Article citation info:
Boualoulou N, Mounia M, Nsiri B, Behoussine Drissi T. A novel Parkinson's disease detection algorithm combined EMD, BFCC, and SVM classifier. Diagnostyka. 2023;24(4):2023404. https://doi.org/10.29354/diag/171712.

DIAGNOSTYKA, 2023, Vol. 24, No. 4

e-ISSN 2449-5220
DOI: 10.29354/diag/171712

A NOVEL PARKINSON'S DISEASE DETECTION ALGORITHM COMBINED EMD, BFCC, AND SVM CLASSIFIER

Nouhaila BOUALOULOU [1, 3*], Miyara MOUNIA[2], Benayad NSIRI [3], Taoufiq BELHOUSSINE DRISSI [1]

[1] Laboratory Electrical and Industrial Engineering, Information Processing, Informatics, and Logistics (GEITIIL). Faculty of Science Ain Chock. University Hassan II, Casablanca, Morocco
[2] Computer Science and Systems Laboratory (LIS), Faculty of Science Ain Chock. University Hassan II, Casablanca, Morocco
[3] Research Center STIS, M2CS, National Higher School of Arts and Craft, Rabat (ENSAM). Mohammed V University in Rabat, Morocco
[*] Corresponding author, boualoulounouha@gmail.com

Abstract

Identifying and assessing Parkinson's disease in its early stages is critical to effectively monitoring the disease's progression. Methodologies based on machine learning enhanced speech analysis are gaining popularity as the potential of this field is revealed. Acoustic features, in particular, are used in a variety of algorithms for machine learning and could serve as indicators of the general health of subjects' voices. In this research paper, a novel method is introduced for the automated detection of Parkinson's disease through speech signal analysis, a support vector machines classifier (SVM) and an Artificial Neural Network (ANN) are used to evaluate and classify the data based on two acoustic features: Bark Frequency Cepstral Coefficients (BFCC) and Mel Frequency Cepstral Coefficients (MFCC). These features are extracted from the denoised signals using Empirical Mode Decomposition (EMD). The most relevant results obtained for a dataset of 38 participants are by the BFCC coefficients with an accuracy up to 92.10%. These results confirm that EMD-BFCC-SVM method can contribute to the detection of Parkinson's disease.

Keywords: EMD, BFCC, MFCC, SVM, Parkinson’s disease

List of Symbols/Acronyms

ANN – Artificial Neural Network;
BFCC – Bark Frequency Cepstral Coefficients;
DCT – Discrete Cosine Transform;
DFT- Discrete Fourier Transform;
DWT – Discrete Wavelet Transform;
EMD – Empirical Mode Decomposition;
FFT – Fast Fourier Transform;
HHT- Hilbert-Huang Transform;
IEDCC- Instantaneous Energy Deviation Cepstral Coefficient;
IMF – Intrinsic Mode Functions;
MFCC – Mel Frequency Cepstral Coefficients;
PD – Parkinson’s Disease;
SVM – Support Vector Machines;
TQWT – Tunable Q factor Wavelet Transform;

1. INTRODUCTION

Following Alzheimer's disease, Parkinson's disease (PD) is a neurodegenerative illness that gets progressively worse over time. Its prevalence rises with age; with 1% of individuals, over 60 being touched and up to 4% of individuals over 80 (1). The main symptoms of PD are akinesia (slow initiation of movement), rigidity and resting tremor, and decreased spontaneous mobility. Unfortunately, these symptoms are not only specific to this disease, which delays the diagnosis. Furthermore, the early manifestation of Parkinson's disease often includes noticeable speech impairment as a prominent symptom, this allows many researchers to focus on voice processing by using different methods of advanced signal processing, and extraction of acoustic coefficients, including various machine learning algorithm to achieve an effective analysis of Parkinson's disease. The extraction of features in various time-frequency fields has received considerable focus. For instance, to differentiate individuals with Parkinson's disease (PWP) from healthy controls (HC), Mel frequency cepstral coefficients (MFCC) were derived from the audio signals using Mel domain analysis. This was achieved by employing a triangular-shaped stacked filter bank that combined cepstrum analysis and spectral area splitting. The goal was to derive the coefficients of Mel cepstrum as distinctive

Received 2023-06-02; Accepted 2023-08-31; Available online 2023-10-09

characteristics of the denoised signals, utilizing various types of wavelets., with 87% accuracy (2). There is also delta delta MFCC features extracted from the approximation a3 (3) and IMFCC features (4), which are analogous to the MFCC extraction technique. Karan et al. proposed the use of IEDCC, which is focused on the Hilbert spectrum and aims to capture the energy variation in cepstral domain. eliminating the need for additional optimization (5). Karan et al. employed time-frequency properties to effectively represent the discontinuity in speech signals. They extracted time-frequency features in combination with non-negative matrix decomposition (NMF), resulting in a vowel accuracy of 92% (6). Zhang et al. proposed utilizing the energy direction properties obtained from empirical mode decomposition (EDF-EMD). They emphasized that the high-frequency component of the speech signal contains supplementary information regarding Parkinson's disease (PD) (7). To capture the variation of derivatives in various time-frequency fields, Zhang et al. suggested local gradient statistical features relying on Mel transform (SFLG-Mel) and the Fourier Transform (SFLG-FT), which yielded significantly higher results than the traditional features (8). Zayrit et al. proposed a novel combination of genetic algorithm and SVM classifier using different features MFCC, LPC, energy, ZCR and wavelet Shannon entropy (9). Based on the analysis of the literature mentioned above, incorporating diverse information from the time-frequency domain of speech signals, along with structural features, has demonstrated impressive performance in the classification of Parkinson's disease. These approaches effectively capture the underlying pathological characteristics associated with PD.

