beneficial to find the proper tradeoff between efficiency and performance.

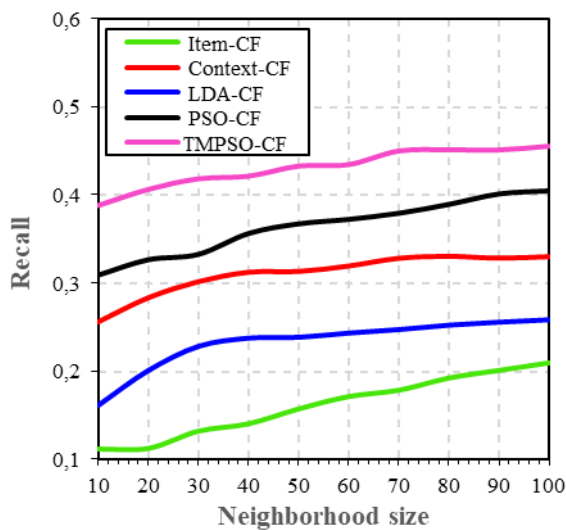


Fig. 4. RECALL with different neighborhood size

VI. CONCLUSION AND FUTURE WORK

This study, proposes a context aware collaborative filtering method to address the sparsity problem and improve the prediction in movie recommender systems. It combines Latent Dirichlet Allocation (LDA) and Particle Swarm Optimisation (PSO) to suggest films that are not only popular in the community, but also ones that are similar in content and context. Experiments on Movielens data set, show that the proposed model achieves significantly better recommendation quality than the baseline collaborative filtering methods. In a future work, we intend incorporating additional contextual information to extend the proposed model.

REFERENCES

- [1] G. Adomavicius, B. Mobasher, F. Ricci, et A. Tuzhilin, « Context-Aware Recommender Systems », *AI Magazine*, vol. 32, p. 67-80, sept. 2011.
- [2] G. Adomavicius et A. Tuzhilin, « Context-Aware Recommender Systems », in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, et P. B. Kantor, Éd. Boston, MA: Springer US, 2011, p. 217-253. doi: 10.1007/978-0-387-85820-3_7.
- [3] Y. Zheng, « Context-Aware Mobile Recommendation By A Novel Post-Filtering Approach », présenté à The Thirty-First International Flairs Conference, mai 2018. Consulté le: 31 mai 2020. [En ligne]. Disponible sur: <https://aaai.org/ocs/index.php/FLAIRS/FLAIRS18/paper/view/17661>
- [4] Z. Lin et H. Chen, « Recommendation over time: a probabilistic model of time-aware recommender systems », *Sci. China Inf. Sci.*, vol. 62, n° 11, oct. 2019, doi: 10.1007/s11432-018-9915-8.
- [5] M. Bazire et P. Brézillon, « Understanding Context Before Using It », in *Modeling and Using Context*, Berlin, Heidelberg, 2005, vol. 3554, p. 29-40. doi: 10.1007/11508373_3.
- [6] J. Herlocker, J. A. Konstan, et J. Riedl, « An Empirical Analysis of Design Choices in Neighborhood-Based Collaborative Filtering Algorithms », *Information Retrieval*, vol. 5, n° 4, p. 287-310, oct. 2002, doi: 10.1023/A:1020443909834.
- [7] V. S. Dixit et P. Jain, « Recommendations with Sparsity Based Weighted Context Framework », in *Computational Science and Its Applications – ICCSA 2018*, Cham, 2018, p. 289-305. doi: 10.1007/978-3-319-95171-3_23.
- [8] P. Moradi et S. Ahmadian, « A reliability-based recommendation method to improve trust-aware recommender systems », *Expert Syst. Appl.*, vol. 42, n° 21, p. 7386-7398, 2015, doi: 10.1016/j.eswa.2015.05.027.
- [9] Y. Zheng, « Interpreting Contextual Effects By Contextual Modeling In Recommender Systems », *IDM Workshop @ CIKM'17*, 2017.
- [10] P. G. Campos, F. Díez, et I. Cantador, « Time-aware recommender systems: a comprehensive survey and analysis of existing evaluation protocols », *User Model User-Adap Inter.*, vol. 24, n° 1, p. 67-119, févr. 2014, doi: 10.1007/s11257-012-9136-x.
- [11] G. Xu, Z. Tang, C. Ma, Y. Liu, et M. Daneshmand, « A collaborative filtering recommendation algorithm based on user confidence and time context », *Journal of Electrical and Computer Engineering*, vol. 2019, n° 6, p. 1-12, 2019, doi: 10.1155/2019/7070487.
- [12] N. Sahoo, param vir singh, et T. Mukhopadhyay, « A Hidden Markov Model for Collaborative Filtering », *Management Information Systems Quarterly*, vol. 36, n° 4, p. 1329-1356, déc. 2012.
- [13] H. Zhang, W. Ni, X. Li, et Y. Yang, « Modeling the Heterogeneous Duration of User Interest in Time-Dependent Recommendation: A Hidden Semi-Markov Approach », *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, n° 2, p. 177-194, févr. 2018, doi: 10.1109/TSMC.2016.2599705.
- [14] M. Braunhofer et F. Ricci, « Selective contextual information acquisition in travel recommender systems », *Inf Technol Tourism*, vol. 17, n° 1, p. 5-29, mars 2017, doi: 10.1007/s40558-017-0075-6.
- [15] X. Zhou, D. Qin, L. Chen, et Y. Zhang, « Real-time context-aware social media recommendation », *The VLDB Journal*, vol. 28, n° 2, p. 197-219, avr. 2019, doi: 10.1007/s00778-018-0524-7.
- [16] E. Frolov et I. Oseledets, « Tensor methods and recommender systems », *WIREs Data Mining and Knowledge Discovery*, vol. 7, n° 3, p. 1-25, 2017, doi: 10.1002/widm.1201.
- [17] M. Abdi, G. Okeyo, et R. Mwangi, « Matrix Factorization Techniques for Context-Aware Collaborative Filtering Recommender Systems: A Survey », *Computer and Information Science*, vol. 11, n° 2, Art. n° 2, mars 2018, doi: 10.5539/cis.v11n2p1.
- [18] Y. Fang et Y. Guo, « A context-aware matrix factorization recommender algorithm », in *2013 IEEE 4th International Conference on Software Engineering and Service Science*, mai 2013, p. 914-918. doi: 10.1109/ICSESS.2013.6615454.
- [19] Y. Ren, M. Tomko, F. D. Salim, J. Chan, C. L. A. Clarke, et M. Sanderson, « A Location-Query-Browse Graph for Contextual Recommendation », *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, n° 2, p. 204-218, févr. 2018, doi: 10.1109/TKDE.2017.2766059.
- [20] M. Al-Ghossein, P.-A. Murena, T. Abdesslem, A. Barré, et A. Cornuéjols, « Adaptive collaborative topic modeling for online recommendation », in *Proceedings of the 12th ACM Conference on Recommender Systems*, British Columbia, Canada, sept. 2018, p. 338-346. doi: 10.1145/3240323.3240363.
- [21] D. M. Blei, A. Y. Ng, et M. I. Jordan, « Latent dirichlet allocation », *J. Mach. Learn. Res.*, vol. 3, n° null, p. 993-1022, mars 2003.
- [22] K.-P. Lin, C.-Y. Shen, T.-L. Chang, et T.-M. Chang, « A Consumer Review-Driven Recommender Service for Web E-Commerce », in *2017 IEEE 10th Conference on Service-Oriented Computing and Applications (SOCA)*, nov. 2017, p. 206-210. doi: 10.1109/SOCA.2017.35.
- [23] S. Yang, Q. Peng, et L. Chen, « The BPSO Based Complex Splitting of Context-Aware Recommendation », in *Computational Intelligence and Intelligent Systems*, Singapore, 2016, p. 435-444. doi: 10.1007/978-981-10-0356-1_46.
- [24] L. Peška, T. M. Tashu, et T. Horváth, « Swarm intelligence techniques in recommender systems - A review of recent research », *Swarm and Evolutionary Computation*, vol. 48, p. 201-219, août 2019, doi: 10.1016/j.swevo.2019.04.003.
- [25] G. Sumathi, S. Sendhil Kumar, et G. S. Mahalakshmi, « Hybrid Recommendation System using Particle Swarm Optimization and User Access Based Ranking », in *Proceedings of the International Conference on Informatics and Analytics*, Pondicherry, India, août 2016, p. 1-9. doi: 10.1145/2980258.2980405.
- [26] A. F. Rozie et H. P. In, « Swarm collaborative filtering through fish school search », *International Journal of Software Engineering and*

- its Applications*, vol. 8, n° 3, p. 251-254, 2014, doi: 10.14257/ijseia.2014.8.3.23.
- [27] P. Choudhary, V. Kant, et P. Dwivedi, « A Particle Swarm Optimization Approach to Multi Criteria Recommender System Utilizing Effective Similarity Measures », in *Proceedings of the 9th International Conference on Machine Learning and Computing*, Singapore, Singapore, févr. 2017, p. 81–85. doi: 10.1145/3055635.3056619.
- [28] R. Katarya et O. P. Verma, « A collaborative recommender system enhanced with particle swarm optimization technique », *Multimed Tools Appl*, vol. 75, n° 15, p. 9225-9239, août 2016, doi: 10.1007/s11042-016-3481-4.
- [29] G. Shani et A. Gunawardana, « Evaluating Recommendation Systems », in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, et P. B. Kantor, Éd. Boston, MA: Springer US, 2011, p. 257-297. doi: 10.1007/978-0-387-85820-3_8.

Implementation of the FIR filter under DSP and FPGA (comparative study)

Farid BENAIDA

Computer Science Department

Badji Mokhtar University,

Annaba, Algeria

benaidafarid@yahoo.fr

Habiba BELEILI

Computer Science Department

Badji Mokhtar University,

Annaba, Algeria

h.belleili@gmail.com

Nacer eddine HAMMAMI

Computer Engineering Dept. (ABET
Accredited)College of Engineering and Computer
Sciences,

Mustaqbal University, Saudi Arabia

Nacereddine.hammami@gmail.com

Radia TOUAHRI

Computer Science Department

Badji Mokhtar University,

Annaba, Algeria

radia_inf@yahoo.fr

Abstract - The purpose of this article is not to compare the execution of a program (written in C language) executed under a general microprocessor and its execution under FPGA in sequential then in parallel. Because it is evident that the FPGA which will win this duel especially in parallel execution. Our aim is to beat a specialized microprocessor such as DSPs which are designed for signal processing by an implementation under FPGA. This article revolves around DSPs, their definition and their optimization tactics, and we have taken as an example the DSP C6000 from the giant Texas Instrument [1].

We chose an FIR filtering algorithm to implement it under DSP. We have proposed an optimization strategy on this DSP to make the most of the reliability of the C6000 DSPs. Then the same algorithm will be carried out and materialized under FPGA in sequential and then in parallel. The results and discussions will come at the end of this article.

I. INTRODUCTION

The most famous processing applied to a signal is filtering. And the more the signals that we are going to process are of an analog nature we need to make assemblies using analog components (resistor, capacitor, coil, etc.) [2]

Analog components are always susceptible and sensitive to temperature, mechanical vibrations, variation in supply voltage and the tolerance of the values of each component.

Each time we are going to have a change in the characteristics of the filter to modify or adjust the transfer function of an analog filter, we are obliged to change the electronic card or some elements of this card because its components are less configurable and more than not reconfigurable.

To fight against this? we have to digitize the analog signal then apply digital filters

Digital filters are invented to make flexible modifications or adjustments [3]; because these are algorithms implemented on a processor dedicated to digital signal processing such as DSPs or reconfiguration and programmable platforms such as CPLD and FPGA circuits (figure 1)

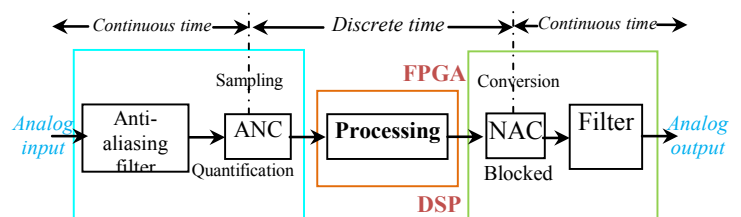


Fig. 1 : Digitization cycle and signal filtering

A. Analog signal digitization

1) *Position of the problem:* We talk about signal processing when we talk about: Speech processing, video processing or image processing.

2) *Need for a specialized μP :* We need micro-processors that can run around 25 MOPS.

TABLE I
PROCESSING CAPACITY FOR ONE μP AT MEDIUM POWER (25 MOPS)

	Frequency of sampling	Period between two samples	μP of 25 MOPS
speech	8 KHZ	125 μs	1250 operations
Audio	44 KHZ	22.7 μs	227 operations
Video	5 KHZ	200 ns	2 operations

The table above (Table 1) gives the approximate number of operations achievable between two samples for a μP of 25 MOPS. As can be seen, the applications that can be implemented may be relatively more complex for signals in the speech band, while the processing of a video signal in real time seems out of reach for such μP [4]. This explains the major efforts made by manufacturers to market architectures capable of supporting multimedia applications.

3) *Need for a specialized μP (DSP):* DSPs are able to work in real time, do not require an operating system to supervise, small in size, architecture adapted for digital signal processing, RISC coding and low power consumption [5]. With the key to digital signal processing algorithms which is the sum of products (SOP).

Is a DSP equipped with a unit that calculates the sum of the products (MAC) more than enough to work comfortably in real time?

So, what are the modifications introduced on the architectural axis of μP to move towards real-time processing?

- From Von-Neumann architecture to HARVARD architecture.
- From single-bus architecture to multi-bus architecture.
- From single DATA space to multi-bank DATA space with the use of cache memories.
- From non-pipelined execution to pipelined execution.
- From the execution of a single instruction per cycle to the execution of a packet of instructions (parallelism).

B. Filter operation

1) *Temporal filtering*: We let the signal pass in moments if necessary.

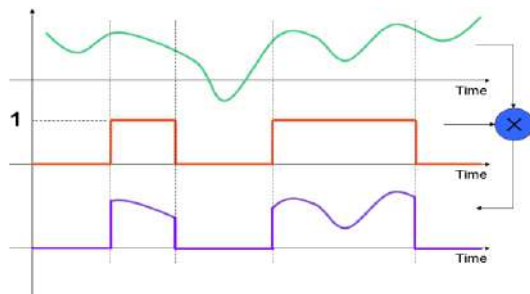


Fig. 2 : temporal filtering

$$y(t) = x(t) \cdot \sum_k \text{rect}(t - t_k)$$

We observe that the temporal filtering is a simple multiplication with rectangles of level "1"

2) *Frequency filtering*: Frequencies are passed through as needed

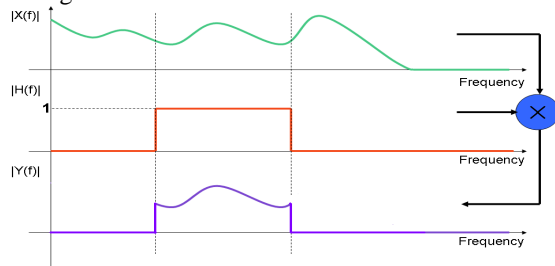


Fig.3 : frequency filtering

On frequency is a simple multiplication, so on time is a convolution

$$\begin{aligned} Y(f) &= X(f) \times H(f) \\ Y(f) &= X(f) \times H(f) \\ &\Downarrow TF^{-1} \end{aligned} \quad (2)$$

$$y(t) = x(t) * h(t) = \int x(\tau) h(t - \tau) d\tau$$

3) *Classes of digital filters*: In the literature of digital filters [2], we find the following two classes:

- Classes of -FIR- filters (with finite impulse response).
- Classes of -IIR- filters (infinite impulse response).

In what follows we will study only the class of FIR filters.

4) *FIR Filter*: The FIR filter is a FEED FORWARD system, its output is given by the following equation (digital convolution).

$$y(n) = \sum_{k=0}^{M-2} b(k)x(n-k)$$

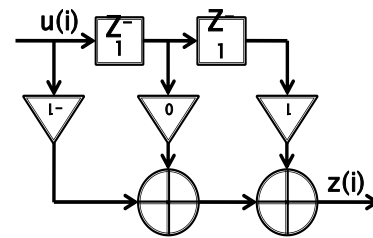
$$y(0) = \sum_{k=0}^2 b(k)x(0-k) = b(0)x(0) + b(1)x(-1) + b(2)x(-2) \quad (3)$$

Causal $x(t)=0$ if $t<0$

$$\begin{aligned} y(1) &= \sum_{k=0}^2 b(k)x(1-k) = b(0)x(1-0) + b(1)x(1-1) + b(2)x(1-2) \\ &= b(0)x(0) = b(0)x(1) + b(1)x(0) \end{aligned} \quad (4)$$

Either the filter FIR $b = \{-1, 0, 1\}$.

- Implementation



We note that after each calculation of $z(i)$ by the SOP formula, the memory elements in Z^{-1} form a buffer with a shift of the samples by $x(i)$.

So the filter behavior is as follows:

- Loading of the 1st buffer element by $x(i)$.
- Calculation of $z(i)$ by the SOP.
- Buffer is shifting by one position.

So, really in C/C++ we will program the behavior of the filter b.

Code in C/C++

```
// load the 1st buffer element by x(i)
buffer = x[i];
// calculation of the output of the filter by the
SOP
z(i) = 0;
for(k=0; k<3; k++) {
    z(i) += buffer[k]*b[k];
}
// shift the buffer by one position
for(k=2; k>0; k--) {
    buffer[k] = buffer[k-1];
}
```

C. IMPLEMENTATION OF THE ALGORITHM UNDER DSP C6X (C6000 CPU):

1) *Presentation of DSP C6x (C6000 CPU)*: To effectively implement the behavior of the FIR filter, it is necessary to work with a specialized processor such as DSPs. In this article, we will use the DSPs of Texas Instrument (TI) class C6x. This class has the following general view:

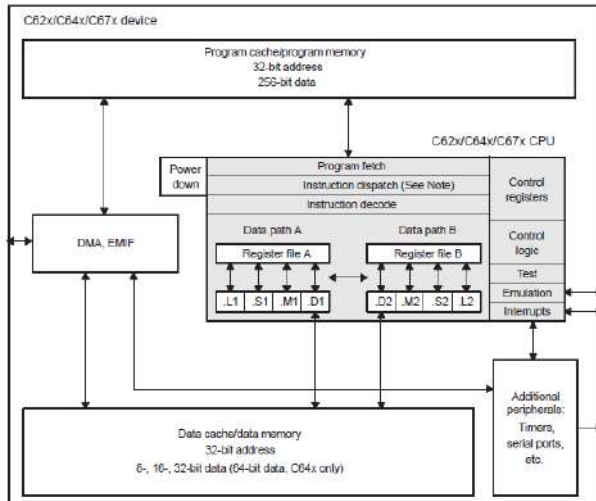


Fig. 5 : Architecture of DSPs RI C6X

C6x DSP (C6000 CPU) has [5] :

- A CPU of two REG FILE A & B banks, 8 units (4 per bank).
- 2 cache memories (DATA CACHE & CODE CACHE).
- A DMA (Direct Memory Access).
- An EMIF (External Memory Interface).
- A set of peripherals (Timers, Serial ports, etc.)

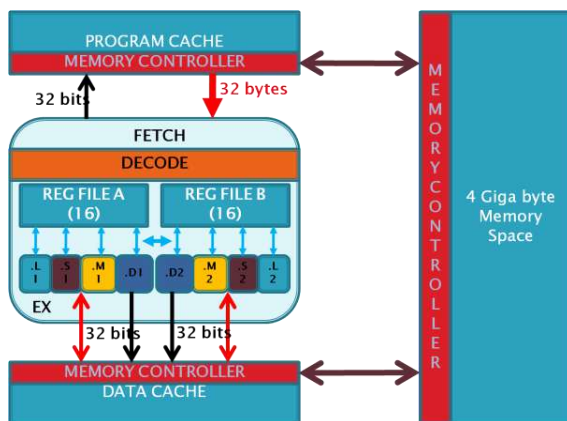


Fig. 6 : programmable unities of DSP C6X

Our CPU (Fig. 6) is a 32-bit extended to 40-bit processor based on :

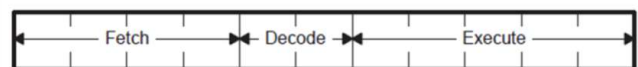
LOAD/STORE technology [7] (RISC type instruction set)

- There are no direct operations between the internal registers and the addressable memory space).

- A memory space of 4 Giga byte addressable in byte.
- 16 32-bit registers for bank A and 16 32-bit registers for bank B.
- 4 units per bank (Load, Memory, Store and Data), therefore a total of 8 units.
- Byte-addressable DATA cache memory.
- Addressable PROGRAM cache memory in 32 bytes.
- Two accesses can be made at the same time to the DATA space.
- Load and Store units perform 32/40 bit operations.
- Memory and Data units perform 32-bit operations only.
- The destination of units L1, S1, M1 and D1 is always a register of REG FILE A.
- The destination of the L2, S2, M2 and D2 units is always a register of REG FILE B.
- The source of the L, S and M units can be from the same REG FILE or a mixture between the two REG FILE (use of the crossed path).
- The source of unit D is always from the same REG FILE.
- The only unit that generates addresses to the DATA space is unit D either through the direct or cross path.

The C6000 CPU uses a pipeline of 11 stages corresponding to the three phases FETCH (4 stages), DECODE (2 stages) and EXECUTE (5 stages). [6]

Fixed-Point Pipeline Stages



The goal is not the study of DSP, but we will exploit it to implement our FIR filter in order to compare with a hardware solution on FPGA circuits.

In this section, we will present how to optimally implement the FIR filter on the Texas Instrument C6x DSP.

2) *Implementation on the C6x DSP*: In this section, we will present how to optimally implement the FIR filter on the Texas Instrument C6x DSP

- The first step: make the C/C++ of the algorithm to implement.

```
N = 20; // nombre de coefs dans le filtre
S = 0;
for (i=0;i<N;i++) {
    S = S + TAB1[i]*TAB2[i];
}
```

The SOP formula is a simple Dot.Prod of two arrays as coded in C/C++

- The second step: make a detailed organization diagram.

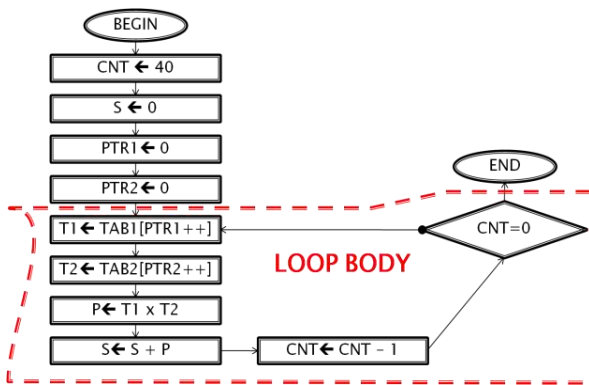


Fig. 7 : Diagram of FIR program

- Third step: translate the flowchart into linear assembler (LOOP BODY ONLY)

```

LOOP : LDH    *PTR1++, T1
      LDH    *PTR2++, T2
      MPY    T1, T2, P
      ADD    P, S, S
      SUB    CNT, 1, CNT
[CN]   B      LOOP
    
```

- Fourth step: make the dependency graph (DEPENDENCY GRAPH)

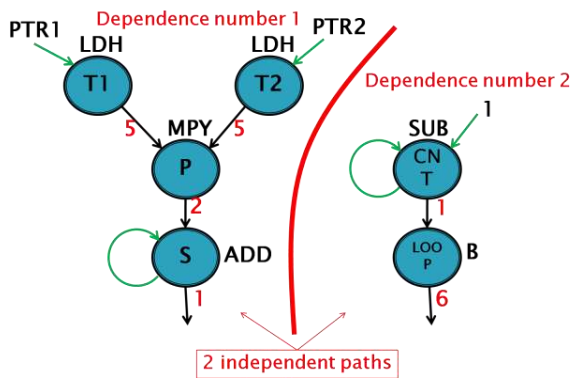


Fig. 8 : dependency graph of the instructions

- Fifth step: calculate the number of cycles per branch:
1st branch = 5 + 2 + 1 = 8 cycles
2nd branch = 1 + 6 = 7 cycles
- Sixth step: select the longest branch and align the independent paths to end at the same cycle by the instruction (SCHEDULING TABLE).

Cycle	.D1	.D2	.M1	.M2	.L1	.L2	.S1	.S2	NOP
1	LDH 1	LDH 2							
2						SUB			
3							B		
4									NOP
5									NOP
6			MPY						
7									NOP
8				ADD					

Fig. 9 : SCHEDULING TABLE

- Seventh step: allocation of registers on the light of the scheduling table.

PTR1, T1 → A4, A6 (.D1)

PTR2, T2 → B4, B6 (.D2)

P → A5 (.M1x)

S → A0 (.L1)

CNT → B0 (.L2)

LOOP → (.S2)

- Eighth step: translates the scheduling table with register allocation to the real C6000 assembler (LOOP BODY ONLY)

```

LOOP:  LDH.D1    *A4++, A6
      || LDH.D2    *B4++, B6 (1)
      SUB.L2     B0, 1, B0 (1)
[B0]   B.S2      LOOP      (1)
      NOP                (2)
      MPY.M1X    B6, A6, A5 (1)
      NOP                (1)
      ADD.L1     A5, A0, A0 (1)
    
```

- Ninth step: calculate the number of cycles to execute the code.

Number of cycles=(1+1+1+2+1+1+1)xN cycles

If N = 40 then 8x40 = 320 cycles. (N is the number of samples to process.)

There are other optimization methods on the DSP C6000 that give a fairly significant gain on the execution speed axis, such as the software pipelining method. Also, Loop Unrolling, Word wide (VLIW), Software pipelining. Software pipelining presents the most efficient method to optimize a code of a loop. And this method is based on the combination between the complementary methods and the maximization of the use of the calculation units per cycle or more (stage).

D. Materialization of the program

Under FPGA (low density) → sequential approach

Under high density FPGA → Parallel approach

1) Sequential approach:

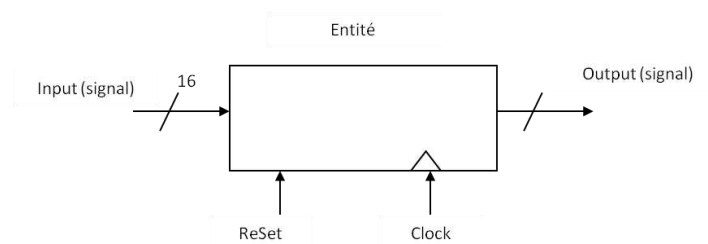


Fig. 10: general view of FIR filter circuit

- Input x[i]
- Pointer (k)
- B constants → ROM
- Accumulator S

The proposed circuit to materialize the FIR filter in its global view is composed of:

- Input $x[i]$ of size 16 bits containing the digitized value of the signal to be processed
- Clock: to synchronize circuit operations
- Reset to return to the initial circuit state
- The 32-bit size signal values after processing by the circuit.

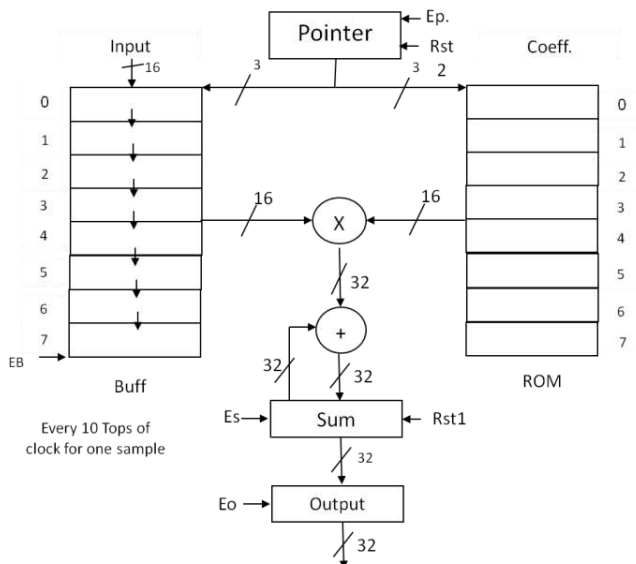
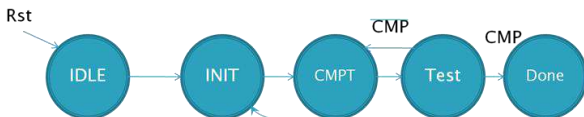


Fig. 11: detailed FIR circuit

The proposed circuit is detailed as follows:

- Input $x[i]$: contains the elements resulting from the digitization of the signal with a size of 16 bits
- Buff: Buffer (BRAM) of size (16×8) to store the signal elements
- A ROM of size (16×8) containing the filter coefficients in this case they are constant
- A Pointer (k) which points both to the buffer and to the ROM
- An accumulator S to store the result of the production of each element by the coefficient and which automatically calculates (accumulates) the sum of the products when it is full (8 multiplications) ($Es=1$)
- The result is routed to the circuit output and the control signal RST1 is reset to 0 to process other signal elements

Finite state machine (FSM):



Estat	Eb	Es	Rst1	Ep	Rst2	EO
IDLE	0	0	0	0	0	0
INIT	1	0	1	0	1	0
CMPT	0	1	0	1	0	0
Test	0	0	0	0	0	0
Done	0	0	0	0	0	1

2) Parallel approach:

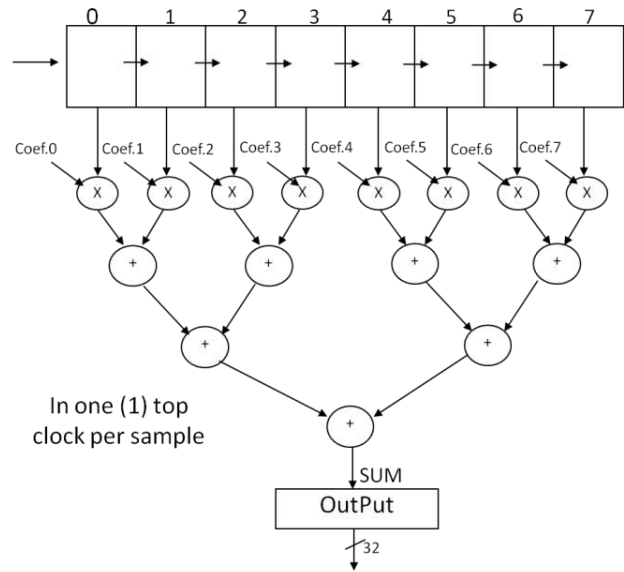


Fig. 12: FIR circuit detailed in Parallel

According to figure 12, the implementation of the FIR filter in parallel (in a case of 8 elements)

Shows a shift buffer where each element is pointed by a coefficient to do the multiplication the result and added “two by two” by an adder at the first level for the four multiplication results then the second level of two elements, the final result is routed to circuit output.

Noted, that at each clock top, an element is loaded into the buffer and when the buffer is full (full buffer=1) the calculation of the result is done by a single clock top.

E. RESULTS AND DISCUSSION

TABLE 2
RESULTS

Implementation	Number of cycles	Gain (Real optimization)
Real code (sequential)	640	$640/640 = 1$
Optimized	320	$640/320 = 2$
Software pipelining	47	$640/47 \approx 13.6$
FPGA sequential	320	$640/320 = 2$
FPGA parallel	40	$640/40 = 16$

TABLE 3
DEVICE UTILIZATION SUMMARY (ESTIMATED VALUES) FOR SEQUENTIAL IMPLEMENTATION FPGA

Logic Utilization	Used	Available	Utilization
Number of Slice Registers	209	54576	0%
Number of Slice LUTs	103	27288	0%
Number of fully used LUT-FF pairs	39	273	14%
Number of bonded IOBs	50	218	22%
Number of BUFG/BUFGCTRLs	1	16	6%

TABLE 4
DEVICE UTILIZATION SUMMARY (ESTIMATED VALUES) FOR PARALLEL
IMPLEMENTATION FPGA

Logic Utilization	Used	Available	Utilization
Number of Slice Registers	66	54576	0%
Number of Slice LUTs	35	27288	0%
Number of fully used LUT-FF pairs	18	83	21%
Number of bonded IOBs	50	218	22%
Number of BUFG/BUFGCTRLs	1	16	6%

REFERENCES

- [1] 1. <https://www.ti.com/lit/ds/symlink/omap-1132.pdf?ts=1648333673253>. 2022.
- [2] 2. Frénéa, M., Notes de cours (version2) Antenne de Bretagne, E.C.A.d. Bretagne, Editor 2013.
- [3] 3. GARNIER, H., Conception de filtres numériques, 2016.
- [4] 4. Ricardo, P., et al., Scalable Sensor Data Processor: Testing and Validation. 2016.
- [5] Seminar Report On The C6000 Family: Architecture, Pipelining and General Trends 2006 – 2007
- [6] 5. Instrument, T., SPRS762E –AUGUST 2011–REVISED JANUARY 2017. 2017.
- [7] 6. <http://www.ti.com/sc/docs/dsp/support.htm>. 2021.

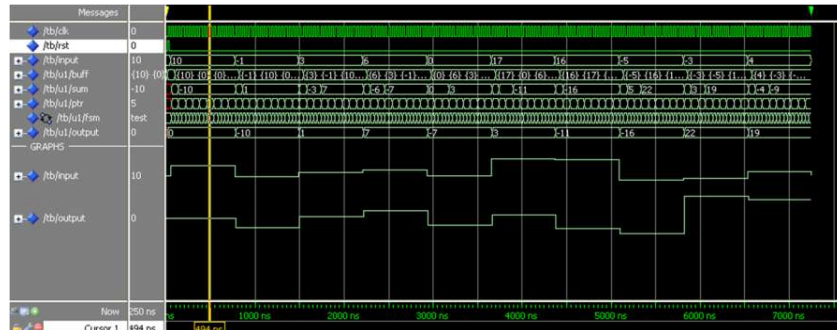
• We note from what has been presented in table 2 that the execution of the FIR algorithm under DSP 6000 gives a satisfactory result, especially after optimization and its specialized instruction set (the MPY) also its super-scalar architecture.

• Under FPGA, although there is a great improvement in parallel design, we note that the consumption of the components in sequential seems high contrary to what was planned (sequential consumes less) this influences the choice of the FPGA type (low or high density).

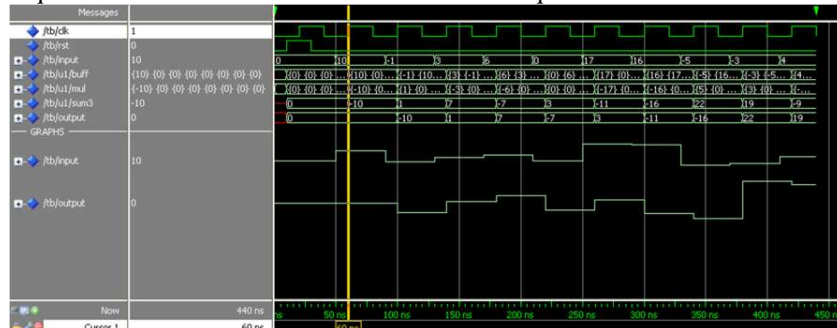
F. HORIZONS:

• We intend to improve this study by relying the most on optimization on the DSP side (other optimization techniques such as: Unrolling loop, Word wide, Software pipeline) and improve the design and implementation on FPGA solutions in parallel order to reduce the number of electronic components and promote its use on low density FPGAs

G. Annexes



Annex 1: screen cut, execution of SOP algorithm on sequential with number of coefficients equal to 8



Annex 2: screen cut, execution of SOP algorithm on sequential with number of coefficients equal to 8

Integrating Reinforcement Learning with Analytic Hierarchy Process for Autonomous Service Web Composition

Selma ZEGGADA

LASTIC laboratory, Department of
Computer Science, University of Batna
2, Pole Fesdis 05110 RN, 05110 Fesdis
Batna, Algeria
s.zeggada@univ-batna2.dz

Sofiane AOUAG

LASTIC laboratory, Department of
Computer Science, University of Batna
2, Pole Fesdis 05110 RN, 05110 Fesdis
Batna, Algeria
sofiane.aouag@gmail.com

Djalal HEDJAZI

LASTIC laboratory, Department of
Computer Science, University of Batna
2, Pole Fesdis 05110 RN, 05110 Fesdis
Batna, Algeria
djalal.hedjazi@gmail.com

Abstract— The service composition process is a highly complex task that presents several problems, mostly associated with continuous development of candidate web services. A large number of existing web services have different Quality of Service (QoS) and often change over time to improve their efficiency. An effective composition must satisfy user requirements. QoS aware service composition has become an important research area to find a near-optimal composition in dynamic environments. To address the aforementioned issues, several existing service composition approaches based on artificial intelligence techniques have emerged, especially in machine learning and deep learning. Various approaches based on reinforcement learning (RL) have substantially contributed to the advancement of the paradigm of service composition process. In this paper, we propose a new autonomous service composition approach based on Multi-Objective Model Decision Process (MOMDP) combining with analytic hierarchy process (AHP) to solve the large-scale adaptive service composition with multiple criteria in uncertain environments.

Keywords— *Service Composition, QoS, Reinforcement Learning, Markov Decision Process, Multi-Objective, Analytic Hierarchy Process*

I. INTRODUCTION

The composition of services implies the ability to select, coordinate, interact, and interoperate existing services. In service computing, the goal of combining multiple services through service composition is to meet complicated user requirements [1]. With the increased development of Web services, many Web services emerge and have the same functional attributes [2]. The research topic led to reinforcement learning (RL). RL is a kind of Machine Learning algorithm in the Artificial Intelligence domain. It uses the observed reward to learn an optimal policy (or approximate optimal policy), to maximise the cumulative reward [3]. RL is a type of trial-and-error learning that has been applied to adaptive service composition scenarios in order to identify the best behaviour approach without prior training samples [4]. In the context of service composition, on the one hand, the environment is constantly changing and some QoS may continuously evolve. On the other hand, there exist increasingly complex composition flows and a huge number of candidate services [1].

To deal with a dynamic environment, many research works introduced RL into the process of service composition and proved that it can optimize the service composition solution and adapt it to the dynamic environment. [5].

The purpose of a multi-objective decision is to select a preferable solution from a collection of choices while

taking into consideration numerous criteria, such as price, quality, appearance, etc.[6]. These criteria can be contradictory (we can hardly maximize the quality and minimize the price at the same time) [7]. To solve this problem, we propose an adaptive service composition based on reinforcement learning which is modeled by a Markov Decision Process (MDP).

Based on many previous studies, several approaches focused on selecting best services among many web services. In this context, multicriteria decision-making (MCDM) techniques are introduced, [8] described how analytic hierarchy process improves (AHP) its efficiency. It is only used during the selection stage, and it can be employed to express the complexity of user preferences and requirements. However, we have found no work using AHP in the whole process of service composition, although many research works have been conducted to satisfy these search strings.

The AHP method's principles are described as follows: (I) Hierarchical structuring in (classes - criteria - weight). (II) Structuring of priorities (under criteria - ranks), (III) Logical consistency and (IV) Semi-quantitative method [5]. The Expert Choice software developed in the US by F. Saaty (1985) is the applied software of the method [8]. In our work, we use the AHP method at selection step to choose a set of appropriate services based on the comparison of the different characteristics, two by two, through its weighting, from the construction of a square matrix [8]. The main goal of using AHP is to reduce the search space, striving for efficient computation. In this work, we propose a multi-level approach to service composition that combines RL and AHP to accelerate web service discovery, selection, and composition.

We propose an approach called MOMDP (Multi-objective Model Decision Process) for selecting an optimal service composition among all possible permutations without knowing the users' preferences in advance. In this paper, we are attempting to enhance the proposed technique for an adaptive composition in an uncertain environment.

The remainder of this work is structured as follows: the next section outlines the background of some recent approaches. In Section 3 we present the problem formulation in detail. Section 4 is dedicated to present the proposed approach. A discussion of the approach in section 5. Finally, the paper is concluded in Section 6.

II. RELATED WORKS

In the literature, a large number of adaptive service composition approaches have been proposed. QoS-aware Web service composition has been widely studied in the areas of Service-Oriented Architecture (SOA) and Service-Oriented Computing (SOC) [9]. Authors in [10] and [11] used Markov Decision Process (MDP) to model service compositions. [2] Presented a mechanism self-adaptivity to enable a service composition to adapt to its environment autonomously by utilizing Reinforcement Learning in dynamic environments.

A typical technology is used by [9] for planning and Optimization that proposed a novel multi-objective approach to handle QoS-aware Web service composition. The proposed approach uses conflicting objectives and various restrictions on quality matrices. Reinforcement learning is employed to deal with the uncertainty characteristic inherent in open and decentralized environments. Gao et al. [3] presented a fast learning method, which is an uncertainty planning method, it is used to compose web services. Invocations of web services are viewed as uncertain actions, while service quality is viewed as partially observable variables.

The main contribution of the work of [6] is to propose the first approach in the literature able to provide a web services composition that explicitly takes into account the users' expectations and preferences on the QoS. Agent technology has gained increasing popularity in service oriented architecture (SOA); A learning model of the service-oriented MAS for the service composition problem is shown in the paper of [12]. It adopts the principle of reinforcement learning and is based on the Markov game and Q-learning. In the same context, [13] suggested a multi-agent Q-learning algorithm, in which each agent would benefit from the advice of other agents in the team.

In recent years, extensive research efforts have been spent on the development and the optimization of service compositions, Wang et al.[14] combined RL with skyline computing where the latter is used to reduce the search space and computational complexity. Authors in [15] defines a complete calculation method for QoS (Quality of Service) of composite services, and uses as a measure of the quality of composite services, by combining Skyline with Ant Colony Algorithm (ACO). Liu et al. [5] proposed an adaptive service composition based on deep reinforcement learning, where recurrent neural network (RNN) is utilized for predicting the objective function.

The work of authors in [8] provides an in-depth investigation of the AHP-based Web Service selection methods presented in the literature (2020). The authors prove that the AHP is another open area that can be focused on in the future. In general, RL is an effective method to solve adaptive service composition. However, this method will become inefficient when facing a large number of candidate services. Our proposed architecture stands out from previous work, by using in the first hand the AHP method to minimize the set of services candidates belonging to the Pareto optimal set and then using Q-learning to find optimal or near optimal service composition.

III. PROBLEM FORMULATION

In general, Markov Decision Process (MDP) represents the environment in Reinforcement Learning (RL). In this section, we describe the theoretical model based on the process of dynamic service composition. Our proposition uses Markov Decision Process (MDP) to define elements for controlling uncertain events. An MDP is a framework that involves creating mathematical formulas and models for sequential decision-making [3] In fact, the process of composition must satisfy user preferences. In our case, QoS-aware service composition is a multi-objective optimization problem. Thus, the Multi-objective Markov Decisional Process (MOMDP) will be the model of our composition approach. We first define the key concepts of the model.

A. Definition 1 (Web Service)

A web service WS is modeled in [14] as a 6-tuple $WS = \langle ID, In, Out, Pr, E, QoS \rangle$, where

- ID : is the identifier of the web service.
- In : is the input of Web service.
- Out : is the output of Web service.
- Pr : is the precondition state that guarantees the successful invocation of the Web service
- E : is the effect to the environment after invoking and running the Web service
- QoS : is an n-tuple $\langle C_1, C_2, \dots, C_n \rangle$, where each C_i denotes a QoS attribute (e.x., Cost, Price, Reliability, Availability, and Response Time) of a WS.

More precisely, we will start by describing the formalism of multi-objective MDP. Therefore, MOMDP is an extension of MDP [13], where decision-making depends on several criteria. These criteria are introduced to the decision process through the reward function, which implies repercussions on the rest of the formalism [11].

B. Definition 2 (Multi-Objective Markov Decision Process (MOMDP)).

Define An MOMDP is defined where an action a_j from a set of actions $A = \{a_1, a_2, \dots, a_m\}$ acts on a number of criteria modeled as a vector criteria $Z = \{z_1, z_2, \dots, z_n\}$. There are ($C > 1$) objectives that the agent wants to achieve, and then it gains the following reward vector from the environment when it moves to the next state [13]. It is therefore appropriate to modify the definition of MDP in order to include the consideration of these criteria, thus, a multi-criterion MDP is defined as:

- a set of states S ,
- a set of actions A ,
- a function of transition $p(s, a, s')$, $s, s' \in S, a \in A$,
- a function of reward $r(s) = \{r_1(s), r_2(s), \dots, r_i(s), \dots, r_n(s)\}$ where each $r_i(s)$ represents the reward obtained in state s for criteria i (in the same way as in single-criteria MDPs).

In the multi-objective case, the reward function returns a vector. More formally,

$$V(S) = \begin{pmatrix} v1(s) \\ v2(s) \\ \vdots \\ vn(s) \end{pmatrix} = \begin{pmatrix} r1(s) \\ r2(s) \\ \vdots \\ rn(s) \end{pmatrix} + \max_a \sum_{s'} p(s, a, s') * \begin{pmatrix} v1(s') \\ v2(s') \\ \vdots \\ vn(s') \end{pmatrix} \quad (1)$$

Where each $v_i, i \in 1 \dots n$ are the values of the different criteria. In the general case, the direct application of the max operator is impossible. The same action does not maximize all the criteria. We are looking here to get a satisfactory policy and not all policies.

C. Definition 3 (MOMDP-Based Web Service Composition (MOMDPWSC)).

An MOMDP-WSC [9] is defined as a 6-tuple: $MOMDP - WSC = \langle S, s_0, S_t, A, P, R \rangle$, where

- S is a finite set of world states.
- $s_0 \in S$ is the initial state that every service composition starts from this state.
- S_t is a set of terminal states, and the service composition terminates.
- $A(s)$ is a set of Web services ws that can be invoked in state $s \in S$, and the precondition must be satisfied by s
- P is a probability that a Web service $ws \in A(s)$ is invoked and a transition is occurs from its current state s to a resulting state s' , where the output of ws is satisfied; the probability of transition denoted as: $P(s'|s, ws)$.
- R is a reward function when a Web service $ws \in A(s)$ is invoked, and the service consumer receives an immediate reward $r = R(s'|s, ws)$.

The selection of web service ws is established with multiple QoS criteria, then the agent receives the following reward vector:

$$Q(s, ws, s') = [Q_1(s, ws, s'), Q_2(s, ws, s'), \dots, Q_M(s, ws, s')] \quad (2)$$

Where each Q_i denotes a QoS attribute of ws .

A MOMDP involves multiple actions and paths for an agent to choose [16]. The solution to an MOMDP-WSC is a decision policy π . A policy is a function or a decision-making process that allows transitions from one state to another [5], which is defined in the process of composition as a service selection $ws \in A$ by agent i in each state s . These policies are actually mappings from states to actions, defined as follow:

$$\pi : S \rightarrow A. \quad (3)$$

IV. PROPOSED APPROACH

After modeling the services' composition environment as a MOMDP-WSC, we are interested in finding a solution for our model. The fully automated composition takes care of the whole composition process and performs it automatically [17], with that, no user intervention is required. Knowing that the composition is a relatively complex process, it goes through several steps. Therefore, any solution proposed for the service composition process must follow the life cycle and the sequence of steps of a composition.

To respond to the requirements above, our proposal is the development of a Service Oriented Architecture (SOA) solution to effectively process the composition of services. Figure 1. Describes the architecture of the proposed approach based on SOA.

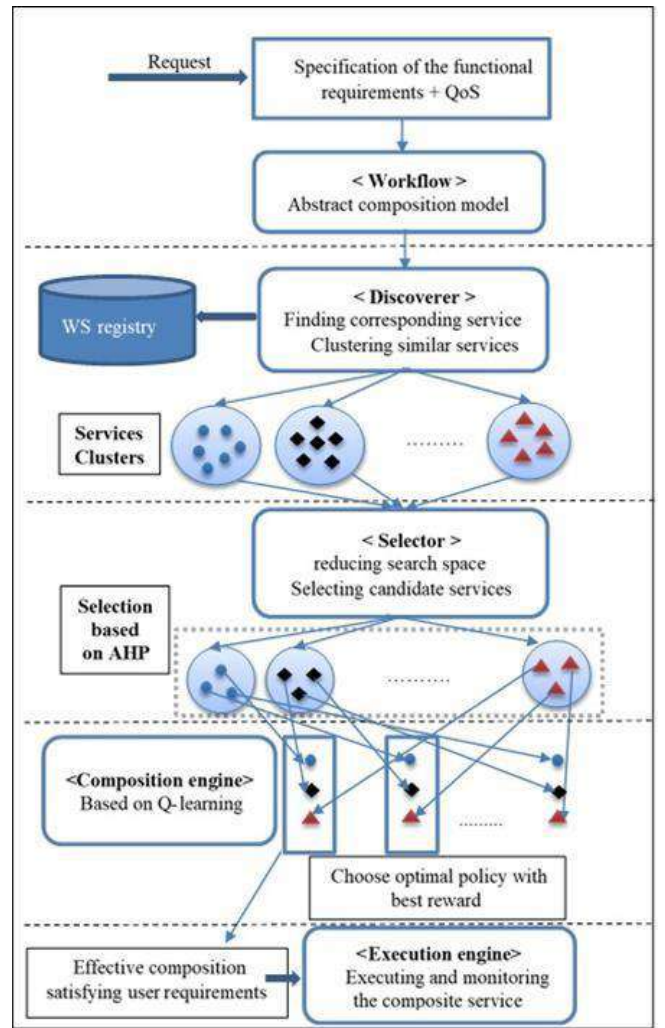


Figure 1: Architecture of the proposed approach based on SOA

A. Integrating reinforcement learning and AHP to solve the model

Based on many previous studies, we propose an approach for multi-objective service composition in dynamic uncertain environments. We combine Q-learning with AHP in order to find an optimal composition. The sequence of our proposed architecture is structured as follows:

1) *Workflow*: After the specification of the functional requirements and QoS criteria, workflow is defined as an abstract model, in order to define the nature of the elements that will be used in the composition.

2) *The Discoverer*: The role of the discoverer is to find the corresponding web services for each of the participating tasks. i.e. the concrete web services which correspond to the different tasks are localized. All the services found for each task will be organized in a Service Cluster (SC). It contains similar web services which performing the same single task, and though they may differ in their quality attributes.

3) *The Selector*: A Web service has more than one QoS attributes and the range of various attributes are not the same, which makes them not directly comparable [6]. Therefore, there is a need for a decision process to choose the most suitable web service among a set of candidates. The main issue of the selector is the comparisons of the several candidate services by evaluating multiple criteria and speeding up the convergence rate of the algorithm [10].

In general, most of the proposed methods in the literature do not reduce the number of candidate services selected in the step of discovery; which makes the process of learning take more time, therefore, for this reason, we propose to use AHP method to limit a large number of candidate services and to reduce the search space. The selector aims at applying the analytic hierarchy process (AHP) on the service cluster (SC) appropriate to solve multicriteria decision problems, based on the prioritization of criteria for comprising several solutions satisfying a set of criteria [12]. Figure 2 shows the step of the hierarchical structuring in AHP method.

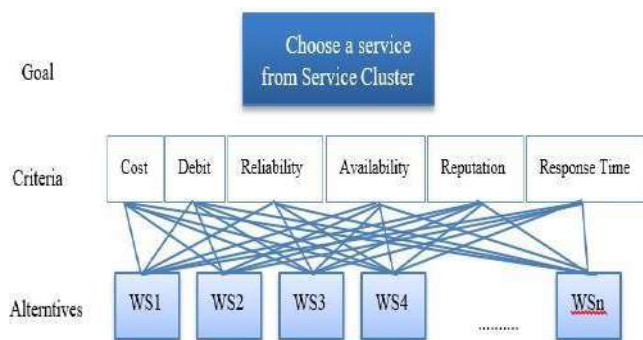


Figure 2: Hierarchical structuring process of AHP method

Once the hierarchy is built, the selector systematically evaluates its different elements by comparing them to each other [8].

