A Novel Approach for Parkinson's Disease Detection Based on Voice Classification and Features Selection Techniques

Asmae Ouhmida[1](✉), Abdelhadi Raihani[1], Bouchaib Cherradi[1,2], Oumaima Terrada[1]
[1] Hassan II University of Casablanca, Mohammedia, Morocco
[2] STIE Team, CRMEF Casablanca-Settat, provincial section of El Jadida, El Jadida, Morocco
asmaeouhmida1995@gmail.com

Abstract—Parkinson's disease (PD) is one of the most widespread diseases that, primarily, affects the motor system of the neural central system. In fact, PD is characterized by tremors, stiffness of the muscles, imprecise gait movements, and vocal impairment. An accurate diagnosis of Parkinson's disease is usually based on many neurological, psychological, and physical investigations despite the fact that its main symptoms cannot be easily decorrelated from other diseases. As such, many automatic diagnostic support systems based on Machine Learning approaches have been recently employed to assist the PD patients' assessment. In the current paper, a comparative analysis was performed on machine learning (ML) techniques for PD identification based on voice disorders analysis. These ML methods included the Support Vector Machine (SVM), K-Nearest-Neighbors (KNN), and Decision Tree (DT) algorithms. In addition, two feature selection techniques; mRMR and ReliefF; are used to further improve the performance of the proposed classifiers. The efficiency of the developed model has been evaluated based on accuracy, sensitivity, specificity and AUC metrics, and it is higher than existing approaches. The simulation results show that the KNN algorithm yielded the best classifier performance in term of accuracy and reached an AUC of 98.26%.

Keywords—Parkinson disease, feature selection, machine learning, acoustic dataset

1 Introduction

Parkinson's disease is a neurodegenerative disease that affects the motor system of the human's body central nervous system, preventing the brain from controlling the movements. The symptoms of Parkinson's disease can be classified into motor symptoms and non-motor symptoms[1]. The motor symptoms cover tremors and slowing of movement (hands/legs) and constipation, difficulties in performing daily activities and shuffling of footsteps while walking. Whereas, the non-motor symptoms include

---

[1] https://www.nhs.uk/conditions/parkinsons-disease/symptoms/

loss or decrease of smell sense, speech problems, constipation, fatigue, insomnia and difficulty memorizing.

Research works have shown that the PD gets started before the motor symptom onset and the voice disorder affects roughly 90% of PD patients [1]. Thus, researchers are looking for improved methods to recognize the non-motor symptoms that appear early and could slow down the course of the disease.

Since other diseases present similar symptoms, PD diagnosis using only some qualitative criterions might further complicate the diagnostic process. So, machine learning-based tools have shown recently a remarkable efficiency in assisting in the PD diagnosis process.

Indeed, Machine Learning (ML) techniques play a main role to improve decision-making performance [2]–[10]. In the medical field, ML helps detecting damages in the purpose of reducing the interaction of human intervention. Compared to traditional classification methods, most of ML algorithms improve the accuracy of disease classification or prediction. However, in many medical applications, the execution time or algorithm complexity is a crucial parameter to consider [11]–[16].

In [17], the authors suggested stacking ensemble learning framework composed of two classification model levels, the first included K-nearest neighbor (KNN), support vector machine (SVM), logistic regression (LR) and artificial neural network (ANN) classifiers. They used multi-modal features (diffusion tensor imaging (DTI) and T1 weighted image (T1W1)) and clinical features to classify PD. They reached an accuracy of 96.88%. In [18], the authors established a model that uses fractional amplitude of low frequency fluctuations (fALFF) and regional homogeneity (ReHo) in the purpose of predicting current and future individual's severity. They utilized the Unified Parkinson's Disease Rating Scale (UPDRS) to calculate scores. They trained and optimized various machine learning models; SVM with linear kernel, ElasticNet regression, Gradient Boosting with decision tree kernel and Random Forest with decision tree kernel. They utilized the UPDRS to predict the individual's severity through four targets; baseline, year 1, year 2 and year 4. The values of 79% and 80% were, respectively, the highest positive predictive and the highest negative predictive values achieved by the model. In [19], the authors presented a model based on handwriting samples for early identification of PD patients. The used method is based on multi-scale architectures of CNN (VGGNet, AlexNet, ResNet and GoogleNet) and various PD handwriting datasets; Parkinson's Drawing Dataset, HandPD, PaHaW dataset and NewHandPD. The model achieved the best performance on combining Parkinson's Drawing, HandPD and NewHandPD datasets.

Speech processing is an intensive field of research. The selection of feature extraction technique has a main role in achieving higher performance, being a needful criterion for good speech processing system [20]. In [21], the authors presented Linear Predictive Coefficient (LPC), Mel Frequency Cepstral Coefficient MFCC, Perceptual Linear Prediction (PLP), Relative Spectral Perceptual Linear Prediction (RASTA-PLP) and Wavelet Transform (WT) feature extraction techniques. Strengths and weaknesses of these techniques are also shown in [22], the authors proposed a medical diagnosis support system (MDSS), based on ANN, AdaBoost and DT machine learning algorithms, in the purpose of predicting atherosclerosis.

Therefore, the current contribution attempts to complete previous works related to PD patients' classification. It aims to investigate and evaluate the performance of three classification methods: Support Vector Machine (SVM), K-Nearest-Neighbors (KNN), and Decision Tree (DT). Our approach is based on feature learning for an automatic identification of Parkinson's Disease using voice recording analysis. The extracted and learned features are then used as inputs to the supervised classification methods. The performance of each method is evaluated using some evaluation metrics like accuracy, sensitivity, specificity, and AUC to outcome the best approach.

The outline of this paper is as follows. In section 2, we review some recent related works. Section 3 is dedicated to materials and methods description. In section 4, we present and discuss our finding and compare our results to the state of the art. In section 5, conclude this paper and give some future works perspectives.

2   Related works

Nowadays, in the medical imaging domain, Machine Learning (ML) is frequently used for Parkinson's disease prediction given its high accuracy of early disease detection. In fact, previous studies have shown the implementation of several machine learning algorithms to classify PD patients and isolate the healthy ones, based on acoustic parameters and feature selection algorithms. Therefore, getting a reliable system of PD classification, that achieves higher accuracy is firmly related to the selection of relevant feature extraction and machine learning algorithms.

