Short Communication

A Hybrid Method for the Diagnosis and Classifying Parkinson's Patients based on Time–frequency Domain Properties and K-nearest Neighbor

Zayrit Soumaya[1], Belhoussine Drissi Taoufiq[1], Nsiri Benayad[2], Benba Achraf[3], Abdelkrim Ammoumou[1]

[1]Laboratory Industrial Engineering, Information Processing and Logistics (GITIL), Faculty of Science Ain Chok. University Hassan II - Casablanca, [2]Laboratory Research Center STIS, M2CS, Higher School of Technical Education of Rabat (ENSET), [3]Electronic Systems Sensors and Nanobiotechnologies (E2SN), ENSET, Mohammed V University in Rabat, Morocco

Abstract

The vibrations of hands and arms are the main symptoms of Parkinson's ailment. Nevertheless, the affection of the vocal cords leads to troubles and defects in the speech, which is another accurate symptom of the disease. This article presents a diagnostic model of Parkinson's disease (PD) and proposes the time–frequency transform (wavelet WT) and Mel-frequency cepstral coefficients (MFCC) treatment for this disease. The proposed treatment is centered on the vocal signal transformation by a method based on the WT and to extract the coefficients of the MFCC and eventually the categorization of the sick and healthy patients by the use of the classifier K-nearest neighbor (KNN). The analysis used in this article uses a database that contains 18 healthy patients and twenty patients. The Daubechies mother WT is used in treatments to compress the vocal signal and extract the MFCC cepstral coefficients. As far as, the diagnosis of Parkinson's ailment is concerned the KNN classifying performance gives 89% accuracy when applied to 52% of the database as training data, whereas when we increase this percentage from 52% to 73%, we reach 98.68% accuracy which is higher than using the support-vector machine classifier. The KNN is conclusive in the determination of the PD. Moreover, the higher the training data is, the more precise the results are.

Keywords: *K-nearest neighbor, Mel-frequency cepstral coefficient, Parkinson's disease, wavelet*

Submitted: 06-Dec-18 **Revision:** 13-Jul-19 **Accepted:** 07-Sep-19 **Published:** 06-Feb-20

Introduction

In 1817, James Parkinson described Parkinson's ailment,[1] which is a neurodegenerative disease of unknown cause, characterized by the progressive destruction of a specific population of neurons.

The loss of dopamine in the midbrain induces the slowness of movement, difficulty with walking, communication trembling, and rigidity which are the most obvious motor symptoms.[2]

The performing neuron analyses and magnetic resonance imaging examination of the brain are employed in the Parkinson's disease (PD) detection. The phonation and articulation means of speech extraction and analysis can give the needed guidance in the spotting of PD.

The PD has a lot of indicators; among them, the vocal impairment which is one of the earliest.[3] Exactly, the phonation is the main part of speech production affected.[4]

Several methods are used for the diagnosis of PD (prediction cepstral coefficient,[2] perceptual linear predictive [PLP],[5] and Mel-frequency cepstral coefficient [MFCC][3,5]). Focusing on the vocal signal, we are interested in the most used in recognition systems which are the MFCC method. To exploit the human auditory system characteristics through the change of frequencies linear scale into Mel scale that allows to make cepstral analysis by passage in the log-spectral domain,[4] the cepstral analysis had been used by Shourie[6] of the electroencephalogram signals in the process of perception observance, and mental imagery proves the impact of artistic expertise and also in the appraising of hypernasality for children affected by cleft palate centered on cepstrum analysis by Akafi *et al*.[7]

The diagnostic of patients affected by PD who undergo a categorization process for appraising and home monitoring of tremor in those patients elaborated by Bazgir

How to cite this article: Soumaya Z, Taoufiq BD, Benayad N, Achraf B, Ammoumou A. A hybrid method for the diagnosis and classifying parkinson's patients based on time–frequency domain properties and K-nearest neighbor. J Med Signals Sens 2020;10:60-6.

