Indonesian Journal of Electrical Engineering and Computer Science
Vol. 33, No. 1, January 2024, pp. 416~424
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v33.i1.pp416-424

Design, security and implementation of learning focal point algorithm in a docker container

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

Article Info

Article history:

Received Oct 23, 2023
Revised Nov 7, 2023
Accepted Nov 11, 2023

Keywords:

Cloud computing
Deep learning
Docker
Oauth2.0
Perceptron

ABSTRACT

Artificial intelligence is not smart enough. This is why we are looking for complementary algorithms in order to increase the performance of machine learning and neural networks (NN). We have innovated an algorithm called learning focal point (LFP). This algorithm will help us increase the intelligence of machine learning and the NN. In this article, we will present the algorithm in detail, starting with the mathematical and theoretical principles, passing through the development and deployment in cloud docker containers, and ending with the security of its application programming interface (API). Finally, we are going to do a test in which we apply it in the case of using tin cans.

Corresponding Author:

Salah Eddine Mansour
Electrical and Industrial Engineering Information Processing IT and logistics (GEIIL)
Faculty of Sciences Ain Chock, Hassan 2 University
Casablanca, Morocco
Email: 19mansour94@gmail.com

1. INTRODUCTION

This work falls under an industrial project that aims to develop the fish canning industry in Morocco using artificial intelligence (AI). Therefor, our primary goal is to create a system based on AI and image processing to supervise the quality of canning in the factory. As a first version, we created a model based solely on neural network (NN). Through a set of tests on the model, it was found that the classification accuracy (CA) was not sufficient due to the quality of the data. That's why we created a complementary algorithm called learning focal point (LFP) which will be the subject of this research paper. This algorithm that takes the images as input and returns the coordinates of the essential divisions of the images. Based on these divisions, we do the training and testing. The results were very wonderful. This is what we will see in the discussion section.

Intuitevely, we searched for algorithms and methods like LFP algorithm that solve the data refinement problem. We found many algorithms like non-negative matrix factorization (NMF) which is a dimensionality reduction and data decomposition technique used in machine learning and data analysis [1]. In addition to these methods, there is local binary patterns (LBP) is a type of feature used in computer vision and image processing [2]. It is a type of feature descriptor which encodes local spatial information of an image by comparing each pixel with its neighbor pixels. Remember that the principal component analysis (PCA) method plays the same role, but in a different way. It is a technique used to reduce the dimensionality of a dataset based on algebraic mathematics [3].

The difference between our method and the previously mentioned methods is that the latter relies heavily on matrix arithmetic and mathematical algebra. While our method relies on the AI perceptron algorithm to calculate the accuracy, and based on this accuracy, the main part of the image is selected. Moreover, existing methods take images or a dataset as input parameter and return an optimal image or dataset, while our method works on a dataset of images and returns the coordinates of the main divisions of the image. One of the advantages of this technique is that we only compute the data once in the training step, while other methods are forced to recalculate at each test for a specific image.

In this paper, in section 2, we will presente related works then in the section 3, we will present the design and mathematical theories on which the LFP algorithm is based. Moreover, we will see application programming interface (API) requests and responses created to allow other developers to use our algorithm. We will also discuss the scurity of API and its containerization in cloud computing using docker. Finally, in section 4, we make a discussion of the results and performance of our own algorithm based on fish can image processing. Followed by the conclusion.

2. RELATED WORKS

Through our research, it turns out that there are many attempts to solve the data refinement problem, because of the importance of this step in creating better and more accurate models. Most of the methods we found are in the area of facial recognition. And we mention NMF is a technique used in machine learning and data mining that is used to identify patterns and uncover the underlying structure of a dataset [4]. It is based on the assumption that the data is composed of non-negative factors that, when multiplied together, can produce the original dataset [5]. NMF works by decomposing the data into two matrices: a basis matrix and a coefficient matrix. The goal of NMF is to find the best possible decomposition of the data into non-negative components that can be used to reconstruct the original dataset [6]. NMF can be used for a variety of tasks including data compression, feature extraction, clustering, and dimensionality reduction. Thanks to this method, many methods based on it have been designed, which we call the NMF family.