This document aims to suggest a PD automatic recognition system using voice analysis. This process relies on the coupling of EMD, enabling signal projection in an adaptive base, and a new algorithm BFCC coefficients, allowing a time-frequency analysis without being constrained by the stationarity and linearity assumptions. These EMD-BFCC techniques were compared with the old MFCC algorithm and applied to a database of 38 recordings, 20 for Parkinson's patients and 18 for healthy patients. This database is used to train and test the extracted features using the SVM and ANN classifiers.

The article structure is as follows: The second part focuses on the dataset used in our study, the third part presents the mathematical tools; the fourth part is devoted to the methodology showing the different steps and presenting the results obtained. Lastly, the article ends with a conclusion and future directions that summarizes these results and reveals the future goals.

2. DATASET

Database1: The dataset suggested by Sakar was gathered by the Department of Neurology at Cerrahpasa Medical School, Istanbul University. The data collection process involved using a TRUST MC-1500 microphone positioned 10 cm away from the mouth, with a sampling rate of 44.1 kHz. The study comprised 18 healthy individuals (10 men and 8 women) and 20 patients diagnosed with PD (12 men and 8 women). The age range of the patients was between 43 and 77 years, with a mean age of 64.86 and a standard deviation of 8.97. The healthy individuals ranged in age from 45 to 83 years, with a mean age of 62.55 and a standard deviation of 10.79. All subjects were invited to state the long-term vowel /a/ in Turkish three times during the collection of this dataset (10). The voice registration was done in stereo and stored in WAV format. The simulation of this method will be performed through a PC hp windows 10 via the software MATLAB R2022b.

Database 2: The PC-GITA Spanish dataset consists of speech recordings from 50 individuals diagnosed with Parkinson's disease and 50 healthy individuals. The recordings were captured at a sampling frequency of 16 kHz and with a resolution of 16 bits. This dataset serves as an independent testing resource for evaluating the performance of models developed using the PC-GITA dataset.

3. METHODOLOGY

This research aims to employ Empirical Mode Decomposition (EMD) for extracting the Bark Frequency Cepstral Coefficient (BFCC) feature from denoised signals. The primary goal is to differentiate between individuals with PD and those who are in a healthy state. Figure 1 illustrates the two crucial phases of the proposed method for diagnosing Parkinson's disease. feature extraction and classification. Each stage is described in detail below.

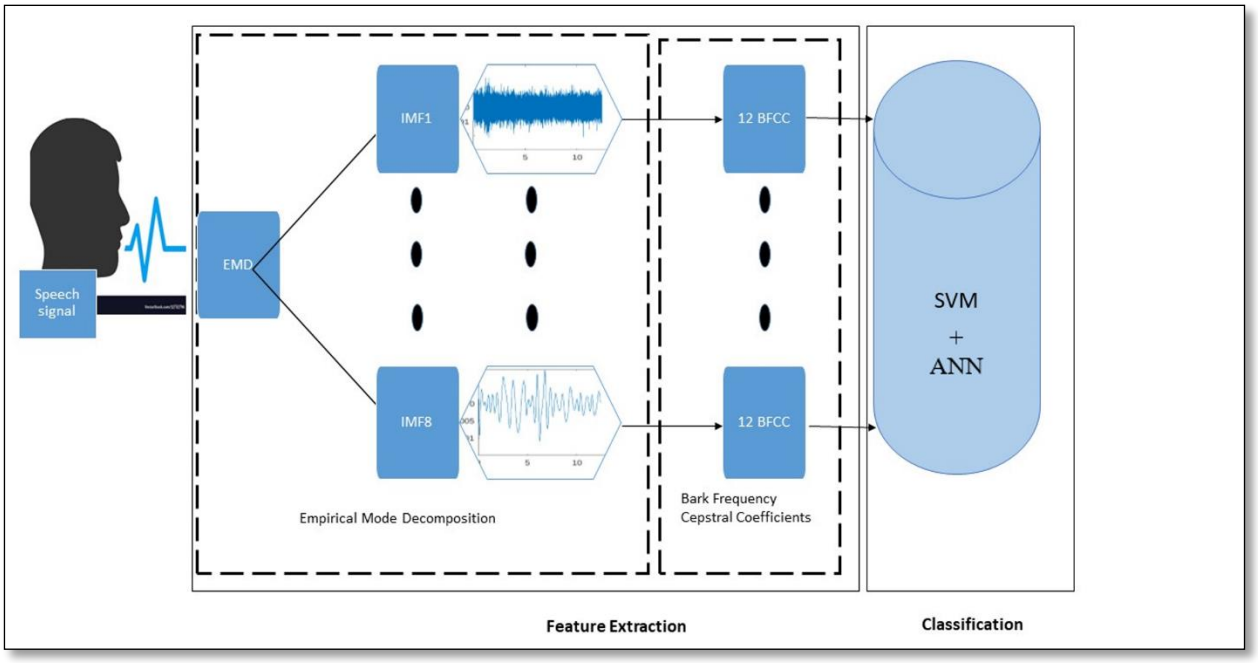

Fig. 1. The principle of Parkinson's disease categorization.