4) *Composition Engine*: After finding the candidate services for all the tasks and linking each task to its chosen web service, the concrete composite service is created. Our model of composition is based on Q-learning which is a technique of RL. It does not require any initial model of the environment. Q is the function that measures the quality of an action performed in a given state of the system [14]. In what follows, we present an algorithm of the composition

engine based on Q-learning in order to calculate the optimal policy. The Q-learning algorithm is an effective RL method. In this paper, we rely on the original work [14], [5] that utilizes Q-learning to solve service composition problems. So, the user satisfaction can be used to describe the reward. QoS is generally used for modeling user satisfaction. The algorithm uses the ϵ -greedy strategy to choose an action for getting a reward [17]

Algorithm: Reinforcement Learning for Autonomous Service Composition with multiple objective

```

1: Initialization
2:  $\forall s, \forall a$  is initialized arbitrarily
3: Repeat
4: For each episode  $ep$  do
5: Initialize state  $s$ 
6: repeat
7: for each step of episode  $ep$  do
8: Select a service from service cluster ( $sc$ )
9: choose  $a \in A(s)$  based on  $\epsilon$ -greedy policy
10: execute service  $a$ , observe reward  $r$  and state  $s'$ 
11:  $Q(s, a) \leftarrow Q(s, a) + \alpha [r + \gamma \max_{a'} Q(s', a') - Q(s, a)]$ 
12:  $s \leftarrow s'$ 
13: end for
14: until state  $s$  reach one terminal state
15: end for
16: until convergence condition is satisfied, algorithm converges

```

5) *Execution engine*: During the run step, the execution engine creates an instance of the process by running the composite service. At this stage, we can define a model for handling exceptions that defines how the exceptional situations that produce the execution of a composite service can be handled, without leading to the termination of the composite service

V. RESULTS AND DISCUSSIONS

As defined in section 4, the architecture of the hybrid proposed approach is based on SOA. At first, an AHP method is introduced in the selection stage to facilitate the efficiency of the composition algorithm and reduce the search space. The main objective of adopting the pre-selection paradigm refers to the large existing number of services. In May 2021, *programmableweb.com*¹ declared over 24.207 APIs in more than 60 fields. AHP improves its efficiency in many recent studies cited in section 2. The core of our framework is the composition engine; it is based on one of the RL algorithms.

The present service composition test employs test data created automatically by the AnyLogic² tool for simulation and testing. So we have applied our algorithm written in Python to compare the different configurations according to the obtained results. This experiment is carried out to verify and evaluate the efficacy of our algorithm. So, we've evaluated the models generated against different

¹ <https://www.programmableweb.com/category>

² <https://www.anylogic.com/> (January 2022)

objective criteria. At each iteration, the AVG time taken to locate randomly selected elements from randomly generated models.

The order preparation of products stored in different depots requires optimizing management of stock. The main target is to optimize delivery date with minimizing costs. This problem is described by stating four objectives:

- **Optimized transport fee:** this is dependent on how product is dispatching in different depots.
- **Minimizing waiting time:** an example can be given as follow: if Depot N°1 has 20 pallets of required products, 10 of which are reserved to be delivered July 05th. Knowing that new supply would arrive on July 1st, a new reservation is made for June 28th. We can confirm the reservation and guarantee that the July 05 reservation will take place on the newly delivered pallets.
- **Optimize the pallet selection based on trailer size:** type of pallet, for example, Stackable pallets feature bottom support and may be stacked on top of another loaded pallet to provide stability during storage and transit. The charging time is variable.
- **Optimize the reservation** in accordance with stock motion planned and possible cancellation of reservation.

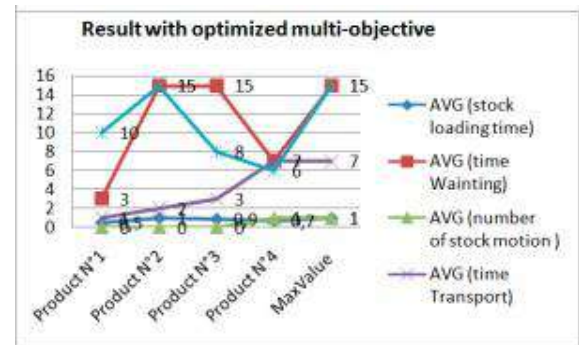


Figure 3: Primary results obtained by application of different parameters

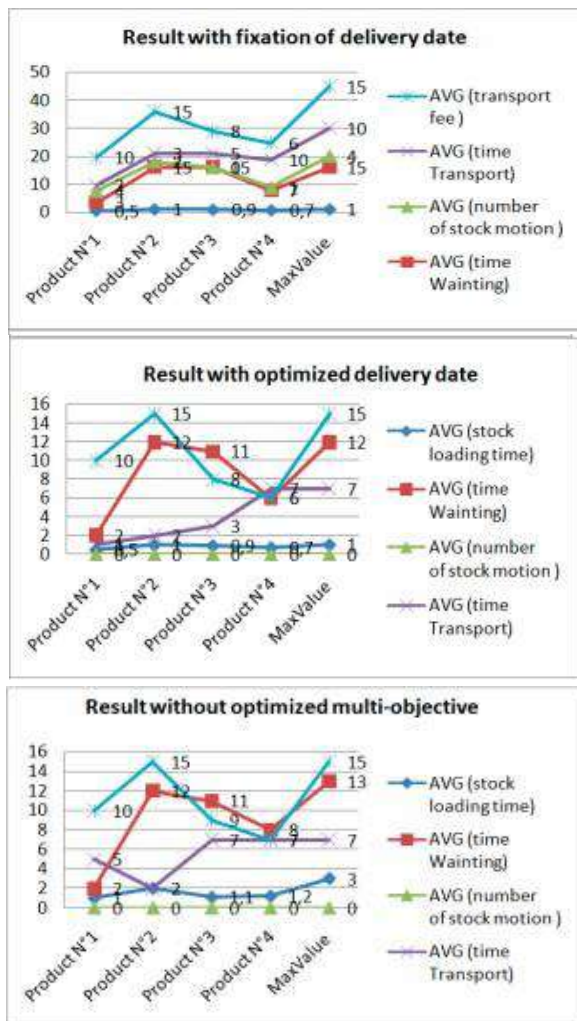
A. Result Analysis

At each iteration, we determined the maximum value acquired from the created model's substantial cost. Besides the fact that our method runs automatically, it is nevertheless adaptable based on the preconditions that may be specified.

When we add max (delivery date) = X as parameter, the results are affected by any canceled reservation. For example, if a reservation for July 1st is canceled, this may result in the cancellation of stock movements, which further optimizes transport costs by canceling scheduled stock movements. It should be noted that our simulation does not address the issue of manufacturing demand; however, we will address this in the future as we enhance our system. The application of our algorithm shows clearly the optimization of the results obtained compared to those given by the simulator (max value is greater than our calculated max time value) by optimizing stock movements in order to reduce the cost of transport. As defined in section 4, the architecture of the hybrid proposed approach is based on SOA. Q-learning has much better usage of web services in all cases in terms of obtaining the optimal solution.

CONCLUSION

We consider our work as an intelligent web assistant; we have described new autonomous systems for service composition based on Multi-Objective Model Decision Process (MOMDP) combining with Analytic Hierarchy Process (AHP) to solve the large-scale adaptive service composition with multiple criteria in uncertain environments. A WSC-MDP model is designed for the proposed framework, which can also deal with a dynamically changing environment. In the literature, most of the works do not use the reduction of the research space. We propose an SOA architecture from the discovery of the service to the effective composition which is not sufficiently studied; they focused on one stage of the process, especially the selection stage. We actually try to improve the efficiency of the proposed method in next regular paper. One of the important directions that we plan to further explore in future research is the use of multi-agent technology to integrate with AHP and deep RL-learning.



REFERENCES

- [1] Doshi, P. et al. (2004) 'Dynamic Workflow Composition using Markov Decision Processes', p. 7.
- [2] Alizadeh, P. et al. (2011) 'Reinforcement Learning for Interactive QoS-Aware Services Composition', IEEE SYSTEMS JOURNAL, p. 11.
- [3] Gao, A. et al. (2005) 'Web Service Composition Using Markov Decision Processes', in Fan, W., Wu, Z., and Yang, J. (eds) *Advances in Web-Age Information Management*. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), pp. 308–319. doi:10.1007/11563952_28.
- [4] Chang, C. W. (2007). *Intelligent Web Services Selection based on AHP and Wiki*. IEEE/WIC/ACM International Conference on Web Intelligence.
- [5] Liu, J.-W. et al. (2019) 'Large-scale and adaptive service composition based on deep reinforcement learning', *Journal of Visual Communication and Image Representation*, 65, p. 102687. doi:10.1016/j.jvcir.2019.102687.
- [6] Jatoth, C. et al. (2019) 'SELCLLOUD: a hybrid multi-criteria decision-making model for selection of cloud services', *Soft Computing*, 23(13), pp. 4701–4715. doi:10.1007/s00500-018-3120-2.
- [7] Lei, Y. et al. (2019) 'LEARNING-BASED WEB SERVICE COMPOSITION IN UNCERTAIN ENVIRONMENTS', p. 19.
- [8] Masdari, M. (2020) 'Service selection using fuzzy multi-criteria decision making: a comprehensive review', p. 32.
- [9] Moustafa, A. and Zhang, M. (2013) 'Multi-Objective Service Composition Using Reinforcement Learning', p. 15.
- [10] Ouadah, A. et al. (2015) 'Combining skyline and multi-criteria decision methods to enhance Web services selection', in 2015 12th International Symposium on Programming and Systems (ISPS). 2015 2th International Symposium on Programming and Systems (ISPS), Algiers, Algeria: IEEE, pp. 1–8. doi:10.1109/ISPS.2015.7244975.
- [11] Uc-Cetina, V., Moo-Mena, F. and Hernandez-Ucan, R. (2015) 'Composition of Web Services Using Markov Decision Processes and Dynamic Programming', *The Scientific World Journal*, 2015, pp. 1–9. doi:10.1155/2015/545308.
- [12] Viriyasitavat, W. (2016) 'Multi-criteria selection for services selection in service workflow', p. 6.
- [13] Wang, H., Wang, X. and Zhou, X. (2012) 'A Multi-agent Reinforcement Learning Model for Service Composition', p. 2. IEEE Ninth International Conference on Services Computing.
- [14] Wang, H. (2020) 'Integrating reinforcement learning and skyline computing for adaptive service composition', *Information Sciences*, p. 20.
- [15] Wang, C., Zhang, X. and Chu, D. (2020) 'Research on Service Composition Optimization Method Based on Composite Services QoS', p. 5.
- [16] Yu, L. W.-L.-M.-S.-T. (2014). *Learning-Based We Service Composition in Uncertain Environment*. *Journal of Web Engineering*, 450-468.
- [17] Wang, H. et al. (2020) 'Adaptive Service Composition Based on Reinforcement Learning', p. 16.
- [18] Boussard, M. B. (2008). *Processus Décisionnels de Markov en Intelligence Artificielle*. Groupe PDMIA.
- [19] Tour, A. A. (2016). *WEB SERVICE COMPOSITION PROCESSES: A COMPARATIVE STUDY*. *International Journal on Web Service Computing (IJWSC)*, p. International Journal on Web Service Computing (IJWSC).
- [20] Vakili, A. J. (2017). *Comprehensive and systematic review of the service composition mechanisms in the cloud environments*. *Journal of Network and Computer Applications*, 24–36.
- [21] Wang, H. Z. (2010). *Adaptive Service Composition Based on Reinforcement Learning*. *ICSOC 2010, LNCS 6470*, 92–107.
- [22] Xu, W. C. (2012). *A Multi-Agent Learning Model for Service Composition*. *IEEE Asia-Pacific Services Computing Conference*.

Intelligent Control of Efficient Bridgeless SEPIC PFC converter Powered BLDC Motor

Salah Eddine HALLEDJ

*Department of Electrical, Electronic
and Automation engineering
National Polytechnic School
Constantine, Algeria
eddinesalah031@gmail.com*

Amar BOUAFASSA

*Department of Electrical, Electronic
and Automation engineering
National Polytechnic School
Constantine, Algeria
amar.bouafassa@gmail.com*

Abstract—This paper proposes an intelligent control of efficient single stage bridgeless single ended primary inductor converter (SEPIC) to achieve an high power indices while powering brushless direct current (BLDC) motor. The proposed AC/DC PFC converter can generate unity power factor (PF) and low voltage stress across the switching device. The overall system is composed of two stages: (1) AC/DC SEPIC PFC system, which is regulated by a cascaded voltage current loop controllers. For voltage loop, it represents either PI or fuzzy logic controller. For current loop is designed by non-linear constant band hysteresis control; (2) a DC/AC voltage source inverter derived BLDC motor. The performances of the proposed application are tested and verified based on power quality indicators. The effectiveness of the controllers are evaluated via SIMULINK/MATLAB environment. The obtained results demonstrate that the present control achieve unity power factor. In addition, it has a significant enhancement of both DC link voltage and BLDC motor's speed performance.

Index Terms—Bridgeless SEPIC, BLDC motor, Fuzzy logic control, hysteresis current control, PI control, Power factor correction.

I. INTRODUCTION

Nowadays, a permanent synchronous brushless DC motor becomes a fundamental device in numerous applications including marine defense [1], robotics [2], pumping systems [3] and domestic appliances. Compared to asynchronous motors, a BLDC motor has a couple of preferable advantages including high efficiency and reliability, long life, fast rate of voltage and current, noiseless [4] and easy to control [5].

Control design of BLDC motor usually uses three common regulation loops, first loop is for PWM or hysteresis current controller [6], [7], second is for speed regulation such as PID and fuzzy logic. A traditional PID can provide an accepted performance but it always suffers at large external disturbances and parameter variations, which the results will be erroneous. Fuzzy control offers an ameliorated speed performance but the robustness is still not touched the requirement range. As solutions of these drawbacks a nonlinear robust controllers have been designed to ensure robustness and smoothness of the speed performance. The third control loop is for electronic commutator device that is used to fill up the gap generated by separated mechanical commutators, three hall sensors or optical encoders are used in this loop control.

Recently, BLDC motors often work with a cascaded PFC converter and VSI system, this PFC converters play a key role in eliminating power loss, improving DC link voltage quality and increasing power quality signals such as power factor, efficiency and total harmonic distortion values (THD) to meet the international electronic commission standards (IEC 61000-3-2) [8].

Generally, the connected full bridge rectifier in front of PFC converter creates many problems in power quality indices as high conduction loss and heat dissipation, especially at low voltage. Hence, the total efficiency decreases. Due to this, diode rectifier becomes unnecessary large and expansive circuitry [9]. To address these drawbacks, a bridgeless PFC converter was suggested and proposed rather than conventional bridge PFC converters [10]. In the process of improving the dynamic performance of AC/DC PFC converters, various bridge/bridgeless topologies were studied in literature. PFC boost converter is one of the most known AC/DC device because of its simplicity and low cost. However, it has a couple of restrictions, such as low efficiency, high voltage stress across the switching device, low efficiency and high inrush current [11]. To avoid the aforementioned drawbacks, a novel bridgeless topologies are designed as bridgeless boost [12] and SEPIC [13]. They provide numerous convenient for AC/DC applications because they successfully reduce THD and restrict the inrush current at starting operations, but it is costly due to the extra passive components. In [9] a modified bridgeless SEPIC were designed and proposed to overcome the previous disadvantages, this PFC device has a high power factor and low stress voltage across the main MOSFET switch. However, it was designed to work in continuous conduction mode. Hence, the control circuit will be complicated to achieve an effective power quality. The authors in [14] propose an efficient bridgeless SEPIC PFC converter; this one has high efficiency with simple control circuit because it is designed to operate in DCM.

Moreover, it is befitting to low cost converter for high

voltage systems [15].

The basic purpose of this article is to use a intelligent cascaded based voltage/current loops to control the proposed efficient bridgeless SEPIC [15] derived BLDC motor in order to improve power factor, DC link voltage as well as speed performance. For voltage loop, PI or fuzzy logic controller are developed to predict the reference current of PFC converter, and a fixed band nonlinear hysteresis controller is designed for current loop to find the appropriate control law.

The remaining parts of this article are arranged as follow. The system description is presented in section 2. In section 3 a control design of the proposed application is discussed. The results/discussion and conclusion are given in section 4 and 5 respectively.

II. THE PROPOSED SYSTEM'S OVERVIEW

The overall system consists of two main blocks, the first one presents efficient bridgeless SEPIC PFC device with its controller and second refers to BLDC motor with its electronic commutator device as illustrated in Fig 1.

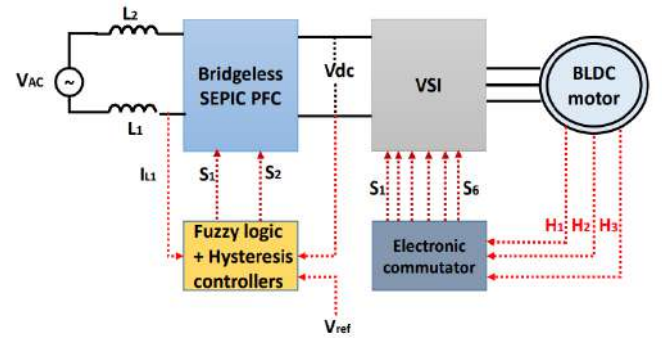


Fig. 1: Overall description of proposed system

A. Proposed SEPIC power factor corrector

The suggested PFC rectifier contains of two power MOSFET switching devices that are activated using the same trigger signal as shown in Fig 2. When MOSFET

switches are OFF the capacitor C_1 charges while switches ON the capacitor C_1 discharges energy to the load.

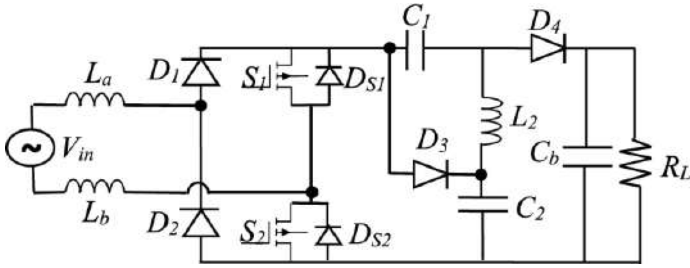


Fig. 2: Efficient bridgeless SEPIC PFC converter

1) *Mode analysis* : The proposed PFC converter has three main operation stages as described in Fig 3. In order to simplify the analysis $L_1 = L_a + L_b$. In order to achieve high power factor, the inductor current L_1 should operate in DCM. Hence, Inductor L_1 will be in phase with source voltage.

Mode 01 $[t_0, t_1]$: During this time, both s_1 and S_2 are activated at the same time as illustrated in Fig 3a. The current i_{L1} rises with rating of $\frac{V_{in}}{L_1}$, while the output inductor current i_{L2} decreases with rating of $-\frac{V_{c2}-V_{c1}}{L_2}$ as described in (1) and (2):

$$i_{L1}(t) = i_{fw}(t) + \frac{V_{in}}{L_1}(t - t_0) \quad (1)$$

$$i_{L2}(t) = i_{fw}(t) - \frac{V_{c2} - V_{c1}}{L_2}(t - t_0) \quad (2)$$

Such that i_{fw} refers to freewheeling current.

Subtracting (1) and (2) gives expression of switch current $i_{s1}(t)$:

$$i_{s1}(t) = \left(\frac{V_{in}}{L_1} + \frac{V_{c2} - V_{c1}}{L_2} \right)(t - t_0) \quad (3)$$

Mode 02 $[t_1, t_2]$: In this mode, both S_1 and S_2 are disactivated and diode D_3 is activated as depicted in Fig 3b. The voltage across main switch S_1 is clamped to V_{c2} and input current i_{L1} drops by $\frac{V_{c2}-V_{c1}}{L_1}(t-t_1)$ from its initial value as:

$$i_{L1}(t) = i_{L1}(t) - \frac{V_{c2} - V_{c1}}{L_1}(t - t_1) \quad (4)$$

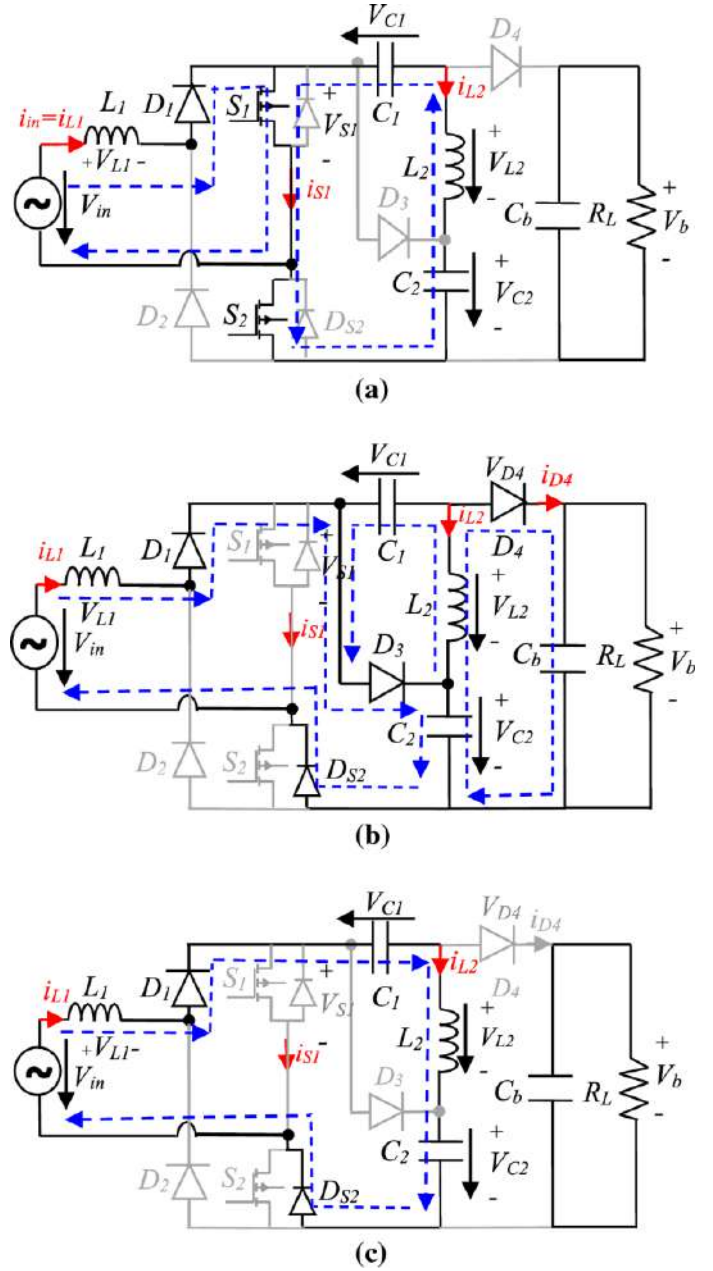


Fig. 3: Operation stages of SEPIC converter

the output inductor current $i_{L2}(t)$ rises with value of $\frac{V_{c1}}{L_2}(t - t_1)$ as:

$$i_{L2}(t) = i_{L2}(t) + \frac{V_{c1}}{L_2}(t - t_1) \quad (5)$$

During this mode, δ_1 is a ratio of sub-interval t_1, t_2 to

switching time T_s and it can be expressed as:

$$\delta_1 = \frac{2V_{in}D}{V_b - V_{in}} \quad (6)$$

Where: $V_b = 2V_{c2} - V_{in}$

Mode 03 [t_2, t_3]: During this mode as shown in Fig 3c, D3 and D4 are disactivated. at t_3 the current i_{L1} and i_{L2} arrive at i_{fw} . this stage keeps unchanged over T_s until the start of new one.

B. BLDC motor drive

Generally, VSIs are utilized as motor drivers. Switching logic signals are generated through the electronic commutator and values are logically based on rotor position. The hall sensor provides a logic signals called hall-effect signals, these signals are converted into six pulses in order to control the IGBTs of VSI [16]. Table 1 describes all the states of these signals in details.

TABLE I: BLDC motor control

position	Hall sensor			switching logic					
	H1	H2	H3	S1	S2	S3	S4	S5	S6
0-60	1	0	1	1	0	0	1	0	0
60-120	0	0	1	1	0	0	0	0	1
120-180	0	1	1	0	0	1	0	0	1
180-240	0	1	0	0	1	1	0	0	0
240-300	1	1	0	0	1	0	0	1	0
300-360	1	0	0	0	0	0	1	1	0

III. CONTROL DESIGN DESCRIPTION

In this work, a voltage/current loops regulation is applied to control the proposed bridgeless PFC converter. PI or fuzzy logic are used for estimating intermediate voltage as shown in Fig 4 while hysteresis controller is utilized as current loop, forcing the output voltage to track the reference voltage.

A. Fuzzy logic controller

Recently, FLCs have become popular solution to control a complicated systems due to its simplicity and no requirements of mathematical information. FLC was inspired from

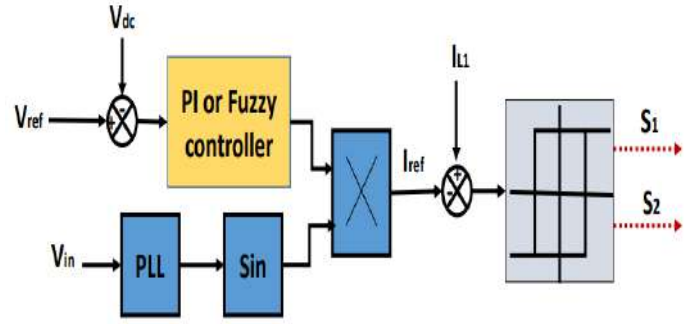


Fig. 4: illustrative schema of Controller paradigm

human thinking in the basis of linguistic variables which are near to human language, these variables are operated via rules that determine a specified knowledge of system's behavior. In this work, FLC has been designed to maintain output voltage constant around the desired value.

The control schema is shown in Fig 5. A 25 rules are used to reach the optimal performance of DCM bridgeless SEPIC PFC device, as illustrated in Fig 6 and Table 2. Fig 7 illustrates the 3D mesh plot of I/O interaction.

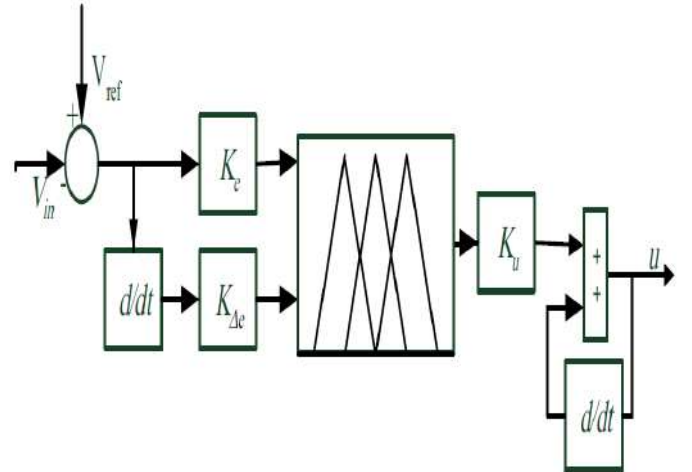


Fig. 5: FLC control schema [17]

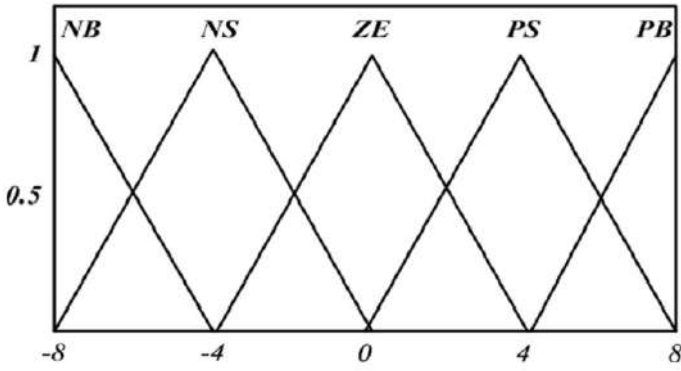


Fig. 6: I/O membership functions

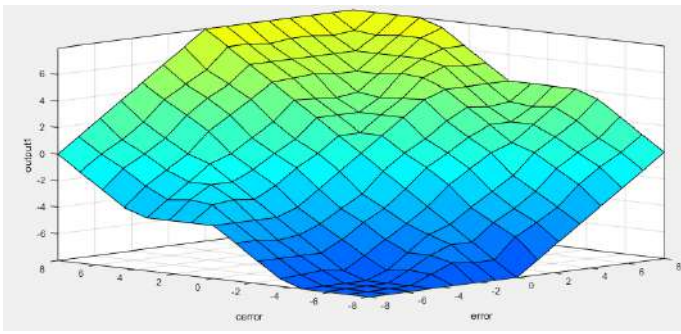


Fig. 7: 3D mesh of the proposed FLC

B. Hysteresis current controller

Hysteresis current control is one of known nonlinear techniques due to its simplicity and robustness. This method forces the current of PFC converter to be within a specific band as illustrated in Fig 8.

$$\begin{aligned} i_{ref} - i &= \frac{\Delta i}{2} \rightarrow s = 1 \\ i_{ref} - i &= -\frac{\Delta i}{2} \rightarrow s = 0 \end{aligned} \quad (7)$$

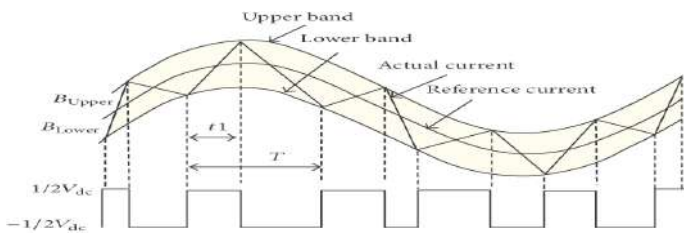


Fig. 8: illustrative schema of HCC

TABLE II: Fuzzy rules

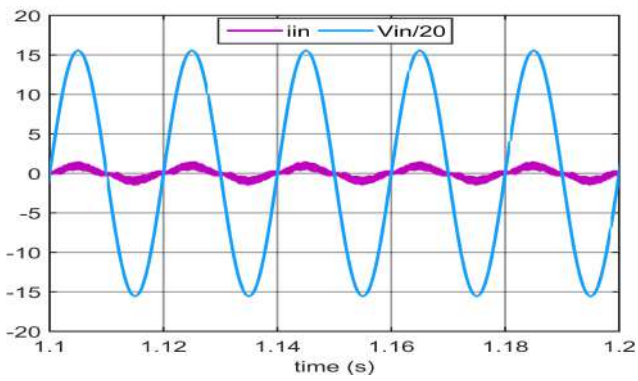
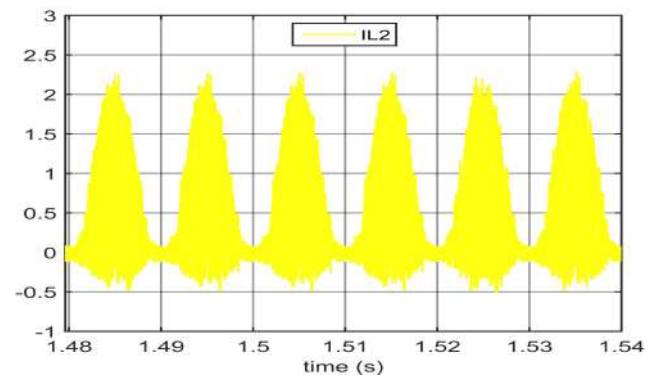
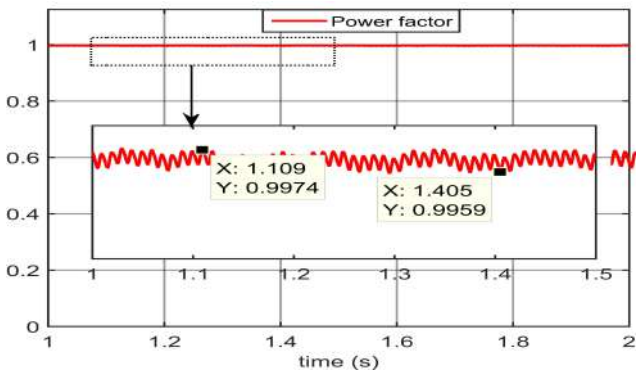
error	change of error				
	NB	NS	ZE	PS	PB
NB	NB	NB	NS	NS	ZE
NS	NB	NB	NS	ZE	PS
ZE	NB	NS	ZE	PS	PB
PS	NS	ZE	PS	PS	PB
PB	ZE	PS	PS	PB	PB

IV. RESULTS AND DISCUSSION

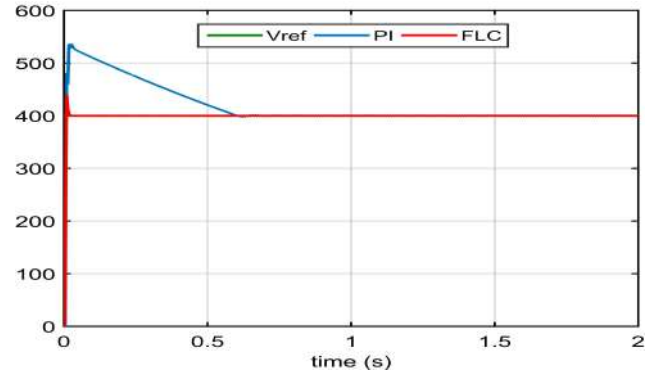
The proposed application as shown in Fig 1 is developed and implemented in MATLAB/SIMULINK environment. Response of DCM bridgeless PFC SEPIC powered BLDC motor is tested by (PI or fuzzy)+ hysteresis controllers. All parameter specifications of bridgeless PFC SEPIC is given in Appendix A.

A. Performance of bridgeless SEPIC PFC powered BLDC motor

Figures 9a, 9b, 9c, 9d depicts respectively the waveforms; input current/voltage I_{in}, V_{in} , inductor current I_{L2} , power factor and output voltage V_{DC} for the proposed efficient bridgeless SEPIC PFC converter when it drives Brushless permanent synchronous motor. As depicted in Fig. 9a, inductor current $I_{La,b}$ is in phase with supply voltage V_{in} , which results a unity power factor, this power factor is bounded between 0.9959 as minimum value and 0.9974 as maximum power factor (Fig. 9c). The inductor current I_{L2} has 2.5(A) as peak value and operates under DCM mode which also helps the system to reach high power factor as presented in Fig. 9b. Intending to enhance output voltage of proposed PFC modified bridgeless SEPIC converter, an intelligent control based fuzzy logic is suggested. As illustrated in Fig. 9d, fuzzy logic controller has been successfully improved the DC-link voltage when it is compared with conventional PI controller. The settling time is enhanced to 0.04 (s) when it was 0.6(s). Moreover, the overshoot is also attenuated from 34% to 10%.


(a) Input current I_{La} , Input voltage V_{in}

(b) Inductor current I_{L2}


(c) Power factor



(d) DC-link voltage

Fig. 9: Performance of Proposed PFC SEPIC fed BLDC motor

B. BLDC motor performance

As it can be observed, Fuzzy logic controller has been successfully improved the dynamic response of BLDC motor (speed, torque, current).

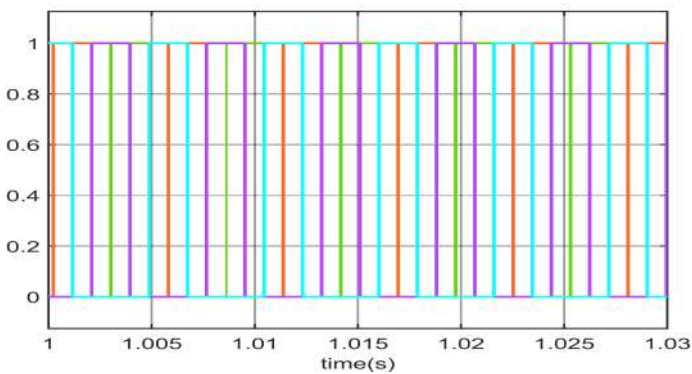


Fig. 10: Switching signals of VSI

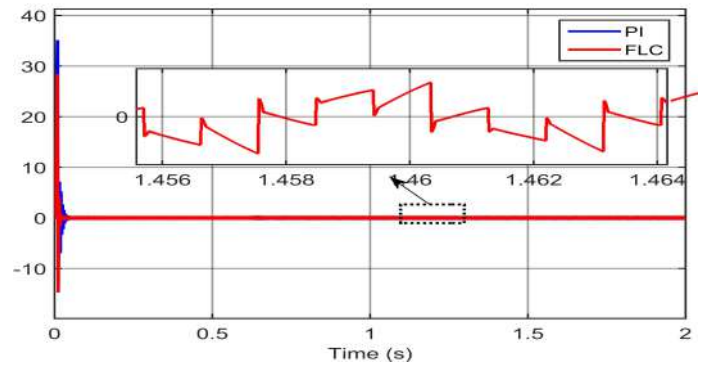


Fig. 11: Stator current (A)

Based on switching gates control (Fig. 10), the speed of BLDC motor ameliorated. FLC has ability of eliminating speed overshoot (0%) and speeding up the transient response (settling time equals to 0.05 s).

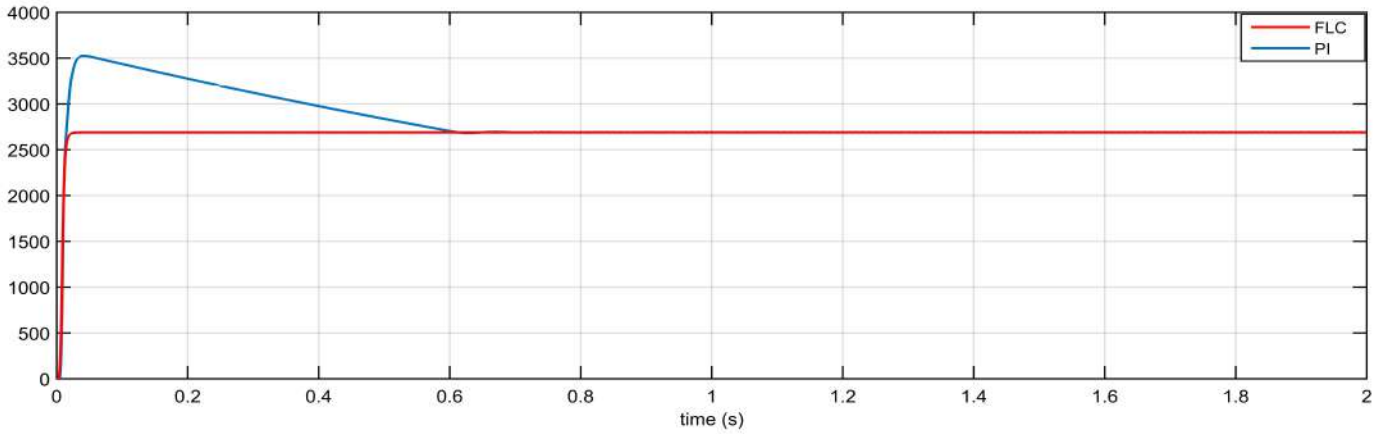


Fig. 12: Speed performance of BLDC motor (rpm)

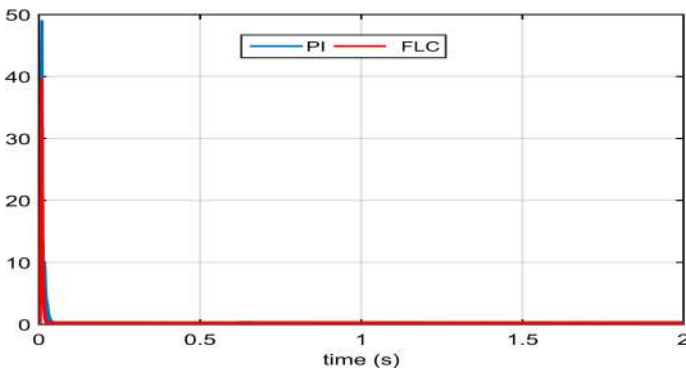


Fig. 13: Electromagnetic torque (N.m)

Furthermore, the proposed controller improved both the stator current (Fig. 11) which reduced peak value to 28 (A), and electromagnetic torque (Fig. 13) with peak value of 38 (N.m).

V. CONCLUSION

This work presented design and simulation of efficient bridgeless PFC SEPIC converter fed BLDC motor. The studied system was controlled via cascaded current/voltage controller, hysteresis regulation was designed for current loop and fuzzy logic controller was developed for voltage loop. this control schema was applied in order to ameliorate the dynamic response of BLDC motor. Moreover, DCM bridgeless PFC SEPIC was proposed to provide unity power

factor and high DC-link voltage without any power loss or voltage stress across the MOSFET device. The proposed system was implemented in MATLAB/SIMULINK tools. the DC-link voltage of PFC SEPIC was improved under intelligent control and power factor was almost unity (0.995). In addition, the results show the improvement of speed, stator current and electromagnetic torque of BLDC motor.

APPENDIX A

SYSTEM'S PARAMETER OF SEPIC CONVERTER

Output voltage $V_{DC} = 400v$, input voltage $V_{in} = 220vrms$, input inductance $L_{a,b} = 1.5mH$ and input capacitance $C_{1,2} = 0.047uF$.

APPENDIX B

CONTROL DESIGN

PI controller: $K_i = 28$, $K_p = 0.1$

Fuzzy controller: $K_e = 0.1$, $K_{\Delta e} = 1$ and $K_u = 8.2$

REFERENCES

- [1] K. Vinida and M. Chacko, "Optimal Tuning of H Infinity Speed Controller for Sensorless BLDC Motor using PSO and its Simulation Study in Underwater Applications," Ph.D. dissertation, Cochin University of Science and Technology, 2019.
- [2] V. M. Hernández-Guzmán and J. Orrante-Sakanassi, "PID control of robot manipulators actuated by BLDC motors," *International Journal of Control*, vol. 94, no. 2, pp. 267–276, 2021.

- [3] N. Priyadarshi, S. Padmanaban, L. Mihet-Popa, F. Blaabjerg, and F. Azam, "Maximum power point tracking for brushless DC motor-driven photovoltaic pumping systems using a hybrid ANFIS-FLOWER pollination optimization algorithm," *Energies*, vol. 11, no. 5, p. 1067, 2018.
- [4] T. D. Krishnan, C. M. C. Krishnan, and K. P. Vittal, "Design of robust H-infinity speed controller for high performance BLDC servo drive," in *2017 International Conference on Smart grids, Power and Advanced Control Engineering (ICSPACE)*. IEEE, 2017, pp. 37–42.
- [5] C.-I. Xia, *Permanent magnet brushless DC motor drives and controls*. John Wiley & Sons, 2012.
- [6] N. Huh, H.-S. Park, M. H. Lee, and J.-M. Kim, "Hybrid PWM Control for Regulating the High-Speed Operation of BLDC Motors and Expanding the Current Sensing Range of DC-link Single-Shunt," *Energies*, vol. 12, no. 22, p. 4347, 2019.
- [7] F. U. Jian, F. U. Yongling, Y. U. Liming, G. U. O. Jianwen, Y. A. N. G. Rongrong, and W. A. N. G. Mingkan, "Digital fixed-frequency hysteresis current control of a BLDC motor applied for aerospace electrically powered actuators," *Chinese Journal of Aeronautics*, vol. 31, no. 6, pp. 1287–1297, 2018.
- [8] A. Prudenzi, U. Grasselli, and R. Lamedica, "IEC Std. 61000-3-2 harmonic current emission limits in practical systems: need of considering loading level and attenuation effects," in *2001 Power Engineering Society Summer Meeting. Conference Proceedings (Cat. No. 01CH37262)*, vol. 1. IEEE, 2001, pp. 277–282.
- [9] V. S. Kommukuri, K. B. Mohanty, A. Chatterjee, and K. Thakre, "Modified bridgeless SEPIC rectifier for power factor correction with reduced switch stress operating in continuous conduction mode," *Journal of Circuits, Systems and Computers*, vol. 27, no. 08, p. 1850127, 2018.
- [10] S. Dutta, S. Gangavarapu, A. K. Rathore, R. Singh, S. Mishra, and V. Khadkikar, "Novel Single-Phase Cuk-derived Bridgeless PFC Converter for On-Board EV Charger with Reduced Number of Components," *IEEE Transactions on Industry Applications*, 2022.
- [11] A. Bouafassa, L. Rahmani, and S. Mekhilef, "Design and real time implementation of single phase boost power factor correction converter," *ISA transactions*, vol. 55, pp. 267–274, 2015.
- [12] M. U. . O. Z. GALEANO, N. et LÓPEZ LEZAMA, and J. María, "Modeling and development of a bridgeless PFC Boost rectifier," 2017.
- [13] M. Mahdavi and H. Farzanehfard, "Bridgeless SEPIC PFC rectifier with reduced components and conduction losses," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 9, pp. 4153–4160, 2010.
- [14] A. Bouafassa, L. M. Fernández-Ramírez, and B. Babes, "Power quality improvements of arc welding power supplies by modified bridgeless SEPIC PFC converter," *Journal of Power Electronics*, vol. 20, no. 6, pp. 1445–1455, 2020.
- [15] S.-W. Lee and H.-L. Do, "Efficient bridgeless PFC converter with reduced voltage stress," *International Journal of Circuit Theory and Applications*, vol. 44, no. 7, pp. 1455–1467, 2016.
- [16] A. Alice Hepzibah and K. Premkumar, "ANFIS current–voltage controlled MPPT algorithm for solar powered brushless DC motor based water pump," *Electrical Engineering*, vol. 102, no. 1, pp. 421–435, 2020.
- [17] A. Bouafassa, L. Rahmani, A. Kessal, and B. Babes, "Unity power factor Converter based on a Fuzzy controller and Predictive Input Current," *ISA transactions*, vol. 53, no. 6, pp. 1817–1821, 2014.

Modeling and Assessment of Energy Harvesting WSNs with Unreliable Sensors for Military Applications

OUKAS Nourredine

LIMOSE laboratory

M'Hamed Bougara University of Boumerdès, Algeria.

LIMPAF Laboratory

Department of Computer Science

Akly Mohand Oulhadj University of Bouira, Algeria

n.oukas@univ-bouira.dz

BOULIF Menouar

Department of Computer Science

LIMOSE laboratory

M'Hamed Bougara University of Boumerdès

Independence Avenue, 35000, Boumerdès, Algeria

boumen7@gmail.com

Abstract—Wireless sensor networks (WSN) are nowadays an active area of research. One of the main reasons of this activity is their military applications. Indeed, armies can place sensor nodes in order to control war zones, monitor borders, supervise battlefields and achieve many other duties with the least amount of casualties. A long-lasting WSN uses sensors that are equipped with an energy harvesting system. Furthermore, in most cases and typically for military applications, sensor nodes are susceptible to breakdowns due to the harsh environment in which they operate. In this paper, we propose a Generalized Stochastic Petri Nets (GSPN) that considers the energy harvesting (EH) capability and that takes into account sensor breakdowns. Our model proves to be able to simulate the actual behavior of the WSNs and hence can estimate the parameter values to ensure the longevity of service.

Index Terms—Wireless sensor network, Military application domain, GSPN modeling, energy harvesting, breakdowns, performance evaluation.

I. INTRODUCTION

A Wireless Sensor Network (WSN) is a large set of small devices named sensor nodes (SNs) that can communicate wirelessly. SNs can detect events pertaining to several phenomenons such as temperature, humidity, pressure, speed, acceleration, etc. They send report packets on these events to a base station (sink) for further or real-time processing. WSNs are used in various applications, such as agriculture, ambient monitoring, health, construction industry, smart cities, military applications, and many others [1]–[4]. WSNs have become an excellent tool for military applications involving intrusion detection, perimeter monitoring, information gathering and smart logistics support in unknown and far-reaching areas [5].

An SN is composed of several units to accomplish its functionality such as the sensing unit, the processing unit, the transceiver, and the energy unit. This last unit is equipped with a battery that can store only a little amount of energy. Hence, the lifetime of SNs is strongly depended on the battery charge.

Among the solutions that enable to extend the SNs lifetime, we can equip them with harvesters to get energy

from their environment (sun, wind, heat, pressure, and others) and convert it to electrical energy. This technique is called energy-harvesting [6], [7] and wireless networks equipped with such SNs are called energy harvesting wireless sensor networks (EH-WSN) [8]–[10].

Usually, when an SN detects an event, it creates a report message and sends it to a neighbor of next-hop to reach the sink. So, a neighbor of next-hop can be considered as a service station. However, in actual life, the service station may fail and requires to be repaired. This breakdown phenomena [11] can be encountered in several areas such as computer networks, sensor network (our case), flexible manufacturing systems and many others. There are two types of breakdowns: Active breakdowns and Idle breakdowns. The first one means that the breakdown occurs when the server is functional. Whereas, the idle breakdown occurs when the server (a neighbor in our case) is idle. Let us define two values of probability: p_{active} and p_{idle} (for active and idle breakdowns respectively). If $p_{active} = p_{idle} > 0$ this discipline is called independent breakdowns and if $p_{active} > 0$ and $p_{idle} > 0$ (generally $p_{active} \neq p_{idle}$) the strategy is called dependent breakdowns (see Fig. 1).

Therefore, the independent breakdown discipline is a special case of the dependent breakdowns approach. In this later, the failure probability depends on the server state. Hence, the values p_{active} and p_{idle} could be equal or not [12].

Modeling approaches allow to evaluate the performances of this kind of networks before an actual construction. GSPN formalism is used to model the retrial phenomenon in [13]. In [14], [15], the authors tackle other circumstances such as the sleeping mechanism and the message losses [15] by considering the sensor-neighbors relationship. A GSPN model with unreliable orbit is introduced by P.Wuchner et al. [16] to evaluate the performance of WSNs. The Authors also consider the sleeping mechanism and present an analysis that shows the influence of the sleep/awake ratio on the response

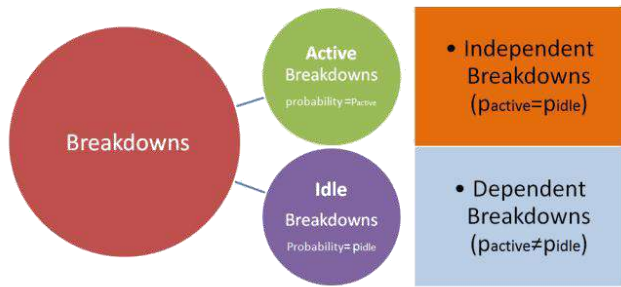


Fig. 1: Breakdown disciplines.

time. In [17], the authors propose a modeling of energy harvesting and energy consumption in an SN. They propose to represent sensor battery by quantum levels and show with experimental analysis the influence of energy harvesting on the mean response time and the mean battery charge. Authors in [18] use GSPN to model the case of clustered networks with two leaders by cluster. In [19] the authors proposed modeling by Colored Generalized Stochastic Petri Nets to show the effect of different packet types on the system performances. In [20], OUKAS and BOULIF proposed a Petri net model for a double sleeping strategy considering the lack of power at night. All these approaches do not consider the possibility of breakdowns which are very likely to occur in military applications.

Given this need, we propose in this paper a GSPN modeling that considers failures and repairs of neighbors and SNs with energy harvesting capability. In addition to these considerations, our modeling takes into account retrieval attempts and sleeping mechanism techniques to address energy conservation. We represent the energy in the SN battery using the quantization approach [17].

The rest of this paper is organized as follows: In section 2, we present some military domain applications. In the next section, we give a short description of the GSPN formalism. In section 4, we construct a GSPN model for EH-WSN. In the next section, we show and discuss some performance evaluation results. Finally, we draw our conclusion as well as some directions for future works.

II. DEFENCE APPLICATIONS OF WSNs

The sensor types and their capabilities determine and constrain the application of WSNs [4], [21]. Implementing WSNs in military applications has triggered various studies. Sensor networks can be used in several areas such as intrusion detection, enemy tracking, target classification, battlefield surveillance (see Fig. 2), battlefield damage assessment, target system, detection of nuclear, biological and chemical attacks, transient signal localization, helicopter detection, sea mine detection and many others [21], [22].

A hybrid, large-scale wireless sensor network for missile defense that consists of terrestrial and satellite nodes is introduced in [24]. They argue that a ballistic missile defense system should be capable of providing a layered defense against multiple ballistic threats of all ranges and in all phases of flight [24], [25]. The proposed network has been considered as a system-of-systems (see Fig. 3) with the objective of real-time tracking of multiple ballistic missile threats. The terrestrial nodes will be both mobile and fixed resources and platforms, such as land-based stations, warships, fixed-wing aircraft, and unmanned air vehicles. The satellite sensor platforms will be Geostationary Earth Orbit (GEO) and Low Earth Orbit (LEO) satellites, carrying search and track IR sensors, respectively [24].

III. GSPN PRESENTATION

Generalized Stochastic Petri Nets are a powerful tool for modeling and evaluating the performance of discrete, stochastic and concurrent systems [26].

A GSPN is a graphical and mathematical formalism composed of places, tokens, transitions, and arcs. Places (represented by circles) may contain marks (or tokens, represented by big dots or by their number). A transition represents a discrete event. There are two types of transitions: Timed transitions and Immediate transitions. The firing of a transition that belongs to the first type needs a time delay. Arcs link places and transitions. The firing of a transition consumes a number of tokens equal to the weights of the input arcs (an input arc links a place to a transition) and produces a number of tokens to the associated output place equal to the weight of their output arcs. An inhibitor arc with a weight equals to w forbids the firing of a transition if the number of tokens in the input place is less than w . Fig.4 represents the graphical building blocks of GSPN modeling.

In general, places can represent resources, the number of marks in a place being the number of copies of this resource. Transitions can represent actions. The firing of a transition corresponds to the realization of the associated action.

Furthermore, each timed transition is associated with a delay parameter that represents the waiting time before the firing can occur. According to their behavior, timed transitions have two semantics: single server (the simple case) and infinite server. Single server semantic models the natural understanding of a restricted resource that can only perform one action at a time (the firing delay parameter is divided by the number of tokens in its input place). The infinite server stands for an arbitrary number of resources, from which one is activated for every token individually [26].

