

An Intelligent Approach based on the Combination of the Discrete Wavelet Transform, Delta Delta MFCC for Parkinson's Disease Diagnosis

BOUALOULOU Nouhaila¹, BELHOSSINE DRISSI Taoufiq²

Laboratory Electrical and Industrial Engineering, Information Processing, Informatics, and Logistics (GEITIL), Faculty of Science Ain Chock. University Hassan II, Casablanca, Morocco

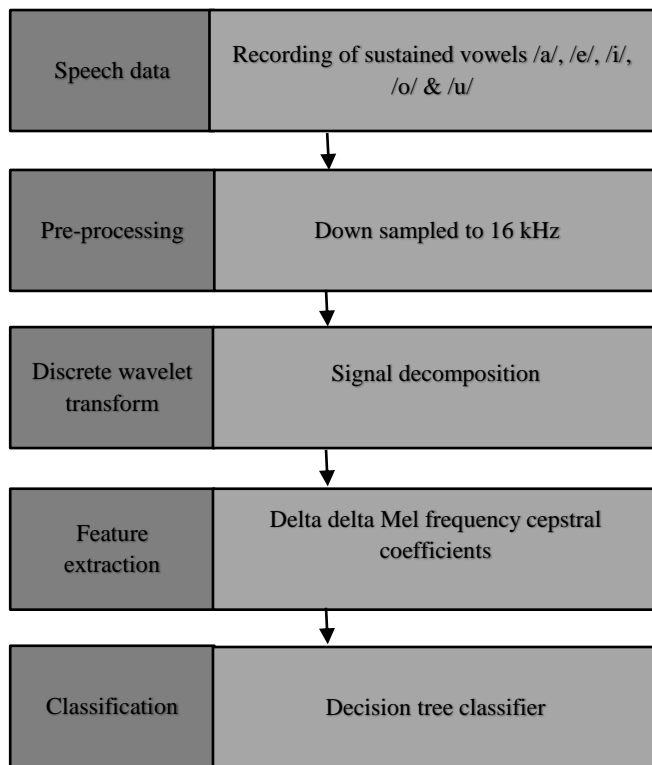
NSIRI Benayad³

Research Center STIS, M2CS, National Higher School of Arts and Craft, Rabat (ENSAM) Mohammed V University in Rabat Morocco

Highlights

- Collection of several types of speech recordings of the vowels /a/, /e/, /i/, /o/ and /u/.
- The decomposition of each voice signal by DWT using the different types of wavelets.
- Extraction of the delta delta MFCC from all voice samples.
- Classification using decision tree classifier along with holdout scheme

Graphical Abstract:



Abstract—To diagnose Parkinson's disease (PD), it is necessary to monitor the progression of symptoms. Unfortunately, diagnosis is often confirmed years after the onset of the disease. Communication problems are often the first symptoms that appear earlier in people with Parkinson's disease. In this study, we focus on the signal of speech to discriminate between people with and without PD, for this, we used a Spanish database that contains 50 records of which 28 are patients with Parkinson's disease and 22 are healthy people, these records contain five types of supported vowels (/a/, /e/, /i/, /o/ and /u/), The proposed treatment is based on the decomposition of each sample using Discrete Wavelet Transform (DWT) by testing several kinds of wavelets, then extracting the delta delta Mel Frequency Cepstral Coefficients (delta delta MFCC) from the decomposed signals, finally we apply the decision tree as a classifier, the purpose of this process is to determine which is the appropriate wavelet analyzer for each type of vowel to diagnose Parkinson's disease.

Keywords—Parkinson's disease; discrete wavelet transform; delta delta MFCC; decision tree classifier

I. INTRODUCTION

Parkinson's disease is a severe health problem. According to the American Parkinson's Disease Association (APDA) [1], more than 10 million people worldwide are affected by Parkinson's disease. Due to the loss of certain groups of brain cells that produce neurotransmitters, including dopamine, causes symptoms such as impaired speech, movement, and sleep, as well as panic and anxiety attacks.

Speech disorders include reduced speech intensity, fluctuating fundamental frequency, and irregular speech articulation which are signs that appear early in people with Parkinson's disease, allowing many studies to use the speech signal for the identification of Parkinson's disease [2-5].

However, the speech signal is one of the most complex signals to characterize, which makes it difficult to develop a system to understand different diseases such as Parkinson's disease [6-8], Alzheimer's disease [9-11], and COVID 19 [12-15], etc. This complexity of the speech signal comes from a combination of several factors, the redundancy of the acoustic signal, the high inter-and intra-speaker variability, the effects

of coarticulation in continuous speech, and the recording conditions. To overcome these difficulties, many methods and mathematical models have been developed, including neural networks [16, 17], Support Vector Machines (SVM) [18, 19], stochastic Markov models, and in particular Hidden Markov Models (HMM) [20, 21].

These methods and models work from information extracted from the speech signal considered relevant. This extraction is performed by an acoustic analysis which leads to gathering this information under the term of the vector of acoustic parameters whose dimension and nature are determinant to reach good performances of the knowledge system of Parkinson's disease which is of interest to us in this article. The different types of acoustic parameters commonly cited in the literature are the coefficients: LPC [22], LPCC [23], LFCC [24], PLP [23], MFCC [6, 23], etc. Generally, the MFCC coefficients are the most used acoustic parameters in speech feature extraction [25-27].

However, for these systems of recognition of Parkinson's disease, research work has studied the improvement of the performance of this system by combining the MFCC coefficients with other types of acoustic parameters such as LPCC [28], PLP [28], energy [29], wavelets [29-31] and Empirical Mode Decomposition (EMD) [32].

In this study, far from static coefficients, for correct detection of Parkinson's disease through voice our contribution consists in proposing a new method of selection of relevant acoustic parameters based on the use of dynamic delta delta MFCC coefficients combined with wavelets, these differential coefficients referred as dynamic parameters provide useful information on the temporal trajectory of the speech signal. This information extracted by discrete wavelet transform and delta delta MFCC will be used in the classification block using a decision tree classifier. To evaluate the performance of this model we applied it to a database of five vowels (/a/, /e/, /i/, /o/, /u/), each vowel includes 28 people with Parkinson's disease and 22 are healthy people.

The structure of the rest of the article is as follows: section II concerns a definition of the methods used, section III presents a description of the process used and the results obtained, and finally the conclusion in section IV.

II. METHOD

A. Feature Extraction

For feature extraction, we are interested in the joint use of DWT and the second derivative of the MFCC (delta delta MFCC). This step allows extracting features that will be used by the decision tree classifier.

1) *Discrete wavelet transform*: For the discrete wavelet transform, it is the discrete version of the Continuous Wavelet Transform (CWT) that used Mallat's algorithm, it is based on the principle of multi-resolution that allows the separation of details and approximations of signals by using a pair of filters H and G that constitute respectively a low-pass filter and a high-pass filter.

