International Journal of Reconfigurable and Embedded Systems (IJRES)
Vol. 13, No. 3, November 2024, pp. 767~773
ISSN: 2089-4864, DOI: 10.11591/ijres.v13.i3.pp767-773

Leveraging the learning focal point algorithm for emotional intelligence

Salah Eddine Mansour[1], Abdelhak Sakhi[1], Larbi Kzaz[2], Abderrahim Sekkaki[1]
[1]Electrical and Industrial Engineering Information Processing IT and Logistics (GEIIL), Faculty of Sciences Ain Chock, Hassan II University, Casablanca, Morocco
[2]De Higher Institute of Commerce and Business Administration, Hassan II University, Casablanca, Morocco

Article Info

Article history:

Received Nov 17, 2023
Revised Jan 26, 2024
Accepted Feb 2, 2024

Keywords:

Convolutional neural network
Emotional intelligence
Learning focal point algorithm
Neural network
Perceptron

ABSTRACT

One of the secrets of the success of the education process is taking into account the learner's feelings. That is, the teacher must be characterized by high emotional intelligence (EI) to understand the student's feelings in order to facilitate the indoctrination process for him. Within the framework of the project to create a robot teacher, we had to add this feature because of its importance. In this article, we create a computer application that classifies students' emotions based on deep learning and learning focal point (LFP) algorithm by analyzing facial expressions. That is, the robot will be able to know whether the student is happy, excited, or sad in order to deal with him appropriately.

Corresponding Author:

Salah Eddine Mansour
Electrical and Industrial Engineering Information Processing IT and logistics (GEIIL)
Faculty of sciences Ain Chock, Hassan II University
Casablanca, Morocco
Email: 19mansour94@gmail.com

1. INTRODUCTION

The main text in the era of rapid technological progress, the fusion of artificial intelligence (AI), and robotics is reshaping various aspects of our lives. One area where this convergence holds enormous promise is education. This research is an integral part of a comprehensive initiative aimed at formulating and improving the pioneering teacher robot, which is an advanced combination of AI and image processing technologies. The primary focus of our research is to harness the power of AI to provide robot teachers with emotional intelligence (EI), which is of great importance in the teaching process [1]. Therefore, our robot has to classify a student's emotions by analyzing his facial expressions.

Understanding the critical role of EI in education unveils a profound connection between a teacher's ability to convey information and a student's capacity to comprehend and engage with it. The dynamics of learning extend beyond the mere transmission of facts; they are deeply intertwined with the emotional landscape within the classroom [2], [3]. Exploring the significance of EI in educators offers a profound insight into how the mastery of emotions shapes the learning environment, influences student-teacher interactions, and ultimately impacts the absorption and retention of knowledge [4].

In this work, we have created a system which make the teacher robot has the emotional intelligence. We tried using convolutional neural network (CNN) with pooling layers, but the results were not satisfactory enough. Therefore, we conducted another experiment by replacing the pooling layers with a learning focal point (LFP) layer, and the results were much better than the first experiment by comparing accuracy and other parameters. Using the LFP layer instead of layer pooling is exactly our scientific contribution in this research and what we will discuss in the rest of this article.

By researching the methods used in AI to classify emotions through facial expressions, we found that most of them rely on deep learning through training neural networks. This means that there are only innovations in deep learning architecture. One of the most used architectural structures in facial recognition is the CNN, because we find many solutions and frameworks such as DeepFace, FaceNet, LightCNN, and others that rely on CNN which in turn relays on pooling layers [5]. Pooling helps to manage computational resources, reduce the number of parameters, and focus on the most important features in the data. CNNs have proven highly effective in various computer vision tasks due to their ability to automatically learn hierarchical representations of features from input data [6]. CNNs derive their name from the convolutional layers, which apply filters or kernels to the input data. These filters slide over the input, capturing local patterns and features [7]. The result is a feature map that highlights relevant structures in the input. Among the core components used in CNN, pooling layers play a crucial role in reducing the spatial dimensionality of the input data, which helps control he network's computational complexity, and parameter count [8]. Pooling is a down-sampling process applied after convolutional layers to retain essential information while reducing spatial resolution. There are two common types of pooling layers: max pooling an average pooling [9]. Max pooling is a popular pooling technique that selects the maximum value from a group of neighboring pixels in the input. The idea is to retain the most important features and reduce the spatial dimensions. Average pooling computes the average value from a group of neighboring pixels.