In [23], the authors evaluate the ability of vocal features in early telediagnosis of PD using ML techniques in two steps. In the first step, they used only the patient's data and the Unified Parkinson's Disease Rating Scale (UPDRS) score to determine the patient group with higher severity of speech impairment. In the second step, they created a new dataset of healthy patients and subjects having less severity of speech impairment. They used Linear and RBF SVM, Extreme Learning Machine (ELM), and KNN classification algorithms. They obtained Mathew's Correlation Coefficient (MCC) of 0.77 and achieved an accuracy of 96.4 %.

Authors in [24] proposed a model based on several feature selection methods such as correlation rates, ROC curves, t-test and, Fisher's Discriminant Ratio to determine whether a patient is suffering from PD or he is not. After selecting the optimal features, they used KNN, SVM, and Discrimination-Function-Based classifiers, and evaluated their performances using accuracy, sensitivity, specificity, and error rate metrics. The best performance was obtained with an accuracy rate of 93.82% using the KNN classifier.

In [25], the authors applied a statistical pooling method to increase the dataset features. Then, they selected the most weighted features using the ReliefF technique. In their proposed model of Parkinson's identification, the highest accuracies obtained were 91.25% and 91.23%, by using SVM and KNN algorithms respectively.

Authors in [26] present a Multiple Feature Evaluation Approach (MFEA) to find the best set of features, based on the use of five classification methods; Decision Tree, Neural Network, Random Forests, Naïve Bayes and Support Vector Machine in the

purpose of enhancing Parkinson's diagnosis. It should be pointed out that the Naïve Bayes machine learning algorithm achieved the highest improvement which is 15.22%.

In [27], the authors applied a wide range of machine learning methods, based on dysphonia symptoms, for PD diagnosis. In fact, they compared the performance of KNN, regression trees (RT), naïve Bayes (NB), linear discriminant analysis (LDA), radial basis function neural network (RBFNN), support vector machine (SVM), and Mahalanobis distance classifier (MDC). The SVM recorded the highest accuracy rate as the best classification result.

In [28], One-against-ALL (OGA) approach was proposed. The authors composed five equal partitions for both, healthy and PD classes. The used classification algorithms were the SVM with medium Gaussian Kernel function, weighted KNN, and Logistic Regression (LR). To assess the experiments of the OGA model, they used confusion matrix, accuracy, and area under the Receiver Operating Characteristic (ROC) curve (AUC). As a result, the weighted KNN produced the best accuracy of 89.46%.

Based on this literature study, we propose a novel hybrid approach for Parkinson's disease detection based on voice classification and advanced features selection techniques. Our contribution consists of adopting features selection algorithms, as a preprocessing phase, for the purpose of ranking and determining the most important ones. Afterward, we applied and selected the most suitable machine learning techniques to classify patients with or without PD. The performance and effectiveness evaluation of each classifier was evaluated using several common metrics.

3 Materials & Methods

3.1 Global overview of the proposed system

In this work, we implemented a methodological approach to produce a more reliable and valid medical diagnosis support system for Parkinson's Disease classification and early detection. Our acoustic dataset [29] has been classified using three various Machine Learning algorithms. And to define the effective features of the problem, FS algorithms have been applied. The flowchart that summarizes the global steps of the adopted system is reported in Figure 1.

3.2 Dataset description

The dataset, used in our study, was created by Max Little [29]. It's composed of 195 sustained vowel phonations obtained from the speech signals of 31 subjects (12 females and 19 males), with ages ranging from 46 to 85 years. 23 of them were diagnosed with Parkinson's disease. Each patient provides an average of 6 phonations recordings. It should be mentioned that there are no missing values. In this PD database, the two classes 'healthy patient' and 'patient with PD' are marked by 0 and 1 respectively in the "status" column.
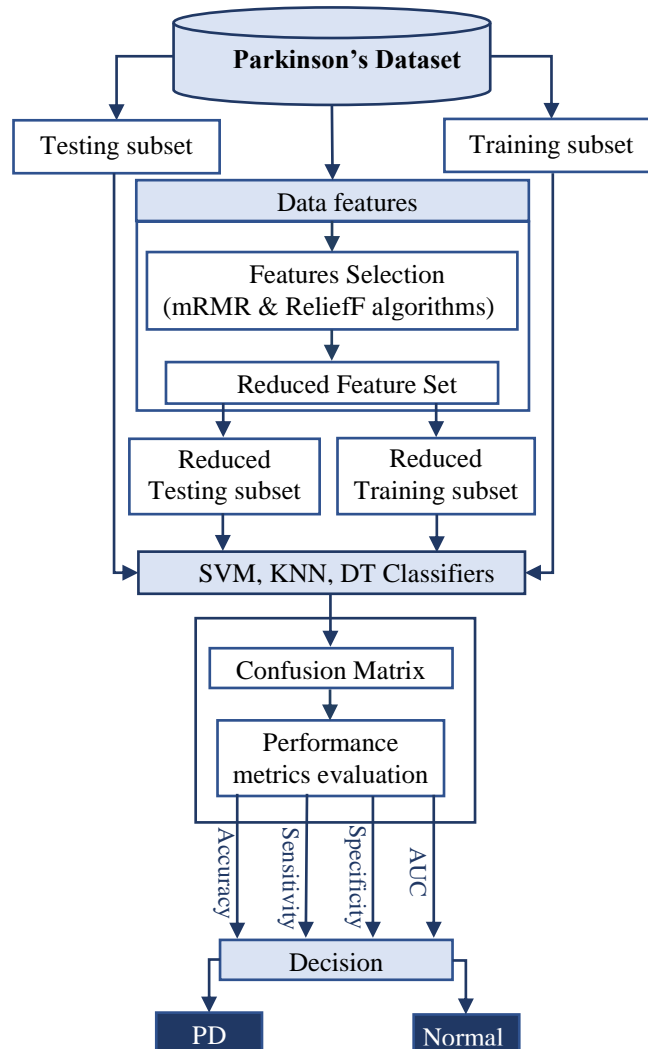
**Fig. 1.** The flowchart of the global steps of the adopted model

Each row represents one of the 195 biomedical voice measurements of individuals (column 'name') and each column represents a particular voice measure. Table 1 shows the 22 voice features used in the experiments. The dataset is divided into healthy phonetics and PD phonetics with a rate of 75% and 25%, respectively, as shown in Figure 2.

3.3 Features selection

Several studies have demonstrated the adequacy and success of feature selection in medical applications. Being a pre-processing technique, it identifies the key feature of the problem. Feature selection helps us to understand the causes of disease, reduces the computational requirements, and prevents degradation in performance [30].