After the construction of a Petri net, two kinds of analysis can be conducted on it. Qualitative analysis that covers, among others, liveness, boundedness, strong connectivity, ergodicity. Quantitative analysis is achieved after the qualitative analysis and consists of calculating the performance parameters after solving the network in the stationary state. After deriving the reachability marking graph, the solution of the network in the equilibrium

Fig. 2: A military surveillance scenario [23]

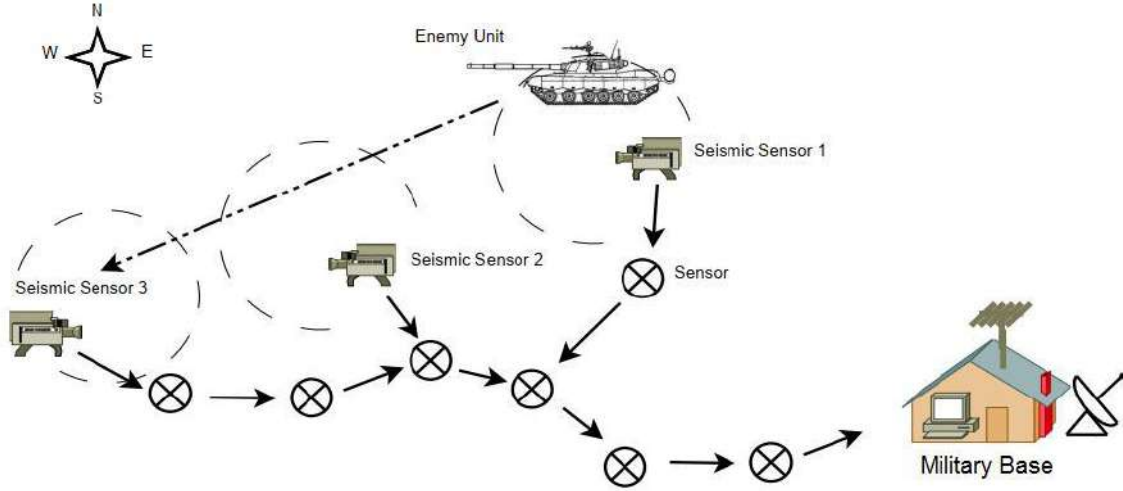


Fig. 3: Hybrid Wireless Sensor Network [24]

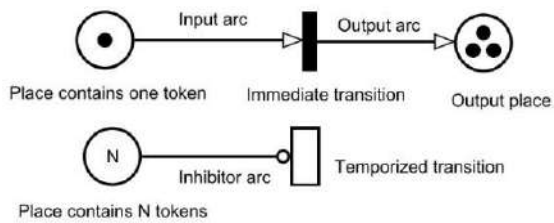
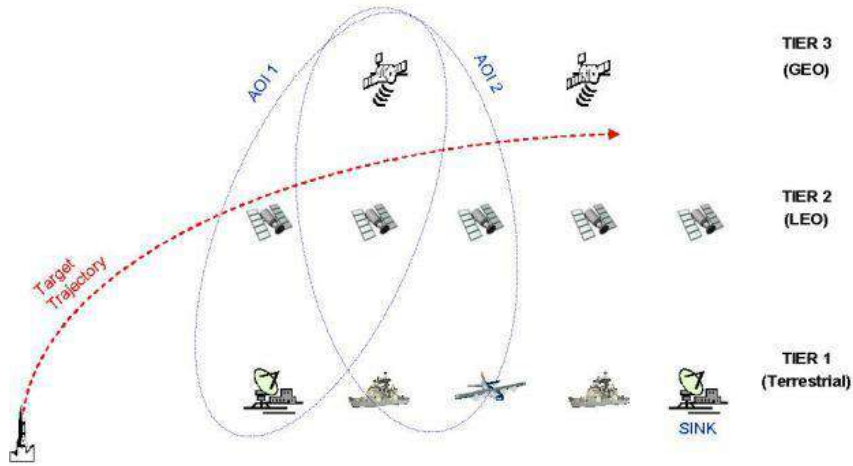


Fig. 4: GSPN graphical building blocks.

state provides the vector of stationary probability denoted .

We present basic performance laws in GSPN modeling [26], [27]:

- The mean number of marks in a place P :

$$\bar{P} = \sum_{i: M_i \in M} M_i(P) \cdot \pi_i \quad (1)$$

- The probability of an event A denoted $\text{prob}(A)$:

$$\text{prob}(A) = \sum_{i: M_i \in M} \pi_i \quad (2)$$

where M_i satisfies A .

- The mean debit of a transition T with rate λ denoted

$$\bar{\lambda} = \lambda \cdot \text{prob}(|P| > 0) \quad (3)$$

where $|P|$ is the number of tokens in the output place P of the transition T .

- The mean sojourn time in a place P denoted $\text{wait}(P)$:

$$\text{wait}(P) = \frac{\sum_{i: M_i \in M} M_i(P) \cdot \pi_i}{\bar{\lambda}} \quad (4)$$

where P is the output place of the timed transition T and λ its mean debit.

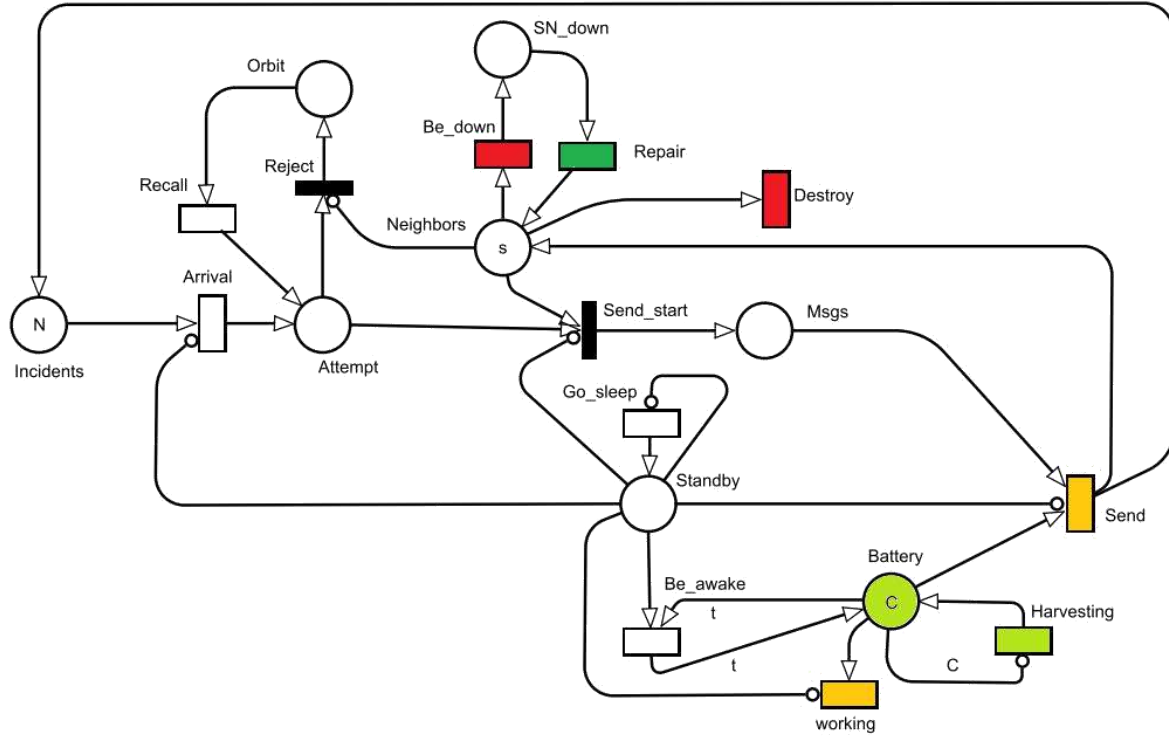


Fig. 5: Petri net Modeling for EH-WSNs with unreliable sensors

IV. PETRI NET MODELING FOR EH-WSNs WITH UNRELIABLE SENSORS

Fig.5 represents our proposed GSPN that describes how does a given node interact with its neighbors. Table I recapitulates the description of all places. When an incident is detected by a sensor, it creates a message and attempts to send it to a next-hop neighbor. Alternatively, if it receives a message from another sensor, it takes care to forward it to a neighbor that is closer to the sink. If there is no idle next-hop neighbor (no token in the place Neighbors), the message joins the orbit (this action is triggered by the firing of the reject transition) for further call attempts. If there is at least one idle neighbor, the transition Send start fires and the message enters the transmission process. The sending process reach its end when the transition Send fires.

TABLE I: Places description

Place	signification	initial marking
Incidents	source of a messages	N
Attempt	message sending attempt	0
Orbit	message waits for an idle next hop	0
Neighbors	idle next hop neighbors	s
Msgs	messages in sending process	0
SN_down	neighbors requiring a repair	0
Battery	energy quantum number in the SN	C
Standby	The SN is sleeping	0

On the other hand, SNs are susceptible to breakdowns. The firing of the transition Be down represents a crash of a

next-hop neighbor. The place SN_down contains the crashed neighbors. After a repair operation, a down neighbor becomes operational and joins the place Neighbors when the firing of the transition Repair occurs. The proposed model is generic and hence it can be customized for military applications by considering the high probability of SNs destruction. Particularly on battlefield, it is known that collocated sensors can be simultaneously destroyed by a single hit (see Destroy transition). For all the timed transitions, The inter-firings times are supposed to be exponentially distributed. Table II lists them with a short description. Table III describes immediate transitions.

In our model, a sensor can switch between two states: Active and Standby. In order to save energy and enhance the battery charge, each sensor stays in the standby sate until the battery charge becomes greater than a given threshold t.

Regarding the energy considerations, the presence of a rechargeable battery is represented by the place Battery that initially contains C quantum of energy. The number of quantum increases when the transition Harvesting fires and it decreases when the sensor is activated (See Working transition). We suppose that a quantum is the amount of energy required by the sending of one message.

V. PERFORMANCE EVALUATION AND DISCUSSION

To analyze our GSPN model, we use the TimeNet tool [26], [28]. The simulation inputs rate values are presented in Table II and we use for the remaining inputs the values:

TABLE II: Timed transitions description

Transition	Signification	Firing rate	Value
Arrival	Arrival of a message	t_{arrive}	0.01
Send	Success-full sending to a next hop	t_{send}	20
Be <u>down</u>	A next-hop neighbor breakdown	t_{down}	0.001
Destroy	Destroy a neighbor	t_{dest}	0.0001
Recall	Retrial of a message	t_{recall}	15
Repair	Repair of a next-hop neighbor	t_{repair}	0.01
Harvesting	Energy recovery	$t_{harvesting}$	10
W orking	Working energy consumption	$t_{working}$	0.01
Go <u>sleep</u>	The main sensor joins the sleeping state	t_{sleep}	0.01
Be <u>awake</u>	Activation of the main sensor	t_{awake}	0.001

TABLE III: Immediate transitions description

Immediate transition	Signification	Priority
Send start	start sending to a next-hop neighbor	2
Reject	message joins the orbit	1

(N; s; C; t)=(10; 3; 20; 5). These values have been chosen to prove the ability of our model to evaluate the EH-WSNs.

TimeNet helps us to derive the reachability graph of the system and obtain the vector of stationary probabilities.

Fig. 6: Mean response time versus harvesting rate

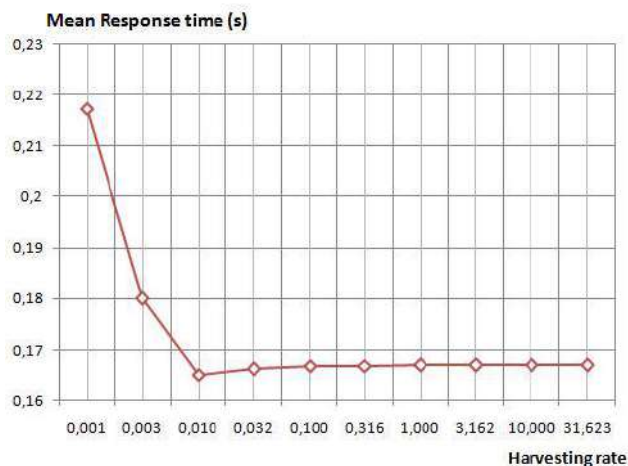


Fig. 6 represents the mean response time versus the harvesting rate. It is clear from the figure that the mean response time decreases when the harvesting rate increases. The mean response time stabilizes near the value of 0:167 seconds.

Fig. 7 represents the mean response time versus the down rate. We vary the breakdown rate to show its influence on the mean response time. When we have a lot of breakdowns among the next-hop neighbors, the mean response time increases because the sensor loses a long time in the orbit.

Fig. 7: Mean response time versus failure rate

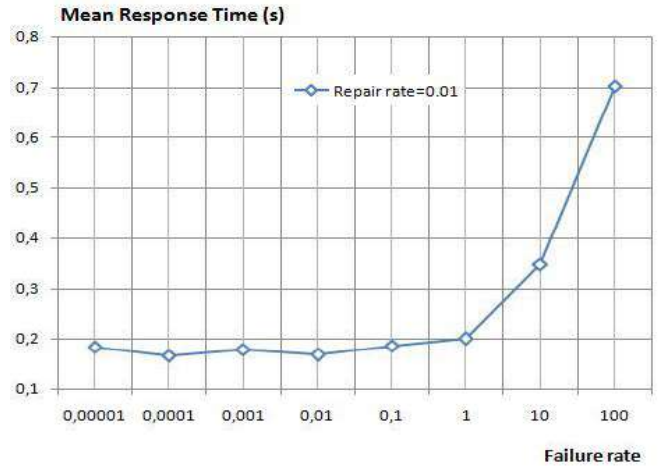


Fig. 8: Mean Battery charge versus harvesting rate

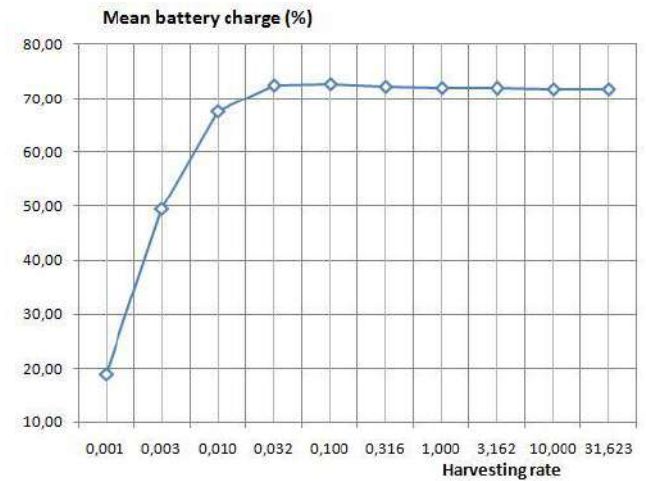
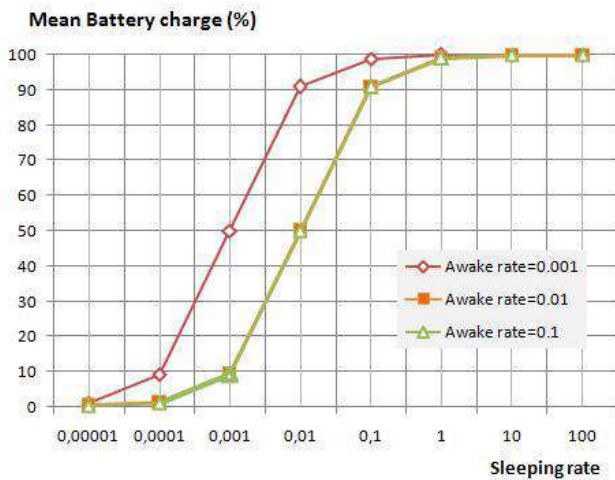


Fig. 8 shows the influence of the harvesting rate on the mean battery charge. We notice that the mean battery charge increases when the harvesting rate increases. These results are logical and coincide with the reality. The mean battery charge stabilizes near the value of 72 % when the harvesting rate value is greater than 0:1.

Fig. 9 shows the influence of the sleeping rate on the mean battery charge. The mean battery charge enhances when the sleeping rate enhances. The SN saves more energy when it is in the sleeping state. The battery of the sensor becomes almost full when the sleeping rate is greater than the value of 1 whatever be the configurations.

Fig. 9: Mean Battery charge versus sleeping rate



VI. CONCLUSION

In this paper, we proposed a Petri net modeling for EH-WSNs with the presence of unreliable sensors. We considered retrial attempts, duty cycles for sensor operating in harsh condition to mimic the actual behavior on battlefields. The analysis of the numerical results by using several scenarios proves that our model allows the network supervisors to adjust the input values to have the suitable behavior of the network before its actual deployment.

In our future work, we want to enhance our modeling by adding awareness of existing protocols in WSNs and by comparing the results to other simulation approaches.

REFERENCES

- [1] S. Shen, L. Sun, Y. Dang, Z. Zou, and R. Wang, "Node localization based on improved pso and mobile nodes for environmental monitoring wsns," *International Journal of Wireless Information Networks*, vol. 25, no. 4, pp. 470–479, Dec 2018. [Online]. Available: <https://doi.org/10.1007/s10776-018-0414-3>
- [2] P. Mohanty and M. R. Kabat, "Energy efficient reliable multi-path data transmission in wsn for healthcare application," *International Journal of Wireless Information Networks*, vol. 23, no. 2, pp. 162–172, Jun 2016. [Online]. Available: <https://doi.org/10.1007/s10776-016-0307-2>
- [3] Z. Xu, Y. Liu, and X. Li, "Special issue on selected technologies and applications in wireless communications for smart city computing," *International Journal of Wireless Information Networks*, vol. 25, no. 3, pp. 229–231, Sep 2018. [Online]. Available: <https://doi.org/10.1007/s10776-018-0413-4>
- [4] M. P. urisi'c, Z. Tafa, G. Dimic, and V. Milutinovic, "A survey of military applications of wireless sensor networks," in *2012 Mediterranean conference on embedded computing (MECO)*. IEEE, 2012, pp. 196–199.
- [5] C. F. Garc'ia-Hernandez, P. H. Ibarguengoytia-Gonzalez, J. Garc'ia-Hernandez, and J. A. Perez-D'iaz, "Wireless sensor networks and applications: a survey," *IJCSNS International Journal of Computer Science and Network Security*, vol. 7, no. 3, pp. 264–273, 2007.
- [6] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: survey and implications," *IEEE Communications Surveys and Tutorials*, vol. 13(3), p. 44361, 2011.
- [7] S. Basagni, M. Y. Naderi, C. Petrioli, and D. Spenza, "Wireless sensor networks with energy harvesting," *Mobile Ad Hoc Networking: Cutting Edge Directions*, pp. 701–736, 2013.
- [8] A. Rashid, F. Khan, T. Gul, F. e Alam, S. A. Khan, and F. K. Khalil, "Improving energy conservation in wireless sensor networks using energy harvesting system," *(IJACSA) International Journal of Advanced Computer Science and Applications*, vol. 9 (1), 2018.
- [9] K. S. Adu-Manu, N. Adam, C. Tapparelo, H. Ayatollahi, and W. Heinzelman, "Energy-harvesting wireless sensor networks (eh-wsns): A review," *ACM Transactions on Sensor Networks (TOSN)*, vol. 14, no. 2, p. 10, 2018.
- [10] N. Ashraf, M. Faizan, W. Asif, H. K. Qureshi, A. Iqbal, and M. Lestas, "Energy management in harvesting enabled sensing nodes: Prediction and control," *Journal of Network and Computer Applications*, vol. 132, pp. 104–117, 2019.
- [11] M. Jain and A. Jain, "Working vacations queueing model with multiple types of server breakdowns," *Applied Mathematical Modelling*, vol. 34, no. 1, pp. 1–13, 2010.
- [12] N. Gharbi and M. Ioualalen, "Gspn analysis of retrial systems with servers breakdowns and repairs," *Applied Mathematics and Computation*, vol. 174, no. 2, pp. 1151–1168, 2006.
- [13] N. Gharbi and L. Charabi, "Wireless networks with retrials and heterogeneous servers: Comparing random server and fastest free server disciplines," *International Journal on Advances in Networks and Services Volume 5, Number 1 & 2*, 2012, 2012.
- [14] N. Oukas and M. Boulif, "Energy-consumption-aware modelling and performance evaluation for eh-wsns," in *(IAM 2019) Proceedings of the second conference on Informatics and Applied Mathematics*. labstic.univ-guelma.dz, 2019, pp. 57–62. [Online]. Available: <http://labstic.univ-guelma.dz/sites/labstic.univ-guelma.dz/files/>
- [15] —, "A petri net modelling for eh-wsn with buffer limitation and mes-age losses," in *1st International Conference on Networking Telecommunications, Biomedical Engineering and Applications (ICNTBA 19)*.
- [16] P. Wuchner, J. Sztrik, and H. de Meer, "Modeling wireless sensor networks using finite-source retrial queues with unreliable orbit," in *International Workshop on Performance Evaluation of Computer and Communication Systems*. Springer, 2010, pp. 73–86.
- [17] N. Oukas and M. Boulif, "A petri net modeling for wsn sensors with renewable energy harvesting capability," in *Smart Energy Empowerment in Smart and Resilient Cities*, M. Hatti, Ed. Cham: Springer International Publishing, 2020, pp. 524–534.
- [18] —, "A new stochastic petri nets modeling for dual cluster heads configuration in energy-harvesting wsns," in *2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH 2020)*, 2021.
- [19] —, "A colored petri net to model message differences in energy harvesting wsns," in *Proceedings of The 4th Conference on Computing Systems and Applications*. Ecole Militaire Polytechnique - Chahid Abderrahmane Taleb (EMP), Algiers, Algeria, 2020.
- [20] —, "A new petri nets for wsns to model the behaviour of solar-energy harvesting sensors with double sleeping strategy," in *Proceedings of the second International Conference on Computing and Information Technology, ICCIT'1442*. University of Tabuk, Kingdom of Saudi Arabia, 2021.
- [21] I. Ahmad, K. Shah, and S. Ullah, "Military applications using wireless sensor networks: A survey," *Int. J. Eng. Sci.*, vol. 6, no. 6, p. 7039, 2016.
- [22] K. Sohraby, D. Minoli, and T. Znati, *Wireless sensor networks: technology, protocols, and applications*. John Wiley & Sons, 2007.
- [23] A. S. Uluagac, R. A. Beyah, and J. A. Copeland, "Secure source-based loose synchronization (sobas) for wireless sensor networks," *IEEE Transactions on Parallel and distributed systems*, vol. 24, no. 4, pp. 803–813, 2012.
- [24] P. Katopodis, G. Katsis, O. Walker, M. Tummala, and J. B. Michael, "A hybrid, large-scale wireless sensor network for missile defense," in *2007 IEEE International Conference on System of Systems Engineering*. IEEE, 2007, pp. 1–5.
- [25] M. M. Shahkaram, "Ballistic missile defense in the middle east," *Naval Postgraduate School Monterey United States*, Tech. Rep., 2019.
- [26] A. Zimmermann, *Stochastic Discrete Event Systems*. Springer, 2007.
- [27] G. Florin, C. Fraize, and S. Natkin, "Stochastic petri nets: Properties, applications and tools," *Microelectronics Reliability*, vol. 31, no. 4, pp. 669–697, 1991.
- [28] A. Zimmermann, "Dependability evaluation of complex systems with timenet," in *Proceedings of the First Workshop on DYNAMIC Aspects in Dependability Models for Fault-Tolerant Systems*. ACM, 2010, pp. 33–34.

Multi-classification of thyroid nodules using Machine learning and deep learning

Linda BELLAL

Biomedical Engineering dept. Faculty of Technology
University of Tlemcen, Algeria
linda.bellal@univ-tlemcen.dz

Kamila KHEMIS

Biomedical Engineering dept. Faculty of Technology
University of Tlemcen, Algeria
Khemiskamila13@gmail.com

Abstract— Histology is considered an essential study in the management of thyroid nodules which are solid or fluid-filled masses that form in the thyroid gland. It is sometimes a truly diagnostic tool for benign or malignant lesions, sometimes a screening tool with a selection of nodules justifying surgery. certain histopathology doctors, in particular the new histopathology doctors, sometimes encounter difficulties in diagnosing these thyroid nodules because their pathological types are complex. This project aims to automatically classify histological images of thyroid nodules to help easily and quickly diagnose histopathology or to help confirm the microscopic diagnosis. We have compared two classification methods on a real basis to make the choice of a good diagnostic aid classifier. The work consisted of four steps, first the acquisition of images, then a step of preprocessing of the acquired images, then a step of feature extraction was carried out, finally, the classification step was carried out with two approaches, the Machine Learning approach (classification with Support Vector Machine (SVM) and K-Nearest Neighbor Method (KNN)) and Deep Learning approach (classification with a simple Convolutional Neural Network (CNN) architecture). The best classification results were obtained with the Machine learning approach by SVM with an accuracy of 95%. Thus the deep CNN method does not bring any added value compared to SVM but this conclusion may relate to the size of the database. Therefore, we will continue these comparisons with a large database.

Keywords— Histology, Thyroid, feature extraction, Classification, Machine learning, Deep learning, CNN.

I. INTRODUCTION

Thyroid nodules are solid or fluid-filled masses that form in the thyroid gland, are common diseases presenting clinically and their pathological types are complex. Thyroid nodules fall into two categories, benign thyroid nodules, and malignant thyroid nodules. Benign thyroid nodules are nodular goiter and thyroid adenoma. As for malignant thyroid nodules, they are represented in papillary carcinomas, follicular, medullary, and anaplastic carcinomas of the thyroid. Thyroid cancer requires surgery, whereas only follow-up is necessary in the case of benign nodules, which is why the differential diagnosis of thyroid nodules is essential. The gold standard for tumor diagnosis is the pathological diagnosis of resected specimens. Currently, the vast

majority of pathological tissue cuts are acquired by pathologists, and long-term collections of samples are used for clinical diagnosis. Nonetheless, manual differential identification of thyroid tumor histological pictures remains difficult for three major reasons: firstly, the ability to correctly diagnose samples is highly dependent on the professional training and experience of the pathologist, and such experience cannot be acquired quickly; secondly, the task is time-consuming, expensive, and boring; thirdly, it is difficult for the human eye to distinguish subtle changes in the tissues; As a result, pathologists may become fatigued, which might lead to an incorrect diagnosis. Thus, determining the correct histological diagnosis of thyroid nodules is difficult[1].

An algorithm is used to allocate test data to specified groups in classification tasks. These algorithms have been increasingly used in the field of medical imaging and have been applied to the pathological diagnosis of different diseases, machine learning algorithms (ML) and deep learning algorithms (DL) were used[2].

In this study, to make an automatic diagnosis of thyroid nodules in histological images we applied automatic learning algorithms (Support Vector Machine SVM and K-nearest neighboring KNN) and also deep learning algorithms (Convolutional Neural Networks CNN, which is a class of deep neural networks). This study aimed to reveal whether the automatic and deep methods have the potential to improve diagnostic efficiency and improve the amount of inter-observer agreement in thyroid nodule categorization.

The remainder of this document is organized as follows: Section 2 details related work; Section 3 details the dataset and the method used; Section 4 presents the results and discusses the obtained results; and Section 5 concludes the study.

II. RELATED WORK

Several new and personalized deep learning models have been proposed by researchers to classify histopathological images of cancers of the following organs: bone (Song et al., 2018)[3], blood (Achi et al., 2018)[4], brain (Xu et al., 2015)[5]; Yonekura et al., 2018[6], Breast (Nejad et al., 2017[7]; Feng et al., 2018[8]; Alom et al., 2019[9]), Colon (Ciompi et al., 2017[10]; Wang et al., 2017a[11]; Xu and Zhang, 2018[12]; Korbar et al., 2017[13]), Lung (Hou et al., 2015[14]), Gastric (Sharma et al., 2017[15]), Pancreas (Chang et al., 2017b)[16] and Prostate (Otálora et al., 2018)[17].

Transfer-learned deep learning models have also been widely used to classify histopathological images of cancers of the following organs: blood (Tambe et al., 2019)[18], brain (Xu et al., 2015)[5], breast (Xie et al., 2019[19]; Bejnordi et al., 2017[20]; Chang et al., 2017a[21]; Spanhol et al., 2017[22]; Gandomkar et al., 2018[23]), cervix (Wang et al., 2017b)[24] and the colon (Ponzio et al., 2018)[25].

Few works have been reported in the literature on the classification of thyroid histopathology images using artisanal features and traditional machine learning techniques (Ozolek et al., 2014 [26]; Huang et al., 2014 [27] ; Jothi and Rajam, 2016 [28], 2017 [29]).

Few works have been reported in the literature on the classification of thyroid histopathology images with the application of deep learning methods (Wang et al., 2019)[1] and transfer learning methods (V. G. Buddhavarapu and A. A. J. J, 2020).[30]

III. MATERIAL AND METHOD

To compare the effectiveness of the deep method to the automatic methods, our work is divided into two axes: the automated approach is treated on the first axis, while the deep approach is treated on the second. In both cases, the images are preprocessed and the test will be done without then with an increase in the database. Regarding the automatic approach, we tested the SVM machine vector support and the K-Nearest Neighbors. We tested a simple CNN Architect for the deep approach. The classification was applied after a feature extraction step. Certain texture properties and fractal dimensions were chosen as parameters. In the deep approach, this step does not exist.

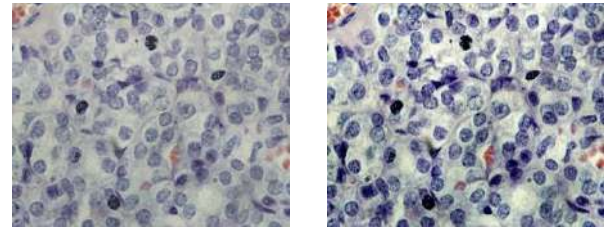
A. Image acquisition

We used an image dataset of 28 histological images of the thyroid. The image dataset was acquired from patients who were diagnosed with thyroid tumors. These images are acquired at the histology laboratory of the Faculty of Medicine of the University of Tlemcen using a microscope Olympus CX41 and a DP12 camera. Pathological slides were stained with hematoxylin and eosin (H&E) for morphological evaluation. All images were scanned at $\times 40$ magnifications. Each image was labeled with one of 3 classes : 0: normal thyroid (NT), 1: follicular adenoma (FA), 2: papillary thyroid carcinoma (PTC). Labeling was done by a pathologist. 10 images were classified as normal thyroid, 10 images were classified as follicular adenoma and 8 images were classified as papillary thyroid carcinoma (PTC).

B. Preprocessing

1) Contrast enhancement

Color images were employed in our work. We noticed that these images are not of good quality this is due to either the poor quality of the blade or the acquisition device (DP12). To improve the quality of the images, we tested some filters (median filter, Gaussian filter), but unfortunately, we didn't get satisfactory results (loss of information), so we opted for a contrast enhancement. Fig 1 shows the result of such processing.



(a) original image

(b) image improved

Fig. 1. Contrast enhancement

2) Data augmentation

deep learning methods need a large database in order to give a better result. In order to slightly increase the number of images in our database (only 28 images). In order to slightly increase the number of images in our database, we used the process of flipping and rotating images.

We augmented the data by flipping and rotating the images. Each image fragment was flipped horizontally and rotated 0° , 90° , 180° , and 270° . Through this process of flipping and rotating, we quadrupled the size of the data.

We obtained a total of 112 images: 40 TN images, 40 AF images, and 32 CP images.

C. feature extraction

Feature extraction is a component of image processing that often goes hand in hand with classification. Indeed, to establish a classification rule (supervised or unsupervised), we generally rely on a set of numerical criteria describing the object or phenomenon observed. We have used textural features to criticize these criteria.

In image processing, texture analysis involves calculating a series of measurements to define a perceived texture on an image. Texture analysis returns information about the spatial arrangement of colors or intensities in all or part of this image texture was measured using some approaches, including co-occurrence matrix, fractals, Gabor filters, and wavelet transform variations[31]. We chose the texture measurement by GLCM and fractal dimension, taking into account what has been reported by researchers in medical image analysis regardless of the image acquisition mode and the investigative organ.

1) The Grey Level Co-Occurrence Matrix (GLCM)

The Grey Level Co-Occurrence Matrix (GLCM)[32] technique is a second-order statistical approach for texture extraction. GLCM is a matrix that represents the relative frequencies of a pair of grayscale images at a certain distance (d) and angle (α). The distance d varies from 1 to the size of the image while α is in four directions, i.e., 0° , 45° , 90° , and 135° . GLCM readings generated from various pairs of angles and distances have a wide range of features. The selection of the optimum angle is critical for extracting textual information from images with strong directional qualities. Typically, the four directions are chosen and the average of the four GLCMs' properties is used[33].

According to Haralick, there are 14 textural characteristics retrieved from the gray level co-occurrence matrix. Contrast,

correlation, homogeneity, energy, and dissimilarity are the most commonly employed. They are defined below[34].

Energy :

$$f1 = \sum_i \sum_j p(i, j)^2 \quad (1)$$

Contrast :

$$f2 = \sum_{n=0}^{Ng-1} n^2 \{ \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} p(i, j) | i - j | = n \} \quad (2)$$

Correlation :

$$f3 = \frac{\sum_i \sum_j p(i, j) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (3)$$

Homogeneity :

$$f4 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j) \quad (4)$$

Dissimilarity :

$$f5 = \sum_i \sum_j |i - j|. p(i, j) \quad (5)$$

2) The fractal dimension

Fractal dimension is a crucial metric in fractal geometry that has a wide range of applications, including image processing. Image analysis is a high-level picture processing approach for determining image attributes including texture, roughness, softness, surface, and robustness[35]. The fractal dimension may be calculated using a variety of methods. The box-counting approach was adopted in this investigation because it is simple, automated, and adaptable to models with or without self-similarity[36]. Each image is covered by a series of descending-size grids in this approach, and two values are recorded for each grid: the number of squares intersected by the image, $N(s)$, and the lateral length of the squares, s . The regression slope D of the line formed by plotting $\log(N(s))$ concerning $\log(1/s)$ indicates the degree of complexity, or fractal dimension, between 1 and 2 ($1 \leq D \leq 2$). An image with a fractal dimension of 1, or 2, is considered completely differentiable or very rough and irregular, respectively. The linear regression equation used to estimate the fractal dimension was[37] :

$$\log(N(s)) = \log(K) + D \log(1/s) \quad (6)$$

where K is a constant and $N(s)$ is proportional to $(1/s)^{-D}$

D. Classification

Image classification, in particular the selection of methodologies and strategies for utilizing the results of image processing and pattern recognition, is a significant problem for image analysis jobs. The results of the picture categorization are then validated using classification methods[38]. The main objective of the classification in our project is to say with high accuracy if a given sample of the thyroid is healthy benignant or

malignant. We, therefore, have three groups and the multiclass classification is the appropriate tool.

We were interested in two classification approaches the machine learning approach with the Support Vector Machine method and the K-nearest neighbors method and the deep learning approach with the convolutional neural network.

1) Multiclass classification using Machine learning

a) Multiclass classification using Support Vector Machine (SVM)

In its simplest type, SVM does not support multi-class classification natively. It supports binary classification and the separation of data points into two classes. Datapoint separation calculations depend on a kernel function. There are different kernel functions: linear, polynomial, Gaussian, radial base function (RBF), and sigmoid. In simple terms, these functions determine the fluidity and efficiency of class separation, and playing with their hyperparameters can lead to over-adjustment or under-adjustment. For multi-class classification, the same principle is used after breaking down the multi-class problem into several binary classification problems.

The three most popular multi-class SVM methods: One Vs Rest, One Vs One, and Directed Acyclic Graph Support Vector Machine[39].

The One-to-One approach breaks down the multi-class problem into several binary classification problems. A binary classifier for each pair of classes.

Another approach that can be used is One-to-Rest. In this approach, the distribution is defined on a binary classifier for each class.

The Directed Acyclic Graph Support Vector Machine approach is built by introducing one vs one classifier into the nodes of a Directed Acyclic Graph (DAG) decision[39].

In our work we have 3 classes (normal thyroid, papillary carcinoma, follicular adenoma), to classify them we have chosen the one-to-rest approach, with the use of two different nucleus functions: linear and RBF.

b) Multiclass classification using k-Nearest Neighbours (KNN)

K-Nearest Neighbors (KNN) is a collection of approaches for classification and regression. As a non-parametric learning algorithm, the nearest K -neighbors are not limited to a set number of parameters. However, their complexity will depend on the size of the learning base.

The KNN algorithm can compete with the most accurate models because it makes extremely accurate predictions. Therefore, it can be used for applications that require high precision but do not require the generation of human-readable models, as KNN does not produce comprehensible models. The K-nearest neighbors have been used in several fields and many applications such as statistical estimation and pattern recognition, prediction of economic events, estimation of the capacity of a lithium-ion battery, distance measurement, text categorization, and multi-label classification[40].

the KNN algorithm will have a very important parameter which is the 'k', 'k' is a parameter that refers to the number of nearest neighbors to include in the majority of the voting process.

In our work we used the KNN to make a multi classification, we chose the value of $K=3$.

2) Multiclass classification using Deep Learning

a) Multiclass classification using Convolutional Neural Network (CNN)

Deep Learning, as a branch of Machine Learning, uses algorithms to process data and mimic the thinking process in a human. Image classification has seen a breakthrough in performance, thanks to the rise of convolutive neural networks. A CNN consists of an input layer, convolution layers, pooling layers, fully connected layers, and an output layer. The input layer of a CNN includes the input image. A convolution layer consists of a set of filters. The filters are convoluted with the input to produce a set of feature maps. This is why we do not need to perform the traditional and separate phase of feature extraction as with SVM or KNN.

In CNN this phase is performed automatically with filters. Feature maps are passed to a non-linear activation function to generate the activation maps. A pooling layer is used to reduce the size of the activation cards. This reduces the number of parameters that can be learned for the following layers. A fully connected layer is a one-dimensional (1D) arrangement of neurons where each neuron in a fully.

connected layer is connected to all other neurons in the previous layer via a weight. The final output (classification) layer of a CNN consists of a set of neurons using which the model generates (predicts) the image class label. A typical CNN may consist of several convolution layers and one pooling layer followed by one or more fully connected layers[30].

In our work, we used a simple CNN architecture. Fig.2 presented the architecture used.

So our architecture is 10 layers. We entered our database of 3 classes with a fixed size of $64 \times 64 \times 3$ (the first layer). We used 3 convolution layers with filters of 3×3 size and with different numbers of filters (32, 64, 128 respectively, with ReLU activation function). After each convolution layer, we put a pooling layer (Max pooling) with filters of size 2×2 . The flattening layer is used after the last pooling layer, then a fully connected layer is defined with 128 units (with the ReLU activation function). In the end, a layer of 3 units is connected (the 3 units represent the 3 classes) the activation function used is softmax because we will have more than two classes.



Fig. 2. CNN architecture.

E. Evaluation metrics

In this section, metrics to evaluate our method have been introduced. In all our experiments, we have divided the dataset

as follows: First, we have tested the classification methods mentioned above (SVM and KNN) with an 80% learning base and a 20% test base. Second, we applied these classifiers with an 85% learning base and a 15% test base. We evaluated the classification performance of histological thyroid images proposed in our dataset in terms of accuracy (ACC) (equation (7)), precision (Pre) (equation (8)), recall equation (9)), F1 score equation (10)). The definitions of evaluation metrics are as follows :

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

$$Pre = \frac{TP}{TP+FP} \quad (8)$$

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

$$F1 = \frac{2.Pre.Recall}{Pre+Recall} \quad (10)$$

In the above formulas, TP, TN, FP and FN are respectively the numbers of true positives, true negatives, false positives and false negatives.

For the CNN assessment, we calculated learning accuracy and test accuracy.

IV. RESULTS AND DISCUSSION

In this section, the results of the classification will be presented. These results were obtained after programming in the Python programming language, the IDE used is PyCharm version 2019.3.1.

As mentioned earlier, we used two learning methods in our work, machine learning and deep learning.

The classification algorithms used in the machine learning method are the Support Vector Machine (SVM) and the K-Nearest Neighbors (KNN). For the SVM there are several kernels, we used two types of the kernel, the RBF kernel, and the linear kernel. For the KNN we set K to 3 ($K=3$).

To see if the number of images in the database (the total number of images) and if the number of images used for training compared to the number of images used for the test influenced the classification results, we went through 4 experiments: classification without augmentation of the database with 85% image for learning and 15% image for the test, classification with augmentation of the database with 85% image for training and 15% image for the test, classification without augmentation of the database with 80% image for training and 20% image for the test, classification with augmentation of the database with 80% image for learning and 20% image for the test.

Table I presents the results of the classification of thyroid histological images, using the SVM and KNN classifiers, with a learning base of 85% and a test base of 15%, without base augmentation and with augmentation. First, we used SVM and KNN classifiers with a database of 28 images (the base before augmentation), and we split this database into 85% images for training (23 images) and 15% images for the test (5 images), we got unsatisfactory results. Then, after augmentation of the

database we got 112 images, we split this database to 85% images for training (95 images) and 15% images for testing (17 images), we obtained satisfactory results with an accuracy of 94% for the SVM classifier (RBF), 94% for the SVM (linear) classifier and 64% for the KNN classifier.

Table II presents the results of the classification of thyroid histological images, using the SVM and KNN classifiers, with a learning base of 80% and a test base of 20%, without base augmentation and with augmentation. we used SVM and KNN classifiers with a database of 28 images (the base before augmentation), and we split this database into 80% images for training and 20% images for the test, we got unsatisfactory results. Then, after augmentation of the database we got 112 images, we split this database into 80% images for training and 20% images for testing, we obtained satisfactory results with 95% accuracy for the SVM(RBF) classifier, 95% for the SVM classifier (linear), and 65% for the KNN classifier.

We noticed that the size of the database influenced the classification results as well as the number of images used for training and testing. In this part, the classifier which gave us the best result is the SVM with an accuracy of 95%. The database used is 112 images, with a learning base of 80% and a test base of 20%.

The classification algorithms used in the deep learning method are the Convolutional Neural Network (CNN). Although the size of our database is very small to make a classification using CNN, despite that, we created a simple architecture of CNN and we have obtained encouraging results.

CNN was tested on an augmented basis with 85% and 15% for learning and testing respectively. We obtained a learning accuracy of 93% and a test accuracy of 75%, a great difference between the two exactitudes. This indicates that we have over-learning, so our classifier can make mistakes, so our classifier is not performing. The over-learning of our architecture may be due to an imbalance between the number of images used for learning that is very large compared to the number of images used for the test.

To improve the results we have increased the number of images used for the test and decreased the number of images used for learning, we have tested CNN modal score on a learning basis of 80% and a test base of 20%. We obtained a learning accuracy of 87% and a test accuracy of 81% (Table III).

According to our results (Table II/Table III), we can say that the Deep Learning approach does not bring a capital gain compared to SVM anyway with an advantage of SVM which consists of the possibility of freely choosing the characterization attributes of our images and a disadvantage of CNN which is the obligation to have a large database. should be noted that the free choice of pretreatment and extraction of characteristics is an important added value because it is the part of the treatment that translates as faithfully as possible (with the help of the histopathologist recommendations), in the form of parameters, the signature (biological/histological...) of images to be treated, something that does not take place in the deep approach (completely automated).

TABLE I. RESULTS OF CLASSIFICATION OF THYROID HISTOLOGICAL IMAGES WITH A LEARNING BASE OF 85% AND A TEST BASE OF 15%, WITHOUT AUGMENTATION IN THE BASE AND WITH AUGMENTATION.

	without augmentation			With augmentation		
	<i>SVM(RBF)</i>	<i>SVM(linear)</i>	<i>KNN(K=3)</i>	<i>SVM(RBF)</i>	<i>SVM(linear)</i>	<i>KNN(K=3)</i>
ACC	0.60	0.60	0.60	0.94	0.94	0.64
F1	0.53	0.53	0.53	0.94	0.94	0.64
Rec all	0.60	0.60	0.60	0.94	0.94	0.64
Pre	0.50	0.50	0.50	0.95	0.95	0.66

TABLE II. RESULTS OF CLASSIFICATION OF THYROID HISTOLOGICAL IMAGES WITH A LEARNING BASE OF 80% AND A TEST BASE OF 20%, WITHOUT AUGMENTATION IN THE BASE AND WITH AUGMENTATION.

	without augmentation			With augmentation		
	<i>SVM(RBF)</i>	<i>SVM(linear)</i>	<i>KNN(K=3)</i>	<i>SVM(RBF)</i>	<i>SVM(linear)</i>	<i>KNN(K=3)</i>
ACC	0.66	0.66	0.66	0.95	0.95	0.65
F1	0.60	0.60	0.60	0.95	0.95	0.64
Rec all	0.66	0.66	0.66	0.95	0.95	0.65
Pre	0.55	0.55	0.55	0.96	0.96	0.65

We can say that the strength of deep approaches emanates from the great capacity of databases, thus reinforcing learning and in conq, the adjustment of attributes (automatic), however, guest the acquisition of a large database is often a true technical obstacle how much even one tries to divert it with logical methods such as the increase of the data (rotation, scaling...).

In the final we compared our classification results with those of the literature, considering only research teams who have worked on the same thyroid pathologies as us) (Table IV).

We can conclude that our results are competitive with those of literature which is for us satisfactory. So this work (proposed tools SVM/CNN) responds to the medical problem posed and remains open to future developments (more pathologies after expansion of database).

TABLE III. RESULTS OF CLASSIFICATION OF THYROID HISTOLOGICAL IMAGES BY CNN.

	learning base of 80% and a test base of 20%	learning base of 85% and a test base of 15%
Accuracy training	87%	93%
Accuracy testing	81%	75%

TABLE IV. COMPARISON BETWEEN OUR RESULTS AND THOSE OF LITERATURE

Classification method	Number of images	Results
Closest Matching Rules (CMR) [28]	219 H&E stained thyroid histopathology images.	Accuracy = 100%
multi-classifier system A combination of SVM-L + SVM-Q + SVM-RBF + CMR and SVM-L + SVM-Q + CMR classifiers [29]	219 H&E stained thyroid histopathology images.	Accuracy of combinations = 99.54%
DCNN deep convolutional neural networks with VGG-19, ResNet-v2. [1]	11,715 H&E stained thyroid histopathology images.	Accuracy VGG-19 = 97.34%
Transfer Learning (VGGNet, ResNet, InceptionNet and DenseNet) [30]	221 H&E stained thyroid histopathology images.	DenseNet
Classification SVM, KNN, CNN.	112 histological images of the thyroid.	SVM Accuracy = 95%

V. CONCLUSION

In this work, classification methods have been proposed to identify histological images of thyroid nodules. In terms of data pre-processing, a pre-processing module including contrast enhancement and base augmentation with the rotation method was applied to process the dataset. As a result, the accuracy of the classification was improved due to the more rich and balanced experimental data available. In terms of characteristic extraction, the GLCM method and fractal dimension were used. Finally, In terms of classification, we trained the SVM classifier with two kernel types 'RBF' and 'linear', we also trained the KNN classifier with $K = 3$. SVM gave better results compared with KNN. And the in-depth learning model with CNN, the results were lower than those of SVM.

Classification results (the best of them) were competitive with those of littérature and may be improved with a bigger database for all methods.

VI. REFERENCES

- [1] Y. Wang *et al.*, "Using deep convolutional neural networks for multi-classification of thyroid tumor by histopathology: a large-scale pilot study," *Ann. Transl. Med.*, vol. 7, no. 18, pp. 468–468, Sep. 2019, doi: 10.21037/atm.2019.08.54.
- [2] L. He, L. R. Long, S. Antani, and G. R. Thoma, "Histology image analysis for carcinoma detection and grading," *Computer Methods and Programs in Biomedicine*, vol. 107, no. 3, pp. 538–556, Sep. 2012, doi: 10.1016/j.cmpb.2011.12.007.
- [3] T.-H. Song, V. Sanchez, H. E. Daly, and N. M. Rajpoot, "Simultaneous cell detection and classification in bone marrow histology images," *IEEE journal of biomedical and health informatics*, vol. 23, no. 4, pp. 1469–1476, 2018.
- [4] H. El Achi *et al.*, "Automated diagnosis of lymphoma with digital pathology images using deep learning," *Annals of Clinical & Laboratory Science*, vol. 49, no. 2, pp. 153–160, 2019.
- [5] Y. Xu *et al.*, "Deep convolutional activation features for large scale brain tumor histopathology image classification and segmentation," in *2015 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, 2015, pp. 947–951.
- [6] A. Yonekura, H. Kawanaka, V. B. Prasath, B. J. Aronow, and H. Takase, "Automatic disease stage classification of glioblastoma multiforme histopathological images using deep convolutional neural network," *Biomedical engineering letters*, vol. 8, no. 3, pp. 321–327, 2018.
- [7] E. M. Nejad, L. S. Affendey, R. B. Latip, and I. Bin Ishak, "Classification of histopathology images of breast into benign and malignant using a single-layer convolutional neural network," in *Proceedings of the International Conference on Imaging, Signal Processing and Communication*, 2017, pp. 50–53.
- [8] Y. Feng, L. Zhang, and J. Mo, "Deep manifold preserving autoencoder for classifying breast cancer histopathological images," *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 17, no. 1, pp. 91–101, 2018.
- [9] M. Z. Alom, C. Yakopcic, M. Nasrin, T. M. Taha, and V. K. Asari, "Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network," *Journal of digital imaging*, vol. 32, no. 4, pp. 605–617, 2019.
- [10] F. Ciompi *et al.*, "The importance of stain normalization in colorectal tissue classification with convolutional networks," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 2017, pp. 160–163.
- [11] C. Wang, J. Shi, Q. Zhang, and S. Ying, "Histopathological image classification with bilinear convolutional neural networks," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 4050–4053.
- [12] Z. Xu and Q. Zhang, "Multi-scale context-aware networks for quantitative assessment of colorectal liver metastases," in *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, 2018, pp. 369–372.
- [13] B. Korbar *et al.*, "Deep learning for classification of colorectal polyps on whole-slide images," *Journal of pathology informatics*, vol. 8, 2017.
- [14] L. Hou, D. Samaras, T. M. Kurc, Y. Gao, J. E. Davis, and J. H. Saltz, "Efficient multiple instance convolutional neural networks for gigapixel resolution image classification," *arXiv preprint arXiv:1504.07947*, vol. 7, pp. 174–182, 2015.
- [15] H. Sharma, N. Zerbe, I. Klempert, O. Hellwich, and P. Hufnagel, "Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology," *Computerized Medical Imaging and Graphics*, vol. 61, pp. 2–13, 2017.
- [16] Y. H. Chang *et al.*, "Deep learning based Nucleus Classification in pancreas histological images," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 672–675.