3.1 Feature extraction method

The feature extraction process is divided into two parts, as described in Figure 1. To obtain each IMF, the speech signals were first processed employing the EMD technique. BFCC was then calculated from each IMF by following the BFCC coefficient extraction steps listed below and then compared these BFCC coefficients with the Mel Frequency Cepstral Coefficients (MFCC).

3.1.1 Mel Frequency Cepstral Coefficients

MFCC are acoustic parameters first introduced in 1980 by Davis and Mermelstein for automatic speech recognition (11). Each of the stages necessary to obtain a characteristic vector of MFCC coefficients is described below:

Pre-emphasis refers to a filter that boosts high frequencies; the following formula presents the pre-emphasis filter's transfer function.

$$H(z) = 1 - kz^{-1} \tag{1}$$

The pre-emphasis coefficient k has been set to 0.97.

Segmentation: Because the speech signal is non-stationary, it is subdivided into segments of N speech samples ranging from 10 to 30 before the parameters are extracted. This step produces a quasi-stationary signal for each segment. By multiplying each segment by a Hamming window, the discontinuity at the end of the segments can be reduced.

Windowing: Since the Hamming window displays each frame to reduce the discontinuity at the end of the frames, the expression of the Hamming window is defined by the next equation:

$$w(n) = 0{,}54 - 0{,}46 \times \cos\left(\frac{2\pi n}{N-1}\right) \tag{2}$$

During the FFT step, each frame undergoes a conversion from the time domain to the frequency domain. This process involves efficiently computing the spectral coefficients by utilizing the DFT. It is determined by the equation below:

$$S_n = \sum_{k=0}^{N-1} S_k \, e^{-j2\pi \frac{kn}{N}} \tag{3}$$

In the Mel filter bank stage, the output of the FFT is multiplied by a triangular filter bank that is spaced based on the Mel scale, using the following formula:

$$\text{Mel} = 2595 \log_{10}\left(1 + \frac{f}{700}\right) \tag{4}$$

Logarithm and Discrete Cosine Transform (DCT): Due to overlapping filters, filter bank energies are correlated, so DCT is calculated to decorrelate the filter bank energies. The cepstral coefficient Ci is obtained directly from the logarithm of the filter bank energies mj (M=20 in this study).

$$C_i = \sqrt{\frac{2}{N}} \sum_{j=1}^{M} m_j \cos\left(\frac{\pi i}{N}(j - 0.5)\right) \tag{5}$$

In the liftering stage, the high-order cepstral coefficients are rescaled to ensure that they have similar magnitudes. This is done using the following formula. In this particular study, a liftering value of L = 22 was employed.

$$C_n' = \left(1 + \frac{L}{2} \times \sin\left(\frac{\pi n}{L}\right)\right) \times C_n \tag{6}$$

3.1.2 Bark Frequency Cepstral Coefficient (BFCC)

BFCC is realized by pre-emphasizing and windowing the entry signals. As shown in Equation 3, the fast Fourier transform of the framed and windowed input signals is computed and then transformed to the bark scale. These vectors are then

logarithmized. Finally, the DCT is used followed by liftring. The first 12 cepstral coefficients are then selected. With the exception of equation (4), which is substituted by equation (7), all formulas for calculating the BFCC are the same as those for the MFCC, from equation (1) to equation (6). Figure 2 illustrates the steps to be followed to obtain the BFCC coefficients.

$$\text{Bark}(f) = \frac{26.81f}{1960+f} + 0.53 \tag{7}$$

3.1.3 Empirical Mode Decomposition (EMD)

EMD is a data-driven approach that divides nonstationary and nonlinear signals into amplitude- and frequency-modulated single-signal constituents referred to as IMF (12).

The EMD plays a crucial role in the HHT. It splits a signal s(t) into a limited set (M) of oscillatory constituents known as Intrinsic Mode Functions (IMFs), namely $IMF_1(t)$, $IMF_2(t)$, ..., $IMF_M(t)$. The original signal can be expressed as the sum of all its IMFs along with a monotonic or constant function called the residual r(t). To qualify as an IMF, a signal must satisfy two criteria: (i) the number of extrema and zero-crossings should be either equal or differ by no more than one, and (ii) at any given point within an IMF, the average value of the envelope formed by the local maxima and the envelope formed by the local minima must be zero. The EMD generates upper and lower envelopes by connecting the local maxima and minima of the signal using cubic splines, respectively. The mean envelope is then calculated, and the residual is obtained by subtracting the mean envelope from the original signal. This result should meet the requirements of an IMF; if not, the operation is reiterated, but this time with the residual calculated as input. Sifting is the name given to this repetitive process. Eventually, once the latter residue has no more than two extremes, the decomposition process is achieved. This algorithm is repetitive and has no purely mathematical foundation because it was developed empirically. The decomposition produces a family of IMFs, each one contains oscillations whose frequency composition is lower than the one of the previous IMF. The lower order IMFs are used to depict fast oscillations, whereas the upper order IMFs are employed to show the slowly oscillating ones.