Table 1. Attributes of Parkinson's dataset.

Voice measure	Description
MDVP: F0 (Hz)	Average vocal fundamental frequency
MDVP: Fhi (Hz)	Maximum vocal fundamental frequency
MDVP: Flo (Hz)	Minimum vocal fundamental frequency
MDVP: Jitter (%)	Fundamental frequency perturbation (%)
MDVP:Jitter (Abs)	Absolute jitter in microseconds
MDVP: RAP	Relative Amplitude Perturbation
MDVP: PPQ	Five-point Period Perturbation Quotient
Jitter: DDP	Average absolute difference of differences between cycles, divided by the average period
MDVP: Shimmer	Shimmer Local amplitude perturbation
MDVP: Shimmer (db)	Local amplitude perturbation (decibels)
Shimmer: APQ3	3-point Amplitude Perturbation Quotient
Shimmer: APQ5	5-point Amplitude Perturbation Quotient
MDVP: APQ	11-point Amplitude Perturbation Quotient
Shimmer: DDA	Average absolute difference between the amplitudes of consecutive periods
NHR	Noise-to-Harmonics Ratio
HNR	Harmonics-to-Noise Ratio
RPDE	Recurrence Period Density Entropy
D2	Correlation Dimension
DFA	Detrended Fluctuation Analysis
Spread1	Fundamental frequency variation
Spread2	Fundamental frequency variation
PPE	Pitch period entropy
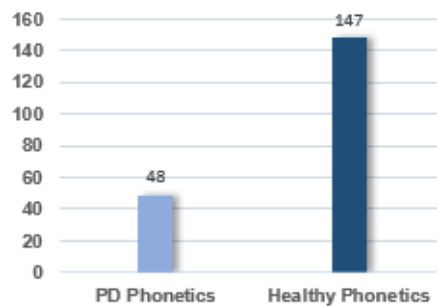
Fig. 2. Class repartition of phonations

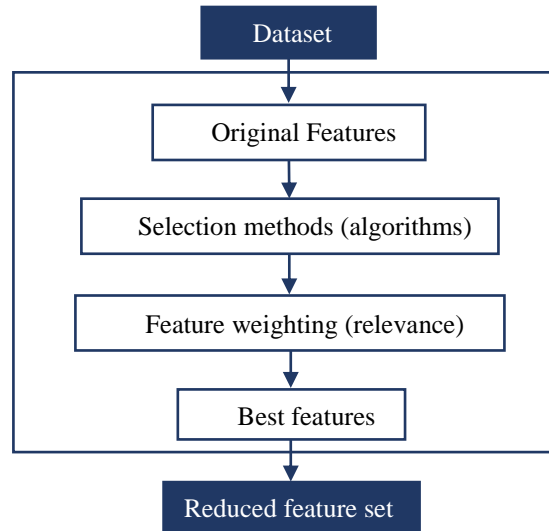Figure 3 depicts the general steps of the feature selection process.

Fig. 3. The general steps of feature selection.

In this study, two feature selection methods are adopted: maximum relevance minimum redundancy (mRMR) and ReliefF.

Maximum relevance minimum redundancy (mRMR). It is a feature selection method that aims to choose the most relevant attributes to the target class (maximum relevance), and provide the feature subset containing less and minimum redundant features as possible (minimum redundancy) [31].

According to mRMR, the optimization condition is given as follows:

$$\underset{x_j \in X - S_{k-1}}{\text{Max}} \left[I(x_j, c) - \frac{1}{k-1} \sum_{x_i \in S_{k-1}} I(x_j, x_i) \right] \qquad (1)$$

Where X is the total set of features, c is the target class, x_i is the ith feature, I is the mutual information, S_{k-1} is the feature set with the selected $k-1$ features in the earlier iterations, $I(x_j, x_i)$ and $I(x_j, c)$ signify the mutual information between features x_i and x_j and the mutual information between feature x_j with class c, respectively.

The mRMR algorithm maximizes the classification performance using a minimal subset of variables by avoiding selecting the redundant features.

ReliefF. It is a feature selection algorithm, based on a statistical method, which detects the features that are relevant to the target concept. The concept of this measure is to penalize the feature which gives different values to a pair of similar examples from the same class and reward the one that gives different values to examples from different classes [32]. At first, the ReliefF algorithm sets all feature weights to 0. A random observation x is selected iteratively. Then, the algorithm finds two nearest neighbors

to x; one from a dissimilar classification NM (the nearest miss) and the other from the same classification NH (the nearest hit). ReliefF updates the weight of the ith feature according to the following formula:

$$w_i = w_i + |x^i - NM^i(x)| - |x^i - NH^i(x)| \tag{2}$$

Where w_i is the weight of the ith feature, $|x^i - NM^i(x)|$ and $|x^i - NH^i(x)|$ denote, in the ith feature, the difference between the observation x^i and its NM and the difference between the observation x^i and its NH, respectively.

The efficiency of this algorithm is expressed by the minimal distance between an observation and its nearest neighbor observation from the same class, and the largest distance to its nearest neighbor observation from the dissimilar class. The feature is considered good when all observations support this rule.

3.4 ML Classifiers

Following the feature selection step, the feature subsets are fed into multiple machine learning algorithms in order to distinguish PD patients from healthy subjects. We have used three classifiers including support vector machine (SVM), decision tree (DT), and k-nearest neighbors (KNN).

Support Vector Machine (SVM). SVM is one of the supervised learning models extensively used for both regression and classification problems. According to the used data, we use the SVM algorithm for binary classification. It has the ability to construct a hyperplane to project the feature to a higher dimensional space, where it separated all observation points of one class from those of the other class [33]. And the one with the largest margin between the two classes is the best hyperplane.

The equation of a hyperplane is given as follows:

$$f(x) = x'\beta + b = 0 \tag{3}$$

Where $\beta \in R^d$ and b is a real number. And the training data is a set of vectors x_j with their categories y_j.