- [17] S. Otálora *et al.*, “Determining the scale of image patches using a deep learning approach,” in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018, pp. 843–846.
- [18] R. Tambe, S. Mahajan, U. Shah, M. Agrawal, and B. Garware, “Towards designing an automated classification of lymphoma subtypes using deep neural networks,” in *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*, 2019, pp. 143–149.
- [19] J. Xie, R. Liu, J. Luttrell IV, and C. Zhang, “Deep learning based analysis of histopathological images of breast cancer,” *Frontiers in genetics*, vol. 10, p. 80, 2019.
- [20] B. E. Bejnordi *et al.*, “Deep learning-based assessment of tumor-associated stroma for diagnosing breast cancer in histopathology images,” in *2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017)*, 2017, pp. 929–932.
- [21] J. Chang, J. Yu, T. Han, H. Chang, and E. Park, “A method for classifying medical images using transfer learning: A pilot study on histopathology of breast cancer,” in *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*, 2017, pp. 1–4.
- [22] F. A. Spanhol, L. S. Oliveira, P. R. Cavalin, C. Petitjean, and L. Heutte, “Deep features for breast cancer histopathological image classification,” in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 1868–1873.
- [23] Z. Gandomkar, P. C. Brennan, and C. Mello-Thoms, “MuDeRN: Multi-category classification of breast histopathological image using deep residual networks,” *Artificial intelligence in medicine*, vol. 88, pp. 14–24, 2018.
- [24] D. Wang, C. Gu, K. Wu, and X. Guan, “Adversarial neural networks for basal membrane segmentation of microinvasive cervix carcinoma in histopathology images,” in *2017 International Conference on Machine Learning and Cybernetics (ICMLC)*, 2017, vol. 2, pp. 385–389.
- [25] F. Ponzio, E. Macii, E. Ficarra, and S. Di Cataldo, “Colorectal cancer classification using deep convolutional networks,” in *Proceedings of the 11th international joint conference on biomedical engineering systems and technologies*, 2018, vol. 2, pp. 58–66.
- [26] J. A. Ozolek *et al.*, “Accurate diagnosis of thyroid follicular lesions from nuclear morphology using supervised learning,” *Medical image analysis*, vol. 18, no. 5, pp. 772–780, 2014.
- [27] H. Huang *et al.*, “Cancer diagnosis by nuclear morphometry using spatial information,” *Pattern recognition letters*, vol. 42, pp. 115–121, 2014.
- [28] J. A. A. Jothi and V. M. A. Rajam, “Effective segmentation and classification of thyroid histopathology images,” *Applied Soft Computing*, vol. 46, pp. 652–664, 2016.
- [29] J. A. A. Jothi and V. M. A. Rajam, “Automatic classification of thyroid histopathology images using multi-classifier system,” *Multimedia Tools and Applications*, vol. 76, no. 18, pp. 18711–18730, 2017.
- [30] V. G. Buddhavarapu and A. A. J. J., “An experimental study on classification of thyroid histopathology images using transfer learning,” *Pattern Recognition Letters*, vol. 140, pp. 1–9, Dec. 2020, doi: 10.1016/j.patrec.2020.09.020.
- [31] S. K. Saha, A. K. Das, and B. Chanda, “CBIR using perception based texture and colour measures,” in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, Cambridge, UK, 2004, pp. 985–988 Vol.2. doi: 10.1109/ICPR.2004.1334424.
- [32] S. Singh, D. Srivastava, and S. Agarwal, “GLCM and Its Application in Pattern Recognition,” p. 6.
- [33] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural Features for Image Classification,” *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973, doi: 10.1109/TSMC.1973.4309314.
- [34] L.-K. Soh and C. Tsatsoulis, “Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices,” *IEEE Trans. Geosci. Remote Sensing*, vol. 37, no. 2, pp. 780–795, Mar. 1999, doi: 10.1109/36.752194.
- [35] P. Shanmugavadivu and V. Sivakumar, “Fractal Dimension Based Texture Analysis of Digital Images,” *Procedia Engineering*, vol. 38, pp. 2981–2986, 2012, doi: 10.1016/j.proeng.2012.06.348.
- [36] “Chaos and Fractals New Frontiers of Science.pdf.”
- [37] “The Fractal Geometry of Nature.pdf.”
- [38] M. Hosseini and M. Zekri, “Review of medical image classification using the adaptive neuro-fuzzy inference system,” *J Med Signals Sens*, vol. 2, no. 1, p. 49, 2012, doi: 10.4103/2228-7477.108171.
- [39] D. R. Don, “Multiclass Classification Using Support Vector Machines,” p. 111.
- [40] K. Dahmane, “Analyse d’images par méthode de Deep Learning appliquée au contexte routier en conditions météorologiques dégradées,” p. 146.

New Inequality Index for Fair Multi Agent Resource Allocation

Betouil Ali Abdelatif

*Department of Computer Science,
Chadli Bendjedid, University, El-Tarf, Algeria,
PB 73, 36000.*

*Laboratory LRS, Badji Mokhtar University,
Annaba, Algeria, PB 12. 23000*

a.betouil@univ-eltarf.dz

Benmachiche Abdelmadjid

*Department of Computer Science,
Chadli Bendjedid, University, El-Tarf, Algeria,
PB 73, 36000.*

*Laboratory LRI, Badji Mokhtar University,
Annaba, Algeria, PB 12. 23000*

benmachiche-abdelmadjid@univ-eltarf.dz

Abstract— Allocating resources among a set of agents having preferences among bundles of indivisible resources is known to be NP-complete problem. In this paper we are interest to fair division that reduces inequality among agents. For this purpose, we model the problem in CSP problem and apply an algorithm to construct all complete sub-graphs between bundles with respect to XOR operator. Among these complete sub-graphs we apply a branch and bound taking into account the compute of Gini index. We propose another inequality index called MxMn index which allows us to prune during the search of allocations.

Keywords—component, formatting, style, styling, insert (key words)

I. INTRODUCTION

Fair division is a mathematical theory based on an idealization of a real problem of dividing equally goods or resources among participants (agents).

The theory gives explicit criteria for different types of fairness. Its purpose is to obtain algorithmic solution leading to a fair share (fair division), or to demonstrate their impossibility and study the properties of these shares, both in theory and practice.

Several criteria can be selected to characterize a fair division: The most fundamental criterion is the Pareto efficiency; an agreement is Pareto efficient (Pareto optimal) if there is no other feasible agreement that would make at least one agent strictly better off while not making any of the others worse off (Endriss 2010). An allocation, assigning a bundle of goods to each agent, is called envy-free if no agent strictly prefers one of the bundles assigned to another agent to their own bundle (Endriss 2010). An agent will be happiest if they receive all the goods, and they may feel entitled to at least $1/n$ of the value of this full bundle. The criterion of proportionality is satisfied when each agent believes that they received a fair share of the goods (Endriss 2010).

A fundamental model in fair division is Cake cutting (Brânzei et al 2013). The mathematicians of the Polish School in the 1940s (Banach, Knaster, Steinhaus) were the first to recognise the problem of fairly dividing a cake between several agents as a question of considerable mathematical interest (Endriss 2010).

In recent years, this problem has received significant attention in computer science (resource allocation, multiagent systems,...) (Chevalerey et al 2006).

Several works exist to study the cake cutting problem (caragiannis et al 2011) (Brams et al 2012) (Cavallo 2012) (Brânzei and Miltersen 2013) (Chen et al 2013). Externalities in the fair division of indivisible goods are considered in (Valez 2011).

Generally, in the literature, several fairness criteria such as: Pigou-Dalton principle (Dalton 1920) (Pigou 1912), Lorenz curve (Lorenz 1905),... and several inequality indices such as: Gini index (Gini 1912), Robin Hood index (Kondor 1971), Atkinson index (Atkinson 1970),... exists. The most widely used inequality index is the Gini index which is based on Lorenz Curve. In general, Lorenz curve and Gini index are used for complete allocations, in other words they need to find all possible allocations (search in all space of solutions) then compare between them.

In this paper, we propose a new index for measuring inequality between agents, the main advantage of this new index is that it can be used in the Branch and Bound without finding all allocations.

II. RELATED WORKS

In (Lumet et al 2010), authors study the problem of fairly allocating a set of indivisible goods to a set of agents having additive preferences. They propose a formal model of this problem, based on the notions of ex-ante and ex-post fairness and propose some algorithms aiming at computing optimal allocations in the sense of ex-post egalitarianism.

(Asadpour and Saberi 2007) propose the first approximation algorithm for the problem of max-min fair allocation of indivisible goods.

In (Ogryczak and Sliwinski 2002), authors introduce and analyze an alternative compromise solution concept of the conditional minimax for equitable approaches to resource allocation problems.

(Lesca and Perny 2010) propose several operational approaches for solving fair allocation problems in the context of multiagent optimization. Authors present and discuss various social welfare functions that might be used to maximize the satisfaction of agents while maintaining a notion of fairness in the distribution.

(Bouveret et al 2010) study the problem of fair division when agents have ordinal preferences over alternative bundles of goods (incomplete preferences).

In (Chevaletre et al 2009) authors present distributed mechanisms for allocating indivisible goods. They study convergence properties for such distributed mechanisms when used as fair division procedures.

In (Weller 1985) (Barbanel 1999) authors prove that for a single heterogeneous good (cake) they exist allocations which are envy-free Pareto optimal. For finding an envy-free and Pareto optimal allocation for multiple homogeneous divisible goods an algorithm is proposed in (Reijnerse and Potters 1998). The proposed algorithm is based on market clearing using Fisher's model. For the same problem (Devanur et al 2002) propose a polynomial time algorithm. In the case of 2 agents, (Brams and Taylor 1996) propose a centralized algorithm based on two methods; the first one divides proportionally each of goods between the two agents, the second method ("adjusted winner") finds a Pareto optimal allocation. For multiple agents an equitable solution which offers some level of truthfulness is proposed in (Brams et al 2006). (Zivan et al 2010) propose a centralized method for reducing untruthful manipulation in envy-free Pareto optimal resource allocation.

III. BACKGROUND

A. Constraint Satisfaction Problem (CSP)

In the literature, combinatorial problems are found as planning problems, assignment, resource allocation... These problems have been formulated as Constraint Satisfaction Problem (CSP) (Nonobe and Ibaraki 1997).

Constraint Satisfaction problem involve a set of variables, associated domain for each variable (values) and a set of constraints. The goal of a CSP is the assignment of values to its variables that will satisfy all constraints (Jones 2000).

A Constraint Satisfaction Problem consists of:

- a set of variables $X = \{X_1, \dots, X_n\}$
- a set of domains $D = \{D_1, \dots, D_n\}$ where D_i is the domain of variable X_i (a domain for each variable)
- a set of constraints $C = \{C_1, \dots, C_m\}$

A solution to a CSP is an assignment of values to variables such that no constraint is violated.

A satisfiable (consistent) problem is problem which has a solution, otherwise is termed unsatisfiable (inconsistent). Several types of constraints exist, for example: unary constraint which specifies one variable, binary constraint which specifies two variables. If in a CSP each constraint is unary or binary the nit termed a binary CSP.

A Constraint Satisfaction Problem can be represented by a constraint graph. Each node in this graph represents a variable and an arc connecting two variables iff exist a

constraint containing these variables. In general, CSPs are NP-hard (Frost 1997).

B. Measuring inequality in MARA

Let $N = \{1, \dots, n\}$ a set of agents and let $R = \{1, \dots, m\}$ a set of resources (goods). Each agent has its own preferences over goods and equipped with utility function.

An allocation A is a function mapping agents to the bundles of goods they receive with $A(i) \cap A(j) = \{\}$ for any $i \neq j$ (each good is allocated to one and only one agent).

For each allocation A we attribute a utility vector $u(A) = (u_1(A), \dots, u_n(A))$; $u_i(A)$ is the utility of agent i in allocation A .

An ordered utility vector of allocation A is a vector $u^*(A)$ which is obtained from $u(A)$ by ordering its elements in ascending manner. In other words, the first element in vector $u(A)$ is the first agent in allocation A and

the first element in vector $u^*(A)$ is the poorest agent for the same allocation.

To compare different allocations (represented by utility vectors) in terms of the inequality, several fairness criteria (Pigou-Dalton principle, Lorenz curve,...) and inequality indices (Gini index, Robin Hood index,...) are used.

Definition 1: *Pigou-Dalton transfer (Endriss 2013)* A move from allocation A to allocation A' is called Pigou-Dalton transfer if there are two agents $i, j \in N$ such that the following conditions are satisfied:

$$A(k) = A'(k) \text{ for all } k \in N \setminus \{i, j\}$$

$$|u_i(A) - u_j(A)| > |u_i(A') - u_j(A')|$$

$$u_i(A) + u_j(A) \leq u_i(A') + u_j(A')$$

In other words, Pigou-Dalton is done only between two agents i and j , such that inequality between these two agents is reduced and total utility is not diminished.

Definition 2: *Lorenz curve (Endriss 2013).* For any allocation A and any agent $k \in N$, let $L_k(A) := \sum_{i=1}^k u_i^*$. Then the vector $(L_1(A), \dots, L_n(A))$ is called the Lorenz curve of A .

According to this definition, $L_1(A)$ is the egalitarian social welfare of A , $L_n(A)$ is the utilitarian social welfare of A and $L_k(A)$ is the sum of the utilities of the k poorest agents in A .

The mean utility of A is $\mu(A) := L_n/n$.

An allocation A is Lorenz dominated by an allocation A' if $L(A) < L(A')$. We write $L(A) < L(A')$ if for all agents

$k \in N$ $L_k(A) < L_k(A')$ and that inequality is strict at least one case.

Definition 3: Gini index (Endriss 2013). The Gini index of allocation A is $G(A) := \frac{\sum_{k=1}^n k \cdot \mu(A) - L_k(A)}{\sum_{k=1}^n k \cdot \mu(A)}$

The Gini index is the most inequality index used for measuring inequality.

An inequality index is a function mapping utility vectors to a value $v \in [0, 1]$, with 0 represents perfect equality and 1 represents complete inequality.

Definition 4: Robin Hood index (Endriss 2013). The Robin Hood index of allocation A is :

$$H(A) := \frac{1}{\mu(A)} \cdot \max \{k \cdot \mu(A) - L_k(A) / k \in N\}$$

The Robin Hood index is the ratio of maximum distance between the line of perfect equality and the Lorenz curve and the mean utility. We call an agent rich if its utility exceeds the mean utility and poor otherwise.

IV. PROBLEM MODELING

Let's take a set of variables (Agents) $A = \{a_1, \dots, a_n\}$.

Each variable a_i is characterized by a set of domains (bids)

$$B = \{b_{1,i}, \dots, b_{m,i}\}.$$

For each variable a_i , we attribute for each value $b_{j,i}$ a table $TAB_{j,i}$ of size n (number of agents). For each index k such as $(k \in \{1, \dots, n\} \text{ and } k \neq i)$: $TAB_{j,i}[k]$ is the number of bids of agent a_k which are compatibles with bid $b_{j,i}$. When $k=i$: $TAB_{j,i}[k] = -1$. FiG(a) represents the data structure used in our approach.

For two agents a_i and a_l , two bids $b_{j,i}$ and $b_{f,l}$ are compatibles iff $b_{j,i} \cap b_{f,l} = \emptyset$.

The first phase of our approach is elimination of dominated bids. A bid $b_{f,i}$ is dominated by another bid $b_{j,i}$ iff

$$v(b_{j,i}) \geq v(b_{f,i}) \text{ and } |b_{j,i}| \leq |b_{f,i}|.$$

The second phase is combination phase: At first, let's take randomly 2 agents a_i and a_l from A . For each bid $b_{j,i}$ of agent a_i , we compute the number of bids of agent a_l which are compatibles with this bid and we put the result in $TAB_{j,i}[l]$. If $TAB_{j,i}[l] = 0$ then the bid $b_{j,i}$ is eliminated (no bid is compatible with it) (see FiG(b)). We do the same process for agent a_l .

In the next iteration, we choose another agent a_t from A . The combination phase is executed for the set of choosed agents (2 by 2). For each agent $a_i \neq a_t$ already choosed and for each bid $b_{j,i}$ of agent a_i we verify: if

$TAB_{j,i}[t] = 0$ then the bid is eliminated, we propagate this information to the set of neighbors of this bid (see FiG(c)) (all compatible bids) for decrementing the counter of compatible bids with this agent (see FiG(d)). A solution (allocation) is represented by a complete sub-graph (see FiG(e)).

At the end of this phase we obtain all possible solutions (all possible sub-graphs) (see FiG(e)).

Our proposed algorithm is as follows

Algorithm Found_Solution

$n := |A|$ (nb agents)

$Q := \{\}$

While ($|Q| < n$) do

 Select a_i from A

 If Make-Compatibility (a_i, Q)=true

 Add (a_i, Q)

 Remove (a_i, A)

 Search_Complete_Subgraphs

END

Make_Compatibility (a_i, Q)

For each $a_j \in Q$ do

 For each $b_{k,j}$ of a_j do

 cpt:=0

 For each $b_{l,i}$ of a_i do

 If compatible ($b_{k,j}, b_{l,i}$)

 Make-arc ($b_{k,j}, b_{l,i}$)

 Neighbor ($b_{k,j}, b_{l,i}$):=true

 cpt:=cpt+1

$TAB_{l,i}[j] := TAB_{l,i}[j] + 1$

$TAB_{k,j}[i] := cpt$

 If ($TAB_{k,j}[i] := 0$)

$Q' := Q - \{a_j\}$

 Propagation ($b_{k,j}, Q'$)

END

Propagation ($b_{k,j}, Q'$)

For m:=1 to n

If $TAB_{k,j}[m] > 0$

For each $b_{f,m}$ of a_m do

If Neighbor ($b_{k,j}, b_{f,m}$):=true

$TAB_{f,m}[j] := TAB_{f,m}[j] - 1$

Neighbor ($b_{k,j}, b_{f,m}$):=false

If $TAB_{f,m}[j] = 0$

$Q' := Q' - \{a_m\}$

Propagation ($b_{f,m}, Q'$)

Remove ($b_{k,j}$)

END

Example

Let's take an example with 3 agents, each agent has 3 bids.

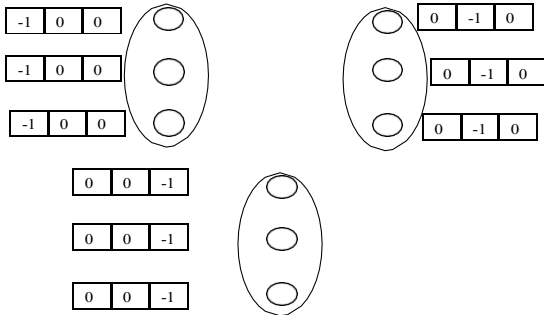


FIG (a)

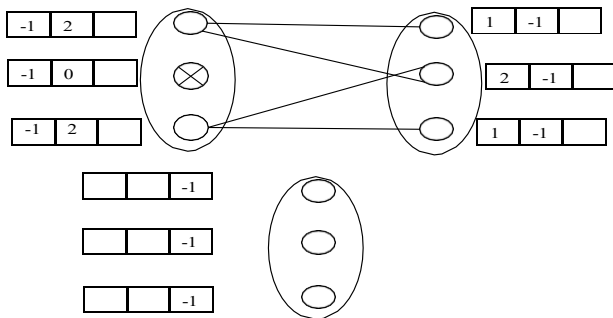


FIG (b)

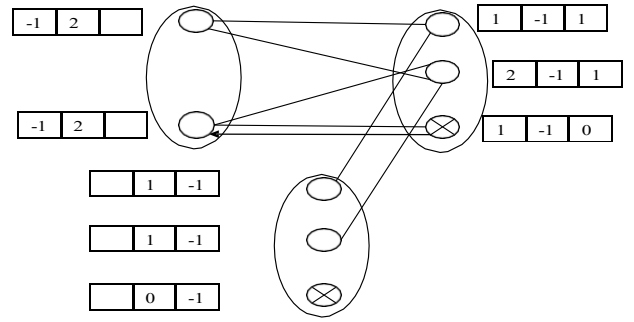


FIG (c)

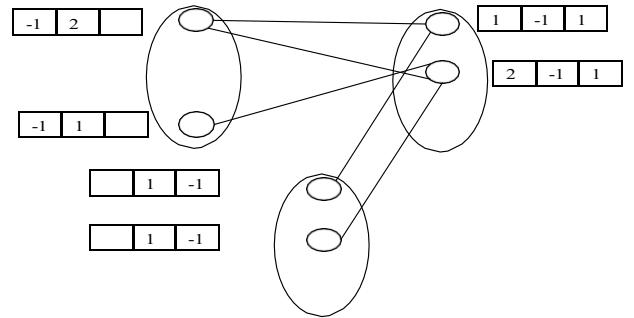


FIG (d)

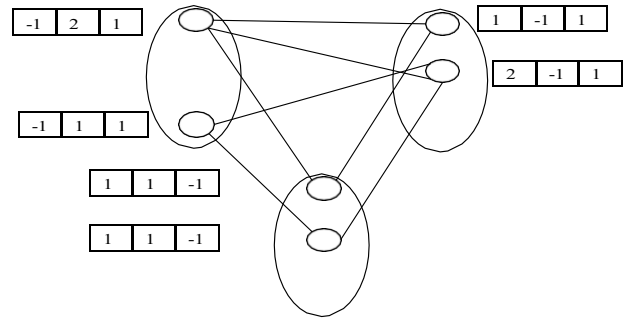


FIG (e)

After we found all solutions, we attribute for each solution its Gini index (calculated as definition 3), the solution which has a minimum Gini index is selectionned as fair allocation. In our example, we have 2 possibles allocations A and B . The utility vectors of these allocations are: $u(A)=(7,4,9)$ and $u(B)=(1,8,4)$.

The ordered utility vectors of A and B are: $u^*(A) = (4,7,9)$ and $u^*(B) = (1,4,8)$.

Now, we compute the Gini index for each allocation.

For allocation A :

$L_1(A)=4$, $L_2(A)=11$, $L_3(A)=20$. The mean

utility of A is: $\mu(A) = 20/3$

The Gini index of A is:

$$GI(A) = \frac{(1 \cdot 20/3 - 4) + (2 \cdot 20/3 - 11) + (3 \cdot 20/3 - 20)}{(1 \cdot 20/3) + (2 \cdot 20/3) + (3 \cdot 20/3)} = \frac{4,99}{39,99} = 0,12$$

For allocation B :

$$L_1(B) = 1, L_2(B) = 9, L_3(B) = 13. \text{ The mean utility}$$

of B is: $\mu(B) = 13/3$

$$GI(B) = 0,26.$$

In this case, allocation A is equitable than allocation B because $GI(A) < GI(B)$ and it is chosen as the solution of our problem.

Using geni index to find the best fair allocation necessitates finding all possible allocations (search in all space of solutions) then compute a geni index for each allocation and the allocation with minimum gi is returned.

we propose another approach which find at first one solution then it find another solution iff its more equitable. This solution is based on novel inequality index called *MxMn* index.

We apply the *MxMn* index on the ordered utility vector of an allocation A , $u^*(A)$. We note $u_i^*(A)$ the i^{th} element in this vector.

The *MxMn* index is defined as follows:

Definition (MxMn index): The *MxMn* index of allocation A , represented by an ordered utility vector $u^*(A) = (u_1^*(A), \dots, u_n^*(A))$, is defined as the sum of difference between two successif elements in $u^*(A)$.

$$\text{Formally } MxMn(A) = \sum_{k=1}^{n-1} (u_{k+1}^*(A) - u_k^*(A))$$

The allocation which has a minimum *MxMn* index between all allocations is considered as the fair allocation.

Proposition 1: The allocation with minimum *MxMn* index is considered as the fair allocation.

Proof: According to definition of *MxMn* index which is the sum of inequalities between agents, intuitively, an allocation with a minimum *MxMn* index represents an allocation which has the minimum sum of inequalities between agents. So its considered as the fair allocation.

In our example, the *MxMn* index of allocation A is: $MxMn(A)=5$ and the *MxMn* index of allocation B is: $MxMn(B)=7$. In this case, the allocation A is the fair allocation.

According to definition of *MxMn* index, it needs a complete allocations and finding all possible solutions (allocations). So we need to compute *MxMn* index with a manner in which we can use it when searching a solution. Our aim is to not use the *MxMn* index in dissociation in the search process but we aim to use it during (INTEGRATED) the search process.

It is enough to note that the *MxMn* index can be calculated in a more direct way without having the utilities of all the agents.

Proposition 2: The *MxMn* of allocation A represented by an ordered utility vector $u^*(A) = (u_1^*(A), \dots, u_n^*(A))$ is

the difference between the maximum utility and the minimum utility in this allocation (vector). Formally

$$\sum_{i=1}^{n-1} (u_{i+1}^* - u_i^*) = u_n^* - u_1^*.$$

Proof : if we suppose that the ordered utility vector $u^*(A)$ represents the distances between points. So

$u_{i+1}^*(A) - u_i^*(A)$ is the distance which separates the point $u_{i+1}^*(A)$ and the point $u_i^*(A)$. The sum of all distances between each two successive points is the distance which separates the last point and the first point.

Let's take an example with 4 agents as the following figure

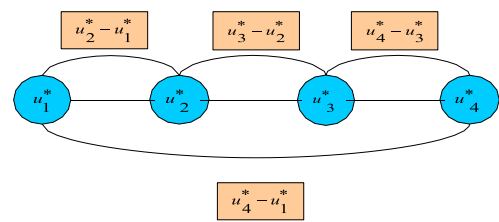


Fig. 1. MxMn Example

The *MxMn* of this allocation A represented by the ordered utility vector $u^*(A) = (u_1^*(A), \dots, u_4^*(A))$ is computed as follows:

$$MxMn(A) = (u_2^* - u_1^*) + (u_3^* - u_2^*) + (u_4^* - u_3^*)$$

which is equal to $u_4^* - u_1^*$.

Definition (MxMn index): The *MxMn* index of allocation A represented by an ordered utility vector $u^*(A) = (u_1^*(A), \dots, u_n^*(A))$ is the difference between the maximum utility and the minimum utility in this allocation. $MxMn(A) = u_n^*(A) - u_1^*(A)$

This approach based on *MxMn* index search at first one solution S and computes its *MxMn* index.

- 1-cpt:=0
- 2-choose randomly a bid b from the first agent a_i ,
- 3-cpt:=cpt+1
- 4- choose a compatible bid b' from an other agent a_j
- 5-cpt:=cpt+1;
- 6-if $(u_{a_i}(b) > u_{a_j}(b'))$ then begin

```

7-          max :=  $u_{a_i}(b)$ 
8-          min :=  $u_{a_j}(b')$ 
9-          End
10-else    begin
11-        max :=  $u_{a_j}(b')$ 
12-        min :=  $u_{a_i}(b)$ 
13-compute  $MxMn(S') = \max - \min$ 
14-if  $MxMn(S') > MxMn(S)$  then prune
15-else begin
16-    While (cpt < n) do
17-        begin
18-            For each agent  $a_k$ 
19-                Begin
20-                    Choose a bid  $b''$  compatible with all
21-                    choosed bids
22-                    Cpt:=cpt+1
23-                    If  $(u_{a_k}(b'') > \max)$  then begin
24-                        goto
25-                    end
26-                    if  $(u_{a_k}(b'') < \min)$  then begin
27-                        min :=  $u_{a_k}(b'')$ 
28-                    end
29-                    goto(13)
30-                end
31-            end
32-        end
33-    end
34-    End

```

allocated after the bid $b_{j,i}$ (noted by $m - |b_{j,i}|$) greater than or equal the remaining number of agents (noted by $n - |E|$). So if $m - |b_{j,i}| < n - |E|$ then eliminate the bid $b_{j,i}$.

Proposition 3: *The allocation obtained at the end of our approach is the more fair possible allocation.*

Proof: In our approach, at first we obtain an allocation A , we compute its MxMn index. The aim is finding an allocation A' more fair than A by combining bids. For each combination, we compute the max and the min for this partial allocation noted by P , these values are updated at each combination. We test if the MxMn index of this partial allocation P is greater or equal than MxMn index of allocation A , formally $MxMn(P) \geq MxMn(A)$. According to proposition 2, $MxMn(P) = \max - \min$ then we prune (according to proposition 1) because it is unnecessary to continue the combination of bids since the MxMn index of the complete allocation it will be greater than or equal the MxMn index of the partial allocation P , else we continue the combination phase. If we obtain a new complete allocation A' (combining without pruning) this allocation is fair than the previous allocation A . In this case $A \leftarrow A'$ and $A' \leftarrow \emptyset$ and the process is repeated until no new complete allocation is possible. The last obtained allocation is the best possible fair allocation.

The advantage of using MxMn index is: more we found new complete allocations more we prune faster and more because each new complete solution has a MxMn index less than the previous allocation.

At the end of this approach, we obtain a solution with minimum MxMn index. The main advantage of this approach is reducing the possible solutions for not searching in the all entire space of solutions.

Before each combination between bids we try to filtering bids for elimination of bids which are not useful.

Filtering bids: let's take n the number of agents and m the number of resources. In an allocation A . E is the set of agents belong to this allocation. The choice of bid $b_{j,i}$ must be made of manner that the number of resources not yet

Object Detection and Classification in Videos via Optical Flow and VGGNet

Merouane Labeni

Department of Computer Science
Mohamed seddik benYahia University
Jijel, Algeria
labmer@gmail.com

Chaouki Boufenar

Department of Computer Science
Algiers I University
Algeria
boufenarc@gmail.com

Mokhtar Taffar

Department of Computer Science
Mohamed seddik benYahia University
Jijel, Algeria
bentaffar@yahoo.fr

Abstract— Object detection and tracking in video streams are a vital field of research that gives appropriate information about objects and their trajectories. An interesting technique to fix these goals is using off-the shelf motion estimation such as optical flow with an image classification approach recently gives well results called ConvNets or CNNs for Convolutional Neural networks. This paper presents a single-fastest object detection and tracking based on a combination of a proposed detection-tracking algorithm and a ConvNet model. An implementation of proposed technique has been performed using python programming language.

Keywords— Image/Video Processing, Optical Flow Analysis, Object Detection, Object Tracking, Deep Learning.

I. INTRODUCTION

Image/video processing is the field of treatment and handling images. There are different properties of an image/frame such as color, pixels correlation and motion, object number, their placements, and other elementary details to extract image features such as edges, objects and contours.

Convolutional Neural Networks (also called ConvNets) is one kind of feed-forward neural networks. It is an efficient recognition process which is widely used in pattern recognition and image processing. It has many features such as simple structure, less training parameters and adaptability. It has become a hot topic in voice analysis and image recognition.

The pre-trained VGGNet [1] ConvNet -used in this paper for image identification and classification- is a deeper convolutional neural network that is trained on the ImageNet database [2].

Figure 1 shows the general architecture of object detection, identification and tracking. The proposed technique consists on a combination of a set of ideas, which are: optical flow analysis, isobarycenter computing, Euclidean distance, slicing, classification and tracking. Section II describes background information, section III lists related works, section IV explains proposed method, section V describes technique illustration and analysis.

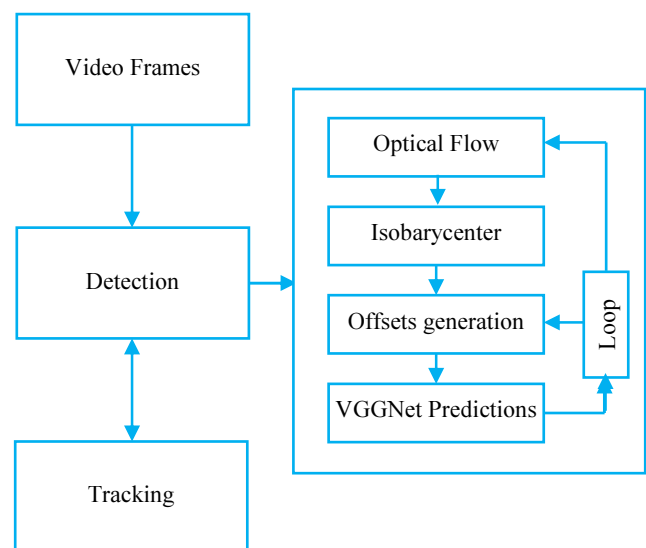


Fig. 1. The whole architecture of the technique.

II. BACKGROUND INFORMATION

A. Python OpenCV

OpenCV is a C++ tool with functions that cover many fields of image and video processing [3]. OpenCV-Python is a Python wrapper around original C++ implementation. So it is an adequate tool for fast prototyping for many computer vision applications [4].

B. Optical Flow

Optical flow or optic flow is the image motion of objects as the objects, camera or scene changes between two successive images [3].

Two assumptions are imposed for using optical flow:

- First, the pixel intensities rest unchanged between two frames.
- Second, the pixels in neighborhood have similar motion.

A pixel $I(x, y, t)$ in initial frame (at time t) moves by distance (dx, dy) in next frame and that after dt time. Subsequently, and for the same pixel $I(x, y, t)$ where its intensity rests unchanged, we can express [4]:

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (1)$$

OpenCV implements several optical flow methods, in our work we use *calcOpticalFlowFarneback* method based on [5]. That is considered a best method for computing optical flow vector [3].

C. Isobarycenter and Euclidean Distance

Isobarycenter

With a set of points $A_1, A_2, \dots, A_n \in E$ and a real number λ , it exists a single point $H \in E$ verifying [6]:

$$H = \lambda \sum_{i=1}^n A_i \quad (2)$$

Where $\lambda = \frac{1}{n}$.

This point is called *Isobarycenter* of A_1, A_2, \dots, A_n points affected by 01 as coefficient, (i.e., the $\frac{1}{n}A_1 + \frac{1}{n}A_2 + \dots + \frac{1}{n}A_n$ point).

Euclidean Distance

Using the Euclidean norm $\|(x, y)\| = \sqrt{x^2 + y^2}$ the distance between two points $P = (x, y)$ and $P' = (x', y')$ is [7]:

$$d(P, P') = \|P - P'\| = \sqrt{(x - x')^2 + (y - y')^2} \quad (3)$$

That is called the Euclidean distance

D. Slicing

Slicing is an operation that allows us to extract an object from an image [8]. An image I can be represented by a 2D array of size $n_1 \times n_2$ (gray scale). A color image is represented by a 3D array of size $n_1 \times n_2 \times 3$ (i.e., *width* \times *height* \times *channels*) [9]. In these two cases a slicing operation can have –within Numeric Python library– one of two following forms:

- $Object = image[a:b, c:d]$ in gray scale;
- $Object = image[a:b, c:d, :]$ in color image.

Where $a \geq 0$, $b \leq n_1$, $c \geq 0$ and $d \leq n_2$.

E. Video Processing and Object Tracking

Videos are a sequence of images and hence dealing with videos is similar to how we dealt with images with a few exceptions (i.e., temporal dimension presents in video sequences), tracking objects in a video is a field of video processing [10], given the primary set of points RoI (Region of Interest), a tracker calculates the motion of these points using the direction of change in the next frame (like the KCF [11] tracker). In every consecutive frame, we try to look for the same set of points in the neighborhood. Once the new positions of these points are identified, we can move the bounding box (i.e., object container) over the new set of points [10].

F. Artificial Intelligence, Deep Learning and ConvNets

Modern artificial intelligence systems and machine learning algorithms have revolutionized approaches to scientific and technological challenges in a variety of fields. We can observe remarkable improvements in the quality of

state-of-the-art computer vision, natural language processing, speech recognition and other techniques [12]. Deep learning DL is a specific subset of ML –Machine Learning– methodologies using ANNs –Artificial Neural Networks– strongly inspired by the connection of neurons located in the brain of human [13].

CNNs –Convolutional Neural Networks– consist on different types of layers (i.e., convolutional, down-sampling and fully connected layers), using these layers, deeper architectures are proposed for classification tasks like: AlexNet (2013), VGGNet (2014), ResNet (2015), SqueezeNet (2016), DenseNet (2017), FractalNet (2017), and others [9]. The ConvNet –CNN– network learns to classify images into categories and to do the final score, each layer learns important features from the previous layers. First layers capture lower features such as colors and edges, and deeper layers compose these lower features into higher features such as parts and shapes. Many companies use these technologies for analyzing and reporting huge amounts of data like GoogleTM, Micro-Soft, Amazon, FaceBook, and others [14].

G. Keras Library

Keras is the important library of Tensorflow v2.0 [15] used for implementing machine learning and deep learning models. It performs modular-building models (i.e., CNNs and RNNs architectures) by stacking different types of layers. Keras is useful for fast prototyping. It works with convolutional and recurrent layers and integration of both, and runs seamlessly on CPUs or GPUs [16]. The VGGNet [1] model used in this paper for classification task is one of the Keras pretrained models.

III. RELATED WORKS

These last years, a lot of works are proposed for object identification and classification with help of image/video processing and deep learning methods, notably in recent ConvNets –CNNs architectures that got an important enhancements, which allow them to surpass the standard computer vision techniques in many computer vision areas [17].

Recently, an interesting ConvNet model is proposed, which is called R-CNNs (for Regional Convolutional Neural Networks), this model combines two tasks that are detection and classification [18]. However, in its initial form, R-CNN network is computationally expensive during training and evaluating processes. To fix this problem Fast R-CNN and SPPnet [19] have been proposed, the idea consists on the introduction of sharing convolutions across proposals. Authors In [20] have proposed the addition of a Region Proposal Network (RPN), which uses convolutional features in the detection process directing to Faster-RCNN. Moreover, authors in [21] have proposed the Region-based Fully Convolutional Networks (R-FCNs), which is based on fully convolutional architectures, like FCN [22] –go further and avoid the per-proposal evaluation of Fast and Faster R-CNN. This change results in a significant speedup over Faster R-CNN on the PASCAL VOC datasets [21].

Some studies like [23] proved that the use of sharing convolutions across proposals can be extremely efficient for object identifications goals. In this work the principal goal is to locate and classify Friesian cattle in areal images. The deep architecture used to fix this challenge is the R-CNN

adapted from the VGGNet-M-1024 network used as complement of several other network architecture proposals [24]. In another approach called YOLO described in [25] the authors formulate object detection as a regression problem for spatially separated Bounding boxes (i.e., objects containers) and related probabilities of different classes, this model imposes rigorous spatial constraints on patch (i.e., bounding box) predictions as each grid location only predicts two boxes and can only have a single class. This constraint influences on the model performance by limiting the number of nearby objects that the model can predict. In [26] an algorithm based YOLO [25] with GMM -Gaussian Mixture Model- was proposed, the idea consists on using the concepts of deep learning, this algorithm has given good accuracy for feature extraction and then object classification.

Authors in [27] have proposed an approach called SqueezeNet that represents a very small CNN network. This model uses a bottleneck approach which it allows to reach the AlexNet [28] level accuracy in detection process.

Comparing the model proposed in this paper to other detection systems like [23][25-28] we can emphasize the main difference, adopted by our approach, that consists on the use of the VGGNet[1] ConvNet power and some mathematical concepts to allow reasonable object detection and tracking. However our approach gives alternate and weaker results in the case of a multi-detection and multi-tracking scenarios.

IV. PROPOSED TECHNIQUE

The methodology proposed in this paper based on an integration of a proposed algorithm for “object detection and tracking” and a pretrained ConvNet for “object identification and classification” to provide better quality and accuracy for single-fastest object identification and tracking in video frames. A deep learning model called VGGNet [1] (that represents a pretrained model of ConvNet architectures) is used to classify detected object, we then pass its classification back to use as object label to provide information about detected object. A tracking algorithm takes place and waits for object classification confidence that surpasses an acceptance threshold for starting object tracking.

The technique includes these steps: Optical flow analysis, offsets computing based on isobarycenter concept, extract sub-image using slicing, VGGNet [1] predictions, and object tracking. Thus, the proposed technique to detect, classify and track object and its illustration will be presented in the rest of the paper.

A. Optical Flow Analysis

We use the *calcOpticalFlowFarneback* method based on [5] for computing optical flow generated between frames (see Fig. 2), this method is used to detect object which is in motion across frames using a camera or video records, with the constraint that the object (foreground) is faster than the camera (background), for that we take $(n+1)$ frames and compute the optical flow between each doublet of frames, by extracting the fastest pixel that has the fastest motion (represented by either dx in X-axis or dy in Y-axis) for each generated flow, we obtain n fastest pixels, then we calculate their isobarycenter.

```
gray = cv2.cvtColor(im,cv2.COLOR_BGR2GRAY)
# compute flow
flow = cv2.calcOpticalFlowFarneback(old_gray,gray,None,\
                                     0.5,3,15,3,5,1.2,0)

old_gray = gray
.....
for i in range(w):
    for j in range(h):
        if (abs(flow[j,i,0]) == fastestmotion or \
            abs(flow[j,i,1]) == fastestmotion ):
            print('x=',i,'y=',j,'dxdy=',flow[j,i,0],flow[j,i,1])
            fastestpixels.append((i,j,fastestmotion))
```

Fig. 2. Code for optical flow computing and fastest pixels extracting.

B. Isobarycenter and Offsets Computing

With help of equation (2) we compute the isobarycenter of n fastest pixels extracted in the previous step. After computing the isobarycenter with help of equation (2), we calculate the sum of Euclidean distances using equation (3) between the isobarycenter and all the fastest pixels extracted above then we it apply a normalization step that allows a reasonable detection. The normalization step consists on scanning the sum for the first pass (this gives first offset): right-to-left, and second pass (this gives second offset): left-to-right applying an exponential formula (4), we suppose that the $sum = abcd$ (a four digit number) and i is initialized to three :

$$\begin{cases} d \cdot e^i + c \cdot e^{i-1} + b \cdot e^{i-2} + a \cdot e^{i-3}, & \text{right to left} \\ a \cdot e^i + b \cdot e^{i-1} + c \cdot e^{i-2} + d \cdot e^{i-3}, & \text{left to right} \end{cases} \quad (4)$$

The goal of the normalization step is decreasing the great sums and increasing the small ones. The results of this step are two offsets used for bounding boxes definition (for both detection and tracking). Figure 3 resumes this process.

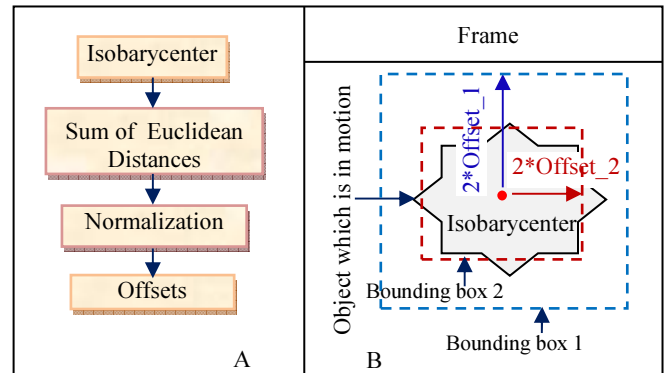


Fig. 3. A: Offsets computing, B: Object detection.

C. Extract Sub-Images Using Slicing

In this step we use slicing (see Fig. 4) for determining the object (sub_images) boundaries based on each offset computed previously, using two offsets we allows to decrease the execution number of the detection process (i.e., by giving another opportunity for object identification before restarting object detection). In our case we use 3d-array images (video color frames) then a slicing operation through these images we gives a 3d-array sub-images.

```

while True:
    k=0
    img=im
    coordvec, im=coordinates(n, w, h)
    for i in range(2):
        x1, y1, x2, y2=coordvec[i][0], coordvec[i][1], coordvec[i][2], coordvec[i][3]
        a, b, c, d=x1, y1, x2-x1, y2-y1
        box=img[y1:y2, x1:x2, :]
        cv2.imwrite('img'+str(k)+'.png', box)

```

Fig. 4. Code for bounding boxes extraction using slicing.

D. The VGG19 model

In this paper we use a variant of VGGNet [1] model called VGG19 (see Fig.5) for predicting extracted sub_images (i.e., detected objects).

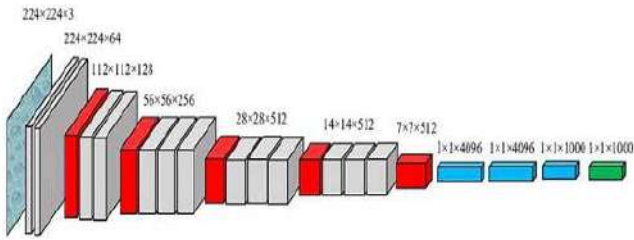


Fig. 5. VGG19 model [29].

E. Object Tracking

If the object classification confidence (explained by a probability) returned by the VGGNet [1] model -using the first offset- surpasses the threshold (see Fig. 6) the system starts tracking by calling the *tracker()* function (this last has five parameters: actual bounding box boundaries, actual frame, fastest pixels number, frame width and height) and shows a classification label in the object front that represents its VGGNet [1] identification. Else we use the second offset for creating a new bounding box (sub-image) and starting new attempt for object identification, if the classification confidence is lower than threshold we restart the detection process (else the system starts tracking). During tracking our model tracks the fastest object, hence in a multi object scenario the detection can pass to another object in the video, for this we restart detection at each *m* times interval.

```

pr=vgg19('img'+str(k)+'.png')
attempt+=1# trials
if pr[0] > threshold :
    tracker(coordvec, im, n, w, h)
    break
k+=1

```

Fig. 6. Code for object detection and tracking.

In this step we use our detection algorithm for object tracking based on optical flow analysis, by activating the bounding boxes definition function (i.e., the *coordinates()* function mentioned in Fig. 4) along frames and showing its new position across frames.

V. TECHNIQUE ILLUSTRATION AND ANALYSIS

For illustrating the value of our technique we execute it in three different scenarios with help of videos records (dynamic camera) and a notebook camera (static camera), these scenarios are based on fastest pixels number *n*, the first scenario parameters are *n* =5, 10 and 15, for the second one the parameters are *n*=5 and *n*=15. For both first and second

scenario the threshold is fixed at 0.5 (i.e., 50%). In the third scenario we choose a set of targets (objects in frames) and we increase the threshold at 0.75 (75% of confidence).

For all scenarios we show the attempts (Trials) number until object successful identification (i.e., identification that surpasses acceptance threshold).

- Scenario 1 : Using video records

1) *n*=5:

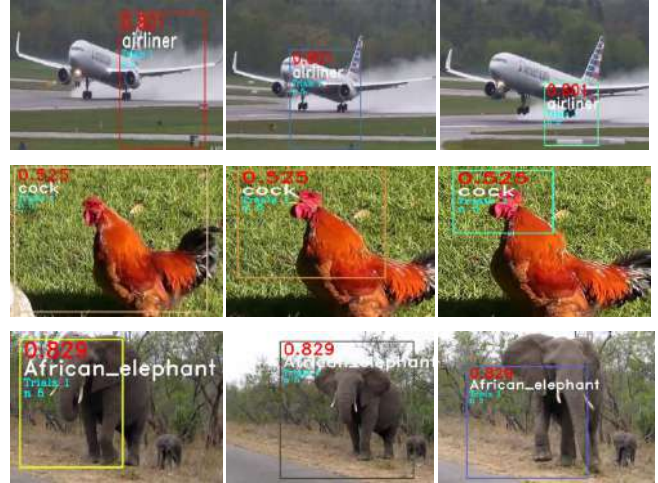


Fig. 7. Detection of airliner (80.1%), cock (52.5%) and elephant (82.9%).

1) *n*=10:

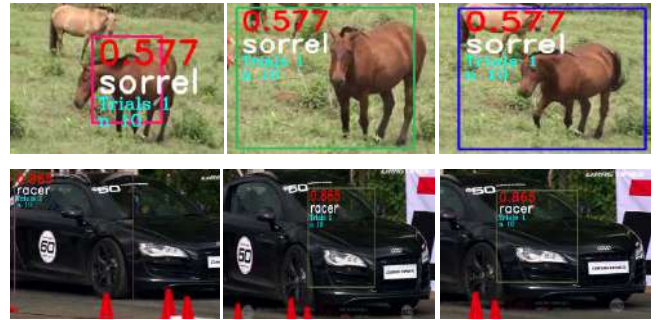


Fig. 8. Detection of sorrel (57.7%) and racer (86.5%).

2) *n*=15



Fig. 9. Detection of pickup (53.5%), cheetah (79.8%) and timber wolf (60.7%).

- Scenario 2 : Using notebook camera

1) $n=5$



Fig. 10. Detection of soccer_ball (99.4%), iPod (51.1%) and notebook (66.1%).

2) $n=15$



Fig. 11. Detection of soccer_ball (95.7%) and sunglass (56.2%).

For both first and second scenario the technique has given well results for object detection, the fastest pixels number n has its utility in the first scenario when background (camera) moves with foreground (object), the great values for n increasing object detection accuracy and object fast detection. In the second scenario small n (e.g., $n=5$) is satisfactory for a good detection. Therefore the object tracking (colored rectangle) has given acceptable results in the two scenarios.

- Scenario 3: Testing the technique capacities

In this scenario we increase detection threshold to 75% and we execute the detection process for each target then we score the target successful identifications under threshold and its identification higher than threshold. We double the n parameter when the background speed is near the foreground speed (i.e., object speed). In these scores each detected object is followed by the trials number until acceptable detection and the associate percentage of confidence.

Figure 12 shows that the technique gives well results when the acceptance threshold has a higher value.

For thresholds that equal or higher then 0.9 (90%) we should use a TensorFlow [15] GPU version, CPU (like our case) version allows a limited detections' number (16 trials) with a slow detection.

Cock: True (7 Trials)
 cock 2 0.517 (51.7%)
 cock 4 0.514 (51.4%)
 cock 6 0.746 (74.6%)
 cock 7 0.824 (82.4%) Ok

Cheetah: True (5 Trials)
 cheetah 2 0.532 (53.2%)
 cheetah 5 0.818 (81.8%) Ok

Taby: True (2 Trials)
 tabby 1 0.451 (45.1%)
 tabby 2 0.757 (75.7%) Ok

Airliner: True (2 Trials)
 airliner 2 0.869 (86.9%) Ok

Duck: False
 ptarmigan 8 0.771 (77.1%)

Kite: False
 kite 4 0.382 (38.2%)
 stingray 6 0.95 (95%)

African_elephant: True (3 Trials)
 African_elephant 3 0.864 (86.4%) Ok

Pickup: False
 pickup 1 0.375 (37.5%)
 pickup 2 0.098 (9.8%)
 pickup 9 0.198 (19.8%)
 pickup 15 0.263 (26.3%)
 Segmentation fault
 (core dumped) at 17th Trial



Fig. 12. Third scenario scores.

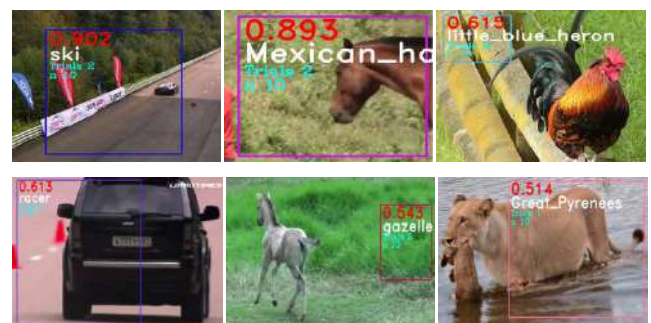


Fig. 13. Erroneous detections with dynamic camera.

Figures 13 and 14 show some system erroneous detections due to the technique constraint overtaking (i.e, background fastest than foreground) or to the VGGNet [1] erroneous predictions.



Fig. 14. Erroneous detections with static camera.

VI. CONCLUSIONS

In this paper an approach using python programming language has been developed. Python is an amusing language that contains a rich library (i.e, Matplotlib, Numpy, Tensorflow, Keras, OpenCV etc.) which can be used to detect, classify and track objects when object detection and tracking are basic data for security tasks.

This work proves that object detection, identification and tracking can be accomplished based on optical flow handling, mathematical concepts, and deep learning methods. This paper has showed, for both detection and tracking goals, that optical flow analysis and ConvNet architectures are well suited towards these goals.

REFERENCES

- [1] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition", CoRR, vol. abs/1409.1556, 2014.
- [2] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. "Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition*, CVPR 2009, *IEEE Conference on*, pages 248–255, IEEE, 2009.
- [3] J. E. Solem, *Programming Computer Vision with Python*, Creative Commons, United States License, USA, 2012.
- [4] A. Mordvintsev, K. Abid, *OpenCV-Python*, Tutorials Documentation, Release 1. Nov 05, 2017.
- [5] G. Farneback. "Two-frame motion estimation based on polynomial expansion", In *Proceedings of the 13th Scandinavian Conference on Image Analysis*, pages 363–370, 2003.
- [6] C. Deschamps, A. Warusfel, F. Moulin, J. F. Ruaud, A. Miquel, J. C. Sifre, "Mathematics All-In-One", French edition, "Mathématiques Tout-En-Un", Dunod, Paris. ISBN 978 2 10 00 7944 5, 2^e edition 2003.
- [7] F. Liret, "Math in Practice", French edition, "Maths en pratique", Dunod, Paris, ISBN 2 10 049629 8, 2006.
- [8] B. Seppke, *Image Processing with Python: An introduction to the use of Python, NumPy, SciPy and matplotlib for image processing tasks*, Hamburg university, Germany, 2013.
- [9] D. Stutz, *Understanding Convolutional Neural Networks*, Seminar Report, August 30, 2014.
- [10] S. Kapur, *Computer Vision with Python 3*, Packt Publishing Ltd, Prod. Ref. 1210817, Livery Place, 35 Livery Street, Birmingham B3 2PB, UK, August 2017.
- [11] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence* (TPAMI), vol. 37, no. 3, pp. 583–596, 2015.
- [12] P. Goldsborough, "A Tour of TensorFlow", Pro-seminar Data Mining, arXiv:1610.01178v1 [cs.LG] 1 Oct 2016.
- [13] A. Golnari, *Introduction to Deep Neural Networks with Keras*, ResearchGate, February 2019.
- [14] A. Gulli, S. Pal, *Deep Learning with Keras- Implement neural networks with Keras on Theano and TensorFlow*, Packt Publishing, Birmingham – Mumbai, 2017.
- [15] The TensorFlow website. [Online]. Available at : <https://www.tensorflow.org/>
- [16] K. Sathrey, *Tutorial on Keras*, CAP 6412 - Advanced Computer Vision, Spring 2018.
- [17] A. Krizhevsky, I. Sutskever, and G. E. Hinton. "Imagenet classification with deep convolutional neural networks", in *F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.
- [18] R. B. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation", CoRR, abs/1311.2524, 2013.
- [19] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition", arXiv preprint arXiv:1406.4729, 2014.
- [20] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", ArXiv 2015, pages 1–10, 2015.
- [21] J. Dai, Y. Li, K. He, and J. Sun, "R-FCN: object detection via region-based fully convolutional networks", CoRR, abs/1605.06409, 2016.
- [22] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3431–3440, 2015.
- [23] W. Andrew, C. Greatwood & T. Burghardt, "Visual Localisation and Individual Identification of Holstein Friesian Cattle via Deep Learning", in *IEEE International Conference of Computer Vision Workshop ICCVW*, 2017, pp. 2850–2859.
- [24] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, "Return of the devil in the details: Delving deep into convolutional nets", in *British Machine Vision Conference*, 2014.
- [25] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, "You only look once: unified, real-time object detection", *IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
- [26] A. P. Jana, A. Biswas, M. Mohana, "YOLO based Detection and Classification of Objects in video records", *IEEE International Conference On Recent Trends In Electronics Information Communication Technology*, (RTEICT), 2018.
- [27] F. N. Iandola, M. W. Moskewicz, K. Ashraf, S. Han, W. J. Dally, and K. Keutzer, "Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <1mb model size". CoRR, abs/1602.07360. 2016.
- [28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks", In the Proceedings of 25th International Conference on Neural Information Processing Systems - Volume 1, NIPS'12, pages 1097–1105, USA, 2012.
- [29] A. V. Ikechukwu, S. Murali, R. Deepu, R. C. Shivamurthy, "ResNet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images", *Global Transitions Proceedings 2*, pages 375–381, ScienceDirect, 2021.