Address for correspondence:
Dr. Zayrit Soumaya,
Faculty of Science,
Ain Chok University,
Hassan II – Casablanca,
Morocco.
E-mail: zayritsoumaya@gmail.com

Access this article online

Website: www.jmssjournal.net

DOI: 10.4103/jmss.JMSS_61_18

Quick Response Code:

et al.[8] are also some other ways to reach the diagnosis of the disease, for example, handwriting.

Filter banks that are devised accordingly to the perceptual criteria of the human ear will be required for the computation of MFCC and PLP features. It is important that this spectrum obtained by computing the discrete Fourier transform (DFT) of the windowed speech frames should be estimated accurately. The procedure involves the estimation of a short-time spectrum.

Whereas, our interest is focused on the diagnosis of PD from vocal disorders detection. The spotting of PD centered on the extraction of the cepstral coefficients of the MFCC from the speech was first proposed by Frail *et al*.[9,10] Shahbakhi *et al*. diagnosed PD[11] by the measures of fundamental frequency disturbance (Jitter), amplitude disturbance (Shimmer), and fundamental frequency F_0. Recent researches using the cepstral coefficients of the MFCC[3,5] and PLP[5] performed by Banba *et al*. and Upadhya *et al*.[12] also conducted a study on the spotting of PD by extracting MFCC and PLP by the use of the Thomson multitaper window technique. Recent studies are based on the works performed by Taoufiq BELHOUSSINE DRISSI *et al*. that deal with wavelet transform and MFCC and the support-vector machine (SVM) classifier.[13] In this work, they transform the speech signals through the sorts of DWT which were tested; then, they extracted MFCC coefficients from the signals and applying the SVM as a classifier.

Among the simplest machine learning algorithms, we find the K-nearest neighbor (KNN) algorithm, which is a robust classification method. This method is widely applied in real-time applications. The SVM principal based on the use of hyper planes to separate the classes. The shape of the decision will change if a different kernel provides, so choosing the kernel is necessary. The choice of good kernel needs to have some knowledge about the data that is not always available. Besides that, more the size of the dataset used for training is big, more the computational time for training grows nonlinearly with it. Whereas, the KNN being based on vector distance concept, so errors are bound to be less. Hence, in this article, we will choose the KNN as classifier in the aim of having higher accuracy.

In this work, we come up with a diagnosis model of the PD based on a time–frequency treatments of speech signals of a database[14] that consists of 18 sound patients and twenty affected by "PD," then extracting the cepstral coefficient of the MFCC, and in the end, a classification will be performed by the KNN classifier. We will create two training bases when the first accounts for 52% of the database and the other 73% and apply the suggested treatment (wavelet, MFCC, and KNN) on the totality of the database.

Continuous time–frequency transform

The continuous time–frequency transform (CWT) was devised by the French geophysicist Morlet in 1980 to study earth tremor signals.[15] Then, Grossmann, Meyer, Mallat, and Daubechies laid their mathematical basis for wavelets.[16] Since that time, WT is more and more used in signal processing.

A wavelet uses two coefficients: a coefficient of scale "a" which permitted to obtain various versions, which were compressed or dilated of windows stemming from the same mother wavelet, this coefficient represents the inverse of the frequency and a coefficient of translation "b" that characterizes the displacement of the window along the axis of time.

The CWT of signal s (t) is defined by:[17]

$$Ws(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} s(t).\psi^{*}\left(\frac{t-b}{a}\right)dt \quad (1)$$

here ψ(t) is the mother wavelet, and ψ*(t) is the conjugate complex ψ(t).

It should be noted that the wavelet transform gives adequate temporal resolution at high frequencies and adequate frequency resolution at low ones.

Discrete wavelet transform

The discrete wavelet transform (DWT) is the discrete version of the continuous time–frequency transform (CWT). It is achieved by the use of the Mallat algorithm[18] that is regarded as a multiresolution analysis. This algorithm is based on the definition of a pair of filters H (low-pass filter) and G (high-pass filter) and whose impulse responses h and g. Several sorts of wavelets are used in literature: Haar, Beylkin, Coiflet, Daubechies, Symmlet, Vaidyanathan, Battle,….

In this work, we will only use the wavelets of Daubechies.