To define the best separating hyperplane, β and b must be specified, in a way that minimizes $\|\beta\|$ for all observation points (x_j, y_j):

$$y_j f(x_j) \geq 1 \tag{4}$$

The support x_j on the boundaries are the support vectors, where $y_j f(x_j) = 1$.
Algorithm 1: Main SVM steps

```
Input: The training and test data.
Output: Test predictions and the calculated performance
evaluation metrics.
Start
Step 1: Train the SVM algorithm and extract the best
model.
```

```
Step 2: Apply the generated model to predict PD pa-
tients from test data.
Step 3: Use confusion matrix values to evaluate the re-
sults of the algorithm test predictions.
Step 4: Display the results.
End
```

Decision Tree (DT). A DT technique is a predictive graphical model that is represented as a hierarchical structure. It primarily contains a starting node (root) and branches (conditions) that conduct to other nodes, until we reach the final decision of the route (leaf). An attribute is tested by the internal node and the attribute values are represented by a branch. Finally, each lead attributes a classification as shown in Figure 4 [34].
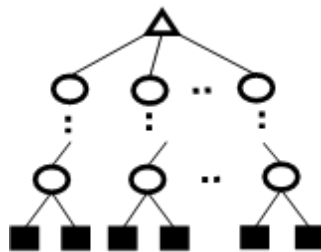

Fig. 4. General structure of decision tree classifier.

Δ, ○ and ■ represent, respectively, root, internal and leaf nodes.

Algorithm 2: Main DT steps

```
Input: The training and test data
Output: Test predictions and the calculated performance
evaluation metrics.
Start
Step 1: Train the DT algorithm and extract the best
trained model.
Step 2: Use the model to predict the response for the
test of the given data.
Step 3: Evaluate the prediction performance of the mod-
el.
Step 4: Display the results
End
```

K-Nearest Neighbors (KNN). KNN is a supervised machine learning algorithm. To apply this classifier, we need to choose the appropriate value for k, the number of nearest neighbors of the data. A new subject is classified according to the plural vote of its neighbors. Then the subject will be assigned to the most common class among its k nearest neighbors, measured by a distance function.

To calculate distance between neighbors, various distance metrics can be used according to the problem, such as Euclidean, Chebychev, Cosine, Mahalanobis, Hamming, Spearman metrics. However, the most appropriate function for our approach is the Cosine distance [35].

The Cosine distance is computed from 1 minus the cosine similarity. And the similarity measures the degree of angle between two vectors.

$$CosD(x, y) = 1 - \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} y_i^2}} \tag{5}$$

x and y are vectors having numerical attributes, where $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_n)$.

Algorithm 3: Main KNN steps

```
Input: k value and the training and data test.
Output: Test predictions and the calculated performance
evaluation metrics.
Start
Step 1: Find the appropriate k value and the best dis-
tance metric.
Step 2: Train the KNN algorithm using the training data
and extract the model.
Step 3: Apply the KNN model to Predict response for the
test data.
Step 4: Calculate evaluation metrics to evaluate the
results of the algorithm.
Step 5: Display the results
End
```

3.5 Performance evaluation metrics

Accuracy, sensitivity, specificity, and AUC are used as evaluation metrics to assess the predictability performances of the classifiers. The employment of these metrics is based on the confusion matrix of a binary classification, which is introduced in Table 2. Where TP is the number of true positive, it means that the prediction is positive while the actual real value is positive; FP, false positive, it means that the prediction is positive while the actual real value is negative; FN, false negative, it means that the prediction is negative while the actual value is positive; and TN, true negative, it means that the prediction is negative while the actual value is negative [36].

Table 2. Confusion Matrix of binary classification.

		Predicted Class	
		Positive	**Negative**
Actual Class	Positive	TP	FN
	Negative	FP	TN

According to the confusion matrix, Accuracy, Sensitivity, and Specificity are computed as:

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \times 100\% \tag{6}$$

$$Sensitivity = \frac{TP}{TP+FN} \times 100\% \tag{7}$$

$$Specificity = \frac{TN}{FP+TN} \times 100\% \tag{8}$$

The receiver operating characteristics (ROC) graph is a probability curve, it's helpful for visualizing the performance of classifiers. To compare classifiers, we may calculate the area under the ROC curve (AUC) [37]. It represents the degree of separability. Its value is always between 0 and 1. A classifier that produces the highest AUC (near to 1) is the better model to distinguish between PD patients and healthy ones. And the AUC of the perfect classifier equals 1. The ROC curve is based on TPR (y-axis) and FPR (x-axis) as shown in Figure 5. And they are defined as follows:

$$TPR = Sensitivity = \frac{TP}{TP+FN} \tag{9}$$

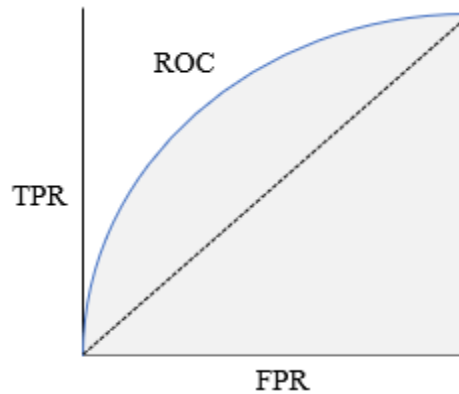$$FPR = 1 - Specificity = \frac{FP}{TN+FP} \tag{10}$$

Fig. 5. The ROC curve (AUC).

4   Results and discussion

The current study is based on a simulation analysis conducted on MATLAB platform. Data were classified using SVM, DT, and KNN. The model is trained according to 75% (147 observations) of the data while 25% (48 observations) has been used for the test for each classifier. To assess the accuracy of the system, the ten-fold cross-validation method had been used in the detection of the PD.

### 4.1 Classification with the whole original dataset

In the SVM classifier case, the linear kernel function was preferred for two-class learning, due to its best performance. While for the KNN classifier, the best accuracy was obtained when k=19, by using the cosine distance.

Based on the evaluation metrics, Table 3 compares the three classifiers. And Figure 6 shows the ROC curves of the three classifiers.

The results indicate that the SVM achieves the best score in terms of accuracy.

Table 3. Comparison of the three classifiers.

Classifiers	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)
SVM	91.67	88.89	100	100
DT	75	75	41.67	37.85
KNN	89	97.22	66.67	94.21
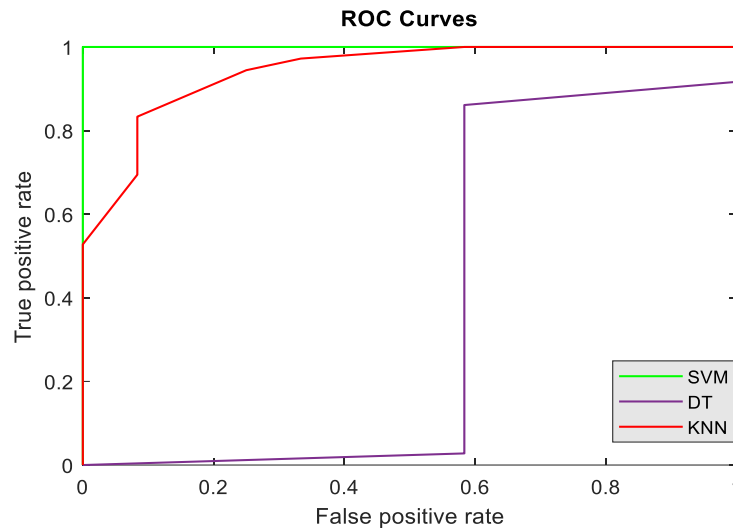
Fig. 6. The ROC curves of the three classifiers.