In this article, in section 2, we will explain the LFP algorithm and how we can use it for emotion classification. In section 3, we will discuss the results and performance of using the LFP algorithm in this case. Followed by the conclusion. we will present the design and mathematical theories on which the LFP algorithm is based.

2. METHOD

2.1. Emotion classification system

Our system that classifies emotions based on facial expressions; it generally consists of three main modules as we see in Figure 1.

– The first module: this unit relies on the open source computer vision (OpenCV) library, so that the student's face is identified and traced back through video or image analysis.
– The second module: it is based on the LFP algorithm to extract the key regions of the face by returning the coordinates of the key squares of the face.
– The third module: its role is to calculate the weights of the neural network to classify facial expressions.

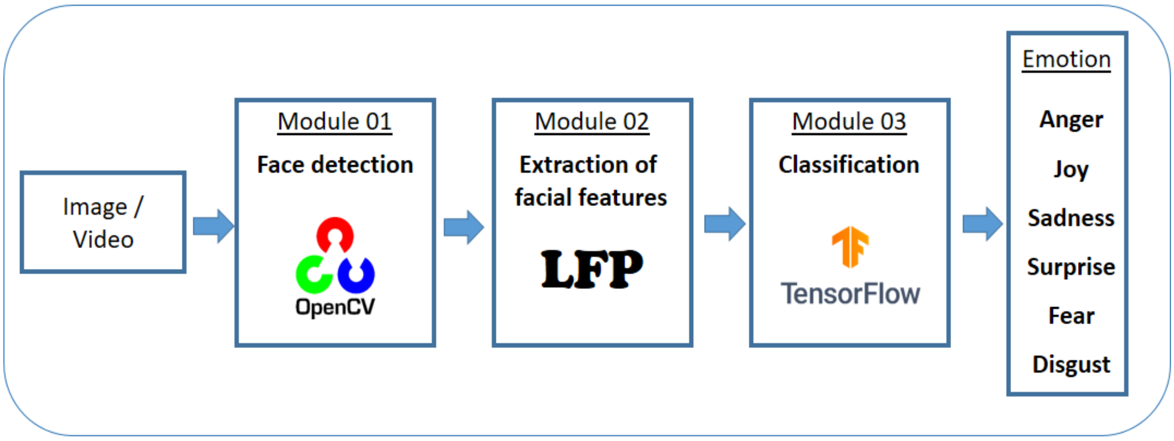

Figure 1. Emotion classification system modules

2.2. Dataset

The dataset is used in “representation learning challenges: facial expression recognition (FER) challenge” on Kaggle. Its main task is to classify facial expressions into different emotion categories. Typically, the dataset consists of images labeled with one of several emotion classes, such as happiness, sadness, anger, surprise, fear, disgust, and neutral. We used this dataset to train the neural network. The specific database used in the Kaggle FER Challenge might be the FER 2013 (FER2013) dataset, which consists of 48×48-pixel grayscale images of faces. Each image is labeled with one of seven emotions. It contains around 35,000 images, categorized into seven different emotions.

2.3. OpenCV

OpenCV library serves as a comprehensive toolbox for performing various image processing tasks, including image filtering, transformation, enhancement, and geometric operations like resizing, cropping, and rotation [10], [11]. It provides a wide range of tools and functionalities for image and video processing, including features for computer vision, machine learning, and image analysis [12]. OpenCV is written in C++ and also has interfaces for Python, Java, and other languages. It provides a comprehensive set of functions for basic to advanced image processing tasks, including image filtering, morphological operations, and color space transformations. OpenCV includes a variety of computer vision algorithms for object detection, feature extraction, image stitching, and camera calibration. It integrates with machine learning frameworks and includes machine learning algorithms for tasks such as classification, clustering, and regression. Also, OpenCV supports deep learning frameworks like TensorFlow and PyTorch, allowing users to work with pre-trained deep learning models and build their own models [13]. In our case, we used OpenCV to detect and extract images of children's faces, as we see in Figure 2, from the cameras that will be installed.