Figure 3 illustrates a set of IMF signals generated by EMD from the pronunciation of "a" by a PD patient. It demonstrates that, because of the EMD operation, the high-frequency part of the original speech signal is obtained first, and the frequency of the following IMFs gradually reduces. This is in aligned with the results of references (7,13), which proves that according to the frequencies of the IMFs, they are grouped in decreasing order. We used the first eight IMFs in this experiment

3.2 Classification

After the feature extraction step, the features were inserted into the SVM and ANN classifiers to differentiate between sufferers of PD and healthy individuals. We examine the BFCC and MFCC coefficients extracted from the IMF_1 to IMF_8 signals to assess the efficiency of the SVM classier on diverse IMFs and to identify the strongest coefficients for discriminating PD from healthy individuals.

3.2.1 Support Vector Machines (SVM)

SVM are popular learning tools in biomedical and health informatics applications (14). An SVM generates an optimal hyperplane during training that can increase the distance of any class from the nearest training data points. The main motivations for machine learning scientists to apply SVMs to their studies are as follows: (1) First, SVMs are very effective in generalizing new data, (2) The second reason is that SVMs rely on a small number of hyperparameters. Let's examine a set of data Ts with S instances, T= {(xi, xj) |xi ε RD, xj ε {-1,1}} Si=1 where xi represents an instance for i^{th}, D is the size of the Parkinson's disease data's original feature space, and xj represents the category labels, i.e., Parkinson's disease, either present or absent. For the Parkinson's disease dataset examined in this paper, the value of D is 38. The SVM modeling generates a hyperplane by f(x)= θT *x + δ, where δ represents the bias and θ the weight vector. Depending on the

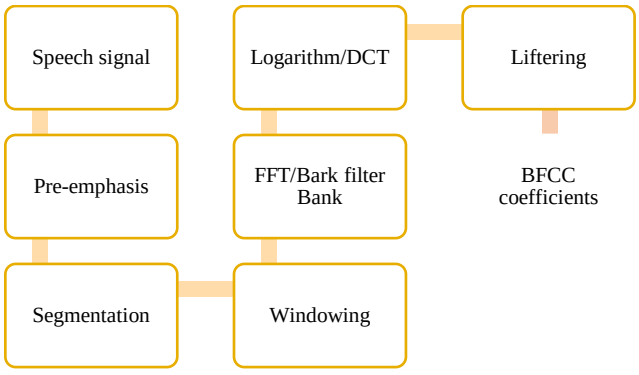

Fig. 2. Steps of calculation of the BFCC coefficients

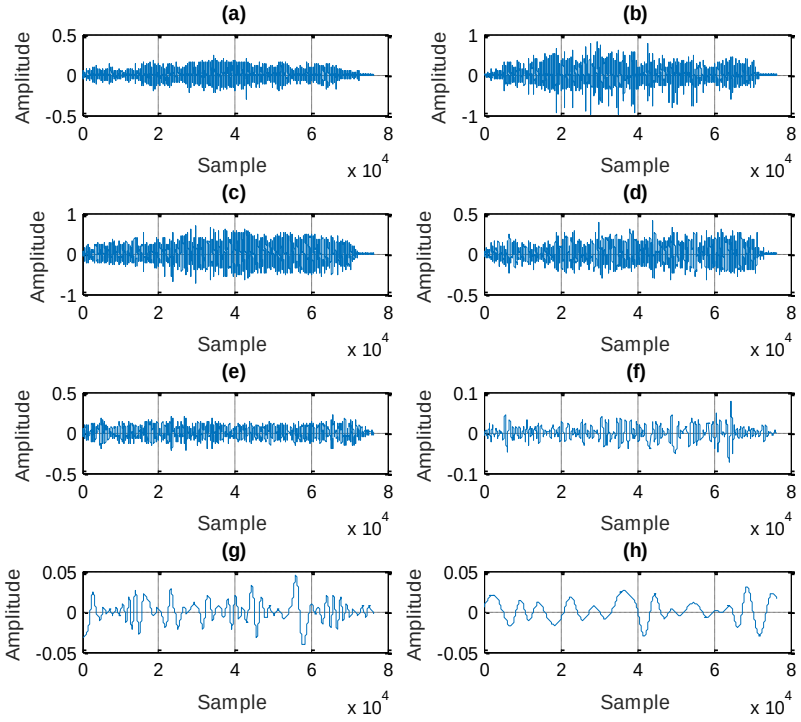

Fig. 3. IMFs obtained by EMD for a PD patient (a: IMF1, b: IMF2, c: IMF3, d: IMF4, e: IMF5, f: IMF6, g: IMF7, h: IMF8).

trained data, The SVM's hyperplane f(x) increases the margin while decreasing (lowering) the classification error. The margin is calculated by adding the distances between the nearest negative and positive instances. In other words, the hyperplane increases the distance of the margin $2/\|\theta\|$. The variables of the SVM are determined by the type of kernel employed. In this paper, a linear kernel function was applied.

3.2.2 Artificial Neural Network (ANN)

An ANN classifier, drawing inspiration from the human brain, is a machine-learning model designed explicitly for classification tasks. Its structure comprises multiple interconnected neuron layers, passing input signals through synaptic weights to generate an output.

In the context of an ANN classifier, input data is presented as vectors, and the neurons in the initial layer process distinct data features. As the input values propagate through subsequent layers, activation functions are utilized to compute the outputs of the neurons. Ultimately, the last layer of the network generates an output that corresponds to the predicted class for the given input data.