4.2 Classification after feature selection

Features selection results.

Maximum relevance minimum redundancy (mRMR): We used the mRMR to choose the most important attributes among the original attributes (22 features).

The most important predictor is the PPE. The difference in score between the first feature and the rest of the features is very large. The software is confident in selecting the most important predictor, because of this large drop and it is depicted in Figure 7.

PPE is a robust measure to noisy acoustic environments and sensitive to PD speech changes, as improved by other researches [38], contrary to traditional measures. Then, to define the PPE measure, entropy is calculated based on the probability distribution

of semitone variations. The probability distribution is built, by converting the speech pitch pattern to a logarithmic semitone measure.
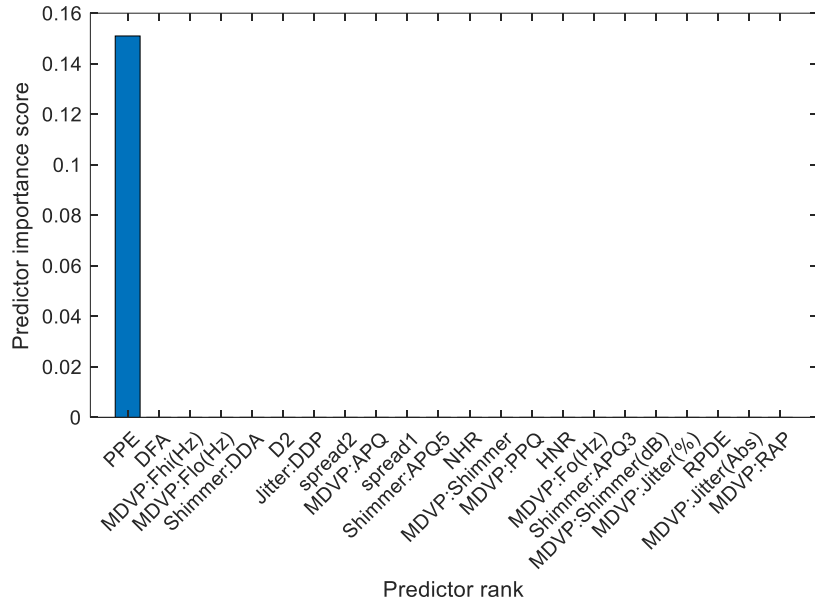

Fig. 7. Feature ranking based on the mRMR algorithm.

ReliefF: The 22 features are represented, in Figure 8, according to their ranking. The bar plot determines the degree of importance weights of all features.

The fourth element is the most important feature and the fifteenth element is the least important one.
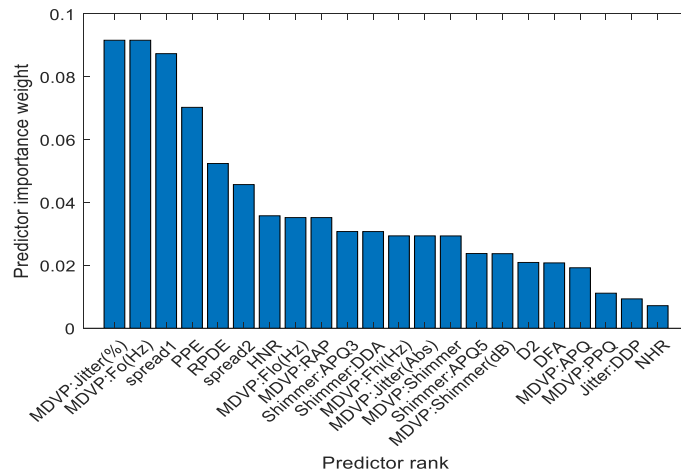

Fig. 8. Feature ranking based on ReliefF algorithm.

After selecting the most informative features from the whole original set, we conduct the experiments in the reduced feature set and apply them to the three classifiers in order to evaluate the performance of the proposed system. Table 4 shows the trends of classification accuracy of SVM, KNN, and DT algorithms over the feature subset. Performance of evaluation results.

Table 4. Performance of evaluation results.

Classifiers	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)
SVM	95.83	94.44	100	100
DT	97.92	97.22	100	98.15
KNN	97.92	97.22	100	98.96

Figure 9 represents the ROC curves of the three classifiers.
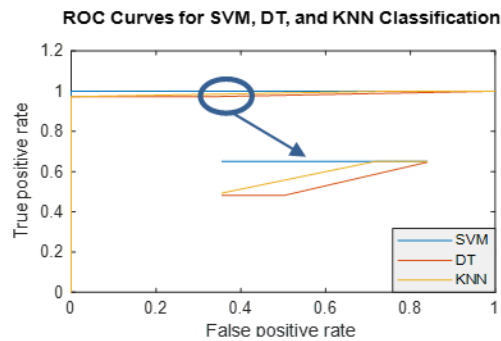
Fig. 9. The ROC curve of the three classifiers with feature selection.

For all the three ML classifiers, the more adequate feature selection is *ReliefF* method.

Regarding the SVM classifier, the best accuracy was achieved with 8 features. The DT classifier achieves the best accuracy when the number of features taken is 6. For the KNN classifier, the best accuracy was obtained when 8 features were adopted, by using the cosine distance and k=23.

The results concerning the three classifiers show that using the feature selection techniques aimed to prevent the degradation of the accuracy and increase the performance. According to the higher accuracy obtained results, The KNN and the DT are, both, the best machine learning classification algorithms.

Table 5 shows the comparison of the proposed model with the state of the art for the classification of PD based on the vocal features. It represents the number of features, the classifiers, the feature selection algorithms, and the performance metrics.

In [41], although the authors obtained high accuracy, the KWELM model used in classification has high complexity and the AABC feature selection algorithm. In [39], the use of the RFE feature selection algorithm gives great accuracy compared to the results of [27], which are based on the whole 22 dataset features.