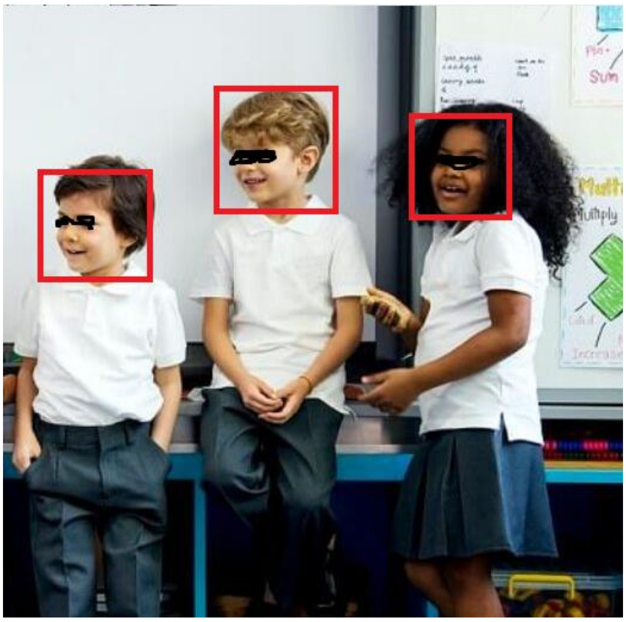

Figure 2. Face extraction and detection using OpenCV

2.4. Learning focal point algorithm

The LFP algorithm helps reduce the number of parameters and focus on the most important features in the data. Through the flowchart in Figure 3, we can understand how the LFP algorithm works. We take a set of images of the same size and then divide each image into squares, as we see in Figure 4. We perform perceptron training on each square, then we calculate their accuracy through the (1), and finally get the coordinates (x, y, width and height) of the high-precision squares. To sum up, LFP algorithm relies on the accuracy of perceptron training on different squares of images to return their coordinates. This means that it has the same role as max pooling and mean pooling used in CNN. As a scientific contribution, we replace the pooling layers with an LFP layer.

The LFP algorithm employs a systematic approach to identify key focal points within input images. Beginning with image segmentation into distinct parts (Div[i]), the algorithm utilizes perceptrons to extract relevant features. Accuracy calculations evaluate focal point identification effectiveness, followed by sorting Div[i] segments based on precision. The algorithm returns coordinates of the segments with the highest precision, signifying the localization of key focal points. This process distills image complexity, facilitating a focused analysis of essential features.

$$Accuracy = \frac{True\ Positive + False\ Positive}{Total} \quad (1)$$

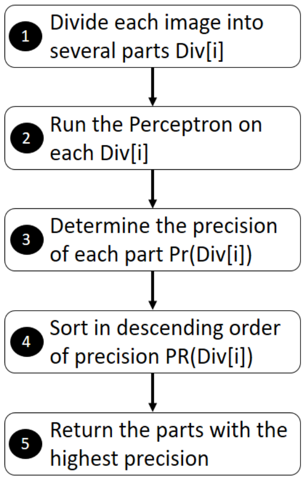

Figure 3. Flowchart of the LFP algorithm

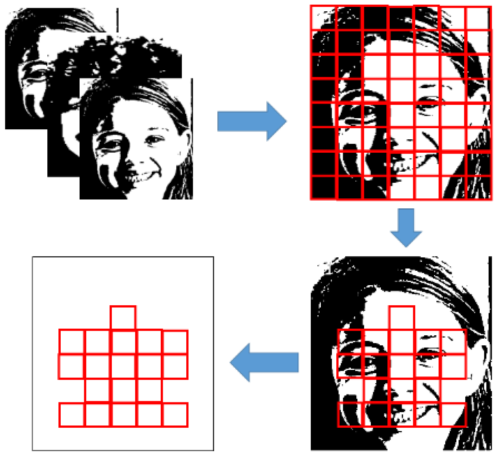

Figure 4. Execution of perceptron on several square

2.5. TensorFlow

TensorFlow is an open-source machine learning framework developed by the Google Brain team [14], [15]. It's designed to facilitate the development and deployment of machine learning models. TensorFlow has strong support for building and training CNN [16]. CNNs are a specific type of neural network architecture that is particularly effective for image-related tasks, such as image classification, object detection, and image segmentation [17]. TensorFlow provides a comprehensive ecosystem for developing and deploying machine learning models. It offers a high-level application programming interface (API) called Keras that simplifies the process of building and training neural networks, including CNNs [18]. TensorFlow allows users to define, train, and deploy complex models efficiently. Keras is an open-source high-level neural networks API that is now tightly integrated into TensorFlow [19]. With Keras, we can easily build and experiment with various neural network architectures, including CNNs, using a clear and user-friendly syntax. Therefore, we train the neural network by using the squares found by LFP algorithm as we see in Figure 5.