Mel-frequency Cepstral Coefficient

MFCCs refer to the parameters that used the most in speech recognition systems. MFCC analysis consists of the adaption of the linear scale of frequencies into the Mel scale to exploit the properties of the human auditory system[19] that give the most effective illustration of the speech signal. The process of extracting the coefficients is shown in Figure 1.[13,20]

Preemphasis

This is a voice signal filtering process (s_n, n = 1,…, N) with a first-order finite impulse response numeric filter s_n given as follows:[19,21]

$$H(z) = 1 - k z^{-1} \quad (2)$$

Where, k is the coefficient of the preemphasis that must be comprised between $0.9 \leq k \leq 1$. In this study, we fixed the

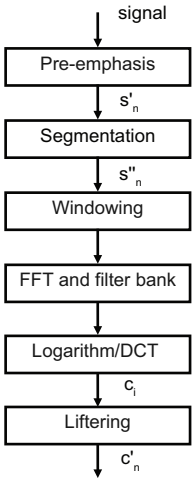

Figure 1: Extraction process of cepstral coefficients of the Mel-frequency cepstral coefficient

parameter k at 0.97.[13,20] In this way, the pre-emphasized signal is related to the signal $s_n^{'}$ by the formula below:

$$s_n^{'} = s_n - k\, s_{n-1} \quad (3)$$

Segmentation

The vocal signal is nonstationary, but the signal processing ways are stationary signals. To solve this problem, we will proceed to the segmentation of the signal into N speech samples of frames in the lapse of 10–30 ms where the voice signal is regarded as stable. To dodge unexpected transitions from frame to frame, the two adjacent frames are overlapped.[3]

Windowing

As a result of the segmentation, some discontinuities are shown at the borders of the frames; in the aim of reducing the revealed discontinuities, we multiply the samples (s'_n, $n = 1,\ldots, N$) of the frame by a Hamming window.[20,22]

$$s_n^{''} = \left\{0.54 - 0.46 \cdot \cos\left(\frac{2\pi(n-1)}{N-1}\right)\right\} \cdot s_n^{'} \quad (4)$$

Where, N is the number of samples in the frame.

The fast Fourier transform

The fast Fourier transform (FFT) application consists of converting every single frame of N samples to the frequency domain instead of the time domain. The FFT is a fast algorithm for implementing DFT.

The definition of the DFT is as follows:[13,22]

$$S_n = \sum_{k=0}^{N-1} s_k e^{\frac{-j2\pi kn}{N}} \quad (5)$$

Mel filtering with a filter bank

The human ear follows a nonlinear scale through an audible spectrum.[20] Consequently, we will use the transformation of the linear scale of frequencies to the Mel scale. The latter is in a linear space under 1000 Hz (low frequencies) and logarithmic above 1000 Hz (high frequencies).

The conversion from the linear scale to the Mel scale[13,20,22] is given as follows:

$$\mathrm{Mel}(f) = 2595 . \log_{10}\left(1 + \frac{f}{700}\right) \quad (6)$$

Logarithm/discrete cosine transform

The MFCC coefficients may be worked out first hand by employing the discrete cosine transform (DCT) on the logarithm of energies coming out of a bank of M triangular filters, apart from according to the Mel scale[13] by the following equation.

$$c_i = \sum_{j=1}^{M} m_j . \cos\left(\frac{\pi i}{M}(j - 0.5)\right) \quad (7)$$

Here, mj is the logarithm of the energy obtained with the triangular filter j, M is the number of filters bank, in our article M was set to 20,[13] and i is the number of coefficients to be extracted.

Liftering

As the higher order of the MFCC coefficients is so small, we have to apply the lifter to lift the cepstrum. Consequently, it is important to increase these amplitudes so that they become quite similar.[20,22] To achieve that, we liftered the cepstral coefficient so that the following equation can be applied:

$$c^{'} = \left(1 + \frac{L}{2} . \sin\left(\frac{n\pi}{L}\right)\right) . c_n \quad (8)$$

Here, L is the parameter of the lifter. In this article, we set L = 22.[13]