With the high-pass filters, we obtain the coefficients of the discrete wavelet decomposition (the details), and with the low-pass filters, we obtain the approximation coefficients. This operation is applied again to the approximation, generating another detail and a new approximation as shown in Figure 1. There are several types of wavelets, in our case; we used the wavelets presented in the table I:

2) *Delta delta MFCC*: Temporal changes in the cepstrum play an important role in human perception and it is through the derivatives of MFCC coefficients that we can measure these changes.

TABLE I. WAVELET FAMILIES USED

| Wavelets families | Wavelets |
|-------------------|-------------------------|
| Daubechies | db1, db2, and db3 |
| Coiflets | coif1, coif2, and coif3 |
| Symplets | sym1, sym2, and sym3 |
| Discrete Meyer | dmey |

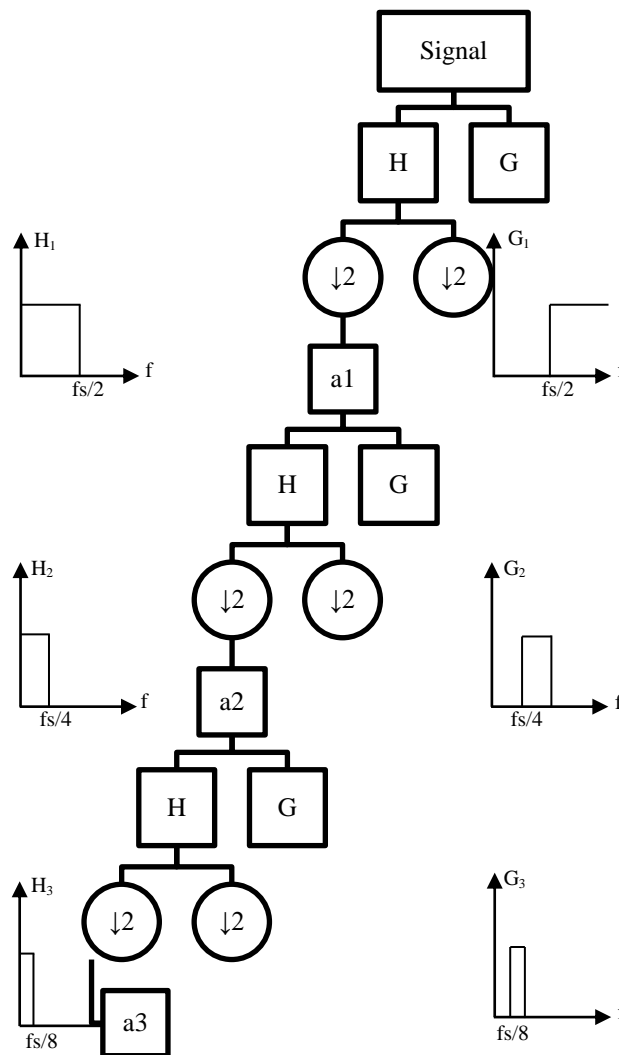


Fig. 1. Representation of the Mallat Algorithm for the Implementation of the DWT Decomposition, H is a Low Pass Filter and G is a High Pass Filter.

Usually, MFCC coefficients are referred to as static parameters, since they contain only the information about a given frame. To improve the frame representation, it is often proposed to introduce new parameters into the parameter vector. The reference [33] proposed the use of dynamic parameters that present cepstral transition information in the speech signal. In particular, it proposed second-order differential coefficients, also called delta delta coefficients, derived from cepstral coefficients. Let d_t be the first-order differential coefficient (delta MFCC) of frame t , then the corresponding second-order differential coefficient (delta delta MFCC) dd_t is calculated by the following formula:

$$dd_t = \frac{\sum_{n=1}^N n(d_{t+n} - d_{t-n})}{2 \sum_{n=1}^N n^2} \quad (1)$$

The coefficients delta delta MFCC, also called acceleration coefficients are obtained using the second derivative of static Mel Frequency Cepstral Coefficients (MFCC), the latter is a representation defined as the discrete cosine transform of the logarithm of the spectrum of the energy of the speech segment. The spectral energy is calculated by applying a bank of evenly spaced filters on a modified frequency scale, called the Mel scale. The Mel scale redistributes the frequencies in a non-linear scale that simulates human perception of sounds. Figure 2 illustrates the steps involved in obtaining the delta delta MFCC coefficients. Based on the results obtained in [26], we use only the first 12 delta delta MFCC coefficients.

B. Feature Classification

1) *The decision tree classifier:* A decision tree is one of the most popular techniques in machine learning. Indeed, decision tree learning is part of supervised learning, it is generally a classifier presented in the form of a tree structure [34].

A decision tree consists of a set of rules allowing a segment of a data set into homogeneous groups. Each rule associates the conjunction of tests with the descriptive variables. The first vertex is called the root of the tree, the following variables, which correspond to non-terminal nodes, are segmentation variables; each branch corresponds to a modality of the variable considered at this level of the tree. This process is repeated on each node of the tree, the nodes that are not pure are segmented until pure leaves are obtained.

Here we try to classify a population of individuals containing healthy people and people with Parkinson's disease pronounce these five vowels (/a/, /e/, /i/, /o/, /u/) into two classes with respect to a label {1 (healthy), 0 (sick)} from the recordings. The decision tree-learning algorithm is described below:

Algorithm 1: Decision tree

Data: a sample Ω of m labeled records

Initialization: empty tree; current node: root; current sample: Ω

- Repeat
- Decide whether the current node is terminal
- If the current node is terminal then
- Label the current node with a leaf
- Otherwise
- Select a test and create the subtree
- End if
- Current node: a node not yet studied
- Current sample: sample reaching the current node
- Until a decision tree is produced

Output: decision tree

2) *Holdout method:* All classification results are obtained using the "holdout" method. This method is a commonly used practice for evaluating machine-learning models. It works by first dividing the data randomly into two parts; one of larger size is used for training and the other part is reserved for error rate estimation. Another version of this method, called "data shuffle", consists in repeating L times the random division of the data into two parts; one for training and the other for testing, and then calculating the average of the L estimates of the error rates evaluated on the test data parts. The advantage of this method is that all data are used for both training and testing. The holdout is a simple method to understand and generally results in a less biased model estimate than other methods.

To assess classifier performance, the following parameters are used:

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FP} + \text{FN}} \quad (2)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (4)$$

Where TP: true positive; TN: true negative; FP: false positive; FN: false negative.

III. METHODOLOGY AND RESULT

We have implemented our algorithm in Matlab version R2019a. The tests were performed on a PC with the following configuration:

CPU: Intel(R) Core(TM) i3-3110M CPU @ 2.40GHz
Memory: 4,00 Go
Operating System: Windows 10 64bit.

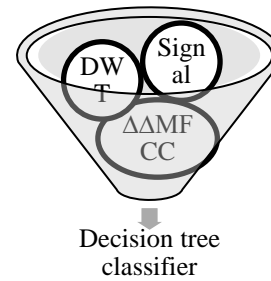


Fig. 3. The Process used to Diagnose Parkinson’s Disease.