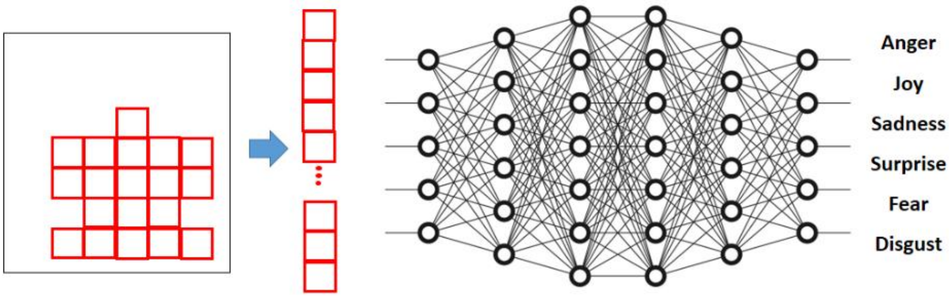

Figure 5. Training the neural network using the squares found by the LFP algorithm

3. RESULTS AND DISCUSSION

After learning about the methods used in this research, we will present the results obtained using the LFP algorithm and compare them with the results of max pooling. We used the FER challenge database on Kaggle. We conducted two experiments as we see in Figure 6 in order to obtain the models: the first we used the LFP algorithm in the CNN and the second we used max pooling. The performance of these models was evaluated based on classification accuracy (CA), precision, recall, F1 score, and receiver operating characteristic-area under the curve (ROC-AUC) values [20]. These indices (these functions) were generated automatically using TensorFlow library. Accuracy in AI refers to how well a machine learning model performs in making correct predictions or classifications compared to the total number of predictions. It's a measure of how often the model is correct. For classification tasks, it measures the percentage of correctly predicted instances among all instances. While accuracy is an essential metric, it might not give a complete picture, especially in cases of imbalanced datasets. Precision in AI, particularly in classification tasks, measures the ratio of true positive predictions to the total predicted positives [21]. It focuses on the relevance of the model's predictions. Precision is about how many of the predicted positive instances are actually relevant. Recall in the context of AI and machine learning refers to the ability of a model to correctly identify all relevant instances or data points within a dataset. It's a measure of a model's completeness in capturing all the relevant information. It calculates the ratio of true positives to the sum of true positives and false negatives [22]. F1 score this metric considers both precision and recall to provide a single score that represents the model's performance. In such cases, the F1 score becomes valuable. It considers both false positives (precision) and false negatives (recall), providing a balanced assessment of the model's effectiveness. It's particularly beneficial when there's a need to avoid either missing positive cases (low recall) or misclassifying negative cases as positive (low precision) [23]. It's a useful metric when you want to balance between precision and recall and need a single value to assess a model's performance. The ROC-AUC is valuable because it evaluates a model's ability to discriminate between positive and negative classes across various threshold values [24], [25]. It's commonly used in binary classification problems and provides a robust assessment of the model's performance regardless of the threshold chosen for classification. A higher AUC generally suggests that the model is better at distinguishing between the classes.

We did two experiments as we see in Figure 6. In the first experiment, we implemented the LFP algorithm with neural network and in the second experiment we just used max pooling. We not in Table 1 the results of these experiments which demonstrate the strength of the LFP algorithm, which presents hight precision. Furthermore, comparing experiments 1 and 2, we observe that we can increase the CA up to 10% if we use the LFP algorithm.

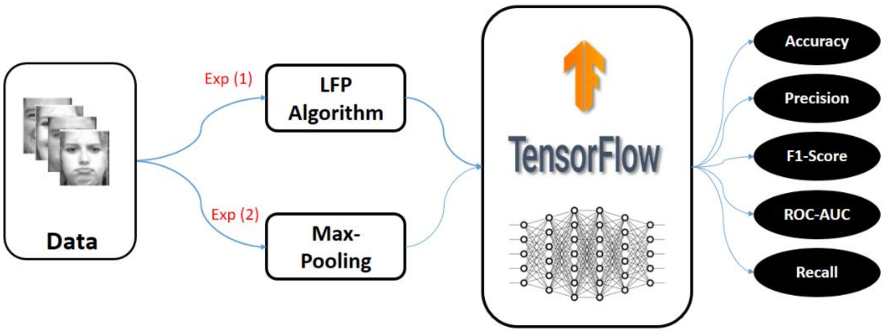

Figure 6. Training the neural network using the squares found by the LFP algorithm