In addition to these methods, there is another method, LBP are a type of feature used in computer vision and image processing [7], [8]. It is a type of feature descriptor which encodes local spatial information of an image by comparing each pixel with its neighbor pixels. The feature is computed by comparing the intensity value of a pixel to the intensity values of its neighbors [9]. The resulting binary pattern is then used to identify the features in the image. LBP is often used for object recognition, facial recognition, and texture analysis. And in this regard, LBP can be considered as maxpooling and min-pooling which are techniques used in convolutional neural networks (CNN) to reduce the size of an input image, while preserving the most important features [10]. The process of max-pooling or min-pooling involves dividing the image into a number of smaller regions, and then taking the maximum or minimum value from each region. This is done to reduce the resolution of an image, while preserving important features. Max-pooling can be used to reduce image size, reduce the computational cost of a CNN, and improve its accuracy. The max-pooling method is similar to our LFP method that we created, except that they differ in the way of choosing the essential regions of the image. Our method depends on calculating the accuracy of each part of the image based on the total number of training images, while the max-pooling method depends on the dominant color of the region.

PCA stands for “principal component analysis,” and it is a statistical technique and dimensionality reduction method used in data analysis and machine learning [11], [12]. PCA is employed to reduce the number of features (variables) in a dataset while preserving the most critical information or patterns in the data. It achieves this by transforming the original features into a new set of uncorrelated variables called principal components [13]. PCA is commonly used for data preprocessing, feature extraction, and visualization in various fields, including data analysis, image processing, and machine learning [14], [15]. It’s especially valuable for reducing the dimensionality of high-dimensional data and removing multicollinearity between variables, which can improve the performance of machine learning models and help with data exploration [16].

All of these mentioned methods rely heavily on arithmetic matrices and mathematical algebra, while LFP algorithm relies on the AI perceptron algorithm to calculate the accuracy, and based on this accuracy, the main part of the image is selected. Furthermore, existing methods take images or a dataset as input parameter and return an ideal image or dataset, while our method operates on a dataset of images and returns the coordinates of the main sections of the image. One of the advantages of this technique is that we only compute the data once in the training step, while other methods are forced to recalculate at each test for a specific image.

Mentioning the perceptron method, there is a nice article that addresses the same topic as ours, which is to choose the essential pixel of an image in an intelligent way based on the multi-layer perceptron (MLP) algorithm in order to recognize facial expressions. Since we are talking about related articles, it is worth mentioning that this is not the first time we are talking about our method, as we have already dealt with this topic in an article of our own to which you can refer. But this work is very special and complete. It can be seen as a continuation or evolution of the previous article, because we will not only limit ourselves to what is theoretical, but also practical, from information technology (IT) development through publishing in cloud computing to the IT security of the API.

3. METHOD
3.1. LFP algorithm

Firstfull, we need to present the inspiration and source of LFP algorithm. By contemplating the work of the brain, we found that before it begins to learn how to classify objects, it begins to identify the differences between these objects. To prove this opinion, we conducted a practical experiment in which we showed 100 students a picture containing a drawing of the bodies of a man and a woman. Then we asked them to identify the gender of each shape, and their answers were all correct. Then we asked them the question: Which area of the body did you rely on to determine gender? as shown in Figure 1, the result was 43% of the students determined the gender by the head and 57% by the shape of the chest. As a result of this experiment, the students' answers were based on noticing the areas of difference in the two bodies. This is what LFP algorithm does. It identifies areas of difference in images that we can later rely on in machine learning.