3.2.3 Evaluation

Three performance indicators have been used in this work, to determine the performance of classifiers on datasets: sensitivity, accuracy and specificity (10). Accuracy is defined as the percentage of valid diagnostic results. Their definitions are as follows:

$$\text{Accuracy} = \frac{TN + TP}{TN + TP + FP + FN} \tag{8}$$

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

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{10}$$

With:

Normal subjects who have been accurately categorized are True Positives (TP).

Parkinson's disease sufferers who were accurately categorized are True Negatives (TN).

Parkinson's disease sufferers who were misclassified are False Positives (FP).

Normal subjects who were misclassified are False Negatives (FN).

4. RESULTS AND DISCUSSION

4.1 Experiment results of BFCC on dataset Sakar

The extraction of voice signal parameters will be performed by the EMD. 8 IMFs are obtained; then the calculation of BFCC and MFCC coefficients from each IMF. Ultimately, the classification phase involves applying the ANN and SVM classifiers

. This process commences by training the vectors extracted from the data utilized for learning. It concludes with a testing phase in which new data is classified based on the trained model. This PD detection model will be applied to 38 different recordings, included 18 normal subjects and 20 sufferers of PD (11). The simulation of this model is carried out using Matlab software R2022b, which allowed us to save a lot of time at the level of programming thanks to the functions, which are incorporated there. The speech signal is split into a

sum of single component signals, resulting in IMFs, in the first step of the EMD method. eight IMFs have been confirmed on these two requirements:

1) The number of extrema minus the number of zero crossings must be less than or equal to one.

2) At any point, the average of the envelopes defined by the local maxima and minima is equal to zero.

In the second step, each IMF will be considered as an BFCC and MFCC block input, from which 12-cepstral coefficients have been recovered for each person. For the speech to be recognized correctly, therefore, it is necessary to model these coefficients by calculating the average value. Finally, a classification phase employing the SVM and ANN classifiers allowed to identify healthy patients from sick patients. The validation technique includes holdout cross-validation to test the generalizability of the system. The holdout is on testing a classifier with 20% of the participant's speech samples in the dataset, while the remaining (80%) of the speech samples are utilized for training, with 20% as test data and 80% as training data being randomly selected.

This experiment is based on a comparative study to select the effeteness coefficients between the BFCC and MFCC. Table 1 displays the sensitivity, accuracy, and specificity of the classification obtained using holdout cross-validation on the Sakar dataset, considering the extraction of 12 BFCC and 12 MFCC coefficients. It is observed that the characteristics derived from the BFCC have the greatest accuracy of 92.10 compared to the MFCC. Table 1 presents the calculated percentages for accuracy, sensitivity, and specificity across all records. Table 2 presents the performance comparison of BFCC and MFCC using ANN. The results indicate that BFCC achieves the highest accuracy of up to 85.71%. These findings demonstrate the effectiveness of the proposed method in Parkinson's disease diagnosis.

4.2 Discussion

The objective of Table 3 is to provide an overview of the effectiveness of the different PD detection methods. In Table 3, The same data is used in this paper for the literature (2,15,16), we performed experiments on our dataset and strengthen the comparisons using the methods described in (17–19).

In references (2) and (15), the focus is on DWT for the diagnosis of Parkinson's disease. The method of literature (2) first applied DWT, especially wavelet Debauchie 2 scale 3 followed by the extraction of MFCC, then the classification step by using an SVM classifier, whereas those of Zayrit et al. Still based on the same wavelet type, they use the SVM classifier (RBF function) to extract the MFCC coefficients (15). Compared the latter with the results of the article (2), the accuracy has decreased. (Results in ref. (2) are higher than those in ref. (15)). Both methods give lower precision results than our work since the EMD decomposition gives an adaptive decomposition of the analyzed signal, EMD is a tool that does not require a priori fixed basis functions as in this case wavelet transform.

.

Table 1. Sensitivity, Accuracy and Specificity for the 8 IMFs using SVM

IMFs	SVM					
	MFCC			BFCC		
	Accuracy (%)	Sensitivity (%)	Specificity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)
IMF1	73.33	85	61.11	71.05	100	38
IMF2	83.33	75	50	73.68	100	44.44
IMF3	83.33	90	66.66	73.68	85	61.11
IMF4	86.67	60	72.22	73.68	70	77.77
IMF5	86.67	100	72.22	**92.10**	95	88.88
IMF6	**90**	85	72.22	81.57	80	83.33
IMF7	71.05	70	50	68.42	65	72.22
IMF8	76.67	90	61.11	68.42	66.66	70

Table 2. Sensitivity, Accuracy and Specificity for the 8 IMFs using ANN

IMFs	ANN					
	MFCC			BFCC		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
IMF1	28,57	33,33	25	42,85	33,33	50
IMF2	42,85	33,33	50	28,57	33,33	25
IMF3	**71,42**	100	50	42,85	33,33	50
IMF4	**71,42**	66,66	75	71,42	66,66	75
IMF5	**71,42**	66,66	75	**85,71**	100	75
IMF6	57,14	66,66	50	71,42	66,66	75
IMF7	57,14	66,66	50	57,14	66,66	50
IMF8	**71,42**	100	50	71,42	100	50