K-nearest Neighbor

KNN classifier is of a simple principle based on the theory of statistical training. First, we give a database that contains the two classes with a label vector is the training phase where the feature space is reached so that the database become separable. At the test phase, the database classified seeks the nearest neighbor given by training database, and according to this, it is classified either in class 1 or class 2. The Euclidean distance was applied to spot the nearest neighbor in the KNN algorithm.[23]

Between the two points x and y, we calculate the Euclidean distance d (x, y) using Eq. 9. Here, N is the number of characteristics such that $x = (x_1, x_2, x_3 \ldots x_N)$ and $y = (y_1, y_2, y_3 \ldots y_N)$.

$$d(x,y) = \sum_{i=1}^{N} \sqrt{x_i^2 - y_i^2} \quad (9)$$

Results

The goal of this study is to determine the KNN performance. Before the bloc of extraction of the MFCC coefficients, DWT block will be injected to achieve a

correct diagnosis of PD as shown in Figure 2.

We apply the database[14] that consists of 18 sound ones and twenty recordings of patients suffering from PD. They all utter the vowel "a."

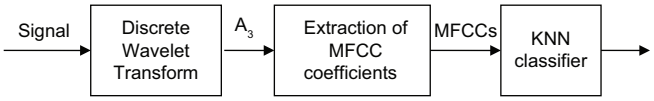

Figure 2: Process of Parkinson's disease diagnosis

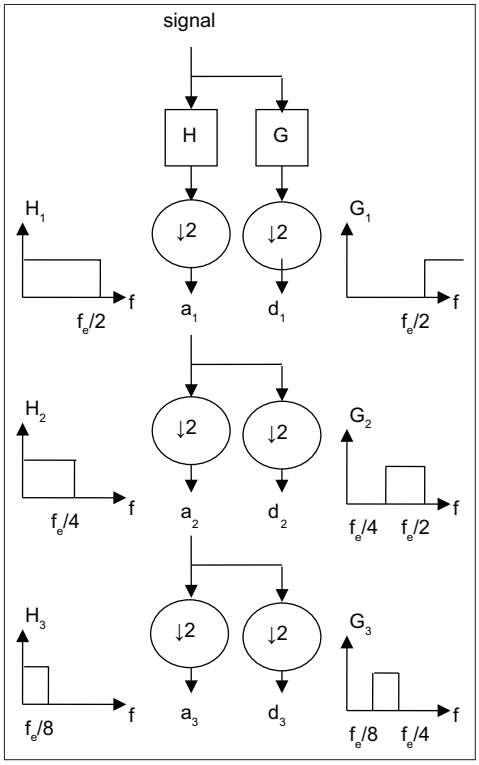

Figure 3: The approximation a_3 of Daubechies wavelet db2 at 3 scale

The algorithm of DWT is centered on the definition of a pair of filters H (low-pass filter) and G (high-pass filter). The filter outputs are subsampled by a factor of 2. The high-pass filter provides DWT coefficients or signal details at a given scale. The low-pass filter gives the coefficients of the approximation of the signal at the same scale. The same operation is again applied to the approximation, thus generating another detail and a new approximation.[24]

A process of PD diagnosis which is similar to our process is applied in the article,[13] the difference between them is the classifier. The DWT gives the higher accuracy at level 2 and the 3rd scale. Hence, in our study, we will work with the Daubechies db2 wavelet at scale 3, and we are interested only in the approximation a_3 [Figure 3].

In the first phase, we transform the vocal recordings by the use of Daubechies wavelet. The vocal signal of PD patient before and after using the Daubechies wavelet is shown in Figure 4. Figure 5 shows a zoom at the two representations of the signal.

In the second phase, we will execute an input of a_3 approximation to the MFCC block to obtain from every single patient, the first 12 MFCC coefficients employing the program "Htk mfcc matlab."[25] These coefficients will be the characteristics that be relied on to get a classification to reach an exact diagnosis. The MFCC is composed of numerous frames that need significant processing time to classify. However, such operation hinders a precise result.[19] To cope with this problem, we had recourse to the calculation of the average value of these images to obtain the voiceprint. The 12 MFCC and voiceprint for a sound patient are featured in Figure 6, as for Figure 7, it illustrates the MFCC and voiceprint for a patient suffered from PD.