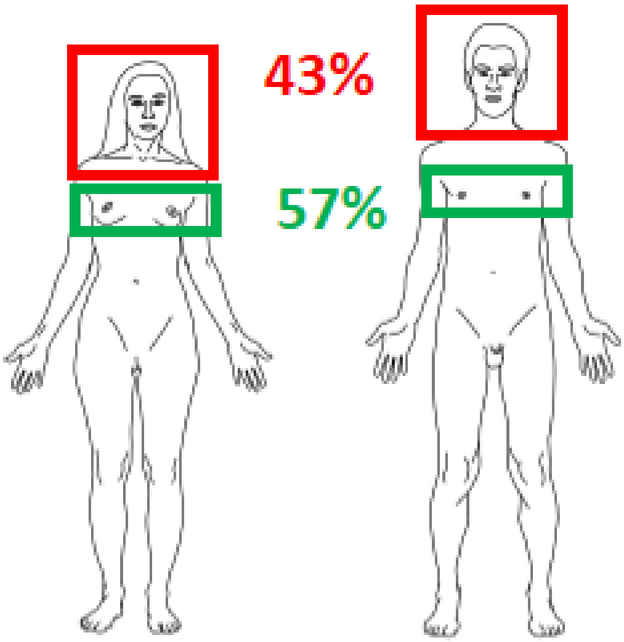

Figure 1. Points of difference between the female and male body

In practice, the LFP algorithm detects core regions using the perceptron algorithm which is a type of linear classifier that can be used to classify data into two classes. Perceptron can be used for binary classification problems, where the output is either a 0 or a 1 [17]. Its main role is to find a hyper plane (W, b) that separates the individuals of a data set $\{(X1,\ y1),\ \ldots,\ (Xn,\ yn)\}$. As we can see in the flowchart in Figure 2, we take a set of images of the same size and then divide each image into squares see Figure 3. We perform perceptron training on each square, then we calculate theire accuracy through the (1), and finally get the coordinates (x, y, width, and height) of the high-precision squares as seen in Figure 2. To sum up, LFP algorithm relies on the accuracy of perceptron training on different squares of images to return their coordinates.

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

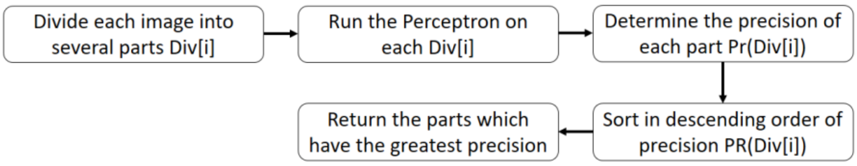

Figure 2. Flowchart of the LFP algorithm

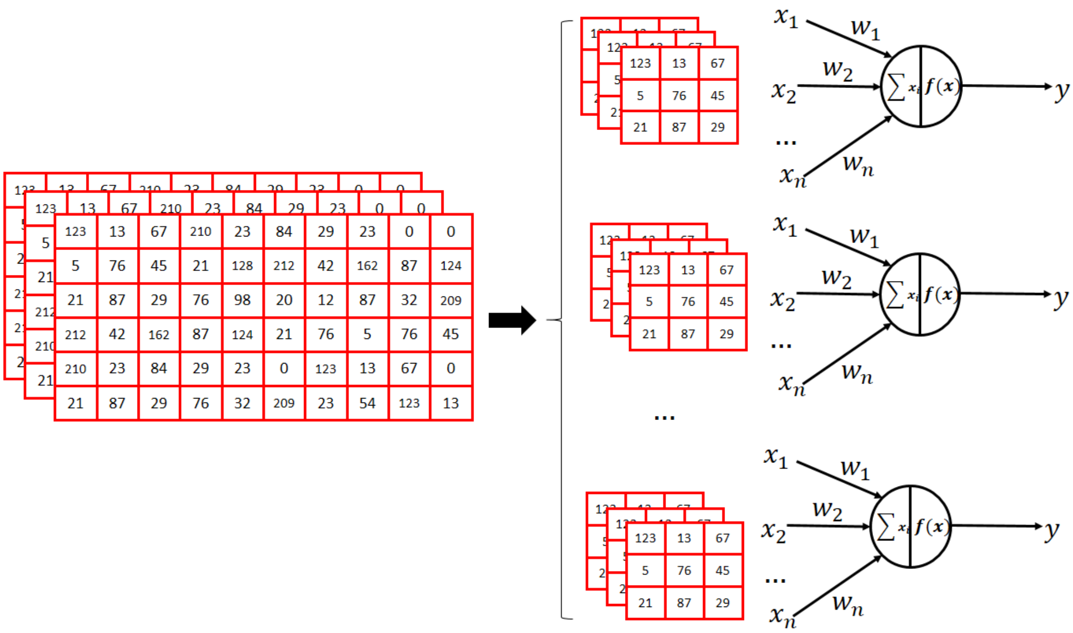

Figure 3. Execution of perceptron on several square

After knowing how LFP algorithm works, we will now move to an important stage, which is how we will share this algorithm with developers community. We created an API that allows developer to use our algorithm easily using hypertext transfer protocol (HTTP) requests. We dissect this point in detail in the following paragraph.