A. Data Collection

This study takes into account two subgroups of this group: the Healthy Control (HC) group with 22 speakers and the Parkinson’s disease (PD) group with 28 speakers. All utterers included in this Italian corpus were registered in Bari (Puglia region), Italy. Each recording session took place in a controlled environment, taking into account factors such as room temperature, distance from the microphone, time of day, and having a conversation with the subject to warm up their vocal muscles. The sampling frequency was 16 kHz; more information is available in [35]. Table II includes the demographic information of the corpus.

B. Methodology

The objective of this part of Parkinson’s disease detection is to design and experiment with a system based on extensive use of wavelet types combined with acceleration coefficients (delta delta MFCC) as features and the decision tree as a classifier, as illustrated in Figure 3.

For the feature selection and extraction domain, which is an important step in the recognition process of Parkinson’s disease, we consider the use of the wavelet transform, which is a time and frequency analysis tool that allows obtaining variable temporal and frequency resolutions.

TABLE II. AGE DEMOGRAPHIC DATA FROM THE ITALIAN CORPUS, ARRANGED BY GENDER AND CLASS

| | Female | | Male | | Total | |
|-----------|--------|-------|-------|-------|-------|-------|
| | PD | HC | PD | HC | PD | HC |
| Subjects | 9 | 12 | 19 | 10 | 28 | 22 |
| Age range | 40-80 | 60-72 | 50-77 | 60-77 | 40-80 | 60-77 |

Our model for the diagnosis of Parkinson’s disease will be applied to a corpus containing 28 subjects with Parkinson’s disease and 22 healthy subjects for each vowel (/a/, /e/, /i/, /o/, /u/). It is based firstly on the decomposition of the signal into two sub-bands of approximate frequency and details by the use of a wide range of wavelets (Debauchies, Coiflets, Symlets, and Discrete Meyer) on the first seven scales.

Susceptible to better characterize our system, extracting the first 12 coefficients delta delta MFCC that are parameters widely used to encode the dynamic information of cepstral parameters. This extraction is only performed on the low-frequency band (approximation) for each type of wavelet and from the first scale up to the seventh scale.

The main objective of the feature selection step is its direct contribution to the performance of the overall system this step allowed us to extract the features that will be then used by the decision tree classifier with the cross-validation method "holdout" that serves to decompose the base of each vowel (/a/, /e/, /i/, /o/, /u/) into 80% as a learning base and the whole base as a test base.

C. Experiments and Results

In this part of our work, we implemented our Parkinson’s disease detection system where we introduced the DWT method and the delta delta MFCC coefficients as a feature vector, we changed the wavelet types for each first seven scales, then we measured the impact of each of these changes on the Parkinson’s disease recognition performance by the decision tree classifier performance evaluation metrics.

A comparative study between the extracted features based on delta delta MFCCs is performed to visualize the distinctive behavior of the subject with Parkinson’s disease and the healthy subject for the five vowels. Figures 4-8 show the 12 delta delta MFCC coefficients as extracted features for a healthy person and a person with Parkinson’s disease. The feature patterns are obtained from 50 speech samples for each vowel (/a/, /e/, /i/, /o/, /u/) from people with Parkinson’s and healthy people. Almost all speech signals follow the same pattern. Figure 4 clearly distinguishes the variation in the speech pattern of a healthy person and a person with Parkinson’s disease for the vowel /a/ based on the wavelet of discrete Meyer at scale 5 the latter is the one that gave us a better Accuracy. Figures 5, 6, 7, and 8 present the delta delta MFCC coefficients respectively for the vowels /e/, /i/, /o/, and /u/, we based on these wavelets with these scales respectively discrete Meyer at scale 5, Coiflets level 3 at scale 6, discrete Meyer at scale 7 and discrete Meyer at scale 6 because it is those that gave us

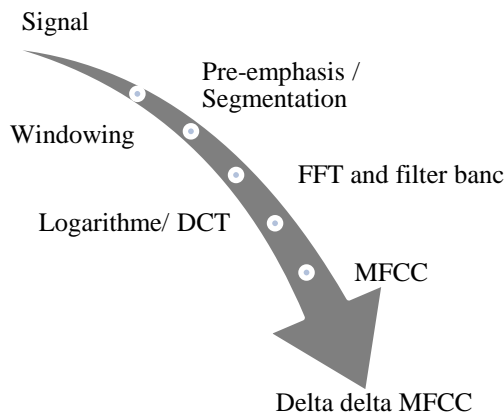


Fig. 2. Delta Delta MFCC Extraction Process.

important results that are displayed in Tables III, IV, V, VI, and VII a notable variation is observed between the healthy speech signal and that of Parkinson's disease. Thus, the proposed feature-based on delta delta MFCC can be a good marker for the prediction of Parkinson's disease.

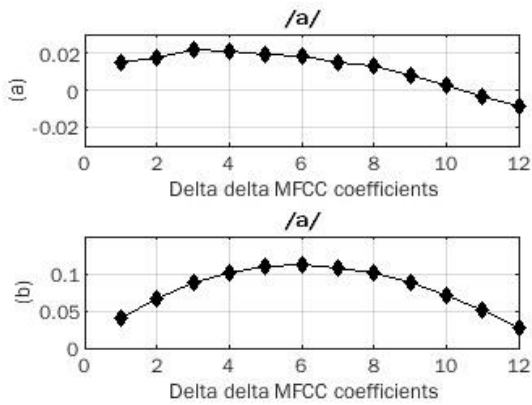


Fig. 4. Delta Delta Coefficients Extracted from the Vowel /a/ for: (a) Healthy Person, (b) Person with Parkinson's Disease.

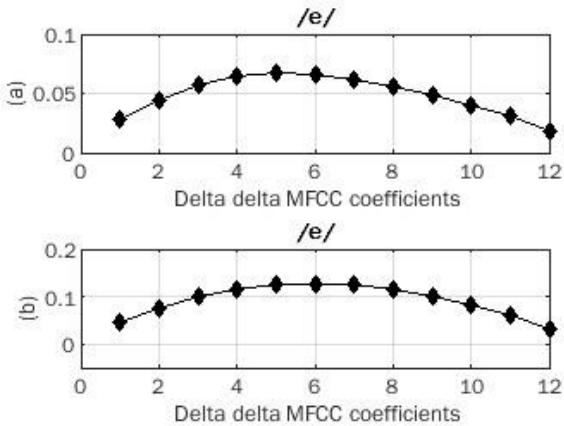


Fig. 5. Delta Delta Coefficients Extracted from the Vowel /e/ for: (a) Healthy Person, (b) Person with Parkinson's Disease.