Table 3. A contrast of the suggested strategy with earlier approaches

Authors (methods source)	Features	Dataset	Accuracy (%)
Belhoussine et al. (2)	DWT-MFCC	Dataset Sakar	86,84
Zayrit et al. (15)	DWT- MFCC	Dataset Sakar	81
Benba et al. (16)	MFCC-SVM	Dataset Sakar	73,53 (polynomial kernel)
Fang et al. (17)	MFCC, Delta MFCC, Delta Delta MFCC	Dataset Sakar	89,29
Solanalavalle et al. (18)	MFCC, Baseline features, TQWT features, WT features	Dataset Sakar	89,66
Sakar et al. (19)	MFCC, Baseline features, TQWT features WT features	Dataset Sakar	81,45
Our method	EMD-BFCC-SVM	Dataset Sakar	**92.10**

Similarly, the results in Ref. (16) show that the accuracy rate of the Sakar dataset is 73.53%, which is focused on the extraction of MFCC coefficients, and afterward, the classification using a polynomial kernel SVM classifier. On the other hand, when experimenting with a deep neural network (DNN) classifier centered on MFCC, delta MFCC and delta delta MFCC features reported in Ref. (17), using the Sakar dataset achieved 89.29% accuracy. This accuracy is lower than that obtained using the EMD-BFCC feature. It is thought that the employment of DNN classifiers poses the problem of expanding the learning parameters, leading to the danger of adapting or falling into a local optimum.

Moreover, the detection of PD relies on MFCC, the baseline features, along with the TQWT and Wavelet Transform (WT) features mentioned in references (18,19). The approach described in reference (18) employed feature selection techniques prior to classification, resulting in an improved accuracy rate. These findings are consistent with the original research, where the results reported in reference (18) outperformed those in reference (19). Since the EMD-BFCC function is a centralized method for extracting information from speech signals, both methods are less accurate than our paper, this is justified by the effectiveness of the EMD method which is characterized by non-linearity, multi-resolution, locality, and self-adaptive as well as for the BFCC coefficients, which have been simulated according to human hearing's auditory model. However, since the basic features extract only the basic information from the speech signal, the extracted information may be incomplete. Consequently, the characteristics obtained from TQWT and MFCC miss significant information about the high frequencies of the signal, resulting in a lower result than our proposed EMD-BFCC feature. This aligns with the conclusion put forth in this paper, which states that significant information about individuals with Parkinson's disease can be extracted from the high-frequency components obtained through the decomposition of speech signals. Based on the previous analysis and Table 1, we can see that the accuracy of BFCC is the highest compared to the MFCC algorithm. For these reasons, BFCC is the most appropriate algorithm for diagnosing patients with PD.

Extensive studies have been conducted to explore the link between Parkinson's disease and voice signals. Citation (20) assessed acoustic and perceptual properties and found that mildly affected patients already had speech and language abnormalities, with a marked deterioration in speech quality and articulation speed. Reference (21) compared the informational content and informational efficacy of oral language spoken by patients with Parkinson's disease to a sound reference group and conclude that the group with Parkinson's disease was different from the reference group concerning both conceptual and vocabulary content and efficacy measures. Reference (22) created a speech-based therapy tool for PD patients. These studies show that speech- and language-based Parkinson's disease diagnosis is meaningful and potentially useful, and they lay the groundwork for PD detection. In the process of extracting characteristics, the voice signal was partitioned into IMFs utilizing EMD method. In order to capture the time-frequency properties of signals at different scales, these IMFs are grouped in descending order of frequency constituents. New characteristics are then obtained from IMFs and ranked. According to Table 1, the BFCCs obtained from the low-order IMFs are more successful in predicting diseased individuals (PD patients) than normal individuals. Furthermore, the high-frequency signals extracted from speech signals encompass more pertinent information about individuals with Parkinson's disease (PD). This observation is in line with the findings reported in reference (23), which emphasizes that characteristics derived from the spectrogram of high-frequency signals contain crucial information. Similarly, reference (19) highlights the substantial variations in the amount of

information present in high-frequency speech samples from individuals with PD. Simultaneously, previous literature (24) has demonstrated that the physiological details of the voice, obtained from the high-frequency segment of the signal, can mirror the physiological changes occurring in the speech organs.

In our study, holdout cross-validation was employed, dividing the data into 80% for training purposes and 20% for testing. This approach effectively mitigates the risks of overfitting and underfitting, as well as preventing inflated predictive accuracy resulting from intricate link between complex predictors of diagnostic and identification status. The classification outcomes indicate that the proposed characteristic selection process in this article is robust and capable of accurate classification.

From all these explanations, we can notice that for a better diagnosis of PD patients, this EMD-BFCC and SVM method with the holdout cross validation using the speech signal gives a better result than MFCC and other techniques.

Table 4 illustrates that even the most promising EMD-based BFCC method was unsuccessful in accurately predicting between Parkinson's and normal subjects. The decline in performance could potentially be attributed to variations in recording conditions or differences in language patterns present within the two datasets.

Table 4. Performance Evaluation of Support Vector Machine for Classifying Parkinson's Disease and Healthy Subjects in Two Independent Databases.