3.2. LFP algorithm API

An API is a set of rules, protocols, and tools that allows different software applications to communicate with each other [18]. It defines the methods and data formats that developers can use to interact with a particular software component, service, or platform, enabling them to access its functionality and data. In the context of APIs, a "request" is a message or command sent by a client (an application or system) to an API endpoint or service, asking for a particular action to be performed or data to be retrieved. Therefor, the client specifies the type of request, such as a GET request to retrieve data, a POST request to create data, a PUT request to update data, or a DELETE request to remove data, among others [19]. In practice, we have created a web service that supports the following request in Figure 4, which gets the file data uniform resource locator (URL) as a parameter. When the server receives this request, it accesses the data to execute the LFP algorithm in order to return the coordinates (x, y, width, and height) of the high-precision squares as a JSON response as we see in Figure 5.

```
POST /api/lfp/coordinates/get
Content-Type : application/json
{
  "url_file" : "https://exemple.com/.../file.txt"
}
```

Figure 4. Request to get coordinates

```
HTTP 202 Accepted
Content-Type : application/json
{
 "coordinates" : [{
          "x" : 12
          "y" : 30
          "width" : 100
           "height" : 100
      },…
   ]
}
```

Figure 5. The response containing the coordinates calculated by the LFP algorithm

After defining the request and response with our API, all developers can easily use the LFP algorithm without seeing the source code. Specifically, developers can integrate our service into a mobile application or website without seeing the source code of the LFP algorithm. On the other hand, we have to protect our service by integrating a protocol of security. This is what we are talking about in the next paragraph.

3.3. API security

There are many security protocols using in API, but we have chosen Oauth2.0 protocol because it's not vulnerable and it is used by many popular websites and applications, such as Google, Facebook, and Twitter. Oauth2.0 an open authorization protocol that allows users to securely access data from other applications without having to share their passwords. To integrate the Oauth2.0 protocol, we need to add an authorization server that grants the client an access token that will be sent on every client request. And then a server with LFP algorithm will work until the access token is validated. Therefore, we modified the API as we see in the following request in Figure 6 by adding the Oauth2.0 protocol attributes like access token.

```
POST /api/lfp/coordinates/get
Content-Type : application/json
{
 "access_token" : "2d8d7bf798d7fnssd",
 "url_file" : "https://exemple.com/.../file.txt"
}
```

Figure 6. The request is protected by an access token

At this point, we are able to protect the LFP algorithm service and manage user permissions using the Oauth2.0 protocol. Therefore, we can collect a lot of information about the use of our service, such as the number of uses, the number of users, and we can also detect errors generated in the LFP algorithm to correct them. Furthermore, we have more control if we want to sell our services through subscriptions.

3.4. Implementation in docker

Now we have reached the paragraph mentioned in the title of this article in which we will see how to build the LFP algorithm and its API in a docker container to deploy it in cloud computing which has many advantages. A docker container is a lightweight, portable, self-contained software package that allows application developers to include all necessary parts, libraries and other dependencies, and ship them all out as one package. Of course, one of the benefits of containers is that developers can ensure that the application runs consistently on any Linux device, regardless of the settings or custom configurations specified on those devices. Docker containers are also extremely flexible, allowing developers to choose the environment and tools that best suit their needs. As we can see in Figure 7, docker infrastructer contains several components including the application container that contains the LFP algorithm. So, to put the source code of our algorithm into a docker container, we have to create a dockerfile that contains all the commands that the user can call on the command line to compile an image. The instructions in the dockerfile will be used by docker to create the image that will be used to create containers, which are runtime instances of the image. Dockerfiles are written in a specific format and contain a set of instructions and arguments that define what the container will look like when it is built and run.