Phishing website detection using deep learning and machine learning.

1st Rania Zaimi

Department of Computer Science
LRS Laboratory, Badji Mokhtar
University
Annaba, Algeria
raniazaimi25@gmail.com

2nd Mohamed Hafidi

Department of Computer Science
LRS Laboratory, Badji Mokhtar
University
Annaba, Algeria
mhafidi@yahoo.fr

3rd Mahnane Lamia

Department of Computer Science
LRS Laboratory, Badji Mokhtar
University
Annaba, Algeria
mahnane_lamia@yahoo.fr

Abstract— the website phishing detection is turning into an important analysis space, attracting augmented focus owing to the growth of several phishing attacks. What is more, as a result of attackers inventing numerous manners, phishing spotting has become a principal preoccupation of developers. A lot of phishing detection schemes are introduced in the literature, such as URL analysis, visual similarity, list-based approach, machine learning...etc. each of them has its benefits and its disadvantages. Recently, deep learning has seen great interest in the phishing detection field and has achieved considerable results.

During this paper, we propose a website phishing detection approach based on URL features and using three classifiers: Convolutional Neural Network (CNN), Logistic Regression, and Decision tree, we also apply various performance measures to evaluate the effectiveness of the proposed solution and compare between the deep learning model (CNN) and the machine learning classifiers. The experiment results show that the CNN model outperformed the machine learning classifiers (Logistic Regression (LG), and Decision tree (DT)) in terms of all performance metrics used and reached the accuracy value of 96.19 %.

Keywords—Phishing detection, CNN, Logistic regression, Decision tree, Cyber security.

I. INTRODUCTION

Nowadays, users' privacy on the web is affected by various cyber threats, among which is a phishing attack. A phishing website is a social engineering threat aiming to get sensitive data like passwords, MasterCard details, and, usernames, typically for malevolent causes, by sitting as a reliable entity in an online transmission [1].

The harms of phishing we face today are increasing rapidly directly affecting the privacy of individuals and causing big economic losses [1]. In keeping with the 2021 4th quarter Anti-phishing unit survey report, the number of phishing attacks has grown to 316,747 in December 2021, and the financial field is the most affected by this attack probably due to the due to the remnants of COVID 19, which confirms that the phishing attack is still active and that the attackers still inventing numerous manners to carry out this attack [2]. Phishing attacks are often prevented by software solutions and educating users to spot phishing websites. There are several software-based anti-phishing solutions developed in the literature such as black/white list, heuristics-based solutions, visual similarity, machine learning, deep learning...etc [1]. Machine learning and deep learning have been largely used in the phishing detection field since it is a binary classification problem. Generally, supervised learning methods are used with URL analysis by extracting

features. Moreover, these approaches have been very successful and achieved very high results in phishing detection because they can be applied to a large-scale dataset, also deep learning models such as CNN can overcome the problem of manual feature extraction [3], [4]. In this paper, we propose a website phishing detection approach based on URL features and using three classifiers: Convolutional Neural Network (CNN), Logistic Regression, and Decision tree, we also apply various performance measures to evaluate the effectiveness of the proposed solution and compare between the deep learning model (CNN) and the machine learning classifiers.

After the introduction, the remainder of the paper is organized in this way: Section two provides the related works, and the proposed solution is presented in section three. Then the experiment and validation are given in section four. Finally, section five gives the conclusion of the paper.

II. RELATED WORKS

A. Machine Learning

Several ML-based approaches are carried out in the phishing field with the combination of heuristic methods, the performance of this solution depends on the algorithm used, the feature number and types, and the volume of the training data [3], [4], [5].

Kumar et al in their study [6] used several classifiers such as Logistic Regression, Naive Bayes Classifier, Random Forest, Decision Tree, and K-Nearest Neighbor based on the attributes extracted from the URL lexical structure. They used the ratio of 7:3 for splitting the dataset into train and test. The Naive Bayes classifier reached the highest accuracy value of 98% with a precision=1, recall=0.95, and F1-score=0.97.

Korkmaz et al [7] introduced a machine-learning-based anti-phishing system by using different algorithms on three different datasets. The algorithms used were Support Vector Machine (SVM), Logistic Regression (LR), K-Nearest Neighbor(KNN), Decision Tree (DT), Naive Bayes (NB), XGBoost, Random Forest (RF), and Artificial Neural Network (ANN). In terms of training time, NB, DT, LR, and ANN algorithms gave better results. Also It was observed that the models using LR, SVM, and NB have a low accuracy rate. They deduced that the RF algorithm or ANN algorithm may be used owing to the reduced training time along with a high accuracy rate.

Chen et al, [8] presented a heuristic anti-phishing method that applies gestalt and decision theory concepts to model perceptual similarity. They used a logistic regression algorithm to normalize the characteristics of the page content.

Alam et al [9] introduced an anti-phishing attack method using Random Forest and Decision Tree. They used 32 features with feature selection algorithms like principal component analysis (PCA) to reduce the redundancy of the data. The Random Forest algorithm reached a maximum accuracy value of 97%.

B. Deep Learning

Recently, various deep learning model-based works such as LSTM, CNN, RNN...etc, are applied in the phishing domain and achieved very high results [4]. Convolutional neural networks (CNN) are one of the more used deep learning models showing good results in phishing website detection [10]. Several CNN-based anti-phishing solutions are carried out in the literature.

Authors in [11] presented a deep learning-based technique for classifying phishing called HTMLPhish. They applied the CNN model with the content of an HTML document on a web page without using manually engineered attributes. The CNN model learns the semantic dependencies of HTML text content. Their approach achieved an accuracy value of 93 %

URLNet is a CNN-based approach to detect malicious URL presented in [12]. Authors implemented Convolutional Neural Networks to both characters and words in URL strings to learn the URL embedding in a collaboratively optimized framework. Their technique worked end-to-end without the need for expert features it also worked on large datasets and significantly improved performance.

Authors in [13] presented a deep learning-based solution for detecting URL phishing websites, they used a character-level Convolutional Neural Network (CNN). The accuracy value achieved by their system is 95.02%.

III. THE PROPOSED METHODOLOGY

Fig. 1 shows an overview of the proposed phishing website detection solution. Our methodology consists of four main stages: Input Data, Feature preprocessing, detection models, and classification output.

In the first stage, the input data consists of a publicly available dataset that was obtained from Kaggle [14]. It contains 87 Features, such as having an IP address, length of URL...etc. Then in the second phase, these features are passed to a pre-processing stage, (data cleaning and removing null sample cases, balancing and shuffling dataset...etc). In the third stage, the feature vector resulting from the preprocessing stage is introduced as input to the detection model, where three classifiers are used such as 1D CNN, Logistic Regression, and Decision tree, each classifier is tested separately with the same dataset. Finally, in the fourth stage, the website is classified as legitimate or phishing.

IV. EXPERIMENT AND VALIDATION

To build the three models we used the Jupyter notebook in the Anaconda 3 environment using a python programming language with the help of the TensorFlow and Keras packages.

The dataset used consists of 87 features and 11430 samples divided into 5715 phishing and 5715 legitimate.

A. Training and test

The dataset is split into two parts with a ratio of 80% for the training dataset and 20% for the testing dataset. The input data are defined as the features and the output as the status index (1 in case of phishing and 0 in case of legitimate).

In our experiment, three classifiers are separately built, trained with the training data, and evaluated with the testing data. The 1D CNN, the logistic regression, and the decision tree to make a comparison between them.

1D CNN model: the CNN model consists of two convolutional layers, two batch normalization layers, two max-pooling layers, flatten layer, and two dense layers. Dropouts are applied after each pooling layer and after the first dense layer.

The parameters used for the CNN model are presented in TABLE 1.

TABLE I. PARAMETERS OF THE CNN MODEL.

Conv 1D Activation function	Conv1D layers filters	Optimizer	Learning rate	Dense layer Activation function	epochs
ReLU	32/64	Adam	0.001	Sigmoid	40/100

B. Performance measures

To evaluate our approach, we used the performance metrics shown in the equations (1), (2), (3), (4), (5), (6), (7), (8) [5].

$$\bullet \quad TPR = \frac{TP}{(TP + FN)} \quad (1)$$

$$\bullet \quad FPR = \frac{FP}{(FP + TN)} \quad (2)$$

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

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

$$\bullet \quad Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$\bullet \text{ Precision} = \frac{TP}{TP+FP} \quad (6)$$

$$\bullet \text{ Recall} = \frac{TP}{TP+FN} \quad (7)$$

$$\bullet \text{ F1-Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

Where:

- TPR is the true positive rate.
- FPR is the false positive rate.
- FNR is the false-negative rate.
- TNR is the true negative rate.

C. Results and Discussion

TABLE 2 shows the results obtained by each classifier with the same dataset using different performance measures. And Fig. 2 and Fig. 3 show the accuracy and loss models of the CNN classifier at different numbers of epochs (40 and 100 respectively).

TABLE II. RESULTS OBTAINED BY THE PROPOSED SOLUTION USING THREE CLASSIFIERS.

Algorithm used	Accuracy(%)	Precision(%) (avg)	Recall(%) (avg)	F1-Score(%) (avg)
CNN	96.19	96	96	96
Logistic Regression	94.40	94	95	94
Decision Tree	94.05	94	94	94

By analyzing TABLE 2 we notice that the CNN classifier gives the highest accuracy value of 96.19% and also gives the highest result values based on the precision, recall, and F1 score.

we also notice that the logistic regression and decision tree classifiers give very similar results according to all the performance measures used. The accuracy value reached by the logistic regression classifier is 94.40% and the accuracy value reached by the decision tree is 94.05%.

By analyzing Fig. 2 and Fig. 3 in the accuracy model we note that there is a direct proportion between the validation accuracy value and the epochs, as the number of epochs increases, the validation accuracy value increases until it starts to stabilize approximately at the epochs 25, 60 (Fig. 2 and Fig. 3 respectively) and achieves the accuracy value of 96.15% at epoch 40 (Fig. 2) and the value of 96.19% at the 100 epoch (Fig. 3). The validation loss value obtained by the CNN model at the 100 epoch is 0.12%.

All these results lead us to deduce that the CNN model performed well in phishing detection compared to logistic regression and decision tree classifiers, which also proves that deep learning classifiers work well than machine learning algorithms in the phishing detection field.

V. CONCLUSION

In this paper, we presented a website anti-phishing solution using URL attributes and testing deep learning and machine

learning classifiers, we used the convolutional neural network (CNN) model, logistic regression, and the decision tree classifiers.

based on the results obtained by the experiments and comparing between the three classifiers we noticed that the CNN model outperformed the machine learning classifiers in terms of all performance metrics and reached the accuracy value of 96.19 %, which proves that the deep learning classifier works better than machine learning algorithms in the phishing detection field.

In the future, we tend to improve the CNN accuracy value of this method and implement it in a web extension to ensure synchronous prediction of phishing websites and protect online user privacy.

ACKNOWLEDGMENT (Heading 5)

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

- [1] Robat D., Mukhter H., Shariful I., Abujarr S.,: Learning a Deep Neural Network for Predicting Phishing Website, PhD Thesis, Brac University, Bangladesh, 2019.
- [2] 4th quarter 2021: Phishing activity trends report. Anti-Phishing Working Group (APWG), 2021. <https://apwg.org/trendsreports/>.
- [3] Büber, E. Phishing URL Detection with ML. In Towards Data Science. (2018, Feb 8). <https://towardsdatascience.com/phishing-domain-detection-with-ml-5be9c99293e5>.
- [4] Huang, Y., Yang, Q., Qin, J., & Wen, W. Phishing URL Detection via CNN and Attention-Based Hierarchical RNN. In *Proceedings of the 18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications*. IEEE. 2019.
- [5] Zaimi, R., Hafidi, M., & Mahnane, L. A literature survey on anti-phishing in websites. In *Proceedings of the 4th International Conference on Networking, Information Systems & Security (NISS2021)*. Association for Computing Machinery. 2021.
- [6] Kumar, J., Santhanavijayan, A., Janet, B., Rajendran, B., & Bindhumadhava, B. S. Phishing website classification and detection using machine learning In *Proceedings of the International Conference on Computer Communication and Informatics (ICCCI 2020)*. IEEE. 2020.
- [7] Korkmaz, M., Sahingoz, O.K., & Diri, B. Detection of phishing websites by using machine learning-based URL analysis. In *Proceedings of the 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT 2020)*. IEEE. 2020.
- [8] Chen.T-C, Stepan.T, Dick.S, and Miller.J. An anti-phishing system employing diffused information. *Journal of ACM Transactions on Information and System Security (TISSEC)*, vol.16, no.4, pp.1-31, 2014.
- [9] Alam, M.N., Sarma, D., Saha, I., Ulfath, R.E., & Sohrab. H. Phishing attacks detection using machine learning approach. In *Proceedings of the 3rd International Conference on Smart Systems and Inventive Technology (ICSSIT 2020)*. 2020.
- [10] Yang, R., Zheng, K., Wu B., Wu C., Wang, X. Phishing Website Detection Based on Deep Convolutional Neural Network and Random Forest Ensemble Learning. *Sensors*, 21(24):8281. 2021. <https://doi.org/10.3390/s21248281>.
- [11] Opara, C., Wei, Bo., & Chen, Y. HTMLPhish: Enabling Phishing Web Page Detection by Applying Deep Learning Techniques on HTML Analysis. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*. IEEE. 2020.
- [12] Hung, Le., Pham, Q., Sahoo, D., Hoi, S.C.H. URLNet: Learning a URL Representation with Deep Learning for Malicious URL Detection. In *Proceedings of ACM Conference 2017*. ACM. 2018.

- [13] Aljofey, A., Jiang, Q., Qu, Q., Huang, M., & Niyigena, J.P. An effective phishing detection model based on character level convolutional neural network from URL. Electronics, 9(9),1514. 2020.
- [14] [Tiwari, S.](#) Web page Phishing Detection Dataset, Detect Phishing in

Web Pages. Kaggle. 2021. <https://www.kaggle.com/shashwatwork/web-page-phishing-detection-dataset>.

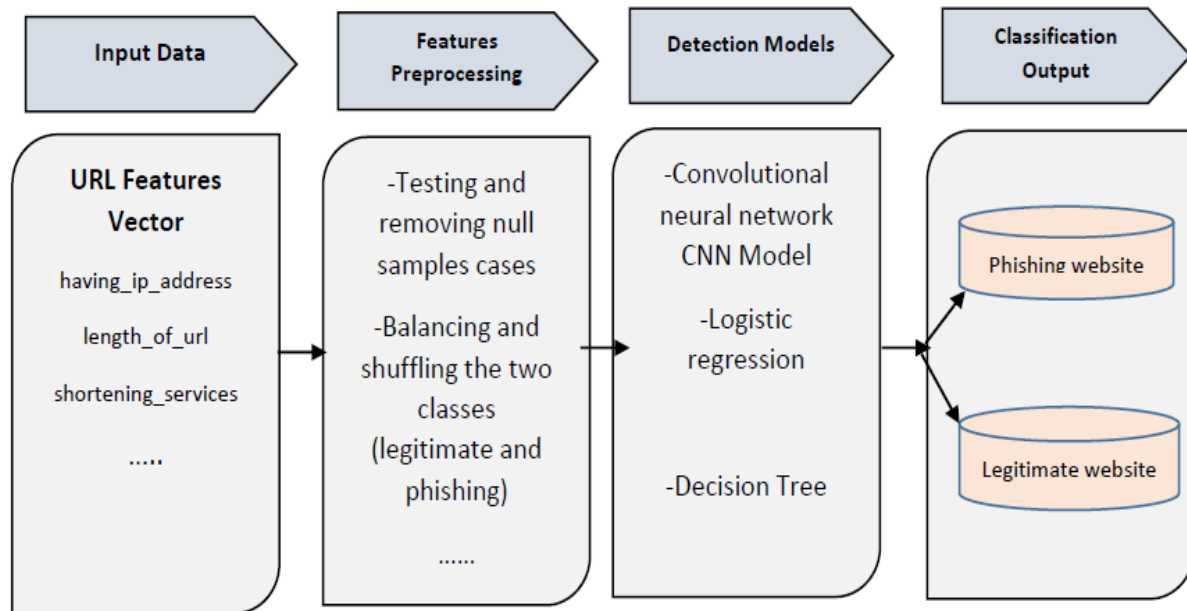


Fig. 1 Overview of the proposed solution.

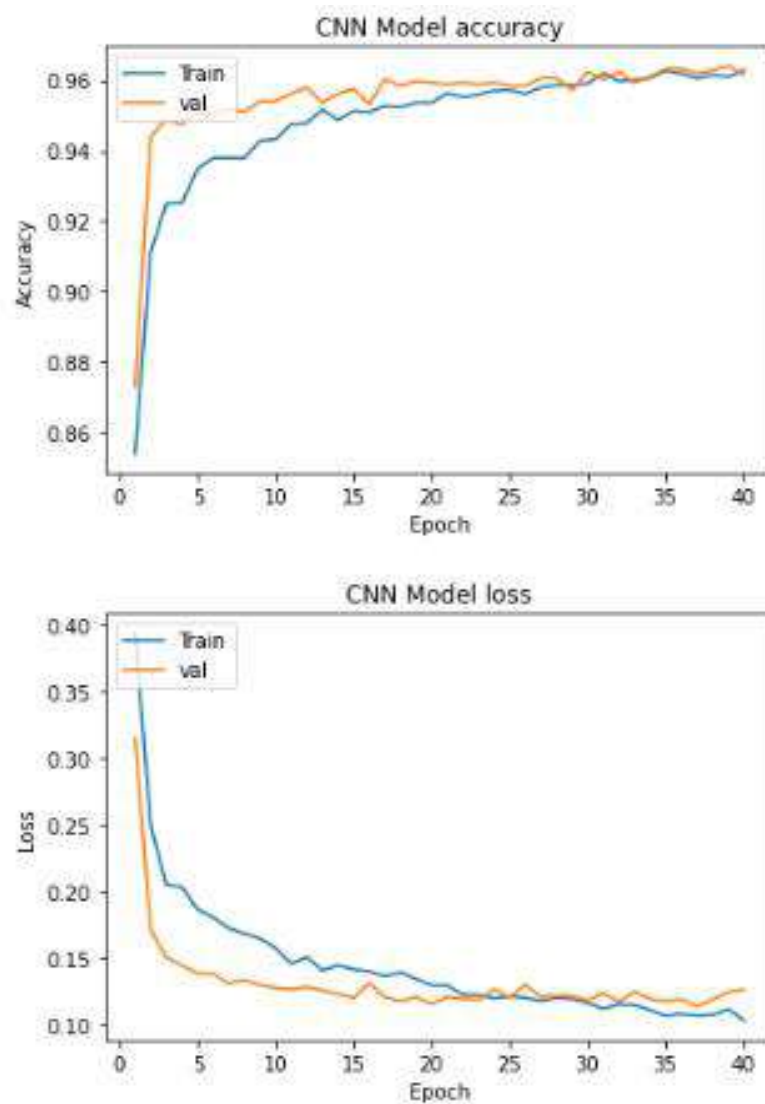


Fig. 2 Accuracy and loss graph of the CNN model in terms of the number of epochs (epochs=40).

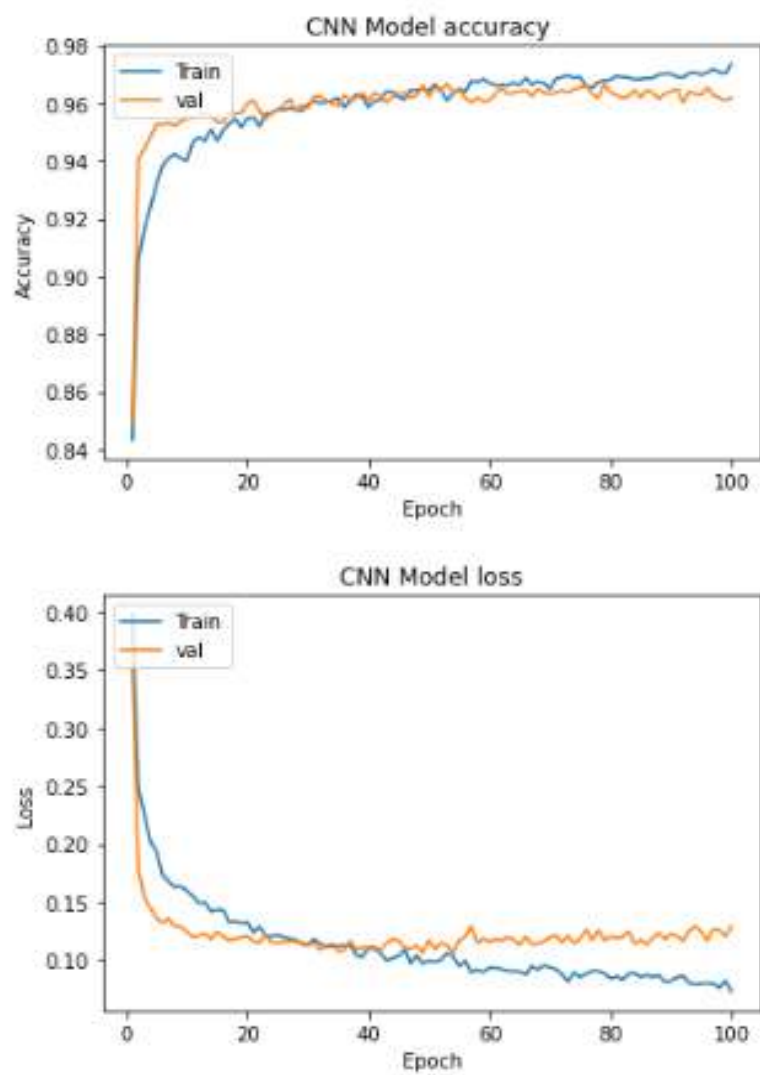


Fig. 3 Accuracy and loss graph of the CNN model in terms of the number of epochs (epochs=100).

RecTulips : A Python Library for Recommendation System

Ouahiba Remadnia

*Laboratory of Computer Science and Mathematics
University of Souk Ahras
Souk Ahras, Algeria
remadniahiba3@gmail.com*

Faiz Maazouzi

*Dept. of Computer Science
University of Souk Ahras
Souk Ahras, Algeria
mazouzi@labged.net*

Houssem Eddine Azzag

*Dept. of Computer Science
ESI - Sidi Bel Abbès
Souk Ahras, Algeria
h.azzage@esi-sba.dz*

Abstract—Recently, the recommendation system has become one of the most famous solutions in the field of e-shopping and social media, even to platforms for viewing movies and series, and many other areas, which are based on the recommendation system to attract more users and make the user spend more time in their application, but the problem here is that not all the developers are familiar with artificial intelligence or recommendations system algorithm, so we proposed a library that help developers to integrate recommendation system in there applications without any previous knowledge about the field, Our library is based on machine learning algorithms in recommendation system, also it was tested in one of the largest movies dataset 'MovieLens 100K' and we obtained with it very satisfactory results.

Key Words : Recommendation system, Collaborative filtering, Python library, Recommender System

I. INTRODUCTION

In recent years, the Internet offers a mass of data very voluminous to the user, and in front of this overload of information, it has become difficult to know what data to search and where to find them to choose the most accurate. In order to solve this problem, computer techniques have been developed to facilitate this search and the extraction of relevant information. The one we focus on is the recommendation system. It is about guiding the user in his exploration of the data so that he finds relevant information. A recommendation system provides a list of items to users based on their past preferences and needs or the preferences of users who have the same characteristics. Recommender system is used in various domains like E-learning [1], search for new movie [2], Book [3], Products [4], Music [5], E-commerce [6], article search [7], Tourism [8], Web and many more. The intention of recommender system is to examine user data and mine necessary data for further detection. The development of such systems requires the creation of well-defined recommendation algorithms by developers. And to accelerate and facilitate this work we approach in this work to provide a library that contains predefined methods in this area serves to help developers by using them directly.

II. RELATED WORK

In the recent years, the philosophy of big data attracts great attention from several official organizations including governments, universities, and industries in which the recommender

systems are introduced to help them to find what they need via a mechanism that can make prediction depending on different criteria. many researchers have tried to make recommendation more understandable and explainable. [9] they propose a novel dimensionality reduction based collaborative filter recommendation algorithm and an improved Big Data collaborative filter recommendation algorithm clustering techniques. They use clustering k-means algorithm and Singular Value Decomposition (SVD) to cluster similar users and reduce the dimensionality, respectively, The experimental dataset was obtained from Netflix also for [10] they developed a neural network framework, which was capable to generate abstractive comments when providing outfit recommendations. They employed a convolutional neural network with a mutual attention function to outfit matching and recurrent neural network with a cross-modality attention function to obtain a concise sentence. [11] they propose a new effective model-based trust collaborative filtering for explainable recommendations that aims not only to improve the quality of recommendation but also to provide an efficient support for explainable recommendations based on trustworthiness modeling. The proposed method is evaluated on Amazon Instant Video dataset in terms of RMSE In [12] they proposed a Deep feedback network (DFN), which considers both explicit/implicit and positive/negative feedback's to learn user unbiased preferences. The significant improvements in offline and online verify the effectiveness and robustness of DFN, [13] this work proposes a new method based on CF recommendation, DDCF, which captures the preference user variations, including the concept of dynamic time decay, DDCF dynamically adjusts the decay function based on user behaviors. , the results indicate that the DDCF performs better than the traditional collaborative filtering with dynamic decay function consideration. Experiments conducted on real datasets have also demonstrated the feasibility of the proposed DDCF as well as [14] They proposed trust matrix measure, which combines user similarity with weighted trust propagation. The recommendation model using different machine learning techniques and then used trustbased filtering to recommend with more accuracy, Results imply that DNN with trust model proved to be the best model with high accuracy of 83% with 0.74 MSE value and can be used for best movie recommendation, [15] In this study they choose the user collaborative

filtering to generate recommendations for online teaching. they use person correlation to calculate proximity between learners, and use rule-based association recommendation to derive top-N recommendations from neighborhood of customers. this system works flows are divided into five sections: Data Collection, Data ETL, Model Generation, Strategy Configuration, and Service Delivery, all of which are essential for any education systems recommendation. The proposed study used a model-based approach matrix factorization, the ALS algorithm associated with a collaborative filtering technique, which fixed cold start, parsimony and scalability issues. the proposed movie recommender system showed 97% accuracy and predicted the top 1000 ratings for movies, [17] This study demonstrates a way to build a large-scale recommender system by simply writing a series of SQL-like queries. they implemented a variety of recommendation algorithms and recommendation functions as user-defined Hive. this study demonstrates that how Apache Hivemall can easily be used to create a scalable recommendation system with satisfying business requirements and for [18] they offers a fast and scalable general-purpose library algorithms of the recommender system, It aims two common collaborative filtering scenarios: diction and prediction of items based solely on positive implicit opinions.

III. COLABORATIVE FILTERING

A recommender system is a form of information filtering that seeks to predict the "preference" a user will give to an item, such as a product, movie, song, etc. therefore a recommender system makes a prediction based on a user's past behaviors and interactions. More specifically, it is designed to predict the user's preference for a set of items based on their experience. these systems are generally categorized into two types based on the data used to make recommendations: the first type is the Content-based filtering, which uses item attributes. The second is the Collaborative filtering, which uses user behavior in addition to item attributes. Collaborative filtering filters information using and data collected by the system from the interactions of other users. It is based on the idea that people who agreed in their assessment of certain items are likely to agree again in the future. These systems focus on the relationship between users and items. The similarity of items is determined by the similarity of ratings of those items by users [19] , Figure 1 shows the recommender system structure.

IV. PROPOSED MODEL

According to the foregoing, we noticed that to put a recommendation system in your application or website, it is required that you have previous knowledge in the field of artificial intelligence and the algorithms of suggestions, etc. This requires a lot of time and effort, so we suggested that we provide a library that provides all the features of the recommendation system, and you can use it and enjoy the features of the recommendation system without any prior knowledge of what was previously mentioned. This will save time and great effort for developers in the future, and to make

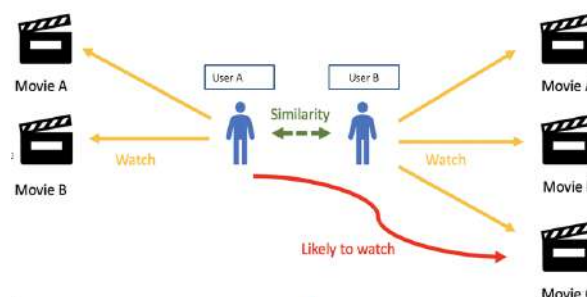


Fig. 1. Basic structure of collaborative filtering .

our library more effective and accurate we focused on having our work already tested on a large dataset and experimenting with several machine learning algorithms.

A. Model Definition

Our model is based on three necessary steps Fig. 5 : First, save the data from users our data structure is very simple it's based on (user_id , item_id , rating , time) Fig. 2 .

	user_id	movie_id	rating	time
0	196	242	3	1997-12-04 15:55:49
1	186	302	3	1998-04-04 19:22:22
2	22	377	1	1997-11-07 07:18:36
3	244	51	2	1997-11-27 05:02:03
4	166	346	1	1998-02-02 05:33:16
5	298	474	4	1998-01-07 14:20:06
6	115	265	2	1997-12-03 17:51:28
7	253	465	5	1998-04-03 18:34:27
8	305	451	3	1998-02-01 09:20:17
9	6	86	3	1997-12-31 21:16:53

Fig. 2. Sample of what the data we use looks like

Second, training our data Fig. 3, in our work we try so many machine learning algorithms as : SVD (Singular Value Decomposition) , SVD++ (Singular Value Decomposition++) , NMF (Non-negative Matrix Factorization) , KNNBaseline (K-Nearest Neighbors Baseline) , etc . and in our library we use the algorithm who gets the highest result when trying it on the selected dataset , and finally the suggestions in this step we give for every user list of suggested item based on what we get from result of training part .

B. Dataset

To ensure the accuracy and effectiveness of our proposal we had to test it on the dataset . So we chose MovieLens-100K dataset , it's contains 716 users that rated 3,952 movies (items) with 100,000 rating from 1-5. You can see the analysis of the ratings Fig. 4 .

movie_id	1	2	3	4	5	6	7	8	9	10
user_id										
1	5.0	3.0	4.0	3.0	3.0	5.0	4.0	1.0	5.0	3.0
2	4.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	2.0
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5	4.0	3.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6	4.0	NaN	NaN	NaN	NaN	NaN	2.0	4.0	4.0	NaN
7	NaN	NaN	NaN	5.0	NaN	NaN	5.0	5.0	5.0	4.0
8	NaN	NaN	NaN	NaN	NaN	NaN	3.0	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	5.0	4.0	NaN	NaN	NaN
10	4.0	NaN	NaN	4.0	NaN	NaN	4.0	NaN	4.0	NaN

Fig. 3. Matrix contain the data used on train

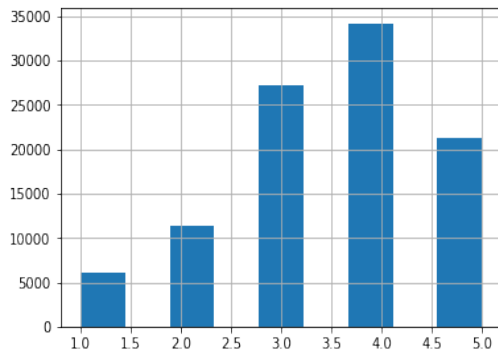


Fig. 4. MovieLens-100k dataset rating histogram

C. Library functions demonstration

There is two main functions in our library .
 Rate(user_id,item_id,rating,time) here every time a user rate an item you call our function to save this rate so we can use it later in suggestions .
 Top10(user_id) and here you prove a user id and we suggest for him 10 items .

D. Example

Here is a simple example showing how to use our library by giving rating data then call Top10 function to suggest items to the user .

```
import RecTulips as RS
```

```
RS.Rate(1,2,5,'10-04-2022 19:50:52')
RS.Rate(1,5,4,'30-03-2022 10:34:10')
RS.Rate(3,4,1,'11-03-2022 19:50:52')
RS.Rate(5,3,2,'12-04-2022 17:05:02')
RS.Rate(5,5,5,'22-04-2022 18:53:40')
RS.Rate(4,1,4,'25-04-2022 20:51:00')
RS.Rate(7,9,4,'20-04-2022 20:20:13')
RS.Rate(7,8,3,'25-03-2022 10:50:17')
```

```
RS.Rate(2,7,5,'15-04-2022 22:25:01')
RS.Rate(6,6,2,'21-04-2022 20:00:00')
RS.Rate(8,10,3,'20-04-2022 20:12:00')
RS.Rate(1,7,5,'04-05-2022 19:00:23')
RS.Rate(1,11,5,'03-04-2022 14:10:52')
RS.Rate(2,6,4,'10-04-2022 18:50:52')
RS.Rate(3,6,5,'11-04-2022 22:50:52')
RS.Rate(3,12,4,'10-03-2022 20:10:02')
RS.Rate(5,7,5,'25-04-2022 15:05:15')
RS.Rate(5,10,5,'10-01-2022 14:51:00')
```

```
Result = RS.TOP10(2) # Will return the id of the 10
item recommended to user (2)
```

```
print(Result) # [2, 12, 11, 1, 9, 5, 10, 3]
```

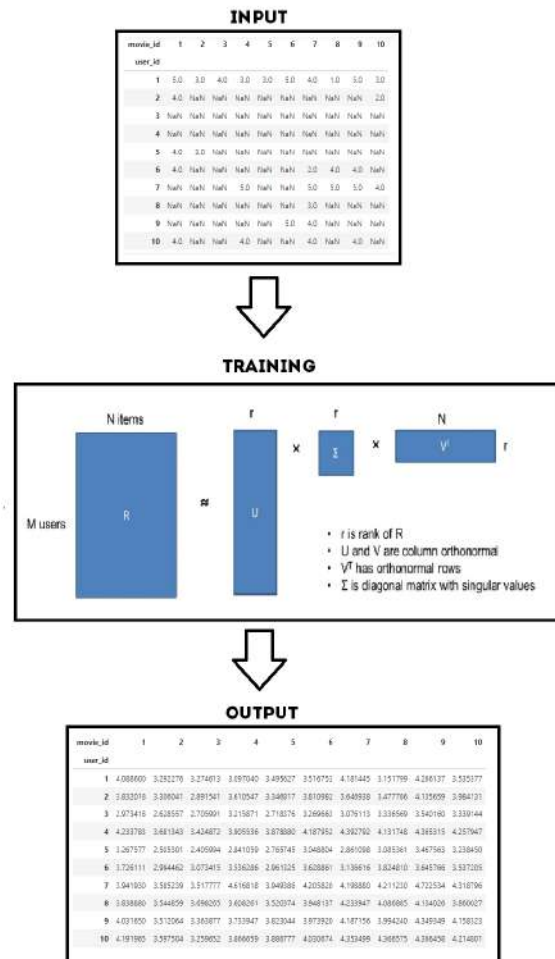


Fig. 5. Concept of the proposed model

V. EXPERIMENTAL RESULTS

The proposed model are implemented in Python using Kaggle platform Test environment configurations provided in Table I , with a help of surprise library [20] .

TABLE I
TEST ENVIRONMENT CONFIGURATIONS

RAM	13.00 GB
GPU	Tesla P100-PCIE
GPU Memory	15.9 GB
Disk	73.1 GB

In these experiments , we work with 4 different machine learning algorithms SVD , SVD++ , KNN , NMF as mention before , and we relied on measuring the results on the following metrics the Root Mean Square Error (RMSE) shown in Equation 1 , Mean Absolute Error (MAE) shown in Equation 2 , Fit time , Test time .

$$RMSE = \sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^n (y_i - x_i)^2} \quad (1)$$

$$MAE = \left(\frac{1}{n}\right) \sum_{i=1}^n |y_i - x_i| \quad (2)$$

TABLE II
COMPARISON OF MACHINE LEARNING METHODS IN MOVIELENS-100K DATASET

Method	RMSE	MAE	Fit time	Test time
SVD	0.9311	0.7312	6.59	0.31
NMF	0.9896	0.7753	5.14	0.41
KNN(n=50)	0.9319	0.7323	0.76	6.59
SVD++	0.9368	0.7380	106.47	7.88

Table II shows the value of RMSE , MAE , Fit time , Test time for all the four algorithms , we can see that the SVD Algorithm is the best in RMSE , MAE Also in Test time so we can concluded that the SVD algorithm is the best algorithm in this case for many reasons the first one is the RMSE score so he is more accurate than any other also for the Fit and Test time SVD make 6.59 sec for Fit 100,000 row and take 0.31 sec for test it's a great result especially that our library will work in real-time so the fit and test time is so important for the speed of our library or the speed of the application that will use our library .

VI. FUTURE WORK

Despite all the efforts made by researchers in this subject, it still has not reached perfection, so it is necessary to work in the future to develop it through the following points:

A. Use Deep learning

Despite the satisfactory results that the current recommendation systems give through machine learning algorithms only,

but we still believe that the results can be developed more, reach more accuracy, and reach more satisfactory results. We can reach these results through the use of deep learning algorithms [21] and others, so we are working in our research In the future, we should provide a recommendation system or a recommendation system library using deep learning [25].

B. Recommendation system dedicated to a specific field

The problem that we have noticed in the recommendation systems or in the offices available for the recommendation systems[20 , 23 , 24], is that it is general and not specific to a specific field. In fact, we can consider this a strength and at the same time we can consider it a weakness that can be a strength point, as it is used in any field, whatever its form or His method and at the same time, of course, will not provide all the characteristics needed for this field, and will not provide the necessary accuracy, as it is not intended for it, so we recommend that connecting systems or offices dedicated to certain fields be provided, for example: E-Commerce , E-Learning , Music , Films , etc .

Providing custom recommendation systems for specific fields by providing functions that represent that field, for example: Video_Watched() , Music_Listened() , Cours_Readed() , etc .

C. More privacy

Recommendation systems often use privacy and sensitive information about users to provide more accurate suggestions about a personality, but the user must have more privacy and that this information is not clear to the owner of the application, this can be provided by several methods such as encrypting the information before using it to be identifiable by The party of the machine only or limit the information to areas [22].

D. Libraries for more programming languages

We noticed that the Libraries currently available for recommender systems are limited to a very limited number of programming languages such as Python and NodeJS, so we suggest that they be generalized more to more programming languages to facilitate the task for a larger number of developers around the world.

E. Hybrid recommendation systems

Most of the research in the field of recommendation systems, if not all, is done using one format, whether collaborative filtering or content-based filtering , especially offices, so we suggest that a new format be used that is a hybrid [26] between the two systems to provide more accurate results.

VII. CONCLUSIONS

Providing a library that provides a recommendation systems service with all its many characteristics and advantages, which facilitates the process a lot for programmers and allows them to use the advantages of recommendation systems without any previous knowledge in the field of artificial intelligence. This is what makes this field more important, so we have provided a complete library that provides a service of recommendation

systems in the programming language Not only that, but we focused on making the results provided by our library very satisfactory, so we worked on experimenting with several algorithms for recommender systems and comparing between them and choosing the best and optimal of them in our case .

We focused in our results on two important factors, namely accuracy (RMSE , MAE) and speed (Fit time , Test time) , accuracy in ensuring that the results given by our library are satisfactory and effective, and speed in giving results to the user as soon as possible to ensure a very smooth user experience.

The experiments we conducted were on a dataset called MovieLens-100k, and we tested a significant number of recommender algorithms to: SVD , SVD++ , NMF , KNN.

And we came out with the conclusion that the Singular Value Decomposition (SVD) algorithm is the best in terms of accuracy and speed .

REFERENCES

- [1] Bhaskaran, Sundaresan, Raja Marappan, and Balachandran Santhi. "Design and analysis of a cluster-based intelligent hybrid recommendation system for e-learning applications." *Mathematics* 9.2 (2021): 197 .
- [2] S. Kumar, K. De, and P. P. Roy, "Movie recommendation system using sentiment analysis from microblogging data," *IEEE Transactions on Computational Social Systems*, 2020
- [3] G. Ramakrishnan, V. Saicharan, K. Chandrasekaran, M. Rathnamma, and V. V. Ramana, "Collaborative Filtering for Book Recommendation System," in *Soft Computing for Problem Solving*, ed: Springer, 2020, pp. 325-338
- [4] S. Akram, S. Hussain, I. K. Toure, S. Yang, and H. Jalal, "ChoseAmobile: A Web-based Recommendation System for Mobile Phone Products," *Journal of Internet Technology*, vol. 21, pp. 1003-1011, 2020.
- [5] R. Wang, X. Ma, C. Jiang, Y. Ye, and Y. Zhang, "Heterogeneous information network-based music recommendation system in mobile networks," *Computer Communications*, vol. 150, pp. 429-437, 2020.
- [6] Hwangbo, H., Kim, Y. S., & Cha, K. J. (2018). Recommendation system development for fashion retail e-commerce. *Electronic Commerce Research and Applications*, 28, 94-101.
- [7] Katarya, R., & Verma, O. P. (2017). An effective collaborative movie recommender system with cuckoo search. *Egyptian Informatics Journal*, 18(2), 105-112.
- [8] Hassannia, R., Vatankeh Barenji, A., Li, Z., & Alipour, H. (2019). Web-based recommendation system for smart tourism: Multiagent technology. *Sustainability*, 11(2), 323.
- [9] Zarzour, H., Al-Sharif, Z., Al-Ayyoub, M., & Jararweh, Y. (2018, April). A new collaborative filtering recommendation algorithm based on dimensionality reduction and clustering techniques. In *2018 9th international conference on information and communication systems (ICICS)* (pp. 102-106). IEEE.
- [10] Lin, Y., Ren, P., Chen, Z., Ren, Z., Ma, J., & De Rijke, M. (2019). Explainable outfit recommendation with joint outfit matching and comment generation. *IEEE Transactions on Knowledge and Data Engineering*, 32(8), 1502-1516.
- [11] Zarzour, H., Jararweh, Y., & Al-Sharif, Z. A. (2020, April). An Effective Model-Based Trust Collaborative Filtering for Explainable Recommendations. In *2020 11th International Conference on Information and Communication Systems (ICICS)* (pp. 238-242). IEEE.
- [12] Xie, R., Ling, C., Wang, Y., Wang, R., Xia, F., & Lin, L. (2021, January). Deep feedback network for recommendation. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence* (pp. 2519-2525).
- [13] Chen, Y. C., Hui, L., & Thaisitukul, T. (2021). A collaborative filtering recommendation system with dynamic time decay. *The Journal of Supercomputing*, 77(1), 244-262.
- [14] Choudhury, S. S., Mohanty, S. N., & Jagadev, A. K. (2021). Multimodal trust based recommender system with machine learning approaches for movie recommendation. *International Journal of Information Technology*, 13(2), 475-482.
- [15] Tan, H., Guo, J., & Li, Y. (2008, December). E-learning recommendation system. In *2008 International conference on computer science and software engineering* (Vol. 5, pp. 430-433). IEEE.
- [16] Awan, M. J., Khan, R. A., Nobanee, H., Yasin, A., Anwar, S. M., Naseem, U., & Singh, V. P. (2021). A Recommendation engine for predicting movie ratings using a big data approach. *Electronics*, 10(10), 1215.
- [17] Kitazawa, T., & Yui, M. (2018, September). Query-based simple and scalable recommender systems with apache hivemall. In *Proceedings of the 12th ACM Conference on Recommender Systems* (pp. 502-503).
- [18] Gantner, Zeno, et al. "MyMediaLite: A free recommender system library." *Proceedings of the fifth ACM conference on Recommender systems*. 2011..
- [19] Sharma, N., Chakrabarti, A., & Balas, V. E. (2019). Data management, analytics and innovation. *Proceedings of ICDMAI*, 1.
- [20] Hug, N. (2020). Surprise: A Python library for recommender systems. *Journal of Open Source Software*, 5(52), 2174.
- [21] Wen, X. (2021). Using deep learning approach and IoT architecture to build the intelligent music recommendation system. *Soft Computing*, 25(4), 3087-3096.
- [22] Yu, X., Zhan, D., Liu, L., Lv, H., Xu, L., & Du, J. (2021). A privacy-preserving cross-domain healthcare wearables recommendation algorithm based on domain-dependent and domain-independent feature fusion. *IEEE Journal of Biomedical and Health Informatics*.
- [23] Guo, G., Zhang, J., Sun, Z., & Yorke-Smith, N. (2015, June). LibRec: A Java Library for Recommender Systems. In *Umap Workshops* (Vol. 4).
- [24] Gantner, Z., Rendle, S., Freudenthaler, C., & Schmidt-Thieme, L. (2011, October). MyMediaLite: A free recommender system library. In *Proceedings of the fifth ACM conference on Recommender systems* (pp. 305-308).

- [25] Sheikh Fathollahi, M., & Razzazi, F. (2021). Music similarity measurement and recommendation system using convolutional neural networks. *International Journal of Multimedia Information Retrieval*, 10(1), 43-53.
- [26] Gulzar, Z., Leema, A. A., & Deepak, G. (2018). Pcrs: Personalized course recommender system based on hybrid approach. *Procedia Computer Science*, 125, 518-524.

Robust Cepstral Feature for Automatic Speech Recognition

Mohamed cherif Amara Korba
Department of Electrical Engineering
Mohamed Cherif Messaadia University
Souk Ahras, Algeria
amara_korba@univ-soukahras.dz

Hakim Doghmane
Department of Electronics and
Telecommunications
8 Mai 1945 - Guelma University
Guelma, Algeria
doghmane.hakim@univ-guelma.dz

Faiz Maazouzi
Department of Mathematics and
Computer Science
Mohamed Cherif Messaadia University
Souk Ahras, Algeria
f.maazouzi@univ-soukahras.dz

Abstract—Automatic Speech Recognition (ASR) systems are vulnerable to environmental noises, which significantly affect their performance. MFCC is the most employed technique for extracting the features from the speech signal, it is very effective for non-noisy environments, but it is very sensitive to noise, which affects dramatically the robustness of the ASR. In this research, the Mel frequency Cepstral Coefficient (MFCC) feature is modified to increase the robustness of the ASR against noises. the modifications made consist of the use of the differential power spectrum (DPS) instead of the power spectrum (PS), and use of the power function instead of the logarithmic non linearity in the algorithm. We used mean and variance normalization (MVN) to reduce cepstral coefficient mismatches between training and testing conditions, followed by the auto-regression moving-average filtering (MVA) in the cepstral domain. The experiments were conducted using LASA database which contains Arabic digits. In the test phase, the samples of the database were corrupted by four background noises at different levels of noise. The experimental results show that the proposed technique significantly improves accuracy recognition compared to the standard MFCC, the relative improvement is more than 28%.

Keywords—Robust speech recognition, MFCC, MVA, Differential power spectrum

I. INTRODUCTION

In automatic speech recognition systems, noise robustness is a critical consideration. The majority of them are sensitive to the environment in which they are employed. When the mismatch between the training and testing conditions is large, the ASR performance degrades significantly. Noise corrupts speech samples and generates a mismatch due to distortion in speech attributes under realistic noise conditions. As a baseline, most ASR systems use the MFCC [1] features. But it has been demonstrated that MFCC should not be used in noisy speech settings. As a result, many MFCC versions have been proposed to improve the robustness of the ASR system [2, 3]. Other features exist based on Perceptual Prediction approaches, such as relative spectra linear prediction (RASTA-PLP) [4] or perceptual linear prediction (PLP), are more effective in noisy environments. To enhance the environmental robustness of ASR systems, a number of compensation techniques have been developed such as cepstral mean normalization (CMN) [5], and variance mean normalization (VMN) [6, 7]. For effective compensation, there is an approach that combines CMN and VMN plus ARMA filtering called (MVA) [8, 9, 10], Recent research has shown that using a deep neural network (DNN) [11] and convolutional neural networks (CNNs) [12, 13] improve significantly robustness of the ASR.

In this study, we present a technique for extracting speech features that enhances the robustness of ASR systems. The main contributions of this paper are as follows:

- For improved acoustical distortion resistance, we employ the cepstrum obtained from differential power spectrum (DPS) rather than power spectrum (PS).
- We employ power-law nonlinearity (PLN) functions instead of the log nonlinearity functions used in MFCC, in order to suppress weak signals and their variations.
- We employ the MVA technique to reduce speech feature distortion; it has been proven in [8, 9, 10, 14] that it enhances the robustness of small-vocabulary ASR tasks greatly.
- We perform a comparison between our proposed feature and a set of robust acoustic features such as: minimum mean-square error (MMSE) spectral amplitude estimator [15], Power Normalized Cepstral Coefficient (PNCC) features [16], and Normalized modulation cepstral coefficients (NMCC) [17] features.

The remainder of this paper is laid out as follows. In section 2, we present the structure of our proposed feature extraction algorithm. In section 3, we present the experimental results and we compare our proposed feature with the most commonly used features in the literature. In section 4, the work is concluded.

II. ARCHITECTURE OF THE PROPOSED FEATURE EXTRACTION TECHNIQUE

In this section, we present the different processing steps leading to the Proposed Robust Cepstral Features (PRCF) coefficients. The structure of PRCF is represented by the figure below.

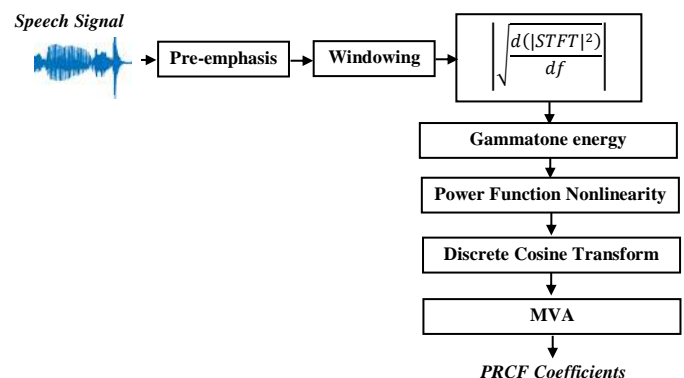


Fig. 1. Structure of the proposed feature extraction algorithm

Firstly, we apply a pre-emphasis filter of the form $H(Z) = 1 - 0.97Z^{-1}$ to input speech signal in the time domain to increase the high frequencies. A short time Fourier transform (STFT) is performed using a Hamming window with frame duration of 25.6ms and a frame shift of 10ms. DPS of speech signal $x(t)$ is obtained by differentiating the power spectrum with respect to frequency, it is defined by

$$D(\omega) = \frac{dX(\omega)}{d\omega} \quad (1)$$

Where ω denotes frequency. There are several discrete counterparts for approximating of the formula above; we chose the difference equation given by

$$d(k) = x(k) - x(k + 1) \quad (2)$$

[18] demonstrated that this form of approximation produces better outcomes. Then, an absolute operation is applied to DPS in order to make its negative parts positive. Finally, a normalized filter-bank based on 40 channel gammatone-shaped is applied to DPS; the filter-bank is applied between 130Hz and 6800Hz, whose center frequencies are linearly spaced in Equivalent Rectangular Bandwidth (ERB) [19].