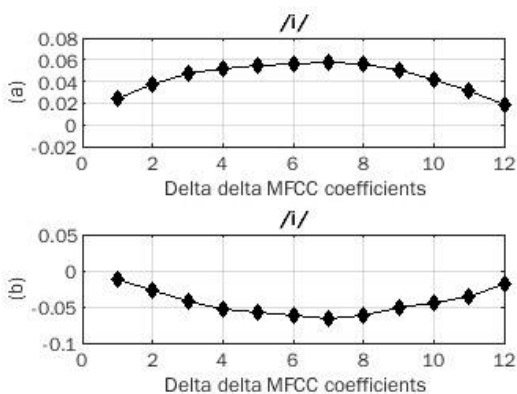


Fig. 6. Delta Delta Coefficients Extracted from the Vowel /i/ for: (a) Healthy Person, (b) Person with Parkinson's Disease.

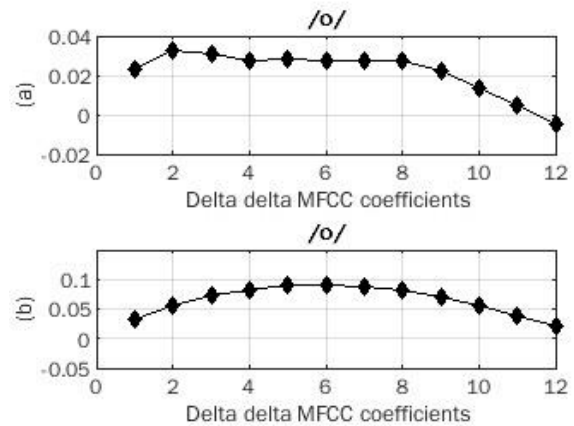


Fig. 7. Delta Delta Coefficients Extracted from the Vowel /o/ for: (a) Healthy Person, (b) Person with Parkinson's Disease.

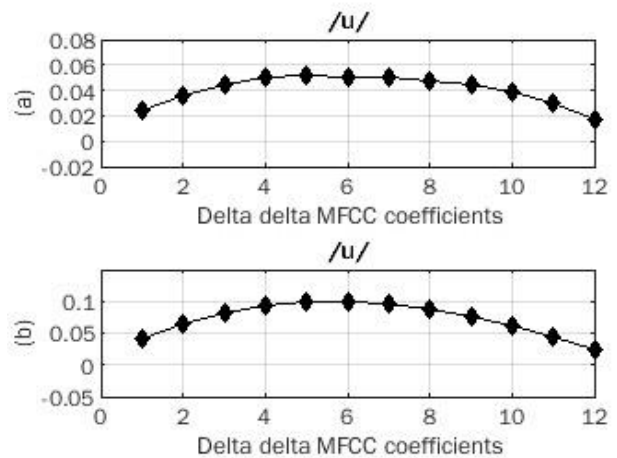


Fig. 8. Delta Delta Coefficients Extracted from the Vowel /u/ for: (a) Healthy Person, (b) Person with Parkinson's Disease.

The following tables show the different results obtained for the five vowels according to accuracy, specificity, and sensitivity.

For the vowel /a/, the results obtained in Table III show that the wavelet by using discrete Meyer at scale 5 offers better modeling for the classification of healthy people and people with Parkinson's disease with high accuracy of 97.5%. While for the vowel /e/, Table IV is still with the same wavelet, and on the same scale of 5, the best accuracy of 92.5% is achieved.

In contrast to the experiment for the vowel /i/, in Table V we register a slight performance low of the accuracy of 87,5% by using the wavelet of Coiflets level 3 at scale 6 and discrete Meyer at scale 7.

For the vowel /o/, according to the results shown in Table VI, we find that the wavelet of discrete Meyer at scale 7 gives accuracy up to 92,5% for the discrimination between healthy patients and patients with Parkinson's disease. For the vowel /u/, Table VII shows an important result with an accuracy of 90% by using the wavelet of discrete Meyer at scale 6.

TABLE III. ACCURACY, SPECIFICITY, AND SENSITIVITY FOR THE VOWEL /A/ USING THE DECISION TREE CLASSIFIER

| Scale wavelet | Accuracy | | | | | | | Specificity | | | | | | | Sensitivity | | | | | | |
|---------------|----------|------|------|------|------|------|------|-------------|------|------|------|------|------|------|-------------|------|------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Db1 | 67,5 | 67,5 | 62,5 | 77,5 | 85 | 83 | 80 | 67,9 | 32,1 | 53,6 | 82,1 | 71,4 | 63,6 | 50 | 4.5 | 72,7 | 45,5 | 63,6 | 63,6 | 75 | 82,1 |
| Db2 | 65 | 67,5 | 70 | 77,5 | 90 | 72,5 | 72,5 | 31,8 | 31,8 | 68,2 | 25 | 72,7 | 72,7 | 45,5 | 68,9 | 45,5 | 39,3 | 77,3 | 67,9 | 60,7 | 75 |
| Db3 | 67,5 | 65 | 70 | 77,5 | 90 | 87,5 | 77,5 | 72,7 | 40,9 | 63,6 | 45,5 | 81,8 | 68,2 | 59,1 | 32,1 | 32,1 | 53,6 | 78,6 | 35,7 | 71,4 | 64,3 |
| Coif1 | 70 | 65 | 75 | 80 | 90 | 89,3 | 77,5 | 50 | 9,1 | 40,9 | 36,4 | 72,7 | 68,2 | 75 | 50 | 89,3 | 82,1 | 82,1 | 78,6 | 78,6 | 50 |
| Coif2 | 65 | 65 | 77,5 | 77,5 | 95 | 87,5 | 90 | 68,2 | 68,2 | 63,6 | 13,6 | 59,1 | 72,7 | 36,4 | 35,7 | 46,4 | 67,9 | 85,7 | 89,3 | 78,6 | 92,9 |
| Coif3 | 70 | 65 | 70 | 80 | 95 | 85 | 80 | 45,5 | 68,2 | 59,1 | 40,9 | 63,6 | 68,2 | 72,7 | 53,6 | 42,9 | 64,3 | 85,7 | 92,9 | 67,9 | 57,1 |
| Sym1 | 67,5 | 67,5 | 65 | 77,5 | 90 | 77,5 | 72,5 | 45,5 | 45,5 | 63,6 | 4,5 | 36,4 | 36,4 | 68,2 | 64,3 | 71,4 | 35,7 | 75 | 85,7 | 25 | 60,7 |
| Sym2 | 67,5 | 67,5 | 70 | 77,5 | 90 | 87,5 | 70 | 59,1 | 72,7 | 36,4 | 50 | 68,2 | 77,3 | 68 | 53,6 | 42,9 | 67,9 | 82,1 | 82,1 | 67,9 | 72 |
| Sym3 | 67,5 | 65 | 70 | 80 | 90 | 80 | 80 | 40,9 | 72,7 | 59,1 | 13,6 | 63,6 | 68,2 | 77,3 | 67,9 | 39,3 | 60,7 | 85,7 | 85,7 | 50 | 28,6 |
| dmey | 65 | 65 | 67,5 | 82,5 | 97,5 | 82,5 | 95 | 27,3 | 54,5 | 36,4 | 54,5 | 54,5 | 68,2 | 59,1 | 71,4 | 50 | 75 | 82,1 | 96,4 | 71,4 | 89,3 |