In the third phase, in which we take a decision based on the categorization of the patients. In this aims, we create two

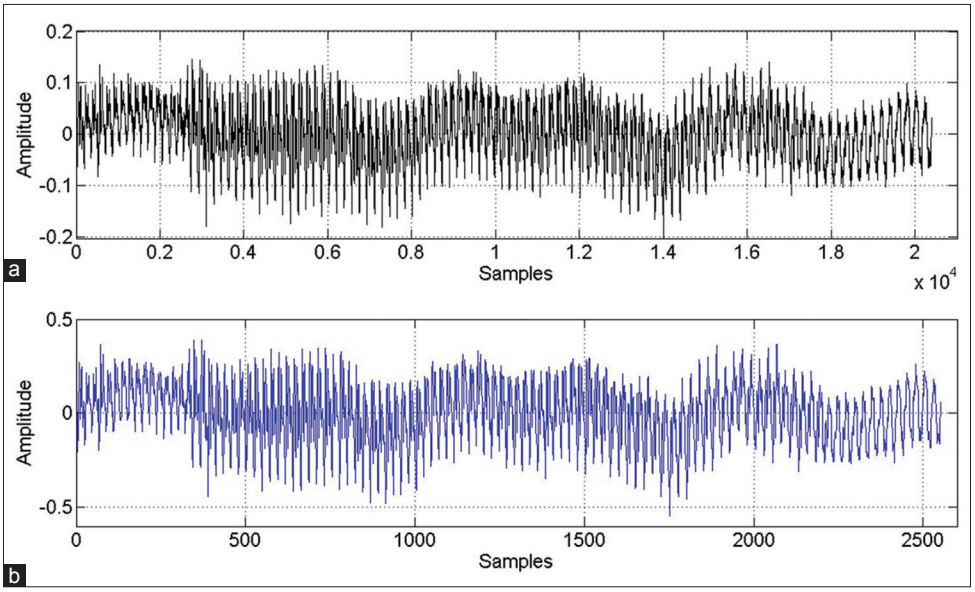

Figure 4: (a) Speech before the transformation. (b) Speech after being transformed through the use of wavelet

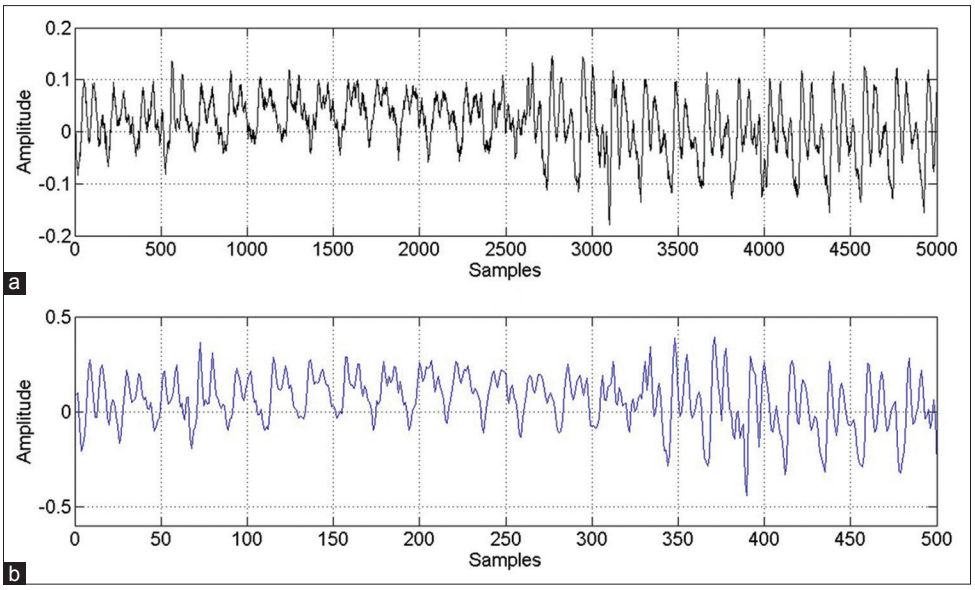

Figure 5: (a) A zoom of speech before the transformation. (b) A zoom of speech after being transformed through the use of wavelet