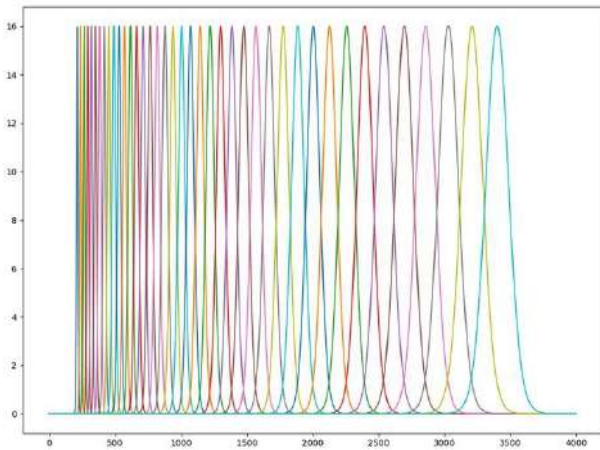


Fig. 2. Gammatone filterbank of 40 dimension

ERB scale is defined by

$$ERBS(f) = 21.4 \cdot \log(0.00437f + 1) \quad (3)$$

The gammatone modeling widely used is a physiologically motivated technique that may be considered as an approximation of human cochlear filter-bank. We use Snaleý's auditory toolbox [20] to get the impulse response of gammatone filter. In each channel the area under the squared transfer function is normalized to unity.

$$\int_{130}^{6800} |H_m(f)|^2 df = 1 \quad (4)$$

Where $H_m(f)$ is the frequency response of the m^{th} gammatone channel. We obtain the short-time spectral power $P(l, m)$ using the squared gammatone summation as below:

$$P(l, m) = \sum_{k=0}^{\left(\frac{K}{2}\right)-1} [|d_l(e^{j\omega_k})| \cdot H_m(e^{j\omega_k})]^2 \quad (5)$$

l frame indices, $d_l(e^{j\omega_k})$ is the short time spectrum of the l^{th} frame of the signal, m gammatone channel indices, K is the discrete Fourier transform size, it is equal to 1024.

Spectral power is enhanced by the expression

$$P_N(l, m) = P(l, m) \cdot 10^\gamma \quad (6)$$

Where γ is empirically defined, it is equal to 4. We apply nonlinearity by exploiting the power-law nonlinearity instead of log nonlinearity. This technique is supposed to eliminate the weak signals and their variability, it provides more robustness to ASR [16]. The nonlinearity is obtained by the following formula

$$P'(l, m) = (P_N(l, m))^{0.1} \quad (7)$$

Cepstral parameters are obtained from the spectral power $P'(l, m)$ using the Discrete Cosine Transform (DCT)

$$C(l, k) = \sqrt{\frac{2}{M}} \sum_{m=1}^M P'(l, m) \cos\left(\frac{\pi k}{M} \left(m - \frac{1}{2}\right)\right), \quad m = 1 \dots M, k = 1 \dots K \quad (8)$$

Where M is the number of gammatone filterbank channels, M and K are identical.

Three phases are included in the MVA post-processing procedure, consisting of mean subtraction, variance normalization and ARMA filtering. For a given utterance, it can be represented by a $N \times T$ matrix of frames, say C , where each colomn represents the feature vector for a given frame.

$$\begin{bmatrix} C_1(1) & \dots & C_1(T) \\ \vdots & \ddots & \vdots \\ C_N(1) & \dots & C_N(T) \end{bmatrix} \quad (9)$$

The first step is mean subtraction (MS) defined by

$$\tilde{C}_n(t) = C_n(t) - \mu_n \quad (10)$$

Where $C_n(t)$ is the n^{th} component of the feature vector at time t , $\tilde{C}_n(t)$ is the mean subtracted feature and μ_n is the mean vector estimated from data as

$$\mu_n = \frac{1}{T} \sum_{t=1}^T C_n(t) \quad (11)$$

The second step is variance normalization (VN) defined by

$$\tilde{\tilde{C}}_n(t) = \frac{\tilde{C}_n(t)}{\sqrt{\sigma_n}} \quad (12)$$

where σ_n is the variance estimated from data as

$$\sigma_n = \frac{1}{T} \sum_{t=1}^T (C_n(t) - \mu_n)^2 \quad (13)$$

Where $\tilde{c}_n(t)$ is the mean-subtracted and variance normalized feature at time t . The third step is the ARMA filtering process. It is used with the objective of making the noisy frame features much similar to clean features. In this study we have used non causal ARMA filter, which is defined as

$$\check{c}_n(t) = \begin{cases} \frac{\sum_{i=1}^Q \tilde{c}_n(t-i) + \sum_{j=0}^Q \tilde{c}_n(t+j)}{2Q+1} & \text{if } Q < t \leq T - Q \\ \tilde{c}_n(t) & \text{Otherwise} \end{cases} \quad (14)$$

Where Q is the order of ARMA filter. The filter structure is based on a study developed in [9], where it has been shown that non causal ARMA filter provides better performance to ASR system with respect to other types of filters. We apply the MVA to static features which are concatenated with their first and second order time derivatives.

III. EXPERIMENTAL RESULTS

This section describes the speech database used in ASR experiments, and it also provides the topology of HMM recognizer. At the end of this section, we present the experimental results to evaluate speech recognition performances of the PRCF features under a variety of acoustical environments.

A. Database description

The experiments were performed using LASA database developed by the LASA laboratory, which was designed for the evaluation of ASR algorithms [23]. The database contains a set of Arabic digit utterances pronounced in an isolated way. It includes 9,000 utterances spoken by 90 adult speakers. Each digit is repeated ten times by each speaker. Utterances in the test set were corrupted by four background noises (white, pink, factory1 and F16) taken from NOISEX-92 database [24]. Noises are artificially added to the LASA database at different SNR levels, the SNR is defined as the ratio of signal to noise energy. Speech database includes two test mode defined as below:

- Test mode A: Each speaker produces 10 utterances for each digit: Six are utilized for training, while the other four are used for testing, which gives 5400 utterances for clean training, and 3600 utterances affected by four additive noises in the range of 6 SNRs values (20, 15, 10, 5, 0 and Clean).
- Test mode B, there are 6000 speech utterances produced by 60 speakers in the training set, the test set contains 3000 utterances from 30 additional speakers who did not participate in the training phase. The test set utterances corrupted by the same noises as in test mode A at 6 SNR levels.

B. The HTK Recognizer

All of the experiments described here use the Hidden Markov Model (HMM) toolkit (HTK) [25]. Each digit HMM

has the same topology, 13 emitting states with 3 Gaussian mixtures per state, this topology is identical regardless of the type of feature used. Acoustic model training is carried out on clean speech utterances in all of our experiments.

C. The effect of filter order on recognition accuracy

To study the effect of the order of the ARMA filter on the robustness of the proposed feature, we varied its order from 1 to 6, in order to find the optimal order of the ARMA filter. Table 1 and Table 2 show that the optimal order of the filter for which the recognition accuracy is better is obtained for the value $Q=2$, this value gives a good balance between robustness to noise and preservation of information.

TABLE I. ACCURACY RECOGNITION (%) WITH DIFFERENT ORDERS OF THE ARMA FILTER FOR TEST MODE A

Filter order (Q)	clean	20-0dB
1	99.22	95.77
2	99.36	96.02
3	99.19	95.49
4	98.78	94.81
5	98.75	94.54
6	98.64	94.24

TABLE II. ACCURACY RECOGNITION (%) WITH DIFFERENT ORDERS OF THE ARMA FILTER FOR TEST MODE B

Filter order	clean	20-0dB
1	97.47	91.70
2	97.07	91.78
3	95.57	91.60
4	96.83	91.28
5	96.47	90.27
6	96.07	89.60

D. Comparison PRCF feature against baseline features

To evaluate the PRCF algorithm's performance, it is compared to the acoustic features widely employed in automatic recognition systems. Table 3 shows the characteristics of the acoustic features used in the experimental phase.

TABLE III. FEATURES PARAMETERS USED FOR EXPERIMENTAL ANALYSIS

Configuration features	MFCC	RASTA-PLP	PNCC	PNRF
window length (ms)	25	25	25.6	25.6
Frame shift (ms)	10	10	10	10
FFT size	200	256	1024	1024
Number of filter-banks	23	/	40	40
Number of coefficients	12	12	13	13
Appended log frame energy	do	do	/	/
Size $\Delta/\Delta\Delta$ window (frames)	3/2	3/2	3/2	3/2
Feature dimension	39	39	39	39

Tables 4 and Table 5 show the detailed experimental results obtained in test mode A and test mode B respectively. the fourth table column labeled (Rel_Imp) contains values of a

Relative Improvement of the proposed algorithm extraction methods. Relative improvement is defined as

$$Rel_Imp = \frac{r_p - r_t}{r_t} \times 100\% \quad (15)$$

r_p is the recognition rate of our proposed algorithm and r_t is the recognition rate of comparison feature extraction methods. The table column denoted by 'Avg 0–20dB' expresses the results recognition rate averaged from 0 dB to 20 dB over a four noise types.

In test mode A, at the SNR 0-20dB, PNRF with ($Q = 2$) achieve an error rate reduction of 21.54% as compared to MFCC baseline, 22.67% as compared to RASTA-PLP, 10.28% as compared to MFCC+MVA and 3.22% as compared to PNCC feature. In test mode B, at the SNR 0-20dB, PNRF with ($Q = 2$) gives 21.22% error rate reduction as compared to MFCC baseline, 21.21% as compared to RASTA-PLP, 9.41% as compared to MFCC+MVA and 1.79% compared to PNCC feature. PNRF gives a relative improvement over the baseline MFCC of 28.92% in the test mode A, and 30.07% in the test mode B. In the clean test, PNRF feature does not affect the performance of the recognition like the other features (MVA, RASTA-PLP and PNCC) on the contrary, it yields more performance as the standard MFCC features for both test modes.

TABLE IV. WORD ACCURACY (%) USING DIFFERENT FEATURE SETS

Features	Clean	AVG 0-20dB	Rel_Imp
MFCC	99.25	74.48	28.92
RASTA_PLP	98.11	73.35	30.90
MFCC + CMN	99.00	82.03	17.05
RASTA_PLP + CMN	98.42	75.73	26.79
MFCC + MVN	98.89	85.48	12.33
MFCC + MVA	99.00	85.74	11.99
PNCC	98.34	92.80	3.46
PRCF	99.36	96.02	—

TABLE V. WORD ACCURACY (%) USING DIFFERENT FEATURE SETS

Features	Clean	AVG 0-20dB	Rel_Imp
MFCC	99.25	74.48	28.92
RASTA_PLP	98.11	73.35	30.90
MFCC + CMN	99.00	82.03	17.05
RASTA_PLP + CMN	98.42	75.73	26.79
MFCC + MVN	98.89	85.48	12.33
MFCC + MVA	99.00	85.74	11.99
PNCC	98.34	92.80	3.46
PRCF	99.36	96.02	—

IV. CONCLUSION

In this paper, a new cepstral features have been proposed by introducing a differential power spectrum and power-law nonlinearity that replaces the existing technique of log nonlinearity used in MFCC processing. The proposed feature used the MVA technique as an integral part of the algorithm and not as post-processing phase. It was verified that the effectiveness of the proposed feature depends on the choice of ARMA filter order, the optimal order of the latter is equal to 2. According to the experimental results conducted on the LASA database, the proposed feature provides a

significant improvement in robustness, it outperforms state-of-the-art algorithms in noisy environments without losing performance in clean environment as well.

REFERENCES

- [1] Davis S., & Mermelstein P., "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," IEEE transactions on acoustics, speech, and signal processing, vol. 28, no 4, pp. 357-366, 1980.
- [2] Hossan M. A., Memon S., & Gregory M. A., "A novel approach for MFCC feature extraction," In Signal Processing and Communication Systems (ICSPCS), 2010 4th International Conference on IEEE, pp. 1-5, 2010.
- [3] Shao Y., Jin Z., Wang D., & Srinivasan S., "An auditory-based feature for robust speech recognition," IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 4625-4628, 2009.
- [4] Hermansky H., & Morgan N., "RASTA processing of speech," IEEE transactions on speech and audio processing, vol. 2, no. 4, pp. 578-589, 1994.
- [5] Furui S., "Cepstral analysis technique for automatic speaker verification," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 29, no. 2, pp. 254-272, 1981.
- [6] Huang X., Acero A., Hon H. W., & Foreword By-Reddy R., Spoken language processing: A guide to theory, algorithm, and system development, Prentice hall PTR, 2001.
- [7] Pujol P., Macho D., & Nadeu C., "On real-time mean-and-variance normalization of speech recognition features," In 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, vol. 1, pp. 773-776, 2006.
- [8] Techini, Elhem, Sakka, Zied, & Bouhleh MedSalim. "Robust front-end based on MVA processing for Arabic speech recognition," Engineering & MIS (ICEMIS), pp. 1-5, 2017.
- [9] Chen C. P., Bilmes J. A., & Kirchhoff, K., "Low-resource noise-robust feature post-processing on Aurora 2.0," In Seventh International Conference on Spoken Language Processing, 2002.
- [10] Chen C. P., & Bilmes, J. A., "MVA processing of speech features," IEEE Transactions on Audio, Speech, and Language Processing, vol. 15, no. 1, pp. 257-270, 2009.
- [11] Miao, Y., Gawayyed, M., & Metze, F., "EESSEN: End-to-end speech recognition using deep RNN models and WFST-based decoding," In Automatic Speech Recognition and Understanding (ASRU), IEEE Workshop on, IEEE, pp. 167-174, 2015.
- [12] Qian, Y., Bi, M., Tan, T., & Yu, K., "Very deep convolutional neural networks for noise robust speech recognition," IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol 24, no 12, pp. 2263-2276, 2016.
- [13] Ovtcharov, K., Ruwase, O., Kim, J. Y., Fowers, J., Strauss, K., & Chung, E. S. "Accelerating deep convolutional neural networks using specialized hardware. Microsoft Research Whitepaper, Vol 2, no 11, 2015.
- [14] Korba M. C. A., Messadeg D., Bourouba H., & Djemili R., "Noise Robust Features Based on MVA Post-processing," In IFIP International Conference on Computer Science and its Applications, vol. 456, pp. 155-166, 2015.
- [15] Ephraim Y., & Malah D., "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 32, no. 6, pp. 1109-1121, 1984.
- [16] Kim C., & Stern R. M., "Power-normalized cepstral coefficients (PNCC) for robust speech recognition," IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), vol. 24, no 7, pp. 1315-1329, 2016.
- [17] Mitra V., Franco H., Graciarena M., & Mandal A., "Normalized amplitude modulation features for large vocabulary noise-robust speech recognition," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4117-4120, 2012.
- [18] Chen J., Paliwal K. K., & Nakamura S., "Cepstrum derived from differentiated power spectrum for robust speech recognition," Speech Communication, vol. 41, no. 2, pp. 469-484, 2003.
- [19] Glasberg B. R., & Moore B. C., "Derivation of auditory filter shapes from notched-noise data," Hearing research, vol. 47, no. 1-2, pp. 103-138, 1990.

- [20] M Slaney, Auditory toolbox version 2. [online]
<https://engineering.purdue.edu/~malcolm/interval/1998-010/>
- [21] Ellis, D. PLP and RASTA in MATLAB. [online], 2006.
<http://labrosa.ee.columbia.edu/matlab/rastamat/>
- [22] PNCC code. online
http://www.cs.cmu.edu/~robust/archive/algorithms/PNCC_IEEETran/PNCC_IEEETran.tar.gz, 2012.
- [23] Korba, Mohamed Cherif Amara, et al. "Robust speech recognition using perceptual wavelet denoising and mel-frequency product spectrum cepstral coefficient features," Informatica journal, vol. 32, no. 3, pp. 283-288, 2008.
- [24] NOISEX-92 Database. [online]
<http://www.speech.cs.cmu.edu/comp.speech/Section1/Data/noisex.html>, 1992.
- [25] Yung S., Evermann G., Gales M., Hain, T., Kershaw D., Moore G., & Woodland P., The HTK book. Cambridge University Engineering Department, Cambridge, 2005.

ROBUST TRACKING SYSTEM OF VEHICLES ON MOTION USING KALMAN FILTER AT NIGHT MODE

Sabrina Bahroun
Bahrounsabrina23@hotmail.com
Departement of computer science
Universite Badji Mokhtar
Annaba, 23000, Algeria

Toufik Sari
sari@labged.net
Departement of computer science
Universite Badji Mokhtar
Annaba, 23000, Algeria

Nabiha azizi
nabiha11@yahoo.fr
Departement of computer science
Universite Badji Mokhtar
Annaba, 23000, Algeria

Farouk Benabbas
benabbas@labged.net
Departement of computer science
Universite Badji Mokhtar
Annaba, 23000, Algeria

Abstract— track the vehicle and provide assistance to the driver is a fundamental research area in the computer vision community and it handles some common challenges, such as occlusion, deformation, motion blur, scale variation, and more. Videos captured at day or night time contains heterogeneous object, interacting object, edge ambiguity, and shape complexity where. The most reliable information at night is headlights. In the proposed study, detecting and tracking moving vehicles on videos with Kalman filter at night is implemented with taking into consideration some preprocessing step and problems of night before a state of the art of works related with this filter .

Keywords— *Vehicle tracking system, Kalman filter, vehicles on motion at night, video of traffic.*

I. INTRODUCTION

Vision-based traffic surveillance an indispensable part of Intelligent Transport System (ITS), has been widely studied over past few years [1] .Many applications, including transportation planning, traffic operating and highway capacity analysis, are based on the vehicle detection, tracking and classification .video monitoring system can obtain useful traffic information from collected video image sequences by using image processing and analyzing technologies, combining with booming pattern recognition, machine learning and other correlation theories. Detection and tracking are both most dynamic research area with number of application including computerized video surveillance, robotic vision, traffic detection, vehicle navigation, object identification and much more. It have been challenging tasks of classical computer vision and image processing research because of the issues such as partial or full occlusion of objects, illusion, camera shaking, extremely high or low-quality picture, distinct weather conditions including rain, snow and wind that complicate the vehicle detection, tracking, and data association processes, and in some cases, these problems make such systems completely fail. To track a vehicle at night is a difficult task, in view of the appearance of the vehicle headlights. In this paper, we extracted frames from video, we located the vehicle before that we treated frames and we tracked target vehicles with Kalman filter.

II. RELATED WORKS

Huda Dheyauldeen Najeeb et Rana Fareed Ghani[2] presented a survey contain all method of detection,representation,feature extraction and tracking in soccer videos wich are obtainable in related work.

Su-rong Qulet. al[3] improved a new kernelized correlation robustness tracking algorithm fixed on kalman filter and motion model to improve the generalization ability for object tracking in complex scenarios(fast moving, severe occlusion).Vannat Rin et Chaiwat Nuthong [4] predicted the collision risk by proposing a reel time detection system of a front moving vehicle using shifting three frames differences and tracking with kalman filter. Vijay Paidi et. al [5] demonstrated a novel methodology to detect the cruising patterns of multiple vehicles moving in an open parking lot using deep learning algorithm and kalman filter for tracking .

Mengxin Li et. al[6]have displayed The advantages and disadvantages of multi-target detection technology and multi-target tracking technology . The performance of multi-target detection and tracking technologies are compared, the future development direction of these two technologies are discussed.

Pin Wang et. al[7] proposed a method based on fuzzy particle filter of vehicle target recognition to introduce nonlinear statistics at night toward realizing process of vehicle target detection and resolved the different problems of night vehicle target recognition and detection .

Vasanthadev Suryakalaet. al[8] have used a modified Mixture Model (GMM) and Kalman filter to detect and track the moving vehicles at a near real time pace using one modality camera sensors,compared detection with GMM at day and nightbtime . Chanhoo Leeet.al[9] described vision-based lane detection algorithm in real-time, gradient cue and a color cue together and a line clustering with scan-line tests to verify the characteristics of the lane markings. It worked until high noise level and road conditions

Xuan Heet.al[10] performed a real-time algorithm during night applying kalman filter for lane tracking where they take into consideration various night scenarios such as rainy weather and tunnel travel.Md. Haidar Sharif [11] highlighted the comparative study of existing approaches

about laser-based detection and tracking algorithms using laser scanners rather than common video recording RGB cameras. R. Yao et al [12] have provided a survey of the video object segmentation and tracking VOST and Video Object Segmentation (VOS), summarized some object representation, image features, motion cues, and more. In addition, they summarized the characteristics of the related video dataset and supply variety of evaluation metrics. Amir Salarpour et. al[13] have described a combinational method for tracking multiple vehicles beyond kalman filter and color feature .It is used to minimize the computation time, and exploited the feature of vehicles for the best tracking. Weiqiang Zhang[14] generated data association process of the tracking objects through computing the likelihood probability. They exercised Kalman filter for detecting and tracking dynamic big and small objects by geometry model and a motion model were utilized to describe tracking static and dynamic object.

Jahongir Azimjonov et Ahmet Ozmen[15] prepared a new vehicle dataset and used two methods of detection Yolo and CNN combined with Yolo .As well as used of Kalman filter to track vehicles and the bounding-box-based vehicle tracking algorithm was developed where they raised classification accuracy.

III. DETECTION PROCESS

In image processing, identify the objects of interest and cluster pixels of these object is the first importante process. Object detection aims to detect the size, direction, and other information for the target object using the geometric and object characteristics of image preprocessing for the purpose of localizing them on a video frame. Detection of the region of interest in frames can be achieved by various methods which include frame difference method, optical flow method and background subtraction where techniques of detection are shown in figure 1.

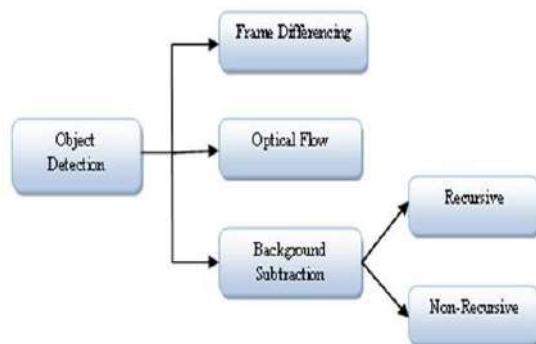
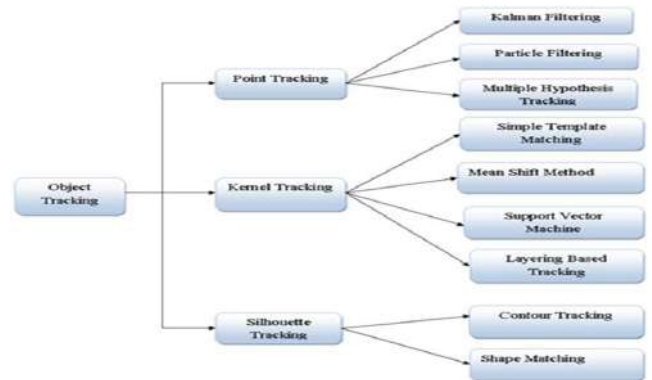


Fig. 1. Object Detection methods

IV. TRACKING PROCESS

Object tracking is a technique for re-identifying of the detected objects and associating them to the best-matched peers through consecutive frames. It is the process of finding any object of interest in the video to get the useful information by keeping tracking of its orientation, motion and occlusion etc, to clarify features notably used : Pixel,



shape, color, and bounding box (Bbox) information are widely used to trace the detected objects and to extract the

object trajectories..The methods including Kalman or Particle Filter tracking algorithms, which use bounding box info, are slightly faster than pixel, shape, or color-based approaches because of coordinated information of the detected objects. Tracking techniques are : point tracking,

kernel tracking and silhouette tracking have been appeared in figure2.

V. PROPOSED METHOD

In this paper we used detection by background subtraction and tracking using Kalman filter before extraction of frames from videos and preprocessing of night images where these steps are discussed later.

A. Used dataset

videos are captured in Annaba city summer season at night in defferents location with 2 types of cameras dome and fixe from defferent field of view. Figure 3 shows some frames from data set.



Fig. 3. .some frames from Data set

B. Extracting frames from videos

we divided videos on sequences of frames, in other words, videos contained around 2min can be fractionated on 3060 frames. Extraction frames from videos are displayed in [13] in a daytime. Figure 4. Indicate cited example.

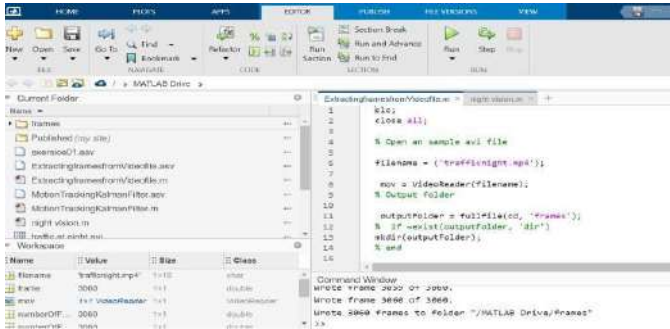


Fig. 4. example of extracting frames from video

C. Preprocessing step

1-Difficulties of lane detection during night

Compared with the uniform illumination at daytime. The difficulties of lane detection during night include :

- (01) The overall gray level of the night image is lower than that of daytime image and a large number of lane boundary points are submerged in the background noise .
- (02) The alternating light spots in image tend to cose edge missing.
- (03) The headlights affect the illumination distribution of the night image .(04) The interference of headlights and street lamp make it difficult for image segmentation

As example [10] have been deleted image noise based on two dimensional Gabor filter. Therefore the major problem in night detection is headlights because they moved and can be detected as the object target.

2-Edge Detection Filters:

It is a common image processing technique, and can be used for a variety of applications such as image segmentation, object detection, and Hough line detection. Use edge detection effectively by using the 'edge' function in MATLAB®, and also explore the different available parameters. There are Different Edge Detect Filter Like prewitt, Sobel, Canny and Robert& more. We discuss Sobel, Canny, log, laplacien and Prewitt Filter In this paper.

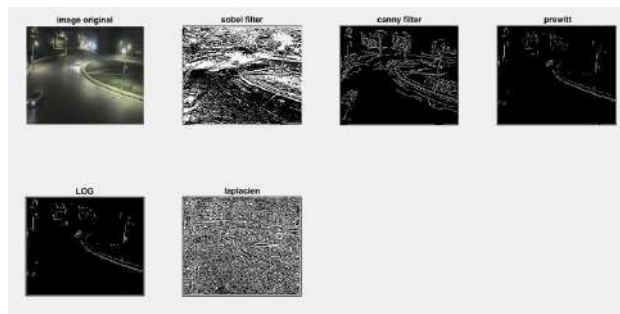


Fig. 5. comparison between original images and images filtered

As shown in figure 5 Sobel detection refers to computing the gradient magnitude of an image using 3x3 filters. Where "gradient magnitude" is, for each a pixel, a number giving the greatest rate of change in light intensity in the direction where intensity is changing fastest. Canny edge detection goes a bit further by removing speckle noise with a low pass filter first, then applying a Sobel filter, and then doing non-

maximum suppression to pick out the best pixel for edges when there are multiple possibilities in a local neighborhood. That's a simplification, but basically its smarter than just applying a threshold to a Sobel filter, but it is still fairly low level processing.

D. DETECTION BY BACKGROUND SUBTRACTION

Background modeling is an initial step for background subtraction method. It consisted on building background model where we can obtained the reference model, it will be useful in background subtraction. Every video sequence is compared with reference frame to identify deference between frames wherever object moving was detected. This detection method is easy, wherever the external environment changes, it is more sensitive. For the background subtraction two types of algorithms are available: recursive and non recursive process.

Recursive process: In recursive technique no buffer is used. Based on each input frame a single background model is updated. So the current model detect a mistake caused *from* frames of the distant past. This decreases the storage space, in this case no necessity of memory to buffer for the data. There are a quantity of recursive process that includes median, Gaussians Mixture and Kalman filtering used in our work.

E. TRACKING BY KALMAN FILTER

The Kalman filter is a cyclic algorithm to estimate the evolving state of a process when making

measurements in the process .It was used to track an object in sequence with captured video [16]. In addition it is a linear approach that operates in two basic phases of prediction and correction update (when the evolution of the case follows the linear motion model and the measurements are linear functions of the case). The prediction phase is accountable for the scoop of the next state and position of the present object. However the correction phase provides the parameters with their instance, they combine the actual measurement with the previous estimate to improve the trajectory where the object information detected in the previous frame is used and provides an estimate of the object's new position, It has the ability to rating the tracking locations with minimal datum on the location of the object steps are explained in figure 6.

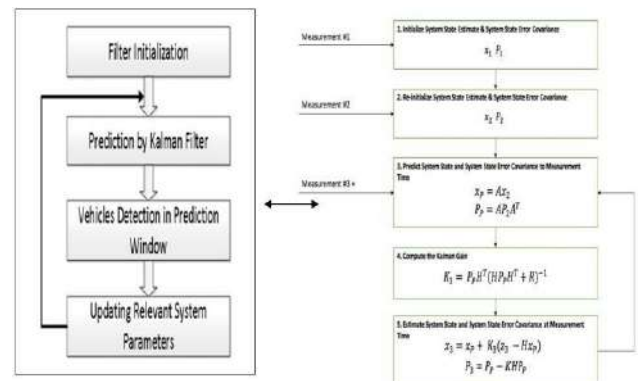


Fig. 6. Kalman tracking steps.

VI. EXPERIMENTAL RESULTS

Detection and tracking of cars on motion at night presented lot of difficulties, the most of them are cars headlights .In

this paper we processed captured pictures from monitoring videos taken on Annaba city located in the east of Algeria with infrared camera on summer season before that we preprocessed images using hybrid filters of canny, sobel and prewitt . We detected by background subtraction and tracked vehicles using kalman filter where the utilized filter facilitates maintaining data association during moving, stationary, and missed detection where. Figure 7 shows results of detection and tracking.



Fig. 7. Results of Kalman filter tracking.



Fig. 8. some frames results by background subtraction

VII. CONCLUSION

In this work we presented our method of detection and tracking on real videos captured of vehicles on motion at night time captured in Annaba city especially in summer season using infrared static and dome cameras ,accordingly the same vehicle can be seen in diverse angles of view. Results appears that canny and laplacian filter are best filters to detect edges of object with more and pertinent details before that kalman filter is used to track cars .

VIII. MATERIAL USED

This videos has been realized under a machine that has the following configuration: Laptop DELL Hard Disk 256 NVME. Windows 10 64-bit operating system, ram: 8, 00 GO. Processor: Intel(R) Core(TM) i5 CPU 8th generation. Images acquired and analyzed using Matlab 2020a version It is associated with toolboxes called TOOLBOX allowing access to functions specific to an application field such as image processing .

IX. PERSPECTIVES

- We will create a new Data set of traffic video captured on Annaba city with deferent conditions.
- We will test a combining filter on images to solve other night problems
- We will write a review article about detection, tracking and classification of vehicles on motions.
- We shall redact a comparative study between all methods existing in classification of cars on motion using deep learning.

REFERENCES

- [1] G. Hinz, C. Guang, M.A. Rajput, "Online multi-object tracking-by-clustering for intelligent transportation system with neuromorphic vision sensor", German/Austrian conference on artificial intelligence, pp. 1322–1341, 2017
- [2] H.D. Najeeb, R .F. Ghani, "a survey on object detection and tracking in soccer videos", Muthanna Journal of Pure Science, vol. 8, no. 2, 2021
- [3] Q. Su-rong, L. Jiangtao, S. Yuan, "Accurate vehicle location and tracking algorithms based on improved kernelized correlation motion model and Kalman filter in intelligent transport surveillance system", Journal of Ambient Intelligence and Humanized Computing, pp.2311–2334, November 2019
- [4] R. Vannat , N. Chaiwat, "Front Moving Vehicle Detection and Tracking with Kalman Filter", IEEE 4th International Conference on Computer and Communication Systems, vol. 28, no. 2, pp. 254– 260, 2019
- [5] P. Vijay, F. Hasan, H. Johan and G. N. Roger, "Tracking Vehicle Cruising in an Open Parking Lot Using Deep Learning and Kalman Filter", Journal of Advanced Transportation, pp. 12 , August 2021
- [6] L. Mengxin, H. Jiali, L. Songang, H. Yu and L. Yiyang, "A Survey of Multi-Target Detection and Tracking Algorithms", International Journal of Computational Intelligence Systems and Applications, Vol. 8, no. 1, pp. 1-11, 2021
- [7] W. Pin, F. En and W. Peng, "Night vehicle target recognition based on fuzzy particle filter", Journal of Intelligent & Fuzzy Systems Vol. 38 ,pp. 3707–3716 , 2020
- [8] S. Vasanthadev, T. Rajalakshmi, S. Kolangiammal and Harshit A. , "A Novel vision based embedded framework system to detect and track dynamic vehicles", IOP Conference Series: Materials Science and Engineering, 1130 012051 , 2021
- [9] L. Chanho, M. Ji-Hyun, "Robust Lane Detection and Tracking for Real-Time Applications" , IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, vol. 25, no. 3, pp. 727–745, 2016
- [10] H. Xuan, D. Zhengyu, C. Chuan, Y. Feng, "Video-Based Lane Detection and Tracking during Night", International Conference of Transportation Professionals, China, July 6–8, 2019
- [11] S.M. HAIDAR, "Laser-Based Algorithms Meeting Privacy in Surveillance: A Survey", journal of IEEE ACCESS, VOL. 9, July 2021.
- [12] Y. RUI, L. GUOSHENG, X. SHIXIONG, Z. JIAQI and Z. YONG, "Video Object Segmentation and Tracking: A Survey", ACM Transactions on Intelligent Systems and Technology, Vol. 11, No. 4, Article 36. , May 2020.
- [13] A. Salarpour, A. Salarpour , M. Fathi and D. MirHossein, "VEHICLE TRACKING USING KALMAN FILTER AND FEATURES" , An International Journal (SIPIJ Signal & Image Processing) Vol. 2, No. 2, June 2011
- [14] W. Zhang, S. Choi, J. Chae, Y. Sung and K. Choi, "object tracking method based on data computing", An International Journal of High-Performance Computer Design, Analysis, and Use, vol. 41, no. 1, pp. 309–332, 2018
- [15] J. Azimjonov, A. Ozmen, "A real-time vehicle detection and a novel vehicle tracking systems for estimating and monitoring traffic flow on highways", journal of Advanced Engineering Informatics, vol. 50, pp. 1474-0346 , 2021
- [16] Q. Li and al. "Kalman filter and its application," In International Conference on Intelligent Networks and Intelligent Systems (ICINIS), pp. 74-77, IEEE, November 2015.

Service request monitoring using Petri nets and model at runtime

Mohammed Mounir Bouhamed
MISC Laboratory

Department of Computer Science and its Applications
University Constantine 2 Abdelhamid Mehri, Constantine,
Algeria.
mohammed.bouhamed@univ-constantine2.dz

Gregorio Diaz
School of Computer Science
University of Castilla-La Mancha,
Campus Universitario
s/n, 02071, Albacete, Spain
gregorio.diaz@uclm.es

Oussama Kamel
MISC Laboratory

Department of Computer Science and its Applications
University Constantine 2 Abdelhamid Mehri, Faculty of
Medicine, University Constantine 3 Salah Boubnider,
Constantine, Algeria.
oussama.kamel@gmail.com

Allaoua Chaoui
MISC Laboratory
Department of Computer Science and its Applications
University Constantine 2 Abdelhamid Mehri, Constantine,
Algeria.
chaoui2001@yahoo.com

Abstract—The service computing paradigm has been widely adopted in various domains, such as in Web applications, Cloud computing and Internet of Things (IoT). Service providers offer services to their clients depending on their functional and non-functional requirements. Generally, the process of using a service consists of three steps discovery, selection, and invocation. This process is subject to many variables. For instance, no service with the required properties was found implying a requirement update, or just the invocation does not work. Therefore, the service monitoring is a necessary feature to be included by service providers. In this paper, we propose a model at-runtime approach to monitor this three-steps-process. We specify the proposed model using the Petri net formalism and the Flexible framework. Our approach allows service managers to monitor the three steps, change requirements when no service can satisfy the current requirements, and select other services when invocations fail. The monitoring system notifies service managers about the request state, the requirement changes required, and the invocation failures based on the Petri net marking. In addition, Python expressions can be used at runtime. Finally, a case study illustrates the approach by the implementation of a compression service.

Index Terms—Service oriented computing, Petri Nets, model@runtime, monitoring

I. INTRODUCTION

Recently, the service computing paradigm has gained more and more popularity [1]. The main advantage of this paradigm is that every part of the hardware and software architecture is offered as a service, which even includes the hardware infrastructure itself where other services are running at the same time. Nowadays, services have been expanded in various areas, such as: cloud computing, Internet of Things (IoT), fog computing, etc.

Clients request services to satisfy their desired needs specified in terms of functional and non-functional requirements. Generally, service clients follow a three-step process to use a service. The first step is the service discovery, search based on a set of functional and non-functional requirements. The usual result is a list of (functional) equivalent services. The second step is the service selection. In this step, Clients may need to come back to the previous step, if an empty list is returned, that is, they must change their combinations of functional and non-functional requirements. Otherwise (a non empty list is returned), the non-functional requirements (such as price, response time and reliability) are used to select the most adequate service [2]–[4]. The last step is the invocation of the selected service. Requesting a service might end without satisfying the client's requirements due to the dynamic nature of the environments where these services are used. For instance, either services discovered might not answer in the agreed response time, or we might even get no response at all when calling them. Consequently, the process of service requesting has to be monitored by service managers to detect any unexpected problem and apply corrective strategies when necessary. Many related works [2]–[12] emphasize the importance of monitoring service requests and their QoS. In general, the approaches proposed [13]–[15] offer a high level of abstraction to monitor service and do not consider the granularity of service requests. We need to verify some properties to monitor a service request, such as deadlock freedom, service delivery assurance and QoS satisfaction. Hence, using a formalism like Petri Nets (PNs) to model the process of a service request is of great importance. PNs and other formalisms are used to model systems but only at build-time without linking them to the concrete implementations of

these systems. Therefore, the adoption of a model@runtime solution to link service models and their invocation is required. A model@runtime is a solution to model and monitor a running system. It is based on computation reflection, in which the system state is reflected on the model and vice versa. In our previous work [16], we developed a model at-runtime framework that models Python applications at runtime using an extension of the PN formalism. In this work, we use this framework to model and monitor a service request.

The contributions of this work are:

- We propose a PN model that specifies the three-step process of using a service. This model considers both the fulfillment and the failure to fulfill the client requirements.
- We implement the proposed model using the model at-runtime framework proposed in our previous work [16]. This framework allows the monitoring of service requests.
- A case study illustrates the benefit of our work.

The next section presents the necessary background to understand this work, and some related works. Section III presents our proposal to model and monitor a service request. Section IV presents a case study that illustrates our approach. In Section V, we discuss our approach. Finally, we conclude this work in Section VI.

II. BACKGROUND AND RELATED WORKS

This section outlines the background concepts used in this work. We focus on the PN formalism used in this work to model a service request. Then, we explain the notion of model@runtime, and we present the architecture of our model at-runtime framework introduced in our previous work [16]. This framework is used to develop the approach presented in this work. Finally, we present some related works.

A. Petri Nets

A PN [17] is a formalism used to specify system behaviors. It is a directed bipartite graph that has two types of nodes: places and transitions, which are connected using weighted arcs. Places are represented as circles where they can hold a finite number of tokens, whereas transitions are represented as rectangles.

Definition II-A.1. Petri Net (PN)

A PN [17] is a tuple $\Gamma = (P, L, F)$ where
 (i) $P = (P_0, P_1 \dots P_m)$ is a finite set of places.
 (ii) $T = (T_0, T_1 \dots T_n)$ is a finite set of transitions.
 (iii) $F : (P \times T) \cup (T \times P) \rightarrow \{0, 1, 2 \dots\}$ is a set of weighted arcs, which represent the flow relation.

A marking m , which is the number of tokens in each place, represents the current state of the system. For instance, assuming $m = [1, 1, 0]$ and $P = \{P_0, P_1, P_2\}$, places P_0, P_1 , and P_2 contain 1, 1, and 0 tokens, respectively.

A Transition T_i is enabled to be fired if its input places contain at least the number of tokens in each associated arc $(P \times \{T_i\})$. If T_i is fired, several tokens that equals $(\{T_i\} \times P)$

is consumed from the output places of T_i . Firing T_i leads to a new marking m' , and it is noted as $m \xrightarrow{T_i} m'$. \square

Figure 1 depicts a PN that represents the behavior of a grading system.

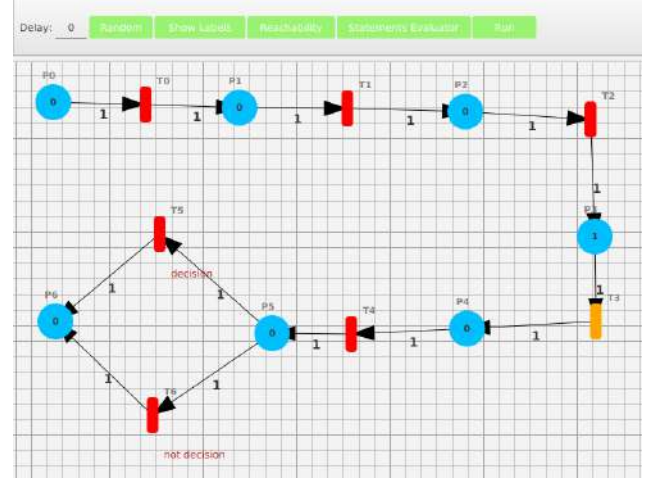


Fig. 1. An example of a PN [16]

B. Model@runtime

MDE is a software engineering paradigm. It assists in software building and maintenance with the aim of increasing productivity. It provides a software development process based on models. A model [18] is a system abstraction that can focus on a specific aspect, such as structure and behavior. Models@runtime [19], [20] is a solution that considers running systems. In this solution, the system and the model are causally connected, meaning that they are updated according to the state of the other one. A benefit of the use of the models@runtime is a richer semantic based on the state of the system, which helps in decision-making [20]. For instance, Figure 1 shows the state of a running system that has fired the transitions T_0 , T_1 and T_2 and is about to fire T_3 , where T_3 is colored in Orange. The transitions represent real actions that calculate a student grade. Therefore, models@runtime will support running systems and decision-making by many things, such as reconfiguring systems [16]. This paradigm offers a model reflection of the running system at runtime. There are many techniques that are followed by researchers to provide models at-runtime, such as reflection [16], [21], heuristics, monitoring [6], model transformation [22], reasoning [23], etc.

In this work, we use a model at-runtime framework that was published in our previous work [16]. This framework, depicted in Figure 2, is used to develop, model and reconfigure Python applications at runtime. Python applications are modelled using an extension of PNs where the transitions of the PN formalism are extended to contain a Python statement and a guard. A statement is executed when its transition is fired. A guard is a Boolean expression that should be satisfied when its transition is fired.

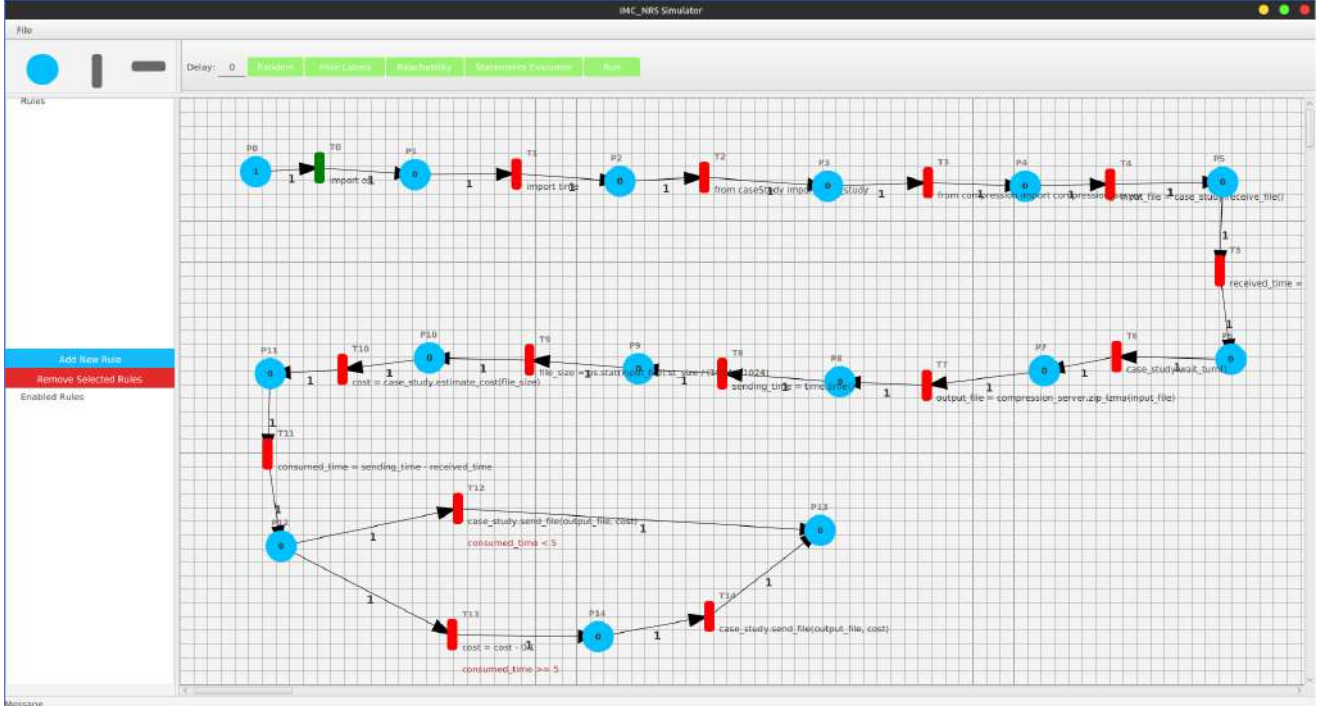


Fig. 2. The main interface of the model at-runtime framework

The framework has two components: the Model Execution Engine (MEE) and the Python Execution Environment (PEE). The former executes the model, whereas the latter executes Python statements that reflect the former execution. Figure 3 shows the architecture of this framework.

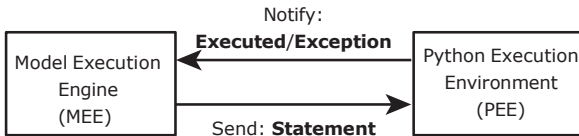


Fig. 3. The architecture of the model at runtime framework

A developer uses this framework starting by modeling the application behavior. Then, he develops the classes and functions that the application needs. Then, he/she enriches the model by using statements and guards. Finally, he runs the application and reconfigures it when necessary.

C. Related works

In the literature, several approaches have been developed to monitor service compositions. Pistore et al. [15] use a technique that is called “Planning as Model Checking” to compose web services and decide which external services have to be monitored. The aim of this work is to improve the internal behavior by dealing with external anomalies that may happen during service requests. Lazovik et al. [5] proposed a framework to assert whether business rules are

met during service requests or not. They proposed a language to specify business rules as sets of assertions, which have to be met during service requests. The framework also provides the ability to monitor service requests and determine when business rules are satisfied. Guzel et al. [7] propose a multi-objective optimization approach that considers QoS, energy consumption and fairness during the request of IoT services in fog-based networks. Based on the concept of edge devices as a service, Hayyolalam et al. [9] proposed a method to adopt and use AI on resource-constrained edge devices. They use meta-heuristics algorithms that consider QoS and QoE during subtask composition. Khan et al. [13] proposed a blockchain-based distributed infrastructure to monitor, in real-time, the violations occurred in cloud-hosted services. Yang [24] proposed a PN model for service composition in IoT environments. The proposed model evaluates cost-effectiveness by considering reliability, response time, and cost. Badshah et al. [14] proposed a third-party monitoring service based on the service level agreement. This service detects violations and manages penalties.

Sun et al. [25] proposed a framework to monitor runtime information during the execution of business processes that are specified using BPEL. Extended message sequence charts are used to store captured information. Boucetti et al. [10] proposed a QoS-aware approach based on genetic algorithms and neural networks to compose IoT services to consume an optimal time during service selection. Osypanka et al. [26] proposed a QoS-aware approach for resource prediction to reduce the costs of cloud computing. The proposed approach

works autonomously with the help of many techniques, such as continuous monitoring, anomaly detection, self-adaptation, machine learning, and the discovery of computing service characteristics. Nikam et al. [27] designed a middleware to provide monitored resources during service compositions in cyber-physical systems. Gomes et al. [28] proposed a model at-runtime approach which was built as a middleware to monitor the satisfaction of non-functional requirements in cloud services.

III. PROPOSAL

This section describes our solution proposed to model service requests at runtime. It presents a PN model that specifies service requests. In addition, it shows how to develop Python methods to bind the proposed model with Python. Finally, it shows the monitoring process.

A. Overview

The proposed approach requests services dynamically using the model-at-runtime framework developed in our previous work [16]. The same framework is used to monitor the whole process of service requests. This process consists of three tasks, namely: discovering, selecting, and invoking, which can all be monitored by the service manager. Figure 4 shows this process. The process of a service request is modeled as a PN, while the execution is performed using the model-at-runtime framework. The service manager monitors the service requests using the PN marking and the expression evaluator.

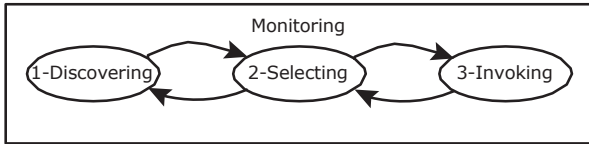


Fig. 4. A three-step process of a service call

Preparing the framework to call external services starts by developing Python methods to discover and select services. This process includes a mechanism to reduce the required non-functional requirements in case no service can satisfy them. Next step includes the service invocation of the service selected. The whole three-step process can be monitored by the service manager. Figure 5 shows the steps followed by the service manager.

B. Service Call as a Petri Net

The PN presented in Figure 6 models an external service call. This process starts by discovering services satisfying the functional requirements and the initial non-functional requirements, such as price and response time. Three different transitions can be fired depending on the outcome. First option is offered when the list of services is empty, but we still can reduce our aspirations regarding the non-functional requirements. The reduction is accounted by adding a token in the Place "Reduction counter", and the discovering process

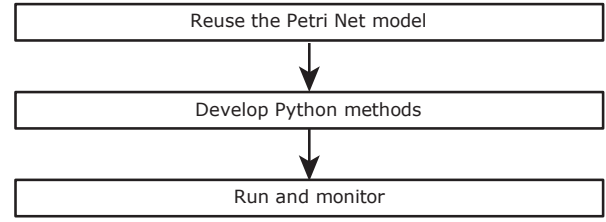


Fig. 5. Steps of configuring a service call

is retaken. Otherwise, the service call terminates in a failure, when we cannot reduce the non-functional requirements, second option. In the last option, a list of services has been discovered satisfying the current non-functional requirement set. In this case, a service is selected from the list and invoked. If the invocation is performed without problems, the service call ends correctly. Otherwise, the failure is accounted in the "fail counter" Place, and the selection process is retaken again discarding this service. If there are no remaining services, the discovery process is restarted.

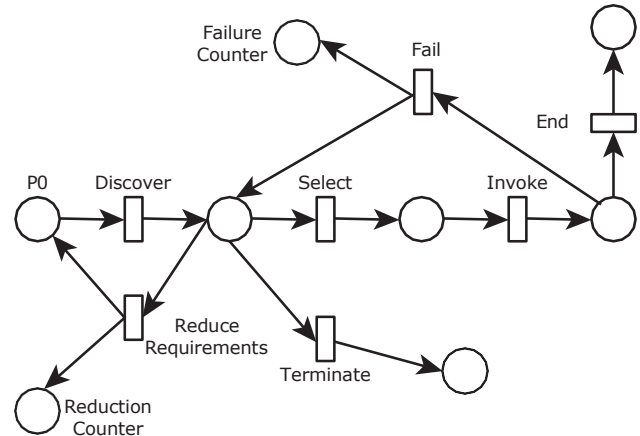


Fig. 6. A Petri net to model a service call

C. Preparing Methods

Table I shows the statements and guards that are associated with each transition of the PN depicted in Figure 6. The transition "Discover" executes the method "discover" with the inputs minimum and maximum requirements. This method has to be developed by the service manager when he has to specify how the discovery is done. Several approaches that can be followed are presented in the literature [12], [29]–[33]. The transition "Select" calls the method "select". Many selection strategies can be implemented, such as the first service on the list. This transition is only fired when exists a list of discovered services, restriction implemented by the guard $len(discovered_services) > 0$ and it implies the selection of a service. Afterwards, the list is updated by removing the service selected. Subsequently, the transition "Invoke" exe-

cutes the “*invoke*” method to invoke it. The transition “Reduce requirements” reduces the aspirations of the non-functional requirements. The service manager should establish a policy to select the requirements to be sacrificed and their order. He/she can delegate this task to an optimization algorithm, for instance, the algorithm presented by Boulesnane et al. [34]–[36] can be considered. This transition is only fired when there is no discovered service with the current requirements, and we can reduce the requirements. The transition “Terminate” is fired, if there is no available service with the minimum requirements to invoke. It should notify the service manager or a dedicated service to act in this case. The transition “Fail” is fired when the invocation fails and executes the “*fail*” method, while the transition “End” is fired otherwise executing the “*end*” method.

D. Monitoring

The service manager can monitor service calls using the marking and the expression evaluator. The PN from Figure 6 shows the number of requirement reductions in the place “Reducing counter”, and the number of invocation failures in the place “Failure counter”. The number of tokens represents the number of reductions and failures. This marking represents the state of the actual execution. For instance, Figure 7 shows that the requirements are reduced two times, the invocation failures are three times, and there is a selected service to be invoked. The service manager could change the proposed PN to monitor other parts. For instance, he could monitor the discovery. He could also change the behavior of the service call by executing the discovery and the selection steps simultaneously, for example, changing the initial marking.