TABLE IV. ACCURACY, SPECIFICITY, AND SENSITIVITY FOR THE VOWEL /E/ USING THE DECISION TREE CLASSIFIER

| Scale wavelet | Accuracy | | | | | | | Specificity | | | | | | | Sensitivity | | | | | | |
|---------------|----------|------|------|------|------|------|------|-------------|------|------|------|------|------|------|-------------|------|------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Db1 | 65 | 62,5 | 62,5 | 62,5 | 87 | 80 | 77,5 | 18,2 | 72,7 | 59,1 | 77,3 | 68,2 | 77,3 | 68,2 | 75 | 35,7 | 35,7 | 42,9 | 67,9 | 64,3 | |
| Db2 | 62,5 | 60 | 60 | 60 | 85 | 80 | 67,5 | 45,5 | 72,7 | 50 | 54,5 | 81,8 | 81,8 | 77,3 | 93,3 | 28,6 | 42,9 | 53,6 | 46,6 | 75 | 39,3 |
| Db3 | 62,5 | 57,5 | 65 | 62,5 | 82,5 | 80 | 70 | 9,1 | 4,5 | 50 | 18,2 | 72,7 | 77,3 | 77,3 | 78,6 | 89,3 | 46,4 | 64,3 | 42,9 | 67,9 | 42,9 |
| Coif1 | 65 | 60 | 57,5 | 62,5 | 85 | 80 | 70 | 13,6 | 50 | 50 | 13,6 | 81,8 | 77,3 | 77,3 | 85,7 | 53,6 | 50 | 89,3 | 60,7 | 67,9 | 35,7 |
| Coif2 | 62,5 | 60 | 60 | 62,5 | 77,5 | 77,5 | 70 | 40,9 | 72,7 | 40,9 | 28,6 | 72,7 | 63,6 | 45,5 | 75 | 35,7 | 53,6 | 72,7 | 64,3 | 64,3 | 57,1 |
| Coif3 | 65 | 62,5 | 62,5 | 62,5 | 82,5 | 82,5 | 77,5 | 77,3 | 54,5 | 72,7 | 31,8 | 68,2 | 81,8 | 77,3 | 32,1 | 64,4 | 39,3 | 60,7 | 74,1 | 46,4 | 53,4 |
| Sym1 | 65 | 60 | 60 | 60 | 87,5 | 80 | 77,5 | 9,1 | 40,9 | 27,3 | 72,7 | 4,5 | 72,7 | 63,6 | 75 | 53,6 | 64,3 | 35,7 | 82,1 | 71,4 | 32,1 |
| Sym2 | 62,5 | 57,5 | 67,5 | 62,5 | 82,5 | 80 | 67,5 | 54,5 | 72,7 | 50 | 18,2 | 72,7 | 72,7 | 54,5 | 42,9 | 32,1 | 46,5 | 70 | 64,3 | 67,9 | 60,7 |
| Sym3 | 65 | 57,5 | 65 | 62,5 | 80 | 80 | 70 | 40,9 | 72,7 | 59,1 | 68,2 | 77,3 | 63,6 | 68,2 | 67,9 | 35,7 | 46,4 | 32,1 | 42,9 | 46,4 | 64,3 |
| dmey | 70 | 70 | 62,5 | 60 | 92,5 | 82,5 | 80 | 45,5 | 45,5 | 68,2 | 54,5 | 72,7 | 36,4 | 68,2 | 78,6 | 78,6 | 39,3 | 32,5 | 35,7 | 85,7 | 89,3 |

TABLE V. ACCURACY, SPECIFICITY, AND SENSITIVITY FOR THE VOWEL /I/ USING THE DECISION TREE CLASSIFIER

| Scale wavelet | Accuracy | | | | | | | Specificity | | | | | | | Sensitivity | | | | | | |
|---------------|----------|------|------|------|------|-------------|-------------|-------------|------|------|------|------|------|------|-------------|------|------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Db1 | 77,5 | 75 | 65 | 72,5 | 77,5 | 77,5 | 72,5 | 54,5 | 54,5 | 63,6 | 63,6 | 72,7 | 59,1 | 77,3 | 67,9 | 71,4 | 53,6 | 60,7 | 57,1 | 57,1 | 42,9 |
| Db2 | 77,5 | 75 | 65 | 70 | 77,5 | 65 | 75 | 59,1 | 13,6 | 18,2 | 77,3 | 54,5 | 45,5 | 68,2 | 53,6 | 85,7 | 75 | 21,4 | 67,9 | 53,6 | 60,7 |
| Db3 | 77,5 | 75 | 62,1 | 65 | 75 | 72,5 | 67,5 | 31,8 | 4,5 | 50 | 59,1 | 77,3 | 63,6 | 45,5 | 78,6 | 67,9 | 46,4 | 57,1 | 21,4 | 53,7 | 60,7 |
| Coif1 | 75 | 67,5 | 70 | 75 | 70 | 75 | 75 | 54,5 | 63,6 | 77,3 | 72,7 | 54,5 | 81,8 | 77,3 | 71,4 | 60,7 | 39,3 | 50 | 57,1 | 53,6 | 42,9 |
| Coif2 | 80 | 75 | 67,5 | 60 | 72,5 | 77,5 | 65 | 59,1 | 13,6 | 18,2 | 77,3 | 54,5 | 63,6 | 54,5 | 53,6 | 85,7 | 75 | 21,4 | 67,9 | 46,4 | 57,1 |
| Coif3 | 77,5 | 75 | 67,5 | 67,5 | 70 | 87,5 | 77,5 | 45,5 | 50 | 36,4 | 50 | 50 | 81,8 | 63,6 | 89,3 | 67,9 | 57,1 | 53,6 | 46,4 | 42,9 | 85,7 |
| Sym1 | 77,5 | 75 | 65 | 70 | 77,5 | 80 | 77,5 | 50 | 9,1 | 54,5 | 63,6 | 77,3 | 81,8 | 63,6 | 85,7 | 60,7 | 57,1 | 67,9 | 39,3 | 35,7 | 67,9 |
| Sym2 | 77,5 | 75 | 65 | 70 | 77,5 | 72,5 | 65 | 59,1 | 50 | 22,7 | 68,2 | 77,3 | 72,7 | 81,8 | 75 | 67,9 | 67,9 | 46,4 | 39,3 | 53,6 | 46,4 |
| Sym3 | 77,5 | 75 | 62,5 | 67,5 | 75 | 75 | 67,5 | 59,1 | 50 | 27,3 | 50 | 72,7 | 22,7 | 13,6 | 75 | 67,9 | 71,4 | 67,9 | 46,4 | 67,9 | 71,4 |
| dmey | 81,3 | 75 | 67,5 | 65 | 77,5 | 77,5 | 87,5 | 13,6 | 63,6 | 63,6 | 68,2 | 59,1 | 40,9 | 68,2 | 82,1 | 60,7 | 39,3 | 39,3 | 67,9 | 50 | 78,6 |