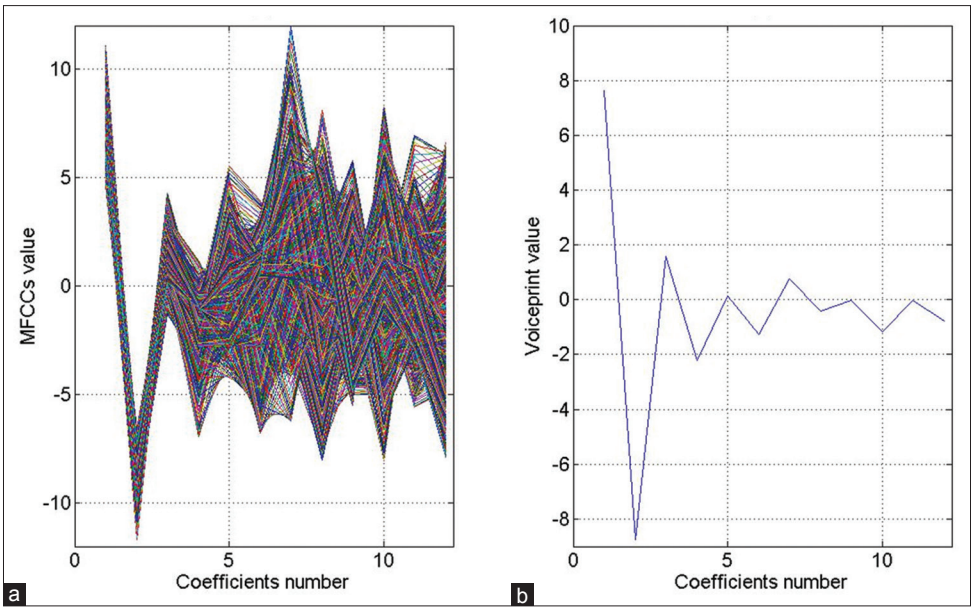

Figure 6: (a) Mel-frequency cepstral coefficient value of a healthy patient. (b) Voiceprint value of a healthy patient

training bases one of 52% and the other 73% of the database. At the first step, we will carry out a test on our database by using the first training base (of 52%) and another test during the use of the second training base (of 73%) classifier.

In a categorization problem, the labels are part of the following possible identities:

Moreover, the task consists of assigning a test example to one of the C classes. KNN classifier is the most used procedure. Moreover, the widely used method is setting $K = 1$ yields the nearest neighbor classification rule.

In spite of it is simplicity, KNN so often gives a good performance mainly for large data sets.

We calculated measures such as accuracy, sensitivity, and specificity by applying the following formula to determine the performance of the classifier:[13,26,27]

$$\text{Accuracy} = \frac{TN + TP}{TP + TN + FP + FN} \quad (10)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (11)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (12)$$

With:

- TP stands for true positive (correctly classified healthy patients)