Substantially, the use of feature selection algorithms [39]–[42]enhances the algorithms' accuracies in the classification process.

Table 5. Comparison of the proposed model with related works.

Author & year	N° of features	Method	FS method	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (%)
Karapinar Senturk (2020) [39]	13 (22)	SVM	RFE	93.84	_	_	_
Depotovic et al.[40]	5 (22)	GPC	ARD	96.92	90	99.29	_
Wang et al.(2017) [41]	14 (22)	KWELM	AABC	98.97	98.62	100	_
Thapa et al.(2020) [42]	13 (22)	TSVM	Correlation Based	93.9	88	99	93.9
Lahmiri and A. Shmuel (2019) [43]	14 (22)	SVM + BO	Wilcoxon	92.13	82.79	95.27	_
Lahmiri et al.(2017) [27]	22	SVM	_	92	95	91	89
Das (2010) [44]	22	NN	_	92.9	_	_	_
Rouzbhani et Daliri (2011) [24]	7 (22)	KNN	T-test, ROC curve FDR	93.82	85	97.74	_
Almeida et al.(2019) [45]	18	KNN	_	94.55	94.55	94.26	87
Yaman et al.(2019) [25]	66 (177)	SVM	ReliefF	91.25	91.25		_
Polat et Nour (2020)[28]	45	weighted kNN	_	89.46	_	_	90.4
Sakar et al (2017) [23]	16	SVM	_	96.4	_	_	_
Proposed	**22**	**KNN**	**mRMR ReliefF**	**97.92**	**97.22**	**100**	**98.26**

5 Conclusion

This study was based on the analysis of three different Machine Learning Algorithms' performance to assist the PD patients' prediction. Those algorithms are applied to the acoustic dataset, using features based on voice signals for both healthy patients and PD patients. The KNN, SVM, and DT techniques are carried out to detect Parkinson's patients and healthy in the early stage. Firstly, feature selection techniques are used to improve the performance of the proposed system, especially the ReliefF algorithm. Afterward, the different classifiers are completed on the resulting dataset. The conducted experiments have shown that the Decision Tree and K-Nearest Neighbours produce a much better accuracy rate of 97.92% compared to the SVM with an accuracy rate of 95.83%. Moreover, sensitivity, specificity, and the ROC curve are used to evaluate the performance of the system. Overall, based on those performance metrics, we did not see a large difference in the performance measures for the different classifiers. However, the AUC suggests that the KNN classifier should be the best technique for Parkinson's disease prediction with very significant and precise results.

Our next objective is to extend this study in different ways; diverse datasets and various machine learning algorithms. In upcoming works, we plan to use different models of deep learning, not only for the classification process but also for regression systems in order to specify the degree of severity of Parkinson's disease.

6   References

[1] C. O. Sakar and O. Kursun, "Telediagnosis of Parkinson's Disease Using Measurements of Dysphonia," *J. Med. Syst.*, vol. 34, no. 4, pp. 591–599, Aug. 2010, https://doi.org/10.1007/s10916-009-9272-y

[2] O. Terrada, B. Cherradi, A. Raihani, and O. Bouattane, "Classification and Prediction of atherosclerosis diseases using machine learning algorithms," in *2019 5th International Conference on Optimization and Applications (ICOA)*, Kenitra, Morocco, Apr. 2019, pp. 1–5. https://doi.org/10.1109/icoa.2019.8727688

[3] O. Terrada, B. Cherradi, A. Raihani, and O. Bouattane, "Atherosclerosis disease prediction using Supervised Machine Learning Techniques," in *2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, Meknes, Morocco, Apr. 2020, pp. 1–5. https://doi.org/10.1109/iraset48871.2020.9092082

[4] O. Terrada, B. Cherradi, A. Raihani, and O. Bouattane, "A fuzzy medical diagnostic support system for cardiovascular diseases diagnosis using risk factors," in *2018 International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, Kenitra, Dec. 2018, pp. 1–6. https://doi.org/10.1109/icecocs.2018.8610649

[5] S. Laghmati, A. Tmiri, and B. Cherradi, "Machine Learning based System for Prediction of Breast Cancer Severity," in *2019 International Conference on Wireless Networks and Mobile Communications (WINCOM)*, Fez, Morocco, Oct. 2019, pp. 1–5. https://doi.org/10.1109/wincom47513.2019.8942575

[6] S. Hamida, O. E. Gannour, B. Cherradi, H. Ouajji, and A. Raihani, "Optimization of Machine Learning Algorithms Hyper-Parameters for Improving the Prediction of Patients Infected with COVID-19," in *2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, Kenitra, Morocco, Dec. 2020, pp. 1–6. https://doi.org/10.1109/icecocs50124.2020.9314373

[7] O. El Gannour, S. Hamida, B. Cherradi, A. Raihani, and H. Moujahid, "Performance Evaluation of Transfer Learning Technique for Automatic Detection of Patients with COVID-19 on X-Ray Images," in *2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, Kenitra, Morocco, Dec. 2020, pp. 1–6. https://doi.org/10.1109/icecocs50124.2020.9314458

[8] O. Daanouni, B. Cherradi, and A. Tmiri, "Self-Attention Mechanism for Diabetic Retinopathy Detection," in *Emerging Trends in ICT for Sustainable Development*, M. Ben Ahmed, S. Mellouli, L. Braganca, B. Anouar Abdelhakim, and K. A. Bernadetta, Eds. Cham: Springer International Publishing, 2021, pp. 79–88. https://doi.org/10.1007/978-3-030-53440-0_10

[9] V. Kate and P. Shukla, "Breast Cancer Image Multi-Classification Using Random Patch Aggregation and Depth-Wise Convolution based Deep-Net Model," *Int. J. Online Biomed. Eng. IJOE*, vol. 17, no. 01, p. 83, Jan. 2021, https://doi.org/10.3991/ijoe.v17i01.18513

[10] Y. Chang and F. Abu-Amara, "An Efficient Hybrid Classifier for Cancer Detection," *Int. J. Online Biomed. Eng. IJOE*, vol. 17, no. 03, p. 76, Mar. 2021, https://doi.org/10.3991/ijoe.v17i03.19683

[11] N. A. Ali, A. E. abbassi, and B. Cherradi, "The performances of iterative type-2 fuzzy C-mean on GPU for image segmentation," *J. Supercomput.*, Jun. 2021, https://doi.org/10.1007/s11227-021-03928-9