TABLE VI. ACCURACY, SPECIFICITY, AND SENSITIVITY FOR THE VOWEL /O/ USING THE DECISION TREE CLASSIFIER

| Scale wavelet | Accuracy | | | | | | | Specificity | | | | | | | Sensitivity | | | | | | |
|---------------|----------|------|------|------|------|------|-------------|-------------|------|------|------|------|------|------|-------------|------|------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Db1 | 82,5 | 75 | 70 | 67,5 | 75 | 77,5 | 77,5 | 63,6 | 63,6 | 18,2 | 68,2 | 68,2 | 81,8 | 50 | 64,3 | 57,1 | 67,9 | 46,4 | 60,7 | 32,1 | 75 |
| Db2 | 80 | 75 | 70 | 70 | 75 | 82,5 | 77,5 | 77,3 | 54,5 | 81,8 | 63,6 | 68,2 | 54,5 | 63,6 | 35,7 | 71,4 | 53,6 | 50 | 57,1 | 64,3 | 67,9 |
| Db3 | 80 | 72,5 | 70 | 65 | 77,5 | 70 | 75 | 22,7 | 59,1 | 59,1 | 63,6 | 59,1 | 36,4 | 77,3 | 85,7 | 64,3 | 39,3 | 25 | 53,6 | 75 | 46,4 |
| Coif1 | 77,5 | 72,5 | 70 | 67,5 | 75 | 77,5 | 72,5 | 63,6 | 54,5 | 81,1 | 54,5 | 77,3 | 72,7 | 81,8 | 60,7 | 32,1 | 35,7 | 50 | 35,5 | 60,7 | 32,1 |
| Coif2 | 72,5 | 75 | 75 | 65 | 77,5 | 80 | 67,5 | 18,2 | 72,7 | 72,7 | 63,6 | 54,5 | 81,8 | 31,8 | 89,3 | 35,7 | 53,6 | 42,9 | 67,9 | 64,3 | 60,7 |
| Coif3 | 75 | 90 | 72,5 | 72,5 | 75 | 82,5 | 79 | 63,6 | 40,9 | 63,6 | 68,2 | 72,7 | 77,3 | 63,6 | 60,7 | 85,7 | 50 | 60,7 | 50 | 53,6 | 67,9 |
| Sym1 | 84 | 75 | 70 | 70 | 75 | 77,5 | 72,5 | 54,5 | 63,6 | 72,7 | 63,6 | 72,7 | 77,3 | 77,3 | 75 | 64,3 | 46,4 | 53,6 | 50 | 60,7 | 46,4 |
| Sym2 | 80 | 75 | 70 | 70 | 77,5 | 82,5 | 77,5 | 54,5 | 77,3 | 45,5 | 59,1 | 40,9 | 77,3 | 59,1 | 64,3 | 21,4 | 64,3 | 46,4 | 64,3 | 67,9 | 64,3 |
| Sym3 | 77,5 | 80 | 70 | 65 | 75 | 75 | 72,5 | 63,6 | 54,5 | 40,9 | 9,1 | 81,8 | 81,8 | 81,8 | 67,9 | 71,4 | 60,7 | 64,3 | 46,4 | 64,3 | 39,3 |
| dmey | 72,5 | 72,5 | 70 | 65 | 77,5 | 80 | 92,5 | 40,9 | 40,9 | 77,3 | 40,9 | 72,7 | 45,5 | 72,7 | 78,6 | 28,6 | 39,3 | 57,1 | 64,3 | 85,7 | 85,7 |

TABLE VII. ACCURACY, SPECIFICITY, AND SENSITIVITY FOR THE VOWEL /U/ USING THE DECISION TREE CLASSIFIER

| Scale wavelet | Accuracy | | | | | | | Specificity | | | | | | | Sensitivity | | | | | | |
|---------------|----------|------|------|------|------|-----------|------|-------------|------|------|------|------|------|------|-------------|------|------|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Db1 | 75 | 75 | 65 | 65 | 67,5 | 77,5 | 72,5 | 72,7 | 50 | 54,5 | 72,7 | 63,6 | 31,8 | 59,1 | 57,1 | 78,6 | 53,6 | 46,4 | 53,6 | 67,9 | 57,1 |
| Db2 | 77,5 | 77,5 | 62,5 | 65 | 75 | 67,5 | 72,5 | 63,6 | 31,8 | 4,5 | 40,9 | 72,7 | 72,7 | 77,3 | 25 | 78,6 | 92,9 | 50 | 57,1 | 50 | 57,1 |
| Db3 | 75 | 77,5 | 62,5 | 62,5 | 72,5 | 80 | 72,5 | 9,1 | 45,5 | 81,8 | 72,7 | 81,8 | 40,9 | 81,8 | 89,3 | 78,6 | 28,6 | 32,1 | 53,6 | 42,9 | 57,1 |
| Coif1 | 77,5 | 77,5 | 62,5 | 65 | 72,5 | 72,5 | 70 | 50 | 54,5 | 40,9 | 68,2 | 50 | 81,8 | 45,5 | 78,6 | 35,7 | 57,1 | 53,6 | 57,1 | 57,1 | 60,7 |
| Coif2 | 72,5 | 72,5 | 62,5 | 62,5 | 72,5 | 72,5 | 70 | 27,3 | 50 | 68,2 | 68,2 | 72,7 | 59,1 | 63,6 | 75 | 75 | 42,9 | 35,7 | 53,6 | 53,6 | 42,9 |
| Coif3 | 72,5 | 67,5 | 62,5 | 62,5 | 70 | 80 | 72,5 | 63,6 | 59,1 | 68,2 | 68,2 | 72,7 | 81,8 | 77,3 | 60,7 | 39,3 | 50 | 35,7 | 57,1 | 53,6 | 57,1 |
| Sym1 | 75 | 75 | 65 | 67,5 | 70 | 77,5 | 70 | 9,1 | 36,4 | 45,5 | 50 | 63,6 | 72,7 | 59,1 | 64,3 | 82,1 | 60,7 | 50 | 50 | 46,4 | 32,1 |
| Sym2 | 77,5 | 77,5 | 62,5 | 65 | 72,5 | 67,5 | 72,5 | 59,1 | 54,5 | 68,2 | 4,5 | 52,2 | 72,7 | 27,3 | 67,9 | 60,7 | 46,4 | 82,1 | 85,7 | 35,7 | 71,4 |
| Sym3 | 75 | 77,5 | 62,5 | 62,5 | 72,5 | 80 | 72,5 | 59,1 | 31,8 | 45,5 | 54,5 | 77,3 | 63,6 | 59,1 | 35,7 | 78,6 | 53,6 | 42,9 | 32,1 | 32,1 | 46,4 |
| dmey | 75 | 80 | 52,5 | 62,5 | 75 | 90 | 72,5 | 27,3 | 50 | 68,2 | 72,7 | 77,3 | 54,5 | 68,2 | 85,7 | 75 | 42,9 | 25 | 53,6 | 64,3 | 85,7 |

In Table VIII we have listed, but not exhaustively, the different works on Parkinson's disease recognition systems. The performances are difficult to compare because they vary according to several elements such as the type of data, the choice of the learning models, the way the parameters are obtained, etc. In this section, we will compare our results with those of some works.