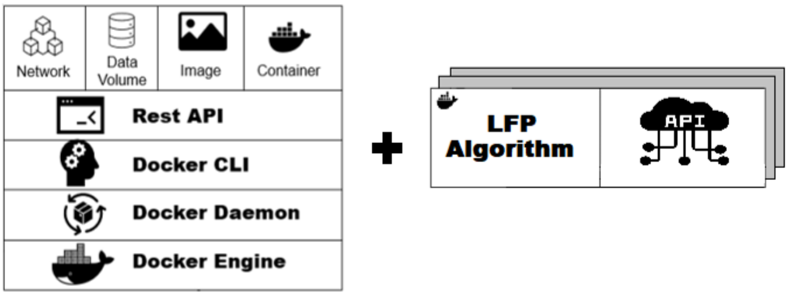

Figure 7. LFP algorithm in the docker container infrastructure

Now that we know how to push our LFP algorithm into docker conatiner, we can now take advantage of the benefits of container cloud computing, such as cost savings, scalability, flexibility, reliability, speed, and efficiency. Cost savings are achieved due to the shared infrastructure and economies of scale associated with cloud computing. Scalability allows for applications to grow or shrink as needed, making them more flexible and cost-efficient. Cloud computing also offers faster deployment times, as applications can be deployed faster than traditional methods. Additionally, cloud computing provides high availability and reliability due to its distributed nature. Finally, cloud computing is highly efficient, as it leverages the power of multiple servers, storage, and networking resources.

4. RESULTS AND DISCUSSION

To examine the performance of our LFP algorithm, we compared it with the methods cited in the related works section: NMF, LBP, and PCA. Process we conducted several training on canned sardine image data containing more than 1,000 individuals see Figure 8. Each time using one of these methods in order to have models. The performance of these models was evaluated based on their classifcation 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 is a measure of how well your model classifies data correctly. It’s often used in classification tasks. Precision is a performance metric that measures the accuracy of the positive predictions made by a model [21], [22]. It is particularly relevant in binary classification problems, where you classify instances into one of two classes: positive (1) or negative (0). Recall, also known as sensitivity or true positive rate, is a performance metric used to evaluate the model’s ability to correctly identify all relevant instances of the positive class [23]. The F1-score is a widely used performance metric in machine learning, particularly for binary classification problems. It combines the precision and recall of a model into a single score and is especially useful when dealing with imbalanced datasets or when there is an uneven cost associated with false positives and false negatives [24]. ROC-AUC is a widely used performance metric for evaluating the quality of a binary classification model [25], [26]. It assesses the model’s ability to discriminate between the positive and negative classes across various classification thresholds.

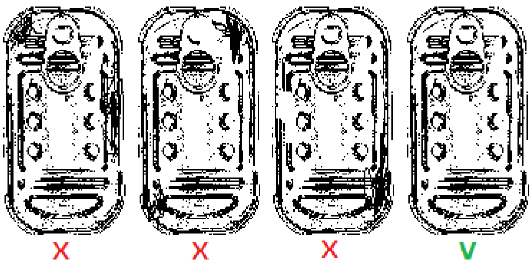

Figure 8. Example of sardine can images

We did five experiments as we see in Figure 9. In the first experiment, we only implemented NN training. For other experiments, we implemented the algorithms: LFP, NMF, LBP, and PCA separately with NN training on the same data. We note in Table 1 following the results of these experiments which demonstrate the strength of the LFP algorithm, which presents high precision. Furthermore, comparing experiments 1 and 2, we observe that we can increase the CA up to 20% if we use the LFP algorithm in the training and testing phases.