The expression evaluator can be used to monitor the value of Python expressions. For instance, Figure 8 shows the value of the grade average variables used by a service in charge of performing the evaluation process.

IV. CASE STUDY

This case study illustrates our approach showing the call and execution of a compression service. In this scenario, we used 20 services able to compress files. Each service has a different compression ratio average, but cannot ensure that files will be compressed with this ratio. “discover” method provides the list of services and their compression ratio. The requirements are initialized with a compression ratio of 5:1, and we consider a reduction step of 1. The lowest accepted compression ratio is 2:1. If the compressed file is lower than this ration, the invocation is considered as a failure.

The simulation shown in Figure 9 shows the final state of the compression process. It ends correctly, but requirements were downgraded twice and five failures occurred. The failures happened either because the service was unavailable or because the compression ratio offered by the compression service selected was inferior to 2:1, which is unacceptable.

V. DISCUSSION

The primary goal of this work is to monitor service requests. Therefore, the proposed approach contributes to the field by:

- proposing a model at-runtime to monitor a service request;
- monitoring the satisfaction of requirements during the service request; and
- monitoring the resources consumed, in terms of discovered, selected and invoked services, during the service request.

We model a service request as a PN, which enables the formal verification of desired properties. The proposed model offer alternatives to service clients at the discovery and selection stages in case their requirements are not satisfied. In addition, it allows to inspect the number of failures produced during invocations. Another feature offered by this model is the downgrading mechanism available to adapt the client requirement to the actual services. For instance, Figure 9 shows the number of failures and requirement downgrading occurred during the execution of the case study. The model has to be completed with different Python methods to implement the three steps: discovery, selection, and invocation. It is worth nothing that this model can be adapted according to the service manager’s needs.

VI. CONCLUSION

This paper introduces a novel approach that adopts model@runtime and Petri nets formalism for monitoring the process of service requests. Petri nets are used to model the three-step process of service requests, namely : discovery, selection, and invocation. The proposed solution enables the verification of the fulfillment of the clients requirements. It enables also the detection of any anomalies that may appear during this process, for instance, either the available services cannot satisfy the non-functional requirements, or the invocation of a selected service cannot be accomplish. The used framework provides also relevant information such as the number of invocation failures, the number of requirements’ changes, etc.

Future work includes the development of a GUI that can deal with business processes in which several instances invoke several services. It could be accomplish by considering Color Petri Nets (CPN) to associate each token with a specific service request. Moreover, we intend to study the problem of service composition using the proposed framework and study different selection and discovery algorithms. Other domains can also be considered with this solution such as real cloud and IoT environments.

ACKNOWLEDGMENT

This research was supported by the Spanish Ministry of Science, Innovation and Universities and the European Union ERDF funds through the FAME project under grant RTI2018-093608-B-C32, and the JCCM project cofinanced with ERDF funds under grant SBPLY/17/180501/000276. It was also supported by DGRSDT of the Algerian Ministry of Higher Education and Scientific Research.

TABLE I
THE TRANSITIONS AND GUARDS OF THE PN THAT MODEL A SERVICE CALL

Transition	Statement	Guard
Discover	$discovered_services =$ $discover(minimum_requirements, maximum_requirements)$	<i>True</i>
Select	$chosen_service = select(discovered_services);$ $discovered_services.remove(chosen_service)$	$len(discovered_services) > 0$
Invoke	$invoked = invoke(chosen_service)$	<i>True</i>
Reduce requirements	$maximum_requirements = minimum_requirements;$ $minimum_requirements =$ $reduce_requirements(minimum_requirements)$	$len(discovered_services) == 0$ and $can_reduce_requirements$ $(minimum_requirements)$
Terminate	$terminate()$	$len(discovered_services) == 0$ and $not\ can_reduce_requirements$ $(minimum_requirements)$
Fail	$fail()$	$not\ invoked$
End	$end()$	$invoked$

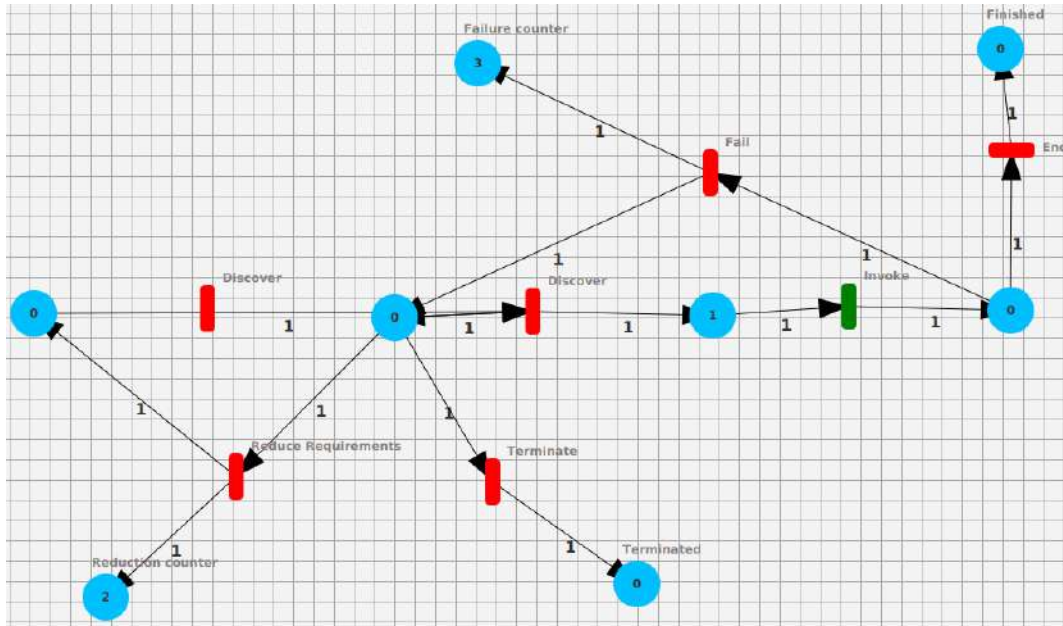


Fig. 7. Monitoring the state of service call

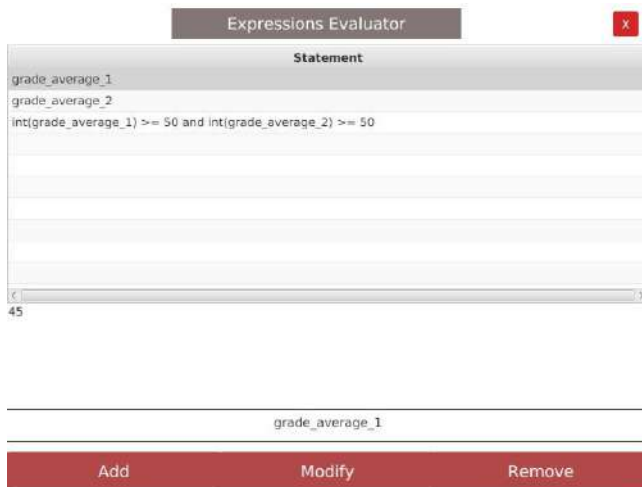


Fig. 8. Monitoring using the expression evaluator

REFERENCES

- [1] A. Bouguettaya, M. Singh, M. Huhns, Q. Z. Sheng, H. Dong, Q. Yu, A. G. Neiat, S. Mistry, B. Benatallah, B. Medjahed *et al.*, "A service computing manifesto: the next 10 years," *Communications of the ACM*, vol. 60, no. 4, pp. 64–72, 2017.
- [2] B. Maati and D. E. Saidouni, "Ciotas protocol: cloudiot available services protocol through autonomic computing against distributed denial of services attacks," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–30, 2020.
- [3] M. Alrifai, D. Skoutas, and T. Risse, "Selecting skyline services for qos-based web service composition," in *Proceedings of the 19th international conference on World wide web*, 2010, pp. 11–20.
- [4] O. Kamel, A. Chaoui, G. Diaz, and M. Gharzouli, "Sla-driven modeling and verifying cloud systems: A bigraphical reactive systems-based approach," *Computer Standards & Interfaces*, vol. 74, p. 103483, 2021.
- [5] A. Lazovik, M. Aiello, and M. Papazoglou, "Associating assertions with business processes and monitoring their execution," in *Proceedings of the 2nd international conference on Service oriented computing*, 2004, pp. 94–104.
- [6] P. Cedillo, E. Insfran, S. Abrahão, and J. Vanderdonckt, "Empirical evaluation of a method for monitoring cloud services based on models at runtime," *IEEE Access*, vol. 9, pp. 55 898–55 919, 2021.

- [33] A. M. Alkalbani and W. Hussain, "Cloud service discovery method: A framework for automatic derivation of cloud marketplace and cloud intelligence to assist consumers in finding cloud services," *International Journal of Communication Systems*, vol. 34, no. 8, p. e4780, 2021.
- [34] A. Boulesnane and S. Meshoul, "Effective streaming evolutionary feature selection using dynamic optimization," in *IFIP International Conference on Computational Intelligence and Its Applications*. Springer, 2018, pp. 329–340.
- [35] —, "Reinforcement learning for dynamic optimization problems," in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 2021, pp. 201–202.
- [36] —, "A modified wind driven optimization model for global continuous optimization," in *International Conference on Hybrid Artificial Intelligence Systems*. Springer, 2015, pp. 294–304.

Sign Language Recognition

Imed Eddine Zeroual

Mathematics and Computer Science Department
Mohamed-Cherif Messaadia University - Souk Ahras
Souk Ahras, Algeria
zeroualimededdine@gmail.com

Houssem Eddine Azzag

Computer Science Department
Ecole Supérieure en Informatique - Sidi Bel Abbe's
Souk Ahras, Algeria
h.azzag@esi-sba.dz

Abstract—Deaf and dumb people suffer a lot in their dealings, not only this, but their families and relatives also suffer from a big problem in communicating with them, and there are not many incentives that make ordinary people ready to learn the language of deaf and dumb people to speak with them in their own language easily, so a systems capable of converting the sign language must be made Referring to a readable or audible language to be a means of linking between a normal person and a person who does not speak sign language, this currency will greatly facilitate transactions with deaf and dumb people and will also give them more opportunities , In this work, I proposed a new system for recognizing sign languages and translating them into a readable and audible language. We also tried many artificial intelligence methods such as YOLO, KNN, CNN, and others, and we were able to get very satisfactory results, and by using our dataset .

Keywords : Sign language recognition, YOLO , Deep learning .

I. INTRODUCTION

Sign language is the official language for the deaf and hard of hearing people or people who have speech problems due to disability. It is a non-vocal language that depends on the movements of the hands and arms and some other movements. There are many sign languages around the world (144 sign languages [1]) that differ from one language to another or From one country to another, Figure . 1. shows the difference between USA and British sign languages.

The difference between them lies in the way the meanings are expressed in different movements. Each number or letter has a specific movement that expresses it in sign language, and these movements are used to spell words Or know the numbers to communicate between its users.

Ordinary people do not learn sign language because it is a language used by a small group of people and cannot be used permanently, as well as difficult to learn, which makes it marginalized and of no importance to them, and this increases the complexity of the matter in terms of communicating with people who only understand sign language, and in fact this is not Just, and makes the deaf and people with special needs in a world isolated and backward from the developed world in which ordinary people live, but with the modern technological development, movement recognition systems appeared through different technologies, gloves equipped with sensors

[2] or pictures and videos. Communication, understanding and interaction through social networking sites.

Artificial intelligence technology interferes with its techniques in machine learning and deep learning, where the machine can learn and recognize movements and translate them from sign language to audio and readable language for the general public, The machine is trained on the movements to be recognized, using an appropriate database that contains images of various types of movements. The data entered is processed and a trained model is created to produce a smart program.

An image captured from the user's device camera is entered into the smart program, the image is processed, and the hand is identified in the image as the first step, then the type of movement is recognized.

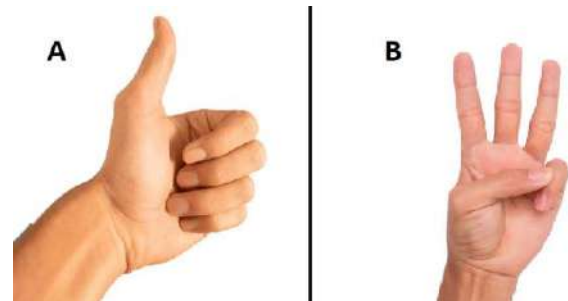


Fig. 1. Sign of number 6 , (A) British and (B) USA sign languages .

II. RELATED WORK

Because of the importance of the topic studied, we find that researchers have presented many approaches to solve some of the problems posed by this topic ,One of the most used approaches is Convolutional Neural Network (CNN) approach [6] [7] they proposed a specific CNN architecture and use it in a selected dataset but the problem faced with this approach that is worked only with selected backgrounds and fixed camera so it will not be used in all cases , In the same context of Deep Learning [7] are tried other method's as deep neural network (DNN) and Long Short-Term Memory (LSTM) and they get acceptable results in the selected dataset,

but the previously mentioned problem is still in doubt , other approaches are using Machine learning algorithms [8] they use K-Nearest Neighbor algorithm to classification phase using various backgrounds and effects and they also get very good results , [9] they choose Support Vector Machine (SVM) algorithm to work with and for result using dataset with 36 ASL and 12 people \times 1 set they get 75.50% accuracy for 26 classes and 67.54% for 36 classes also [10] they used the SVM algorithm using dataset with 10 ASL gestures and 14 people \times 10 sets for results they get 80.86% accuracy , for other approaches [11] used MultiLayer Perceptron (MLP) with 32 American Sign Language gestures for 146 people every people have 1 set and they get 90% accuracy.

III. PROPOSED MODEL

A. Model Definition

1) **Object Detection** : object detection Techniques that detect and identify the object through the image work to detect the objects in the image, after which they are classified to identify that object. This process takes place through two main stages, first, determining the coordinates and dimensions of the object in the image and creating a box surrounding the object in the image, second The process of classifying the extracted frame to identify the class of the object in it.

2) **YOLO** : You Only Look Once (YOLO) is a modern algorithm used in many object detection applications. YOLO is known for its speed and accuracy in real-time object detection. Yolo uses Convolutional Neural Networks (CNNs) to discover objects in real time. Yolo needs one stage of CNN to detect objects with a colored box and classify them Figure. 2 shows this process .that by divides the image into parts and predicts probabilities for each part. This technology has several versions: YOLO V2, YOLO V3, YOLO V4, YOLO V5 and more [5]. The yolo architecture consists of 24 convolutional layers and two fully connected layers. Figure 3 shows the full architecture.

3) **convolutional neural network (CNN)** : A convolutional neural network (CNN) is a widely used type of neural network. The CNN is made up of neurons that have weight, bias, and activating functions.

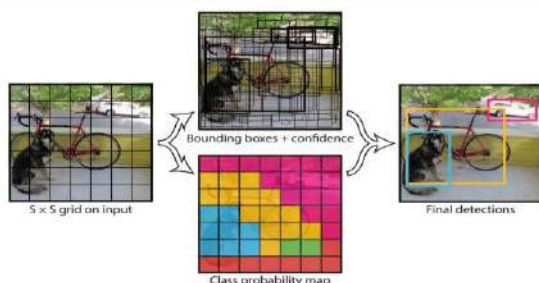


Fig. 2. Object Detection and Classification using YOLO [4]

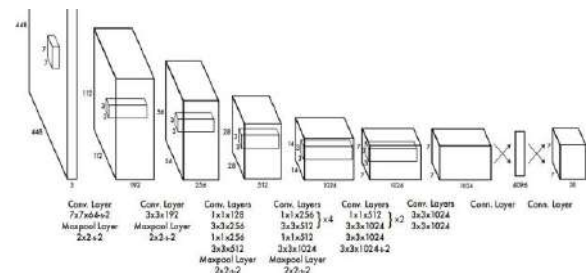


Fig. 3. YOLO architecture [4].

4) **Training details** : We have applied YOLOv3 technology to train a model that determines the position of the hand in the image and recognizes the type of signal. In the input, YOLO technology needs an image data-set in the form of a training part and a testing part. Also, each photo must be attached to a file with the same name with the photo that contains the location of the photo's hand frame in the original image, and its width and height, and a private key define the class of this sign.

After the training process, we get a ready -to -use trainer model, in the prediction part, we only enter a picture into the model, and the outputs are in the form of a colored box that determines the location of the sign in the image as well as the private key that represents the sign type, After that, we may get the desired, we define each special key with an audio file and a text file, so that they are displayed in the last with the image , Figure . 4 . shows the General proposed architecture

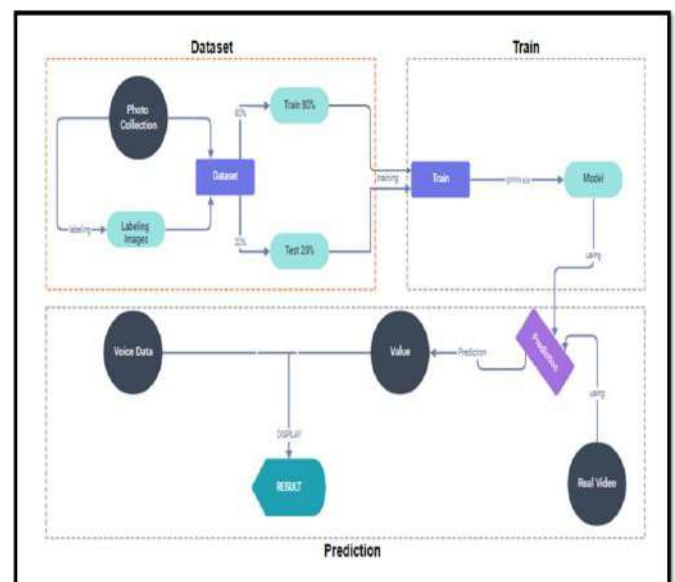


Fig. 4. General proposed architecture .

B. Dataset

Due to the scarcity of sign language datasets for YOLO, we had to create our own datasets. We have created a database of numbers from 0 to 9. The creation of this data set is based on taking pictures of a number of people while they make references to the numbers. Then the hand is identified in each image using a special program called LabelImg (Figure 5), where we determine the position of the hand in each image, and the application creates a file with the same name as the image and contains the width and height of the bounding box of the hand as well as the location of the box in the image (X, Y), (Figure 6). Finally, a private key is placed for each digital image. About 200 images (Figure 7 8) were taken for each digit, nearly 2,000 images for digit recognition. We identified the hand in each image and gave it a private key for the number in the file. Attached text with the same name as the image, using LabelImg.



Fig. 5. labeling image application.



Fig. 6. File created through the labelImg application.

C. Classification process

We use 80 % to use the training stage for the process of pre -processed data, and it uses 20 % for the test.

We divide the data set so that we guarantee the availability of all conditions in the training part as well as in the test part. We made adjustments from the original Yolo structure, as we change the number of channels and the number of filters in the tafiye class. In the training process, we use the training part of the data to build a Yolo model. Using Darknet53 (Figure. 9 . shows the full architecture) Brereated Weight [13-14] The training process , which was trained on imagenet, is done (Figure. 10 .) . The training will stop when the value of the loss stops to get off .

Specification	Value
Resolution	1280x720
Extension	.jpg
Number of images	2000
Number of class	10
Number of images per class	180-220
File size	50-500KB
Channel	3 (rgb)

Fig. 7. Dataset Specification.



Fig. 8. : A group of people doing different signaling movements.

IV. EXPERIMENTAL RESULTS

We implement our approach using Python Programming language , for training part we are work with a help of Google Colab train configuration showed in Table I and for testing part we work with configuration showed in Table II.

TABLE I
TRAIN ENVIRONMENT CONFIGURATION

RAM	12.68 GB
GPU	NVIDIA Tesla K80
GPU Memory	12GB / 16GB
Disk	78.19 GB
GPU Memory Clock	0.82GHz / 1.59GHz
Performance	4.1 TFLOPS / 8.1 TFLOPS
CPU Model Name	Intel(R) Xeon(R)
CPU Freq.	2.30GHz
No. CPU Cores	2

We work in this experiments with YOLO Algorithm for object detection and we get a very satisfying results in our case using our dataset .

Type	Filters	Size	Output
Convolutional	32	3 × 3	256 × 256
Convolutional	64	3 × 3 / 2	128 × 128
1x Convolutional	32	1 × 1	
1x Convolutional	64	3 × 3	
Residual			128 × 128
Convolutional	128	3 × 3 / 2	64 × 64
2x Convolutional	64	1 × 1	
2x Convolutional	128	3 × 3	
Residual			64 × 64
Convolutional	256	3 × 3 / 2	32 × 32
8x Convolutional	128	1 × 1	
8x Convolutional	256	3 × 3	
Residual			32 × 32
Convolutional	512	3 × 3 / 2	16 × 16
8x Convolutional	256	1 × 1	
8x Convolutional	512	3 × 3	
Residual			16 × 16
Convolutional	1024	3 × 3 / 2	8 × 8
4x Convolutional	512	1 × 1	
4x Convolutional	1024	3 × 3	
Residual			8 × 8
Avgpool		Global	
Connected		1000	
Softmax			

Fig. 9. Darknet [13].



Fig. 10. Training process.

V. CONCLUSIONS

Almost complete absence of people who are deaf and hard of hearing and who suffer from speech problems due to disability from using social networking sites and various fields of automated media and other fields. This made us think about looking for solutions to help them, and with the help of the latest artificial intelligence technologies, YOLO, we were able to create a smart program that allows translating numbers from one to nine from sign language into a readable and audible text.

We also created a special database of 2000 photos, containing numbers from 0 to 9 in the British sign language for YOLO technology, and the first of its kind as it consists of pictures

TABLE II
TEST ENVIRONMENT CONFIGURATION

RAM	4.00 GB
GPU	NVIDIA Geforce 920M
GPU Memory	2 GB
Disk	250 GB
Operating system	Windows 10 Pro
CPU	Intel(R) Core(TM) i5-5200U CPU @ 2.20GHz 2.20 GHz



Fig. 11. : Results.

of people who represent signal movements, Each image has an attached file with the same name that contains the location of the movement in the image, its width, its height and the type of movement .

We are aware that there are some open aspects of discussion, including: (1) Does this data-set achieve the evaluation measures, (2) Does this data-set include a complete sign language, (3) Does this data-set study all the obstacles that it may face on land Reality .

Given the massive progress in this field, we are determined to produce more versions, this is not the last version. We look at the expansion of the data-set to include all sign language as well as improvement in the accuracy of the results and achieve all standards. The upcoming versions will be more expanded for this first version.

REFERENCES

- [1] Ethnologue. 2019.Languages of the World. Available at http://www.ethnologue.com/15/show_family/90008/ (accessed April 1, 2019)
- [2] Ahmed, M. A., Zaidan, B. B., Zaidan, A. A., Salih, M. M., & Lakulu, M. M. B. (2018). A review on systems-based sensory gloves for sign language state of the art between 2007 and 2017. *Sensors*, 18(7), 2208.
- [3] Alawwad, R. A., Bchir, O., & Ismail, M. M. B. Arabic Sign Language Recognition using Faster R-CNN.

- [4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779–788.
- [5] Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, 30(11), 3212-3232
- [6] Rao, G. A., Syamala, K., Kishore, P. V. V., & Sastry, A. S. C. S. (2018, January). Deep convolutional neural networks for sign language recognition. In 2018 Conference on Signal Processing And Communication Engineering Systems (SPACES) (pp. 194-197). IEEE.
- [7] Bantupalli, K., & Xie, Y. (2018, December). American sign language recognition using deep learning and computer vision. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 4896-4899). IEEE.
- [8] Haque, P., Das, B., & Kaspy, N. N. (2019, February). Two-handed bangla sign language recognition using principal component analysis (PCA) and KNN algorithm. In 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE) (pp. 1-4). IEEE.
- [9] Chong, T. W., & Lee, B. G. (2018). American sign language recognition using leap motion controller with machine learning approach. *Sensors*, 18(10), 3554.
- [10] Yang, L., Chen, J. A., & Zhu, W. (2020). Dynamic hand gesture recognition based on a leap motion controller and two-layer bidirectional recurrent neural network. *Sensors*, 20(7), 2106.
- [11] Mapari, R. B., & Kharat, G. (2016, March). American static signs recognition using leap motion sensor. In *Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies* (pp. 1-5).
- [12] Daniels, S., Suciati, N., Fathichah, C. (2021, February). Indonesian Sign Language Recognition using YOLO Method. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1077, No. 1, p. 012029). IOP Publishing.
- [13] Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
- [14] Misra, Diganta. "Mish: A self regularized non-monotonic activation function." *arXiv preprint arXiv:1908.08681* (2019).

Towards a Conceptual Architecture for Enabling Self-Adaptive IoT-Based Systems: A Service-Based Approach

Merabet Fatima Zohra

Department of Software Technologies and
Information Systems.
University of Constantine 2 Abdelhamid
Mehri

Constantine, Algeria

fatima.merabet@univ-constantine2.dz

Benmerzoug Djamel

Department of Software Technologies and
Information Systems.
University of Constantine 2 Abdelhamid
Mehri

Constantine, Algeria

djamel.benmerzoug@univ-constantine2.dz

Abstract— In a dynamic and ever-evolving environment like the Internet of Things (IoT), the original services may become unavailable or be interrupted by their providers. Thus, new services with different quality of service (QoS) values may emerge. Furthermore, IoT devices are generally limited in resources, and QoS values are not consistent and vary depending on the time and location of the call. So, there is a high probability that the service's QoS will fail or deteriorate during execution. That means this service must be replaced by another functionally equivalent one. In our research, we advocate the use of the services-oriented systems paradigm for building next-generation services using IoT, and the main objective of our work is to ensure the self-adaptation of these complex systems. We stress thinking of adaptation at the design phase, starting from the first step in the BPM (Business Process Management) life cycle, by proposing a BPMN extension named BPMN4SAS (BPMN for the self-adaptive system). The proposed extension is illustrated through an example dealing with accident management in a smart city system. The second contribution of this paper concerns the optimization phase, where we propose a global architecture to ensure the runtime self-adaptation of IoT services. The proposed architecture offers optimal services to the runtime system, which uses them to replace the currently functioning services in case of failure or degradation of their Quality-of-Service (QoS) values.

Keywords— *Services-Oriented Systems (SoS), Internet of Things (IoT), Quality of Services (QoS), Adaptation, Self-adaptive system, BPMN extension*

I. INTRODUCTION

The number of connected objects is growing to the extent of a billion objects in the near future [6]. According to statistics³, showing the growth of connected objects from 2015 to 2025, approximately 3 billion new objects are connected each year around the world. Indeed, each of these objects offers one or more services through sensors (temperature, motion, pollution, etc.), actuators (remotely controllable devices acting on the environment), and other entities. They can be considered new services in a service-oriented computing vision [2].

According to several studies [23], [24], the rise of connected objects will result in a multiplicity of services that

can provide data from multiple sources, such as surveillance cameras, sensors, and actuators. This will lead to new applications in many areas, such as home applications, industrial automation, medical aids, and traffic management. These new application services are complex spaces characterized by the openness, heterogeneity, uncertainty, and dynamics of the entities that constitute them [22]. Thus, these characteristics raise challenges such as uncertainty management, event responsiveness, context sensitivity, and self-adaptation to random changes that may take place in an IoT environment.

Due to this dynamism and all these environmental changes, the system's functioning will become more complex, and the quality of service will be uncertain. One promising way to reduce these problems is self-adaptation. Self-adaptation has emerged as a method for enabling software systems capable of responding autonomously to environmental changes to maintain the required QoS [17]. That means the system can detect changes (ex: QoS degradation in executing services) and then act correctly (replace the failed services with corresponding ones from the candidate services) while satisfying user goals and QoS constraints. The question is how the IoT system can respond to environmental changes and user goals during its execution while maintaining correct behavior and overall functioning. A self-adaptive system can modify itself automatically by adjusting attributes or artifacts in response to changes in the system or its operating environment [14] [9]. In recent years, self-adaptive systems have seen increasing interest in different research areas such as Pervasive Computing, Autonomic Computing [9], and Nature-Inspired (Organic) Computing [11]. This paper makes the following contributions according to the BPMN life cycle, as shown in Fig 1:

Fig 1 depicts a global overview of the proposed approach according to the BPM life cycle. This approach is based on different models, including the BPMN extension (BPMN4SAS) and self-adaptive system architecture.

- We propose a BPMN4SAS extension in the early modeling phases to simplify adaptation during system

execution, whose principal goal is to collect the maximum amount of information about the system.

- We propose a self-adaptive architecture in the optimization phase, which contains several interconnected components working together to achieve specific goals when detecting events in an IoT environment, whose principal objective is to ensure self-adaptation of the system during its execution, taking into account the user context on the one hand and the variable preferences of users on the other.

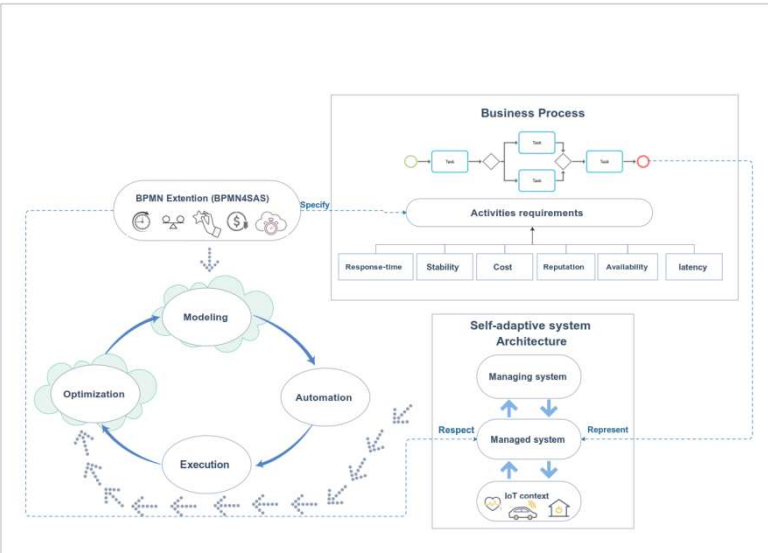


Fig. 1 Overview of the proposed approach according to the BPM life cycle

The remainder of this paper is structured as follows. Section 2 discusses our first contribution, which is an extension of BPMN for a self-adaptive system. In section 3, a global architecture that represents the second contribution will be presented. Finally, Section 4 concludes this paper.

II. RELATED WORK

In this section, we present the main related works, including service-oriented systems and self-adaptive systems, to show the originality of our research.

Adaptation is an important requirement for many software systems, such as cloud-based, cyber-physical [1], and mobile systems. These systems are designed to satisfy different user needs and expectations, especially the need to take advantage of service permanently. Such a system should be able to adapt its structure and/or behavior at runtime to cope with changes in its operating environment goals and user needs [8]. In the context of self-adaptation, several works are proposed to solve the problem of environmental change and user needs. Several works have been proposed that deal with self-adaptation in the design phase [3], [15], [16], and [7]. Their main objective was to provide designers with the right tools they needed to model all relevant concerns in the context [10]. Some authors introduce the notion of BPMN and the MAPE-K (Monitor, Analyze, Plan, Execute, Knowledge) approach, such as in [15] and [16], when the authors introduce an Adaptation Engine (AE) that supports self-adaptation of an approach from autonomic computing for self-adaptation. More specifically, they describe how autonomous managers responsible for self-adaptation of process instances implement the MAPE control

loop. First, in [15], they defined the architecture of the AE. They have discussed the interaction between the AE and the operating environment through sensors and between the AE and the process engine through connectors. Second, they have presented the internal architecture of the autonomic process. While it was presented in [16] that the MAPE control loop aims to identify the adaptation needs for running business process instances (i) adaptation by deviation, (ii) adaptation by variability, and (iii) adaptation by evolution. MAPE-K defines and executes the operations required to deal with these needs, while the K is the knowledge needed for the MAPE control loop. However, their contributions are incomplete as the recommended solution has only dealt with the architecture of the AE, notably the internal architecture of the autonomic managers. In [7], the authors propose a hybrid approach for the development of self-adaptive systems that uses both the local knowledge base in the vehicle and the global knowledge base provided via a Web service. The global knowledge base is shared and evolved by multiple vehicles through the Web service. While in [3], the authors present a conceptual SaSs modeling approach consisting of a meta-model and a modeling process. The process defines how the met model can be instantiated from requirements specifications to create conceptual SaSs models. They performed a controlled experiment with subjects to evaluate the effectiveness of their modeling approach. Also in [10], the authors present a new architecture that uses context-aware management to manage the auto-adapted system and uses the intermediate context model to redefine useful data exported from the BPMN tool at design time. Its goal is to provide a solution for using specific views at design time to simplify the expert's task. Moreover, it also opens the possibility of using many independent views to separate concerns and improve responsiveness in a dynamic context-aware system. Of all the previous works, some authors are based only on the self-adaptation technique like in [15], [16], [7], [3], and [10]. However, other works are based only on the extensions in their research, like in [4], in which the authors proposed an extension to the BPMN language standard that allows specifying the requirements of business process (BP) activities in terms of security, compliance, cost, performance, and storage. BP activity requirements help select the most appropriate cloud/fog service provider to outsource their extension. However, they were not discussing the adaptation mechanism, although this is especially necessary for a dynamic environment like the IoT, where there are many changes to which the system must respond in a self-adaptive manner.

Process self-adaptation has been the focus of several contributions, such as [1], [13], [12], [5], [19], [20], [25], or [26]. Some of these papers recommend the use of the MAPE-K approach for monitoring activities (sub-processes or tasks) or processes. For example, [1] recommended a solution for managing process variants at design time and self-adjusting them at run-time. At design time, the solution makes it possible to model process variability by describing process variants using three models: the base model for specifying process fragments (i.e., consistent parts of processes) shared by all process variants; the variation model for specifying the replacement fragments that alternatively exist for the fragments of the base model; and the resolution model for specifying the context conditions that define the usage conditions for the replacement fragments. At run-time, the recommended solution monitors fragments and

dynamically adapts the base model according to the variation model. Thus, it compares the context conditions of the monitored fragments with the context events recorded in the context model and finally defines an adaptation plan containing actions for fragment replacement.

[18] decomposes the mechanism of adaptation into several processes: monitoring software entities (self-awareness) and the environment (context-awareness), analyzing significant changes, planning how to react, and executing so that the decisions take effect. In most existing solutions, the adaptation processes are assigned to an external adaptation manager that is separate from the application logic. An adaptation manager implements the four processes mentioned above to control the behavior of adaptive software.

Other than that, to make an efficient self-adaptation, it is necessary to consider criteria defined at design time in the business process activities during its execution, contrary to what the authors have done in previous works [15], [16], [7], [3], and [10]. The adaptation is guaranteed at the time of execution-only if we reflect on it from the conceptual level. Thus, in this paper, we have proposed an approach with two complementary facades. The first one represents a BPMN extension named BPMN4SAS that includes some criteria such as availability, reliability, cost, reputation, latency, stability, and response time. BPMN4SAS is a preparation for the adaptation to simplify it at the executing time. The second part represents the global architecture of a self-adaptive system based on MAPE-K and a service recommendation component. These components aid the executed system to be self-adaptive by monitoring the failed services and replacing them with the optimally selected ones.

III. BPMN4SAS: A BPMN EXTENSION FOR SELF-ADAPTIVE SYSTEM

Our BPMN extension for self-adaptive systems has been demonstrated through an illustrative example of a business process, followed by a set of BPM tools that allow its implementation.

A. BPMN extension principle

Modeling different criteria of such an IoT system is not a simple task because of several factors that will influence the overall behavior of these systems. Our BPMN4SAS extension is based on some criteria (as defined in the following section). BPMN4SAS explains different criteria that must be taken into account during process execution. This extension was introduced to provide a standard visualization mechanism for designing, controlling, monitoring, and fluidly communicating ubiquitous business processes. The extension mechanism offered by BPMN ensures the validity of the basic elements as well as the integrity of the concepts, adding to the specificity of the domain. This is an abstract structural representation used to define the different components of the BPMN 2.0 language [21].

B. QoS criteria used in BPMN4SAS extension

In our BPMN4SAS extension, a business process is seen as a set of activities organized over time to produce a specific result. Our extension allows each process activity to represent the following criteria: availability, latency, response time, cost, reliability, reputation, and stability.

Criteria are defined for all activities, and their values change according to the needs and priorities of each activity.

C. A Meta-Model of our extension

Based on the meta-model presented in Fig 2, we have proposed a BPMN extension for self-adaptive systems. The goal is to collect as much information as possible on the process to simplify its adaptation during execution. We have added new concepts (previously specified criteria) related to our context. The new BPMN extension must take into consideration the rules to make the complete design as [21]:

- The basic flow elements (activity, event, and gateway) must not be modified.
- It is not possible to add new flow elements.

D. The new graphic elements introduced in BPMN4SAS extension

The BPMN specification model provides insufficient graphical notations for the representation of extensions. However, new graphic forms can be added as artifacts¹. Therefore, a set of graphical elements is proposed to enable the annotation of BP activities. Furthermore, the BPMN specification 1 requires that elements added retain their basic look and feel so that any viewer can understand them. Accordingly, we completed the model by adding new graphical elements into the used modeling tool. Fig 3 depicts the meta-model of our proposed extension after incorporating the new reliability, availability, reputation, cost, response time, latency, and stability criteria.

Figure 4 presents the new graphical representations of our extensions over the activity element and their corresponding descriptions.






QoS attributes	Description	Graphic element
Cost	This element is used to indicate that an activity is very costly.	
Reputation	This element represents the fidelity of a service. It mainly depends on the experience of the users of this service.	
Stability	This element is used to indicate that the activity must have the same values several times under the same conditions.	
Latency	This element represents the time between sending and receiving the query to the cloud and fog by the activity.	
Execution time	This element represents the required execution time specified by the activity.	

Fig. 4 The graphical elements of the proposed BPMN extension

IV. A GLOBAL ARCHITECTURE FOR DYNAMIC ADAPTATION OF SERVICES BASED IOT APPLICATION

The design of service compositions is not sufficient. So, it is essential to supervise the operation of services throughout their execution and to act in the case of QoS parameter violations or malfunctions to ensure the proper operation of applications and their QoS values. In this

¹ Research Google from the website Business Process Model and Notation (BPMN) Version 2.0.2. <https://www.omg.org/spec/BPMN/2.0.2/>

section, we will present a generic design of the smart environment architecture. We intend to provide a quality architecture that satisfies the requirements mentioned earlier. Therefore, our architecture was implemented to develop a self-adaptive system to manage autonomously the dynamics associated with IoT systems with a service-oriented vision. The whole architecture may be split into three layers (context, software system, and adaptation), with each layer being a set of different components that interact to achieve a specific goal.

Layer 1 (Context): The main goal of this layer is to capture and collect external information about the environment and transfer it to the next layer.

Layer 2 (Software System): is a suitable set of activities represented with BPM. It represents our managed system with specified constraints to be respected during its execution.

Layer 3 (Adaptation Layer): It aims to adapt services according to the context changes detected by the preceding layer. It is responsible for analyzing the context and deciding upon actions to execute. This latter contains several interconnected modules that work together in a collaborative process to achieve goals in detecting events in an IoT environment. These layers exhibit strong cohesion and weak coupling, which are suitable for the development of smart system applications.

A. Overview of the proposed architecture

The general architecture supports the self-adaptation of running micro-services. It gives us a global view of the organization, the roles, and the functions of the different modules. This section defines these modules and describes the interactions between them.

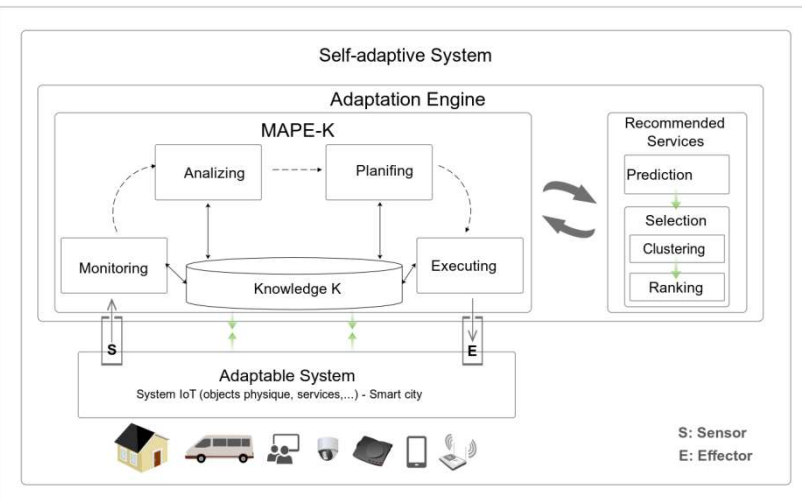


Fig. 5 a global architecture proposed for self-adaptive service-oriented systems

Fig 5 shows the architecture of our self-adaptive system. The idea is to be able to replace failed services with the most appropriate proposal. In our case, the system is based on a service-oriented vision to facilitate interaction with the supervised system, either through the use of platforms (middleware, cloud platforms, etc.) or directly with the concerned objects. In the presented system, the monitoring interacts with both the knowledge base and the IoT system.

As shown in Fig 5, our self-adaptive system is conceptually divided into two main modules: the adaptable

system (representing the managed system) and the adaptation engine (representing the system's manager).

B. Adaptation engine

Is the kernel element in our architecture. It represents the high-level entity responsible for monitoring and managing the underlying parts. It represents our self-adaptive system in conjunction with the adaptable system (a managed system). The communication between those subsystems in a self-adaptive system is performed via well-defined sensor and effector interfaces provided by the managed subsystem.

- Sensors (or probes) are software or hardware components responsible for collecting information about the current state of the managed subsystem or the environment where it operates.
- Effectors (or actuators) are also software or hardware components that provide a way to operate on the managed subsystem and enable self-adaptation.

Usually, the adaptation engine model is based on the MAPE-K architecture. In this work, and according to its responsibilities and definitions, we summarized this model into two essential processes: the MAPE-K control loop introduced in [9] and the recommended services model that we have proposed. These two models are explained in the rest of this section.

a. MAPE-K control loop

This layer consists of a MAPE-K loop that solves the system's unavailability by reconfiguring (i.e., adapting) the weather service. The MAPE-K feedback loop adapts the system architecture autonomously to fulfill the system's requirements. It monitors the sensors, analyzes the monitored data, plans actions, and executes these actions via the effectors of the managed system. The managed system is monitored and adapted via sensors (S) and effectors (E).

- Monitor: The monitor is in charge of knowing what is happening in the supervised system. It will be able to collect events that occur in the supervised system and generate symptoms according to them with the help of sensors. (i. e., if one of the currently working services is about to fail or its quality degrades, especially in critical domains such as healthcare and emergency response).

- Analyzer: The analyzer is in charge of making decisions about what actions to take. The analyzer will have the ability to analyze the symptoms and generate requests for change if necessary. Send these requests to the recommended services to get optimum ones.

- Planner: it is responsible for implementing the changes requested by the analyzer. Thus, the planner generates execution plans from the RFCs to be executed on the monitored system using the recommended services.

- Executor: is responsible for execution. Thus, this component executes plans provided by the scheduler on the monitored system. It uses the recommended services to replace one with another at run-time when a service fails or a new, more

efficient one is available. The goal is to build a performing system and ensure its correct execution.

- Knowledge Base (KB): it collects information and models needed to process information and analyze symptoms, etc. It can include static information about devices, appliances, or users, such as configuration information, or dynamic information exposed to frequent changes, such as the location of the users or devices in the environment.

Several components involved in the development and instantiation of this architecture will be detailed below.

b. Recommender System

In our context, the use of recommended services is an axis of the solution because it brings elements of response to the problems that can occur in the monitored system. It is composed of two steps.

Prediction: In a real scenario, the number of services invoked by users is quite limited, which leads to missing QoS values for these services. So, we need to predict the missing QoS values based on the historical QoS records of services invoked by users to facilitate their selection and recommendation (see our work in [28]).

Selection: users aim to select the best services, adequate to their needs, to use them in their systems. However, choosing an inappropriate service may cause many problems. (e.g., suit the performance of the process resulting). So, a performed selection of services is based on two complementary steps:

- Clustering: In this step, we use k-means, the well-known algorithm for clustering, to group services with the same functionality into the same clusters after specifying the ideal number of k clusters.

- Ranking: Based on the results from the previous step, we rank services in the same clusters from the best to the worst (services with a low score of QoS values are considered the best) to recommend them to the executing model in MAPE-K. This latter kind will use them in the supervised system.

V. CASE STUDY

To demonstrate the benefits of our approach, we will apply it to accident management in a smart city [27]. This example has been modeled as a business process (BP) using the BPMN standard, with various activities to be performed with different QoS constraints defined in our proposed BPMN4SAS extension. There are thirteen (13) activities in this business process that are managed by an accounting software service. Each activity may have one or more candidate services. The various candidate services are as follows [27].

A1: collect photos of the accident. Candidate services for Activity 1 are: • S11: Cameras around the area • S12: cameras deployed on vehicles in the same area as the distressed vehicle • S13: cameras of different users • S14: drone with a camera A2: analyze recovered images. The candidate services in Activity 2 are: • S21: software deployed on Cloud Computing. • S22: software deployed on

the Fog Computing. A3: transport injured persons. Candidate services for Activity 3 are: • S31: ambulance service 1. • S32: ambulance service 2. • S33: vehicles. • S34: an emergency drone that can transport the sick. A4: Conduct a pre-diagnosis and send it to the medical facility. Candidate services for Activity 4 are: • S41: a mobile application. • S42: a connected object. A5: care for the injured. Candidate services for Activity 5 are: • S51: the hospital 1. • S52: the hospital 2. • S53: the clinic. A6: make a medical report and send it to the police automatically. Candidate services for Activity 6 are: • S61: software deployed on Cloud Computing. • S62: software deployed on Fog Computing. A7: extinguish the fire. Candidate services for Activity 7 are: • S71: fire service. • S72: firefighter robots. • S73: firefighter drone. A8: Involve law enforcement and investigate. Candidate departments for Activity 8 are: • S81: police center 1. • S82: police center 2. A9: Collected information from people in the accident area and on all drivers. Candidate services for Activity 9 are: • S91: mobile application 1. • S92: mobile application 2. A10: Collect information regarding road conditions. Candidate services for Activity 10 are: • S101: connected object 1. • S102: connected object 2. A11: Collect weather information. Candidate services for Activity 11 are: • S111: software deployed on Cloud Computing. • S112: software deployed on Fog Computing. A12: Make an automatic report of the accident. Candidate services for Activity 12 are: • S121: software deployed on Cloud Computing. • S122: software deployed on Fog Computing. A13: remove wrecked automobiles. Candidate services for activity 13 are: • S131: actuator 1. • S132: actuator 2.

A. Constraints

To obtain services that fully satisfy the conditions imposed by the activities in the business process, users make service requests, and these services requests must meet the following constraints: • Availability: true for all services • Stability: at least 60 • Latency: maximum 1 hour • Response time: maximum 3 hours • Reputation: at least 3 for all services • Cost: 800 euros maximum

B. Application of the BPMN4SAS extension

We modeled this example using the BPMN standard by specifying different activities that need to be performed with different QoS constraints considered in our proposed BPMN4SAS extension. Fig. 6 and 7 show the proposed Accident Management Services Systems modeled using the BPMN standard and our BPMN4SAS extension, respectively.

C. Normalization

The normalization follows two main categories of quality of service:

Category 1: higher non-functional attribute values produce higher satisfaction (the higher, the better), e.g., reputation and stability non-functional attributes. This type of category represents attributes to be maximized:

$$Normalize(q_i) = \frac{\max(q_i) - (q_i)}{\max(q_i) - \min(q_i)} \quad (1)$$

Category 2: lower non-functional attribute values produce higher satisfaction (the lower, the better), e.g., response time, latency, and cost non-functional attributes. This type of category represents attributes to be minimized:

$$Normalize(q_i) = \frac{(q_i) - \min(q_i)}{\max(q_i) - \min(q_i)} \quad (2)$$

Where minimum and maximum represent the highest and lowest levels of satisfaction with the user-specified non-functional attribute, respectively.

After normalization and to calculate the score of the quality of service, we apply the following formula:

$$Score(S1q_i) = \sum_{i=1}^M Normalize(q_i) \cdot \lambda_i \quad (3)$$

Where:

- The score value in this example is between 0.2 and 4.53.
- M is the sum of the quality of services (in this case, m = 5).
- q_i represents the quality of service for a web service i.

$$\lambda_i = \begin{cases} 0, & \text{if } q_i = \begin{cases} \text{attribute specified in the activity extension} + \\ \text{value not specified constraint} \end{cases} \\ 1, & \text{else} \end{cases} \quad (4)$$

For activity A1, the score values for service 1 for T0 can be computed as follows:

Attributes to be maximized:

$$Normalize(Stability) = (70 - 0.12) / (70 - 0.12) = 69.88 / 69.88 = 1$$

$$Normalize(Reputation) = (4.1 - 2) / (4.1 - 1) = 2.1 / 3.1 = 0.68$$

Attributes to be minimized:

$$Normalize(Response\ time) = (22 - 22) / (140 - 22) = 0$$

$$Normalize(Latency) = (2 - 2) / (60 - 2) = 0$$

$$Normalize(Cost) = (50 - 17) / (63 - 17) = 33 / 46 = 0.72$$

$$Score(S1q_i) = Normalize(Stability) \cdot \lambda_i + Normalize(Reputation) \cdot \lambda_i + Normalize(Response\ time) \cdot \lambda_i + Normalize(Latency) \cdot \lambda_i + Normalize(Cost) \cdot \lambda_i$$

$$Score(s1q_i) = 1*0 + 0.68*0 + 0*1 + 0*1 + 0.72*1 \quad Score(s1q_i) = 0.72$$

Similarly, we compute the score values for the remaining times T1 to T3 for services 11 on activity A1. Again, the score values for services 12 to 14 can be computed, and based on these values the services will be ranked as shown in table 1b.

In the same way, the total QoS values of activities A2 to A13 can be calculated as shown in tables 2a to 13a

The web services were then ranked according to their QoS score values, leading to a number of ordered lists. The results displayed in Tables 1b through 13b indicate that the recommended services for selection depend on score values of Quality of Services. For instance, the top-2 recommended services for activity A1 in times t0, t1 and t2 are service 12, and service 13. Those for t3 are service 12, and service 14 respectively.

VI. CONCLUSION

In this paper, an approach has been proposed to make a self-adaptive system. This approach was carried out in a service-oriented system by following two essential aspects of

the business process life cycle: modeling and optimization. First, we have proposed a BPMN extension for self-adaptive systems: our extension allows us to collect as much data as possible about the process to facilitate its adaptation during execution. In the second step, we have presented a global architecture of service-oriented systems that can be designed to build a performed self-adaptive system. Our architecture can control and adapt composite services, taking into account the user context on the one hand and the variable preferences of users on the other. We evaluate it using an illustrative example in the smart city.

ACKNOWLEDGMENT

THE AUTHORS WOULD LIKE TO ACKNOWLEDGE THE ICASA'22-EL-TARF EDITOR AND REVIEWERS FOR THE THOUGHTFUL READING OF THIS MANUSCRIPT AND FOR THEIR RELEVANT COMMENTS THAT HELPED US IMPROVE THE QUALITY OF THE PAPER

REFERENCES

- [1] Clara Ayora, Victoria Torres, Vicente Pelechano, and Germán H Alferez. Applying cvl to business process variability management. In Proceedings of the VARIability for You Workshop: Variability Modeling Made Useful for Everyone, pages 26–31, 2012.
- [2] Sylvain Cherrier and Yacine M Ghamri-Doudane. The “object-as-a-service” paradigm. In 2014 Global Information Infrastructure and Networking Symposium (GIIS), pages 1–7. IEEE, 2014.
- [3] João Pablo S da Silva, Miguel Ecar, Marcelo S Pimenta, Fabio Natanael Kepler, Gilleanes TA Guedes, and Carlos Michel Betemps. Improving self-adaptive systems conceptual modeling. In Proceedings of the 33rd Annual ACM Symposium on Applied Computing, pages 1292–1299, 2018.
- [4] Louar Fadila, Zarour Karim, and Benmerzoug Djamel. Modelling business processes for outsourcing into the fog and cloud computing. Proceedings of the 8th International Symposium on Data-driven Process Discovery and Analysis (SIMPDA 2018), Seville, Spain, 2018.
- [5] Sheila Katherine Venero Ferro and Cecilia Mary Fischer Rubira. An architecture for dynamic self-adaptation in workflows. In Proceedings of the International Conference on Software Engineering Research and Practice (SERP), page 35. The Steering Committee of The World Congress in Computer Science, Computer . . . , 2015.
- [6] Guillaume Garzone. Approche de gestion orientée service pour l'Internet des objets (IoT) considérant la Qualité de Service (QoS). PhD thesis, Institut National des Sciences Appliquées de Toulouse (INSA de Toulouse), 2018.
- [7] Dhrgam AL Kafaf and Dae-Kyoo Kim. A web service-based approach for developing self-adaptive systems. Computers & Electrical Engineering, 63:260–276, 2017.
- [8] Slim Kallel, Ismail Bouassida Rodruigez, and Khalil Drira. Editorial adaptive and reconfigurable software systems and architectures, 2016.
- [9] Jeffrey O Kephart and David M Chess. The vision of autonomic computing. Computer, 36(1):41–50, 2003.
- [10] Stéphane Lavirotte, Gaëtan Rey, Nhan Le Thanh, Jean-Yves Tigli, et al. From bpmn to live application: How the context can drive an auto-adapted system. 2019.
- [11] Christian Müller-Schloer, Hartmut Schmeck, and Theo Ungerer. Organic computing—a paradigm shift for complex systems. Springer Science & Business Media, 2011.
- [12] Karolyne Oliveira, Jaelson Castro, Sergio España, and Oscar Pastor. Multi-level autonomic business process management. In Enterprise, Business-Process and Information Systems Modeling, pages 184–198. Springer, 2013.
- [13] Karolyne Oliveira, Jaelson Castro, Emanuel Santos, Robson Fidalgo, Sergio España, and Oscar Pastor. A multi level approach to autonomic business process. In 2012 26th Brazilian Symposium on Software Engineering, pages 91–100. IEEE, 2012.
- [14] Peyman Oreizy, Michael M Gorlick, Richard N Taylor, Dennis Heimbigner, Gregory Johnson, Nenad Medvidovic, Alex Quilici,

David S Rosenblum, and Alexander L Wolf. An architecture-based approach to self-adaptive software. IEEE Intelligent Systems and Their Applications, 14(3):54–62, 1999.