TABLE VIII. COMPARISON OF PERFORMANCE WITH RECENTLY PUBLISHED WORK

| Study | Method | Accuracy |
|-------------------------------|--|---|
| Zayrit et al [29] | DWT-genetic algorithm and SVM | 91,18 (k-fold cross-validation) |
| Benba et al [26] | MFCC- MLP kernel of SVM | 82,50 (LOSO) |
| Sakar et al [36] | KNN+SVM | 55,00 (LOSO on training database) 68,45 (LOSO on the testing database) |
| Belhoussine drissi et al [30] | DWT-MFCC-SVM | 86,84 |
| Eskidere et al [37] | Random subspace classifier | 74,17 (10-fold CV) |
| Li et al [38] | SVM+FS | 82,5 |
| Ali et al [39] | Multimodal approach | 70 |
| Ali et al. [40] | LDA-NN-GA | 95,00 |
| Proposed method. | DWT-delta delta MFCC- decision tree classifier | vowel /a/ 97,5 |

Zayrit et al [29] evaluated the /a/ vowel in the Turkish corpus by a vector of 21 prosodic features including LPC, ZCR, energy, Shannon entropy, and MFCC. The recognition accuracy of PD using the SVM classifier was around 91.18%. The researchers in [26] also conducted experiments using a Turkish corpus where the accuracy was 82.50%. This accuracy rate was shown for cepstral features applied with the SVM classifier. Sakar et al [36] reported a PD recognition accuracy of 68.45% using prosodic features in an SVM-based classification. They used the vowel /a/ from the Turkish corpus. Belhoussine drissi et al [30] reported a recognition accuracy of 86.84% for PD using cepstral features in an SVM-based classification. They used a database of 38 recordings, 18 of which were from healthy individuals and 20 from patients with Parkinson's disease from the Turkic corpus. References [37], [38], [39], and [40] extracted the spectral, and prosodic features used respectively Random subspace classifier, Support Vector Machines (SVM) as classifiers, Multimodal approach and LDA-NN-GA classifiers, they achieved recognition accuracy of PD respectively 74.17%, 82.5%, 70%, and 95%. In this article, the results show that the joint use of the wavelets and the delta delta MFCC coefficients as features brings a significantly important improvement to the performance of the Parkinson's disease diagnostic system with an accuracy of up to 97,5%.

IV. CONCLUSION

The work presented in this article is part of a project to recognize Parkinson's disease from the voice, in an educational context. The objective is to detect the state of each person if he/she is healthy or suffering from Parkinson's disease. To achieve this goal, we first proposed a comprehensive and

efficient system for automatic recognition of people's states from a Spanish corpus of the five sustained vowels (/a/, /e/, /i/, /o/, and /u/) produced by 28 subjects with Parkinson's disease and 22 healthy subjects. Our process starts with the transformation of the speech signals by several types of DWT based on the approximation of the first seven scales which will be injected into the delta delta MFCC block to extract the 12 coefficients at each time. These coefficients are applied in the classification using the decision tree classifier.

The results show that the proposed feature is superior, providing a maximum accuracy of 97.5% for the database that contains the vowel /a/. There is a significant improvement over recent studies. The complete study showed that the proposed combination of wavelets with delta delta MFCC could be used to effectively detect PD.