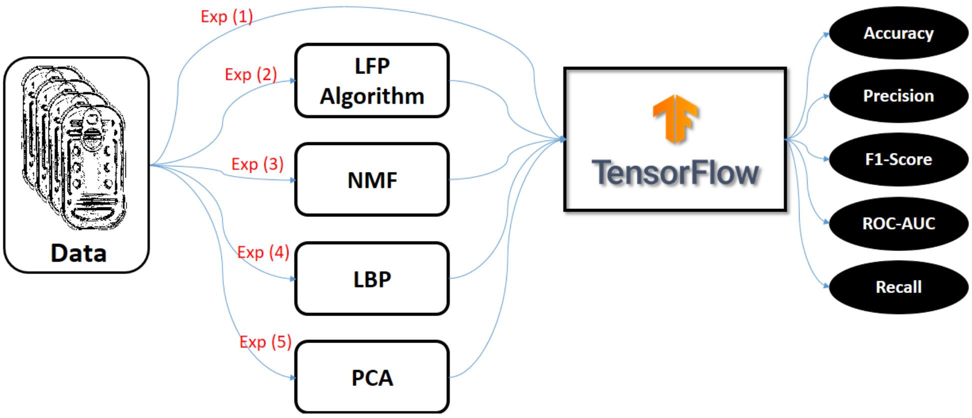

Figure 9. Diagram of the five experiences

Table 1. Results of five experiments

Experiment N°	Algorithm	CA	Precision	Recall	F1-score	ROC-AUC
1	NN	0.706	0.705	0.704	0.704	0.779
2	LFP+NN	0.921	0.921	0.920	0.920	0.934
3	NMF+NN	0.903	0.904	0.903	0.903	0.910
4	LBP+NN	0.870	0.870	0.870	0.870	0.873
5	PCA+NN	0.912	0.912	0.910	0.910	0.916

5. CONCLUSION

In this work, we presented the theoretical principle of the LFP algorithm. Then we talked about the API that allows other developers to use our algorithm without seeing the source code. We also talked about the Oauth security protocol, and how we can integrate it into our API. Next, we introduced how to set up our application in the cloud container by creating a dockerfile. Using sardine can image data, we compare the LFP algorithm and existing methods such as NMF, LBP, and PCA 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 competitors in terms of accuracy (0.921), which means increased CA up to 20%. 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.

ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to the national center for scientific and technical research (CNRST)-Morocco for their generous support and assistance throughout the duration of this research project. Their commitment to advancing scientific knowledge and their dedication to achieving our research goals were instrumental in the success of this work, which is part of the Al-Khwarizmi program.

REFERENCES

[1] Y. Li and A. Ngom, "The non-negative matrix factorization toolbox for biological data mining," *Source Code for Biology and Medicine*, vol. 8, no. 1, p. 10, Dec. 2013, doi: 10.1186/1751-0473-8-10.

[2] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, "Local binary patterns and its application to facial image analysis: a survey," *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 41, no. 6, pp. 765–781, Nov. 2011, doi: 10.1109/TSMCC.2011.2118750.

[3] IEEE Staff, "2008 First International Workshops on Image Processing Theory, Tools and Applications," IEEE, 2008.

[4] Y. Cai, H. Gu, and T. Kenney, "Learning microbial community structures with supervised and unsupervised non-negative matrix factorization," *Microbiome*, vol. 5, no. 1, p. 110, Dec. 2017, doi: 10.1186/s40168-017-0323-1.

[5] P. Paatero, "Least squares formulation of robust non-negative factor analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 37, no. 1, pp. 23–35, May 1997, doi: 10.1016/S0169-7439(96)00044-5.

[6] Y.-X. Wang and Y.-J. Zhang, "Nonnegative matrix factorization: a comprehensive review," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 6, pp. 1336–1353, Jun. 2013, doi: 10.1109/TKDE.2012.51.

[7] B. Yang and S. Chen, "A comparative study on local binary pattern (LBP) based face recognition: LBP histogram versus LBP image," *Neurocomputing*, vol. 120, pp. 365–379, Nov. 2013, doi: 10.1016/j.neucom.2012.10.032.

[8] M. Xi, L. Chen, D. Polajnar, and W. Tong, "Local binary pattern network: a deep learning approach for face recognition," in *2016 IEEE International Conference on Image Processing (ICIP)*, Sep. 2016, vol. 2016-Augus, pp. 3224–3228, doi: 10.1109/ICIP.2016.7532955.

[9] P. Banerjee, A. K. Bhunia, A. Bhattacharyya, P. P. Roy, and S. Murala, "Local neighborhood intensity pattern-a new texture feature descriptor for image retrieval," *Expert Systems with Applications*, vol. 113, pp. 100–115, Dec. 2018, doi: 10.1016/j.eswa.2018.06.044.

[10] Z. Liu, Z. Yuan, C.-L. Ong, and S. Ang, "Influence of data patterns on reader performance at off-track reading," *IEEE Transactions on Magnetics*, vol. 50, no. 11, pp. 1–4, Nov. 2014, doi: 10.1109/TMAG.2014.2329813.