- [15] Jamila Oukharjane, Imen Ben Said, Mohamed Amine Chaabane, Eric Andonoff, and Rafik Bouaziz. A new adaptation engine for self-adaptation of bpmn processes instances. 2019.
- [16] Jamila Oukharjane, Imen Ben Said, Mohamed Amine Chaabane, Eric Andonoff, and Rafik Bouaziz. Towards a new adaptation engine for self-adaptation of bpmn processes instances. In Proceedings of the 14th International Conference on Evaluation of Novel Approaches to Software Engineering, pages 218–225. SCITEPRESS-Science and Technology Publications, Lda, 2019.
- [17] Mazeiar Salehie and Ladan Tahvildari. Self-adaptive software: Landscape and research challenges. ACM transactions on autonomous and adaptive systems (TAAS), 4(2):1–42, 2009.
- [18] Mazeiar Salehie and Ladan Tahvildari. Towards a goal-driven approach to action selection in self-adaptive software. Software: Practice and Experience, 42(2):211–233, 2012.
- [19] Ronny Seiger, Steffen Huber, Peter Heisig, and Uwe Assmann. Enabling self-adaptive workflows for cyber-physical systems. In Enterprise, Business-Process and Information Systems Modeling, pages 3–17. Springer, 2016.
- [20] Ronny Seiger, Steffen Huber, Peter Heisig, and Uwe Assmann. Toward a framework for self-adaptive workflows in cyber-physical
- [21] Stephen A White and Conrad Bock. BPMN 2.0 Handbook Second Edition: Methods, Concepts, Case Studies and Standards in Business Process Management Notation. Future Strategies Inc., 2011.
- [22] Ali Yachir. Composition dynamique de services sensibles au contexte dans les systèmes intelligents ambiants. PhD thesis, Université Paris-Est Créteil Val-de-Marne - Paris 12 - France, 2014.
- [23] Al-Fuqaha, Ala et al. (2015). “Internet of things: A survey on enabling technologies, protocols, and applications”. In: IEEE communications surveys & tutorials 17.4, pp. 2347–2376.
- [24] Bellavista, Paolo et al. (2013). “Convergence of MANET and WSN in IoT urban scenarios”. In: IEEE Sensors Journal 13.10, pp. 3558–3567.
- [25] Alfonso, Iván and Garcés, Kelly and Castro, Harold and Cabot, Jordi(2021). “Self-adaptive architectures in IoT systems: a systematic literature review”. In: Journal of Internet Services and Applications 12.1, pp. 1–28.
- [26] Saputri, Theresia Ratih Dewi and Lee, Seok-Won. (2020). “The application of machine learning in self-adaptive systems: A systematic literature review”. In: IEEE Access Journal 8, pp. 205948–205967.
- [27] Hadjab Selma, Benmerzoug Djamel, Guermouche Nawal, Guidara Ikbel. (2018). “Approche prédictive d’adaptation dynamique des systèmes orientés services : application à l’Internet des objets ”.
- [28] Merabet, Fatima Zohra and Djamel Benmerzoug (2021). “Qos prediction for service selection and recommendation with a deep latent features autoencoder”. In: Computer Science and Information Systems 00, pp. 54–54.

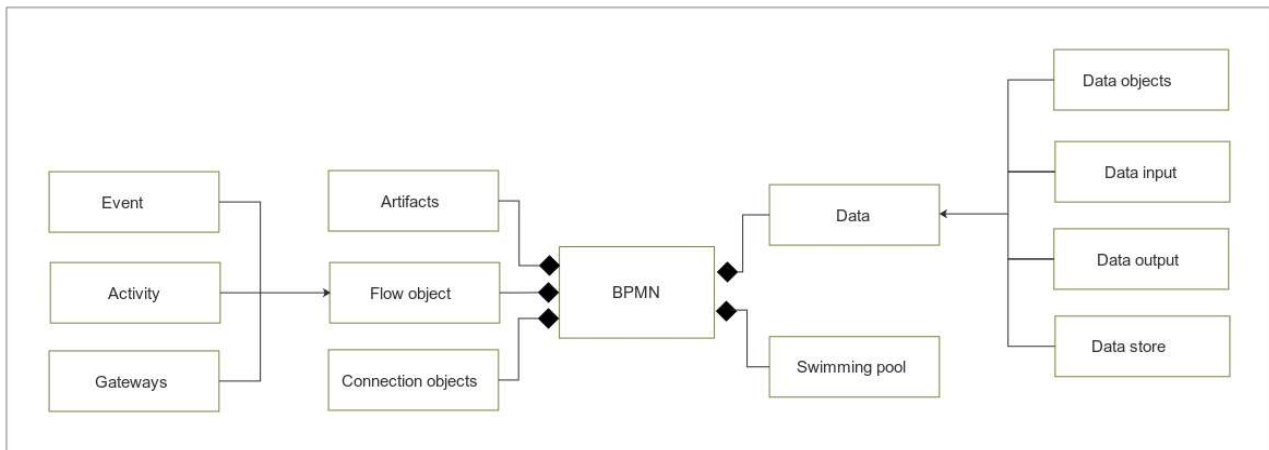


Fig. 2 Meta model for structuring BPMN language elements

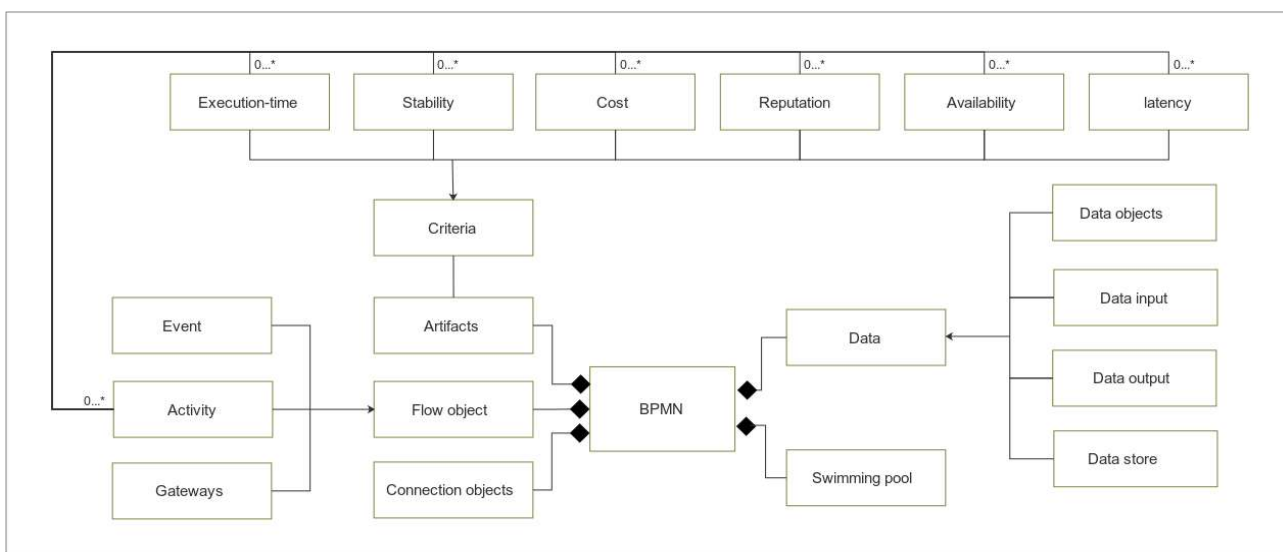


Fig. 3 Meta model for structuring BPMN4SAS extension elements

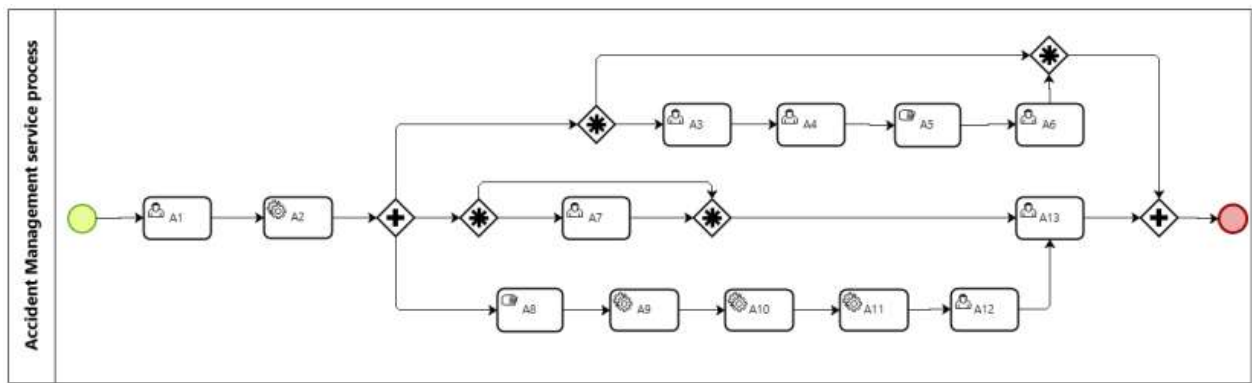


Fig. 6 Accident management service process modeling with BPMN

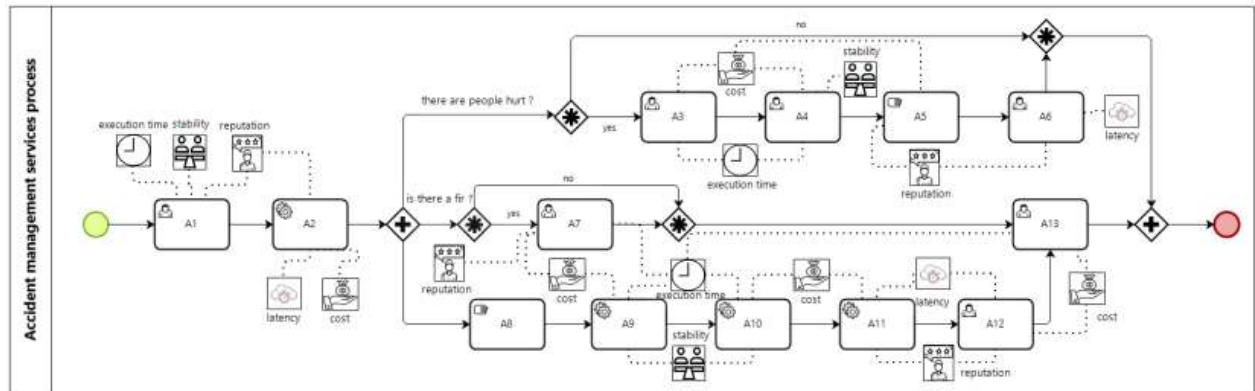


Fig. 7 Accident management service process modeling with BPMN4SAS extension

Table 1.a: This table shows the candidates services with their QoS parameters and their score value before ranking of activity A1

Activities	Candidates services	Availability	Response time	Stability	Latency	Cost	Reputation	Score
A1	S11	T0	22	0.12	2	50	2	0.72
		T1	30	0.55	10	50.5	2.2	0.94
		T2	24	0.51	11	40	3	1.03
		T3	26	0.40	8	51	1.8	0.87
	S12	T0	45	0.60	15	60	3.5	1.53
		T1	100	0.50	60	61	3	2.97
		T2	140	0.44	60	62	3.2	3.27
		T3	70	0.30	20	63	3	2.06
	S13	T0	56	0.50	3	31	3.25	0.88
		T1	120	0.25	2	30.5	3	1.47
		T2	100	0.35	4	20	3.5	1.22
		T3	40	0.25	2	19.2	4.1	0.2
	S14	T0	23	0.30	8	24	1	0.26
		T1	40	0.24	8	23	1.8 5	0.38
		T2	55	0.70	9	17	2	0.34
		T3	105	0.15	11	26	1.5	1.05

Table1. b: This table shows the possible scenario of ranking candidates services based on their scores values in t0, t1, t2, t3

	T0	T1	T2	T3
1	S12 (1.53)	S12 (2.97)	S12 (3.27)	S12 (2.06)
2	S13 (0.88)	S13 (1.47)	S13 (1.22)	S14 (1.05)
3	S11 (0.72)	S11 (0.94)	S11 (1.03)	S11 (0.87)
4	S14 (0.26)	S14 (0.38)	S14 (0.34)	S13 (0.2)

Tumor Detection in Mammography Images using Discrete Wavelet Transform and Evidence Fusion Technique

Abdelkader Zitouni
Department of Electronics
University Amar Telidji
Laghouat, Algeria
a.zitouni@lagh-univ.dz

Fatiha Benkouider
Department of Electronics
University Amar Telidji
Laghouat, Algeria
fbenkouider@gmail.com

Fatima Chouireb
Department of Electronics
University Amar Telidji
Laghouat, Algeria
f.chouireb@lagh-univ.dz

Mourad Reggab
Department of Electronics
University Amar Telidji
Laghouat, Algeria
m.reggab@lagh-univ.dz

Abstract: This research presents supervised classification algorithm based on information fusion for detecting masses in mammography images. Discrete wavelet transform can preserve information regarding both high and low frequencies and offer great discriminatory power between areas with strong similarities. This motivates us to use this type of features to improve image segmentation. So, in the first stage, the suggested technique used this feature extraction approach on mammography images in order to obtain additional information. After that in the second stage, estimated feature vector of each pixel is sent to a neural network classifier for initial-labeling. Then, in the third stage of the suggested technique, Evidence fusion method is used to combine the scores, within a sliding window, obtained by the neural network for each pixel. The performance of the proposed segmentation algorithm was evaluated on mammography images from Mammography Image Analysis Society (MIAS) dataset. The achieved classification results by the proposed fusion system leads to higher classification precision in detecting masses on mammography images, which are one of breast cancer signs.

Keywords: Image segmentation, Masses Detection, Breast Cancer, Neural Network, Wavelets, Evidence fusion.

I. INTRODUCTION

One of the leading causes of death worldwide among women is breast cancer [1]. Studies on breast cancer have demonstrated that early detection of these abnormalities plays a very important factor in cancer treatment and allows better recovery for most patients [2].

Medical imaging is a robust and reliable diagnostic method for the breast related diseases and it can be produced from various equipment in the medical field, such as Ultrasound (USG), MRI, CT-Scan / CAT-Scan, and Mammography [3].

Mammography is the major screening tool which is carried out for detection of breast cancer at early stage, and several images processing techniques have been used for mammograms interpretation in order to assist radiologists while detecting or identifying eventual abnormalities [4-18].

Generally, masses (space occupying lesions) and calcifications (tiny flecks of calcium, like grains of salt) are the two abnormalities present in the mammogram images. The pre-processing and feature extraction process is an important stage in identifying the presence of tumors. So, referring to the advantages and disadvantages of the methods and algorithms that have been developed by previous researchers, in this paper we attempt to use one type of feature extraction techniques with the purpose of detecting masses on mammography images. This feature is a structural feature obtained by using the wavelet transform coefficients. For the pre-processing process, which finds to improve image quality such as contrast enhancement to obtain a better image visualization [1], there are different types of filtering techniques. In this study the contrast of the mammogram images was regulated by histogram adjustment which improves the contrast of the output image by spreading out the intensity values.

So, in the first stage of our work, we have used the discrete wavelet transform as feature extraction strategy in order to get more information that enables the classifiers to discriminate between the different areas in the mammography image. In the second stage, an appropriate classification algorithm is applied using the set of extracted features obtained from the previous stage. The Backpropagation Artificial Neural Networks classifier is chosen among the most well-known classifiers, it was initiated in [19-20]. The estimated feature vector of each pixel is sent to the neural networks classifier for initial-labeling.

A sliding window, whose class is assigned to its central pixel, is used. However, this central pixel belongs to other window neighbors that may be classified into other classes. Consequently, in the third stage, in order to obtain a more precise segmentation result, Evidence fusion method is used for each pixel to combine the scores results of several windows that contain this central pixel. The proposed segmentation algorithm performance was verified on mammography images from MIAS dataset [21]. The obtained results lead to higher classification precision in detecting masses which are one of breast cancer signs.

The rest of this manuscript is divided in three sections. The first one describes the background theory of several techniques used in this paper. Then, section 2 is dedicated to give in details the suggested segmentation process and the reached performance of the proposed fusion technique. Finally, in section 3, we conclude and recommend possibilities for future work.

II. Feature Extraction Algorithm and Fusion Theory

A. The wavelet analysis

Since the work of Grossman and Morlet [22], the wavelets transform has appeared as a powerful tool to solve problems in different application. The wavelets transform decomposes the input signal into a series of wavelet functions $\Psi_{a,b}(t)$ derived from a mother function $\Psi(t)$ given by dilatation (factor a) and translation (factor b) operations. Fig.1 illustrates some examples of wavelets that are generally used in image processing [23].

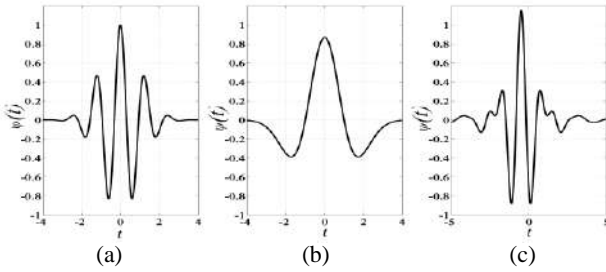


Fig.1 Examples of wavelets.

(a) MorletWavelet, (b) Mexican hat Wavelet, (c) Meyer Wavelet

Wavelet analysis transforms a finite energy signal in the spatial domain into another finite energy signal in the spatio-frequency domain. The components $C_{a,b}$ of this novel signal, which are described in equation (1), are called wavelet coefficients. In an image, these coefficients offer information on the local variation of the grey levels around a given pixel. The more significant is this variation, the higher they are [24].

$$C_{a,b} = \int_{-\infty}^{+\infty} x(t)\Psi_{a,b}(t)dt, \quad (1)$$

where

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right) \quad \text{with } a \neq 0 \quad (2)$$

The most important advantage of wavelets as compared to other frequency methods, as Fourier transform, is that it offers both frequency and spatial locality [25]. In 1989, Mallat [26] suggested a multi-resolution decomposition algorithm based on wavelets transform. The algorithm decomposes an input image into a set of detail images and an approximation image using a filter bank comprising a high pass filter (HP) and a low pass filter (LP). At each decomposition level the size of the transformed images is reduced by a factor of two [24]. The discrete wavelet transform of a 2-D image can be obtained by performing the filtering consecutively along horizontal and vertical directions (separable filter bank) [27]. Four images are then created at each level. Fig. 2 shows an example of decomposition of the image on one level.

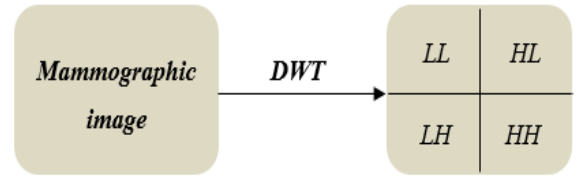


Fig. 2 Use of 1-stage discrete wavelet transform (DWT)

In Fig. 2, the DWT decomposes the image into 4 orthogonal sub-band: low-low (LL), high low (HL), low high (LH), and high-high (HH) consisting of approximation, horizontal, vertical, and diagonal information. The approximations image is the smoothed version of the original image and it contains global information that is similar to the original image with the number of rows and the number of columns being half of the original image. Horizontal, vertical, and diagonal contain the detail and represent the fluctuations of the pixel intensity in horizontal, vertical, and diagonal directions and they have low-intensity areas, whereas areas with high intensity are only found on the edges of the image object.

The values of transformed coefficients in detail and approximation images (sub-band images) represent the necessary features that capture useful discrimination information for masses segmentation [28].

1) *Wavelet's Choice*: In our research, we have used a second order biorthogonal spline wavelet. This wavelet is used for masses analysis due to its excellent location in the frequency and spatial domains and its sensitivity to local singularity and correlation of the image [24].

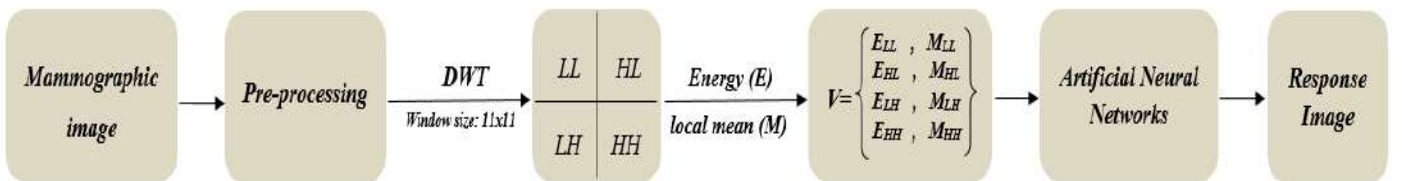


Fig. 3 Wavelet features extraction stage

2) *Indices' Calculation:* One of the most used indices for characterizing the masses in the spatio-frequency plane is the measurement of energy. Because the transformed images have different frequencies, scales and orientations, the energy index is a local measure of the wavelet coefficient distribution according to the scale, the orientation and the frequency. It has been used successfully for segmentation and classification of masses [24].

The expression of energy is given by [24]:

$$E = \frac{1}{N} \sum_R C(i, j)^2 \quad (3)$$

The second index used in conjunction with energy is the measurement of the local mean of the wavelet coefficients given by [24]:

$$M = \frac{1}{N} \sum_R |C(i, j)|, \quad (4)$$

where N denotes the number of pixels, designated by the indices (i, j) , and enclosed in the area R .

The calculation of these indices was done on a sliding window W . The local mean and the energy on the sliding window are calculated from the resultant sub-band images. So the feature vector of each window is made of eight parameters $V = [E_{LL}, E_{LH}, E_{HL}, E_{HH}, M_{LL}, M_{LH}, M_{HL}, M_{HH}]$, as seen in Fig. 3.

Several tests were carried out on a series of window sizes going from 5×5 to 25×25 . The highest good classification rate was reached for a window of dimension 11×11 . The produced features vector of each window is used as an input to the neural network classifier for a primary labeling, and the score for the window delivered by the neural network is assigned to its mid pixel.

B. Evidence Fusion Theory

Evidence fusion is one of the best methods used to make the combination of images with decision-making. Evidence theory, also known as Dempster-Shafer theory or theory of belief functions, was proposed by Dempster in 1967 [29] and then mathematically formulated by Shafer [30]. This theory was taken up in 1990 by Smets [31] under the name of transferable belief model.

1) Math Basics:

a) *Discernment framework:* A discernment framework is the all possible hypotheses set of the studied problem (that is to say the set of classes of the fused image). Let Ω denote this set and suppose that it is composed of N different hypotheses.

$$\Omega = \{H_1, H_2, \dots, H_N\} \quad (5)$$

Only one hypothesis of this set is considered true. The power set of the discernment set is the collection which contains all possible combinations of hypotheses. The power set is defined as follows [32]:

$$2^\Omega = \{A/A \subseteq \Omega\} \\ = \{\emptyset, \{H_1\}, \{H_2\}, \dots, \{H_1, H_2\}, \{H_1, H_3\}, \dots, \{H_1, \dots, H_N\}\} \quad (6)$$

"A" may be a simple hypothesis or a class of hypotheses or even an empty set.

b) *Mass function:* Generally, a mass function, also known as a "belief mass", is simply noted m . This mass is calculated for any element of the 2^Ω set and is defined as follows:

$$\forall A \in 2^\Omega, m(A) \in [0, 1],$$

$$\sum_{A \in 2^\Omega} m(A) = 1 \quad (7)$$

Information that is used to model masses is often a priori information, and observations on hypotheses [32].

c) *Mass transformations:* From the mass function m we can deduce other functions such as credibility and plausibility to better represent the knowledge:

-*Credibility:* The credibility (or belief) denoted by bel measures the total belief that can be attributed to a given element. It is defined by [32]:

$$\forall A \in 2^\Omega, \quad bel(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B) \quad (8)$$

-*Plausibility:* The plausibility denoted by Pl measures the maximum belief that can potentially be attributed to a given element. It is defined by [32]:

$$\forall A \in 2^\Omega, \quad Pl(A) = \sum_{A \cap B \neq \emptyset} m(B) \quad (9)$$

2) *Combination:* We assume to have N different sources noted: S_1, S_2, \dots, S_N . Masses which come from each source, attributed to an element A of the discernment set, are noted respectively: $m_1(A), m_2(A), \dots, m_N(A)$.

a) *Orthogonal law:* The orthogonal law is the normalization of conjunctive law [32]. Let there be N sources of information, with A an element of the power set:

$$\forall A \in 2^\Omega, \\ m(A) = \frac{1}{1 - k} \sum_{B_1 \cap B_2 \cap \dots \cap B_N = A} \left(\prod_{j=1}^N m_j(B_j) \right) \quad (10)$$

with:

$$\forall A \in 2^\Omega, \quad k = \sum_{B_1 \cap B_2 \cap \dots \cap B_N = \emptyset} \left(\prod_{j=1}^N m_j(B_j) \right) \quad (11)$$

k is a law normalization term also known as a conflict term. This law has the properties of commutativity and associativity.

3) *Decision:* The final fusion step is the decision. This is the step where only one hypothesis is chosen among the discernment set hypotheses after the fusion of knowledge brought by different sources. The evidence theory offers a set of different decision rules which do not necessarily give the same result:

a) *Maximum of credibility*: Credibility (bel) is interpreted as the lower bound of probability. Decision based on maximum of credibility is therefore a pessimistic decision, which is expressed as follows [32]:

$$Dec = \arg \max_{H \in \Omega} bel(H), \quad H \in \Omega \quad (12)$$

bel is the credibility function associated to the mass of the hypothesis $H \in \Omega$ corresponding.

b) *Maximum of plausibility*: Plausibility (Pl) is interpreted as the upper bound of probability. The decision of plausibility's maximum is then an optimistic decision:

$$Dec = \arg \max_{H \in \Omega} Pl(H), \quad H \in \Omega \quad (13)$$

Pl denotes the plausibility function associated to the mass of the hypothesis H [32].

III. The Proposed Segmentation Algorithm

Data obtained from the mammographic images is often, noisy, incomplete, inconsistent, and have low contrast. Therefore, pre-processing is needed in the medical image processing to improve image quality, remove unwanted noise, preserves the edges within an image, and make the feature extraction phase more reliable [1].

There are different types of filtering techniques in the pre-processing. So, in the first step of our work, the contrast of the mammogram images was regulated by histogram adjustment which increases and improves the contrast of the output image by spreading out the intensity values.

After that in the second step, the proposed segmentation method uses Wavelets transform as feature extraction strategy on mammography images in order to get more information in this data set. The parameters of the feature set were selected as mentioned in the previous sections. After a proper features' extraction, each estimated feature vector of each pixel is sent to the neural networks classifier for primary labeling.

Using a sliding window, the class for this window is assigned to its central pixel. However, this central pixel belongs to other window neighbors that may be classified into other classes. Consequently, in order to achieve a more precise segmentation result, for each pixel Evidence fusion method is used to combine the scores results of several windows that contain this central pixel:

Consider I to be the segmented image comprising the scores S_{ij} of each pixel (the output of the neural networks classifier):

$$I = S_{ij} \text{ with } i = 1 \dots n, j = 1 \dots m,$$

where n and m represent the sizes of the mammography image.

We perused the images by using a sliding window of size $M \times M$, so that every pixel is surrounded by $M^2 - 1$ pixels. Each central pixel $P_{ij,l}$ of window W_l with score $S_{ij,l}$ belongs to the $M^2 - 1$ window in the surrounding windows before the classification process. However, each central pixel $P_{ij,z}$ of the window z , with $z = 1 \dots M^2 - 1$ produced different scores $S_{ij,z}$.

For example, in the case of pixel $P_{3,3}$ with score $S_{3,3}$ and $M = 3$, the central pixel is surrounded by eight pixels, that are the center of the eight windows which pixel $P_{3,3}$ belonged to, (see Fig. 4).

From the above example, we joined the scores produced by the current block and its eight neighboring ones for the wavelet features: $\{S_{33}, S_{32}, S_{34}, S_{23}, S_{24}, S_{22}, S_{42}, S_{44}, S_{43}, S_{33}, S_{32}, S_{34}, S_{23}, S_{24}, S_{22}, S_{42}, S_{44}, S_{43}\}$ (see Fig. 5). In this work, a sliding window of size 9×9 is used, so the central pixel is surrounded by 80 pixels, that are the center of the 80 windows which pixel $P_{5,5}$ belonged to.

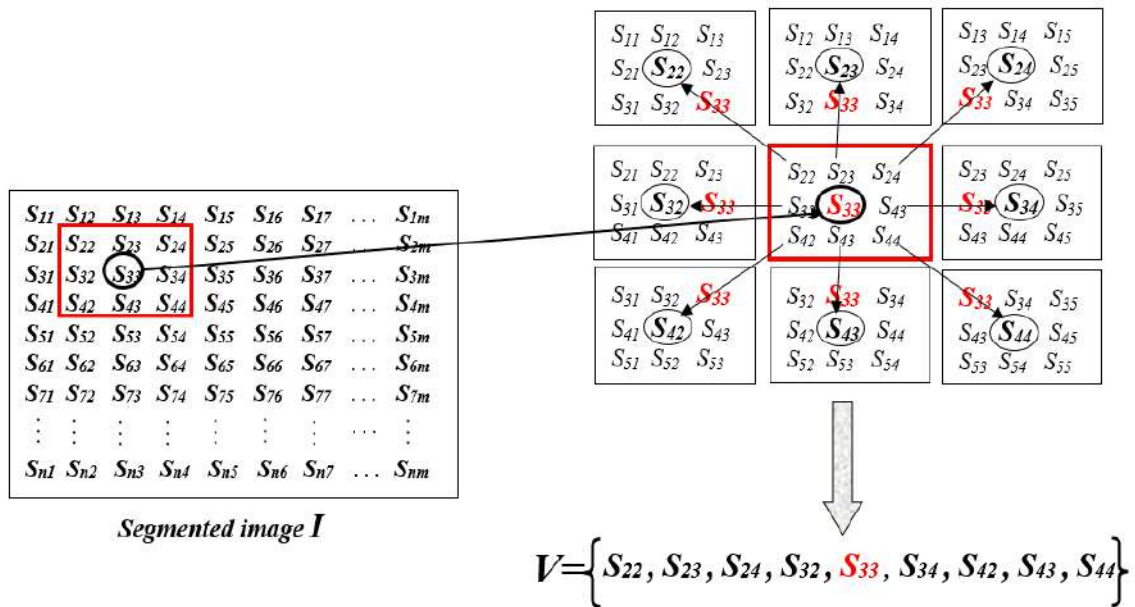


Fig. 4 Composition of the fusion vector

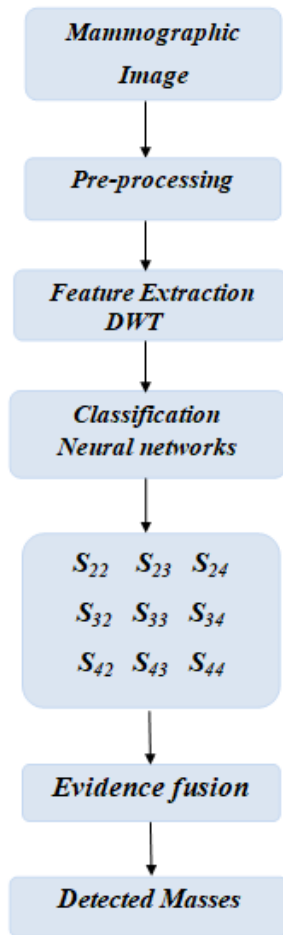


Fig. 5 Classification stages block diagram

The essential fusion algorithm steps are summarized as follows:

Step1:	Pre-processing of the mammography image.
Step2:	Feature extraction via DWT.
Step3:	Neuronal classification of the estimated feature vector. We obtain:
Step4:	<p><i>Scores1</i>: for DWT.</p> <p>while the recognition rate is changed do</p> <p> for each pixel do</p> <p> Evidence Fusion of Score1 and the neighboring scores.</p> <p> end</p> <p> Calculate the recognition rate.</p> <p>end</p>

The performance of the proposed algorithm for segmenting mammography images is assessed using many images from the MIAS (Mammographic Image Analysis Society) database [21] containing 322 mammograms sized 1024 x 1024 pixels. The images are arranged in pairs: those with even-numbers correspond to left MLO (medio-lateral oblique Mammograms) and those with odd-numbers are right MLO.

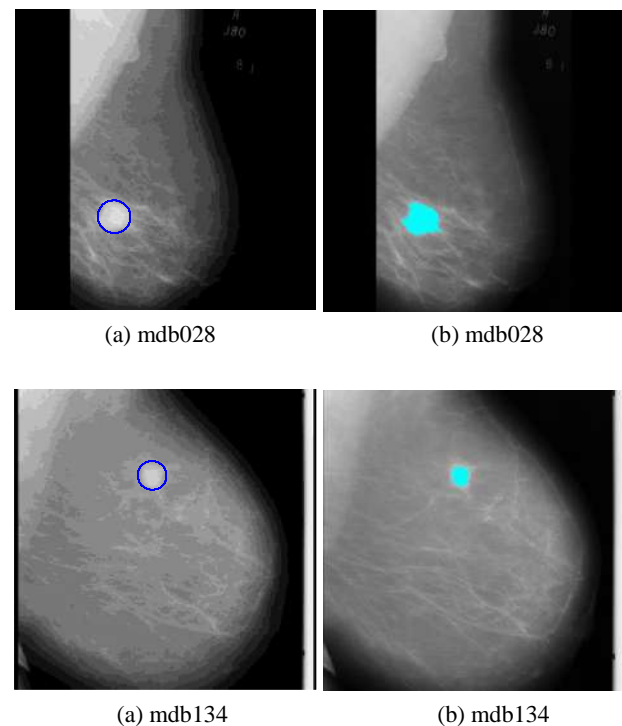
For the learning phase, in this work we started the training by picking two to four different images randomly from the dataset and finally we ended up with two images mdb028 and mdb005 that we believe from the good training results that they contain all the complexity zones and features common in mammography in all dataset images, and in order to just validate our algorithm, we have tested one or two images from each kind of class of abnormality present, CIRC, SPIC ... etc:

- mdb025 and mdb132, for Well-defined/circumscribed masses (CIRC) [33].
- mdb184, for Spiculated masses (SPIC) [33].
- mdb134, mdb271 and mdb274, for Other, ill-defined masses (MISC) [33].
- mdb136 and mdb310, for Normal breast (NORM) [33].

Figure 6 illustrates the obtained results (detected masses are displayed in cyan color) compared to expert decision (masses centers coordinates and radiuses shown in blue color). Table 1 demonstrates the Jaccard index that was achieved by utilizing this fusion algorithm.

So, as shown in figure 6 and table 1, the results obtained on these taken images from MIAS database are promising and they show the effectiveness of our fusion algorithm for the masses segmentation on mammography images.

We can see clearly that, for these taken images, the proposed algorithm gives a good result whatever the kind of class of abnormality present, CIRC, SPIC, ... etc. And also, we don't have any false detection in the cases of normal breast. So, the proposed method has the potential to identify the presence of any masses in the taken mammogram images.



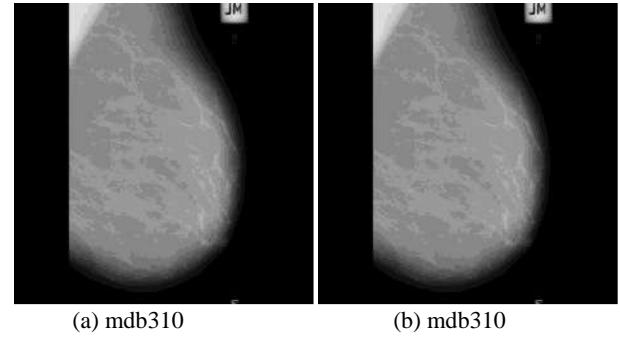
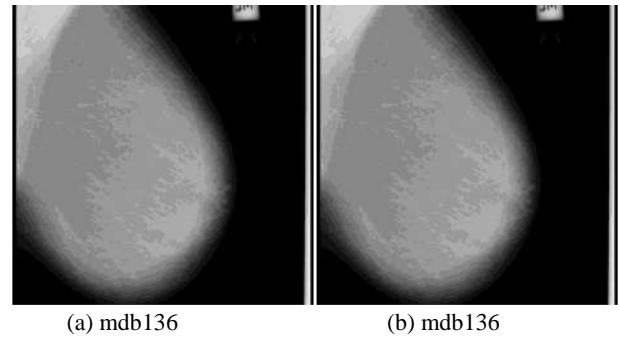
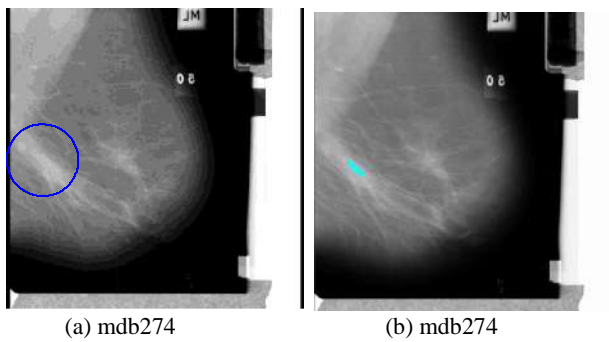
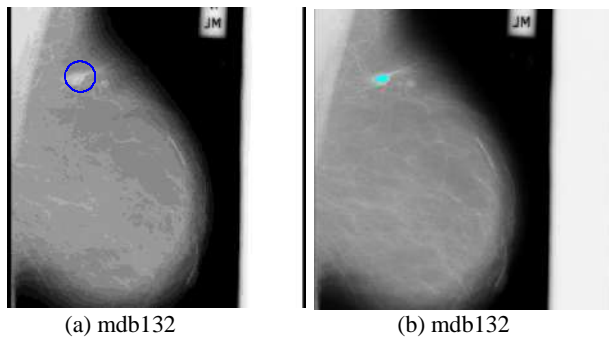
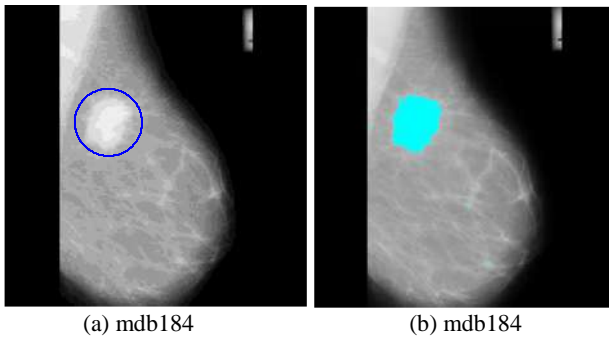
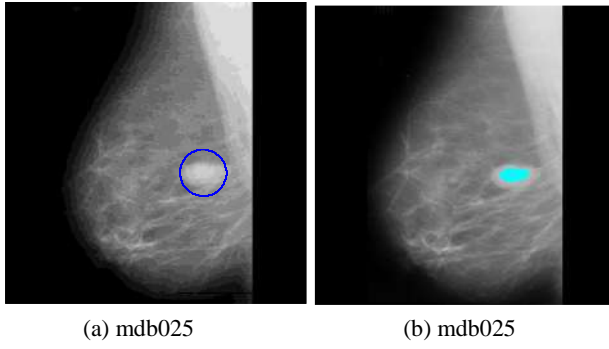
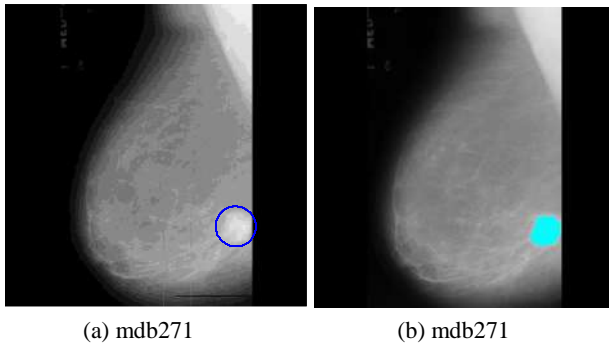


Fig. 6 Experimental results on MIAS database: The first column (images a) contains the original mammograms with expert masse location. The second column (images b) represents the classification results using our approach.

Table 1. Jaccard index of the fusion algorithm.

Image	jaccard index (%)
mdb028	99.84
mdb025	99.44
mdb132	99.83
mdb184	99.69
mdb134	99.79
mdb271	99.86
mdb274	99.49
mdb136	100
mdb310	100

Figure 7 illustrates a comparison of our proposed approach, on MIAS images mdb184 and mdb028, with another unsupervised techniques proposed by Kanta Maitra et al [34], based on Divide and conquer algorithm, and Boulehmi Helaet al [2], based on Generalized Gaussian Density.

As seen in figure 7, the proposed approach has the advantage of being simple and precise; we have exactly detected the shape of the present masses.

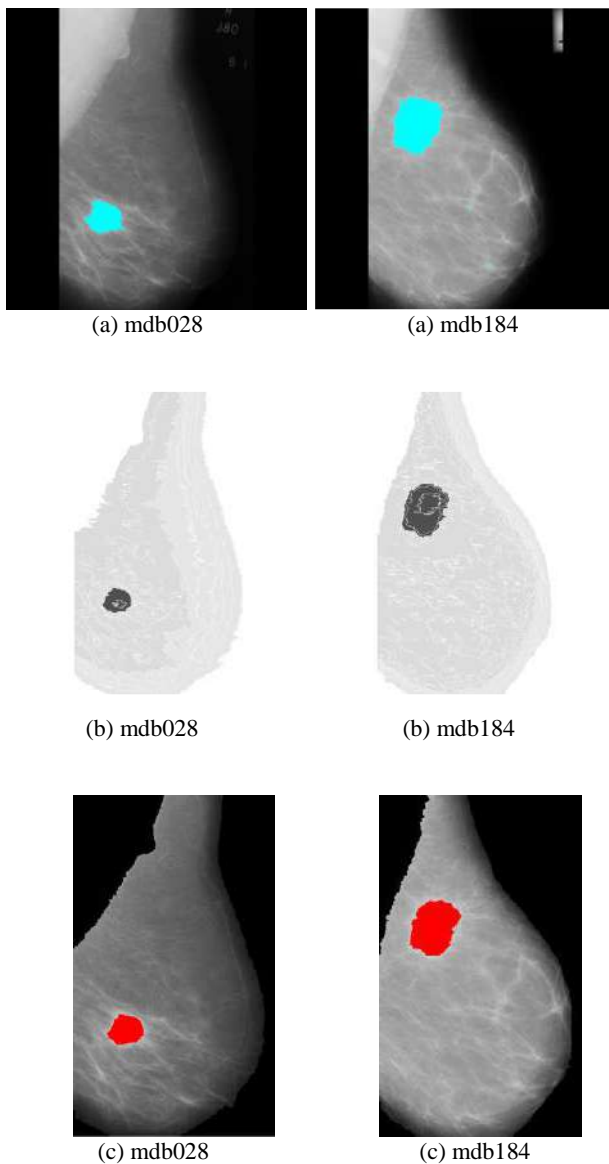


Fig. 7 Comparison between results from:
(a) Our approach, (b) [34], (c) [2].

The outcomes of our contribution demonstrate that it is possible to reach excellent fusion performance by neatly selecting the best fusion method. We also note that by our fusion method, the segmentation results of the mammography images are much improved as compared to other works.

VI. Conclusion

In this article, we have presented and discussed a new approach for the segmentation of mammography images based totally on information fusion. We started by extracting the features using the wavelet transform. After that, the estimated vector of features for every pixel was sent to the neural network classifier for primary labeling. Next, a new fusion model for improving decision-making is used, it consists of combining the scores of each pixel within a sliding window. The proposed fusion algorithm was tested on mammography images from MIAS dataset.

This research has shown that this method is very effective for the automatic detection of abnormalities in digital mammogram.

As perspective we will complete the masses detection system by classifying abnormal mammographic images into benign and malignant.

REFERENCES

- [1] Wisudawati Lulu M, Sarifuddin M, Wibowo Eri P, Abdullah Arman A (2020) "Feature Extraction Optimization with Combination 2D-Discrete Wavelet Transform and Gray Level Co-Occurrence Matrix for Classifying Normal and Abnormal Breast Tumors". *Modern Applied Science*, 14 (15): 5.
- [2] Boulehmi H, Mahersia H, Hamrouni K (2013) "Unsupervised Masses Segmentation Technique Using Generalized Gaussian Density". *International Journal of Image Processing and Graphics (IJIPG)*, 1, (2).
- [3] Fahnun B U, Mutiara A B, Wibowo E P, Arlan J, Latief A (2018) "Filtering techniques for noise reduction in liver ultrasound images". Phd thesis, Program Doktor Teknologi Informasi Universitas Gunadarma, 261–266. <https://doi.org/10.1109/EIConCIT.2018.8878547>
- [4] Abdelsamea M M, Mohamed M H, Bamatraf M (2019) "Automated classification of malignant and benign breast cancer lesions using neural networks on digitized mammograms". *Cancer Informatics*, 18, 1–3. Sage Journal. <https://doi.org/10.1177/1176935119857570PMid:31244522> PMID:PMC6580711
- [5] Lestari D, Madenda S, Massich J (2015) "A Segmentation Algorithm for Breast Lesion Based on Active Contour Model and Morphological Operations". *Advanced Science, Engineering and Medicine*, 7, 920-924. 10.1166/asem.2015.1786. <https://doi.org/10.1166/asem.2015.1786>
- [6] Al-antari M A, Al-masni M A, Park SU et al (2018) "An Automatic Computer-Aided Diagnosis System for Breast Cancer in Digital Mammograms via Deep Belief Network". *J. Med. Biol. Eng.* 38, 443–456. <https://doi.org/10.1007/s40846-017-0321-6>
- [7] Pratiwi M, Alexander, Harefa J, Nanda S (2015) "Mammograms classification using gray-level co-occurrence matrix and radial basis function neural network". <https://doi.org/10.1016/j.procs.2015.07.340>
- [8] Biswas R, Nath A, Roy S (2016) "Mammogram classification using gray-level co-occurrence matrix for diagnosis of breast cancer". 161–166. <https://doi.org/10.1109/ICMETE.2016.85>
- [9] Ergin S, Esener I, Yuksel T (2016) "A genuine glcm based feature extraction for breast tissue classification on mammograms". *International Journal of Intelligent Systems and Applications in Engineering*, 124–124. <https://doi.org/10.18201/ijisae.269453>
- [10] Ucar K, Kocer H E (2017) "Breast cancer classification with wavelet neural network". *International Artificial Intelligence and Data Processing Symposium (IDAP)*, 1–5. <https://doi.org/10.1109/IDAP.2017.8090347>
- [11] Putra A J (2018) "Mammogram classification scheme using 2d discrete wavelet and local binary pattern for detection of breast cancer". *Journal of Physics: Conference Series*, <https://doi.org/10.1088/1742-6596/1008/1/012004>
- [12] Pawar M, Talbar S (2018) "Local entropy maximization-based image fusion for contrast enhancement of mammogram". *Journal of King Saud University Computer and Information Sciences*. <https://doi.org/10.1016/j.jksuci.2018.02.008>
- [13] Pawar M, Talbar S, Dudhane A (2018) "Local binary patterns descriptor based on sparse curvelet coefficients for false-positive reduction in mammograms". *Journal of healthcare engineering*. <https://doi.org/10.1155/2018/5940436PMid:30356422> PMID:PMC6178513
- [14] Shachor Y, Greenspan H, Goldberger J (2018) "A mixture of views network with applications to the classification of breast microcalcifications". *Computer Vision and Pattern Recognition*.

- [15] Bekker A J, Shalhon M, Greenspan H, Goldberger J (2016) "Multi-view probabilistic classification of breast microcalcifications". In: Proceedings of the IEEE Transactions on Medical Imaging.
- [16] Dhungel N, Carneiro G, Bradley A P (2017) "Fully automated classification of mammograms using deep residual neural networks". In: Proceedings of the 14th International Symposium on Biomedical Imaging (ISBI 2017), IEEE.
- [17] Li Y, Chen H, Cao L, Ma J (2016) "A survey of computer-aided detection of breast cancer with mammography". J. Health Med. Inform. 7 (4).
- [18] Geras K J, Wolfson S, Shen N W Y, Kim S.G, Kim E, Heacock L, Parikh U, Moy L, Cho K (2017) "High-resolution breast cancer screening with multi-view deep convolutional neural networks". In: Proceedings of the Computer Vision and Pattern Recognition.
- [19] Hérault J, JuttenCh(1994) "Réseaux neuronaux et traitement du signal". Hermès, Paris
- [20] Jodouin J F (1994) "Les réseaux de neurones. Principes et définitions". Hermès, Paris
- [21] Suckling J, Parker J, Dance D, Astley S, Hutt I, Boggis C, Ricketts I et al (2015) "Mammographic Image Analysis Society (MIAS) database v1.21 [Dataset]". <https://www.repository.cam.ac.uk/handle/1810/250394>
- [22] Grossman A, Morlet J (1984) "Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape". SIAM Journal on Mathematical Analysis, 15, (4), pp. 723-736.
- [23] Sahbani M H, Hamrouni K (2005) "Segmentation d'images texturées par transformée en ondelettes et classification C-moyenne floue". Proc. 3rd Int. Conf. Sciences of Electronic, Technologies of Information and Telecommunications, Tunisia, March 27-31.
- [24] Iftene T, Safia A (2004) "Comparaison Entre La Matrice De Cooccurrence Et La Transformation En Ondelettes Pour La Classification Texturale Des Images Hrv (Xs) De Spot". Télédétection, 4, (1), pp. 39-49.
- [25] Anibou Ch, Saidi M N, Aboutajdine D (2015) "Classification of Textured Images Based on Discrete Wavelet Transform and Information Fusion". J. Inf. Process. Syst, 11, (3), pp. 421-437.
- [26] Mallat S (1989) "A theory of multiresolution signal decomposition: the wavelet representation". IEEE Trans. Pattern Analysis and Machine Intelligence, 11, (7), pp. 674-693.
- [27] Scheunders P, Livens S, Van de Wouwer G et al (1997) "Wavelet-based Texture Analysis". Int. J. Computer Science and Information Management.
- [28] Arivazhagan S, Ganesan L (2003) "Texture segmentation using wavelet transform". Pattern Recognition Letters, 24, (16), pp. 3197-3203.
- [29] Dempster A P (1967) "Upper and lower probabilities induced by a multivalued mapping". The annals of mathematical statistics, pp. 325-339.
- [30] Shafer G A (1976) "mathematical theory of evidence". vol. 1: Princeton university press Princeton.
- [31] Smets P (1990) "The combination of evidence in the transferable belief model". IEEE Transactions on pattern analysis and machine intelligence, vol. 12, pp. 447-458.
- [32] Zitouni A (2020) "Image Processing Methodology and Textures Analysis for their Segmentation". PhD. Dissertation, University Amar Telidji of Laghouat, Algeria.
- [33] The mini-MIAS database of mammograms, <http://peipa.essex.ac.uk/info/mias.html>. Accessed 20 Dec 2020.
- [34] Kanta I, Maitra S, Nag S, Bandyopadhyay K (2011) "Detection of Abnormal Masses using Divide and Conquer Algorithm in Digital Mammogram". Int. J. Emerg. Sci., 1(4), 767-786.