[12] N. A. Ali, B. Cherradi, A. El Abbassi, O. Bouattane, and M. Youssfi, "New parallel hybrid implementation of bias correction fuzzy C-means algorithm," in *2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, Fez, May 2017, pp. 1–6. https://doi.org/10.1109/atsip.2017.8075519

[13] N. A. Ali, B. Cherradi, A. El Abbassi, O. Bouattane, and M. Youssfi, "Parallel Implementation and Performance Evaluation of some Supervised Clustering Algorithms for MRI Images Segmentation," in *Proceedings of the 4th International Conference on Big Data and Internet of Things*, Rabat Morocco, Oct. 2019, pp. 1–7. https://doi.org/10.1145/3372938.3373007

[14] N. Aitali, B. Cherradi, A. El Abbassi, O. Bouattane, and M. Youssfi, "GPU based implementation of spatial fuzzy c-means algorithm for image segmentation," in *2016 4th IEEE International Colloquium on Information Science and Technology (CiSt)*, Tangier, Morocco, Oct. 2016, pp. 460–464. https://doi.org/10.1109/cist.2016.7805092

[15] N. Ait Ali, B. Cherradi, A. El Abbassi, O. Bouattane, and M. Youssfi, "GPU fuzzy c-means algorithm implementations: performance analysis on medical image segmentation," *Multimed. Tools Appl.*, vol. 77, no. 16, pp. 21221–21243, Aug. 2018, https://doi.org/10.1007/s11042-017-5589-6

[16] N. Aitali, B. Cherradi, O. Bouattane, M. Youssfi, and A. Raihani, "New fine-grained clustering algorithm on GPU architecture for bias field correction and MRI image segmentation," in *2015 27th International Conference on Microelectronics (ICM)*, Casablanca, Morocco, Dec. 2015, pp. 118–121. https://doi.org/10.1109/icm.2015.7438002

[17] Y. Yang, L. Wei, Y. Hu, Y. Wu, L. Hu, and S. Nie, "Classification of Parkinson's disease based on multi-modal features and stacking ensemble learning," *J. Neurosci. Methods*, vol. 350, p. 109019, Feb. 2021, https://doi.org/10.1016/j.jneumeth.2020.109019

[18] K. P. Nguyen *et al.*, "Predicting Parkinson's disease trajectory using clinical and neuroimaging baseline measures," *Parkinsonism Relat. Disord.*, vol. 85, pp. 44–51, Apr. 2021, https://doi.org/10.1016/j.parkreldis.2021.02.026

[19] I. Kamran, S. Naz, I. Razzak, and M. Imran, "Handwriting dynamics assessment using deep neural network for early identification of Parkinson's disease," *Future Gener. Comput. Syst.*, vol. 117, pp. 234–244, Apr. 2021, https://doi.org/10.1016/j.future.2020.11.020

[20] D. Purbohadi, S. Afriani, N. Rachmanio, and A. Dewi, "Developing Medical Virtual Teaching Assistant Based on Speech Recognition Technology," *Int. J. Online Biomed. Eng. IJOE*, vol. 17, no. 04, Art. no. 04, Apr. 2021. https://doi.org/10.3991/ijoe.v17i04.21343

[21] M. A. Mazumder and R. A. Salam, "Feature Extraction Techniques for Speech Processing: A Review," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 1.3, pp. 285–292, Aug. 2019, https://doi.org/10.30534/ijatcse/2019/5481.32019

[22] O. Terrada, S. Hamida, B. Cherradi, A. Raihani, and O. Bouattane, "Supervised Machine Learning Based Medical Diagnosis Support System for Prediction of Patients with Heart Disease," *Adv. Sci. Technol. Eng. Syst. J.*, vol. 5, no. 5, pp. 269–277, 2020, https://doi.org/10.25046/aj050533

[23] B. Sakar, G. Serbes, and C. O. Sakar, "Analyzing the effectiveness of vocal features in early telediagnosis of Parkinson's disease," *PLOS ONE*, vol. 12, p. e0182428, Aug. 2017, https://doi.org/10.1371/journal.pone.0182428

[24] H. Karimi Rouzbahani and M. R. Daliri, "Diagnosis of Parkinson's Disease in Human Using Voice Signals," *Basic Clin. Neurosci.*, vol. 2, no. 3, pp. 12–20, Apr. 2011.

[25] O. Yaman, F. Ertam, and T. Tuncer, "Automated Parkinson's disease recognition based on statistical pooling method using acoustic features," *Med. Hypotheses*, vol. 135, p. 109483, Feb. 2020, https://doi.org/10.1016/j.mehy.2019.109483

[26] S. A. Mostafa *et al.*, "Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson's disease," *Cogn. Syst. Res.*, vol. 54, pp. 90–99, May 2019, https://doi.org/10.1016/j.cogsys.2018.12.004

[27] S. Lahmiri, D. Dawson, and A. Shmuel, "Performance of machine learning methods in diagnosing Parkinson's disease based on dysphonia measures," *Biomed. Eng. Lett.*, vol. 8, pp. 1–11, Oct. 2017, https://doi.org/10.1007/s13534-017-0051-2

[28] K. Polat and M. Nour, "Parkinson Disease Classification using one against all based data sampling with the acoustic features from the speech signals," *Med. Hypotheses*, vol. 140, p. 109678, Mar. 2020, https://doi.org/10.1016/j.mehy.2020.109678

[29] M. Little, P. Mcsharry, S. Roberts, D. Costello, and I. Moroz, "Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection," *Biomed. Eng. Online*, vol. 6, p. 23, Feb. 2007, https://doi.org/10.1186/1475-925x-6-23

[30] B. Remeseiro and V. Bolon-Canedo, "A review of feature selection methods in medical applications," *Comput. Biol. Med.*, vol. 112, p. 103375, Sep. 2019, https://doi.org/10.1016/j.compbiomed.2019.103375

[31] Hanchuan Peng, Fuhui Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005, https://doi.org/10.1109/tpami.2005.159

[32] K. Kira and L. A. Rendell, "A Practical Approach to Feature Selection," in *Machine Learning Proceedings 1992*, D. Sleeman and P. Edwards, Eds. San Francisco (CA): Morgan Kaufmann, 1992, pp. 249–256. https://doi.org/10.1016/b978-1-55860-247-2.50037-1

[33] S. Misra, H. Li, and J. He, *Machine learning for subsurface characterization*, 1st ed. San Diego: Elsevier, 2019.