Table 1. Results of two experiments

Experiment N°	Algorithm	CA	Precision	Recall	F1-Score	ROC-AUC
1	LFP algorithm	0.931	0.931	0.930	0.930	0.944
2	Max pooling	0.826	0.825	0.824	0.824	0.835

4. CONCLUSION

In this work, we presented the strength of LFP algorithm to classify the emotions by ana lysing the facial expressions. And this we help to develop our robot teacher project by adding to it the emotional intelligence. We also present how we can change CNN architecture by replace max pooling and average pooling by LFP algorithm. In two experiments, we have compared two methods by using some estimation indices including AUC, CA, precision, F1 score, and recall. The numerical output result shows that the classifier based on the LFP algorithm performs better than its competitor in terms of accuracy (0.931), which means increased CA up to 10%. The LFP algorithm will open our eyes to a long road of scientific research because our algorithm is based on a single layer of sensory perception. However, we have had great results. So, if we use other machine learning algorithms that are better than perceptron. It is necessary to obtain increased CA.

REFERENCES

[1] K. Ahmad *et al.*, "Artificial intelligence in education: a panoramic review," *ML in Education and Information with authors*, pp. 1–51, 2020, doi: 10.35542/osf. io/zvu2n.
[2] N. Humphrey, A. Curran, E. Morris, P. Farrell, and K. Woods, "Emotional intelligence and education: a critical review," *Educational Psychology*, vol. 27, no. 2, pp. 235–254, Apr. 2007, doi: 10.1080/01443410601066735.
[3] K. Keefer, J. D. A. Parker, and D. H. Saklofske, *Emotional intelligence in education*. in The Springer Series on Human Exceptionality. Cham: Springer International Publishing, 2018. doi: 10.1007/978-3-319-90633-1.
[4] B. L. Kennedy and R. Junker, "The evolution of 'loaded moments' toward escalation or de-escalation in student–teacher interactions," *Review of Educational Research*, Oct. 2023, doi: 10.3102/00346543231202509.
[5] F. Demir, A. M. Ismael, and A. Sengur, "Classification of lung sounds with CNN model using parallel pooling structure," *IEEE Access*, vol. 8, pp. 105376–105383, 2020, doi: 10.1109/ACCESS.2020.3000111.
[6] S. Srinivas, R. K. Sarvadevabhatla, K. R. Mopuri, N. Prabhu, S. S. S. Kruthiventi, and R. V. Babu, "A taxonomy of deep convolutional neural nets for computer vision," *Frontiers Robotics AI*, vol. 2, Jan. 2016, doi: 10.3389/frobt.2015.00036.
[7] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5455–5516, Dec. 2020, doi: 10.1007/s10462-020-09825-6.
[8] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, no. 1, pp. 1–74, Mar. 2021, doi: 10.1186/s40537-021-00444-8.
[9] N. Akhtar and U. Ragavendran, "Interpretation of intelligence in CNN-pooling processes: a methodological survey," *Neural Computing and Applications*, vol. 32, no. 3, pp. 879–898, Feb. 2020, doi: 10.1007/s00521-019-04296-5.
[10] D. K. Mishra, S. E. Umbaugh, N. Lama, R. Dahal, D. J. Marino, and J. Sackman, "Image processing and pattern recognition with CVIPtools MATLAB toolbox: automatic creation of masks for veterinary thermographic images," in *Applications of Digital Image Processing XXXIX*, A. G. Tescher, Ed., Sep. 2016. doi: 10.1117/12.2238183.
[11] K. Pulli, A. Baksheev, K. Kornyakov, and V. Eruhimov, "Realtime computer vision with OpenCV," *Queue*, vol. 10, no. 4, pp. 40–56, 2012, doi: 10.1145/2181796.2206309.
[12] D. Ireri, E. Belal, C. Okinda, N. Makange, and C. Ji, "A computer vision system for defect discrimination and grading in tomatoes using machine learning and image processing," *Artificial Intelligence in Agriculture*, vol. 2, pp. 28–37, Jun. 2019, doi: 10.1016/j.aiia.2019.06.001.
[13] S. A. Sanchez, H. J. Romero, and A. D. Morales, "A review: comparison of performance metrics of pretrained models for object detection using the TensorFlow framework," *IOP Conference Series: Materials Science and Engineering*, vol. 844, no. 1, pp. 1–16, Jun. 2020, doi: 10.1088/1757-899X/844/1/012024.