Train database	Test database	Features	ACC (%)
Database 1	Database 2	BFCC	56
Database 2	Database 1	BFCC	52,63

5. CONCLUSION AND FUTURE DIRECTIONS

In the present study, we propound a novel methodology founded on combination of EMD-BFCC as a feature extraction step. We test its performance on a database of 38 records, 18 of which are sound patients and 20 of which are PD ones. The voice signal transformation is processed by the EMD method from which we obtained 8 IMFs; Each IMF will be infused into the BFCC block to retrieve 12 coefficients, which will then be applied to the SVM classifier with an 80% training base. Finally, we perform a test with 20% of the records. We achieve an accuracy of 92.10%. The experiments show that the composed EMD-BFCC can effectively diagnose PD patients. The goal of future work is to prove the effectiveness of this proposed model by applying it to a large dataset and other types of sounds as well as to the classification of different neurodegenerative diseases based on the speech signal.

Author contributions: *research concept and design, N.B., M.M., B.N., T.B.D.; Collection and/or assembly of data, N.B., M.M., B.N., T.B.D.; Data analysis and interpretation, N.B., M.M., B.N., T.B.D.; Writing the article, N.B., M.M., B.N., T.B.D.; Critical revision of the article, (E); Final approval of the article, N.B., M.M., B.N., T.B.D.*

Declaration of competing interest: *The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.*

REFERENCES

1. de Lau LM, Breteler MM. Epidemiology of Parkinson's disease. Lancet Neurology 2006; 5(6): 525–35. https://doi.org/10.1016/S1474-4422(06)70471-9.
2. Drissi TB, Zayrit S, Nsiri B, Ammoummou A. Diagnosis of Parkinson's Disease based on Wavelet Transform and Mel Frequency Cepstral Coefficients. International Journal of Advanced Computer Science and Applications 2019;10(3). https://dx.doi.org/10.14569/IJACSA.2019.0100315.
3. Nouhaila B, Taoufiq BD, Benayad N. An intelligent approach based on the combination of the discrete wavelet transform, delta delta MFCC for Parkinson's Disease diagnosis. International Journal of Advanced Computer Science and Applications 2022;13(4). http://dx.doi.org/10.14569/IJACSA.2022.0130466.
4. Karan B, Sahu SS, Mahto K. Parkinson disease prediction using intrinsic mode function-based features from speech signal. Biocybernetics and Biomedical Engineering 2020; 40(1): 249–64. https://doi.org/10.1016/j.bbe.2019.05.005.
5. Karan B, Sahu SS, Orozco-Arroyave JR, Mahto K. Hilbert spectrum analysis for automatic detection and evaluation of Parkinson's speech. Biomedical Signal Processing and Control 2020; 61: 102050. https://doi.org/10.1016/j.bspc.2020.102050.
6. Karan B, Sahu SS, Orozco-Arroyave JR, Mahto K. Non-negative matrix factorization-based time-frequency feature extraction of voice signal for Parkinson's disease prediction. Computer Speech & Language 2021;69:101216. https://doi.org/10.1016/j.csl.2021.101216.
7. Zhang T, Zhang Y, Sun H, Shan H. Parkinson disease detection using energy direction features based on EMD from voice signal. Biocybernetics and Biomedical Engineering 2021; 41(1): 127–41. https://doi.org/10.1016/j.bbe.2020.12.009.
8. Zhang T, Lin L, Zhang Y, Niu X. Statistical analysis of local gradient in Mel transform domain for Parkinson's dysphonia. Journal of Frontiers of Computer Science and Technology. 2022; 16(10): 2345-2356. https://doi.org/10.3778/j.issn.1673-9418.2102055.
9. Soumaya Z, Drissi Taoufiq B, Benayad N, Yunus K, Abdelkrim A. The detection of Parkinson disease