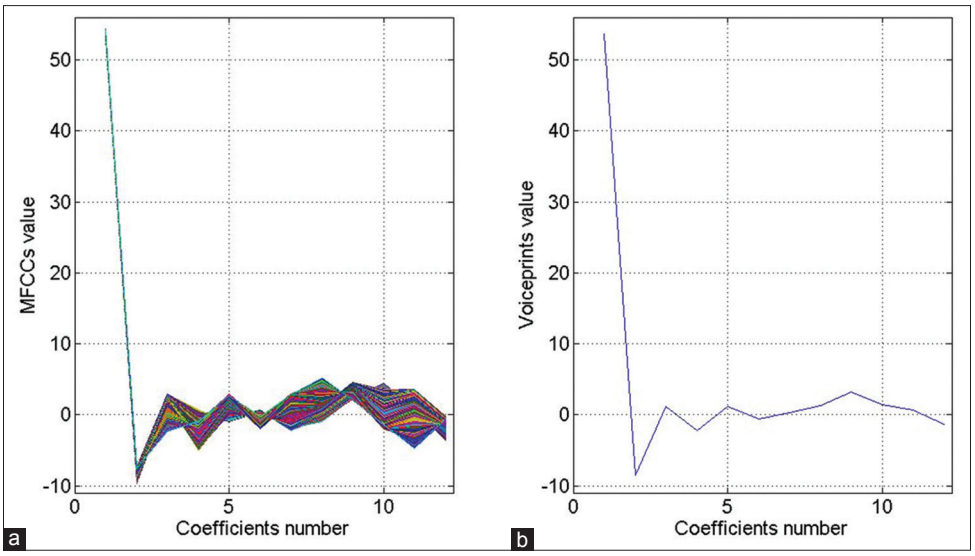

Figure 7: (a) Mel-frequency cepstral coefficient value of a sick patient. (b) Voiceprint value of a sick patient

- TN stands for true negative (correctly classified patients)
- FP stands for false positive (incorrectly classified patients)
- FN stands for false negative (incorrectly classified sound patients).

The percentage of the test is accuracy, sensitivity as well as specificity of all the recordings by the use of the 52% training is shown in Table 1, and then their percentage by using the training base of 73% in the test of all the recordings (including 6 sick and 4 healthy patients) is shown in Table 2.

Conclusion

We have presented in this article, a sample of diagnosis based on PD that is centered on the signal treatment, in which we will employ the wavelet transform and the MFCC using a database of recordings of sick patients and healthy ones while they pronounce the vowel "a." The change of speech signals is treated by Daubechies wavelet by the third-scale approximation and then, we will recover the 12 cepstral coefficients after injecting the approximation into the MFCC bloc. To make a decision on which one is sick or healthy, we work with the KNN classifier by using two-learning bases, one is 52% and the other 73% of the database. When you work with the database of 52%, one obtains an accuracy of 89% which is higher than the accuracy obtained by using the classifier SVM with the database of 73%, and when we increase the percentage of the database to 73%, we get an accuracy of 98.68% and from that one can conclude that the increase of the base of data gives us better results by increasing the accuracy of the classifier and that the KNN is more accurate than the SVM classifier.

Financial support and sponsorship

None.

Table 1: Results of diagnosis by using 52% training base

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
KNN	89	100	80

KNN – K Nearest Neighbor

Table 2: Results of the diagnosis by using 73% training base

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
KNN	98.68	98.14	99.16

KNN – K-nearest neighbor

Conflicts of interest

There are no conflicts of interest.

References

1. Parkinson J. An Essay on the Shaking Palsy. London: Whittingham and Rowland for Sherwood, Neely, and Jones; 1817.
2. Orozco-Arroyave JR, Arias-Londoño JD, Vargas-Bonilla JF, Nöth E. Perceptual analysis of speech signals from people with Parkinson's disease. In: Ferrández Vicente JM, Álvarez Sánchez JR, de la Paz López F, Toledo Moreo FJ. Editors. Natural and Artificial Models in Computation and Biology. IWINAC 2013. Lecture Notes in Computer Science. Vol. 7930, Berlin, Heidelberg; Springer; 2013b.
3. Benba A, Jilbab A, Hammouch A, Sandabad S. Voiceprints analysis using MFCC and SVM for detecting patients with Parkinson's disease. IEEE 1st International Conference on Electrical and Information Technologies ICEIT'2015. 2015. p. 300-4.
4. Rabiner LR, Schafer RW. Introduction to digital speech processing. Foundat Trends Signal Process 2007;1:1-194.
5. Benba A, Jilbab A, Hammouch A. Discriminating between patients with Parkinson's and neurological diseases using cepstral analysis. IEEE Trans Neural Syst Rehabil Eng 2016;24:1100-08.
6. Shourie N. Cepstral analysis of EEG during visual perception and mental imagery reveals the influence of artistic expertise. J Med Signals Sens 2016;6:203-17.
7. Akafi E, Vali M, Moradi N, Baghban K. Assessment of