[14] G. Al-Bdour, R. Al-Qurran, M. Al-Ayyoub, and A. Shatnawi, "A detailed comparative study of open source deep learning frameworks," *arXiv e-prints*, p. arXiv:1903.00102, 2019, doi: 10.48550/arXiv.1903.00102.
[15] A. I. Khan and A. Al-Badi, "Open source machine learning frameworks for industrial internet of things," *Procedia Computer Science*, vol. 170, pp. 571–577, 2020, doi: 10.1016/j.procs.2020.03.127.
[16] B. Pang, E. Nijkamp, and Y. N. Wu, "Deep learning with tensorflow: a review," *Journal of Educational and Behavioral Statistics*, vol. 45, no. 2, pp. 227–248, Apr. 2020, doi: 10.3102/1076998619872761.
[17] H. Lin, Z. Shi, and Z. Zou, "Fully convolutional network with task partitioning for inshore ship detection in optical remote sensing images," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1665–1669, Oct. 2017, doi: 10.1109/LGRS.2017.2727515.
[18] B. T. Chicho and A. B. Sallow, "A comprehensive survey of deep learning models based on keras framework," *Journal of Soft Computing and Data Mining*, vol. 2, no. 2, pp. 49–62, Oct. 2021, doi: 10.30880/jscdm.2021.02.02.005.
[19] D. Grattarola and C. Alippi, "Graph neural networks in tensorflow and keras with spektral [application notes]," *IEEE Computational Intelligence Magazine*, vol. 16, no. 1, pp. 99–106, Feb. 2021, doi: 10.1109/MCI.2020.3039072.
[20] J. Miao and W. Zhu, "Precision–recall curve (PRC) classification trees," *Evolutionary Intelligence*, vol. 15, no. 3, pp. 1545–1569, Sep. 2022, doi: 10.1007/s12065-021-00565-2.
[21] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing and Management*, vol. 45, no. 4, pp. 427–437, Jul. 2009, doi: 10.1016/j.ipm.2009.03.002.
[22] P. A. Flach and M. Kull, "Precision-recall-gain curves: PR analysis done right," *Advances in Neural Information Processing Systems*, pp. 838–846, 2015.
[23] C. Cao, D. Chicco, and M. M. Hoffman, "The MCC-F1 curve: a performance evaluation technique for binary classification," *arXiv preprint*, pp. 1–17, 2020.
[24] L. A. Dalton, "Optimal ROC-based classification and performance analysis under bayesian uncertainty models," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 13, no. 4, pp. 719–729, Jul. 2016, doi: 10.1109/TCBB.2015.2465966.
[25] A. M. Carrington *et al.*, "A new concordant partial AUC and partial C statistic for imbalanced data in the evaluation of machine learning algorithms," *BMC Medical Informatics and Decision Making*, vol. 20, no. 1, pp. 1–12, Dec. 2020, doi: 10.1186/s12911-019-1014-6.

BIOGRAPHIES OF AUTHORS

Salah Eddine Mansour completed his higher education in Casablanca, Morocco. He started his Ph.D. in artificial intelligence in 2020. Currently, he is a professor of informatics at the Ministry of Education in Morocco. He can be contacted at email: 19mansour94@gmail.com.

Abdelhak Sakhi completed his higher education in Casablanca, Morocco. He started his Ph.D. in artificial intelligence in 2020. Currently, he is a professor of mathematics at the Ministry of Education in Morocco. He can be contacted at email: sakhi442@gmail.com.

Larbi Kzaz received doctorate degree in computer science from Lille University of Science and Technology, France, in 1985. He is professor in information systems at ISCAE Business School, Casablanca, Morocco. His research interests include semantic integration, data warehouse design, machine learning and big data applications in digital. He can be contacted at email: kzaz.larbi@gmail.com.

Abderrahim Sekkaki completed his graduate studies in Toulouse, France. After having passed his master's degree and his DEA in computer science, he began his Ph.D. in network management, defended in 1991. He crowned his career in computer science with a state thesis in 2002. Currently, he is a full professor in University Hassan II of Casablanca. He can be contacted at email: sekkabd@gmail.com.