using the genetic algorithm and SVM classifier. Applied Acoustics 2021; 171:107528. https://doi.org/10.1016/j.apacoust.2020.107528.
10. Sakar BE, Isenkul ME, Sakar CO, Sertbas A, Gurgen F, Delil S, Apaydin H, Kursun O. Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. IEEE Journal of Biomedical and Health Informatics 2013; 17(4): 828–834. https://doi.org/10.1109/JBHI.2013.2245674.
11. Davis SB, Mermelstein P. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. in: readings in speech recognition. IEEE Transactions on Acoustics, Speech, and Signal Processing 1990; 28(4): 65–74. https://doi.org/10.1109/TASSP.1980.1163420.
12. Shrivastava Y, Singh B. A comparative study of EMD and EEMD approaches for identifying chatter frequency in CNC turning. European Journal of Mechanics - A/Solids 2019; 73: 381–393. https://doi.org/10.1016/j.euromechsol.2018.10.004.
13. Mondal A, Banerjee P, Tang H. A novel feature extraction technique for pulmonary sound analysis based on EMD. Computer Methods and Programs in Biomedicine 2018 :159:199–209. https://doi.org/10.1016/j.cmpb.2018.03.016.
14. Mansour RF, Amraoui A El, Nouaouri I, Diaz VG, Gupta D, Kumar S. Artificial intelligence and Internet of Things enabled disease diagnosis model for smart healthcare systems. IEEE Access 2021; 9: 45137–45146. https://doi.org/10.1109/ACCESS.2021.3066365.
15. Soumaya Z, Taoufiq BD, Nsiri B, Abdelkrim A. Diagnosis of Parkinson disease using the wavelet transform and MFCC and SVM classifier. 2019 4th World Conference on Complex Systems (WCCS) 2019; 1–6. https://doi.org/10.1109/ICoCS.2019.8930802.
16. Benba A, Jilbab A, Hammouch A, Sandabad S. Voiceprints analysis using MFCC and SVM for detecting patients with Parkinson's disease. 2015 International Conference on Electrical and Information Technologies (ICEIT) 2015; 300–304. https://doi.org/10.1109/EITech.2015.7163000.
17. Fang SH, Tsao Y, Hsiao MJ, Chen JY, Lai YH, Lin FC. Detection of pathological voice using cepstrum vectors: a deep learning approach. Journal of Voice 2019; 33(5): 634–641. https://doi.org/10.1016/j.jvoice.2018.02.003.
18. Solana-Lavalle G, Galán-Hernández JC, Rosas-Romero R. Automatic Parkinson disease detection at early stages as a pre-diagnosis tool by using classifiers and a small set of vocal features. Biocybernetics and Biomedical Engineering 2020; 40(1): 505–516. https://doi.org/10.1016/j.bbe.2020.01.003.
19. Sakar CO, Serbes G, Gunduz A, Tunc HC, Nizam H, Sakar BE. A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform. Applied Soft Computing 2019; 74: 255–263. https://doi.org/10.1016/j.asoc.2018.10.022.
20. Skodda S, Grönheit W, Mancinelli N, Schlegel U. Progression of voice and speech impairment in the course of Parkinson's Disease: a longitudinal study. Parkinsons Dis. 2013; 2013: 1–8. https://doi.org/10.1155/2013/389195.
21. Roberts A, Post D. Information content and efficiency in the spoken discourse of individuals with Parkinson's disease. Journal of Speech, Language, and Hearing Research 2018; 61(6): 2259–74. http://dx.doi.org/10.1044/2018_JSLHR-L-17-0338.
22. Mühlhaus J, Frieg H, Bilda K, Ritterfeld U. Game-based speech rehabilitation for people with Parkinson's Disease. Universal Access in Human-Computer Interaction. Human and Technological Environments 2017 :76–85. https://doi.org/10.1007/978-3-319-58700-4_7.
23. Sharma R, Mahadeva Prasanna SR. A better decomposition of speech obtained using modified Empirical Mode Decomposition. Digital Signal Processing 2016;58:26–39. https://doi.org/10.1016/j.dsp.2016.07.012.
24. Hayakawa S, Itakura F. Text-dependent speaker recognition using the information in the higher frequency band. Proceedings of ICASSP '94 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE 1994; p. I/137-I/140. https://doi.org/10.1109/ICASSP.1994.389336.

BOUALOULOU Nouhaila. Was born in DAR BOUAZZA CASABLANCA, Morocco on August 8, 1997. She graduated in Electronics, Electrotechnics, Automatic, and Industrial Computing from Faculty of Science Ain Chock. University Hassan II - Casablanca, Morocco. In 2020, she is a research student in Research Laboratory in Industrial and Electrical Engineering, Information Processing, Informatics and Logistics (GEITIIL). Faculty of Science Ain Chok, University Hassan II - Casablanca, Morocco. Her interests are in speech processing for detecting people with neurological disorders.
e-mail: boualoulounouha@gmail.com

MOUNIA Miyara, received the Ph.D. degree in Analysis and Optimization jointly supervised by the Faculty of Sciences, University of Mohamed V, Rabat, Morocco, and the Faculty of Sciences of Montpellier, France, in 1988. She is currently a professor of computer science in Mathematics and Computer Science Department, Faculty of Sciences Ain Chock, Casablanca, Morocco. She is also the vice dean for education at the faculty. She is member of Computer Science and Systems Laboratory (LIS), Faculty of Science Ain Chock, University Hassan II of Casablanca, Morocco. Her current research interests include Information Systems, Big Data, Artificial intelligence and e-learning.
e-mail: miyara12@gmail.com

Nsiri Benayad. Held MBI degree in computer sciences from Telecom Bretagne, in 2005, and Ph.D. degree in signal processing from Telecom Bretagne, in 2004. He received D.E.A (French equivalent of a M.Sc. degree) in electronics from Occidental Bretagne University, in 2000. Currently, he is a Full Professor in the National School of Arts and Crafts of Rabat (ENSAM),), Mohammed V University; a member of Research Center STIS, M2CS, Mohammed V University; and a member associate in Researcher,

Industrial Engineering, data processing and logistic Laboratory, Hassan II University. He was a Professor in the Faculty of Sciences Ain Chock, Hassan II University. Benayad NSIRI has advised and co-advised more than 15 Ph.D. thesis and contributed to more than 80 articles in regional and international conferences and journals. His research interests include but are not restricted to computer science, telecommunication, signal and image processing, adaptive techniques, blind deconvolution, MCMC methods, seismic data, and higher-order statistics.
e-mial: nsiri2000@yahoo.fr

Belhoussine Drissi Taoufiq. was born in Oujda, Morocco in 1978 he graduated with a Ph.D. degree in acoustics in 2009 at the university of le Havne (France) Since 2011 he has been an assistant professor at the sciences faculty of Ain chock University Hassan II, Casablanca. His scientific interest lies in the research of non-destructive testing and signal treatment.
e-mail: taoufiq_belhoussine_drissi@yahoo.fr