[34] A. Mashat, M. Fouad, P. Yu, and T. Gharib, "A Decision Tree Classification Model for University Admission System," *J. Adv. Comput. Sci. Appl.*, vol. 3, Oct. 2012, https://doi.org/10.14569/IJACSA.2012.031003

[35] H. A. Abu Alfeilat *et al.*, "Effects of Distance Measure Choice on K-Nearest Neighbor Classifier Performance: A Review," *Big Data*, vol. 7, no. 4, pp. 221–248, Dec. 2019, https://doi.org/10.1089/big.2018.0175

[36] O. Asmae, R. Abdelhadi, C. Bouchaib, S. Sara, and K. Tajeddine, "Parkinson's Disease Identification using KNN and ANN Algorithms based on Voice Disorder," in *2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, Apr. 2020, pp. 1–6. https://doi.org/10.1109/iraset48871.2020.9092228

[37] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006, https://doi.org/10.1016/j.patrec.2005.10.010

[38] M. Little, P. McSharry, E. Hunter, J. Spielman, and L. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," *Nat. Preced.*, pp. 1–1, Sep. 2008, https://doi.org/10.1038/npre.2008.2298.1

[39] Z. Karapinar Senturk, "Early diagnosis of Parkinson's disease using machine learning algorithms," *Med. Hypotheses*, vol. 138, p. 109603, May 2020, https://doi.org/10.1016/j.mehy.2020.109603

[40] V. Despotovic, T. Skovranek, and C. Schommer, "Speech Based Estimation of Parkinson's Disease Using Gaussian Processes and Automatic Relevance Determination," *Neurocomputing*, vol. 401, pp. 173–181, Aug. 2020, https://doi.org/10.1016/j.neucom.2020.03.058

[41] Y. Wang, A.-N. Wang, Q. Ai, and H.-J. Sun, "An adaptive kernel-based weighted extreme learning machine approach for effective detection of Parkinson's disease," *Biomed. Signal Process. Control*, vol. 38, pp. 400–410, Sep. 2017, https://doi.org/10.1016/j.bspc.2017.06.015

[42] S. Thapa, S. Adhikari, A. Ghimire, and A. Aditya, Feature Selection Based Twin-Support Vector Machine for the Diagnosis of Parkinson's Disease. 2020. https://doi.org/10.1109/r10-htc49770.2020.9356984

[43] S. Lahmiri and A. Shmuel, "Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine," *Biomed. Signal Process. Control*, vol. 49, pp. 427–433, Mar. 2019, https://doi.org/10.1016/j.bspc.2018.08.029

[44] R. Das, "A comparison of multiple classification methods for diagnosis of Parkinson disease," *Expert Syst. Appl.*, vol. 37, no. 2, pp. 1568–1572, Mar. 2010, https://doi.org/10.1016/j.eswa.2009.06.040.

[45] J. S. Almeida *et al.*, "Detecting Parkinson's disease with sustained phonation and speech signals using machine learning techniques," *Pattern Recognit. Lett.*, vol. 125, pp. 55–62, Jul. 2019, https://doi.org/10.1016/j.patrec.2019.04.005

7   Authors

Asmae Ouhmida received her Master's degree in 2018 in distributed information systems at ENSET of Mohammedia, Hassan II University of Casablanca, Morocco. Currently, she is a Ph.D. student at the SSDIA Laboratory of ENSET Mohammedia of the same University. Her research interest is Biomedical signal processing and Artificial Intelligence (Email: asmaeouhmida1995@gmail.com).

Abdelhadi Raihani Professor Abdelhadi RAIHANI was appointed as a professor in Electronics Engineering at Hassan II University of Casablanca, ENSET Institute, Mohammedia Morocco since 1991. He received the B.S. degree in Electronics in 1987 and the M.S. degree in Applied Electronics in 1991 from the ENSET Institute. He has his DEA diploma in information processing from the Ben M'sik University of Casablanca in 1994. He received the Ph.D. in Parallel Architectures Application and image processing from the Ain Chock University of Casablanca in 1998. His current research interests are in the medical image processing areas, electrical engineering fields, particularly in renewable energy, energy management systems, power and energy systems control. He has more than 46 journal publications and over 60 international conference papers. He is an active member in national research programs with IRESEN under the grant "Green INNO Project / UPISREE". He supervised several Ph.D. and Engineers students in these topics (Email: raihani@enset-media.ac.ma).

Bouchaib Cherradi was born in 1970 at El Jadida, Morocco. He received the B.S. degree in Electronics in 1990 and the M.S. degree in Applied Electronics in 1994 from the ENSET Institute of Mohammedia, Morocco. He received the DESA diploma in Instrumentation of Measure and Control (IMC) from Chouaib Doukkali University at El Jadida in 2004. He received his Ph.D. in Electronics and Image processing from the Faculty of Science and Technology, Mohammedia. His research interests include applications of Massively Parallel Architectures, Cluster Analysis, Pattern Recognition, Image Processing, Fuzzy Logic systems, Artificial Intelligence, Machine Learning and Deep Learning in Medical and educational data analysis. Dr. Cherradi

works actually as an Associate professor in CRMEF-El Jadida. In addition, he is associate researcher member of Signals, Distributed Systems and Artificial Intelligence (SSDIA) Laboratory in ENSET Mohammedia, Hassan II University of casablanca (UH2C), and LaROSERI Laboratory on leave from the Faculty of Science El Jadida (Chouaib Doukali University), Morocco. He is a supervisor of several PhD students (Email: bouchaib.cherradi@gmail.com).

Oumaima Terrada was born in Safi, Morocco in 1989. She has her Ph.D. degree in 2020 in computer science and artificial intelligence from Faculty of sciences and technics Mohammedia at Hassan II University of Casablanca. She is associate member of Signals, Distributed Systems and Artificial Intelligence (SSDIA) Laboratory in ENSET Mohammedia. She received her bachelor's degree in Mathematical Sciences, Computer Science and Applications (SMIA) from Poly Disciplinary Faculty of Safi in 2012 at Cadi AYYAD University, Morocco. She holds a Master's degree in Computer Science, Computer Graphics and Imaging (M3I) from Faculty of Sciences Dhar Mahraz in 2014 at Sidi Mohamed Ben Abdellah University Fez, Morocco. During her studies, she worked in medical diagnosis systems using several machine learning algorithms focus on cardiovascular diseases, which resulted in published papers and international conferences presentations (Email: oumaima.terrada@gmail.com).