REFERENCES

- [1] Standaert, D. G., Saint-Hilaire, M. H., & Thomas, C. A. (2015). Parkinson's disease handbook. American Parkinson Disease Association, New York, USA.
- [2] Rusz, J., Tykalová, T., Novotný, M., Zogala, D., Růžička, E., & Dušek, P. (2022). Automated speech analysis in early untreated Parkinson's disease: relation to gender and dopaminergic transporter imaging. *European Journal of Neurology*, 29(1), 81-90.
- [3] Mollaei, F., Shiller, D. M., Baum, S. R., & Gracco, V. L. (2022). Early auditory responses to speech sounds in Parkinson's disease: preliminary data. *Scientific reports*, 12(1), 1-8.
- [4] Tykalova, T., Novotny, M., Ruzicka, E., Dusek, P., & Rusz, J. (2022). Short-term effect of dopaminergic medication on speech in early-stage Parkinson's disease. *npj Parkinson's Disease*, 8(1), 1-6.
- [5] Ahmed, I., Aljahdali, S., Khan, M. S., & Kaddoura, S. (2022). Classification of parkinson disease based on patient's voice signal using machine learning. *Intelligent Automation and Soft Computing*, 32(2), 705-722.
- [6] Karan, B., Sahu, S. S., & Mahto, K. (2020). Parkinson disease prediction using intrinsic mode function based features from speech signal. *Biocybernetics and Biomedical Engineering*, 40(1), 249-264.
- [7] Sakar, C. O., Serbes, G., Gunduz, A., Tunc, H. C., Nizam, H., Sakar, B. E., ... & Apaydin, H. (2019). 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*, 74, 255-263.
- [8] Polat, K. (2019, April). A hybrid approach to Parkinson disease classification using speech signal: the combination of smote and random forests. In *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)* (pp. 1-3). Ieee.
- [9] König, A., Satt, A., Sorin, A., Hoory, R., Toledo-Ronen, O., Derreumaux, A., ... & David, R. (2015). Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 1(1), 112-124.
- [10] Pulido, M. L. B., Hernández, J. B. A., Ballester, M. Á. F., González, C. M. T., Mekyska, J., & Směkal, Z. (2020). Alzheimer's disease and automatic speech analysis: a review. *Expert systems with applications*, 150, 113213.
- [11] Pappagari, R., Cho, J., Joshi, S., Moro-Velázquez, L., Zelasko, P., Villalba, J., & Dehak, N. (2021). Automatic detection and assessment of Alzheimer Disease using speech and language technologies in low-resource scenarios. In *Proc. Interspeech* (pp. 3825-3829).
- [12] Deshpande, G., & Schuller, B. (2020). An overview on audio, signal, speech, & language processing for covid-19. *arXiv preprint arXiv:2005.08579*.
- [13] Dash, T. K., Mishra, S., Panda, G., & Satapathy, S. C. (2021). Detection of COVID-19 from speech signal using bio-inspired based cepstral features. *Pattern Recognition*, 117, 107999.
- [14] Deshpande, G., Batliner, A., & Schuller, B. W. (2022). AI-Based human audio processing for COVID-19: A comprehensive overview. *Pattern recognition*, 122, 108289.
- [15] Quatieri, T. F., Talkar, T., & Palmer, J. S. (2020). A framework for biomarkers of COVID-19 based on coordination of speech-production subsystems. *IEEE Open Journal of Engineering in Medicine and Biology*, 1, 203-206.
- [16] Hireš, M., Gazda, M., Drotár, P., Pah, N. D., Motin, M. A., & Kumar, D. K. (2022). Convolutional neural network ensemble for Parkinson's disease detection from voice recordings. *Computers in biology and medicine*, 141, 105021.
- [17] Alissa, M., Lones, M. A., Cosgrove, J., Alty, J. E., Jamieson, S., Smith, S. L., & Vallejo, M. (2022). Parkinson's disease diagnosis using convolutional neural networks and figure-copying tasks. *Neural Computing and Applications*, 34(2), 1433-1453.
- [18] Moradi, S., Tapak, L., & Afshar, S. (2022). Identification of Novel Noninvasive Diagnostics Biomarkers in the Parkinson's Diseases and Improving the Disease Classification Using Support Vector Machine. *BioMed Research International*, 2022.
- [19] Mansour, A., & Lachiri, Z. (2017). SVM based emotional speaker recognition using MFCC-SDC features. *International Journal of Advanced Computer Science and Applications*, 8(4), 538-544.
- [20] Khorasani, A., & Daliri, M. R. (2014). HMM for classification of Parkinson's disease based on the raw gait data. *Journal of medical systems*, 38(12), 1-6.
- [21] Roth, N., Küderle, A., Ullrich, M., Gladow, T., Marxreiter, F., Klucken, J., ... & Kluge, F. (2021). Hidden Markov Model based stride segmentation on unsupervised free-living gait data in Parkinson's disease patients. *Journal of neuroengineering and rehabilitation*, 18(1), 1-15.
- [22] Wyrobnik, M., van der Meer, E., & Klostermann, F. (2022). Altered event processing in persons with Parkinson's disease. *Psychophysiology*, e14021.
- [23] Orozco-Arroyave, J. R., Arias-Londoño, J. D., Vargas-Bonilla, J. F., & Nöth, E. (2013, June). Perceptual analysis of speech signals from people with Parkinson's disease. In *International Work-Conference on the Interplay Between Natural and Artificial Computation* (pp. 201-211). Springer, Berlin, Heidelberg.
- [24] Hlavnička, J., Čmejla, R., Tykalová, T., Šonka, K., Růžička, E., & Rusz, J. (2017). Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. *Scientific reports*, 7(1), 1-13.
- [25] Soumaya, Z., Taoufiq, B. D., Nsiri, B., & Abdelkrim, A. (2019, April). Diagnosis of Parkinson disease using the wavelet transform and MFCC and SVM classifier. In *2019 4th World Conference on Complex Systems (WCCS)* (pp. 1-6). IEEE.
- [26] Benba, A., Jilbab, A., Hammouch, A., & Sandabad, S. (2015, March). Voiceprints analysis using MFCC and SVM for detecting patients with Parkinson's disease. In *2015 International conference on electrical and information technologies (ICEIT)* (pp. 300-304). IEEE.
- [27] Benba, A., Jilbab, A., & Hammouch, A. (2016). Discriminating between patients with Parkinson's and neurological diseases using cepstral analysis. *IEEE transactions on neural systems and rehabilitation engineering*, 24(10), 1100-1108.
- [28] Ung, Q. W., Basah, S. N., Muthusamy, H., Vijejan, V., & Lee, H. (2018, March). Evaluation of short-term cepstral based features for detection of Parkinson's Disease severity levels through speech signals. In *IOP Conference Series: Materials Science and Engineering* (Vol. 318, No. 1, p. 012039). IOP Publishing.
- [29] Soumaya, Z., Taoufiq, B. D., Benayad, N., Yunus, K., & Abdelkrim, A. (2021). The detection of Parkinson disease using the genetic algorithm and SVM classifier. *Applied Acoustics*, 171, 107528.
- [30] Drissi, T. B., Zayrit, S., Nsiri, B., & Ammoummou, A. (2019). Diagnosis of Parkinson's disease based on wavelet transform and mel frequency cepstral coefficients. *Int. J. Adv. Comput. Sci. Appl*, 10, 125-132.
- [31] Soumaya, Z., Taoufiq, B. D., Benayad, N., Achraf, B., & Ammoummou, A. (2020). A hybrid method for the diagnosis and classifying parkinson's

- patients based on time–frequency domain properties and K-nearest neighbor. *Journal of Medical Signals and Sensors*, 10(1), 60.
- [32] Boualoulou, N., Nsiri, B., Drissi, T. B., & Zayrit, S. (2022). Speech analysis for the detection of Parkinson's disease by combined use of empirical mode decomposition, Mel frequency cepstral coefficients, and the K-nearest neighbor classifier. In *ITM Web of Conferences* (Vol. 43, p. 01019). EDP Sciences.
- [33] Furui, S. (1986). Speaker-independent isolated word recognition using dynamic features of speech spectrum. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(1), 52-59.
- [34] Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1(1), 81-106.
- [35] Dimauro, G., Di Nicola, V., Bevilacqua, V., Caivano, D., & Girardi, F. (2017). Assessment of speech intelligibility in Parkinson's disease using a speech-to-text system. *IEEE Access*, 5, 22199-22208.
- [36] Sakar, B. E., Isenkul, M. E., Sakar, C. O., Sertbas, A., Gurgun, F., Delil, S., ... & Kursun, O. (2013). Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. *IEEE Journal of Biomedical and Health Informatics*, 17(4), 828-834.
- [37] Eskidere, Ö., Karatutlu, A., & Ünal, C. (2015, September). Detection of Parkinson's disease from vocal features using random subspace classifier ensemble. In *2015 Twelfth International Conference on Electronics Computer and Computation (ICECCO)* (pp. 1-4). IEEE.
- [38] Li, Y., Zhang, C., Jia, Y., Wang, P., Zhang, X., & Xie, T. (2017, October). Simultaneous learning of speech feature and segment for classification of Parkinson disease. In *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)* (pp. 1-6). IEEE.
- [39] Ali, L., Khan, S. U., Arshad, M., Ali, S., & Anwar, M. (2019, July). A multi-model framework for evaluating type of speech samples having complementary information about Parkinson's disease. In *2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)* (pp. 1-5). IEEE.
- [40] Ali, L., Zhu, C., Zhang, Z., & Liu, Y. (2019). Automated detection of Parkinson's disease based on multiple types of sustained phonations using linear discriminant analysis and genetically optimized neural network. *IEEE Journal of Translational Engineering in Health and Medicine*, 7, 1-10.