hypernasality for children with cleft palate based on cepstrum analysis. J Med Signals Sens 2013;3:209-15.
8. Bazgir O, Habibi SAH, Palma L, Pierleoni P, Nafees S. A classification system for assessment and home monitoring of tremor in patients with Parkinson's disease. J Med Signals Sens 2018;8:65-72.
9. Frail R, Godino-Llorente JI, Saenz-Lechon N, Osma-Ruiz V, Fredouille C. MFCC-based Remote Pathology Detection on Speech Transmitted Through The Telephone Channel. Proceedings Biosignals, Porto; 2009.
10. Jafari A. Classification of Parkinson's disease patients using nonlinear phonetic features and Mel-frequency cepstral analysis. Biomed Eng Appl Basis Communication 2013;25:1350001.
11. Shahbakhi M, Far DT, Tahami E. Speech analysis for diagnosis of Parkinson's disease using genetic algorithm and support vector machine. J Biomed Sci Eng 2014;7:147-56.
12. Upadhya SS, Cheeranb AN. Nirmalc JH. Thomson Multitaper MFCC and PLP voice features for early detection of Parkinson disease Biomed Signal Process Control 2018;46:293-301.
13. Belhoussine T, Zayrit S, Nsiri B, Ammoummou A. Diagnosis of Parkinson's disease based on wavelet transform and Mel Frequency Cepstral Coefficients. Int J Adv Comput Sci Appl 2019;10:125-32.
14. Sakar BE, Isenkul ME, Sakar CO, Sertbas A, Gurgen F, Delil S, *et al.* Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. IEEE J Biomed Health Inform 2013;17:828-34.
15. Mallat S. A Wavelet Tour of Signal Processing. The Sparse Way. 3rd ed. Academic Press, Inc. Orlando, FL: USA; 2009.
16. Daubechies I. Ten Lectures on Wavelets, of CBMS-NSF Regional Conference Series in Applied Mathematics. Vol. 61. SIAM, Philadelphia, PA; 1992.
17. Chui CK. An Introduction to Wavelets. Boston: Academic Press; 1992.
18. Mallat S. Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Trans Paerntt Anal Mack Intell 1989;11:674-93.
19. Hacine-Gharbi A. Selection of Relevant Acoustic Parameters for Speech Recognition. University Orléans; 2012.
20. Young S, Evermann G, Gales M, Hain T, Kershaw D, Liu X. The HTK Book. Trumpington St, Cambridge CB2 1PZ: Cambridge University Engineering Department; 2006.
21. Rabiner LR. Juan BH. Hidden Markov Models for Speech Recognition. Englewood Cliffs NJ, editors. Fundamentals of Speech Recognition. USA: Prentice Hall; 1993.
22. Benba A, Jilbab A, HammouchA. Voice Analysis for Detecting Persons with Parkinson's Disease Using MFCC and VQ. The 2014 International Conference on Circuits, Systems and Signal Processing. Saint Petersburg, Russia: Saint Petersburg State Polytechnic University; September 23-25, 2014.
23. Benba A, Jilbab A, Hammouch A. Voice Analysis for Detecting Persons with Parkinson's Disease Using MFCC and VQ. The 2014 International Conference on Circuits, Systems and Signal Processing. Saint Petersburg, Russia: Saint Petersburg State Polytechnic University; September 23-25, 2014.
24. Bahoura M. Analysis of Respiratory Acoustic Signals: Contribution to the Automatic Detection of Sibilants by Wavelet Packages. PhD Thesis, University de Rouen. Defended on; 1999.
25. Wojcicki K. writehtk. In: Voicebox Toolbox; 2011. Available from: http://www.mathworks.com/matlabcentral/fileexchange/32849-htk-mfcc-matlab/content/mfcc/writehtk.m. [Last accessed on 2018 Mar 14].
26. Benba A, Jilbab A, Hammouch A. Hybridization of best Acoustic Cues for Detecting Persons With Parkinson's Disease. 2nd World Conference on Complex System (WCCS'14). IEEE, Agadir; November 10-12, 2014.
27. Benba A, Jilbab A, Hammouch A. Voiceprint Analysis Using Perceptual Linear Prediction and Support Vector Machines. For Detecting Persons With Parkinson's Disease. The 3rd International Conference on Health Science and Biomedical Systems (HSBS '14) Florence. Italy; November 22-24, 2014.