[11] A. Radhika and M. S. Masood, "Effective dimensionality reduction by using soft computing method in data mining techniques," *Soft Computing*, vol. 25, no. 6, pp. 4643–4651, Mar. 2021, doi: 10.1007/s00500-020-05474-7.

[12] M. Sudharsan and G. Thailambal, "Alzheimer's disease prediction using machine learning techniques and principal component analysis (PCA)," *Materials Today: Proceedings*, vol. 81, pp. 182–190, 2023, doi: 10.1016/j.matpr.2021.03.061.

[13] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, p. 20150202, Apr. 2016, doi: 10.1098/rsta.2015.0202.

[14] Y. Guo, Y. Zhou, and Z. Zhang, "Fault diagnosis of multi-channel data by the CNN with the multilinear principal component analysis," *Measurement*, vol. 171, p. 108513, Feb. 2021, doi: 10.1016/j.measurement.2020.108513.

[15] N. Verbeeck, R. M. Caprioli, and R. Van de Plas, "Unsupervised machine learning for exploratory data analysis in imaging mass spectrometry," *Mass Spectrometry Reviews*, vol. 39, no. 3, pp. 245–291, May 2020, doi: 10.1002/mas.21602.

[16] G. Rorbach, O. Unold, and B. M. Konopka, "Distinguishing mirtrons from canonical miRNAs with data exploration and machine learning methods," *Scientific Reports*, vol. 8, no. 1, p. 7560, May 2018, doi: 10.1038/s41598-018-25578-3.

[17] S. -E. Mansour, A. Sakhi, L. Kzaz, O. Tali and A. Sekkaki, "Focal Point of Learning," in *2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, 2021, pp. 0522-0526, doi: 10.1109/UEMCON53757.2021.9666524.

[18] S. P. Ong *et al.*, "The materials application programming interface (API): a simple, flexible and efficient API for materials data based on representational state transfer (REST) principles," *Computational Materials Science*, vol. 97, pp. 209–215, Feb. 2015, doi: 10.1016/j.commatsci.2014.10.037.

[19] C. Pautasso, "RESTful web service composition with BPEL for REST," *Data & Knowledge Engineering*, vol. 68, no. 9, pp. 851–866, Sep. 2009, doi: 10.1016/j.datak.2009.02.016.

[20] Z. A. Khan, M. Adil, N. Javaid, M. N. Saqib, M. Shafiq, and J.-G. Choi, "Electricity theft detection using supervised learning techniques on smart meter data," *Sustainability*, vol. 12, no. 19, p. 8023, Sep. 2020, doi: 10.3390/su12198023.

[21] R. Caruana and A. Niculescu-Mizil, "Data mining in metric space: an empirical analysis of supervised learning performance criteria," *KDD-2004 - Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 69–78, 2004.

[22] C. Liu, M. White, and G. Newell, "Measuring the accuracy of species distribution models: a review," *18th World IMACS Congress and MODSIM 2009 - International Congress on Modelling and Simulation: Interfacing Modelling and Simulation with Mathematical and Computational Sciences, Proceedings*, pp. 4241–4247, 2009.

[23] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," *PLoS ONE*, vol. 10, no. 3, p. e0118432, Mar. 2015, doi: 10.1371/journal.pone.0118432.

[24] F. Thabtah, S. Hammoud, F. Kamalov, and A. Gonsalves, "Data imbalance in classification: experimental evaluation," *Information Sciences*, vol. 513, pp. 429–441, Mar. 2020, doi: 10.1016/j.ins.2019.11.004.

[25] S. Halligan, D. G. Altman, and S. Mallett, "Disadvantages of using the area under the receiver operating characteristic curve to assess imaging tests: a discussion and proposal for an alternative approach," *European Radiology*, vol. 25, no. 4, pp. 932–939, Apr. 2015, doi: 10.1007/s00330-014-3487-0.

[26] J. Huang and C. X. Ling, "Using AUC and accuracy in evaluating learning algorithms," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 3, pp. 299–310, Mar. 2005, doi: 10.1109/TKDE.2005.50.

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.

Oussama Tali completed his higher education in Casablanca, Morocco. Founder of AIDA-Software, and IT project manager at Inspheris Morocco. He can be contacted at email: oussama.tali@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.